

# The QBF Gallery 2013

A Non-Competitive Evaluation of QBF Tools

Florian Lonsing<sup>1</sup>   Martina Seidl<sup>2</sup>   Allen Van Gelder<sup>3</sup>

<sup>1</sup>Vienna University of Technology

<http://www.kr.tuwien.ac.at/staff/lonsing>

<sup>2</sup>Johannes Kepler University, Linz, Austria

<http://fmv.jku.at/seidl>

<sup>3</sup>University 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.

**No competition, no winners (nor losers!), no prizes**



**Goal:**

- Evaluate the state-of-the-art in practical QBF research...
- ...by running QBF tools (any kind!) in an organized and centralized manner...
- ...and by collecting and evaluating data...
- ...in a community-driven manner with interaction / intervention opportunities during the runs.

### Organizational Details

- 1st QBF Gallery ever (complementary to biannual QBFEval competition).
- 4 (strongly related) showcases: Preprocessing, Solving, Applications, Certificates.
- 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).

### Submissions:

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

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

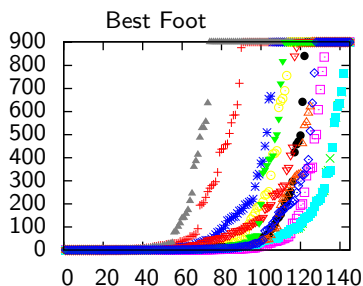
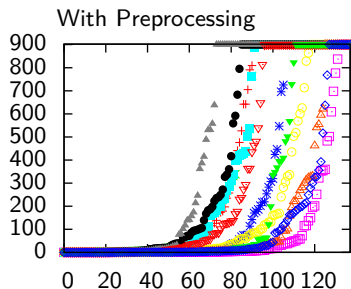
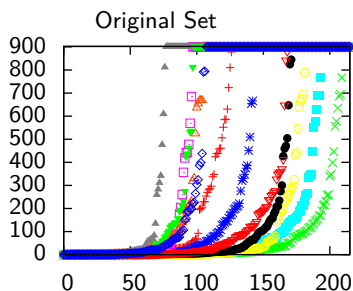
- Comparison of individual preprocessors and combinations.
- Evaluation of solving power of preprocessors.
- Time-limited preprocessing
  - ... in multiple rounds
  - ... with different execution sequences
  - ... and fixpoint detection.
- Effects of preprocessing on solver performance.

	hiqqr3e			Bloqqr			hiqqr3p			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

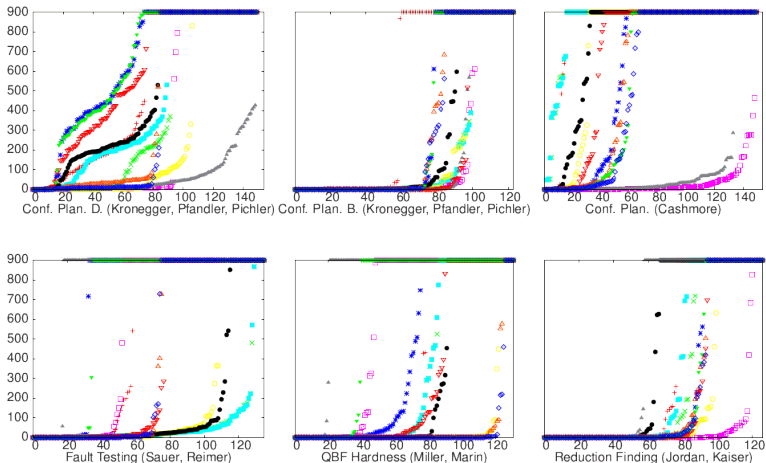
Individual preprocessors: solved instances ( $t$ ), solved satisfiable ( $s$ ) and solved unsatisfiable instances ( $u$ ).

## Excerpt of Showcase Solving (+ Preprocessing)

- 345 formulas.
- 69 solved by preprocessor Bloqqer.
- Solvers run on the remaining 276 formulas.
- Question: is preprocessing always beneficial?
- Best foot evaluation (virtual experiment): let solvers choose whether to use Bloqqer.



## Excerpt of Showcase Applications (+Solving)



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

### Observation:

At least one solver is good for one set (but it is not always the same!).

- Small, but very important showcase:
  - Only one solver and two tool suites submitted.
  - Urgently needed for practical applications.
- Additional experiments with publicly available tools not submitted by their authors.
- Requirements:
  - Need for standard proof formats and checkers.
  - More proof generating solvers.
  - Proof compression techniques.
  - Support from preprocessors.

	eval2012r2	
Workflow	Solved	Certified
DepQBF and QBFcert	91 (34 s, 57 u)	67 (20 s, 47 u)
DepQBF and ResQu <sup>1</sup>	91 (34 s, 57 u)	63 (22 s, 41 u)

---

<sup>1</sup>Workflow involves proof format conversion.

## 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.

## More details in the Poster Session



## 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, M. Reimer, S. Sauer



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