# Decision Theory: Single Decisions

CPSC 322 Lecture 31

March 27, 2006
Textbook §12.2

# Lecture Overview

Recap

Decision Problems

Single Decisions

# Decisions Under Uncertainty

- An agent's decision will depend on:
  1. what actions are available
  2. what beliefs the agent has
     - note: this replaces "state" from the deterministic setting
  3. the agent's goals
     - a richer notion of goals: rating how happy the agent is in different situations

# Preferences Over Outcomes

If $o_1$ and $o_2$ are outcomes

- $o_1 \succeq o_2$ means $o_1$ is at least as desirable as $o_2$.
  - read this as "the agent weakly prefers $o_1$ to $o_2$"
- $o_1 \sim o_2$ means $o_1 \succeq o_2$ and $o_2 \succeq o_1$.
  - read this as "the agent is indifferent between $o_1$ and $o_2$."
- $o_1 \succ o_2$ means $o_1 \succeq o_2$ and $o_2 \not\succeq o_1$
  - read this as "the agent strictly prefers $o_1$ to $o_2$"

# Lotteries

- An agent may not know the outcomes of his actions, but may instead only have a probability distribution over the outcomes.
- A lottery is a probability distribution over outcomes. It is written
$$[p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]$$
where the $o_i$ are outcomes and $p_i > 0$ such that
$$\sum_i p_i = 1$$
- The lottery specifies that outcome $o_i$ occurs with probability $p_i$.
- We will consider lotteries to be outcomes.

# Preference Axioms

1. Completeness
2. Transitivity
3. Monotonicity
4. Continuity
5. Decomposability
6. Substitutivity

# Utility

If preferences satisfy the preceding axioms, then preferences can be measured by a function

$$U : outcomes \rightarrow [0, 1]$$

such that

1. $o_1 \succeq o_2$ if and only if $U(o_1) \geq U(o2)$.
2. Utilities are linear with probabilities:

$$U([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i \times U(o_i)$$

# Lecture Overview

Recap

Decision Problems

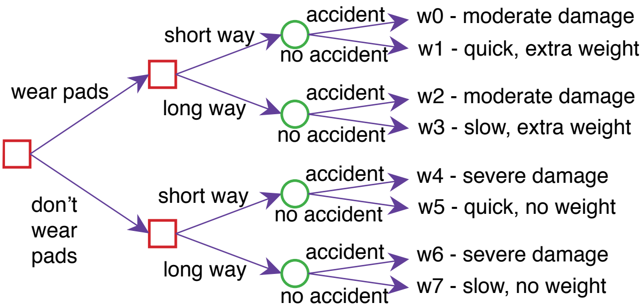
Single Decisions

# Decision Variables

- Decision variables are like random variables that an agent gets to choose the value of.
- A possible world specifies the value for each decision variable and each random variable.
- For each assignment of values to all decision variables, the measures of the worlds satisfying that assignment sum to 1.
- The probability of a proposition is undefined unless you condition on the values of all decision variables.

# Decision Tree for Delivery Robot

- The robot can choose to wear pads to protect itself or not.
- The robot can choose to go the short way past the stairs or a long way that reduces the chance of an accident.
- There is one random variable indicating whether there is an accident.



# Utility

- Utility: a measure of desirability of worlds to an agent.
  - Let $U$ be a real-valued function such that $U(\omega)$ represents an agent's degree of preference for world $\omega$.
- Simple goals can be specified by a boolean utility function:
  - worlds that satisfy the goal have utility $1$
  - other worlds have utility $0$
- Often utilities are more complicated. For example, in the delivery robot domain:
  - some function of the amount of damage to a robot
  - how much energy is left
  - what goals are achieved
  - how much time it has taken.

# Expected Utility

- The expected value of a function of possible worlds is its average value, weighting possible worlds by their probability.
- Suppose $U(w)$ is the utility of world world $w$.
  - The expected utility is

$$\mathcal{E}(U) = \sum_{\omega \in \Omega} P(\omega) \times U(\omega).$$

  - The conditional expected utility given $e$ is

$$\mathcal{E}(U|e) = \sum_{\omega \models e} P(\omega|e) \times U(\omega).$$

# Lecture Overview

Recap

Decision Problems

Single Decisions

# Single decisions

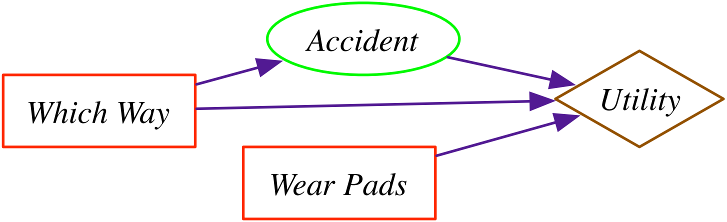
- Given a single decision variable, the agent can choose $D = d_i$ for any $d_i \in dom(D)$.
- The expected utility of decision $D = d_i$ is $\mathcal{E}(U|D = d_i)$.
- An optimal single decision is the decision $D = d_{max}$ whose expected utility is maximal:

$$d_{max} = \underset{d_i \in dom(D)}{\arg\max}\ \mathcal{E}(U|D = d_i).$$

# Single-stage decision networks

Extend belief networks with:

- Decision nodes, that the agent chooses the value for. Domain is the set of possible actions. Drawn as rectangle.
- Utility node, the parents are the variables on which the utility depends. Drawn as a diamond.




This shows explicitly which nodes affect whether there is an accident.

# Finding the optimal decision

- Suppose the random variables are $X_1, \ldots, X_n$, and utility depends on $X_{i_1}, \ldots, X_{i_k}$

$$\begin{aligned}\mathcal{E}(U|D) &= \sum_{X_1,\ldots,X_n} P(X_1,\ldots,X_n|D) \times U(X_{i_1},\ldots,X_{i_k}) \\ &= \sum_{X_1,\ldots,X_n} \prod_{i=1}^{n} P(X_i|pX_i) \times U(X_{i_1},\ldots,X_{i_k})\end{aligned}$$

# Finding the optimal decision

- Suppose the random variables are $X_1, \ldots, X_n$, and utility depends on $X_{i_1}, \ldots, X_{i_k}$

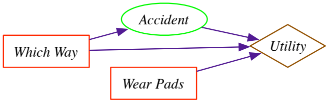
$$\mathcal{E}(U|D) = \sum_{X_1,\ldots,X_n} P(X_1, \ldots, X_n|D) \times U(X_{i_1}, \ldots, X_{i_k})$$

$$= \sum_{X_1,\ldots,X_n} \prod_{i=1}^{n} P(X_i|pX_i) \times U(X_{i_1}, \ldots, X_{i_k})$$

To find the optimal decision:

- Create a factor for each conditional probability and for the utility
- Sum out all of the random variables
- This creates a factor on $D$ that gives the expected utility for each $D$
- Choose the $D$ with the maximum value in the factor.

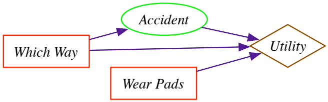
# Example Initial Factors




| Which Way | Accident | Probability |
|---|---|---|
| long | true | 0.01 |
| long | false | 0.99 |
| short | true | 0.2 |
| short | false | 0.8 |

| Which Way | Accident | Wear Pads | Utility |
|---|---|---|---|
| long | true | true | 30 |
| long | true | false | 0 |
| long | false | true | 75 |
| long | false | false | 80 |
| short | true | true | 35 |
| short | true | false | 3 |
| short | false | true | 95 |
| short | false | false | 100 |

# Example Initial Factors




| Which Way | Accident | Probability |
|---|---|---|
| long | true | 0.01 |
| long | false | 0.99 |
| short | true | 0.2 |
| short | false | 0.8 |

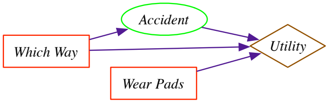
| Which Way | Accident | Wear Pads | Utility |
|---|---|---|---|
| long | true | true | 30 |
| long | true | false | 0 |
| long | false | true | 75 |
| long | false | false | 80 |
| short | true | true | 35 |
| short | true | false | 3 |
| short | false | true | 95 |
| short | false | false | 100 |

Sum out Accident:

| Which Way | Wear pads | Value |
|---|---|---|
| long | true | 0.01*30+0.99*75=74.55 |
| long | false | 0.01*0+0.99*80=79.2 |
| short | true | 0.2* 35+ 0.8*95=83 |
| short | false | 0.2*3+0.8*100=80.6 |

Thus the optimal policy is to take the short way and wear pads, with an expected utility of 83.

# Example Initial Factors




| Which Way | Accident | Probability |
|---|---|---|
| long | true | 0.01 |
| long | false | 0.99 |
| short | true | 0.2 |
| short | false | 0.8 |

| Which Way | Accident | Wear Pads | Utility |
|---|---|---|---|
| long | true | true | 30 |
| long | true | false | 0 |
| long | false | true | 75 |
| long | false | false | 80 |
| short | true | true | 35 |
| short | true | false | 3 |
| short | false | true | 95 |
| short | false | false | 100 |

Sum out Accident:

| Which Way | Wear pads | Value |
|---|---|---|
| long | true | 0.01*30+0.99*75=74.55 |
| long | false | 0.01*0+0.99*80=79.2 |
| short | true | 0.2* 35+ 0.8*95=83 |
| short | false | 0.2*3+0.8*100=80.6 |

Thus the optimal policy is to take the short way and wear pads, with an expected utility of 83.