Planning with Uncertainty

RYAN SPRING

12/2/2014

Introduction

- Markov Decision Process (MDP) (S, A, T, R)
  - **Markov Property** – The future is dependent only on the current state and action
  - A set of states, S
  - A set of action, A
  - Transition Probability Function $T(s'|s,a)$
  - Reward Function R(s)

A policy for action uncertainty instead of a plan from start to goal

A policy specifies an action that will lead to the goal from a given state $\pi\colon S \rightarrow A$

Problem

Steerable Medical Needle

Dubins Car with a bang-bang controller

Turn left and right

Move forward distance δ every time step

Action Uncertainty is represented using a Gaussian distribution

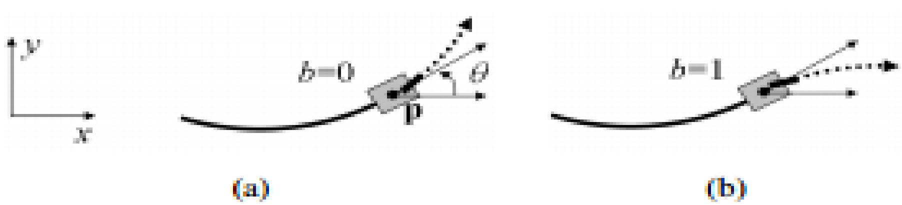
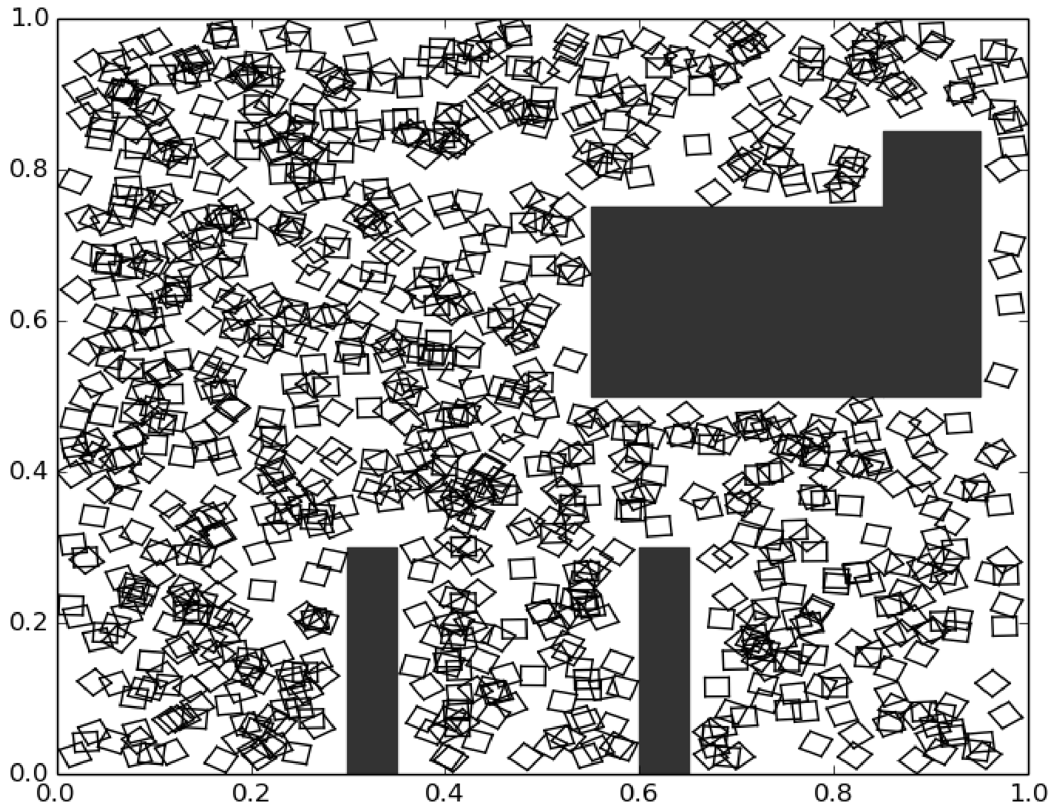


Fig. 3. The state of a bang-bang steering car is defined by point **p**, orientation θ, and turning direction b (a). The car moves forward along an arc of constant curvature and can turn either left (a) or right (b).

Stochastic Motion Roadmap

1. Sample configuration space
2. Generate an approximation of the action uncertainty
3. MDP Value Iteration



Build SMR Graph

Sample n states from the configuration space to form the vertices of the graph – n = 50,000

Transition Model

Generate a sample from a node i using an action U_i

If the sample is valid, find the nearest node j in the SMR graph

If the sample has a collision, then there is a transition to a special obstacle state

Sample each action m times – m = 20

$P_{ij}(U_i) = \frac{v}{m}$ where v is the number of times the transition from I to J occurs

MDP Value Iteration

Reward 1 for Goal States

Reward 0 for Obstacle States

Use an epsilon ε threshold to terminate value iteration - $\varepsilon = \sim 10^{-7}$

P_{ij} - Transition Probability from I to J

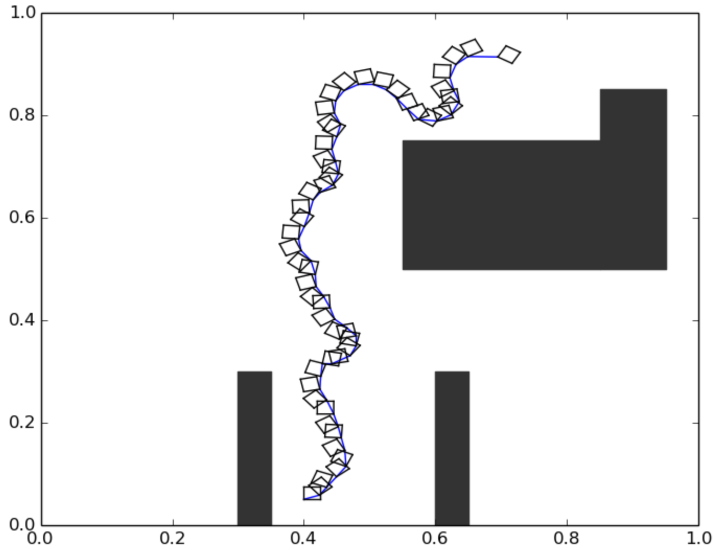
$g(i, u_i, j)$- Reward for transitioning from I to J in state u_j

- A negative reward creates a preference for shorter paths

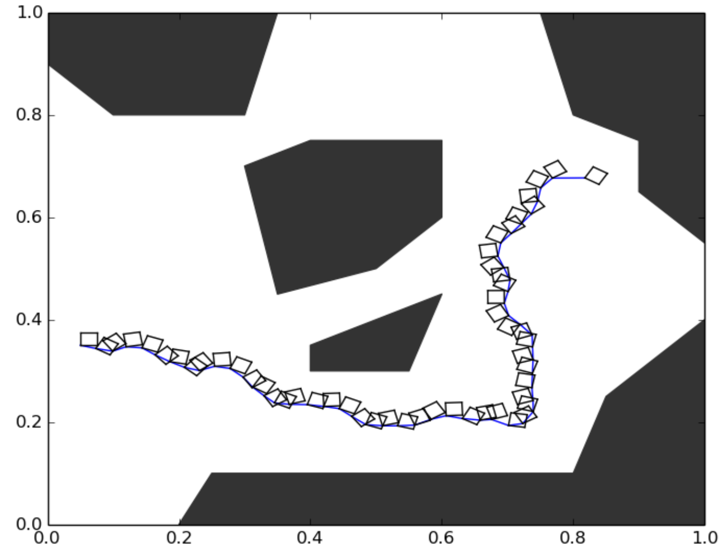
$J^*(j)$ - Probability of success for state j

$$J^*(i) = max_{U_i} \sum_{j \epsilon V} P_{ij}(U_i)\,(g(i, u_i, j) + J^*(j))$$

Results

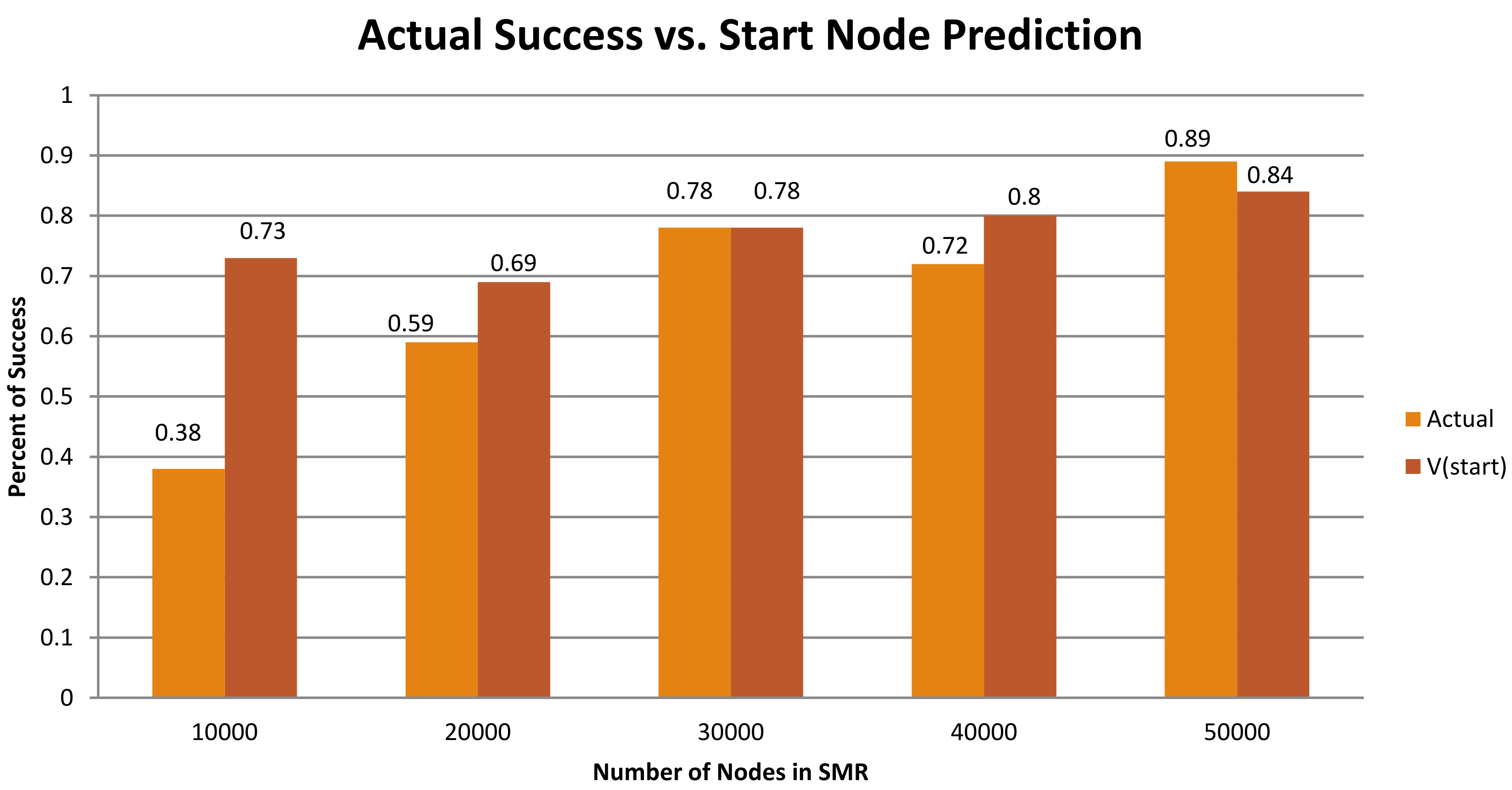


Environment 1



Environment 3

Results

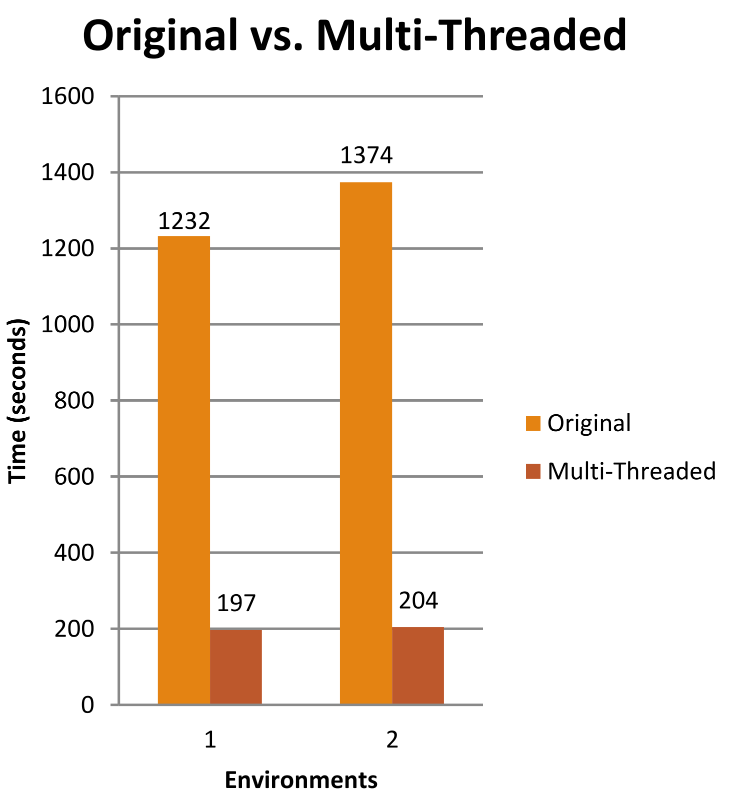


Actual Success vs. Start Node Prediction
Percent of Success
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.38
0.73
0.59
0.69
0.78
0.78
0.72
0.8
0.89
0.84
10000
20000
30000
40000
50000
Number of Nodes in SMR
Actual
V(start)

Extensions

The main bottleneck was creating the transition model and performing the MDP Value iteration

About 6x performance improvement with Multi-threaded Support



Motivation/Challenges

MDPs are everywhere!

- i.e. Robotics, Machine Learning, Sensing, Control Theory

Sanity Check:

- First, solve the problem without action uncertainty
- Then, incrementally add uncertainty to the problem