The 5th Answer Set Programming Competition

Francesco Calimeri, Martin Gebser, Marco Maratea, Francesco Ricca

The 30th International Conference on Logic Programming
Vienna, Austria
Outline

1 The Fifth ASP Competition
2 Setup, Scoring and Benchmark Suite
3 Results
The Fifth ASP Competition

An event moved towards maturity

Departed from the usual timeline, in order to:

- be part of the Vienna Summer of Logic (VSL)
- push the standard ASP-Core-2
  (that was not satisfactorily supported by most participants)
The Fifth ASP Competition

An event moved towards maturity

Departed from the usual timeline, in order to:
- be part of the Vienna Summer of Logic (VSL)
- push the standard ASP-Core-2
  (that was not satisfactorily supported by most participants)

Goals
- measure the progress of the state of the art
- draw a more complete picture of approaches to problems with different features
- improve benchmarks w.r.t. modeling
The 5th Competition Setting

- Benchmarks from past editions, mainly from 2013 edition
- System competition only and modeling competition on site
- New benchmark classification based on language features
- Redesigned problem encodings
- Updated versions of the solvers, and newcomers
System Competition Format

Two Categories
- Single-Processor (restricted to 1-CPU Core)
- Multi-Processor (up to 8-CPU Cores)

Tracks based on language features
- **Track 1 (Basic)**: normal LP + simple built-ins
- **Track 2 (Advanced)**: + choices, aggregates, HCF disjunction
- **Track 3 (Optimization)**: + weak constraints
- **Track 4 (Unrestricted)**: + non-HCF disjunction
The competition featured 16 systems coming from three teams:

- **Aalto Team**, Aalto University (9 solvers):
  - LP2SAT3+GLUCOSE, LP2SAT3+LINGELING, LP2SAT3+LINGELING-MT,
  - LP2BV2+BOOLECTOR, LP2GRAPH, LP2MAXSAT+CLASP, LP2MIP2,
  - LP2MIP2-MT, LP2NORMAL2+CLASP

- **Potassco Team**, University of Potsdam (2 solvers):
  - CLASP, CLASP-MT

- **Wasp Team**, University of Calabria (5 solvers):
  - WASP-1, WASP-2, WASP-1.5, WASP-WMSU1-ONLYWEAK,
  - WASP-WPM1-ONLY-WEAK
Outline

1. The Fifth ASP Competition
2. Setup, Scoring and Benchmark Suite
3. Results
Setup

System Inputs
- Fixed input in ASP-Core-2
- Solvers run with fixed settings
- 20 instances per domain, randomly selected

System Environment
- Debian Linux server with Intel Xeon X5365 CPUs
- Time limit: 10 minutes
- Memory Limit: 6 GB
- Multi-processor track: up to 8 cores
- Performance measured using the pyrunlim tool
Scoring

Simplified Scoring

- Consider number of solved instances for decision problems
- Rank solvers on optimization problems by solution quality
- Runtime for tiebreaker

Decision and Query Problems

\[ \text{Score}(\text{Solver}, \text{Problem}) = \#\text{Solved}(\text{Solver}) \times 5 \]

Optimization Problems

\[ \text{Score}(\text{Solver}, \text{Problem}) = \sum_{\text{Instances}} \frac{\#\text{NotBetter}(\text{Solver}, I) \times 5}{\#\text{Participants}} \]
Scoring

Simplified Scoring

- Consider number of solved instances for decision problems
- Rank solvers on optimization problems by solution quality
- Runtime for tiebreaker

Additional Criteria

- Problems are equally weighted up to 100 points each
- Incorrect answers: disqualification on per problem basis
- Final scores by summing over all problems
## Benchmark Suite

<table>
<thead>
<tr>
<th>Domain</th>
<th>P</th>
<th>2013 Encoding</th>
<th>2014 Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labyrinth</td>
<td>D</td>
<td>basic, non-tight</td>
<td>basic, non-tight</td>
</tr>
<tr>
<td>Stable Marriage</td>
<td>D</td>
<td>basic</td>
<td>basic</td>
</tr>
<tr>
<td>Bottle Filling</td>
<td>D</td>
<td>aggr</td>
<td>aggr, choice</td>
</tr>
<tr>
<td>Graceful Graphs</td>
<td>D</td>
<td>choice#</td>
<td>choice#</td>
</tr>
<tr>
<td>Graph Colouring*</td>
<td>D</td>
<td>disj</td>
<td>basic</td>
</tr>
<tr>
<td>Hanoi Tower*</td>
<td>D</td>
<td>disj</td>
<td>basic</td>
</tr>
<tr>
<td>Incremental Scheduling</td>
<td>D</td>
<td>aggr, choice#</td>
<td>aggr, choice#</td>
</tr>
<tr>
<td>Knight Tour with Holes*</td>
<td>D</td>
<td>disj, non-tight</td>
<td>basic, non-tight</td>
</tr>
<tr>
<td>Nomystery</td>
<td>D</td>
<td>aggr, choice#</td>
<td>choice#</td>
</tr>
<tr>
<td>Partner Units</td>
<td>D</td>
<td>aggr, disj, non-tight</td>
<td>aggr, choice</td>
</tr>
<tr>
<td>Permutation Pattern Matching</td>
<td>D</td>
<td>choice#</td>
<td>choice</td>
</tr>
<tr>
<td>Qualitative Spatial Reasoning</td>
<td>D</td>
<td>choice#, disj</td>
<td>disj</td>
</tr>
<tr>
<td>Reachability</td>
<td>Q</td>
<td>non-tight</td>
<td>n/a</td>
</tr>
<tr>
<td>Ricochet Robots</td>
<td>D</td>
<td>choice#</td>
<td>aggr, choice#</td>
</tr>
<tr>
<td>Sokoban</td>
<td>D</td>
<td>aggr, choice#</td>
<td>choice#</td>
</tr>
<tr>
<td>Solitaire</td>
<td>D</td>
<td>choice#</td>
<td>aggr, choice#</td>
</tr>
<tr>
<td>Visit-all*</td>
<td>D</td>
<td>aggr, choice#</td>
<td>basic</td>
</tr>
<tr>
<td>Weighted-Sequence Problem</td>
<td>D</td>
<td>choice#</td>
<td>aggr, choice</td>
</tr>
<tr>
<td>Connected Still Life</td>
<td>O</td>
<td>aggr, choice#, non-tight</td>
<td>aggr, choice, non-tight</td>
</tr>
<tr>
<td>Crossing Minimization</td>
<td>O</td>
<td>disj</td>
<td>aggr, choice</td>
</tr>
<tr>
<td>Maximal Clique</td>
<td>O</td>
<td>disj</td>
<td>basic</td>
</tr>
<tr>
<td>Valves Location</td>
<td>O</td>
<td>aggr, choice#, non-tight</td>
<td>aggr, choice#, non-tight</td>
</tr>
<tr>
<td>Abstract Dialectical Frameworks</td>
<td>O</td>
<td>aggr, disj, level, non-tight</td>
<td>aggr, disj, level, non-tight</td>
</tr>
<tr>
<td>Complex Optimization</td>
<td>D</td>
<td>choice, disj, non-tight</td>
<td>choice, disj, non-tight</td>
</tr>
<tr>
<td>Minimal Diagnosis</td>
<td>D</td>
<td>disj, non-tight</td>
<td>disj, non-tight</td>
</tr>
<tr>
<td>Strategic Companies</td>
<td>Q</td>
<td>disj, non-tight</td>
<td>n/a</td>
</tr>
</tbody>
</table>
## Benchmark Suite

<table>
<thead>
<tr>
<th>Domain</th>
<th>P</th>
<th>2013 Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Labyrinth</strong></td>
<td>D</td>
<td>basic, non-tight</td>
</tr>
<tr>
<td><strong>Stable Marriage</strong></td>
<td>D</td>
<td>basic</td>
</tr>
<tr>
<td><strong>Bottle Filling</strong></td>
<td>D</td>
<td>aggr</td>
</tr>
<tr>
<td><strong>Graceful Graphs</strong></td>
<td>D</td>
<td>choice#</td>
</tr>
<tr>
<td><strong>Graph Colouring</strong></td>
<td>D</td>
<td>disj</td>
</tr>
<tr>
<td><strong>Hanoi Tower</strong></td>
<td>D</td>
<td>disj</td>
</tr>
<tr>
<td><strong>Incremental Scheduling</strong></td>
<td>D</td>
<td>aggr, choice#</td>
</tr>
<tr>
<td><strong>Knight Tour with Holes</strong></td>
<td>D</td>
<td>disj, non-tight</td>
</tr>
<tr>
<td><strong>Nomystery</strong></td>
<td>D</td>
<td>aggr, choice#</td>
</tr>
<tr>
<td><strong>Partner Units</strong></td>
<td>D</td>
<td>aggr, disj, non-tight</td>
</tr>
<tr>
<td><strong>Permutation Pattern Matching</strong></td>
<td>D</td>
<td>choice#</td>
</tr>
<tr>
<td><strong>Qualitative Spatial Reasoning</strong></td>
<td>D</td>
<td>choice#, disj</td>
</tr>
<tr>
<td><strong>Reachability</strong></td>
<td>D</td>
<td>choice#</td>
</tr>
<tr>
<td><strong>Binary Budget</strong></td>
<td>Q</td>
<td>non-tight</td>
</tr>
<tr>
<td><strong>Dingbat</strong></td>
<td>D</td>
<td>choice#</td>
</tr>
</tbody>
</table>
Outline

1. The Fifth ASP Competition
2. Setup, Scoring and Benchmark Suite
3. Results
Results: Multi Processor Track

Multi Processor Track

- clasp-mt
- lp2sat3+plingeling-mt
- lp2mip2-mt
Results: Track 1 - Basic
Results: Track 1 - Basic

![Bar Chart for Track 1]

- wasp1.5
- wasp2
- lp2graph
- lp2maxsat+clasp
- lp2sat3+glucose
- wasp1
- lp2sat3+lingeling
- lp2bv2+boolector
- lp2mip2
Results: Track 1 - Basic

Track 1

- lp2normal2+clasp
- wasp1.5
- wasp2
- lp2graph
- lp2maxsat+clasp
- lp2sat3+glucose
- wasp1
- lp2sat3+lingeling
- lp2bv2+boolector
- lp2mip2
Results: Track 1 - Basic

The bar chart shows the performance of different tools in Track 1 of the ASP competition. The tools are ordered from top to bottom as follows:

- clasp-st
- lp2normal2+clasp
- wasp1.5
- wasp2
- lp2graph
- lp2maxsat+clasp
- lp2sat3+glucose
- wasp1
- lp2sat3+lingeling
- lp2bv2+boolector
- lp2mip2

The x-axis represents the score range from 0 to 200, while the y-axis lists the tools.
Results: Track 2 - Advanced
Results: Track 2 - Advanced

![Bar chart showing results for Track 2]

- lp2sat3+lingeling
- lp2maxsat+clasp
- lp2sat3+glucose
- wasp1.5
- lp2graph
- wasp2
- lp2bv2+boolector
- wasp1
- lp2mip2

Francesco Calimeri, Martin Gebser, Marco Maratea, Francesco Ricca
Results: Track 2 - Advanced

clasp-st
lp2sat3+lingeling
lp2maxsat+clasp
lp2sat3+glucose
wasp1.5
lp2graph
wasp2
lp2bv2+boolector
wasp1
lp2mip2
Results: Track 2 - Advanced

The chart shows the performance of different solvers on Track 2. The solvers are listed as follows:

- lp2normal2+clasp
- clasp-st
- lp2sat3+lingeling
- lp2maxsat+clasp
- lp2sat3+glucose
- wasp1.5
- lp2graph
- wasp2
- lp2bv2+boolector
- wasp1
- lp2mip2

The x-axis represents the performance metric, ranging from 0 to 1000. The performance of each solver is depicted by the length of the bar.
Results: Track 3 - Optimization
Results: Track 3 - Optimization
Results: Track 3 - Optimization

![Bar chart showing results for various components in Track 3.](chart.png)
Results: Track 3 - Optimization

![Bar chart showing results for Track 3]

- clasp-st
- wasp1
- wasp2
- wasp1.5
- lp2normal2+clasp
- lp2maxsat+clasp
- lp2mip2
- wasp_wpm1_only_weak
Results: Track 4 - Unrestricted
Results: Track 4 - Unrestricted

![Bar Chart for Track 4](chart.png)

- wasp1
- wasp1.5
- wasp_wpm1_only_weak
Results: Track 4 - Unrestricted
Results: Track 4 - Unrestricted

Track 4

- clasp-st
- lp2normal2+clasp
- wasp1
- wasp1.5
- wasp_wpm1_only_weak
On the impact of new encodings

![Bar chart showing the gain of different encodings.](chart.png)
Results: ASP Competition 2014 - Overall

The chart shows the overall results for different tracks in the ASP Competition 2014. The x-axis represents the score, ranging from 0 to 1800. The y-axis lists the tracks labeled as Track1, Track2, Track3, Track4, and Track4Opt. Each track is represented by a different color and the segments show the performance across different benchmarks or tasks.
Results: ASP Competition 2014 - Overall
Results: ASP Competition 2014 - Overall

The results for ASP Competition 2014 are shown in the diagram. The x-axis represents the score, ranging from 0 to 1800. Each track (Track1, Track2, Track3, Track4, Track4Opt) is represented by a different color. The names of some datasets, such as "lp2mip2" and "wasp_wpm1_only_weak", are listed below the chart.
Results: ASP Competition 2014 - Overall

- lp2graph
- lp2sat3+lingeling
- lp2sat3+glucose
- lp2bv2+boolector
- lp2mip2
- wasp_wpm1_only_weak

The 5th Answer Set Programming Competition
Results: ASP Competition 2014 - Overall

![Bar chart showing results for different systems and tracks.]

- Systems: wasp1, lp2maxsat+clasp, wasp2, lp2graph, lp2sat3+lingeling, lp2sat3+glucose, lp2bv2+boolector, lp2mip2, wasp_wpm1_only_weak
- Tracks: Track1, Track2, Track3, Track4, Track4Opt

Francesco Calimeri, Martin Gebser, Marco Maratea, Francesco Ricca
Results: ASP Competition 2014 - Overall

Francesco Calimeri, Martin Gebser, Marco Maratea, Francesco Ricca

The 5th Answer Set Programming Competition
Results: ASP Competition 2014 - Overall

- clasp-st
- lp2normal2+clasp
- wasp1.5
- wasp1
- lp2maxsat+clasp
- wasp2
- lp2graph
- lp2sat3+lingeling
- lp2sat3+glucose
- lp2bv2+boolector
- lp2mip2
- wasp_wpm1_only_weak

Legend:
- Track1
- Track2
- Track3
- Track4
- Track4Opt

Francesco Calimeri, Martin Gebser, Marco Maratea, Francesco Ricca
Suggestions for future ASP events

Simplify Output
- Unify output for tasks
- Reduce number of exit codes

Instance Selection
- Process for discarding very easy/hard
- More ASP-oriented real-world applications
- Enforce classification by language features
- Non-ground and ground tracks?
- Cautious/Brave Reasoning?

Modeling Competition
- Interactive event? Challenges? . . .