Directing a Portfolio with Learning

We leverage learned models of success and runtime to rank a suite of planning systems and allocate runtime to them.

Mark Roberts  Adele Howe
markroberts@cs.colostate.edu howe@cs.colostate.edu

Grant IIS-0138690

We leverage learned models of success and runtime to rank a suite of planning systems and allocate runtime to them.

Experiment Design

23 classical planners from 1991 to 2006:
- 6 Partial Order
- 3 Model-checking
- 9 Heuristic Search (Relaxed Graphplan)
- 3 Graphplan
- 2 Hybrid

Experiment Design

Over 4000 benchmark problems from competitions, previous research, and planner distributions

Almost 60 features:
- Domain/Problem instance
- Operator Interaction
- State-space topology

Model runtime and success for each planner

Feature Selection

Q: Can we minimize feature cost and maintain model accuracy?

<table>
<thead>
<tr>
<th>Feature Cost</th>
<th>Problem Features</th>
<th>Language Features</th>
<th>Operator Features</th>
<th>State Space Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>13</td>
<td>5</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

High cost features do not increase accuracy
All models use the low cost features (in dark blue)

Modeling Runtime

Q: How accurately can runtime be modeled?

Alternative Models:
- Decision Trees - 94.3% average accuracy on log(time) bins
- Linear Regression - average $r^2 = .41$

Modeling Success

Q: How accurately can success be modeled?

Success is a binary class
Decision Trees achieve 95.5% average accuracy
- Estimate probability using leaf node counts
BayesNets achieve 90.1% average accuracy
- Faster to build and gives direct probability

Portfolio Construction

1) Remove dominated planners and rank by Decision Tree success models
2) Schedule ranked planners:
   - Round 1: run best five planners for 10 seconds
   - Rounds 2+: run all live planners 100 more seconds
3) Terminate if:
   - planner succeeds (portfolio success)
   - time runs out (portfolio failure)
   - no planners are alive to run (portfolio failure)

Portfolio Performance

The portfolio
- Solves 88% of the problems it could have solved with unlimited time
- Is an average of 6 seconds faster than the average planner, as shown in figure
- Is an average of 3 seconds slower than best planner

Portfolio Performance

The portfolio
- Solves 88% of the problems it could have solved with unlimited time
- Is an average of 6 seconds faster than the average planner, as shown in figure
- Is an average of 3 seconds slower than best planner

Extending the Portfolio

More sophisticated learning /prediction techniques
- Automatic, cost-sensitive feature selection
- Extend problems, features, and planners
- Extend ranking and allocation strategies

Extending the Portfolio

More sophisticated learning /prediction techniques
- Automatic, cost-sensitive feature selection
- Extend problems, features, and planners
- Extend ranking and allocation strategies

Our long-term goal is to identify dependencies between features, domains, heuristics, and algorithms to explain why one algorithm is favored over another.