Decision Theory: Markov Decision Processes

CPSC 322 – Decision Theory 3

Textbook §12.5

Lecture Overview

1. Recap
2. Finding Optimal Policies

Sequential decision problems

- A sequential decision problem consists of a sequence of decision variables $D_1, \ldots, D_n$.
- Each D_i has an information set of variables pD_i, whose value will be known at the time decision D_i is made.

- What should an agent do?
  - What an agent should do at any time depends on what it will do in the future.
  - What an agent does in the future depends on what it did before.

Policies

- A policy specifies what an agent should do under each circumstance.
- A policy is a sequence $\delta_1, \ldots, \delta_n$ of decision functions

$$\delta_i : dom(pD_i) \rightarrow dom(D_i).$$

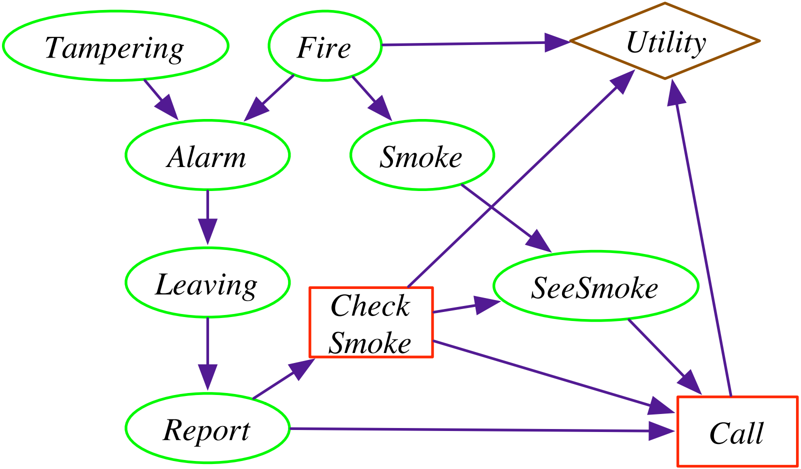
  This policy means that when the agent has observed $O \in dom(pD_i)$, it will do $\delta_i(O)$.
- The expected utility of policy δ is

$$\mathbb{E}(U|\delta) = \sum_{\omega \models \delta} P(\omega)U(\omega)$$

- An optimal policy is one with the highest expected utility.

Lecture Overview

Decision Network for the Alarm Problem



Finding the optimal policy

- Remove all variables that are not ancestors of a value node
- Create a factor for each conditional probability table and a factor for the utility.
- Sum out variables that are not parents of a decision node.
- Select a variable D that is only in a factor f with (some of) its parents.
    - this variable will be one of the decisions that is made latest
- Eliminate D by maximizing. This returns:
    - the optimal decision function for D, $\arg\max_D f$
    - a new factor to use in VE, $\max_D f$
- Repeat till there are no more decision nodes.
- Sum out the remaining random variables. Multiply the factors: this is the expected utility of the optimal policy.

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there?

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there? 2^k

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there? 2^k
- If there are b possible actions, how many different decision functions are there?

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there? 2^k
- If there are b possible actions, how many different decision functions are there? b^{2^k}

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there? 2^k
- If there are b possible actions, how many different decision functions are there? b^{2^k}
- If there are d decisions, each with k binary parents and b possible actions, how many policies are there?

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there? 2^k
- If there are b possible actions, how many different decision functions are there? b^{2^k}
- If there are d decisions, each with k binary parents and b possible actions, how many policies are there? $\left(b^{2^k}\right)^d$

Complexity of finding the optimal policy

- If a decision D has k binary parents, how many assignments of values to the parents are there? 2^k
- If there are b possible actions, how many different decision functions are there? b^{2^k}
- If there are d decisions, each with k binary parents and b possible actions, how many policies are there? $\left(b^{2^k}\right)^d$
- Doing variable elimination lets us find the optimal policy after considering only $d \cdot b^{2^k}$ policies
  - The dynamic programming algorithm is much more efficient than searching through policy space.
  - However, this complexity is still doubly-exponential—we'll only be able to handle relatively small problems.