Goals and Preferences

Alice . . . went on "Would you please tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to," said the Cat.
"I don't much care where —" said Alice.
"Then it doesn't matter which way you go," said the Cat.

Lewis Carroll, 1832–1898
Alice's Adventures in Wonderland, 1865
Chapter 6

Overview

At the end of the class you should be able to:

- model a user's preferences and utility when there is uncertainty
- build a simple model that includes actions, uncertainty and utilities.

Single decisions

- Single decisions: agent makes all decisions before acting
- The agent can choose a value for each decision variable
- Lets combine all decision variables into a single variable D
- The expected utility of decision $D = d_i$ is

$$\mathcal{E}(u \mid D = d_i) = \sum_{\omega \in \Omega} P(\omega \mid D = d_i) \times u(\omega)$$

  where $u(\cdot)$ is the utility function
  Ω is the set of all worlds

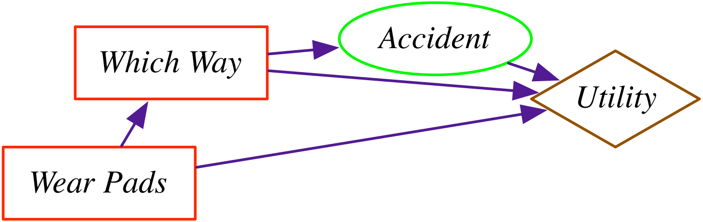
- An optimal single decision is a decision $D = d_{max}$ whose expected utility is maximal:

$$\mathcal{E}(u \mid D = d_{max}) = \max_{d_i \in domain(D)} \mathcal{E}(u \mid D = d_i).$$

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, whose parents are the variables on which the utility depends. Drawn as a diamond.



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

Single-stage decision networks

A single-stage decision network consists of:

- DAG with three sorts of nodes: decision, random, utility.
  Random nodes are the same as the nodes in a belief network.
- A domain for each decision variable and each random variable.
- A unique utility node.
  The utility node has no children and no domain.

A single-stage decision network has the factors:

- A utility function is a factor on the parents of the utility node
- A conditional probability for each random variable given its parents
- (No tables associated with the decision nodes.)

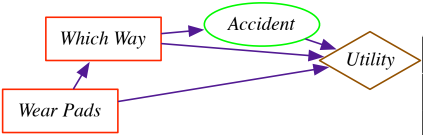
Finding an optimal decision

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

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

  To find an 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 value in the domain of D
  - Choose the D with the maximum value in the factor.

Example Initial Factors



Which Way	Accident	Value
long	true	0.01
long	false	0.99
short	true	0.2
short	false	0.8

Which Way	Accident	Wear Pads	Value
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

After summing out Accident

Which Way	Wear Pads	Value
long	true	74.55
long	false	79.2
short	true	83.0
short	false	80.6

Decision Networks

- flat or modular or hierarchical
- explicit states or features or individuals and relations
- static or finite stage or indefinite stage or infinite stage
- fully observable or partially observable
- deterministic or stochastic dynamics
- goals or complex preferences
- single agent or multiple agents
- knowledge is given or knowledge is learned
- perfect rationality or bounded rationality

Sequential Decisions

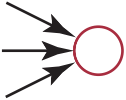
- An intelligent agent doesn't carry out a multi-step plan ignoring information it receives between actions.
- A more typical scenario is where the agent:
  observes, acts, observes, acts, ...
- Subsequent actions can depend on what is observed.
  What is observed depends on previous actions.
- Often the sole reason for carrying out an action is to provide information for future actions.
  For example: diagnostic tests, spying.

Sequential decision problems

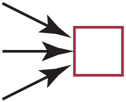
- A sequential decision problem consists of a sequence of decision variables $D_1, \ldots, D_n$.
- Each D_i has an information set of variables $parents(D_i)$, whose value will be known at the time decision D_i is made.

Decisions Networks

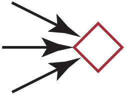
A decision network is a graphical representation of a finite sequential decision problem, with 3 types of nodes:



- A random variable is drawn as an ellipse. Arcs into the node represent probabilistic dependence. Each random variable has a domain and an associated factor.

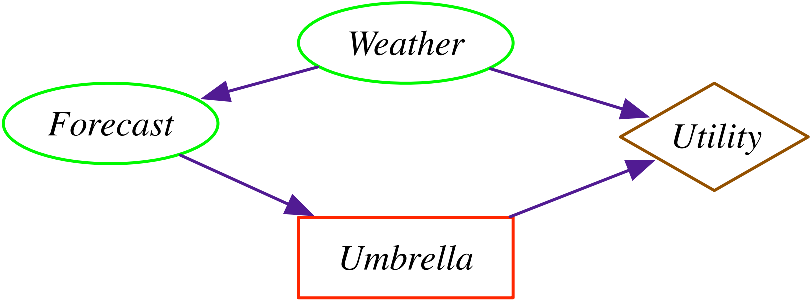


- A decision variable is drawn as an rectangle. Arcs into the node represent information available when the decision is make. Each decision variable has a domain, but no associated factor.



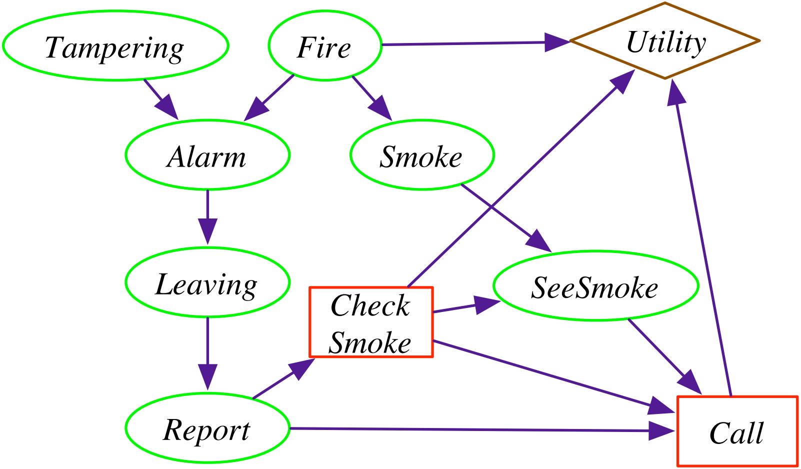
- A utility node is drawn as a diamond. Arcs into the node represent variables that the utility depends on. The utility node has no domain, and a factor on the parents of the node.

Umbrella Decision Network



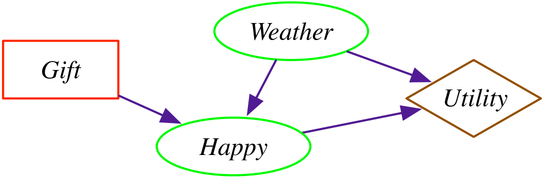
- The agent has to decide whether to take its umbrella.
- It observes the forecast.
- It doesn't observe the weather directly.
- The forecast is a noisy sensor of the weather.
- The utility depends on the weather and whether the agent takes the umbrella.

Decision Network for the Alarm Problem



Clicker Question

The decision network



requires which probabilities to be specified:

A $P(Utility \mid Weather, Happy)$, $P(Weather)$, $P(Happy \mid Weather, Gift)$, $P(Gift)$

B $P(Weather)$, $P(Happy \mid Weather, Gift)$, $P(Gift)$

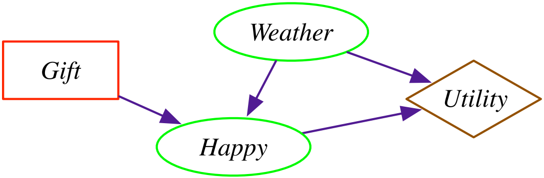
C $P(Utility \mid Weather, Happy)$, $P(Weather)$, $P(Happy \mid Weather, Gift)$,

D $P(Weather)$, $P(Happy \mid Weather, Gift)$

E $P(Weather)$, $P(Happy \mid Weather)$

Clicker Question

The decision network



requires how many factors be specified initially:

A 2

B 3

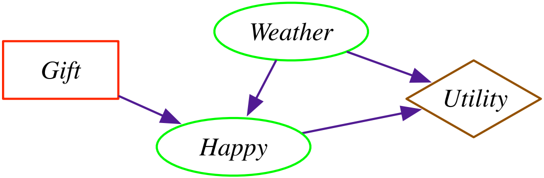
C 4

D 5

E 6

Clicker Question

The decision network



The initial factor that isn't a (conditional) probability is a factor on which variables?

A *Gift*, *Weather*, *Happy*, *Utility*

B *Gift*, *Weather*, *Happy*

C *Weather*, *Happy*, *Utility*

D *Weather*, *Happy*

E *Gift*

No-forgetting

A No-forgetting decision network is a decision network where:

- The decision nodes are totally ordered. This is the order the actions will be taken.
- All decision nodes that come before D_i are parents of decision node D_i. Thus the agent remembers its previous actions.
- Any parent of a decision node is a parent of subsequent decision nodes. Thus the agent remembers its previous observations.

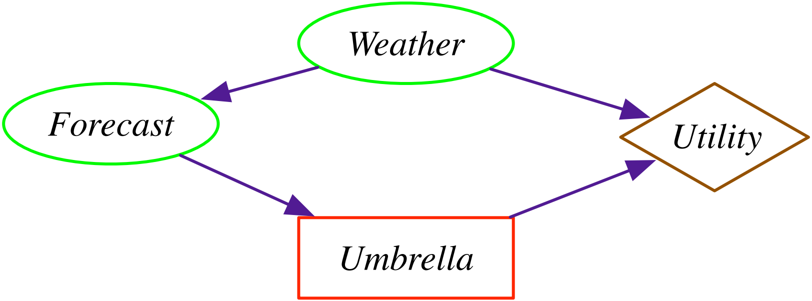
What should an agent do?

- 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 decision function for decision node D_i is a function π_i that specifies what the agent does for each assignment of values to the parents of D_i.
  When it observes O, it does $\pi_i(O)$.
- A policy is a sequence of decision functions; one for each decision node.

Umbrella Decision Network

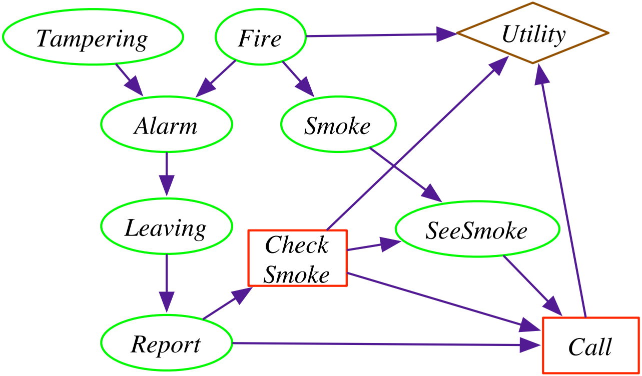


$domain(Forecast) = \{sunny, cloudy, rainy\}$
$domain(Umbbrella) = \{take, leave\}$
Some policies:

- take if cloudy else leave
- always take
- always leave

There are $2^3 = 8$ policies

Decision Network for the Alarm Problem



All variables are Boolean. Some policies:

- Never check. Call iff report.
- Check iff report. Call iff report and see smoke.
- Always check. Always call.

There are $2^2 * 2^8 = 1024$ policies.

Expected Utility of a Policy

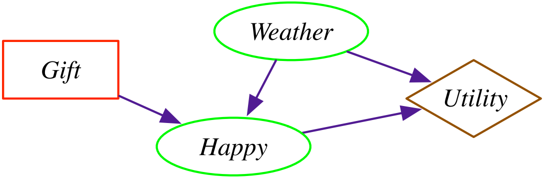
- Possible world ω satisfies policy π if ω assigns the value to each decision node that the policy specifies.
- The expected utility of policy π is

$$\mathcal{E}(u \mid \pi) = \sum_{\omega \text{ satisfies } \pi} u(\omega) \times P(\omega)$$

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

Clicker Question

Consider the decision network



where each variable is Boolean (domain is $\{True, False\}$). How many policies are there?

A 1

B 2

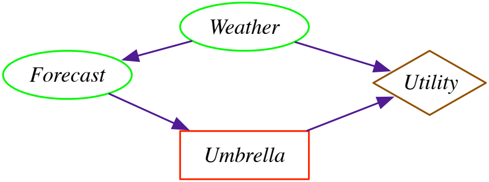
C 2^3

D 2^4

E There is not enough information to tell.

Clicker Question

Consider the decision network



where $domain(Weather) = \{Sunshine, Rain\}$,
$domain(Forecast) = \{Sunny, Cloudy, Rainy\}$,
$domain(Umbrella) = \{Take, Leave\}$.
How many policies are there?

A 2

B 2^3

C 3^2

D 2^4

E There is not enough information to tell.