Decision Networks

Alice Gao

Lecture 16

Readings: RN 16.5 - 16.6. PM 9.1 - 9.2.

Outline

Learning Goals

By the end of the lecture, you should be able to

- Model a one-off decision problem by constructing a decision network containing nodes, arcs, conditional probability distributions, and a utility function.
- Determine the optimal policy of a decision network by computing the expected utility of every policy.
- Determine the optimal policy of a decision network by applying the variable elimination algorithm.

Decision Theory

- Decision theory = Probability theory + Utility theory.
  How should an agent act in an uncertain world?
- Probability theory:
  What should the agent believe based on the evidence?
- Utility theory:
  What does the agent want?

The principle of maximum expected utility:
A rational agent should choose the action that maximizes the agent's expected utility.

Decision Networks

Decision networks
$=$ *Bayesian network* $+$ *actions* $+$ *utilities*

A mail pick-up robot

The robot must choose its route to pick up the mail. There is a short route and a long route. On the short route, the robot might slip and fall. The robot can put on pads. Pads won't change the probability of an accident. However, if an accident happens, pads will reduce the damage. Unfortunately, the pads add weight and slow the robot down. The robot would like to pick up the mail as quickly as possible while minimizing the damage caused by an accident.

What should the robot do?

Variables

What are the random variables?

A : whether an accident occurs or not.

What are the decision variables (actions)?

S: whether the robot chooses the short route.

P: whether the robot puts on pads.

Nodes in a Decision Network

- Chance nodes
  represent random variables (as in Bayesian networks).
- Decision nodes
  represent actions (decision variables).
- Utility node
  represents agent's utility function on states (happiness in each state).

Robot decision network

chance nodes: Accident

decision nodes: Short

Pads

utility node: Utility

Arcs in the Decision Network

How do the random variables and the decision variables relate to one another?

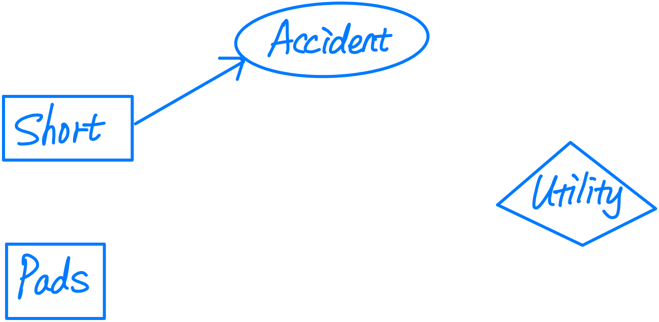
- Short affects Accident.
  On the long route, an accident won't occur.
  On the short route, an accident may occur.

- Pads does not affect Accident.

Robot decision network

$$\boxed{\begin{array}{l} P(A \mid \neg S) = 0 \\ P(A \mid S) = q \end{array}}$$

$0 \leq q \leq 1$



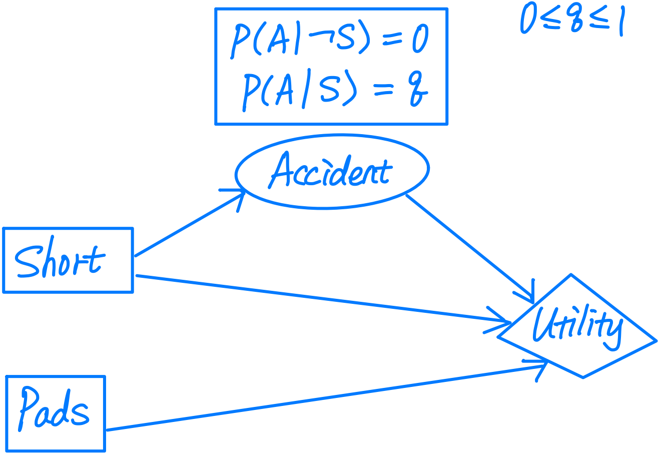
CQ: The robot's happiness

CQ: Which variables directly influence the robot's happiness?

(A) P only
(B) S only
(C) A only
(D) Two of (A), (B), and (C)
(E) All of (A), (B) and (C)

The robot must choose its route to pick up the mail. There is a short route and a long route. On the short route, the robot might slip and fall. The robot can put on pads. Pads won't change the probability of an accident. However, if an accident happens, pads will reduce the damage. Unfortunately, the pads add weight and slow the robot down. The robot would like to pick up the mail as quickly as possible while minimizing the damage caused by an accident.

Robot decision network



CQ The robot's utility function

CQ: When an accident does NOT happen, which of the following is true?

(A) The robot prefers not wearing pads than wearing pads.

(B) The robot prefers the long route over the short route.

(C) Both (A) and (B) are true.

(D) Both (A) and (B) ~~is~~ are false.

The robot's utility function

	State	$U(w_i)$
$\neg P, \neg S, \neg A$	w_0 slow, no weight	6
$\neg P, \neg S, A$	w_1 impossible	
$\neg P, S, \neg A$	w_2 quick, no weight	10
$\neg P, S, A$	w_3 severe damage	0
$P, \neg S, \neg A$	w_4 slow, extra weight	4
$P, \neg S, A$	w_5 impossible	
$P, S, \neg A$	w_6 quick, extra weight	8
P, S, A	w_7 moderate damage	2

The robot's utility function

How does the robot's utility/happiness depend on the random variables and the decision variables?

- When an accident does not happen, does the robot prefer not wearing pads or wearing pads? The robot prefers not wearing pads because it can move faster.
  $U(\neg P \land \neg S \land \neg A) > U(P \land \neg S \land \neg A)$, $U(\neg P \land S \land \neg A) > U(P \land S \land \neg A)$
- When an accident does not happen, does the robot prefer the short route or the long route?
  The robot prefers the short route because it can reach the mail faster.
  $U(P \land S \land \neg A) > U(P \land \neg S \land \neg A)$
  $U(\neg P \land S \land \neg A) > U(\neg P \land \neg S \land \neg A)$

The robot's utility function

How does the robot's utility/happiness depend on the random variables and the decision variables?

- When an accident occurs, does the robot prefer the short route or the long route? The robot must have taken the short route. There is no utility for $\neg P \wedge \neg S \wedge A$ and $P \wedge \neg S \wedge A$.
- When an accident occurs, does the robot prefer not wearing pads or wearing pads?

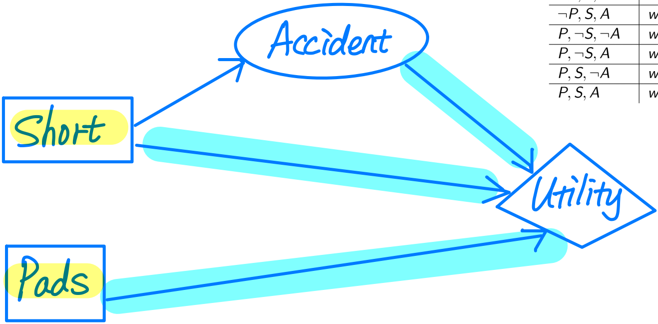
  The robot prefers wearing pads because pads reduce the severity of the damage.

  $$U(P \wedge S \wedge A) > U(\neg P \wedge S \wedge A)$$

Robot decision network

$0 \leq q \leq 1$

$$\boxed{\begin{aligned} P(A \mid \neg S) &= 0 \\ P(A \mid S) &= q \end{aligned}}$$



	State	$U(w_i)$
$\neg P, \neg S, \neg A$	w_0 slow, no weight	6
$\neg P, \neg S, A$	w_1 impossible	
$\neg P, S, \neg A$	w_2 quick, no weight	10
$\neg P, S, A$	w_3 severe damage	0
$P, \neg S, \neg A$	w_4 slow, extra weight	4
$P, \neg S, A$	w_5 impossible	
$P, S, \neg A$	w_6 quick, extra weight	8
P, S, A	w_7 moderate damage	2

Evaluating a decision network

How do we choose an action?

1. Set evidence variables for current state
2. For each possible value of decision node
   (a) set decision node to that value
   (b) calculate posterior probability for parent nodes of the utility node
   (c) calculate expected utility for the action
3. Return action with highest expected utility

Calculating the expected utilities (1/4)

What is the agent's expected utility of
not wearing pads and choosing the long route?

$$EU(\neg P \wedge \neg S) = P(W_0 | \neg P \wedge \neg S) * U(W_0)$$
$$+ P(W_1 | \neg P \wedge \neg S) * U(W_1)$$
$$= P(\neg P \wedge \neg S \wedge \neg A | \neg P \wedge \neg S) * U(W_0)$$
$$+ P(\neg P \wedge \neg S \wedge A | \neg P \wedge \neg S) * U(W_1)$$
$$= P(\neg A | \neg P \wedge \neg S) * U(W_0) + P(A | \neg P \wedge \neg S) * U(W_1)$$
$$= P(\neg A | \neg S) * U(W_0) + P(A | \neg S) * U(W_1)$$
$$= (1)(6) + (0)(-) = 6$$

Calculating the expected utilities (2/4)

What is the agent's expected utility of
not wearing pads and choosing the short route?

Sol:

$$\begin{aligned}
EU(\neg P, S) =& P(w_2|\neg P \wedge S) * U(w_2) \\
&+ P(w_3|\neg P \wedge S) * U(w_3) \\
=& P(\neg P \wedge S \wedge \neg A|\neg P \wedge S) * U(w_2) \\
&+ P(\neg P \wedge S \wedge A|\neg P \wedge S) * U(w_3) \\
=& P(\neg A|\neg P \wedge S) * U(w_2) \\
&+ P(A|\neg P \wedge S) * U(w_3) \\
=& P(\neg A|S) * U(w_2) \\
&+ P(A|S) * U(w_3) \\
=& (1 - q)(10) + (q)(0) \\
=& 10 - 10q
\end{aligned}$$

Calculating the expected utilities (3/4)

What is the agent's expected utility of wearing pads and choosing the long route?

Sol:

$$\begin{aligned}
EU(P, \neg S) =& P(w_4 | P \wedge \neg S) * U(w_4) \\
&+ P(w_5 | P \wedge \neg S) * U(w_5) \\
=& P(P \wedge \neg S \wedge \neg A | P \wedge \neg S) * U(w_4) \\
&+ P(P \wedge \neg S \wedge A | P \wedge \neg S) * U(w_5) \\
=& P(\neg A | P \wedge \neg S) * U(w_4) \\
&+ P(A | P \wedge \neg S) * U(w_5) \\
=& P(\neg A | \neg S) * U(w_4) \\
&+ P(A | \neg S) * U(w_5) \\
=& (1)(4) + (0)(-) \\
=& 4
\end{aligned}$$

Calculating the expected utilities (4/4)

What is the agent's expected utility of wearing pads and choosing the short route?

Sol:

$$
\begin{aligned}
EU(P, S) = & P(w_6 | P \wedge S) * U(w_6) \\
& + P(w_7 | P \wedge S) * U(w_7) \\
= & P(P \wedge S \wedge \neg A | P \wedge S) * U(w_6) \\
& + P(P \wedge S \wedge A | P \wedge S) * U(w_7) \\
= & P(\neg A | P \wedge S) * U(w_6) \\
& + P(A | P \wedge S) * U(w_7) \\
= & P(\neg A | S) * U(w_6) \\
& + P(A | S) * U(w_7) \\
= & (1 - q)(8) + (q)(2) \\
= & 8 - 6q
\end{aligned}
$$

What should the robot do?

- Should it wear pads or not?
- Should it choose the short or the long route?

$10 - 10q = 6 \Rightarrow q = \frac{2}{5}$

$EU(\neg P \land \neg S) = 6$

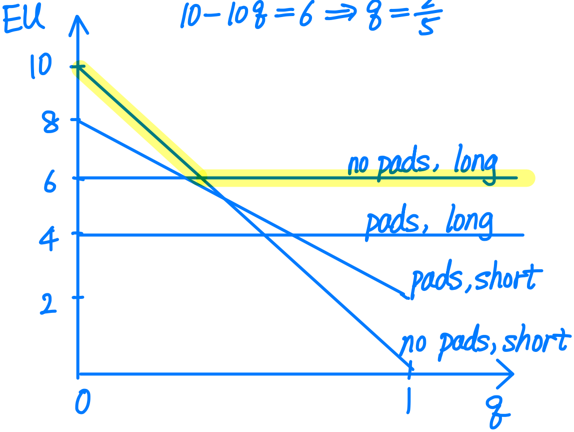
$EU(\neg P \land S) = 10 - 10q$

$\cancel{EU(P \land \neg S) = 4}$

$EU(P \land S) = 8 - 6q$

Optimal policy:

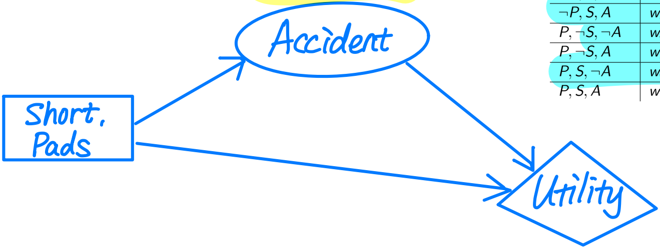
- if $q \le \frac{2}{5}$, no pads, short
- if $q > \frac{2}{5}$, no pads, long



Robot decision network

$0 \le q \le 1$

$$\boxed{\begin{aligned} P(A \mid \neg S) &= 0 \\ P(A \mid S) &= q \end{aligned}}$$



	State	$U(w_i)$
$\neg P, \neg S, \neg A$	w_0 slow, no weight	6
$\neg P, \neg S, A$	w_1 impossible	
$\neg P, S, \neg A$	w_2 quick, no weight	10
$\neg P, S, A$	w_3 severe damage	0
$P, \neg S, \neg A$	w_4 slow, extra weight	4
$P, \neg S, A$	w_5 impossible	
$P, S, \neg A$	w_6 quick, extra weight	8
P, S, A	w_7 moderate damage	2

Variable Elimination for a Single-Stage Decision Network

- Prune all the nodes that are not ancestors of the utility node.
- Sum out all chance nodes.
- For the single remaining factor, return the maximum value and the assignment that gives the maximum value.

Define the Factors

Accident

$f_1(A, S \wedge P)$:

A	$S \wedge P$	val
t	$S \wedge P$	q
f	$S \wedge P$	$1-q$
t	$S \wedge \neg P$	q
f	$S \wedge \neg P$	$1-q$
t	$\neg S \wedge P$	0
f	$\neg S \wedge P$	1
t	$\neg S \wedge \neg P$	0
f	$\neg S \wedge \neg P$	1

Utility function

$u(A, S \wedge P)$:

A	$S \wedge P$	val
t	$S \wedge P$	2
f	$S \wedge P$	8
t	$S \wedge \neg P$	0
f	$S \wedge \neg P$	10
t	$\neg S \wedge P$	–
f	$\neg S \wedge P$	4
t	$\neg S \wedge \neg P$	–
f	$\neg S \wedge \neg P$	6

Sum out all chance nodes

Multiply the two factors.

$f_2(A, S \wedge P)$:

A	$S \wedge P$	val
t	$S \wedge P$	$2q$
f	$S \wedge P$	$8 - 8q$
t	$S \wedge \neg P$	0
f	$S \wedge \neg P$	$10 - 10q$
t	$\neg S \wedge P$	0
f	$\neg S \wedge P$	4
t	$\neg S \wedge \neg P$	0
f	$\neg S \wedge \neg P$	6

Sum out A from f_2.

$f_3(S \wedge P)$:

$S \wedge P$	val
$S \wedge P$	$8 - 6q$
$S \wedge \neg P$	$10 - 10q$
$\cancel{\neg S \wedge P}$	$\cancel{4}$
$\neg S \wedge \neg P$	6

Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Model a one-off decision problem by constructing a decision network containing nodes, arcs, conditional probability distributions, and a utility function.
- Determine the optimal policy of a decision network by computing the expected utility of every policy.
- Determine the optimal policy of a decision network by applying the variable elimination algorithm.