

The QBF Gallery 2013

A Non-Competitive Evaluation of QBF Tools

Florian Lonsing¹ Martina Seidl² Allen Van Gelder³

¹Vienna University of Technology
<http://www.kr.tuwien.ac.at/staff/lonsing>

²Johannes Kepler University, Linz, Austria
<http://fmv.jku.at/seidl>

³University of California at Santa Cruz, USA
<http://www.cse.ucsc.edu/~avg>



This work is supported by the Austrian Science Fund (FWF) under grants S11409-N23 and S11408-N23 as well as by the Vienna Science and Technology Fund (WWTF) under grant ICT10-018.

Recent Progress in QBF Solving:

- Powerful preprocessing techniques.
- New solvers.
- Certificate extraction.
- New encodings for various applications.

Purpose of the QBF Gallery 2013:

- Evaluate the state-of-the-art in QBF solving and related aspects.
- **No competition, no winners, no prizes: collect and analyze data.**
- Community-driven organization, much interactions during runs.
- Ultimate goal: Identify promising research directions.

Details: <http://www.kr.tuwien.ac.at/events/qbfgallery2013/>

Introduction (2/2)

Showcases:

1. Preprocessing: individual and combined effects of different preprocessors.
2. Solving: diverse performance of solvers.
3. Applications: identify user requirements of QBF tools.
4. Certificates: solver performance vs. performance of certification tools.

Participants:

- 23 contributors from 8 countries.
- 14 CNF-solvers, 1 Non-CNF-solver, 3 2QBF-solvers.
- 4 preprocessors.
- 2 certification tools.
- 5 new benchmark sets.

General Statistics:

- 1st QBF Gallery ever (complementary to biannual QBFEval competition)
- Experiments on
 - FMV Cluster @ JKU Linz
 - Infosys Cluster @ TU Vienna
- > 7000 considered formulas (from QBFLIB and new benchmarks)
- > 114.000 runs in 3.92297e+07 seconds (11.000 hours)

Introduction (2/2)

Showcases:

1. Preprocessing: individual and combined effects of different preprocessors.
2. Solving: diverse performance of solvers.
3. Applications: identify user requirements of QBF tools.
4. Certificates: solver performance vs. performance of certification tools.

Participants:

- 23 contributors from 8 countries.
- 14 CNF-solvers, 1 Non-CNF-solver, 3 2QBF-solvers.
- 4 preprocessors.
- 2 certification tools.
- 5 new benchmark sets.

General Statistics:

- 1st QBF Gallery ever (complementary to biannual QBFEval competition)
- Experiments on
 - FMV Cluster @ JKU Linz
 - Infosys Cluster @ TU Vienna
- > 7000 considered formulas (from QBFLIB and new benchmarks)
- > 114.000 runs in 3.92297e+07 seconds (11.000 hours)

Introduction (2/2)

Showcases:

1. Preprocessing: individual and combined effects of different preprocessors.
2. Solving: diverse performance of solvers.
3. Applications: identify user requirements of QBF tools.
4. Certificates: solver performance vs. performance of certification tools.

Participants:

- 23 contributors from 8 countries.
- 14 CNF-solvers, 1 Non-CNF-solver, 3 2QBF-solvers.
- 4 preprocessors.
- 2 certification tools.
- 5 new benchmark sets.

General Statistics:

- 1st QBF Gallery ever (complementary to biannual QBFEval competition)
- Experiments on
 - FMV Cluster @ JKU Linz
 - Infosys Cluster @ TU Vienna
- > 7000 considered formulas (from QBFLIB and new benchmarks)
- > 114.000 runs in 3.92297e+07 seconds (11.000 hours)

Benchmarks (1/2)

- **eval2012r2:**
345 formulas sampled from the collection of formulas available from QBFLIB at <http://www.qbflib.org/>. This set was also used for QBFEVAL 2012-R2.
- **eval2012r2-inc-preprocessed:**
234 + 241 instances in two sets, respectively, resulting from the benchmark set eval2012r2 by incremental preprocessing.
- **eval2012r2-bloqqer:**
276 instances preprocessed with Bloqqer (69 were solved by Bloqqer itself).
- **sampling-set:**
7 × 445 formulas randomly selected (about 6 samples per family) from QBFLIB, which currently contains about 19.000 formulas.
- **eval10:**
568 formulas sampled from the collection of formulas available from QBFLIB. This set was used in the QBFEVAL 2010.
- **eval10-bloqqer:**
420 instances preprocessed with Bloqqer (148 were solved by Bloqqer itself).

- **reduction-finding-full-set-params-k1c3n4 (NEW):**

4608 formulas generated using a generator for instances from reduction finding, contributed by Charles Jordan and Lukasz Kaiser.

- **conformant-planning (NEW):**

1750 instances from a planning domain with uncertainty in the initial state, contributed by Martin Kronegger, Andreas Pfandler, and Reinhard Pichler.

- **planning-CTE (NEW):**

150 instances from QBF-based planning, contributed by Michael Cashmore.

- **sauer-reimer (NEW):**

924 instances from QBF-based test generation, contributed by Matthias Sauer, Sven Reimer.

- **qbf-hardness (NEW):**

198 instances from bounded model checking of incomplete designs, contributed by Paolo Marin and Christian Miller.

Solvers and Tools

Tool Name	Submitter(s)
<i>Preprocessors (4)</i>	
hiqqer3e	Allen Van Gelder
hiqqer3p	Allen Van Gelder
Bloqqer	Martina Seidl and Armin Biere
squeezebf	Massimo Narizzano
<i>Solvers (14)</i>	
squeezebf1.2-qube3.0	Massimo Narizzano
ghost	Will Klieber
ghost-CEGAR	Will Klieber
bGhost-CEGAR	Will Klieber
rareqs	Mikolás Janota
hiqqer3	Allen Van Gelder
ooq	Alexandra Goultiaeva
dual_ooq	Alexandra Goultiaeva
sDual_ooq	Alexandra Goultiaeva
nenoFex	Florian Lonsing
DepQBF	Florian Lonsing
depqbf-lazy-qup	Florian Lonsing
free2qbf(2QBF)	Sam Bayless
mini2qbf(2QBF)	Sam Bayless
mini2qbf with ext. (2QBF)	Sam Bayless
<i>Certification Tools (2)</i>	
QBFcert	Aina Niemetz and Mathias Preiner
ResQu	Valeriy Balabanov and Jie-Hong Roland Jiang

Showcase: Preprocessing (1/4)

	A hiqqer3e			B Bloqqer			C hiqqer3p			D squeezebf		
	t	s	u	t	s	u	t	s	u	t	s	u
eval2012r2	19	0	19	69	33	36	77	35	42	11	3	8
qbf-hardness	0	0	0	49	12	37	51	12	39	12	0	12
sauer-reimer	81	0	81	137	24	113	153	29	124	78	9	69
planning-CTE	0	0	0	3	2	1	7	6	1	0	0	0
conf.-planning	646	0	646	489	11	478	486	12	474	48	0	48
red.-finding	176	0	176	1496	837	659	1650	924	726	674	326	348

Individual preprocessors: solved instances (t), solved satisfiable (s) and solved unsatisfiable instances (u).

	A + B + C + D		
	t	s	u
eval2012r2	87	36	51
qbf-hardness	51	12	39
sauer-reimer	158	29	129
planning-CTE	7	6	1
conf.-planning	757	12	745
red.-finding	1679	940	739

	(ABCD) ⁶		
	t	s	u
eval2012r2	107	44	63
qbf-hardness	51	12	39
sauer-reimer	180	31	149
planning-CTE	11	8	3
conf.-planning	938	13	925
red.-finding	1855	936	919

Overall individually solved instances (left) and instances solved by incremental preprocessing in six rounds with 120 sec. time limit per individual call (right). Additionally, we ran all 24 possible execution sequences of A, B, C, and D on the set eval2012r2.

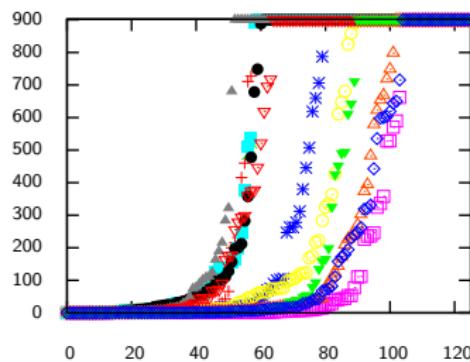
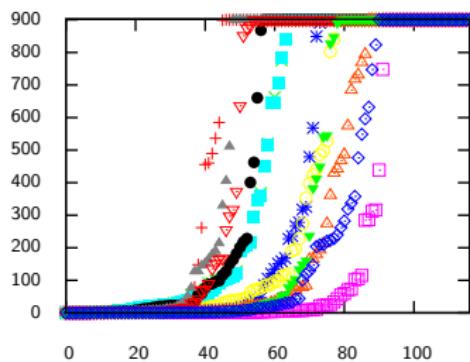
Showcase: Preprocessing (2/4)

eval2012r2			
	t	s	u
$(A^2B^2C^2D^2)^6$	111	46	65
$(A^2B^2D^2C^2)^6$	111	45	66
$(A^2D^2B^2C^2)^6$	104	42	62
$(B^2C^2D^2A^2)^6$	103	42	61
$(D^2A^2B^2C^2)^6$	102	38	64
$(ABCD)^6$	107	44	63
A	19	0	19
B	69	33	36
C	77	35	42
D	11	3	8
$A + B + C + D$	87	36	51

“A” is hiqqer3e, “B” is Bloqqr, “C” is hiqqer3p, and “D” is squeezebf.
Instances solved by time-limited incremental preprocessing in six rounds with various execution sequences (upper part) and individual preprocessing (lower part).

Showcase: Preprocessing (3/4)

- Left: eval2012r2 preprocessed with execution sequence AABBCCDD in six rounds.
 - 111 solved in preprocessing, total solved instances (best solver) of 345: 203.
- Right: eval2012r2 preprocessed with execution sequence AADDBBCC in six rounds.
 - 104 solved in preprocessing, total solved instances (best solver) of 345: 208.



squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqqr3	□	sDual_oog	▽
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qupup	◊

Observations:

- Individual preprocessors show diverse performance.
- Incremental preprocessing is more powerful than individual one.
- Vast majority of instances is solved in the first round.
- Running more than three rounds of preprocessing yields only minor improvements.
- Preliminary experiments with solving: solvers are sensitive to the execution sequence of preprocessors.
- The execution sequence which solved the largest number of benchmarks does not allow solvers to perform best on preprocessed but unsolved benchmarks.

Showcase Solving

- Probably the track closest to traditional competition organization.
- Evaluation could have done in competition style ...
- ...but we were not interested in crowning a winner, instead we considered:
 - the sensitivity of solvers to the benchmarks.
 - the impact of preprocessing.
 - the quality of some benchmarks.

Considered Sets

eval2012r2:

- 345 formulas
- 69 solved by preprocessor Bloqqer
- families: ≈ 60 (major selection criterion: equal representation of families)
- avg. no. equivalences: 580
(after bloqqer: 0)
- avg. number of implications: 2305
(after bloqqer: 2161)

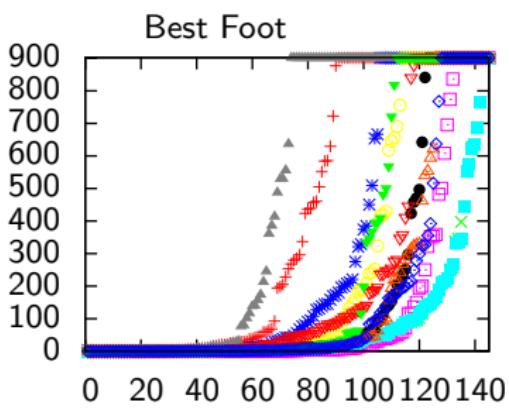
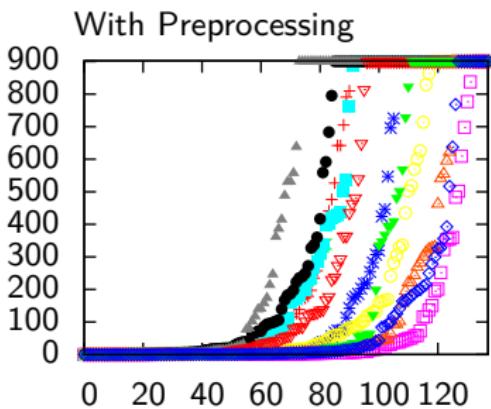
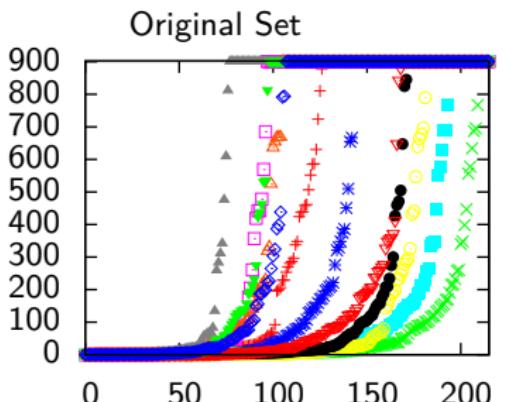
eval10:

- 568 formulas
- 148 solved by preprocessor Bloqqer
- families: ≈ 70 (Problem: some families have more formulas than others)
- avg. no. equivalences: 1278
(after bloqqer: 0)
- avg. number of implications: 4510
(after bloqqer: 854)

Thanks to A. Goultiaeva and W. Klieber for their tools to count gates in CNF formulas.

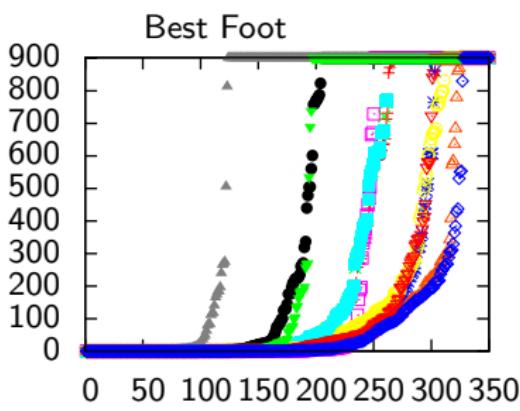
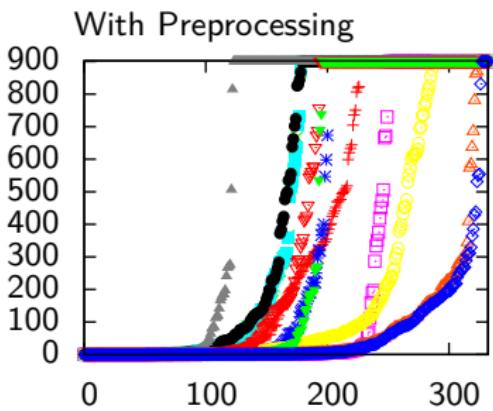
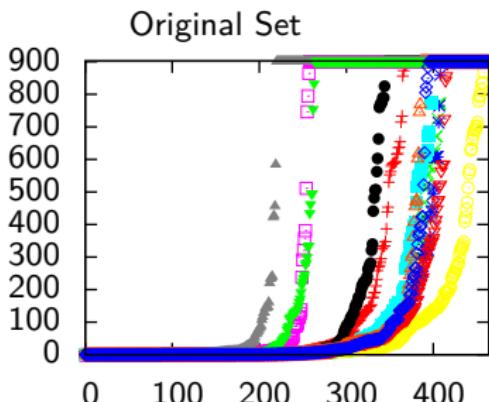
Showcase Solving + Preprocessing (1/2)

- Question: Is preprocessing always beneficial?
- Best foot evaluation (virtual experiment): solver chooses whether to use Bloqqer.



Showcase Solving + Preprocessing (2/2)

- Question: Is preprocessing always beneficial?
- Best foot evaluation (virtual experiment): solver chooses whether to use Bloqqer.



Showcase Solving (Sampling Experiment)

- Goal: evaluate the quality of benchmark set.
- 7 trials, 8 solvers, 1900 different formulas.
- Timeout per formula: 200s.
- Anonymized solver names.
- Stratification: 6 samples per family.
- Future work:
 - Longer timeouts, different sampling.

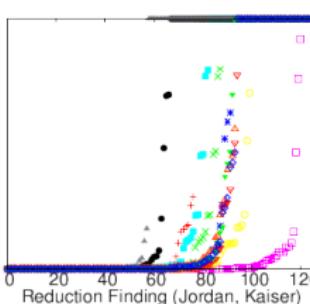
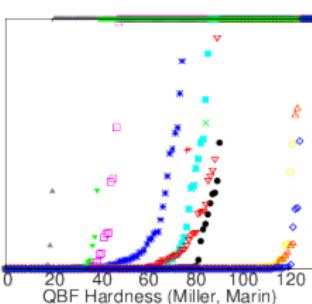
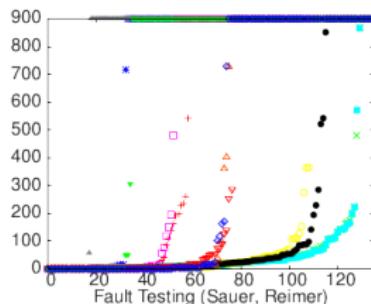
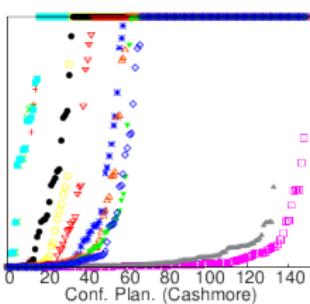
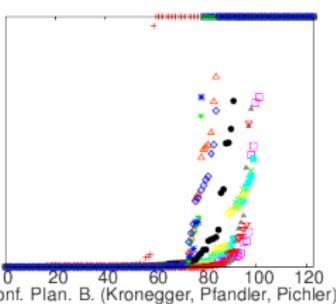
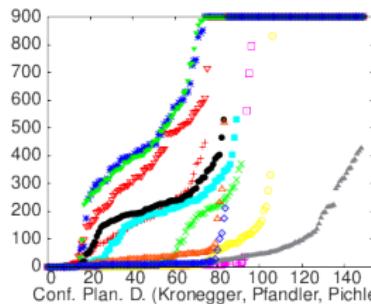
Seven trials ranked by number solved (out of 455)

BB	312	BB	317	BB	316	BB	315	BB	315	BB	310	BB	319
HH	293	HH	298	HH	300	HH	298	HH	291	HH	290	HH	299
AA	284	AA	285	AA	274	AA	272	AA	279	AA	275	AA	284
CC	274	CC	270	CC	265	CC	270	CC	270	CC	273	CC	274
DD	241	DD	237	DD	236	DD	242	DD	232	DD	241	DD	239
EE	211	EE	211	EE	205	EE	216	EE	208	EE	215	EE	211
GG	204	GG	201	GG	200	GG	202	GG	194	GG	194	GG	207
FF	159	FF	162	FF	161	FF	162	FF	159	FF	156	FF	171

Seven trials ranked by PAR10 score on 455 instances

BB	289789	BB	279394	BB	281633	BB	284161	BB	283213	BB	293765	BB	275484
HH	329163	HH	317815	HH	315260	HH	319013	HH	331889	HH	333963	HH	316156
AA	345379	AA	342510	AA	365017	AA	368546	AA	355440	AA	362225	AA	344501
CC	363203	CC	370937	CC	381240	CC	371138	CC	370969	CC	365632	CC	363675
DD	430920	DD	438119	DD	440782	DD	428847	DD	448715	DD	430942	DD	434698
EE	490114	EE	489906	EE	502105	EE	480172	EE	495984	EE	482456	EE	490329
GG	504751	GG	510314	GG	513311	GG	509006	GG	524914	GG	524489	GG	498584
FF	593122	FF	587675	FF	589655	FF	587346	FF	593564	FF	599697	FF	569816

Showcase Applications (+Solving)



- 6 formula sets with 150 formulas each.
- Not in QBFLIB.
- 900s timeout, 7GB memory limit.

Observation:
For each set there is at least one good solver.

Showcase: Certificates

Note:

- Only one proof-producing solver and two certification tools were officially submitted.
- Additionally, we ran a subset of publicly available tools of non-participating authors.
- 600 seconds time limit, 3 GB memory limit, different hardware than in other showcases.

Workflow	Solved	Certified
<i>Submitted Tools</i>		
DepQBF and QBFcert	91 (34 s, 57 u)	67 (20 s, 47 u)
DepQBF and ResQu ¹	91 (34 s, 57 u)	63 (22 s, 41 u)
<i>Non-Submitted Tools</i>		
sKizzo and ozziKs	88 (36 s, 52 u)	35 ² (35 s, 0 u)
squolem and qbv	38 (19 s, 19 u)	38 (19 s, 19 u)
squolem and ResQu	38 (19 s, 19 u)	19 (0 ³ s, 19 u)
qube-cert and checker	80 (25 s, 55 u)	32 (11 s, 21 u)
qube-cert and ResQu	80 (25 s, 55 u)	52 (17 s, 35 u)

¹Workflow involves proof format conversion.

²ozziKs is designed to verify satisfiable instances only.

³Format errors reported on 19 instances.

Lessons Learned

- If this had been a competition, there would not be a clear winner.
- Preprocessing strongly influences solving.
- Preprocessors are powerful (but incomplete) solvers.
- QBF solvers are not blackboxes, some use built-in preprocessing.
- Benchmark selection and scoring methods strongly influence rankings.
- Community-driven organization is challenging, but fruitful.

What's next?

- More analysis of the available data.
- Establish fair benchmark sets for competitions and evaluations.
- More emphasis on special tracks (formulas needed!)
- Tighter integration of certificate generation.
- Common standards for input formats and testing workflows.

Thanks to the Contributors !!!!!!

- **Solvers:** S. Bayless, A. Goultiaeva, M. Janota, W. Klieber, F. Lonsing, M. Narizzano, A. Van Gelder
- **Preprocessors:** A. Biere, M. Narizzano, M. Seidl, A. Van Gelder
- **Certificates:** V. Balabanov, J.R. Jiang, A. Niemetz, M. Preiner
- **Applications:** M. Cashmore, L. Kaiser, M. Kronegger, C. Jordan, P. Marin, A. Pfandler, R. Pichler, S. Reimer, M. Sauer



<http://www.kr.tuwien.ac.at/events/qbfgallery2013/>

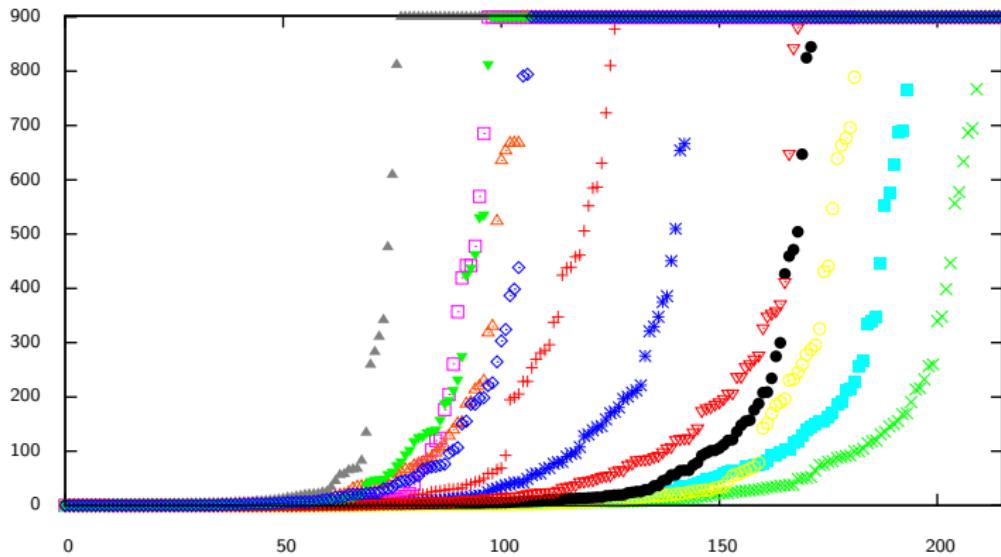
- Olympic Games @ FLoC 2014 in Vienna.
- QBF Workshop 2014.
- How to organize competitions, evaluations, galleries?
- How to select good (fair, expressive) benchmark sets?
- Where shall we store collected data and tools? How to ensure provenance?

Showcase: Solving (1/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
bGhost-CEGAR	210	111	99	135	0	369	127592
ghost-CEGAR	194	103	91	151	0	412	142191
hiqqer3	182	93	89	163	6	452	156058
ghost	172	87	85	173	2	456	157587
sDual_oq	169	80	89	176	5	484	167314
dual_oq	143	66	77	202	0	550	189840
squeezebf1.2-qube3.0	127	60	67	218	0	552	190569
depqbf-lazy-qpup	107	43	64	238	0	640	220919
DepQBF	105	42	63	240	0	648	223714
oq	98	34	64	247	0	662	228450
rareqs	97	34	63	248	5	649	224207
nenofex	77	34	43	268	5	295	102014

Table: QBF Eval 2012

Showcase: Solving (2/12)



squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqer3	○	sDual_oog	▼
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qupup	◊

Showcase: Solving (3/12)

	squeezeblf1.2-qube3.0	bGhost-CEGAR	dual_oq	rareqs	ghost-CEGAR	hiqqer3	ghost	DepQBF	nenoex	sDual_oq	oq	depqbf-lazy-qupup
Ansotegui (20)	11	5	6	6	5	10	7	10	0	11	10	10
Ayari (12)	6	5	6	5	0	7	0	0	6	6	4	0
Basler (18)	3	13	1	0	13	6	4	0	0	8	0	0
Biere (16)	5	14	9	1	14	5	14	3	3	6	3	3
gelder (12)	9	12	11	6	12	12	12	11	3	11	10	11
Gent-Rowley (4)	1	1	1	0	1	1	1	1	0	1	1	1
Herbsttritt (12)	10	9	10	7	9	9	10	6	1	10	6	6
jiang (12)	2	5	6	1	5	6	5	5	1	6	6	4
Katz (12)	2	0	0	0	0	2	0	0	1	3	0	0
Kontchakov (18)	12	8	16	0	8	18	0	18	0	17	1	18
Lahiri-Seshia (3)	0	2	0	0	2	0	2	0	0	0	0	0
Letombe (6)	6	6	6	6	6	6	6	6	2	6	6	6
Ling (2)	0	2	2	2	2	2	1	2	2	2	2	2
Mangassarian-Veneris (11)	2	4	4	6	4	6	3	5	7	4	4	6
Messinger (6)	0	0	0	0	0	0	0	0	0	0	0	0
Miller-Marin (18)	14	15	12	17	15	17	13	15	12	16	12	16
Mneimneh-Sakallah (38)	15	31	18	0	31	9	31	0	0	16	0	0
Palacios (16)	2	9	5	14	9	5	4	5	9	5	5	4
Pan (48)	10	33	19	6	30	39	30	8	15	22	18	8
Rintanen (12)	2	8	6	10	8	6	8	7	10	5	6	9
sauer_reimer (6)	2	6	2	2	6	4	5	1	0	2	2	1
Scholl-Becker (25)	2	13	3	8	13	3	15	2	4	4	2	2
Wintersteiger (18)	11	9	0	0	1	9	1	0	1	8	0	0
total (345)	127	210	143	97	194	182	172	105	77	169	98	107

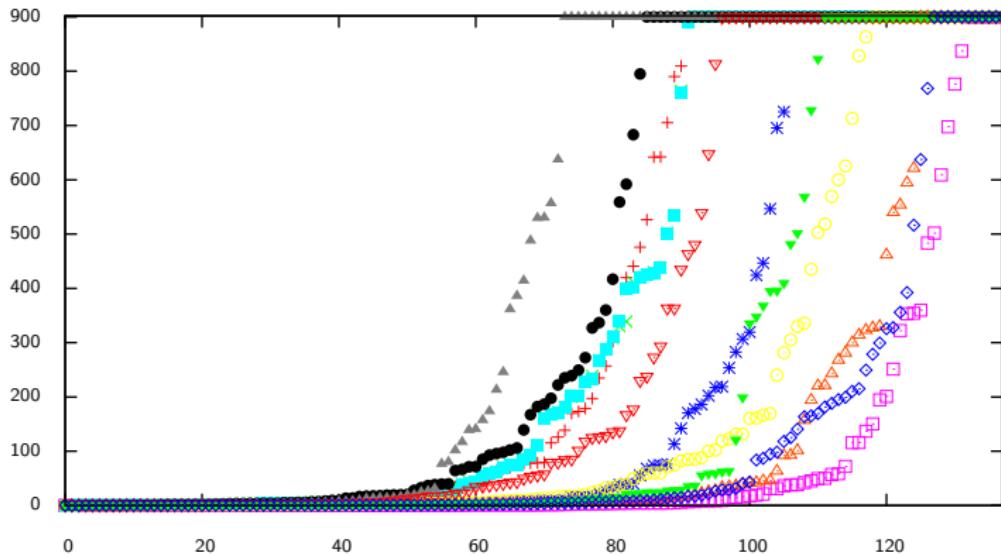
Table: eval12

Showcase: Solving (4/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
rareqs	132	66	66	144	7	492	136060
depqbf-lazy-qpup	127	66	61	149	0	511	141191
DepQBF	125	66	59	151	0	517	142863
hiqqer3	118	59	59	158	3	550	151970
ooq	111	58	53	165	3	560	154815
dual_ooq	106	57	49	170	2	570	157363
sDual_ooq	97	54	43	179	0	610	168481
ghost-CEGAR	92	54	38	184	0	599	165391
bGhost-CEGAR	92	54	38	184	0	599	165523
squeezebf1.2-qube3.0	91	52	39	185	0	592	163431
ghost	85	50	35	191	0	639	176438
nenofex	73	39	34	203	8	330	91303

Table: QBF Eval 2012 with bloqqer

Showcase: Solving (5/12)



squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoex	▲
bGhost-CEGAR	X	hiqer3	○	sDual_oog	▽
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qupup	◊

Showcase: Solving (6/12)

	squeezebf1.2-qube3.0	bGhost-CEGAR	dual_oocq	rateqs	ghost-CEGAR	hqqr3	ghost	DepQBF	nenoex	sDual_oocq	oocq	depqbf-lazy-qupup
Ansotegui (20)	12	9	6	7	9	14	11	12	2	7	13	12
Ayari (6)	0	0	0	0	0	1	0	0	4	0	0	0
Basler (12)	0	0	0	3	0	1	0	0	0	0	5	0
Biere (14)	2	3	3	5	3	3	3	3	3	3	4	4
gelder (6)	6	6	6	6	6	6	6	6	6	6	6	6
Gent-Rowley (4)	1	1	1	0	1	1	1	1	0	1	1	1
Herbsttritt (11)	7	0	7	11	0	8	0	9	0	6	8	9
jiang (12)	3	6	6	5	6	6	6	4	2	6	6	6
Katz (12)	1	1	3	0	1	2	2	3	1	3	2	3
Kontchakov (18)	9	2	2	1	2	12	3	18	0	1	1	18
Lahiri-Seshia (3)	0	0	0	0	0	0	0	0	0	0	0	0
Letombe (6)	5	6	6	6	6	6	6	6	1	6	6	6
Ling (2)	2	2	2	2	2	2	2	2	2	2	2	2
Mangassarian-Veneris (10)	3	2	4	5	2	5	3	5	9	4	4	5
Messinger (6)	0	0	0	0	0	0	0	0	0	0	0	0
Miller-Marin (11)	8	11	9	11	11	10	9	9	7	9	9	9
Mneimneh-Sakallah (37)	14	14	30	26	14	9	6	16	3	24	19	15
Palacios (16)	2	5	5	11	5	5	5	5	10	5	5	5
Pan (20)	11	6	6	11	6	12	7	10	7	6	7	9
Rintanen (12)	2	9	4	10	9	8	7	8	10	3	7	9
sauer_reimer (4)	2	3	2	2	3	2	3	2	0	2	2	2
Scholl-Becker (25)	1	6	4	9	6	5	5	6	6	3	4	6
Wintersteiger (9)	0	0	0	1	0	0	0	0	0	0	0	0
total (276)	91	92	106	132	92	118	85	125	73	97	111	127

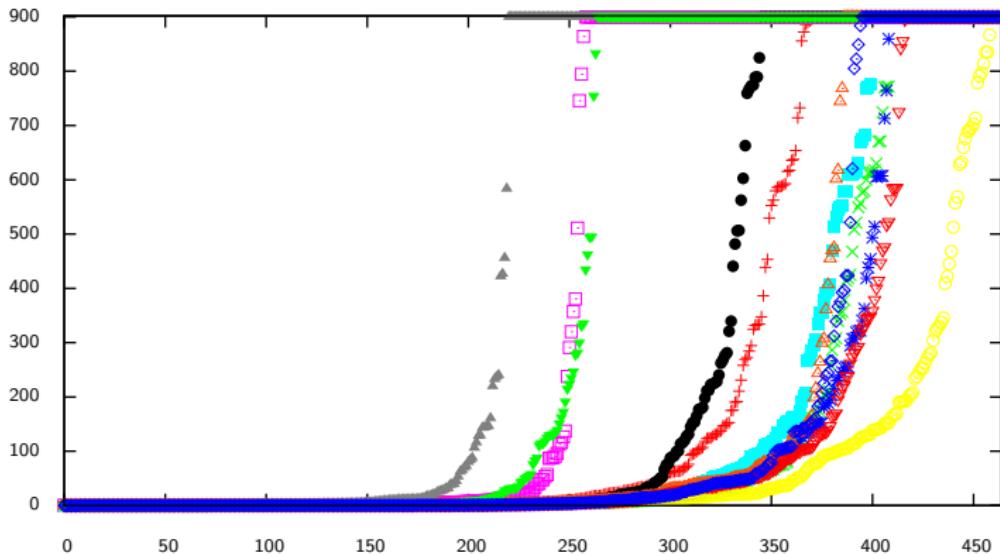
Table: Eval12 Round 2 with bloqqer

Showcase: Solving (7/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
hiqqr3	459	221	238	109	5	223	126741
sDual_ooc	417	190	227	151	0	272	154946
bGhost-CEGAR	409	195	214	159	0	286	162688
dual_ooc	409	185	224	159	0	269	152988
ghost-CEGAR	400	191	209	168	0	301	171482
depqbf-lazy-qpup	395	171	224	173	0	299	169849
DepQBF	386	167	219	182	0	308	175253
squeezebf1.2-qube3.0	371	161	210	197	0	335	190654
ghost	345	172	173	223	0	375	213280
ooc	264	99	165	304	0	488	277215
rareqs	258	96	162	310	3	499	283712
nenofex	220	106	114	348	7	192	109443

Table: QBF Eval 2010

Showcase: Solving (8/12)



squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqer3	○	sDual_oog	▼
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qupup	◊

Showcase: Solving (9/12)

	squeezef1.2-qub3.0	bGhost-CEGAR	dual_ooc	rareqs	ghost-CEGAR	hidqer3	ghost	DepQBF	nenofex	sDual_ooc	ooc_q	depqbflazy-qcup
Ansotegui (22)	13	5	7	7	5	13	7	11	0	12	10	12
Ayari (19)	4	5	6	4	1	9	1	2	14	6	3	2
Biere (42)	34	39	23	12	39	22	40	14	8	27	14	14
Castellini (37)	37	37	37	37	37	37	37	37	37	37	37	37
Gent-Rowley (11)	8	8	8	7	8	8	8	8	5	8	8	8
Herbstritt (61)	54	37	45	41	37	45	39	53	11	48	46	54
Katz (3)	3	0	0	0	0	3	0	0	1	3	0	0
Kontchakov (136)	89	75	133	19	75	134	22	136	0	120	13	136
Letombe (52)	50	50	50	50	50	50	51	49	41	51	50	50
Ling (3)	1	3	3	3	3	3	2	3	3	3	3	3
Mangassarian-Veneris (23)	6	14	13	17	13	16	8	12	17	12	13	13
Messinger (3)	1	2	1	2	2	2	1	2	2	1	2	2
Mneimneh-Sakallah (19)	14	18	16	0	18	10	18	0	2	13	0	0
Palacios (15)	3	10	5	14	10	5	4	5	9	5	5	5
Pan (80)	33	74	37	13	71	77	74	27	41	44	36	30
Rintanen (18)	9	16	13	16	15	13	16	16	18	13	13	17
Scholl-Becker (24)	12	16	12	16	16	12	17	11	11	14	11	12
total (568)	371	409	409	258	400	459	345	386	220	417	264	395

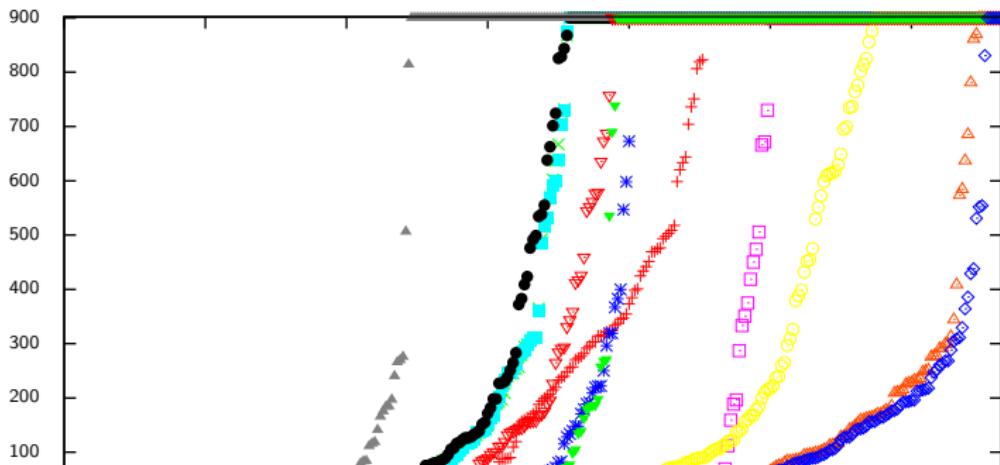
Table: QBF Eval 2010

Showcase: Solving (10/12)

solver	total	sat	unsat	to	unique	avg	sum
depqbf-lazy-qpup	327	164	163	93	0	237	99866
DepQBF	324	166	158	96	0	247	104151
hiqqer3	287	131	156	133	2	348	146534
rareqs	250	117	133	170	12	379	159381
squeezebf1.2-qube3.0	227	94	133	193	1	461	193857
dual_oog	201	96	105	219	0	482	202455
oog	196	97	99	224	0	492	206703
sDual_oog	194	97	97	226	0	514	215988
ghost	179	90	89	241	0	553	232587
ghost-CEGAR	179	103	76	241	0	436	183140
bGhost-CEGAR	179	103	76	241	0	436	183437
nenofex	123	73	50	297	7	322	135495

Table: QBF Eval 2010 with bloqqer

Showcase: Solving (11/12)



squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqqr3	○	sDual_oog	▼
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qupup	◊

Showcase: Solving (12/12)

	squeezef1.2-qub3.0	bGhost-CEGAR	dual_oog	rareqs	ghost-CEGAR	hqqr3	ghost	DepQBF	nenoex	sDual_oog	oog	depqbf-lazy-qup
Ansotegui (22)	13	12	7	16	12	14	14	17	3	7	11	17
Ayari (13)	1	1	1	2	1	2	1	1	6	1	1	2
Biere (29)	10	8	8	13	8	13	9	11	5	9	11	12
Castellini (27)	27	27	27	27	27	27	27	27	27	27	27	27
Gent-Rowley (4)	2	1	1	0	1	1	1	1	0	1	1	1
Herbstritt (50)	37	10	38	50	10	39	11	35	1	36	38	37
Katz (3)	3	1	2	1	1	3	2	3	1	2	2	3
Kontchakov (136)	57	23	25	26	23	101	32	136	1	19	14	136
Letombe (29)	27	26	27	28	26	25	27	26	9	27	26	26
Ling (3)	3	3	3	3	3	3	3	3	3	3	3	3
Mangassarian-Veneris (17)	2	9	10	12	9	10	7	10	14	9	10	9
Messinger (3)	1	2	2	3	2	2	1	2	3	2	2	2
Mneimneh-Sakallah (17)	13	11	16	15	11	8	6	12	5	16	14	12
Palacios (14)	2	7	4	11	7	4	5	4	7	4	4	4
Pan (15)	12	9	9	11	9	12	10	11	9	9	9	11
Rintanen (14)	6	13	8	14	13	10	9	11	14	9	10	10
Scholl-Becker (24)	11	16	13	18	16	13	14	14	15	13	13	15
total (420)	227	179	201	250	179	287	179	324	123	194	196	327

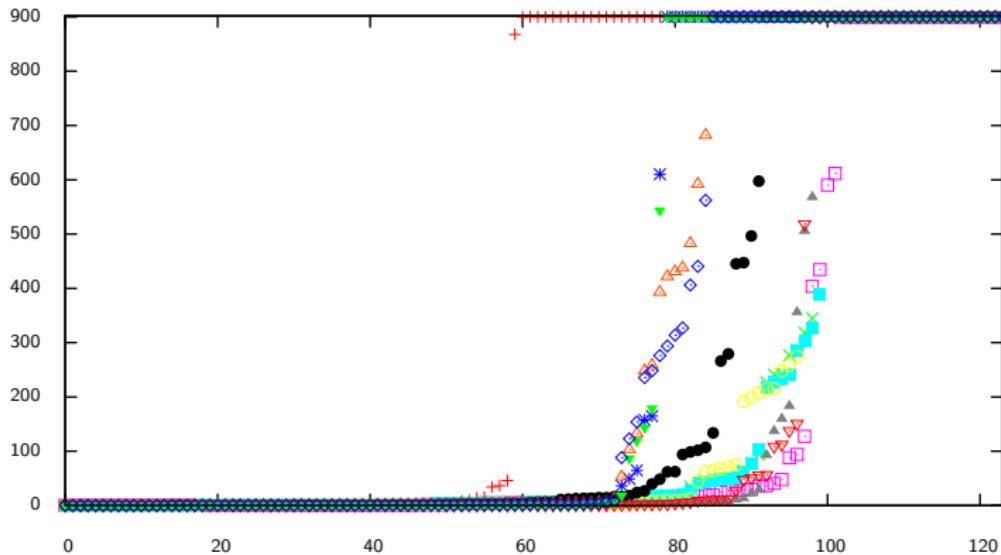
Table: QBF Eval10 with bloqqer

Showcase: Applications (1/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
rareqs	102	37	65	48	5	227	34196
ghost-CEGAR	100	50	50	50	0	303	45466
bGhost-CEGAR	100	50	50	50	0	305	45808
nenoFex	99	56	43	51	0	320	48080
sDual_ooq	98	56	42	52	0	321	48207
hiqqer3	97	56	41	53	0	333	50059
ghost	92	51	41	58	0	353	53074
DepQBF	85	43	42	65	0	419	62868
depqbf-lazy-qpup	85	43	42	65	0	413	62042
ooq	79	37	42	71	0	433	65046
dual_ooq	79	37	42	71	0	433	65082
squeezebf1.2-qube3.0	60	24	36	90	0	547	82073

Table: conf_planning_bomb

Showcase: Applications (2/12)



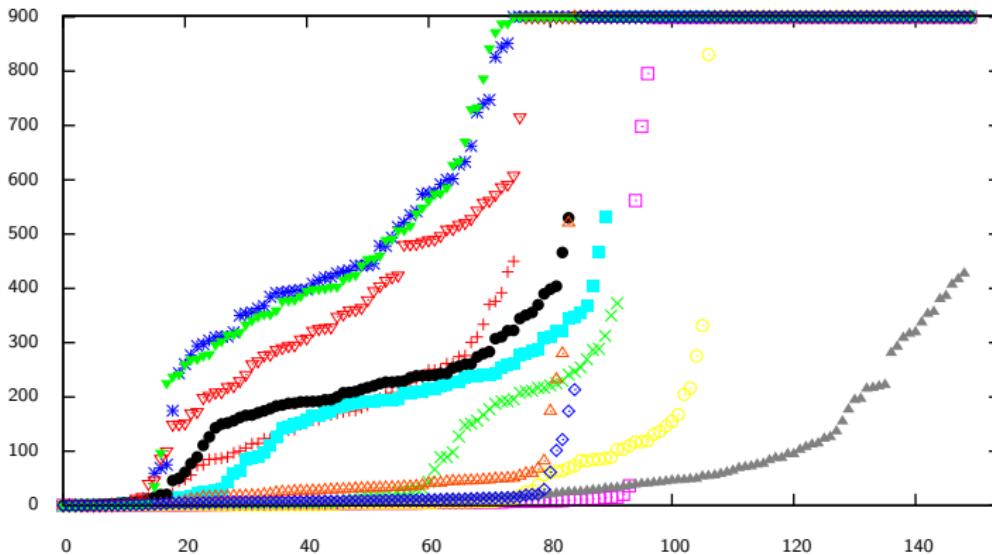
squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqqr3	○	sDual_oog	▼
dual_oog	*	ghost	●	oog	▽
rareqs	□	DepQBF	△	depqbf-lazy-qpup	◊

Showcase: Applications (3/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
nenofex	149	18	131	1	36	69	10472
hiqqer3	107	18	89	43	0	287	43182
rareqs	97	18	79	53	0	334	50244
bGhost-CEGAR	92	17	75	58	0	393	58952
ghost-CEGAR	90	17	73	60	0	454	68152
depqbf-lazy-qpup	85	18	67	65	0	398	59813
DepQBF	84	17	67	66	0	417	62651
ghost	84	16	68	66	0	497	74627
sDual_ooq	76	9	67	74	0	561	84226
squeezebf1.2-qube3.0	75	6	69	75	0	521	78262
ooq	74	8	66	76	0	617	92668
dual_ooq	74	8	66	76	0	615	92281

Table: conf_planning_dungeon

Showcase: Applications (4/12)



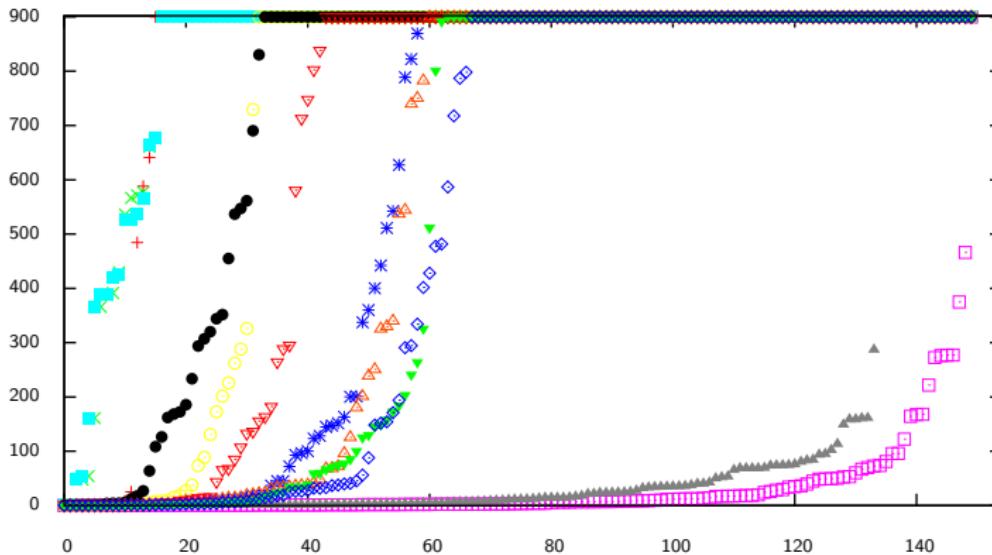
squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqer3	□	sDual_oog	▼
dual_oog	×	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qupup	◊

Showcase: Applications (5/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
rareqs	149	40	109	1	1	36	5411
nenofex	134	39	95	16	0	30	4553
depqbf-lazy-qpup	67	31	36	83	0	545	81833
ooq	63	29	34	87	0	556	83540
DepQBF	60	30	30	90	0	581	87176
dual_ooq	59	26	33	91	0	596	89541
sDual_ooq	43	27	16	107	0	680	102067
ghost	33	19	14	117	0	740	111139
hiqqer3	32	17	15	118	0	725	108840
bGhost-CEGAR	16	9	7	134	0	836	125461
ghost-CEGAR	16	9	7	134	0	837	125676
squeezebf1.2-qube3.0	15	10	5	135	0	821	123277

Table: planning_CTE

Showcase: Applications (6/12)



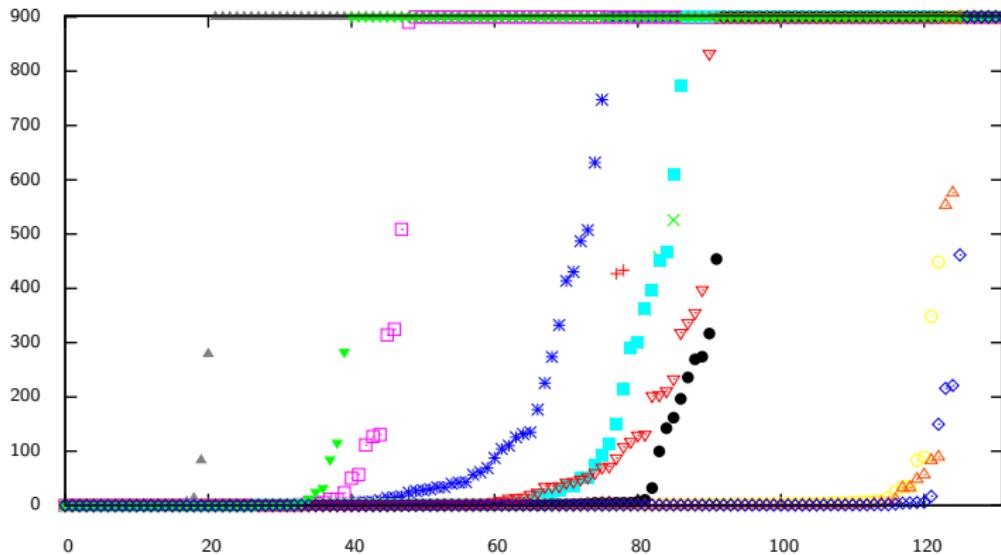
squeezebf1.2-qube3.0	+	ghost-CEGAR	ghost	nenoex	▲
bGhost-CEGAR	×	hiqer3	●	sDual_oog	▼
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qpup	◊

Showcase: Applications (7/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
depqbf-lazy-qpup	126	8	118	24	3	151	22748
DepQBF	125	8	117	25	2	160	24070
hiqqer3	123	11	112	27	2	170	25538
ghost	92	10	82	58	0	363	54497
sDual_oog	91	8	83	59	0	382	57433
ghost-CEGAR	87	10	77	63	0	408	61319
bGhost-CEGAR	86	10	76	64	0	409	61364
squeezebf1.2-qube3.0	79	8	71	71	0	163	24495
dual_oog	76	8	68	74	0	480	72134
rareqs	49	8	41	101	0	623	93503
oog	40	8	32	110	0	663	99585
nenofex	21	8	13	129	0	79	11889

Table: qbf_hardness

Showcase: Applications (8/12)

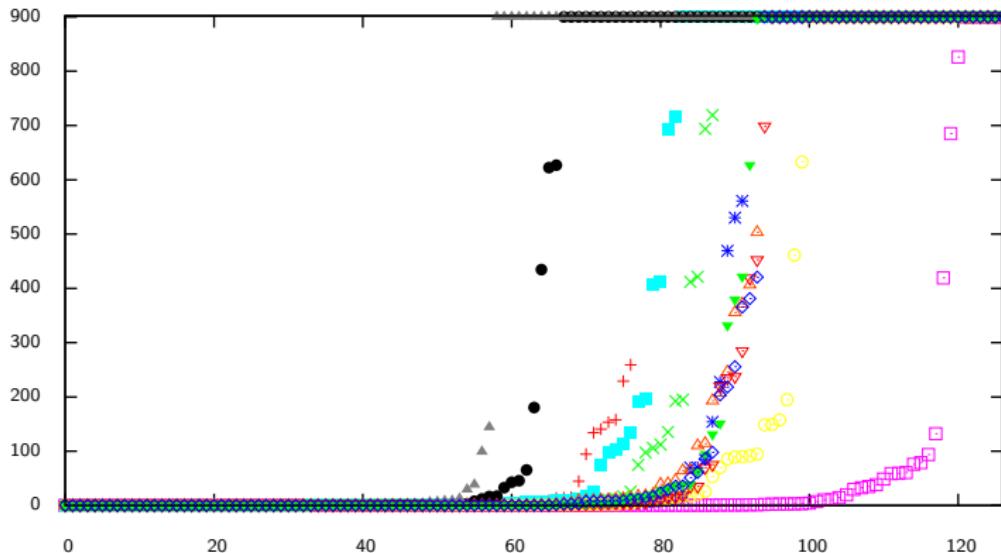


Showcase: Applications (9/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
rareqs	121	53	68	29	10	192	28908
hiqqer3	100	50	50	50	0	316	47507
sDual_ooc	95	44	51	55	0	349	52409
DepQBF	94	45	49	56	0	355	53358
depqbf-lazy-qpup	94	46	48	56	0	352	52811
ooc	94	42	52	56	1	340	51114
dual_ooc	92	41	51	58	0	345	51797
bGhost-CEGAR	88	35	53	62	0	394	59123
ghost-CEGAR	83	31	52	67	0	424	63619
squeezebf1.2-qube3.0	77	42	35	73	0	446	67008
ghost	67	32	35	83	0	512	76851
nenofex	58	29	29	92	0	220	33115

Table: red_finding

Showcase: Applications (10/12)



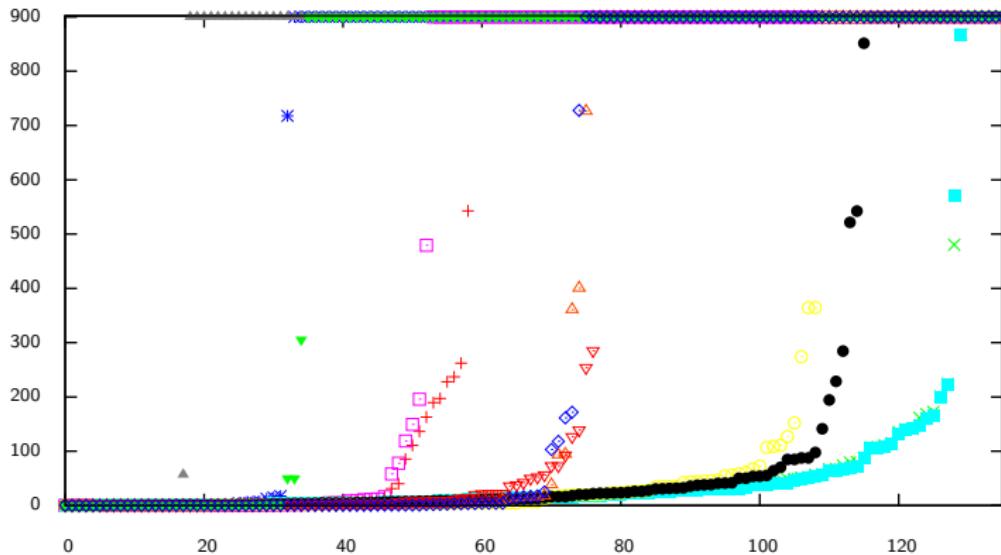
squeezebf1.2-qube3.0	+	ghost-CEGAR	■	nenoFex	▲
bGhost-CEGAR	×	hiqqr3	○	sDual_oog	▼
dual_oog	*	ghost	●	oog	▼
rareqs	□	DepQBF	△	depqbf-lazy-qup	◊

Showcase: Applications (11/12)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
ghost-CEGAR	130	102	28	20	0	95	14335
bGhost-CEGAR	130	102	28	20	0	96	14478
ghost	116	89	27	34	0	176	26497
hiqqr3	109	79	30	41	3	265	39774
sDual_ooc	77	48	29	73	0	398	59724
DepQBF	76	51	25	74	0	456	68498
depqbf-lazy-qupup	75	51	24	75	0	459	68946
squeezebf1.2-qube3.0	59	33	26	91	0	206	30960
rareqs	53	27	26	97	0	407	61118
ooc	35	6	29	115	0	693	103950
dual_ooc	33	5	28	117	0	707	106130
nenofex	18	7	11	132	0	479	71917

Table: sauer_reimer

Showcase: Applications (12/12)



<code>squeezebf1.2-qube3.0</code>	<code>+</code>	<code>ghost-CEGAR</code>	<code>ghost</code>	<code>nenoFEX</code>	<code>▲</code>
<code>bGhost-CEGAR</code>	<code>×</code>	<code>hiqqr3</code>	<code>●</code>	<code>sDual_oog</code>	<code>▼</code>
<code>dual_oog</code>	<code>*</code>	<code>ghost</code>	<code>○</code>	<code>oog</code>	<code>▼</code>
<code>rareqs</code>	<code>□</code>	<code>DepQBF</code>	<code>△</code>	<code>depqbf-lazy-qupup</code>	<code>◊</code>

Showcase: Preprocessing (1/3)

	eval2012r2												
	1	2	3	4	5	6	7	8	9	10	11	12	Σ
$(ABCD)^{12}$ fixpoints	3	25	62	64	38	5	4	2	0	2	0	1	206
$(ABCD)^{12}$ solved	93	9	3	1	1	1	0	0	0	0	0	0	108

“A” is hiqqr3e, “B” is Bloqqer, “C” is hiqqr3p, and “D” is squeezebf.

12 rounds of time-limited incremental preprocessing. Numbers of solved instances and detected fixpoints. Running 24 rounds does not change the results: no instance solved and no fixpoint detected in rounds 13-24.

	Fixpoints per round: eval2012r2						
	1	2	3	4	5	6	Σ
$(A^2B^2C^2D^2)^6$	2	52	76	60	14	2	206
$(A^2B^2D^2C^2)^6$	1	16	113	58	16	3	207
$(A^2D^2B^2C^2)^6$	0	16	137	53	12	0	218
$(B^2C^2D^2A^2)^6$	2	46	81	66	14	2	211
$(D^2A^2B^2C^2)^6$	2	12	141	49	10	4	218

	Solved per round: eval2012r2						
	1	2	3	4	5	6	Σ
$(A^2B^2C^2D^2)^6$	100	8	2	1	0	0	111
$(A^2B^2D^2C^2)^6$	101	7	2	1	0	0	111
$(A^2D^2B^2C^2)^6$	96	8	0	0	0	0	104
$(B^2C^2D^2A^2)^6$	96	5	1	1	0	0	103
$(D^2A^2B^2C^2)^6$	93	8	1	0	0	0	102

Six rounds of time-limited incremental preprocessing. Numbers of solved instances and detected fixpoints in each round.

Showcase: Preprocessing (2/3)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
depqbf-lazy-qpup	104	59	45	137	0	548	132263
rareqs	104	59	45	137	7	533	128489
DepQBF	102	59	43	139	0	556	134010
hiqqer3	90	52	38	151	0	600	144827
ooq	90	56	34	151	2	586	141278
dual_ooq	80	54	26	161	0	624	150491
sDual_ooq	64	43	21	177	0	684	164969
ghost	61	43	18	180	0	683	164641
bGhost-CEGAR	59	40	19	182	0	678	163438
ghost-CEGAR	59	40	19	182	0	677	163345
squeezebf1.2-qube3.0	58	37	21	183	0	657	158500
nenofex	52	34	18	189	3	352	84859

Table: AADDBBCC_eval12 (241 formulas)

Showcase: Preprocessing (3/3)

solver	number of solved formulas					runtime (sec)	
	total	sat	unsat	to	unique	avg	sum
rareqs	92	57	35	142	8	554	129819
depqbf-lazy-qpup	90	52	38	144	0	587	137459
DepQBF	87	50	37	147	0	601	140857
ooq	79	49	30	155	4	622	145581
hiqqer3	78	45	33	156	0	631	147679
dual_ooq	74	48	26	160	0	636	149013
bGhost-CEGAR	64	46	18	170	0	667	156217
ghost-CEGAR	64	46	18	170	0	665	155745
ghost	57	43	14	177	0	679	159017
sDual_ooq	54	37	17	180	0	713	166865
nenoFEX	48	32	16	186	3	343	80278
squeezebf1.2-qube3.0	46	33	13	188	1	696	163057

Table: AABBCCDD_eval12 (234 formulas)