Planning Under Uncertainty: Decision Theory 1 (One-off Decisions)

Shafiq Joty

Nanyang Technological University

srjoty@ntu.edu.sg

April 22, 2019

Overview

What is Planning?

How to **select and organize** a sequence of actions/decisions to achieve a given **goal**.

What is Planning?

How to **select and organize** a sequence of actions/decisions to achieve a given **goal**.

- **Goal:** A possible world/state in which some **propositions** are true (goal-based agent) or a state with a desired utility (utility-based agent).

What is Planning?

How to **select and organize** a sequence of actions/decisions to achieve a given **goal**.

- **Goal:** A possible world/state in which some **propositions** are true (goal-based agent) or a state with a desired utility (utility-based agent).
- If you have taken CZ 3005, we have already seen planning in deterministic environments:
  1. Search & CSP ⇒ single (macro) decisions.
  2. STRIPS ⇒ sequential decisions.

What is Planning Under Uncertainty?

How to select and organize a sequence of actions/decisions to **maximize the probability** of achieving a given goal.

What is Planning Under Uncertainty?

How to select and organize a sequence of actions/decisions to **maximize the probability** of achieving a given goal.

- In uncertain domains, we've so far only considered how to **represent** and **update** beliefs about domains.
  $\Rightarrow$ BayesNets, MRFs, Neural Nets, etc.
- What if an agent has to make decisions in a domain that involves uncertainty?
  $\Rightarrow$ (this is likely) One of the main reasons to represent the world probabilistically is to be able to use these beliefs as the basis for making decisions

What is Planning Under Uncertainty?

An agent's decision will depend on:

- what actions are available
- the agent's goals
- what beliefs (prob. distributions) the agent has (note: this replaces "state" from the deterministic setting)

What is Planning Under Uncertainty?

Representations:

An agent's decision will depend on:

- Beliefs ⇒ Belief Nets (BayesNets, MRFs, Neural Nets)
- Actions ⇒ Decision variables
- Goals ⇒ we'll move from all-or-nothing goals to a richer notion: rating **how happy** the agent is in different situations.
  ⇒ Utility based agent.

Putting these together, we'll extend belief networks to make a new representation called decision networks.

"Single" Action vs. Sequence of Actions

- **"Single" Action:** Set of primitive decisions that can be treated as a single macro (overall) decision to be made before performing the actions. Also known as **One-off** decisions
$\Rightarrow$ This lecture

"Single" Action vs. Sequence of Actions

- **"Single" Action:** Set of primitive decisions that can be treated as a single macro (overall) decision to be made before performing the actions. Also known as **One-off** decisions
  $\Rightarrow$ This lecture
- **Sequential Decisions:** Do the following three steps in iteration: (*i*) Observe, (*ii*) Make a decision, (*iii*) Act
  $\Rightarrow$ Next 3-4 lectures

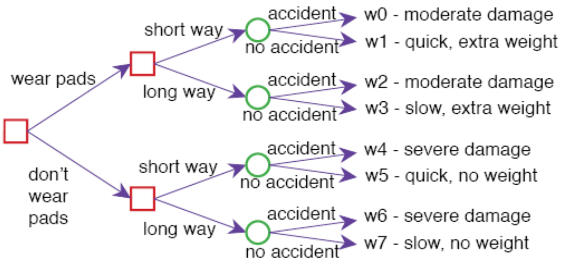
One-off decisions: Example

Delivery Robot: Robot needs to reach a certain room.

- Going through stairs may cause an accident.
- It can go the short way through long stairs, or the long way through short stairs (that reduces the chance of an accident but takes more time)
- It can choose to wear pads to protect itself or not (to protect itself in case of an accident) but pads slow it down
- If there is an accident, the robot does not get to the room

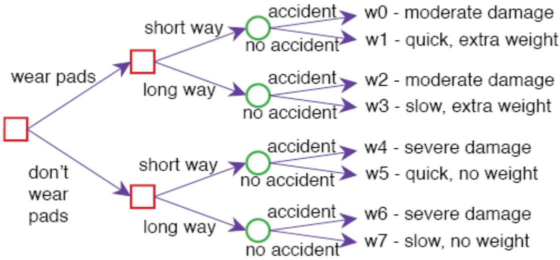
Decision and Random Variables

Delivery Robot: Decition Tree



Decision and Random Variables

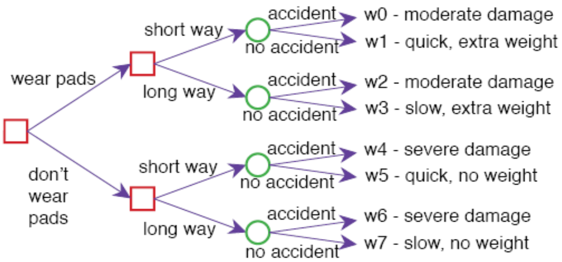
Delivery Robot: Decition Tree



- A sequence of decisions (square nodes) to make; or a macro-action that the agent can perform.

Decision and Random Variables

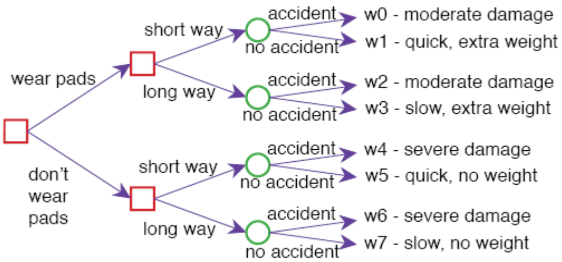
Delivery Robot: Decition Tree



- A sequence of decisions (square nodes) to make; or a macro-action that the agent can perform.
- Decisions can influence random variables (circle nodes).

Decision and Random Variables

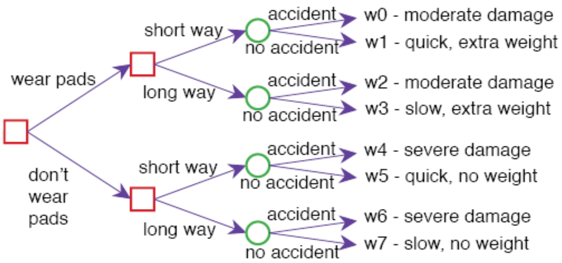
Delivery Robot: Decition Tree



- A sequence of decisions (square nodes) to make; or a macro-action that the agent can perform.
- Decisions can influence random variables (circle nodes).
- Decisions have **probability distributions** over outcomes (value of random variables and the resulting utilities).

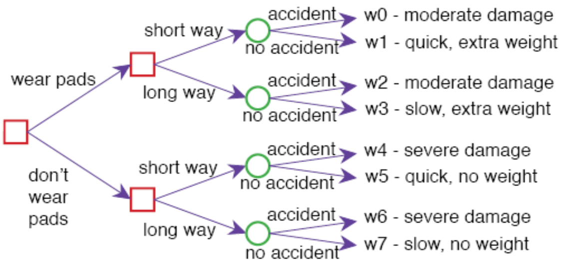
Decision and Random Variables

Delivery Robot: Decition Tree



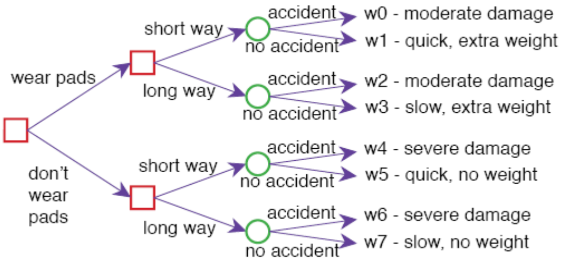
- Note that in this scenario, all decisions **precede** the random variable(s). We treat them as a single macro action.
- Random variables express probability distributions over outcomes/utility values (e.g., severe damage).

Decision and Random Variables



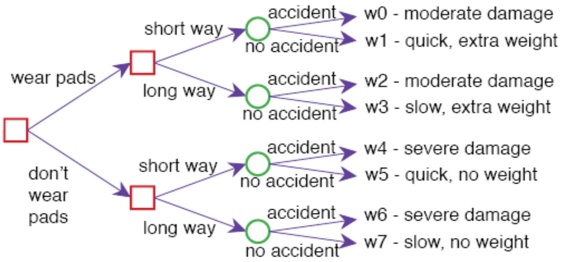
- Decision variables are like random variables whose value an agent gets to set.
- A **possible world** specifies a value for each random variable and each decision variable.

Decision and Random Variables



- For each assignment of values to all decision variables, the probabilities of the worlds satisfying that assignment sum to 1.
- In the example, $p(w_0) + p(w_1) = 1$; $p(w_2) + p(w_3) = 1$, etc.

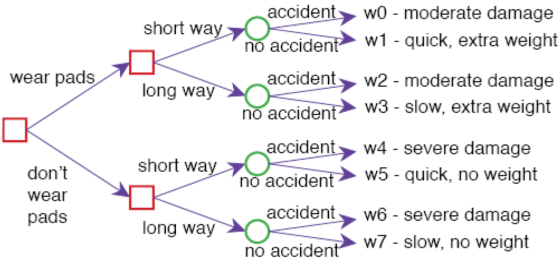
Decision and Random Variables



- For each assignment of values to all decision variables, the probabilities of the worlds satisfying that assignment sum to 1.
- In the example, $p(w_0) + p(w_1) = 1$; $p(w_2) + p(w_3) = 1$, etc.
- The probability of a proposition is undefined unless you condition on the values of all decision variables.

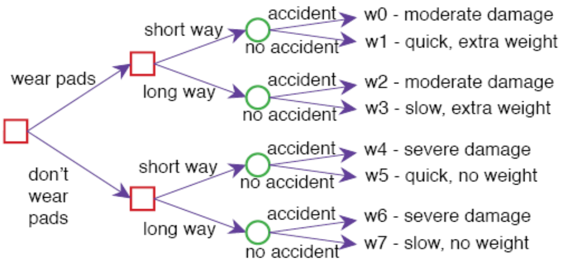
Role of Utility

What are the optimal decisions for our robot?



Role of Utility

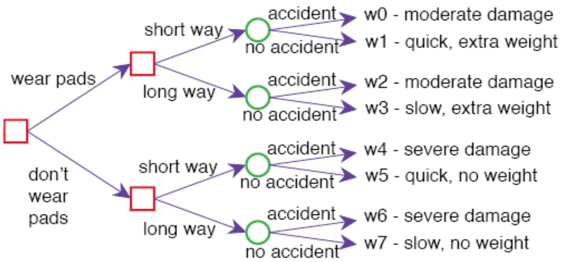
What are the optimal decisions for our robot?



- Depends on how happy the agent is in different situations
  $\Rightarrow$ utility values

Role of Utility

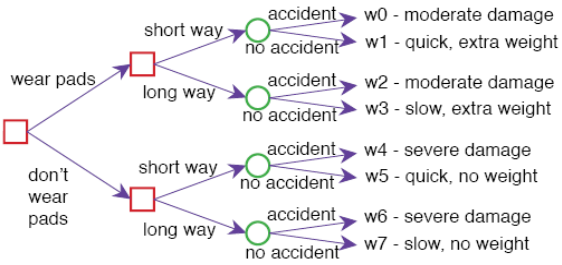
What are the optimal decisions for our robot?



- Depends on how happy the agent is in different situations
  $\Rightarrow$ utility values
- For sure, getting to the room is better than not getting there, but we need to consider other factors.

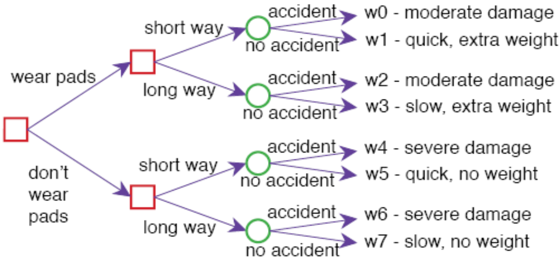
Role of Utility

Utility: A **measure of desirability** of possible worlds to an agent. Let U be a real-valued function such that $U(w)$ represents an agent's degree of preference for world w.



Role of Utility

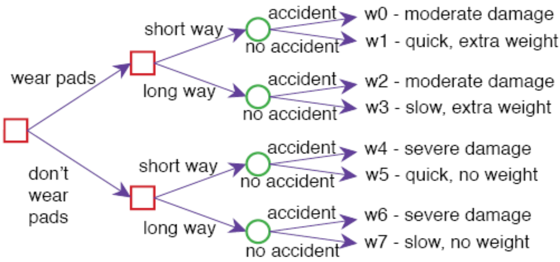
Utility: A **measure of desirability** of possible worlds to an agent. Let U be a real-valued function such that $U(w)$ represents an agent's degree of preference for world w.



$U(w_0) = 35$, $U(w_1) = 95$, $U(w_2) = 30$, $U(w_3) = 75$, $U(w_4) = 3$, $U(w_5) = 100$, $U(w_6) = 0$, $U(w_7) = 80$

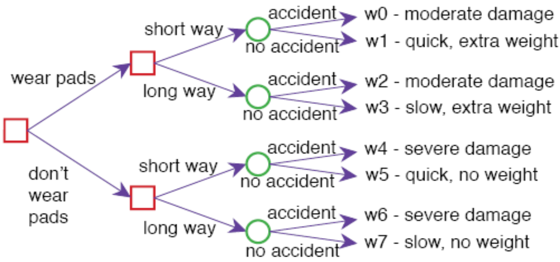
Role of Utility

How can the simple (boolean) goal "reach the room" be specified?



Role of Utility

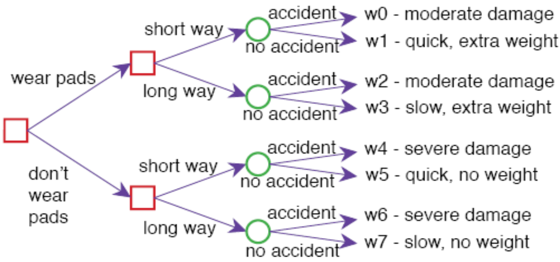
How can the simple (boolean) goal "reach the room" be specified?



Using a boolean utility function: worlds that satisfy the goal have utility 100 other worlds have utility 0.

Role of Utility

How can the simple (boolean) goal "reach the room" be specified?

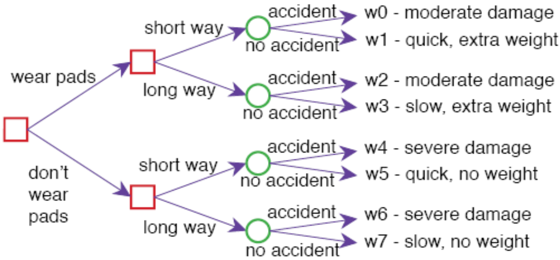


Using a boolean utility function: worlds that satisfy the goal have utility 100 other worlds have utility 0.

$U(w_1) = 100$, $U(w_3) = 100$, $U(w_5) = 100$, $U(w_7) = 100$

Optimal Decisions

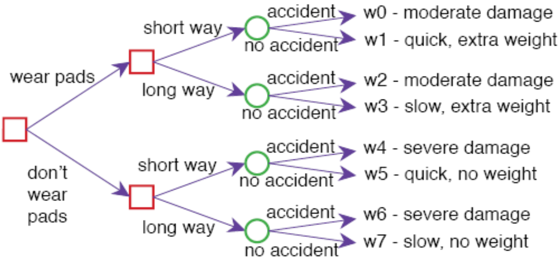
How to Combine Utility with Probability?



What is the utility of achieving a certain probability distribution over possible worlds?

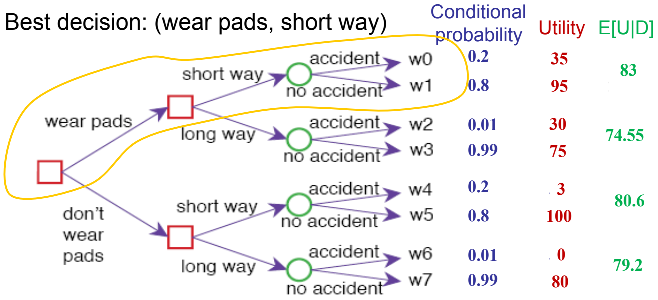
Optimal Decisions

How to Combine Utility with Probability?



What is the utility of achieving a certain probability distribution over possible worlds? $\Rightarrow$ It is its **expected utility** (average utility) by weighting possible worlds by their probability.

Optimal Decisions



What is the utility of achieving a certain probability distribution over possible worlds? $\Rightarrow$ It is its **expected utility** (average utility) by weighting possible worlds by their probability.

- Expected utility of Wear Pads and Short Way is: $p(w_0) * U(w_0) + p(w_1) * U(w_1) = 35 * 0.2 + 95 * 0.8 = 83$

Optimal Decisions

Formally, the expected utility of a decision sequence $\mathbf{d}_i \in \mathrm{dom}(d_1) \times \ldots \times \mathrm{dom}(d_n)$ is:

$$\mathbb{E}(U|\mathbf{d}_i) = \sum_j P(w_{i,j}) \times U(w_{i,j})$$

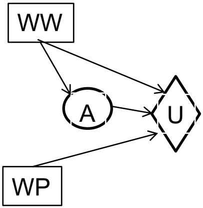
where $d_1, \ldots, d_n$ are the decision variables, and $w_{i,j}$ is a possible world for a decision sequence $\mathbf{d}_i$.

- The **optimal decision** is: $\mathbf{d}_k = \mathrm{argmax}_{\mathbf{d}_i}\, \mathbb{E}(U|\mathbf{d}_i)$

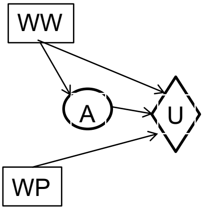
Single-Stage Decision Networks: Representation

Extend belief networks (contains random variables) with:

- **Decision nodes**, that the agent chooses the value for. Drawn as rectangle.
- **A Utility node**, the parents are the variables on which the utility depends. Drawn as a diamond.



Single-Stage Decision Networks: Representation



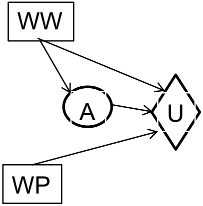
Compact and explicit representation

- **Compact:** each random/decision variable only occurs once
- **Explicit:** Shows explicitly which decision nodes affect random variables

Single-Stage Decision Networks: Reasoning

We can use Variable Elimination to find the optimal decision:

- Create a factor for each conditional probability and for the utility. *e.g.*, $P(A|WW)$ and $P(U|A, WW, WP)$.
- Multiply factors and sum out all of the random variables
- Choose the one with the maximum value in the factor



Single-Stage Decision Networks: Reasoning

Formally, suppose the random variables are $X_1, \ldots, X_n$

$$\mathbb{E}(U|D) = \sum_{X_1,\ldots,X_n} P(X_1,\ldots,X_n)U(pU) \quad (1)$$

$$= \sum_{X_1,\ldots,X_n} \prod_{i=1}^{n} P(X_i|pX_i)U(pU) \quad (2)$$

Single-Stage Decision Networks: Reasoning

Formally, suppose the random variables are $X_1, \ldots, X_n$

$$\mathbb{E}(U|D) = \sum_{X_1,\ldots,X_n} P(X_1,\ldots,X_n)U(pU) \tag{1}$$

$$= \sum_{X_1,\ldots,X_n} \prod_{i=1}^{n} P(X_i|pX_i)U(pU) \tag{2}$$

To find the optimal decision:

- Create a factor for each cond. probability and for the utility.
- Multiply factors and sum out all of the random variables.
- This creates a factor on D that gives the expected utility for each $\mathbf{d}_i \in D$.
- Choose the one with the maximum value in the factor

Learning Goals

- Compare and contrast stochastic **single-stage (one-off)** decisions vs. **multistage** decisions.
- Define a **utility function** on possible worlds
- Define and compute optimal one-off decision
- Represent one-off decisions as single stage decision networks and compute optimal decisions by Variable Elimination

The End