Learning and Optimizing Probabilistic Models for Planning under Uncertainty

R. van Bekkum

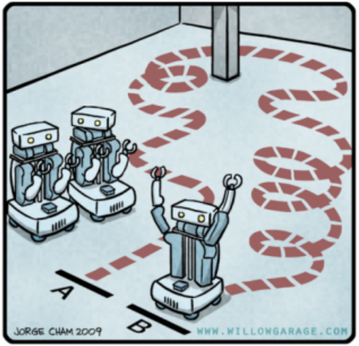
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

September 27, 2017

Planning under Uncertainty

Domain Characteristics

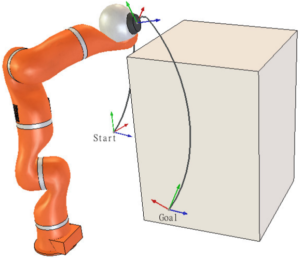
- A system is controlled by one or more *agents*
- *Uncertain domain dynamics*, i.e. uncertainty may be present in:
  - Execution of actions (e.g., robot may slip)
  - Exogenous factors (e.g., doors open/closed)
  - Percepts (e.g., sensor noise)
- *Sequential decision making*
  - Non-myopic agents
  - Selection of actions with high future pay-off



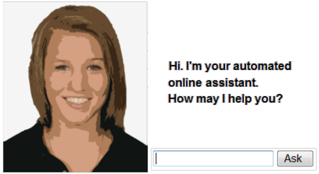
Planning under Uncertainty

Example Domains

Path planning



Motion planning

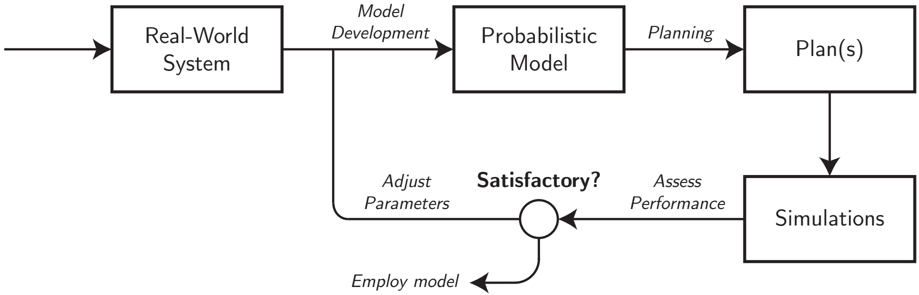


Dialog management systems

Image credits: Hawes et al. [1], Iocchi et al. [2], V-REP Manual, Bemidji State University

Planning under Uncertainty

Typical Development Routine



Planning under Uncertainty

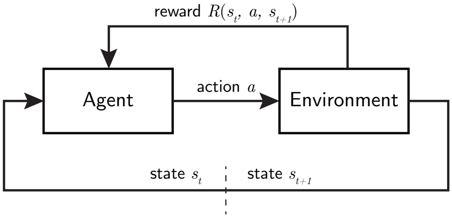
Markov Decision Processes (MDPs)

In DTP systems are modeled by probabilistic models, e.g. MDPs:

Definition

An MDP is a 5-tuple $\mathcal{M} = (\mathcal{S}, s_0, A, \delta, R)$:

- $\mathcal{S}$ is the state-space, $s_0 \in \mathcal{S}$ the initial state
- A is the action-space
- $\delta : \mathcal{S} \times A \times \mathcal{S} \mapsto [0, 1]$ is the transition function
- $R : \mathcal{S} \times A \times \mathcal{S} \mapsto \mathbb{R}$ is the reward function



Planning under Uncertainty

Model Development

Problem: How to find a suitable MDP?

Planning under Uncertainty

Model Development

Problem: How to find a suitable MDP?

Classical approach: Model development by a *human designer*, however:

- Requires significant effort (e.g., trial-and-error)
- Typically demands knowledge/experience, accompanied by high costs

Planning under Uncertainty

Model Development

Problem: How to find a suitable MDP?

Classical approach: Model development by a *human designer*, however:

- Requires significant effort (e.g., trial-and-error)
- Typically demands knowledge/experience, accompanied by high costs

Alternative: Use Reinforcement Learning instead of Planning, however:

- Requires direct interaction with environment
- One might require *reusable* models, applicable for multiple tasks

Problem Description

Problem Statement

Problem Description

Related Work

Problem Description

Research Questions

Main Research Question

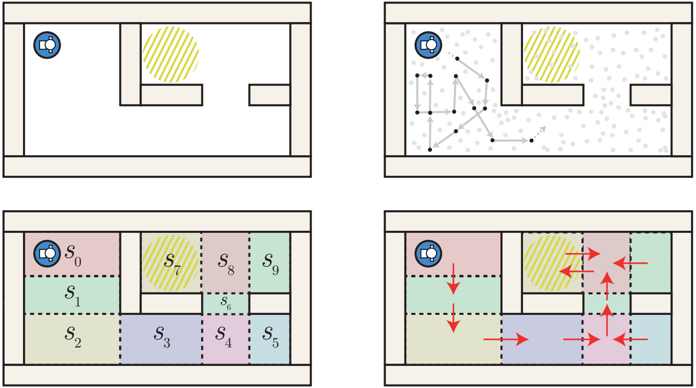
How can the task of obtaining an MDP that maximizes the yielded performance of executing plans that are derived from it, given a dataset about the system under consideration, be automated?

Proposed Solution

General Idea

Proposed Solution

Running Example: Path Planning for Mobile Robot Navigation



Proposed Solution

Performance Measure

- Value Function
- Simulations
- Lorem ipsum dolor sit amet

$$V_{\mathcal{M}} = \frac{\sum_{t \in T_{\mathcal{M}}} \beta \cdot V_{DTP,t} + (1 - \beta) \cdot V_{SIM,t}}{|T_{\mathcal{M}}|}$$

Proposed Solution

Base Framework

Proposed Solution

Multi-Phase Framework

Experimental Results

Conclusions and Recommendations

Learning and Optimizing Probabilistic Models for Planning under Uncertainty

R. van Bekkum

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

September 27, 2017

References

N. Hawes, C. Burbridge, F. Jovan, L. Kunze, B. Lacerda, L. Mudrová, J. Young, J. Wyatt, D. Hebesberger, T. Körtner, et al.
The strands project: Long-term autonomy in everyday environments.
arXiv preprint arXiv:1604.04384, 2016.

L. Iocchi, L. Jeanpierre, M. T. Lazaro, and A.-I. Mouaddib.
A practical framework for robust decision-theoretic planning and execution for service robots.
In *ICAPS*, pages 486–494, 2016.

D. Nikovski and I. R. Nourbakhsh.
Learning probabilistic models for decision-theoretic navigation of mobile robots.
In *ICML*, pages 671–678. Citeseer, 2000.

A. Stolcke and S. M. Omohundro.
Best-first model merging for hidden markov model induction.
arXiv preprint cmp-lg/9405017, 1994.