Unsolvability IPC Track

Christian Muise $^1$   Nir Lipovetzky $^2$

$^1$Computer Science and Artificial Intelligence Laboratory  
Massachusetts Institute of Technology, USA  
cjmuise@mit.edu

$^2$Department of Computing and Information Systems,  
University of Melbourne, Australia  
nir.lipovetzky@unimelb.edu.au
Before we begin...

Running an IPC is hard (now we know!)

Unsolvability is (thankfully still) hard

Kudos to those who made it easier:

Florian Pommerening & Jendrik Seipp (& the Basel cluster admins) for lab and server support / consultation

Malte Helmert for involved FD diagnosis

ICAPS for sponsoring the awards

All who submitted planners and domains (to be mentioned below) – many who put up with very long email threads!

Muise and Lipovetzky: Unsolvability IPC Track
Before we begin...

- Running an IPC is hard (*now we know!*)

Muise and Lipovetzky: Unsolvability IPC Track
Before we begin...

- Running an IPC is hard (*now we know!*)
- Unsolvability is (thankfully still) hard
Before we begin...

- Running an IPC is hard (*now we know!*)

- Unsolvability is (thankfully still) hard

- Kudos to those who made it easier:
  - Florian Pommerening & Jendrik Seipp (& the Basel cluster admins) for lab and server support / consultation
  - Malte Helmert for involved FD diagnosis
  - ICAPS for sponsoring the awards
  - All who submitted planners and domains (to be mentioned below) – many who put up with very long email threads!
Motivation

Satisficing is too easy. To promote techniques that are dedicated at detecting if a planning problem is unsolvable.

Applications

- System verification / diagnosis
- Planning with avoidable deadends
- Identifying human error in encodings
- Components in cellular automata proofs
Scoring

- Primary focus was coverage on unsolvable instances
- Solver disqualified on a domain for an incorrect result

Didn’t happen once!
### Scoring
- Primary focus was coverage on unsolvable instances
- Solver disqualified on a domain for an incorrect result

Didn’t happen once!

### Ideal Benchmark Properties
1. Mix of solvable and unsolvable instances
2. No syntactic distinction between them
3. At least some solvable problems are hard
### Scoring
- Primary focus was coverage on unsolvable instances
- Solver disqualified on a domain for an incorrect result

*Didn’t happen once!*

### Ideal Benchmark Properties
1. Mix of solvable and unsolvable instances
2. No syntactic distinction between them
3. At least some solvable problems are hard

### Benchmarks
Sourced from deadend states of classical problems, width-based classical encodings, oversubscribed problems, new domains, etc.
## Submissions

<table>
<thead>
<tr>
<th>Domains (15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 3 from existing set</td>
</tr>
<tr>
<td>- 12 newly submitted (or modified) domains</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Planners (11+1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 14 teams expressed interest</td>
</tr>
<tr>
<td>- 11 teams submitted</td>
</tr>
<tr>
<td>- 4 labs represented</td>
</tr>
</tbody>
</table>
Reused Unsolvable Domains

- **bottleneck**: Move around a grid; never revisit a spot
- **pegsol**: Classic game – hop a peg, remove the hopped
- **sliding-tiles**: Classic game – slide tiles to open space.
  Reversing two tiles causes the problem to be unsolvable
Thanks
Marcel Steinmetz, Hootan Nakhost, Jörg Hoffmann, Martin Mueller

Idea
Constrains the problem via a set of tunable parameters. Most importantly for the contest, “constrainedness” < 1 makes the problem unsolvable (i.e., 0.999 is very nearly solvable)

Input Domains
nomystery, tpp, rovers
Thanks  Patricia Riddle and Mike Barley

Idea  Reformulates the problem using a bagged representation

Input Domains  transport, gripper, barman
• **document-transfer** (*Thanks to Jordan Douglas*): Need to deliver documents, but can also use them to fuel the travel.

• **diving** (*Thanks to Charles Gretton, Nathan Robinson, CM*): Need to photograph certain locations, but air tanks are a precious resource that aren’t always there.

• **diagnosis** (*Thanks to Patrik Haslum*): Encodings of diagnosis problems for control of power plans / autonomous helicopters.
Baffling Board Games

- **tetris** (*Thanks to Mauro Vallati*): Moving tetris pieces around until the board is cleared.

- **pegsol-row5** (*Thanks to Florian Pommerening and Malte Helmert*): Must get 5 rows up on a board, but the task is mathematically impossible.

- **chessboard-pebbling** (*Thanks to Florian Pommerening*): Must clear the bottom corner of an infinite chessboard.
Planners (1/3)

- **ReachLunch** (*Tomas Balyo, Martin Suda*)
  6min DFS phase, followed by Property Directed Reachability

- **iProverPlan** (*Konstantin Korovin, Martin Suda*)
  Theorem prover for a lifted (1st-order) SAT-as-planning encoding

- **SymPA** (*Alvaro Torralba, Jörg Hoffmann*)
  Symbolic search using PDBs and perimeters
  - **irr** variant removes irrelevant operators found with M&S

- **M+S** (*Alvaro Torralba, Jörg Hoffmann*)
  M&S with linear merges and perfect shrinking (optionally **irr**)
• **SimDominance** *(Alvaro Torralba, Jörg Hoffmann)*
  M&S with multiple merge strategies. A* with $h^1$ / dominance pruning

• **DECS** *(Daniel Gnad, Alvaro Torralba, Jörg Hoffmann, Martin Wehrle)*
  Decoupled search with an X-shape topology using A* and $h^{\text{max}}$

• **Django** *(Daniel Gnad, Marcel Steinmetz, Jörg Hoffmann)*
  Performs incremental red-black search using A* and $h^{\text{max}}$.

• **CLone** *(Marcel Steinmetz, Jörg Hoffmann)*
  DFS using Critical-Path Driven Clause Learning
**Aidos** *(Jendrik Seipp, Florian Pommerening, Silvan Sievers, Martin Wehrle, Chris Fawcett, Yusra Alkhazraji)*

Portfolio using stubborn sets and combination of (1) DE PDBs (2) LP with potentials (3) projected “depletable resources”

- **v1** distributes time according to experiments; **v2** distributes uniformly; and **v3** distributes to maximize coverage / time

**h**++ *(Patrik Haslum)*

Incrementally improves lower bounds until h+ is unsolvable

**DE-PDB** *(Florian Pommerening, Jendrik Seipp)*

$\frac{1}{2}$ time spent on building DE PDB and $\frac{1}{2}$ spent on DFS with pruning

**blind**

Eager blind search using the latest FD
Results (out of 340 problems)

<table>
<thead>
<tr>
<th>Runner Up</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SymPA</td>
<td>239</td>
</tr>
<tr>
<td>Alvaro Torralba</td>
<td></td>
</tr>
<tr>
<td>base</td>
<td>227</td>
</tr>
<tr>
<td>Jörg Hoffmann</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Winner</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aidos</td>
<td>262</td>
</tr>
<tr>
<td>Jendrik Seipp</td>
<td></td>
</tr>
<tr>
<td>v1</td>
<td>259</td>
</tr>
<tr>
<td>Florian Pommerening</td>
<td></td>
</tr>
<tr>
<td>v2</td>
<td>250</td>
</tr>
<tr>
<td>Silvan Sievers</td>
<td></td>
</tr>
<tr>
<td>v3</td>
<td></td>
</tr>
<tr>
<td>Martin Wehrle</td>
<td></td>
</tr>
<tr>
<td>Chris Fawcett</td>
<td></td>
</tr>
<tr>
<td>Yusra Alkhazraji</td>
<td></td>
</tr>
</tbody>
</table>

Muise and Lipovetzky: Unsolvability IPC Track
## Results (out of 340 problems)

### Runner Up – SymPA

<table>
<thead>
<tr>
<th></th>
<th>Score</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>irr</td>
<td>239</td>
<td>Alvaro Torralba</td>
</tr>
<tr>
<td>base</td>
<td>227</td>
<td>Jörg Hoffmann</td>
</tr>
</tbody>
</table>

### Winner

- Aidos
- Jendrik Seipp
- v1 262
- v2 259
- v3 250
- Chris Fawcett
- Yusra Alkhazraji

Muise and Lipovetzky: Unsolvability IPC Track
## Results (out of 340 problems)

### Runner Up – SymPA

<table>
<thead>
<tr>
<th>irr</th>
<th>239</th>
<th>Alvaro Torralba</th>
</tr>
</thead>
<tbody>
<tr>
<td>base</td>
<td>227</td>
<td>Jörg Hoffmann</td>
</tr>
</tbody>
</table>

### Winner – Aidos

<table>
<thead>
<tr>
<th>v1</th>
<th>262</th>
<th>Jendrik Seipp</th>
</tr>
</thead>
<tbody>
<tr>
<td>v2</td>
<td>259</td>
<td>Florian Pommerening</td>
</tr>
<tr>
<td>v3</td>
<td>250</td>
<td>Silvan Sievers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Martin Wehrle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chris Fawcett</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yusra Alkhazraji</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th>bottleneck</th>
<th>pegsol</th>
<th>over-tpp</th>
<th>over-nomystery</th>
<th>over-rovers</th>
<th>sliding-tiles</th>
<th>doc-transfer</th>
<th>tetris</th>
<th>diagnosis</th>
<th>bag-transport</th>
<th>chess-pebbling</th>
<th>bag-grripper</th>
<th>bag-barman</th>
<th>pegsol-row5</th>
<th>cave-diving</th>
</tr>
</thead>
</table>

Aidos1
Aidos2
Aidos3
SymPA (irr)
SymPA
SimDominance
M+S (irr)
DECS
M+S
DE-PDB
CLone
Django
blind
ReachLunch
h++
iProverPlan
## Results

<table>
<thead>
<tr>
<th>bottleneck</th>
<th>pegsol</th>
<th>over-tpp</th>
<th>over-nomystery</th>
<th>over-rovers</th>
<th>sliding-tiles</th>
<th>doc-transfer</th>
<th>tetris</th>
<th>diagnosis</th>
<th>bag-transport</th>
<th>chess-pebbling</th>
<th>bag-grripper</th>
<th>bag-barman</th>
<th>pegsol-row5</th>
<th>cave-diving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Aidos1    |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| Aidos2    |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| Aidos3    |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| SymPA (irr) |      |          |                |              |               |              |        |            |                |                |              |             |             |             |
| SymPA     |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| SimDominance |     |          |                |              |               |              |        |            |                |                |              |             |             |             |
| M+S (irr) |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| DECS      |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| M+S       |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| DE-PDB    |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| CLone     |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| Django    |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| blind     |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| ReachLunch |      |          |                |              |               |              |        |            |                |                |              |             |             |             |
| h++       |        |          |                |              |               |              |        |            |                |                |              |             |             |             |
| iProverPlan |     |          |                |              |               |              |        |            |                |                |              |             |             |             |
Results

<table>
<thead>
<tr>
<th>bottleneck</th>
<th>pegsol</th>
<th>over-tpp</th>
<th>over-nomystery</th>
<th>over-rovers</th>
<th>sliding-tiles</th>
<th>doc-transfer</th>
<th>tetris</th>
<th>diagnosis</th>
<th>bag-transport</th>
<th>chess-pebbling</th>
<th>bag-gripper</th>
<th>bag-barman</th>
<th>pegsol-row5</th>
<th>cave-diving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aidos1
Aidos2
Aidos3
SymPA (irr)
SymPA
SimDominance
M+S (irr)
DECS
M+S
DE-PDB
CLone
Django
blind
ReachLunch
h++
iProverPlan
## Unknown Known Results

### Table: Unknowns solved (Part 1)

<table>
<thead>
<tr>
<th>Domain Team</th>
<th>SymPA (irr)</th>
<th>SimDominance</th>
<th>SymPA</th>
<th>CLone</th>
<th>DECS</th>
</tr>
</thead>
<tbody>
<tr>
<td>document-transfer</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>diagnosis</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td><strong>7</strong></td>
<td><strong>4</strong></td>
<td><strong>3</strong></td>
<td><strong>2</strong></td>
<td><strong>2</strong></td>
</tr>
</tbody>
</table>

### Table: Unknowns solved (Part 2)

<table>
<thead>
<tr>
<th>Domain Team</th>
<th>Aidos3</th>
<th>Aidos2</th>
<th>Aidos1</th>
<th>Django</th>
<th>M+S (irr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>document-transfer</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>diagnosis</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td><strong>1</strong></td>
<td><strong>1</strong></td>
<td><strong>1</strong></td>
<td><strong>1</strong></td>
<td><strong>1</strong></td>
</tr>
</tbody>
</table>
Next Steps

To be released on unsolve-ipc.eng.unimelb.edu.au

- Git repo of solvers / domains / eval setup / etc
- DataJoy project of all the statistics / analysis
- Planner descriptions (extended abstracts)
Next Steps

To be released on unsolve-ipc.eng.unimelb.edu.au

- Git repo of solvers / domains / eval setup / etc
- DataJoy project of all the statistics / analysis
- Planner descriptions (extended abstracts)

Unsolve-IPC-2018

Any takers?
Thanks again to all who participated, submitted domains, provided server support, played devil’s advocate\(^1\), encouraged the endeavour, etc.

\(^1\)if LAMA doesn’t solve in 30min; return "unsolvable"