Motivation

Task: Stochastic shortest path
- reaching some goal state when the effects of actions are stochastic
- special case of planning
- subclass of Markov Decision Problems
- medium size: fully enumerated state space

Here: Multiple SSPs with the same domain

Goal: Speed up using Abstractions
- construct a multi-level hierarchy of progressively simpler abstractions
- find a policy for the most abstract level, then recursively refine into a solution to the original problem.

Features:
- Options-based abstraction
- Multiple levels
- Deterministic abstractions

Algorithm

Building abstractions

Planning
- Build a region around the goal, solve at the ground level
- Plan in the abstract graph
- Follow the ground options to execute the plan
- Follow the ground solution when entering the goal region

Motivation

Related work

<table>
<thead>
<tr>
<th>Problem</th>
<th>Diff.</th>
<th>HIP</th>
<th>SDP</th>
<th>LASP</th>
<th>ASAP</th>
<th>MDLearner</th>
<th>Asymptotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost aggregation</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Automatically built abstractions</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Options</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Option discovery</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lifted policy perf. bound</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Deterministic abstractions</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Experiments with</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Path planning for agents in commercial video games (uncertainty modeled by map congestion)

Theory

Result

Processes:
Lifting a policy π of M to M: π → L(π)

Theorem 1: The expected error of the lifted value function of the abstract policy π, relative to the base level policy π, is small if (i) aggregation does not lose details of π and (ii) the projected costs and transitions underlying π are matched by the costs (resp., transitions) associated with L(π).

Result

Theorem 2 (Simulation):
Interpretation: We can accurately simulate π with some policy L(π) of the abstract MDP provided all of the terms are small.

Abstractions

Definition: M = (X, A, p, c, g) original MDP
- M = (A, p, c, g) abstract MDP

Suboptimality versus the solution time ratio as compared to IPS for different parameter configurations. The dominant configurations are shown for different levels of abstraction.

Suboptimality versus the solution time ratio as compared to IPS for different values of δ

Cost versus solution time for IPS and abstraction at different values of δ

Solution times for several game maps.