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
- 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)
<table>
<thead>
<tr>
<th>Solver</th>
<th>Institution</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>abc</td>
<td>University of California, Berkely</td>
<td>R. Brayton, M. Case, A. Hurst and A. Mishchenko</td>
</tr>
<tr>
<td>abmc</td>
<td>University of California, Berkely</td>
<td>A. Hurst and A. Mishchenko</td>
</tr>
<tr>
<td>aigtrav</td>
<td>JKU Linz</td>
<td></td>
</tr>
<tr>
<td>mcaiger</td>
<td>JKU Linz</td>
<td>A. Biere</td>
</tr>
<tr>
<td>mcaigerbmc</td>
<td>JKU Linz</td>
<td></td>
</tr>
<tr>
<td>nflbmc</td>
<td>CMU Pittsburgh</td>
<td>H. Jain</td>
</tr>
<tr>
<td>nflsmv2qbf</td>
<td>CMU Pittsburgh</td>
<td></td>
</tr>
<tr>
<td>nusmvbmc</td>
<td>IRST/FBK Trento</td>
<td>A. Cimatti, M. Roveri et.al.</td>
</tr>
<tr>
<td>pdtravbdd</td>
<td>Politecnico di Torino</td>
<td>G. Cabodi, S. Quer, S. Nocco, M. Murciano and L. Garcia</td>
</tr>
<tr>
<td>pdtravbmc</td>
<td>Politecnico di Torino</td>
<td></td>
</tr>
<tr>
<td>pdtravcbq</td>
<td>Politecnico di Torino</td>
<td></td>
</tr>
<tr>
<td>pdtravitp</td>
<td>Politecnico di Torino</td>
<td></td>
</tr>
<tr>
<td>tipbmc</td>
<td>Chalmers Gothenburg</td>
<td>N. Sörensson, K. L. Claessen</td>
</tr>
<tr>
<td>tipidi</td>
<td>Chalmers Gothenburg</td>
<td></td>
</tr>
<tr>
<td>tipids</td>
<td>Chalmers Gothenburg</td>
<td></td>
</tr>
<tr>
<td>tipind</td>
<td>Chalmers Gothenburg</td>
<td></td>
</tr>
</tbody>
</table>

2 were running last year

14 new submission!
Setup

- 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-ipt last year)
  - 12 old unsolved HWMCC’07 benchmarks are solved this year
  - 14 new benchmarks still remain unsolved

- smallest unsolved: cmu.dmel.B.aig
  - 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

<table>
<thead>
<tr>
<th>SAT + UNSAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. abc 552</td>
</tr>
<tr>
<td>2. tipind 522</td>
</tr>
<tr>
<td>3. pdtravitp 517</td>
</tr>
</tbody>
</table>

winner
HWMCC’07

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

<table>
<thead>
<tr>
<th>SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. tipbmc 247</td>
</tr>
<tr>
<td>2. mcaigerbmc 243</td>
</tr>
<tr>
<td>3. nusmvbmc / pdtravbmc 239</td>
</tr>
</tbody>
</table>

nusmvbmc slightly faster

<table>
<thead>
<tr>
<th>UNSAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. abc 314</td>
</tr>
<tr>
<td>2. tipids 307</td>
</tr>
<tr>
<td>3. tipidi 303</td>
</tr>
<tr>
<td>4. pdtravitp 302</td>
</tr>
</tbody>
</table>
Old vs. New Version of PDTRAVITP on HWMCC’07 Benchmarks

cross below diagonal: new version faster
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!**