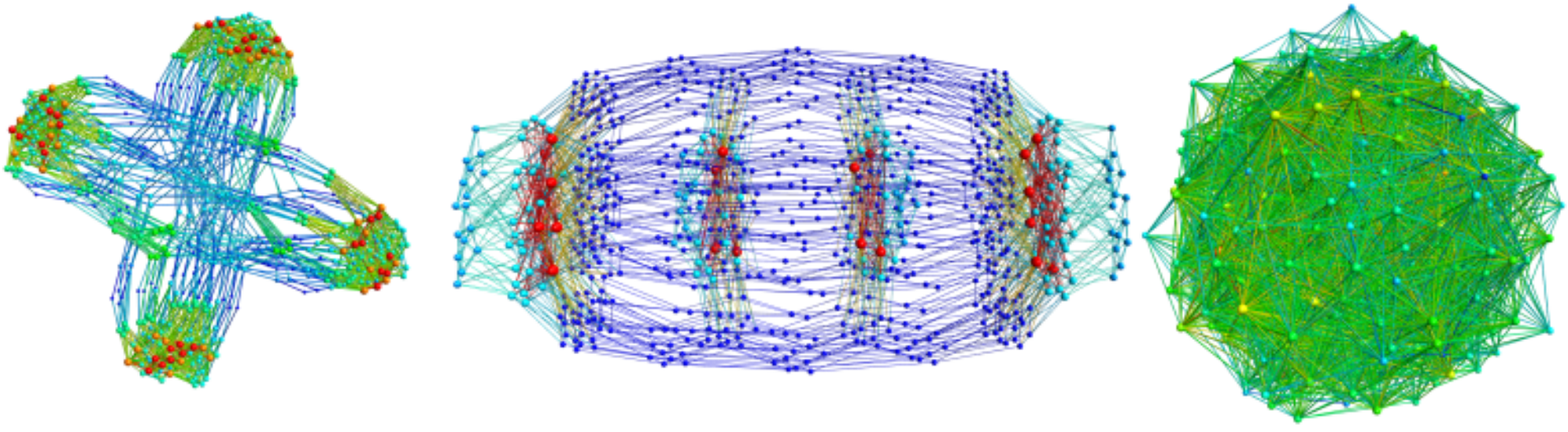


Adrian Balint Anton Belov
Marijn Heule Matti Järvisalo

July 14, 2013

SAT Competition 2013

Tracks

1. Core solvers, Sequential, Random SAT+UNSAT
2. Core solvers, Sequential, Random certified UNSAT
3. Core solvers, Sequential, Random SAT
4. Core solvers, Sequential, Hard-combinatorial SAT+UNSAT
5. Core solvers, Sequential, Hard-combinatorial certified UNSAT
6. Core solvers, Sequential, Hard-combinatorial SAT
7. Core solvers, Sequential, Application SAT+UNSAT
8. Core solvers, Sequential, Application certified UNSAT
9. Core solvers, Sequential, Application SAT
10. Core solvers, Sequential, MiniSAT Hack, Application SAT+UNSAT
11. Core Solvers, Parallel, Hard-combinatorial SAT+UNSAT
12. Core Solvers, Parallel, Application SAT+UNSAT
13. Open track

Rules (excerpt)

1. Organizers can compete but had to make the MD5 of their code available before the opening of the submission system
2. Internal agreement: An organizer can not participate in the instances selection of the track where he is competing (conflict of interests)
3. Not more than 2 submission per author in a track (several violations that resulted in "non-competing" tagging of several solvers, i.e. not eligible to win a price - this decision was taken by the judge board)
4. Results from the SAT+UNSAT tracks were exported to the SAT tracks unless that resulted in violating the 2 submission per author track.
5. Every solver and benchmark has to be described in a short solver/benchmark description (worked much better than last year)

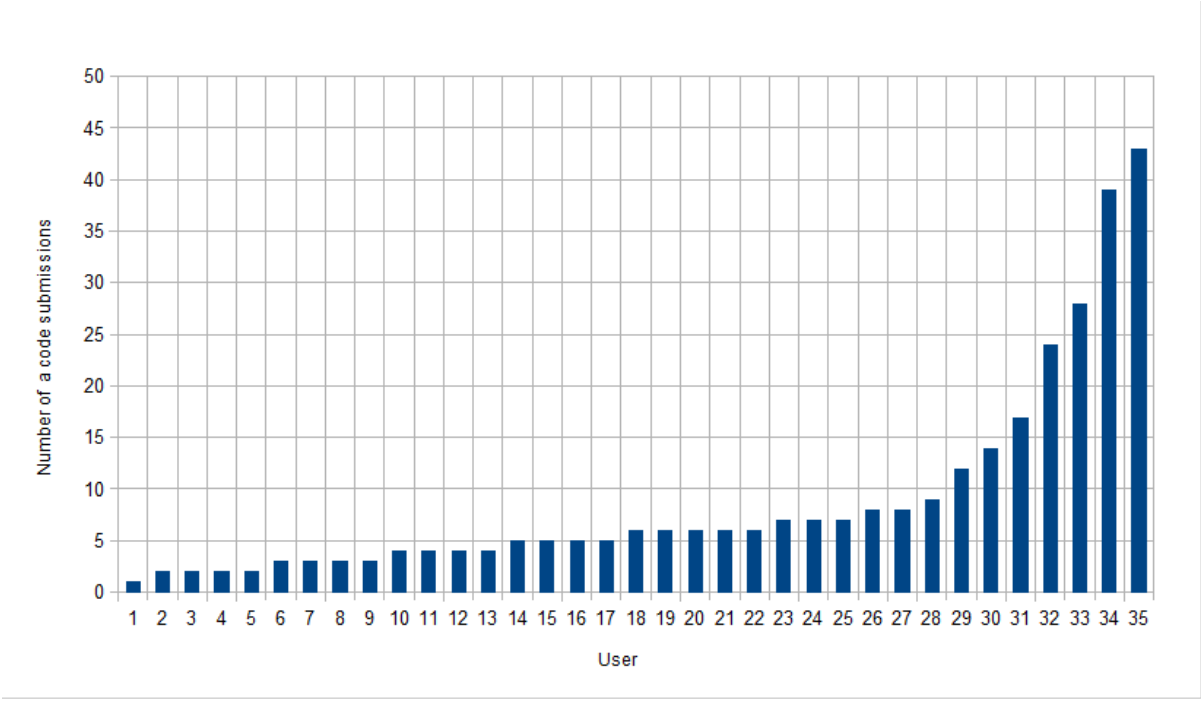
New - Certified UNSAT Tracks

- New Checker available online - allowed formats: TraceCheck, RUP, DRUP, BRUP
- Should prevent problems like this one:

arcfour_initialPermutation_5_32.cnf	0.313951	UNSAT due to a bug
arcfour_initialPermutation_6_14.cnf	1.43778	UNSAT due to a bug
arcfour_initialPermutation_6_15.cnf	1.42478	UNSAT due to a bug
arcfour_initialPermutation_6_16.cnf	1.43878	UNSAT due to a bug
arcfour_initialPermutation_6_24.cnf	1.43978	UNSAT due to a bug
arcfour_initialPermutation_6_40.cnf	1.42378	UNSAT due to a bug
arcfour_initialPermutation_6_56.cnf	1.43878	UNSAT due to a bug
arcfour_initialPermutation_6_64.cnf	1.51677	UNSAT due to a bug

- None other solver solved these UNSAT instances

Code Submission Statistics



Number of times a user has submitted a code in EDACC.

- 310 code packcages submitted for 93 solvers $\rightarrow$ every solver was submitted 3.3 times until it worked

Resource Usage - Lower Bound

- Used ≈**100.000h** of CPU time on a 8 core machine using only 1/4 of the machine
- Blocked ≈**4000.00h** of CPU time of resources
- Would cost ≈**50.000€** on Amazon EC2 ← **666€** per solver author
- Power consumption ≈**6000kWh** ≅ **5** person household /year
- CO_2 production ≈ **3t** ≈ **the weight of the submitting authors**

Is this resource usage worth the results?

Execution Procedure

- One phase competition
- Automatized testing phase for competitors to test their solvers on the executions system
- No further changes possible after the testing phase
- Cluster used:

  bwGrid (2x Quad-Core Intel Xeon E5440, 2.83 GHz with 16GB RAM)
- Execution System: EDACC
  - Simple and transparent execution of solvers on distributed clusters
  - Automatic collection and (statistical) analysis of the results
  - Web front end provides a competition mode (with user management)
- **Daniel Diepold and Simon Gerber** *worked heavily* on the execution

SAT Competition 2013 Proceedings

1. All submitted solver and benchmark descriptions
2. Descriptions of benchmark selection and generation procedures
3. Permanent URI: `http://hdl.handle.net/10138/40026`
4. ISBN 978-952-10-8991-6
5. Solver description for each solver also available through the EDACC web front-end

Benchmarks

Application and Hard-Combinatorial tracks

- Many new submissions (5 Application, 10 Hard-Combinatorial).
- 1/2 of selected Application an 2/3 of selected Hard-Combinatorial benchmarks are new.
- Large diversity: 19 sources ("buckets") in Application; 35 in Hard-Combinatorial.

Random tracks

- SAT benchmarks: k-SAT for $k = 3, \ldots, 7$
  - "Threshold" — around the threshold, up to 13000 vars.
  - "Huge" — under threshold, up to 1000000 vars.
- UNSAT benchmarks: All at the phase-transition, all with different size. If the an instance was SAT a new one of the same size is generated.

Winners — Minisat Hack Track

1. SINNminisat 1.0.0 (206) — Takeru Yasumoto
2. minisat_bit 1.0 (189) — Jingchao Chen
3. MiniGolf prefetch (175) — Norbert Manthey

Winners — Open Track

1. CSHCpar8 (234) — Yuri Malitsky, Ashish Sabharwal, Horst Samulowitz and Meinolf Sellmann.
2. MIPSat (231) — Sergio Núñez, Daniel Borrajo and Carlos Linares López
3. GlucoRed+March r531 (186) — Siert Wieringa

Winners — Parallel Tracks

Application SAT+UNSAT

1. Plingeling aqw (271) — Armin Biere
2. Treengeling aqw (260) — Armin Biere
3. PeneLoPe 2013 (247) — Gilles Audemard, Benoît Hoessen, Saïd Jabbour, Jean-Marie Lagniez and Cédric Piette

Hard-Combinatorial SAT+UNSAT

1. Treengeling aqw (253) — Armin Biere
2. Plingeling aqw (242) — Armin Biere
3. pmcSAT 1.0 (219) — Ricardo Marques, Luís Guerra e Silva, Paulo Flores and Luís Miguel Silveira

Winners — Random Tracks

SAT+UNSAT

1. CSHCrandMC (179) — Yuri Malitsky, Ashish Sabharwal, *et al.*
2. MIPSat random sat_unsat (151) — Sergio Núñez *et al.*
3. march_vflip 1.0 (120) — Jingchao Chen

SAT

1. probSAT SC13 (99) — Adrian Balint and Uwe Schöning
2. sattime2013 2013 (92) — Chu Min Li and Yu Li
3. Ncca+ V 1.0 (91) — Djamal Habet, Donia Toumi and André Abramé

Certified UNSAT

- dk-SAT11 unsat (76) — Donald Knuth
- march_br unsat (72) — Marijn Heule

Winners —- Hard-Combinatorial Tracks

SAT+UNSAT

1. BreakIDGlucose 1 (208) — Jo Devriendt and Bart Bogaerts
2. gluebit_clasp 1.0 (208) — Jingchao Chen
3. glucose 2.3 (202) — Gilles Audemard and Laurent Simon

SAT

1. glucose 2.3 (109) — Gilles Audemard and Laurent Simon
2. gluebit_clasp 1.0 (109) — Jingchao Chen
3. BreakIDGlucose 1 (109) — Jo Devriendt and Bart Bogaerts

Certified UNSAT

1. Riss3g cert (92) — Norbert Manthey
2. glucose 2.3 (certified) (91) — Gilles Audemard and Laurent Simon
3. forl drup-nocachestamp (83) — Mate Soos

Winners — Application Tracks

SAT+UNSAT

1. Lingeling aqw (231) — Armin Biere
2. Lingeling 587f (212) — Armin Biere
3. ZENN 0.1.0 (208) — Takeru Yasumoto

SAT

1. Lingeling aqw (119) — Armin Biere
2. ZENN 0.1.0 (113) — Takeru Yasumoto
3. satUZK 48 (110) – Alexander van der Grinten *et al.*

Certified UNSAT

1. glucose 2.3 (certified) (94) — Gilles Audemard and Laurent Simon
2. glueminisat-cert-unsat 2.2.7j (91) — Hidetomo Nabeshima *et al.*
3. Riss3g cert (85) — Norbert Manthey

Thanks

Thanks to all the submitters of benchmarks and solvers!

All results are available on the EDACC system:

`http://edacc4.informatik.uni-ulm.de/SC13/`