Objective: Develop a fast, scalable, offline HighProb probabilistic planner.

Approach: Leverage state-of-the-art techniques in Fully Observable Non-Deterministic Planning (FOND) planning.

Contribution: HighProb probabilistic planner ProbPRP that outperforms previous state of the art.

Exploring High-Likelihood Plans

State-action π = Σ1, 2, 3, ..., n, 1, 2, 3

Likelihood of π: \( L_π = \prod_{i=1}^{n} P(s_i, a_i, s_{i+1}) \)

• Search bias in GenerateWeakP to favour exploration of high-likelihood plans

Final Full Exploration

• Best quality policy P found so far is selected
• All states reachable by P are fully explored

ProbPRP extends core FOND algorithm PRP to obtain high-quality HighProb solutions to probabilistic planning problems.

PRP \implies ProbPRP

Safety Belt

The safety belt mechanism gradually disables a feature when it is detected not to be beneficial.
• A safety belt is used for strong-cyclic detection.

Early Deadend Detection

Discard plan π = Σ1, 2, 3, ..., n when one of the outcomes of s_i by a_i is a recognised deadend.

Example: The River Problem [Little and Thiebaux, 2007]

The agent has two options to cross a river: (i) traverse a path of slippery rocks with a 25% chance of success, (ii) swim from one side of the river to the other, with a 50% chance of reaching a small island. In the latter case, she can swim towards the other side of the river with an 80% chance of success and a 20% chance of drowning; (ii) swim from one side of the river to the other, with a 50% chance of success, and a 50% chance of falling in.

Question: Which strategy maximizes the probability of success?

High-Quality Solutions

ProbPRP finds high-quality solutions that outperform the state of the art.

Optimality ProbPRP is guaranteed to find optimal HighProb solutions when all of the dead-ends are avoidable.

Small Size ProbPRP inherits the compact policy representation from PRP.

Short Plans ProbPRP solutions do not rely on highly improbable events to reach the goal.

Conclusions

• Highlighted and explored core similarities between Probabilistic Planning and FOND
• Introduced HighProb planner ProbPRP
• ProbPRP outperforms previous state of the art:
  • offline policy computation
  • improved optimality guarantees.
  • improved performance and scalability
    • (including run time, coverage, policy size, plan length)

Future Work

• Anytime behaviour
• Improved probabilistic reasoning


http://www.haz.ca/research/probprp