Local Search Topology
Implications for planner performance

Mark Roberts
Adele Howe
Colorado State University
Fort Collins, Colorado
Hoffmann’s Topological Analyses: The Taxonomy

**MINIMA**

- M1
- M0
- Mx

**NO MINIMA**

- med < c

<table>
<thead>
<tr>
<th>M1</th>
<th>BW400ap</th>
<th>Schedule</th>
<th>Mystery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hanoi</td>
<td></td>
<td>Mprime</td>
</tr>
<tr>
<td></td>
<td>BW-30p</td>
<td></td>
<td>Mic-ADL</td>
</tr>
<tr>
<td></td>
<td>Fridge</td>
<td></td>
<td>Freecell</td>
</tr>
<tr>
<td></td>
<td>Breitcase</td>
<td></td>
<td>Assembly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>M0</th>
<th>Logistics</th>
<th>TSP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ferry</td>
<td>Movie</td>
</tr>
<tr>
<td></td>
<td>Gripper</td>
<td>TSP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mx</th>
<th>Mic-SIM</th>
<th>Mic-STR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Movie</td>
<td>TSP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DC</th>
<th>DH</th>
<th>DR</th>
<th>DU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undirected</td>
<td>Harmless</td>
<td>Recognized</td>
<td>Unrecognized</td>
</tr>
</tbody>
</table>

Hoffmann’s Topological Analyses: Results

- Empirical results show FF performance followed the taxonomy very well
- Theoretical results prove the taxonomy class for each domain under $h+$ and $h_{ff}$
Hoffmann’s Topological Analyses: Limitations

- Computing $h^+$ is *NP*-hard
  - Problem instances had to be small
- Applied to a single planner: FF

The findings are convincing for FF
Do they transfer to other planners?

Questions

- Does the taxonomy distinguish performance for HSh+ planners?
- What about non-HSh+ planners?
Approach

- Collect
  - publicly available planners
  - benchmark problems in the 20 domains
  - performance of the planners on problems
- Analyze taxonomy effect on performance of HS planners using $h^+$
- Analyze taxonomy effect on performance of non-HS planners

Setup: Variables

- Predictors (Independent)
  - 910 Problems (290 “challenging” problems)
  - 28 Planners (10 HSh+, 18 non-HSh+)
  - Taxonomy Category
- Responses (Dependent)
  - Runtime : [0,1800] seconds
  - Success : { yes, no }
Setup: Success Ratio

- Construct a contingency table
- Perform a G-test (exact version of $\chi^2$)

\[
\begin{array}{c|cc|c|cc}
   & S & F & & S & F \\
MX & 249 & 39 & & DC & 177 & 29 \\
M0 & 10 & 2 & & DH & 159 & 38 \\
M1 & 516 & 94 & & DR & 150 & 0 \\
DU & 289 & 68 & & & & \\
\end{array}
\]

FF 2.3, all problems

G = 0.58, p = 0.75
G = 55.81, p << 0.001

Setup: Runtime

- Split by planner, taxonomy, and time
- Construct ANOVA: TTC, TTS, TTF

\[
\frac{SSR/df}{SSE/df} = \frac{MSR}{MSE} = F^*
\]
Setup: Runtime

- Significant ANOVA justifies pair-wise analysis
- TukeyHSD determines differences

<table>
<thead>
<tr>
<th></th>
<th>DC</th>
<th>DH</th>
<th>DR</th>
<th>DU</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC</td>
<td>--</td>
<td>cs</td>
<td>cs</td>
<td></td>
</tr>
<tr>
<td>DH</td>
<td>cf</td>
<td>--</td>
<td>cs</td>
<td></td>
</tr>
<tr>
<td>DR</td>
<td>c</td>
<td>c</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>DU</td>
<td></td>
<td></td>
<td></td>
<td>--</td>
</tr>
</tbody>
</table>

c=complete
s=success
f=failure

The presence of a letter shows a similarity.

Taxonomy versus HSh+

- All G-tests and ANOVAs significant
- Pair-wise comparison
  - TTF predictable regardless of taxonomy
  - Challenging problems: TTS does not depend on dead-end type
  - Extremes of taxonomy distinguish performance (except point two above)

Provisionally state that the taxonomy does distinguish performance for HSh+
Taxonomy versus non-HSh+

- Taxonomy lacks effect for success ratio
- TTS does not depend on taxonomy
- Taxonomy effect for TTF on both categories

Provisionally state that the taxonomy does **not** distinguish performance for non-HSh+

Limitations

- Planner
  - Switching search methods
  - Optimal versus satisficing
  - Grouped versus individual planners
- Problem
  - Inter-domain difficulty
  - Low cell counts
Summary

- Applied statistical hypothesis testing to determine effect of a model in explaining performance
  - Taxonomy explains HSh+
  - Taxonomy fails to explain non-HSh+
- Further work to deal with limitations

Future Work

- Extend to newer problems
  - Problem generators
- Probe inter-domain difficulty
- Better control across planner families
- Link to results on domain complexity
et al.

- Funding from CSU, NSF, AFOSR
- MEPS group: Landon Flom, Christie Williams, Crystal Redman
- CSU Student Group: Mark Rogers, Andrew Sutton, Artem Sokolov, Keith Bush, Laura Barbulescu, and others
- ICAPS 2005/6 attendees: Too many names
- The International Planning Competition, public planners and problems

Planner Success Sammon Map for all data

[Diagram showing data points and labels]