# The 4th Reactive Synthesis Competition - SYNTCOMP 2017

Swen Jacobs
Saarland University
Reactive Systems

Roderick Bloem
TU Graz
SCOS
Systematic Construction of Correct Systems

22 July 2017 – SYNT Workshop, Heidelberg

# *SYNTCOMP: Goals*

Make reactive synthesis tools comparable:

- establish **benchmark format**
- collect **benchmark library**
- provide platform for **fair and comprehensive evaluation**

Guide development of reactive synthesis tools:

- encourage implementation of mature, **push-button tools**
- improve state of the art through **challenging benchmarks**

# *SYNTCOMP: History*

**First Call**: 2013 (discussion at SYNT 2013, St. Petersburg)

**Design Choices**:

- low entry-barrier → only safety properties, low-level format (AIGER)
- need verifiable correctness → hardware model checkers
- output quality is important → ranking based on solution size

**First Competition**: 2014 (SYNT/FLoC, Vienna Summer of Logic)

5 participating groups, >500 benchmarks collected

**Second Competition**: 2015 (SYNT/CAV, San Francisco)

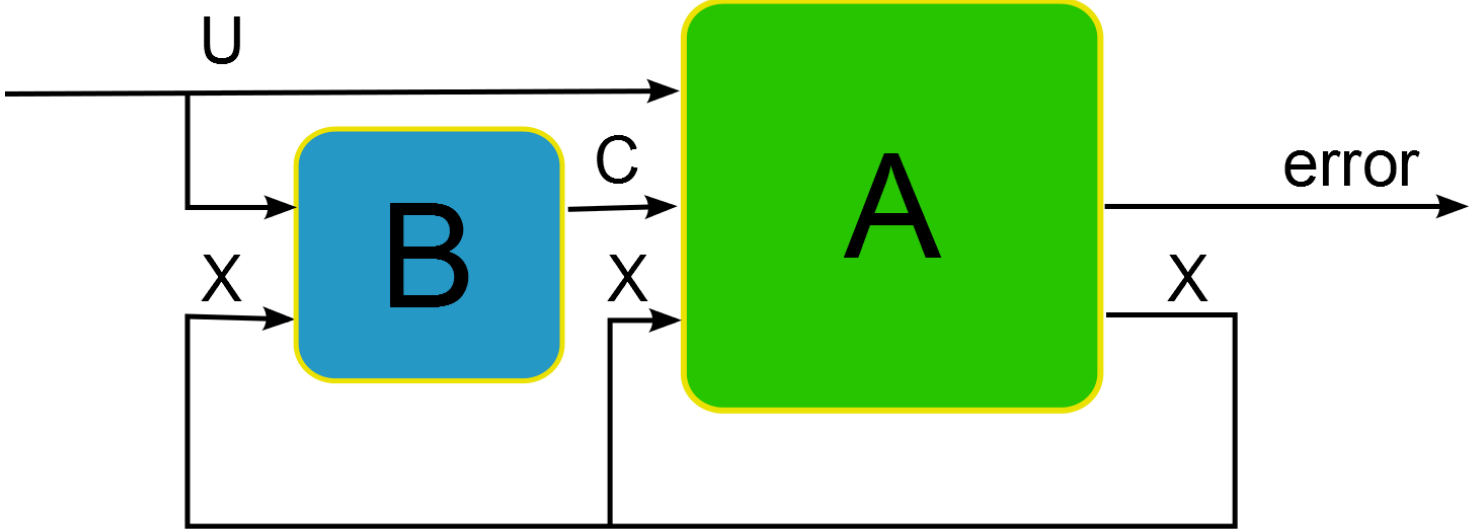
essentially same setup, >2000 benchmarks collected, comparison to 2014 tools

**Third Competition**: 2016 (SYNT/CAV, Toronto)

**extension to LTL specs** in TLSF, parameterized benchmarks, 6 new tools (3 safety, 3 LTL)

# *AIGER-based Safety Track of SYNTCOMP*

- **synthesis problem** defined by AIGER circuit A, with controllable (C) and uncontrollable (U) inputs, and single output error
- **solution** of synthesis problem is an AIG that includes original AIG A, and adds control structure B for inputs C such that resulting system never raises error



# *TLSF-based LTL Track of SYNTCOMP*

- **synthesis problem** defined by TLSF specification (LTL formula, input/output signals, meta-information and parameterization)
- **solution** of synthesis problem is an AIG
- **model checking** by combination with an AIG for the spec

# *Lessons Learned from Previous Competitions*

- SYNTCOMP with verified results is (in principle) feasible
- verification not always easy
  ⇒ safety tools can give winning region (since 2016)

- SYNTCOMP is well-received (8 safety tools, 3 LTL tools until 2016)

- safety tools improved significantly from 2014 to 2016

# *Mostly Stable Rules for SYNTCOMP 2017*

- No new tracks
- No new solver features
- Quality ranking re-introduced. New scheme:
  - every benchmark has a reference size $ref$
    ($\sim$best known solution, measured in gates of AIG)
  - if solver returns a correct solution of size $n$, then it gains

$$2 - \log_{10} \frac{n+1}{ref+1}$$ **quality points**.

# SYNTCOMP 2017: AIGER/SAFETY TRACK

# *AIGER/Safety: Participants 2017*

**Re-entry:**

- **Demiurge** (Könighofer, Seidl): SAT-based, different cooperating strategies [VMCAI14]
- **SafetySynth** (Tentrup): BDD-based, implements all optimizations from SYNTCOMP 2014 except abstraction [STTT17]
- **Simple BDD Solver** (Walker, Ryzhyk): BDD-based, abstraction, CUDD 3.0.0
- **TermiteSAT** (Legg, Narodytska, Ryzhyk, (Walker)): SAT-based, portfolio with Simple BDD Solver, hybrid mode that shares information [CAV16a]

**Updated:**

- **AbsSynthe** (Brenguier, Perez, Raskin, Sankur): BDD-based, compositional, abstraction; **updates on BDD minimization and compositionality** [SYNT14,15]

Tools that only support realizability check (no synthesis): Simple BDD Solver, TermiteSAT

# *AIGER/Safety Results 2017: Realizability*

**Number of Benchmarks: 234**

## Sequential execution mode:

| Rank | Tool (conf) | Solved | Unique |
|---|---|---|---|
| **1** | **Simple BDD Solver (w/ Abstraction)** | **171** | 0 |
| 2 | **SafetySynth** | 165 | **1** |
| | SafetySynth (alt.) | 165 | 0 |
| | Simple BDD Solver | 165 | 0 |
| | Simple BDD Solver (w/ Abstraction 2) | 165 | 0 |
| 6 | **AbsSynthe (SC3)** | 160 | **3** |
| 7 | AbsSynthe (SC2) | 156 | 0 |
| 8 | AbsSynthe (SC1) | 148 | 0 |
| 9 | **Demiurge (D1real)** | 127 | **11** |
| 10 | **TermiteSAT** | 101 | **6** |

**Not solved: 22**

## Parallel execution mode:

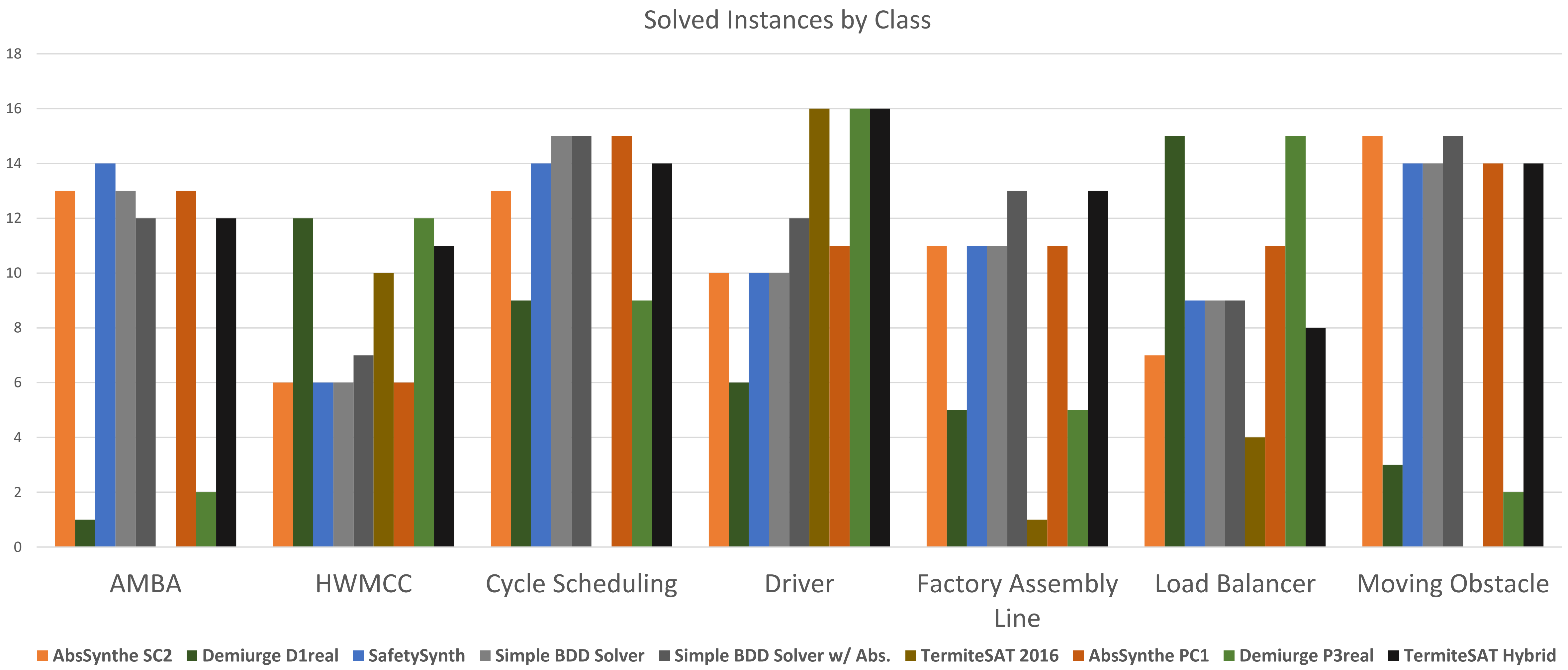
| Rank | Tool (conf) | Solved | Unique* |
|---|---|---|---|
| **1** | **TermiteSAT (Hybrid)** | **186** | 0 |
| 2 | TermiteSAT (Portfolio) | 185 | 0 |
| 3 | **AbsSynthe (PC1)** | 177 | 0 |
| 4 | **Demiurge (P3real)** | 161 | **1** |
| 5 | AbsSynthe (PC3) | 156 | 0 |
| 6 | AbsSynthe (PC2) | 147 | 0 |

*: including sequential solutions

**Not solved: 15**

# *AIGER/Safety Results: By Benchmark Class*

A few benchmark classes (not necessarily representative):



# *AIGER/Safety Results 2017: Synthesis*

**Number of benchmarks: 234**

## Sequential execution mode:

| Rank | Tool (conf) | Solved | Unique | MC TO | Quality |
|---|---|---|---|---|---|
| **1** | **SafetySynth** | **155** | **2** | 0 | **236** |
| 2 | SafetySynth (alt.) | 152 | **1** | 0 | 232 |
| 3 | **AbsSynthe (SSC2)** | 149 | 0 | 1 | 191 |
| 4 | AbsSynthe (SSC3) | 147 | 0 | 1 | 195 |
| 5 | AbsSynthe (SSC1) | 141 | **2** | 0 | 183 |
| 6 | **Demiurge (D1Synt)** | 118 | **20** | 1 | 175 |

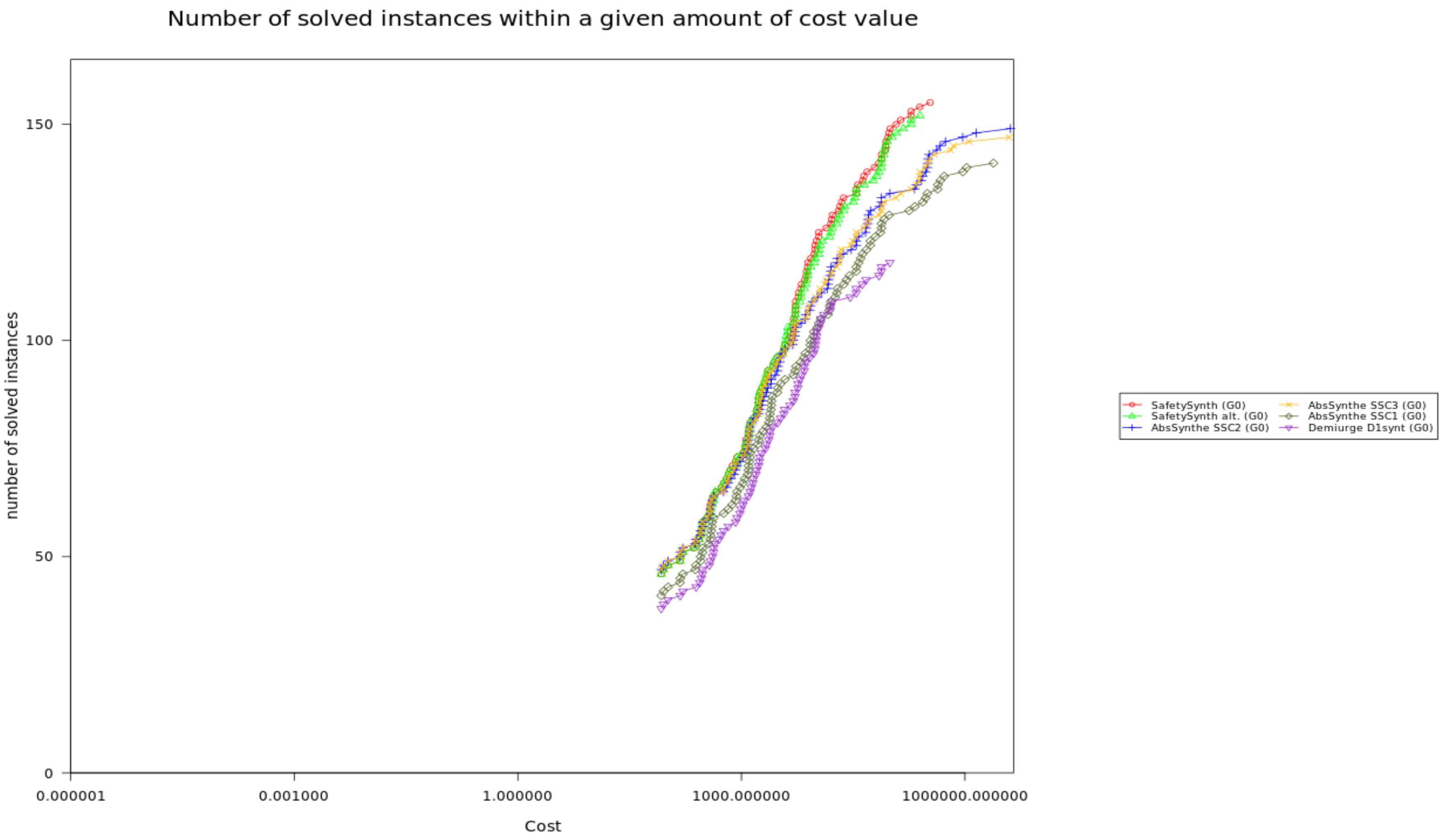
**not solved: 48**

## Parallel execution mode:

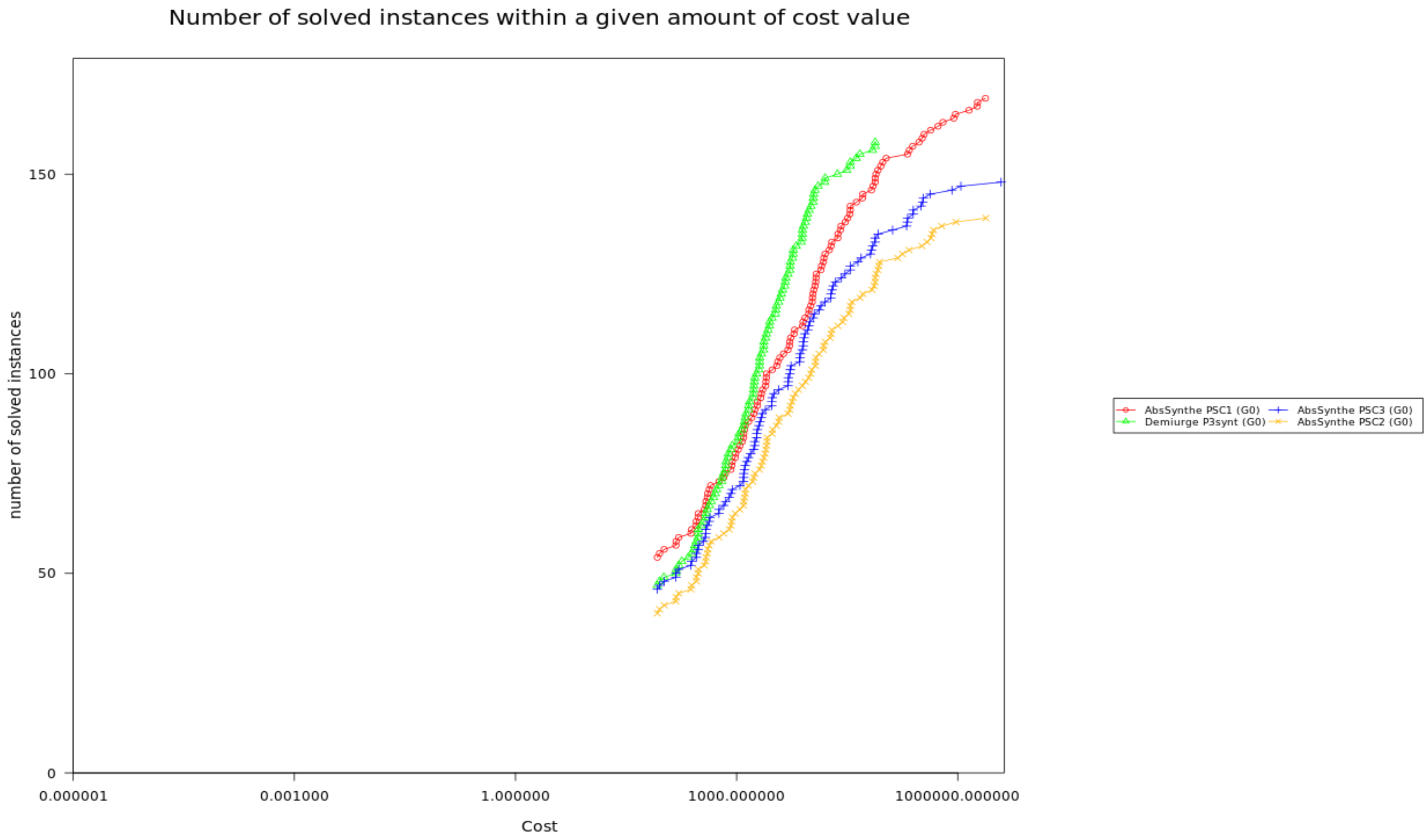
| Rank | Tool (conf) | Solved | Unique | MC TO | Quality |
|---|---|---|---|---|---|
| **1** | **AbsSynthe (PSC1)** | **169** | 0 | 2 | 210 |
| 2 | **Demiurge (P3Synt)** | 158 | 0 | 0 | **266** |
| 3 | AbsSynthe (PSC3) | 148 | 0 | 2 | 198 |
| 4 | AbsSynthe (PSC2) | 139 | 0 | 1 | 179 |

**not solved: 27**

# *AIGER/Safety Results 2017: Synthesis*



# *AIGER/Safety Results 2017: Synthesis*



# SYNTCOMP 2017: TLSF/LTL TRACK

# *TLSF/LTL: New Benchmarks 2017*

Parameterized benchmarks:

- **Decomposed AMBA** benchmarks from Felix Klein
- **Unrealizable variants of existing benchmarks** (additional requirement forces violation after fixed number of steps, or eventually)

# *TLSF/LTL: Participants 2017*

**Updated:**

- **BoSy** (Tentrup): bounded synthesis, SAT/QBF solving,
  **new: automata translation by spot** [CAV17]
- **Acacia4Aiger** (Brenguier, Perez, Raskin, Sankur): reduction to safety game, solved by antichains, uses compositionality, **new: parallel mode**
- **Party** (Khalimov): bounded synthesis, SMT solving,
  **new modes: reduction to AIGER/safety game (kid aiger), portfolio of three configurations**

**New entrants:**

- **Bowser (Klein):** bounded synthesis, SAT solving, additional bounded cycle synthesis [CAV16b] and aggressive circuit optimizations
- **spot-ltlsynt (Colange, Michaud):** translation to turn-based parity games with edge acceptance conditions, solving by Zielonka or Calude algorithm, BDD solving

# *TLSF/LTL Results 2017: Realizability*

**Number of benchmarks: 244**

**Sequential Execution Mode:**

| Rank | Tool (conf) | Solved | Unique |
|---|---|---|---|
| **1** | **Party (kid aiger)** | **219** | **7** |
| 2 | **spot-ltlsynt** | 195 | **3** |
| 3 | **Bosy (spot)** | 181 | 0 |
| 4 | Bosy (ltl3ba) | 172 | 0 |
| 5 | **Party (int)** | 169 | 0 |
| 6 | **Bowser** | 165 | 0 |
| | Party (bool) | 165 | 0 |
| 8 | **Acacia 4 Aiger** | 142 | **4** |

**Not Solved: 13**

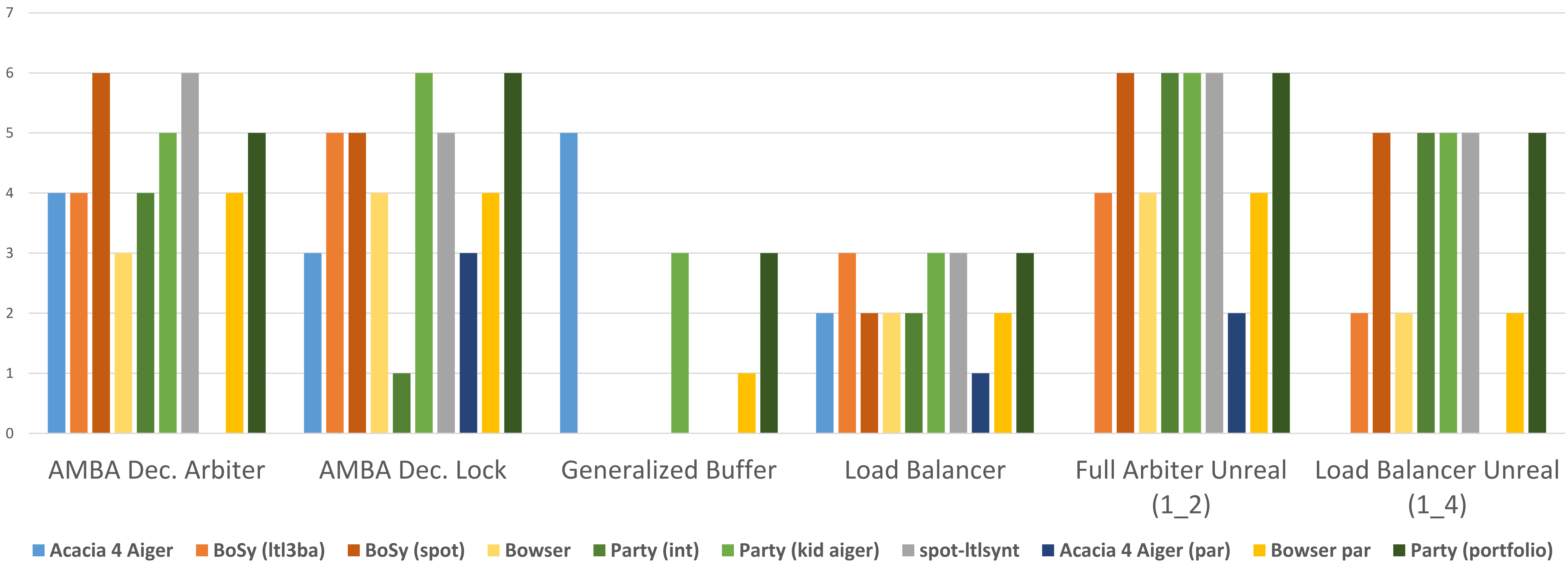
**Parallel Execution Mode:**

| Rank | Tool (conf) | Solved | Unique |
|---|---|---|---|
| **1** | **Party (portfolio)** | **224** | **2** |
| 2 | Bosy (spot, par) | 181 | 0 |
| 3 | **Bowser (par)** | 173 | 0 |
| 4 | Bosy (ltl3ba, par) | 170 | 0 |
| 5 | **Acacia 4 Aiger (par)** | 153 | 0 |
| | | | |

**Not Solved: 10**

# TLSF/LTL Results 2017: Parameterized Benchmarks

Number of solved instances for increasing parameter



# *TLSF/LTL Results 2017: Synthesis*

**Number of benchmarks: 244**

## Sequential execution mode:

| Rank | Tool (conf) | Solved | Unique | MC TO | Quality |
|---|---|---|---|---|---|
| **1** | **Party (kid aiger)** | **200** | **4** | 20 | 219 |
| 2 | **spot-ltlsynt** | 182 | **3** | 13 | 180 |
| 3 | **BoSy (spot)** | 181 | **3** | 0 | **298** |
| 4 | Party (int) | 167 | 0 | 0 | 249 |
| 5 | BoSy (ltl3ba) | 165 | 0 | 0 | 277 |
| 6 | Party (bool) | 163 | **1** | 0 | 222 |
| 7 | **Bowser (c0)** | 162 | 0 | 0 | 273 |
| 8 | **Bowser (c1)** | 155 | 0 | 0 | 260 |
| 9 | **Acacia 4 Aiger** | 110 | **2** | 17 | 91 |
| 10 | **Bowser (c2)** | 93 | **0** | 0 | 141 |

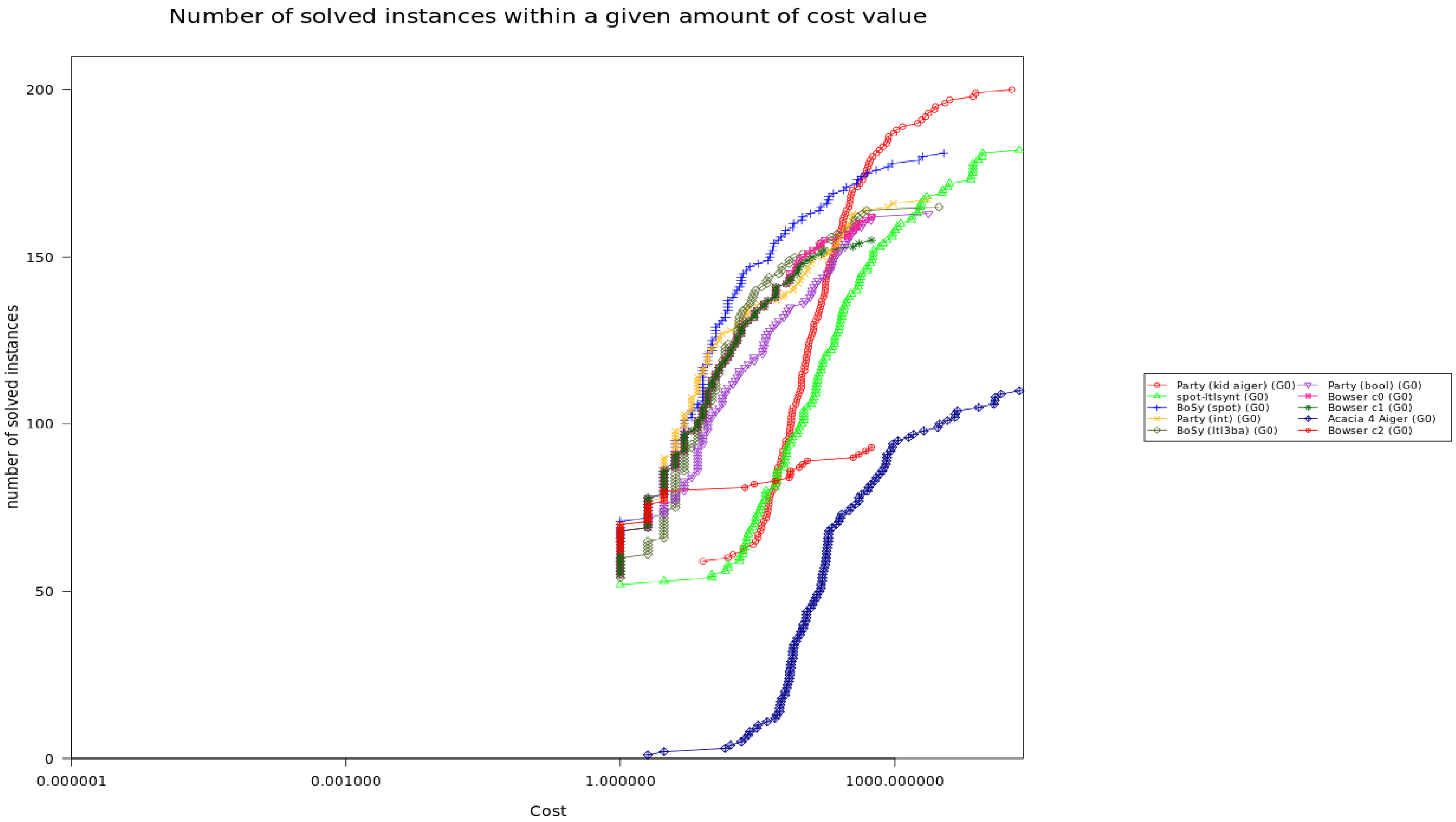
**not solved: 29**

## Parallel execution mode:

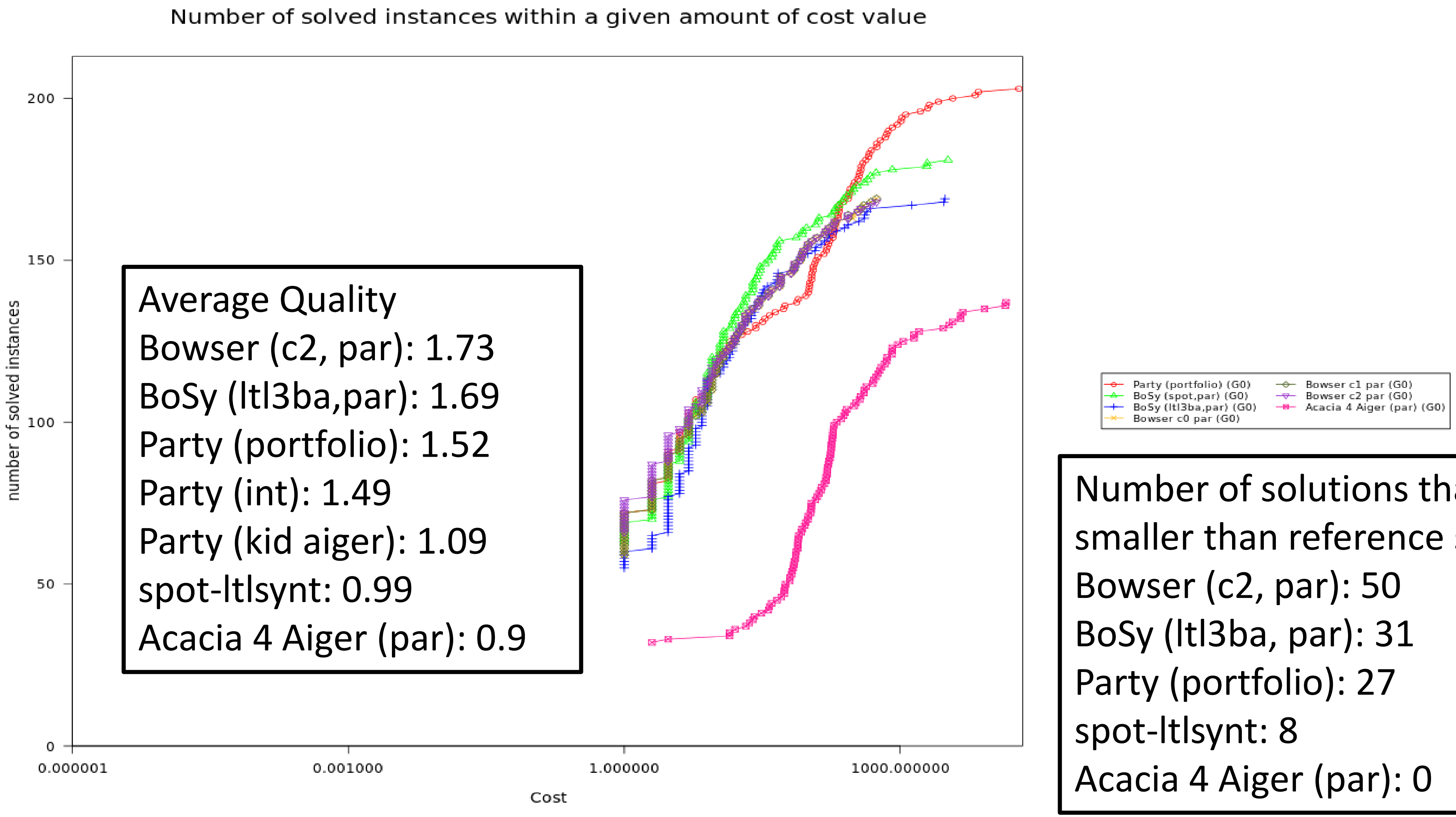
| Rank | Tool (conf) | Solved | Unique | MC TO | Quality |
|---|---|---|---|---|---|
| **1** | **Party (portfolio)** | **203** | 0 | 18 | **308** |
| 2 | BoSy (spot, par) | 181 | 0 | 0 | 297 |
| 3 | BoSy (ltl3ba, par) | 169 | 0 | 0 | 286 |
| | Bowser (c0, par) | 169 | 0 | 0 | 285 |
| | Bowser (c1, par) | 169 | 0 | 0 | 285 |
| 6 | Bowser (c2,par) | 168 | 0 | 0 | 290 |
| 7 | Acacia 4 Aiger | 137 | 0 | 5 | 123 |

**not solved: 27**

# *TLSF/LTL Results 2017: Synthesis*



# *TLSF/LTL Results 2017: Synthesis*




Average Quality
Bowser (c2, par): 1.73
BoSy (ltl3ba,par): 1.69
Party (portfolio): 1.52
Party (int): 1.49
Party (kid aiger): 1.09
spot-ltlsynt: 0.99
Acacia 4 Aiger (par): 0.9

Number of solutions that are smaller than reference size:
Bowser (c2, par): 50
BoSy (ltl3ba, par): 31
Party (portfolio): 27
spot-ltlsynt: 8
Acacia 4 Aiger (par): 0

# SYNTCOMP 2017: Summary

Winners of AIGER/Safety Track:

| | Sequential | Parallel |
|---|---|---|
| Realizability | Simple BDD Solver | TermiteSAT |
| Synthesis | SafetySynth | AbsSynthe |
| Synthesis Quality | SafetySynth | Demiurge |

Winners of TLSF/LTL Track:

| | Sequential | Parallel |
|---|---|---|
| Realizability | Party | Party |
| Synthesis | Party | Party |
| Synthesis Quality | BoSy | Party |

Detailed data online: http://syntcomp.cs.uni-saarland.de/syntcomp2017/experiments/

News and announcements on http://www.syntcomp.org

# References


[CAV13] Khalimov, A., Jacobs, S., & Bloem, R. PARTY parameterized synthesis of token rings. In CAV 2013, LNCS 8044 (pp. 928-933).

[CAV16a] Legg, A., Narodytska, N., & Ryzhyk, L. A SAT-Based Counterexample Guided Method for Unbounded Synthesis. In CAV 2016 (II), LNCS 9780 (pp. 364-382).

[CAV16b] Finkbeiner, B., & Klein, F. Bounded cycle synthesis. In CAV 2016 (I), LNCS 9779 (pp. 118-135).

[CAV17] Faymonville, P., Finkbeiner, B. & Tentrup, L. BoSy: An experimentation framework for Bounded Synthesis. In CAV 2017 (II), LNCS 10427 (pp. 325-332).

[SYNT14] Brenguier, R., Pérez, G. A., Raskin, J. F., & Sankur, O. AbsSynthe: abstract synthesis from succinct safety specifications. In SYNT 2014, EPTCS 157 (pp. 100-116) .

[SYNT15a] Brenguier, R., Pérez, G. A., Raskin, J. F., & Sankur, O. Compositional Algorithms for Succinct Safety Games. In SYNT 2015, EPTCS 202 (pp. 98-111).

[SYNT15b] Jacobs, S., Bloem, R., Brenguier, R., Könighofer, R., Pérez, G. A., Raskin, J. F., ... & Walker, A. The second reactive synthesis competition (SYNTCOMP 2015). In SYNT 2015, EPTCS 202 (pp. 27-57)

[SYNT16] Jacobs, S., Bloem, R., Brenguier, R., Khalimov, A., Klein, F., Könighofer, R., ... & Raskin, J. F. The 3rd reactive synthesis competition (SYNTCOMP 2016): Benchmarks, participants & results. In SYNT 2016, EPTCS 229 (pp. 149-177).

[VMCAI14] Bloem, R., Könighofer, R., & Seidl, M. SAT-based synthesis methods for safety specs. In VMCAI 2014, LNCS 8318 (pp. 1-20).

[STTT17] Jacobs, S., Bloem, R., Brenguier, R., Ehlers, R., Hell, T., Könighofer, R., ... & Seidl, M. The first reactive synthesis competition (SYNTCOMP 2014). In STTT, *19*(3), 2017 (pp. 367-390).