HWMCC'08

Hardware Model Checking Competition 2008

Armin Biere

Institute for Formal Models and Verification
Johannes Kepler University Linz, Austria

CAV'08

Princeton, USA

July 13, 2008

Organization

- Chair
  - Armin Biere (JKU, Linz, Austria)
- Committee
  - Alessandro Cimatti (IRST, Trento, Italy)
  - Koen Lindström Claessen (Chalmers, Gothenburg, Sweden)
  - Toni Jussila (OneSpin Solutions, Munich, Germany)
  - Ken McMillan (Cadence Berkeley Labs, Berkeley, USA)
  - Fabio Somenzi (University of Colorado, Boulder, USA)

Motivation

- advance model checking technology and research:
  - generate a large set of public available benchmarks
  - encourage researchers to work on novel model checking engines
  - platform for comparison
- repeat success story of SAT competition:
  - exponential improvement of SAT solvers
  - enhance visibility and generate more applications
- first things first: synchronous gate level models
  - AIGER format http://fmv.jku.at/aiger

645 Benchmarks: 344 old 301 new

- 344 "old" HWMCC'07 benchmarks (L2S, TIP, Intel, AMBA'07) AIGER
- 18 from Intel (submitted last year by Zurab Khasidashvili) SMV
- 35 new benchmarks from Barbara Jobstmann AIGER
- 207 modified VIS benchmarks from Politecnico di Torino BLIF
- 28 equivalence checking problems from Politecnico di Torino BLIF
- 13 model checking problems from NEC BLIF
- removed some "redundant" benchmarks (still have trivial ones)
- could not use/translate all benchmarks (Intel, BJ, Torino, incl. liveness)

16 Solvers

+	abc	University of California, Berkely	R. Brayton, M. Case,
+	abmc	University of California, Berkely	A. Hurst and A. Mishchenko
	aigtrav	JKU Linz	
+	mcaiger	JKU Linz	A. Biere
+	mcaigerbmc	JKU Linz	
+	nflbmc	CMU Pitsburgh	H. Jain
+	nflsmv2qbf	CMU Pitsburgh	
	nusmvbmc	IRST/FBK Trento	A. Cimatti, M. Roveri et.al.
+	pdtravbdd	Politecnico di Torino	G. Cabodi, S. Quer, S. Nocco,
+	pdtravbmc	Politecnico di Torino	M. Murciano and L. Garcia
+	pdtravcbq	Politecnico di Torino	
+	pdtravitp	Politecnico di Torino	
+	tipbmc	Chalmers Gothenburg	N. Sörensson, K. L. Claessen
+	tipidi	Chalmers Gothenburg	
+	tipids	Chalmers Gothenburg	2 were running last year
+	tipind	Chalmers Gothenburg	14 new submission!

- identical to last year

- 15 node cluster running Ubuntu Linux 7.04
  - Fully Automatic Install (FAI)
  - Sun's Grid Engine (SGE)

- identical nodes with Intel Pentium IV, 3 GHz, 2 GB main memory

- limits enforced by resource sampling `run` utility
  - time limit: **900 seconds**
  - space limit: **1.5 GB**

Unsolved Instances

- 48 unsolved instances:
  - 34 HWMCC'07 benchmarks (9 solved by "old" pdtrav-itp last year)
  - 12 old unsolved HWMCC'07 benchmarks are solved this year
  - 14 new benchmarks still remain unsolved

- smallest unsolved: `cmu.dme1.B.aig`

M	I	L	O	A
379	54	61	1	264

  same as last year

- the Intel suite is still the hardest one
  - there seems to be an issue with SMV to AIGER/BLIF translation
  - related to "transparent latches": `next(latch) := next(signal);`

Winners

ranked by number of solved instances

SAT + UNSAT		
1.	abc	552
2.	tipind	522
3.	pdtravitp	517

winner
HWMCC'07

ptdtravitp'08 is an improved version
of pdtravind'07 not of ptdtravitp'07

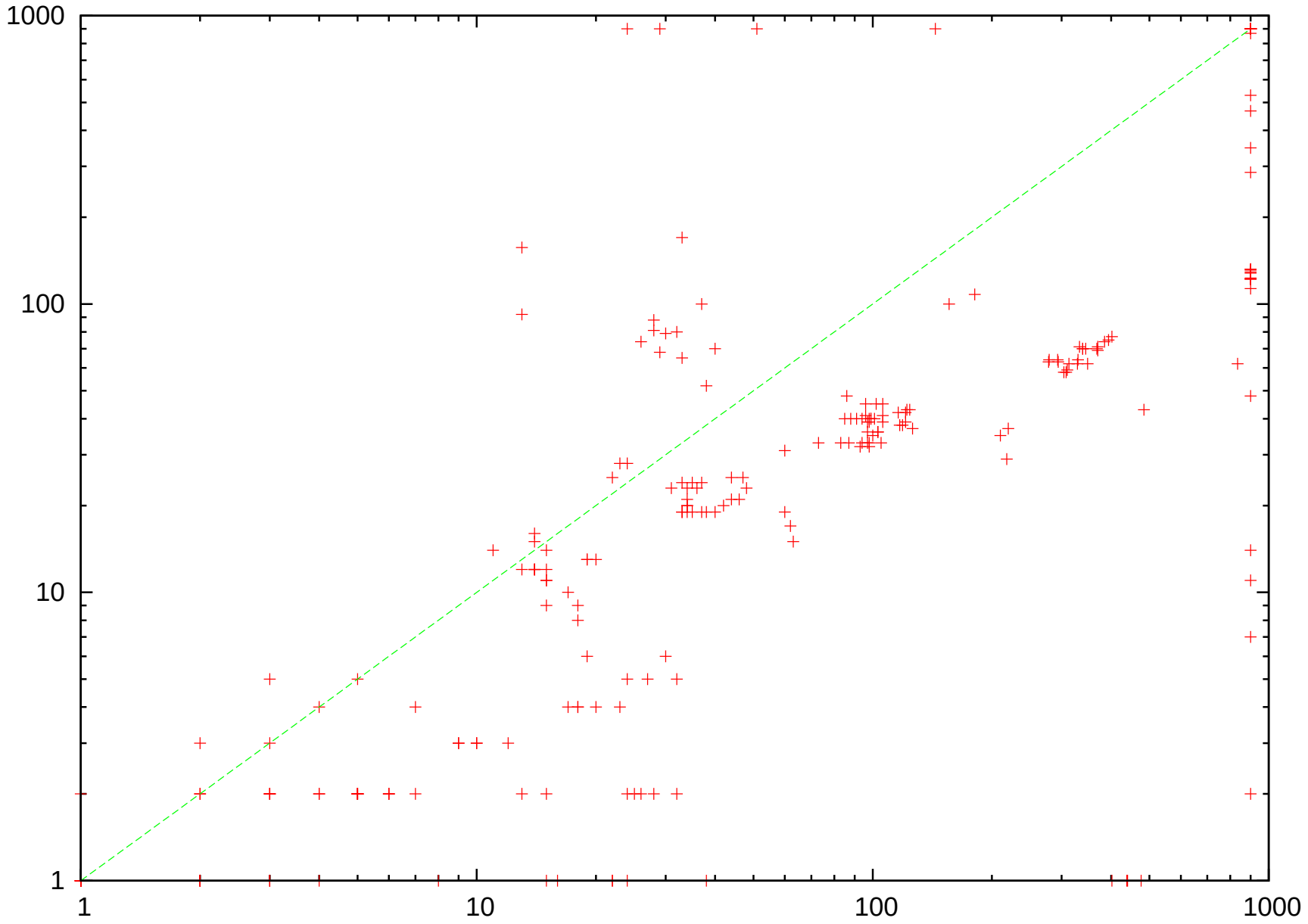
SAT		
1.	tipbmc	247
2.	mcaigerbmc	243
3.	nusmvbmc / pdtravbmc	239

nusmvbmc slightly faster

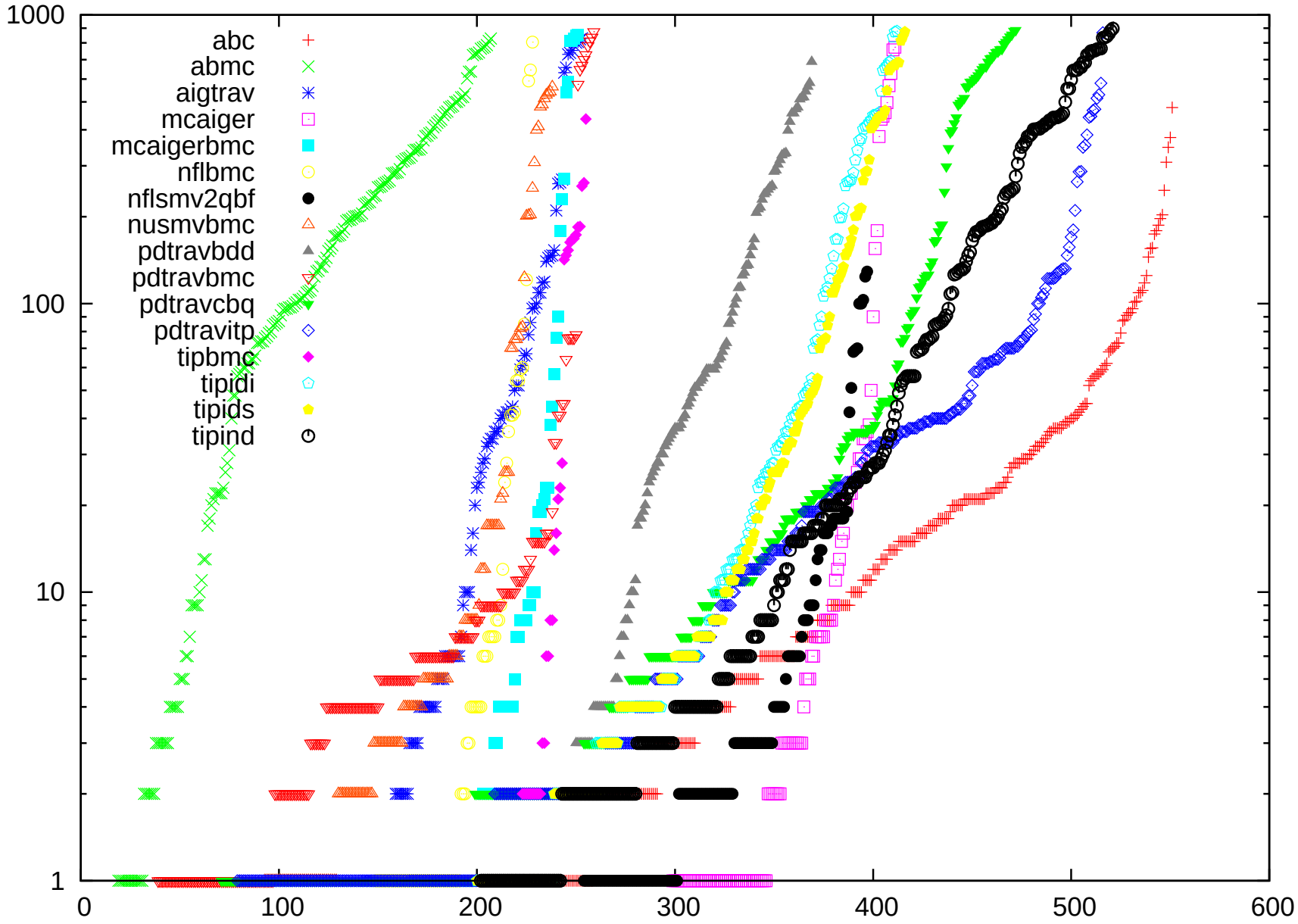
UNSAT		
1.	abc	314
2.	tipids	307
3.	tipidi	303
4.	pdtravitp	302

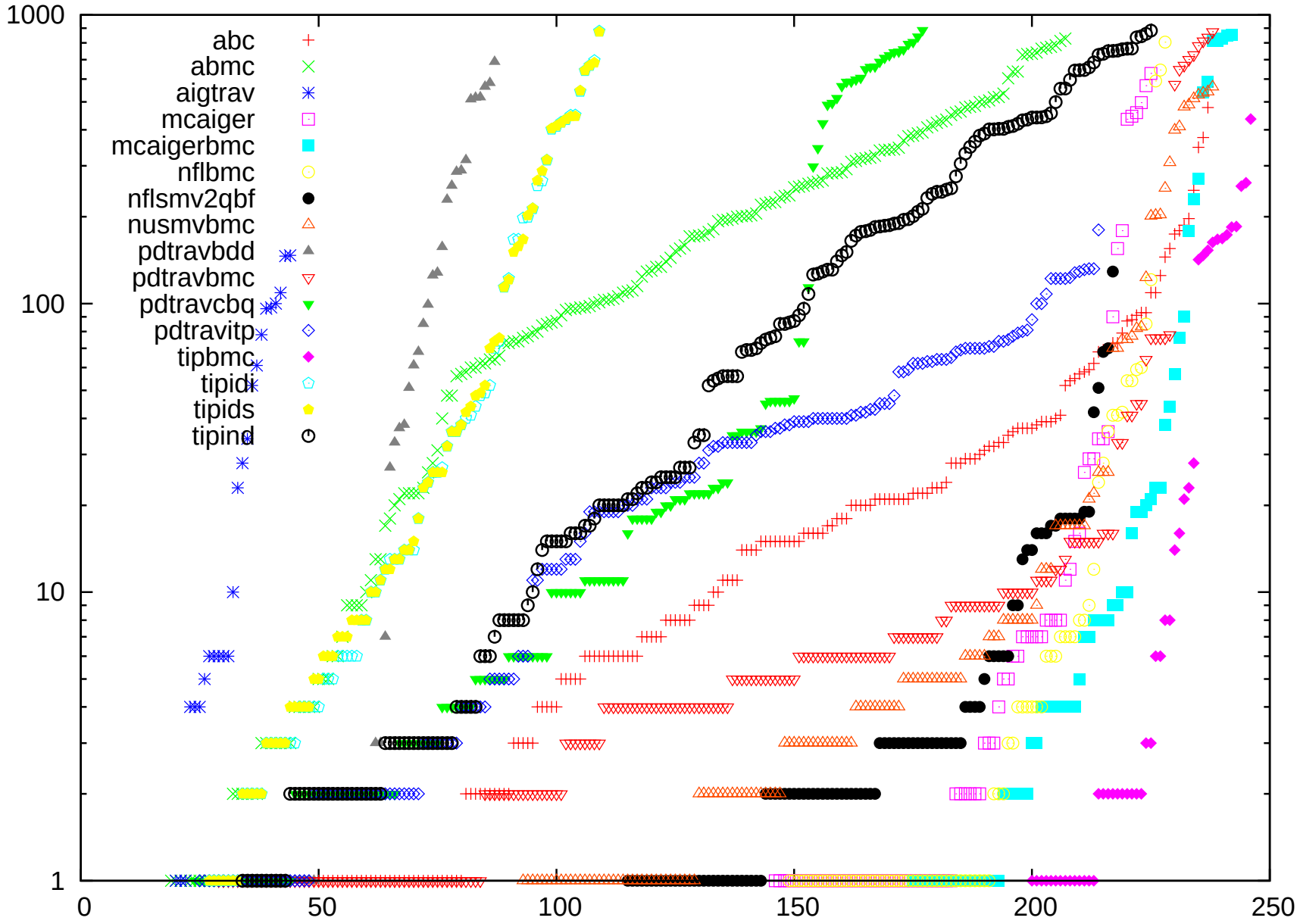
Old vs. New Version of PDTRAVITP on HWMCC'07 Benchmarks

cross below diagonal: new version faster

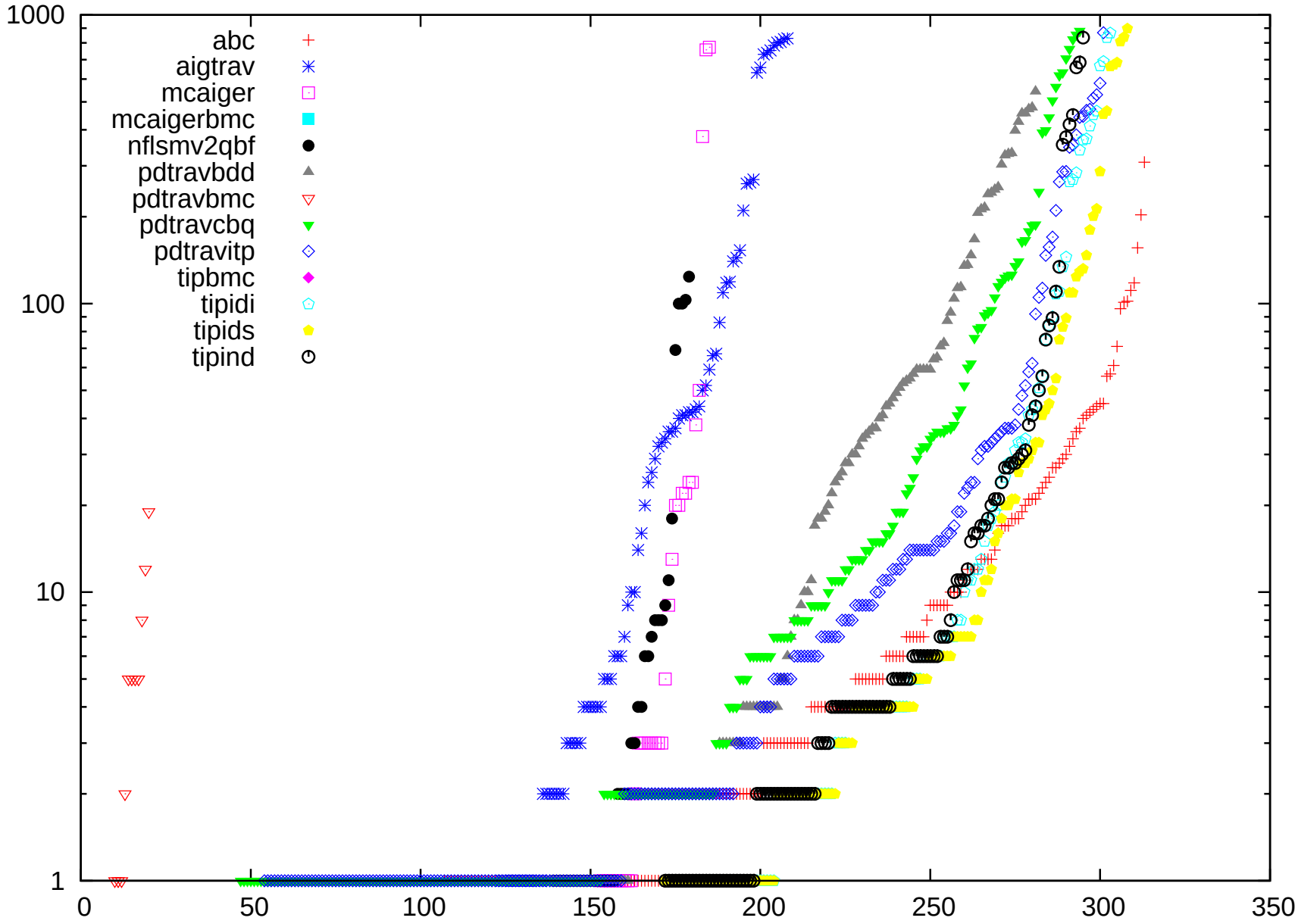


SAT + UNSAT


1000
100
10
1
0
100
200
300
400
500
600
abc
abmc
aigtrav
mcaiger
mcaigerbmc
nflbmc
nflsmv2qbf
nusmvbmc
pdtravbdd
pdtravbmc
pdtravcbq
pdtravitp
tipbmc
tipidi
tipids
tipind


1000
100
10
1
0
50
100
150
200
250
abc
abmc
aigtrav
mcaiger
mcaigerbmc
nflbmc
nflsmv2qbf
nusmvbmc
pdtravbdd
pdtravbmc
pdtravcbq
pdtravitp
tipbmc
tipidi
tipids
tipind

UNSAT


abc
aigtrav
mcaiger
mcaigerbmc
nflsmv2qbf
pdtravbdd
pdtravbmc
pdtravcbq
pdtravitp
tipbmc
tipidi
tipids
tipind
1000
100
10
1
0
50
100
150
200
250
300
350

Conclusion

- 14 new solvers
- 301 new benchmarks (actually many more in the pipe)
- solvers improved considerably
- more solvers natively read AIGER format
- some solvers (sometimes) provide witnesses for checking correctness
- no discrepancies in the (really) final runs

What is next?

- AIGER Format 2.0 (still on the agenda)
  - binary header (currently in ASCII even though rest is binary)
  - secondary outputs: constraints, multiple properties, fairness
- OS and I/O conformance of solvers (still on the agenda)
  - ideally: single statically linked binary, temporary files in `/tmp`
  - forbidden: processes, environment assumptions
  - clearly: witnesses / counterexample traces
- resolve “transparent latch” problem with Intel benchmarks
- start working on your solver now and send more benchmarks!