CSCE A405/A605 (Adv) Artificial Intelligence

Making Simple Decisions

Ref: Artificial Intelligence: A Modern Approach, 4th ed by Stuart Russell and Peter Norvig, chapter 16

Instructor: Masoumeh Heidari (mheidari2@Alaska.edu)

UAA College of Engineering
UNIVERSITY of ALASKA ANCHORAGE

Decision making under uncertainty

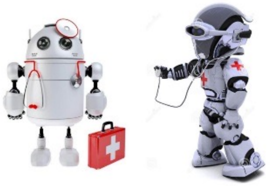
How an agent should make decisions so that it gets what it wants—on average, at least.

Reminder: A goal-based agent has a binary distinction between good (goal) and bad (non-goal) states, while a decision-theoretic agent assigns a continuous range of values to states, and thus can more easily choose a better state even when no best state is available.

Decision making under uncertainty

Accelerating?
Braking?
Talking to passengers?

Increasing amount of production?
Decreasing price?



Prescribing additional tests?
Recommending treatments?
Do nothing?

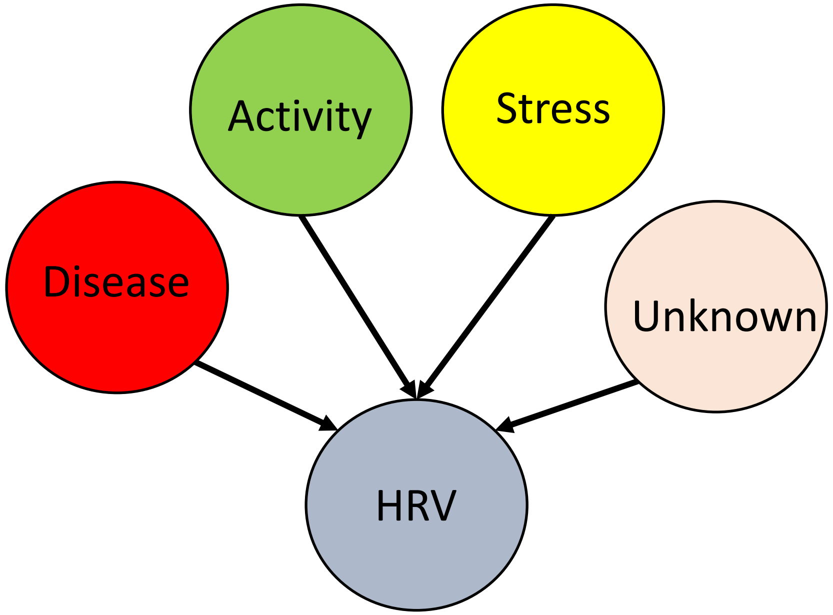
Giving feedbacks?
Asking questions?
Covering a new topic?

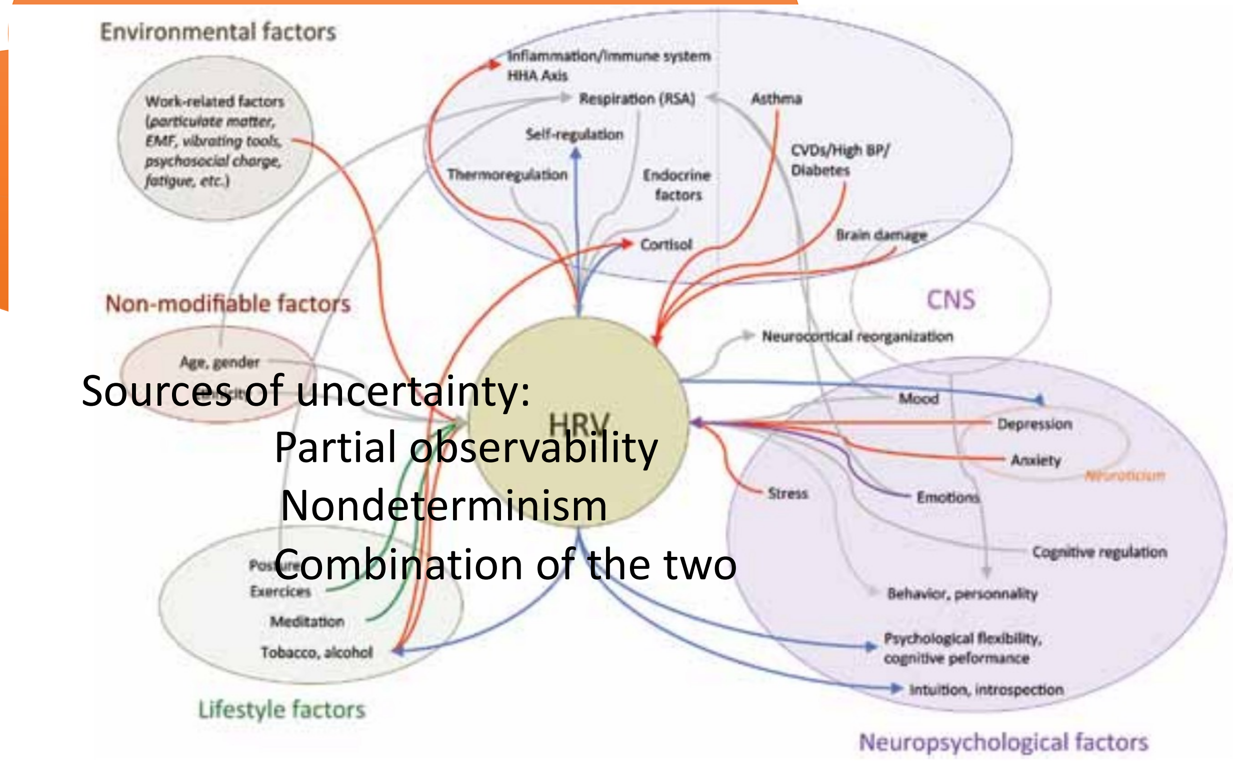
Uncertainty

Probability provid... wa... summarizing the uncertainty
... ignorance, thereby solving the

I will reach to the airport:

Where is the 1,000,000$?





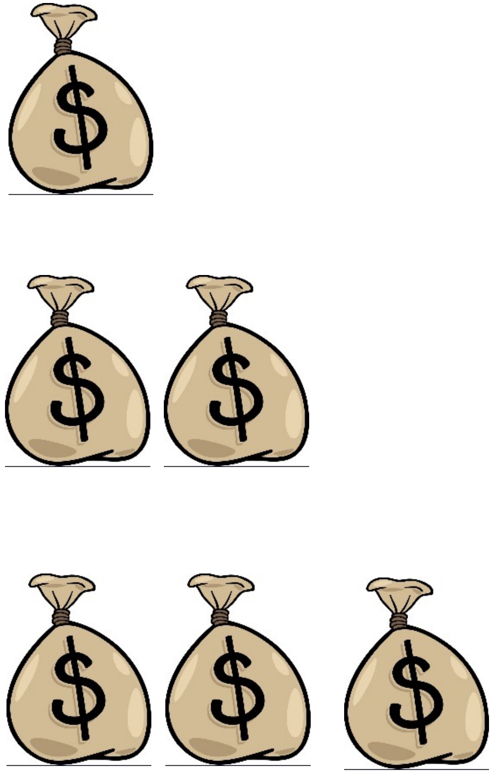
Sources of uncertainty:

Partial observability

Nondeterminism

Combination of the two

Preferences



To make such choices, an agent must first have **preferences** between the different possible **outcomes** of the various plans.

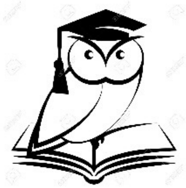
An outcome is a completely specified state, including factors such as whether the agent arrives on time and the length of the wait at the airport.

Uncertainty and Rational Decisions

- We use utility theory to represent and reason with preferences.
- Preferences, as expressed by utilities, are combined with probabilities in the general theory of rational decisions called decision theory:

Decision theory = probability theory + utility theory

Utility



Utility is just a number which says how good something is!
(in average)

- Fundamental idea of decision theory is that an agent is rational if and only if it chooses the action that yields the highest expected utility, averaged over all the possible outcomes of the action.

- This is called the principle of maximum expected utility (MEU).

- “Expected” means the “average,” or “statistical mean” of the outcomes, weighted by the probability of the outcome.

Maximum Expected Utility-Example 1

5 $

500,000 $

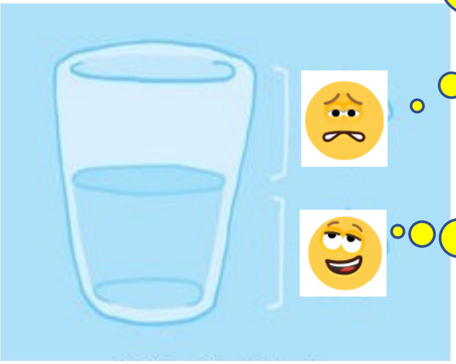
Current state: 0 $

Possible actions: {Deal , No Deal}

Possible outcomes: {5$, 500,000$, 300,000$}

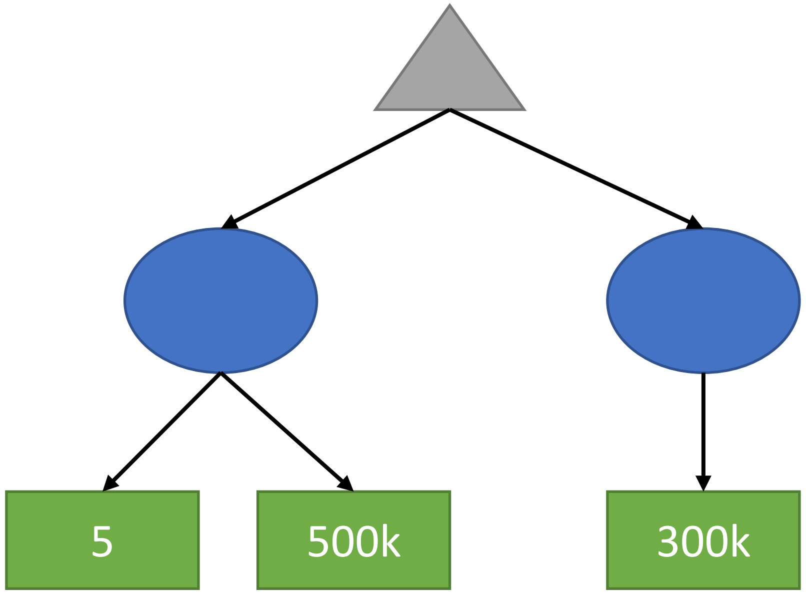
Banker's offer: 300,000$

Maximum Expected Utility-Example 1

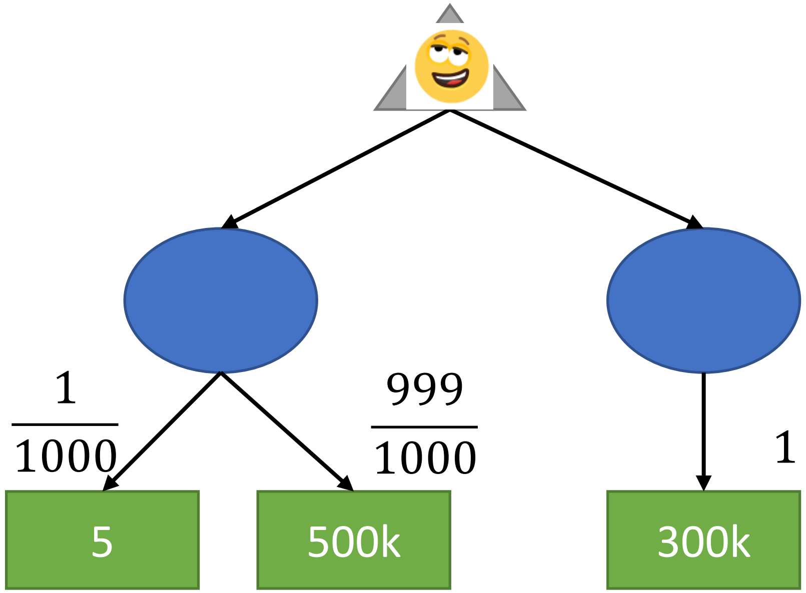


5$ is in my case.
a=Deal!

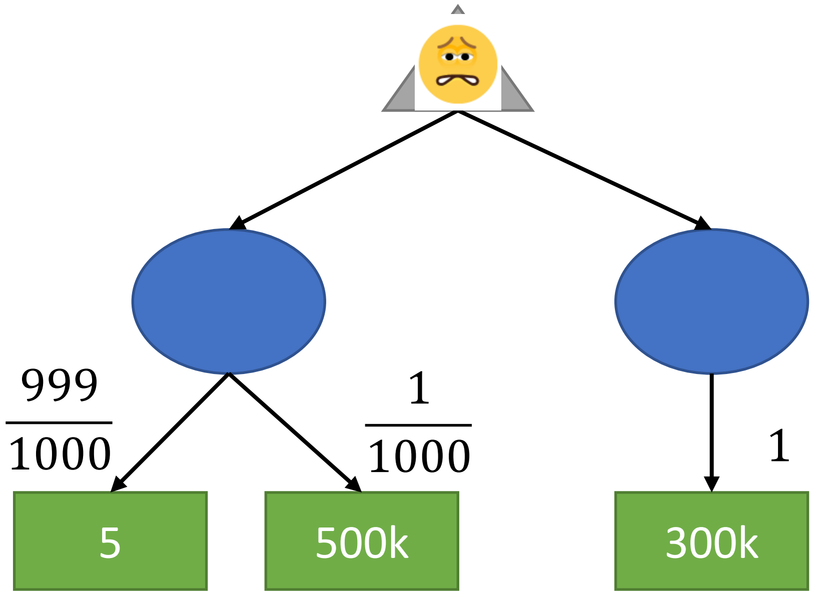
500k$ is in my case.
a = No Deal!



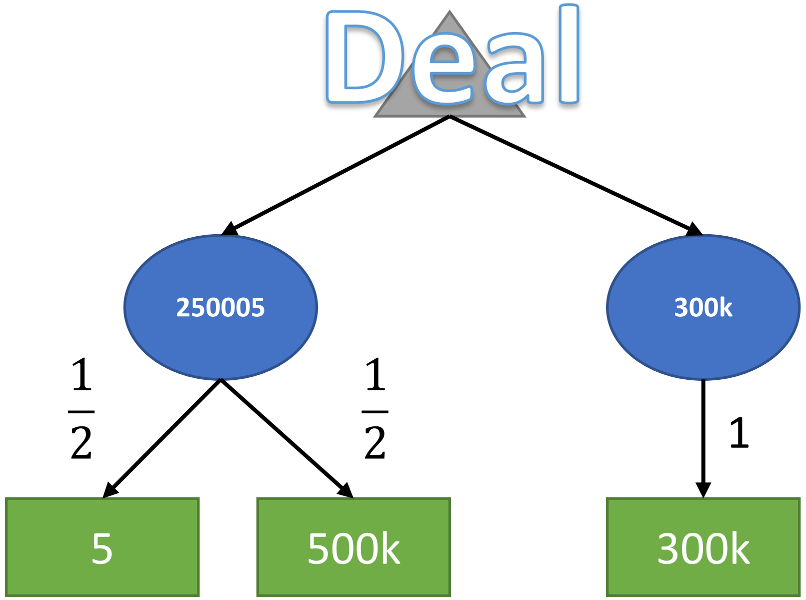
Maximum Expected Utility-Example 1



Maximum Expected Utility-Example 1



Maximum Expected Utility-Example 1



5 $

500,000 $

Banker's offer: 300,000$

Combining Beliefs and Desires under Uncertainty

- The agent's preferences are captured by a utility function, U(s), which assigns a single number to express the desirability of a state.

- There may be uncertainty about the current state, so we'll assume that the agent assigns a probability *p(s)* to each possible current state *s*.

- There may also be uncertainty about the action outcomes; the transition model is given by *p(s'|s,a)*, the probability that action *a* in state *s* reaches state *s'*.

Combining Beliefs and Desires under Uncertainty

- We are interested in the outcome s', in other words, *p(Result(a)=s')*.

The probability of reaching *s'* by doing *a* in the current state, whatever that is.

$$p(Result(a) = s') = \sum_{s} p(s)p(s'|s, a)$$

- Decision theory, in its simplest form, deals with choosing among actions based on the desirability of their immediate outcomes; that is, the environment is assumed to be episodic.

Combining Beliefs and Desires under Uncertainty

- The expected utility of an action given the evidence, $EU(a)$, is just the average utility value of the outcomes, weighted by the probability that the outcome occurs:

$$EU(a) = \sum_{s'} p(Result(a) = s')U(s')$$

- The principle of maximum expected utility (MEU) says that a rational agent should choose the action that <u>maximizes the agent's expected utility</u>:

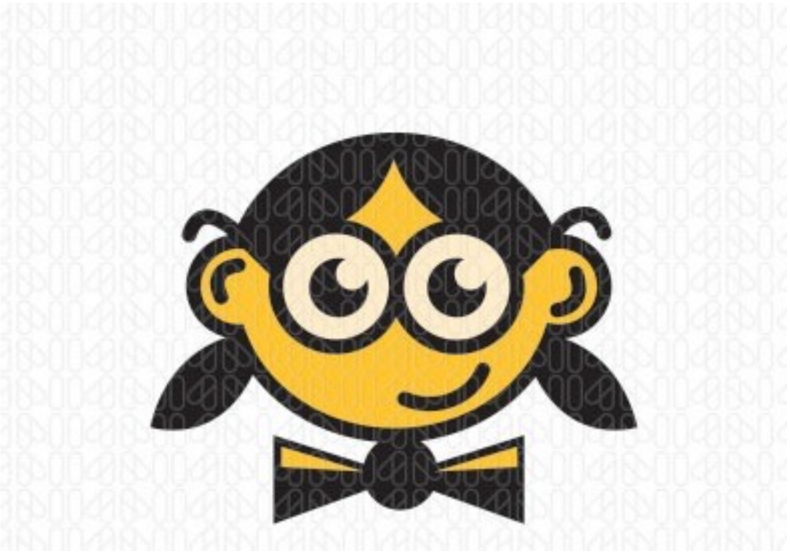
$$action = arg \max_{a} EU(a)$$

- In a sense, the MEU principle could be seen as a prescription for intelligent behavior.

Maximum Expected Utility-Example 2

$$EU(a|\mathbf{e}) = \sum_{s'} P(\text{RESULT}(a) = s' \mid a, \mathbf{e})\, U(s')$$

$$action = \underset{a}{\operatorname{argmax}}\, EU(a|\mathbf{e})$$



Emma is a smart student at UAA. She knows that AI is the most important thing to learn! But she does not know anything about AI.

Maximum Expected Utility-Example 2

s = Emma doesn't know anything about AI.

a_1= Take the Artificial Intelligence course

a_2= Don't take the Artificial Intelligence course

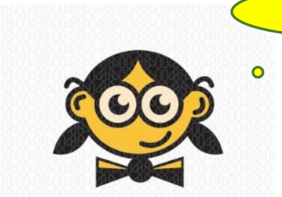
s_1‘ = Emma knows nothing about AI.

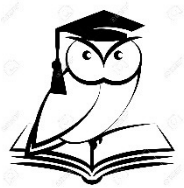
s_2‘ = Emma knows a few things about AI but she is not interested to know more.

s_3‘ = Emma has become an AI researcher, she has build her first intelligent agent and AI is her favorite topic.

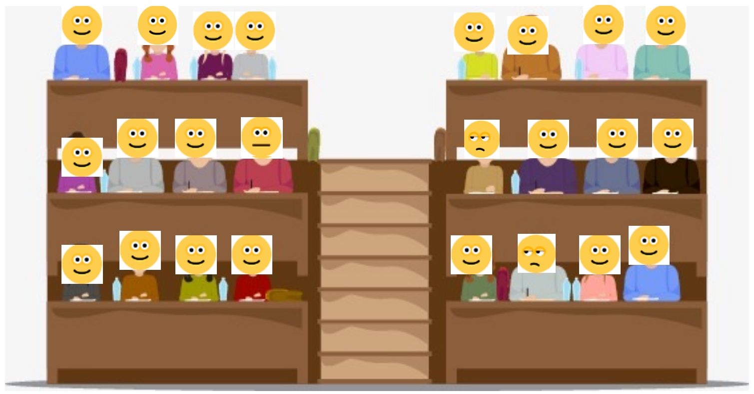
Maximum Expected Utility-Example 2

$P(s_1'|a_1) = P(s_2'|a_1) = P(s_3'|a_1) = \frac{1}{3}$

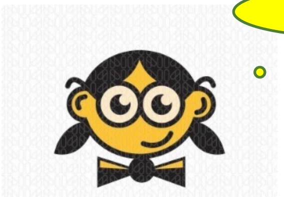




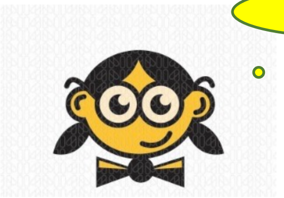
This course was offered last year.



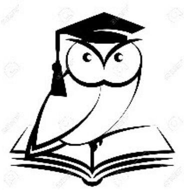
Maximum Expected Utility-Example 2



$P(s_1'|a_1) = P(s_2'|a_1) = P(s_3'|a_1) = \frac{1}{3}$



$P(s_1'|a_1,e) = \frac{2}{24}$, $P(s_2'|a_1,e) = \frac{1}{24}$, $P(s_3'|a_1,e) = \frac{21}{24}$



$$EU(a|\mathbf{e}) = \sum_{s'} P(\textsc{Result}(a) = s' \mid a, \mathbf{e})\, U(s')$$ and

$$action = \operatorname*{argmax}_a EU(a|\mathbf{e})$$

Maximum Expected Utility-Example 2

$P(s_1'|a_2,e) = 1$, $P(s_2'|a_2,e) = 0$, $P(s_3'|a_3,e) = 0$

$P(s_1'|a_1,e) = \frac{2}{24}$, $P(s_2'|a_1,e) = \frac{1}{24}$, $P(s_3'|a_1,e) = \frac{21}{24}$

$U(s_1') = 0$

$U(s_2') = 1$

$U(s_3') = 1000$

$$EU(a|\mathbf{e}) = \sum_{s'} P(\text{RESULT}(a) = s' \mid a, \mathbf{e})\, U(s')$$

$$action = \underset{a}{\operatorname{argmax}}\, EU(a|\mathbf{e})$$

Maximum Expected Utility-Example 2

$P(s_1'|a_2,e) = 1$, $P(s_2'|a_2,e) = 0$, $P(s_3'|a_3,e) = 0$

$P(s_1'|a_1,e) = \frac{2}{24}$, $P(s_2'|a_1,e) = \frac{1}{24}$, $P(s_3'|a_1,e) = \frac{21}{24}$

$U(s_1') = 0$

$U(s_2') = 1$

$U(s_3') = 1000$

$EU(a_1|e) = \frac{21001}{24} \sim 875$

$EU(a_2|e) = 0$

$$action = \underset{a}{\operatorname{argmax}}\, EU(a|\mathbf{e})$$

UAA College of Engineering
UNIVERSITY *of* ALASKA ANCHORAGE

Maximum Expected Utility-Example 2

$P(s_1' | a_2, e) = 1$, $P(s_2' | a_2, e) = 0$, $P(s_3' | a_3, e) = 0$

$P(s_1' | a_1, e) = \frac{2}{24}$, $P(s_2' | a_1, e) = \frac{1}{24}$, $P(s_3' | a_1, e) = \frac{21}{24}$

$U(s_1') = 0$

$U(s_2') = 1$

$U(s_3') = 1000$

$EU(a_1 | e) = \frac{21001}{24} \sim 875$

$EU(a_2 | e) = 0$

action = $a1$

UAA College of Engineering
UNIVERSITY of ALASKA ANCHORAGE

Maximum expected utility (MEU)

- The MEU principle formalizes the general notion that an intelligent agent should "do the right thing," but does not operationalize that advice.
- Estimating the probability distribution *p(s)* over possible states of the world, which folds into *p(Result(a)=s')*, requires perception, learning, knowledge representation, and inference.
- Computing *p(Result(a)=s')* itself requires a causal model of the world.
- There may be many actions to consider, and computing the outcome utilities *U(s')* may itself require further searching or planning because an agent may not know how good a state is until it knows where it can get to from that state.
- An AI system acting on behalf of a human may not know the human's true utility function, so there may be uncertainty about *U.*

Maximum expected utility (MEU)

- In summary, decision theory is not a panacea that solves the AI problem—but it does provide the beginnings of a basic mathematical framework that is general enough to define the AI problem.

- The MEU principle has a clear relation to the idea of performance measures introduced in Chapter 2.

- If an agent acts so as to maximize a utility function that correctly reflects the performance measure, then the agent will achieve the highest possible performance score (averaged over all the possible environments).

Utility functions

- An agent might prefer to have a prime number of dollars in its bank account; in which case, if it had \$16 it would give away \$3. This might be unusual, but we can't call it irrational.

- Preference elicitation. If we want to build a decision-theoretic system that helps the agent make decisions or acts on his or her behalf, we must first work out what the agent's utility function is.

Utility functions

- Suppose you have triumphed over the other competitors in a television game show.
- The host now offers you a choice: either you can take the $1,000,000 prize or you can gamble it on the flip of a coin.
- If the coin comes up heads, you end up with nothing, but if it comes up tails, you get $2,500,000.
- What Do you Do?

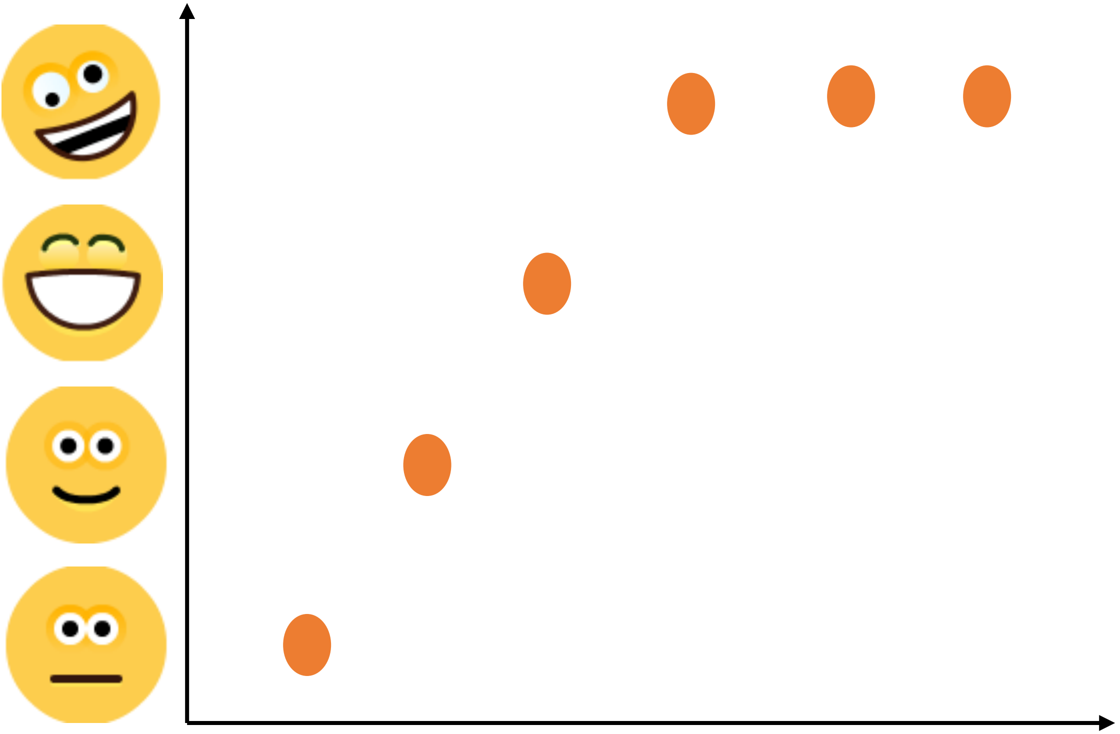
Utility functions

- Assuming the coin is fair, the expected monetary value (EMV) of the gamble is 0.5*($0)+0.5*($2,500,000) = $1,250,000, which is more than the original $1,000,000.

- Is accepting the gamble a better decision?

Utility functions

- Utility is not directly proportional to monetary value.



Utility functions

- The utility for your first million is very high (or so they say), whereas the utility for an additional million is smaller.

- Suppose you assign a utility of 5 to your current financial status (S_k), a 9 to the state $S_{k+2,500,000}$, and an 8 to the state $S_{k+1,000,000}$.

- Then the rational action would be to decline, because the expected utility of accepting is only 7 (less than the 8 for declining).

- On the other hand, a billionaire would most likely have a utility function that is locally linear over the range of a few million more, and thus would accept the gamble.

Utility function

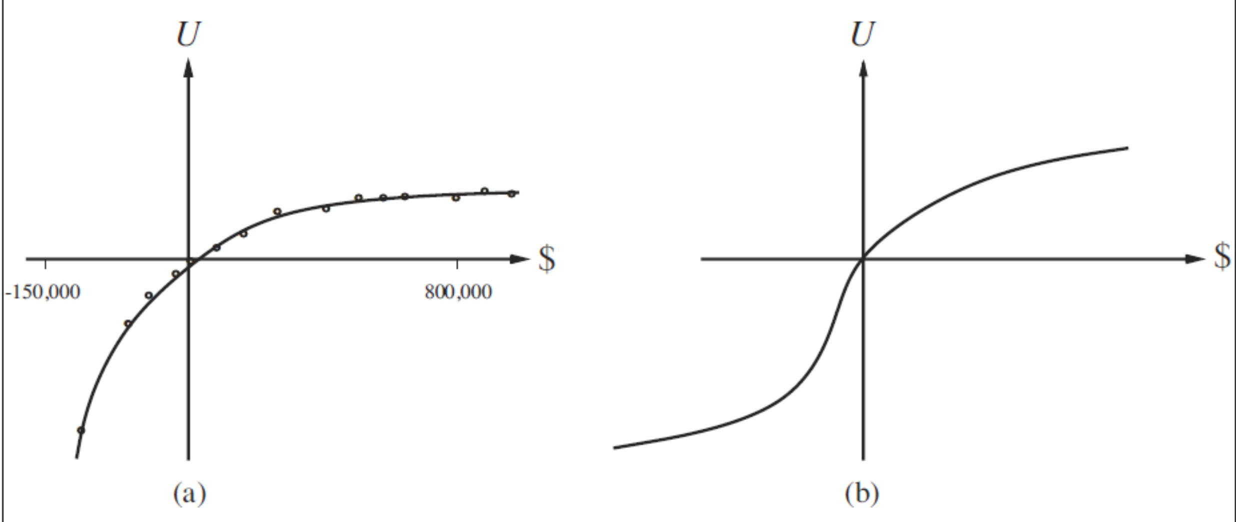
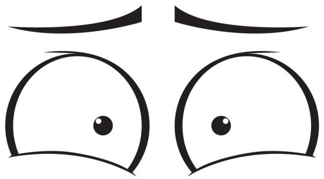


Figure 16.2 The utility of money. (a) Empirical data for Mr. Beard over a limited range. (b) A typical curve for the full range.

Insurance Premium

- The difference between the EMV of a lottery and its certainty equivalent is called the insurance premium.



People’s point of view: would rather pay a small insurance premium than gamble the price of their house against the chance of a fire.

Risk-averse

Insurance company’s point of view: the price of the house is very small compared with the firm’s total reserves. This means that the insurer’s utility curve is approximately linear over such a small region, and the gamble costs the company almost nothing.

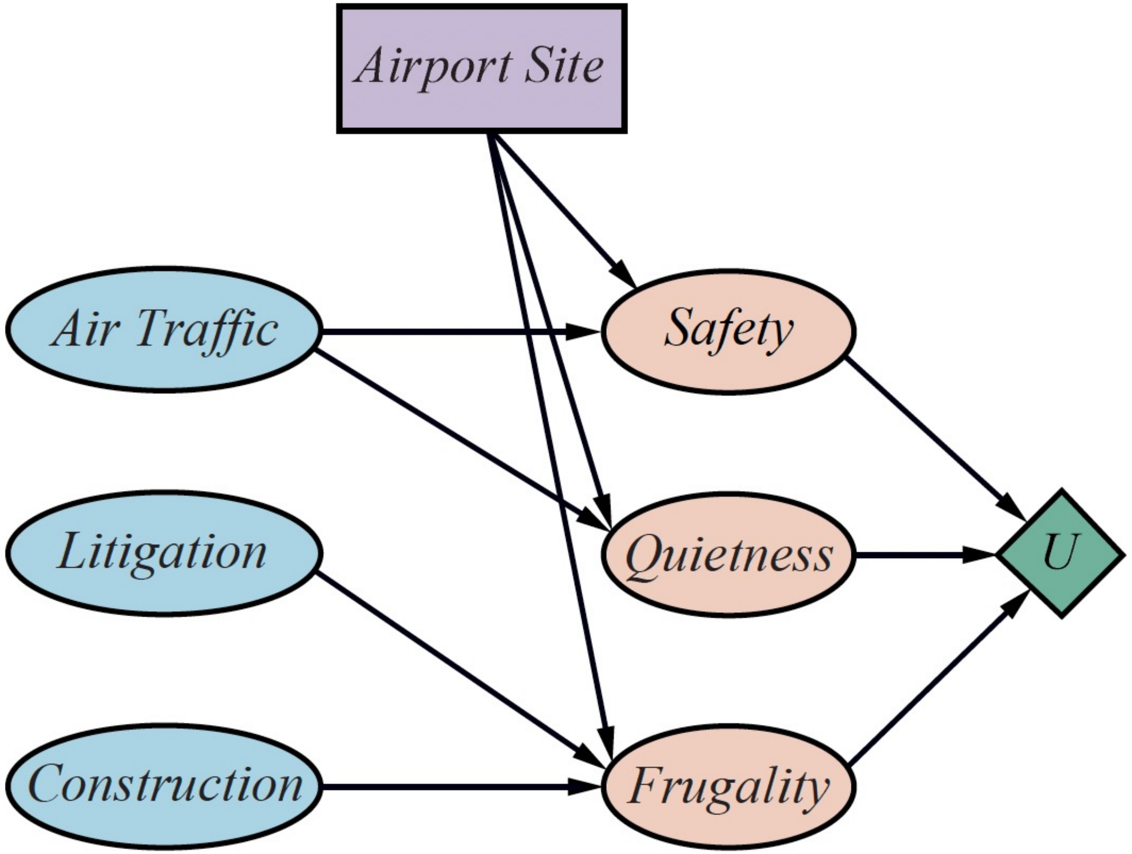
Risk-neutral

Decision networks

- In this section, we look at a general mechanism for making rational decisions. The notation is often called an influence diagram (Howard and Matheson, 1984), but we will use the more descriptive term decision network.

- Decision networks combine Bayesian networks with additional node types for actions and utilities.

Decision networks



Chance nodes (ovals) represent random variables, just as they do in Bayesian networks. Each chance node has associated with it a conditional distribution that is indexed by the state of the parent nodes. The parent nodes can include decision nodes as well as chance nodes.

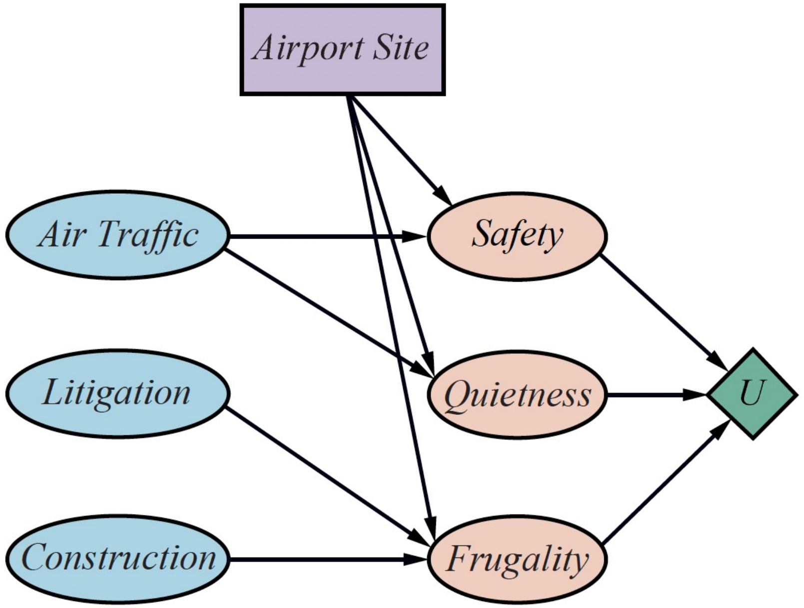
Decision nodes (rectangles) represent points where the decision maker has a choice of actions. (This chapter deals with a single decision node while next chapter deals with cases in which more than one decision must be made.)

Utility nodes (diamonds) represent the agent’s utility function. The utility node has as parents all variables describing the outcome that directly affect utility. Associated with the utility node is a description of the agent’s utility as a function of the parent attributes. The description could be just a tabulation of the function, or it might be a parameterized additive or linear function of the attribute values.

Decision networks

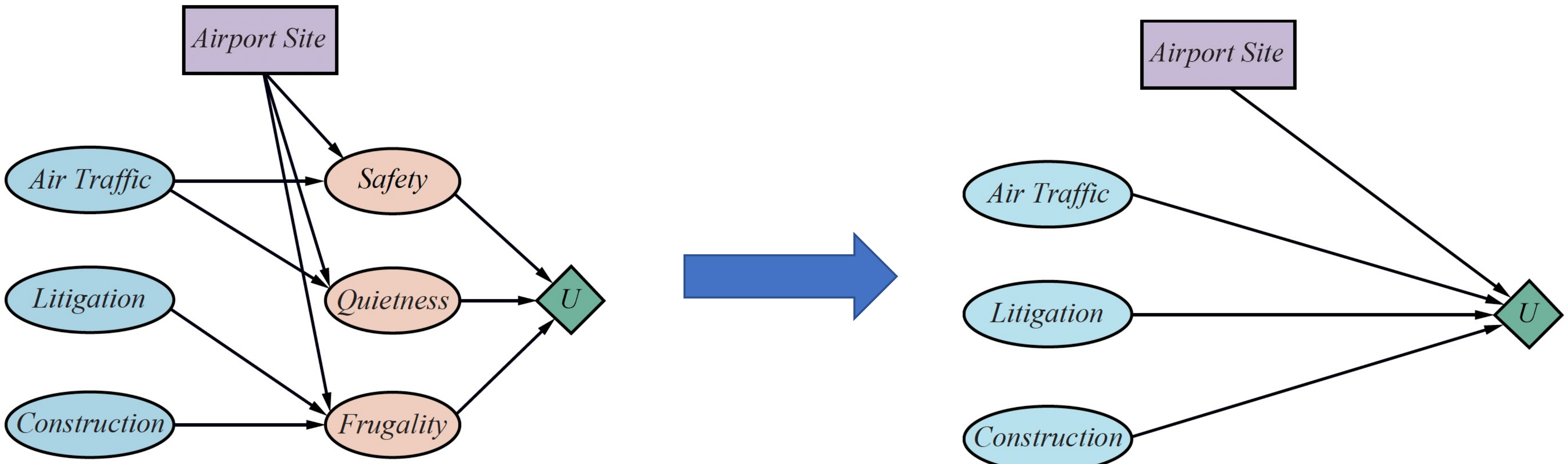
- In its most general form, a decision network represents information about the agent’s current state, its possible actions, the state that will result from the agent’s action, and the utility of that state.

- Note that each of the current-state chance nodes could be part of a large Bayesian network for assessing construction costs, air traffic levels, or litigation potentials.



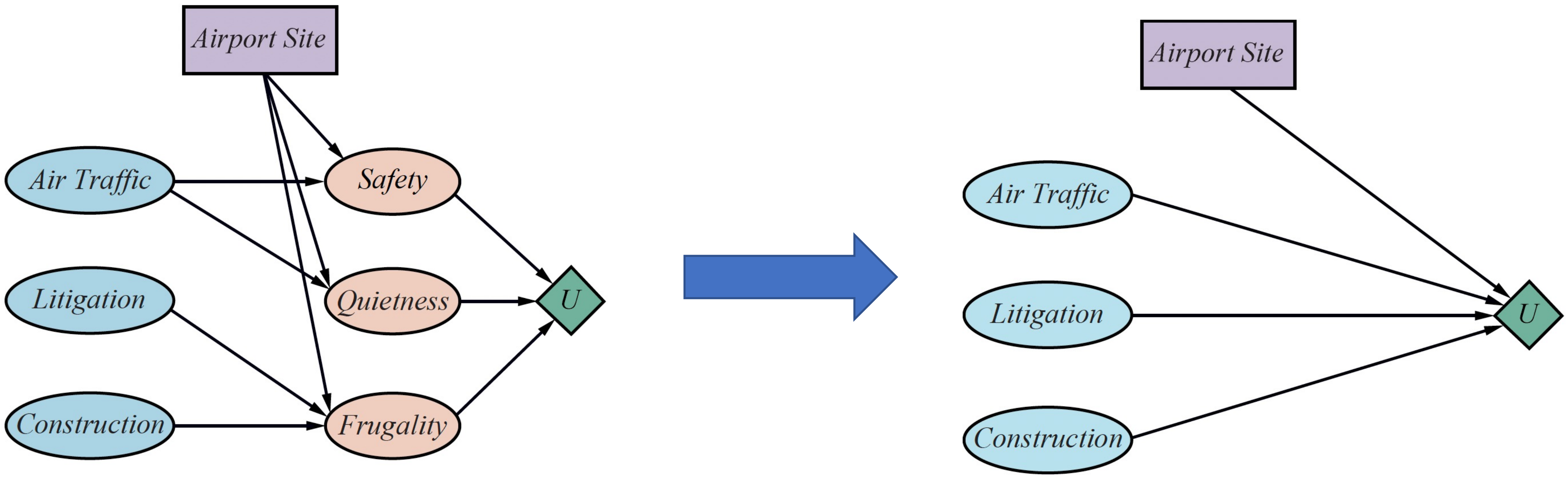
Decision networks

- Sometimes to simplify the network, chance nodes describing the outcome state are omitted and the utility node is connected directly to the current-state nodes and the decision node.



Decision networks

In this case, rather than representing a utility function on outcome states, the utility node represents the *expected* utility associated with each action; that is, the node is associated with an **action-utility function** (also known as a **Q-function** in reinforcement learning).



Decision networks

- Notice that, because the Quietness, Safety, and Frugality chance nodes refer to future states, they can <u>never</u> have their values set as evidence variables.

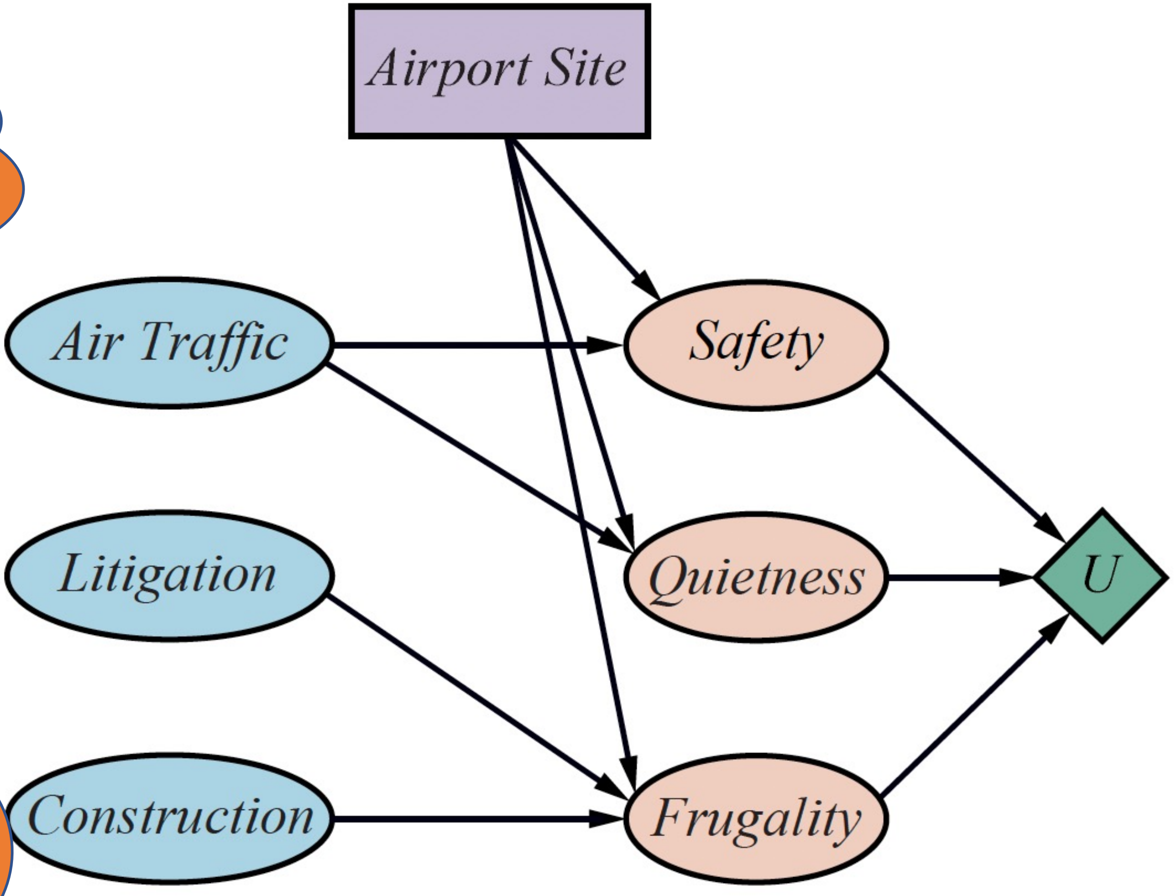
- Thus, the simplified version that omits these nodes can be used <u>whenever</u> the more general form can be used.

- Although the simplified form contains fewer nodes, the omission of an explicit description of the outcome of the siting decision means that it is less flexible with respect to changes in circumstances.

Decision networks

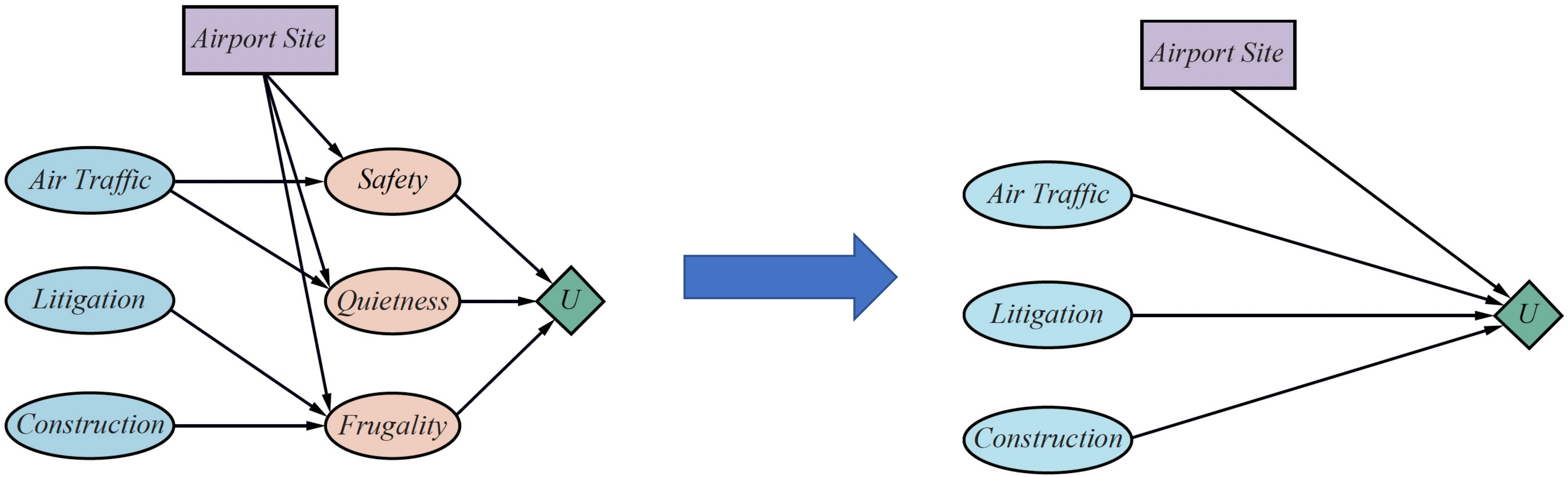
A new generation of aircraft is being used. There is a change in aircraft noise levels!

There are a lot of complaints about noise pollution.
There is a change in the weight accorded to noise pollution in the utility function!



Decision networks

- In the action-utility diagram, on the other hand, all such changes have to be reflected by changes to the action-utility table.



UAA College of Engineering
UNIVERSITY of ALASKA ANCHORAGE

Decision networks

- Actions are selected by evaluating the decision network for each possible setting of the decision node.
- The algorithm for evaluating decision networks is the following:

1. Set the evidence variables for the current state.
2. For each possible value of the decision node:
   (a) Set the decision node to that value.
   (b) Calculate the posterior probabilities for the parent nodes of the utility node, using a standard probabilistic inference algorithm.
   (c) Calculate the resulting utility for the action.
3. Return the action with the highest utility.

Decision networks

- This is a straightforward approach that can utilize any available Bayesian network algorithm and can be incorporated directly into the agent design.

- We will see in the next chapter that the possibility of executing several actions in sequence makes the problem much more interesting.

Recap

- Preferences
- Utility functions
- Maximum utility principle
- Combining belief and desires under uncertainty
- Decision networks

The value of Information

- In the preceding analysis, we have assumed that all relevant information, or at least all available information, is provided to the agent before it makes its decision.

- In practice, this is hardly ever the case. One of the most important parts of decision making is <u>knowing what questions to ask</u>.

- For example, a doctor cannot expect to be provided with the results of all possible diagnostic tests and questions at the time a patient first enters the consulting room.

The value of information

- For example, a doctor cannot expect to be provided with the results of all possible diagnostic tests and questions at the time a patient first enters the consulting room.

- Tests are often expensive and sometimes hazardous (both directly and because of associated delays).

- Their importance depends on two factors: whether the test results would lead to a significantly better treatment plan, and how likely the various test results are.

The value of information

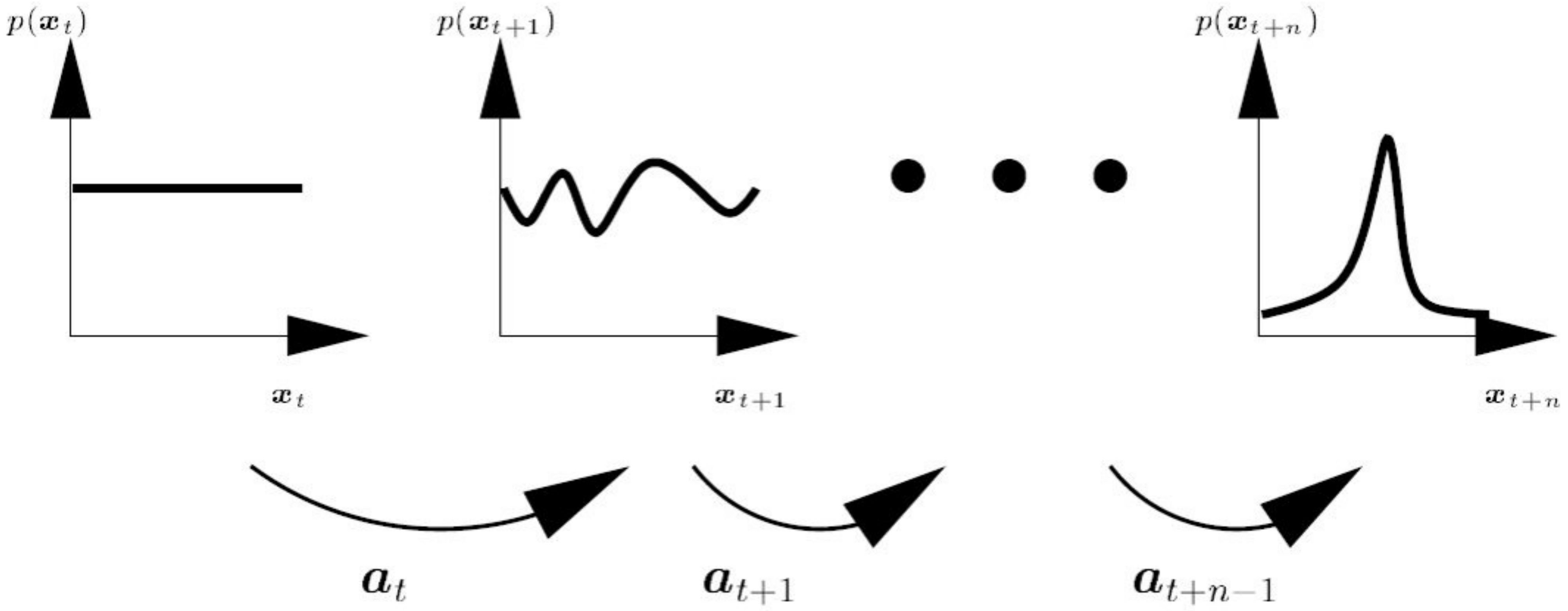
- Information value theory enables an agent to choose what information to acquire.

- We assume that, prior to selecting a "real" action represented by the decision node, the agent can acquire the value of any of the potentially observable chance variables in the model.

- Thus, information value theory involves a simplified form of sequential decision making—simplified because the observation actions affect only the agent's belief state, not the external physical state.

- The value of any particular observation must derive from the potential to affect the agent's eventual physical action; and this potential can be estimated directly from the decision model itself.

The value of information



UAA College of Engineering
UNIVERSITY of ALASKA ANCHORAGE

The value of information

- Let the agent's initial evidence be *e*. Then the value of the current best action *a* is defined by $$EU(\alpha|\mathbf{e}) = \max_a \sum_{s'} P(\textsc{Result}(a) = s' \mid a, \mathbf{e})\, U(s')$$

- The value of the new best action (after the new evidence $E_j = e_j$ is obtained) will be $$EU(\alpha_{e_j}|\mathbf{e}, e_j) = \max_a \sum_{s'} P(\textsc{Result}(a) = s' \mid a, \mathbf{e}, e_j)\, U(s')$$

- But E_j is a random variable whose value is currently unknown, so to determine the value of discovering E_j, given current information *e* we must average over all possible values e_{jk} that we might discover for E_j, using our current beliefs about its value:

$$VPI_{\mathbf{e}}(E_j) = \left(\sum_k P(E_j = e_{jk}|\mathbf{e})\; EU(\alpha_{e_{jk}}|\mathbf{e}, E_j = e_{jk}) \right) - EU(\alpha|\mathbf{e})$$

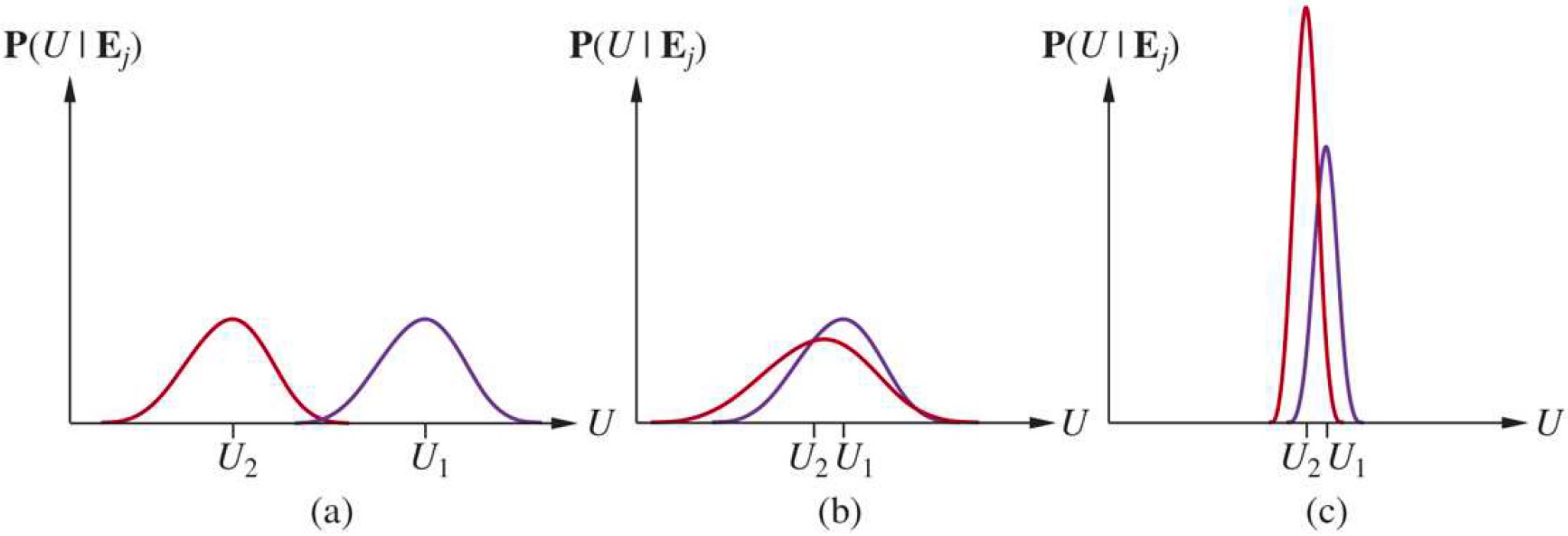
The value of information

$$VPI_{\mathbf{e}}(E_j) = \left(\sum_k P(E_j = e_{jk} | \mathbf{e}) \; EU(\alpha_{e_{jk}} | \mathbf{e}, E_j = e_{jk}) \right) - EU(\alpha | \mathbf{e})$$

where VPI stands for **value of perfect information.**

Information has value to the extent that it is likely to cause a change of plan and to the extent that the new plan will be significantly better than the old plan.

The value of information



Implementation of an information-gathering agent

- Assumptions:

1. With each observable evidence variable E_j, there is an associated cost, $Cost(E_j)$, which reflects the cost of obtaining the evidence through tests, consultants, questions, or whatever.
2. The agent requests the most efficient observation in terms of utility gain per unit cost.
3. The result of the action $Request(E_j)$ is that the next percept provides the value of E_j . If no observation is worth its cost, the agent selects a "real" action.

Implementation of an information-gathering agent

```
function INFORMATION-GATHERING-AGENT(percept) returns an action
    persistent: D, a decision network

    integrate percept into D
    j ← the value that maximizes VPI(E_j) / Cost(E_j)
    if VPI(E_j) > Cost(E_j)
        return REQUEST(E_j)
    else return the best action from D
```

Recap

- Sometimes, solving a problem involves finding more information before making a decision.
- The value of information is defined as the expected improvement in utility compared with making a decision without the information; it is particularly useful for guiding the process of information-gathering <u>prior to</u> making a final decision.