Frank Hutter, **Adrian Balint**,
Sam Bayless, Holger Hoos,
Kevin Leyton-Brown

08.03.2013

The Configurable SAT Solver Challenge 2013

Motivation

- Within a given SAT application domain, instances are often similar

  $\rightarrow$ designers can improve SAT solvers by parameter tuning

  $\rightarrow$ aim for automated performance improvements

- The CSSC emphasizes this application-specific view. It

  $\rightarrow$ assesses peak performance of SAT solvers with parameters

  $\rightarrow$ uses parameter settings that can be identified automatically

  $\rightarrow$ increases fairness: all solvers use same configuration process (same time budget)

- Side effects:

  $\rightarrow$ Encourage solver developers to expose more parameters

  $\rightarrow$ Assess efficacy of current configuration procedures

Participants

- Clasp: Benjamin Kaufmann, Torsten Schaub and Marius Schneider
- Riss3g: Norbert Manthey
- gNovelty$^+$GCa, gNovelty$^+$GCwa, gNovelty$^+$PCL: Thach-Thao Duong and Duc Nghia Pham
- SAT4J: Daniel Le Berre and Stéphanie Roussel
- Solver43: Valeriy Balabanov
- for1-nodrup: Mate Soos
- lingeling: Armin Biere
- Simpsat: Cheng-Shen Han and Jie-Hong Roland Jiang
- riss3gExt (closed source, ineligible for medals): Norbert Manthey

Benchmarks

- Application

  → BMC08: BMC unrolling of HWMCC'08 instances.

  → IBM: BMC set from Zarpas (2005)

  → SWV: generated with Calysto static checker (Babić & Hu, 2007)

  → Circuit Fuzz: SAT encoding of randomly generated circuits (Brummayer et al, 2010)

- Crafted

  → GI: Graph isomorphism

  → LABS: Peak Side Lobe coding to SAT

- Random:

  → K3-300s: 3-SAT at the phase transition

  → SAT5500: satisfiable 5-SAT with v=500, c=10000

  → unsat-unif-k5: unsatisfiable 5-SAT with v=50, c=1056

Automated configuration process

- Configurators used:

  → ParamILS (Hutter et al. 2009), 5 runs independent runs

  → SMAC (Hutter et al. 2011), 5 runs independent runs

  → GGA (Ansótegui et al. 2009), 1 run with 5 workers
  (Difficulties running this, continuing post-competition assessment)

- Configuration budget: 48 CPU hours / 50 wall clock hours
- Cutoff time per solver run during configuration: 300 seconds
- Overall budget of the competition: roughly 5 000 CPU days

Determination of winners

- For each benchmark, solver + configuration space:

  $\rightarrow$ Assess final configurations of each run on full training set

  $\rightarrow$ Select the one with best training performance

  $\rightarrow$ Assess that configuration on the test set
  (same time bound of 300s per run as during configuration)

Application Track

Solver	**CSSC Results**		**Solver defaults**	
	#(timeouts)	avg. time for solved	#(timeouts)	avg. time for solved
riss3gExt	82	8.25	123	10.53
lingeling	115	12.78	136	16.35
riss3g	117	10.18	122	10.80
Solver43	127	13.17	127	14.46
forl-nodrup	128	15.07	152	19.01
simpsat	128	20.27	134	19.59
clasp-cssc	130	9.97	163	11.21
sat4j	176	19.46	184	21.37
gnovelty+GCwa	1090	23.88	1131	7.30
gnovelty+PCL	1099	9.62	1101	14.07
gnovelty+GCa	1104	10.97	1129	14.81

Table 1: Final results for CSSC track Application

Winners - Application Track

- riss3gExt (not competing - cloused source)
- Gold: lingeling
- Silver: riss3g
- Bronze: Solver43

Hard Combinatorial Track

Solver	**CSSC Results**		**Solver defaults**	
	#(timeouts)	avg. time for solved	#(timeouts)	avg. time for solved
riss3gExt	44	4.71	148	7.09
clasp-cssc	96	13.87	139	7.10
forl-nodrup	98	15.27	135	10.01
lingeling	107	10.75	148	7.51
riss3g	131	7.52	148	7.52
simpsat	149	9.86	149	9.86
Solver43	152	8.79	156	10.75
sat4j	161	7.44	172	9.42
gnovelty+GCwa	334	13.34	375	4.33
gnovelty+GCa	353	7.13	423	4.74
gnovelty+PCL	361	12.02	378	12.02

Table 2: Final results for CSSC track Hard Combinatorial

Winners - Hard Combinatorial Track

- riss3gExt (not competing - cloused source)
- Gold: clasp-cssc
- Silver: forl-nodrup
- Bronze: lingeling

Random Track

Solver	**CSSC Results**		**Solver defaults**	
	#(timeouts)	avg. time for solved	#(timeouts)	avg. time for solved
clasp-cssc	250	1.58	261	14.10
lingeling	250	4.20	258	13.24
riss3gExt	250	7.20	261	11.00
riss3g	250	7.68	260	11.51
Solver43	253	12.32	256	11.34
simpsat	254	13.85	254	13.95
sat4j	255	14.96	257	16.33
forl-nodrup	258	11.61	289	14.75
gnovelty+GCwa	375	13.40	382	22.64
gnovelty+GCa	378	19.58	537	48.11
gnovelty+PCL	385	44.46	624	0.11

Table 3: Final results for CSSC track Random

Winners - Random Track

- Gold: clasp-cssc
- Silver: lingeling
- Bronze: riss3g

The Configurable SAT Challenge 2014 - It will happen

Solver Preparation

- Do not take premature design decision
- Expose your parameters beginning with the design

Benchmark Preparation

- Please also submit interesting instance distributions
- Homogeneous instances
- Best 600 or more instances (300+ for train, 300+ for test)
- At least some instances solvable with solver defaults in 300s