

A POMDP Formulation of Preference Elicitation Problems

Craig Boutilier



Discussion Layout

- Definitions
- Preference Elicitation
- Problem Definition
- Validity of Assumptions
- Alternative Approaches
- Utility Functions and Game Theoretic Extensions

Definitions

- **Markov Decision Process (MDP):** “the process of calculating an optimal policy in an accessible, stochastic environment with a known transition model”
- **Fully Observable Markov Decision Processes** We get a fully observable Markov Decision Process if the agent is acting in an
 - accessible environment
 - at each step, the agent knows what state it is in
 - and what actions are available to it

Partially observable Markov Decision Process

- A Markov Decision Process operating in an inaccessible environment
- Sequential Decision Process
- Agent tries to maximize some Utility Function WITHOUT knowledge of the system
- Optimization/Learning (NN)?

More definitions

- **Discount Factor:** a constant $0 \leq \gamma \leq 1$ “that determines the relative value of delayed vs. immediate rewards”

Preference Elicitation

- What
 - Knowledge about User's preferences
- Why
 - Applications

Problem Definition

- Given to the System
 - Set of Choices (e.g. actions, policies, recommendations)
 - Effects of the choices
- Aim of the System
 - Take a decision with MAX expected utility
- Problem translated to:-
 - Finding “enough” about $\langle U \rangle$

An Example

- MP3 Player
- Car

PE as POMDP

- vector $\langle U \rangle$ never changes
- Gamble Queries
- Effective? Alternative Querying types?

Alternative approach

- POMDP vs EVOI
- Worth the Tradeoff? (computational complexity vs accuracy)

Assumption Space

- Value (utility) functions – Continuous
- POMDP state space – discrete
- Solution?
- Validity?
- Practical approximations (component pruning based on weights sec4.2)

Utility

- System Utility rather than User Utility?
- Game Theoretic Extensions?