COMP 424 - Artificial Intelligence
Lecture 22: Utility Theory

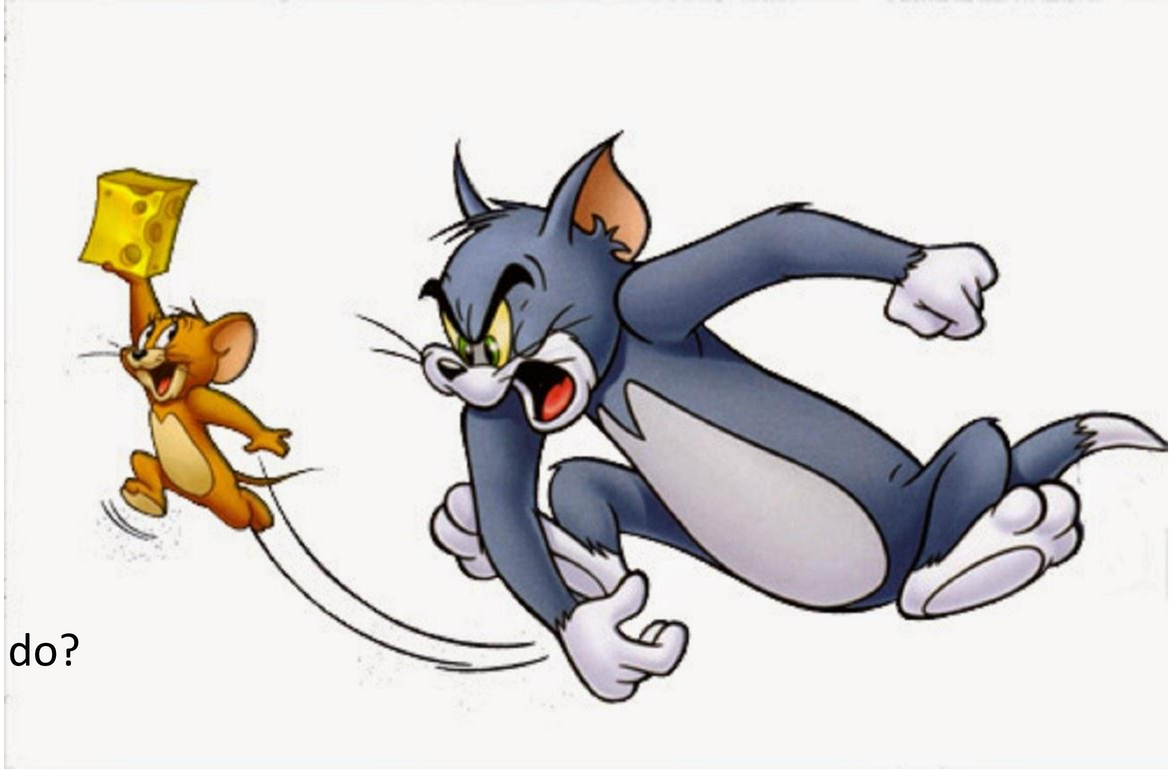
Instructor: Jackie CK Cheung (jcheung@cs.mcgill.ca)

Readings: R&N Ch 16

From probabilities to decisions

Probability theory

- What is the world like, accounting for uncertainty?
- Where is the location of the cheese to steal?



Utility theory

- What do agents want?
- The cheese!
- Not to get caught!

Decision theory

- What should an agent do?
- Keep running or hide?
- Make rational choice based on probability and utility theory

Bernoulli's puzzle

- You have the opportunity to play a game in which a fair coin is tossed repeatedly until it comes up heads.

- If the head appears on the nth toss, you win 2^n dollars.

- **Question: How much would you pay to play the game?**

Actions and consequences

- Intelligent agents should not only be *observers*, but also *actors*

  i.e., they should choose actions in a rational way.

- Most often, actions produce *consequences*, which cause changes in the world.

- Decision-making should maximize the overall **utility** of the agent's actions.

Preferences

- Actions have consequences. We call the consequences of an action **payoffs** or **rewards**
- A rational method would be to evaluate the *benefit* (desirability, value) of each consequence and *weigh* it by its *probability*

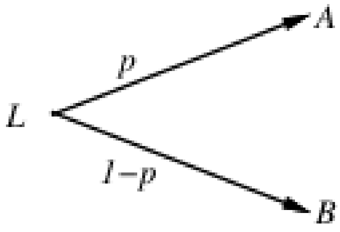
- To compare different actions, we need to know for each:
  - set of consequences $C_a = \{ c_1, \ldots, c_m \}$
  - probability distribution over consequences $P_a(c_i)$, s.t. $\sum_i P_a(c_i) = 1$.

Lotteries

- A pair $L_a = (C_a, P_a)$ is called a **lottery** (Luce and Raiffa, 1957).
  - A lottery is usually represented as a list of pairs

    e.g., $L_a = [\ A, p; B, (1\text{-}p)\ ]$.

    or as a tree-like diagram:



- Choosing between actions corresponds to **choosing between lotteries** corresponding to these actions.
- Agents have preferences over consequences:

  A > B : A preferred to B

  A ~ B : indifference between A and B

  A $\gtrsim$B : B not preferred to A

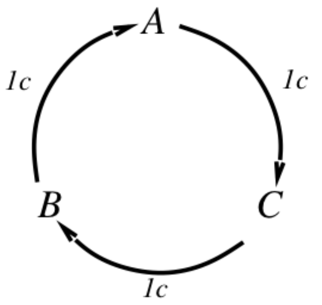
The axioms of utility theory

- For an agent to act rationally, *its preferences have to obey certain constraints*.

- These axioms are called the **axioms of utility theory**.
  1. Orderability
  2. Continuity
  3. Substitutability
  4. Monotonicity
  5. Reduction of compound lotteries

The axioms of utility theory (2)

1. **Orderability**: A linear and transitive preference relation must exist between the prizes of any lottery.
   - Linearity: (A > B) v (B > A) v (A ~ B)
   - Transitivity: $(A > B) \wedge (B > C) \Rightarrow (A > C)$



Suppose an agent with following preferences:

B>C, A>B, C>A and it owns C.

- If B>C then the agent would pay (say) 1 cent to get B.
- If A>B then the agent (who now has B) would pay (say) 1 cent to get A.
- If C>A then the agent (who now has A) would pay (say) 1 cent to get C.

The agent loses money forever. (Not rational behaviour!)

The axioms of utility theory (3)

2. **Continuity**: If A>B>C, then there exists a lottery L with prizes A and C that is equivalent to receiving B for sure: $\exists p, L = [p, A; (1-p)\ C] \sim B$

   The probability *p* at which equivalence occurs can be used to compare the merits of B w.r.t. A and C.

3. **Substitutability**: Adding the same prize with the same probability to two equivalent lotteries does not change the preference between them.

4. **Monotonicity**: If two lotteries have the same prizes, the one producing the best prize most often is preferred.

5. **Reduction of compound lotteries** ("No fun in gambling"): Two consecutive lotteries can be compressed into a single equivalent lottery.

Reminder: Expected value

- For a discrete-valued random variable X, with n possible values $\{x_1, \ldots, x_n\}$, occurring with probabilities $p_1, \ldots, p_n$ respectively.

- Then the expected value (mean) of X is:

$$E[X] = \Sigma_{i=1:n}\, p_i\, x_i$$

Utilities

- **Utilities map outcomes (or states) to real values.**

- Given a preference behaviour, the utility function is *non-unique*.

  e.g., behaviour is invariant w.r.t. additive linear transformations:

  $$U'(x) = k_1 U(x) + k_2, \text{ where } k_1 > 0$$

- With deterministic prizes only (no lottery choice), only **ordinal utility** matters. i.e., total order on prizes.
- Utilities don't need to obey the same laws as expected values.

Money

- Suppose you had to choose between these two lotteries:
  - L_1: win \$1M for sure.
  - L_2: win \$5M with prob. 0.1
    win \$1M with prob. 0.89
    win \$0 with prob 0.01.

- Which would you choose?

Money (2)

- Suppose you had to choose between these two lotteries:
  - L_1: win \$1M for sure.
  - L_2: win \$5M with prob. 0.1
    win \$1M with prob. 0.89
    lose \$1M with prob 0.01.

- Which would you choose?

Money (2)

- Suppose you had to choose between these two lotteries:
  - L_1: win \$1M for sure.
  - L_2: win \$5M with prob. 0.1
    win \$1M with prob. 0.89
    lose \$1M with prob 0.01.

- Which would you choose?
- What if you were the head of the bank of Canada?

Money (3)

- Suppose you had to choose between these two lotteries:
  - L_1: win \$5M with prob 0.1
    win \$0 with prob 0.9.
  - L_2: win \$1M with prob. 0.3
    win \$0 with prob 0.7.

- Which would you choose?

Money (3)

- Suppose you had to choose between these two lotteries:
  - L_1: win \$5M with prob 0.1
    win \$0 with prob 0.9.
  - L_2: win \$1M with prob. 0.3
    win \$0 with prob 0.7.

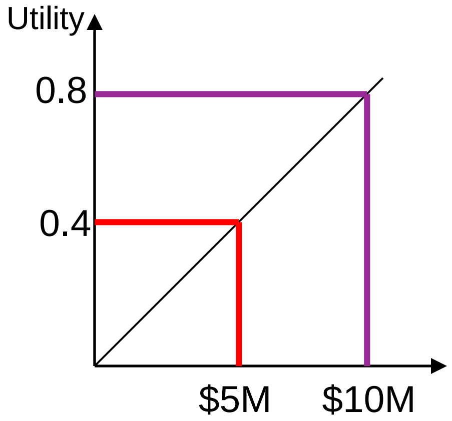
- Which would you choose?

Most people are risk-averse!

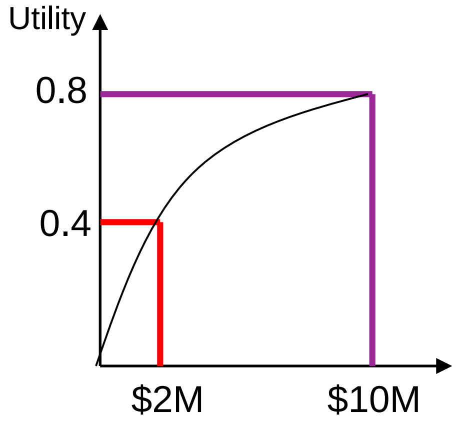
Utility models

- Capture preferences for rewards and resource consumption.
- Capture risk attitude

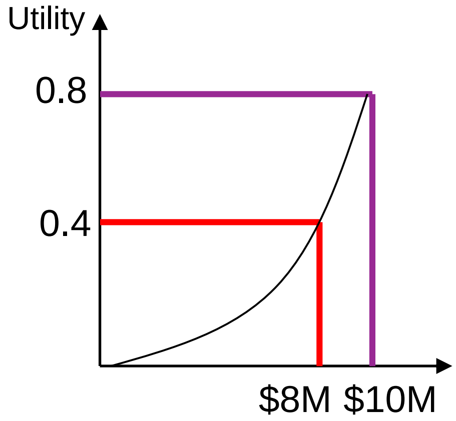
E.g. If risk-neutral, getting \$5M has half the utility of getting \$10M.



Risk Neutral
(Utility= Expected reward)

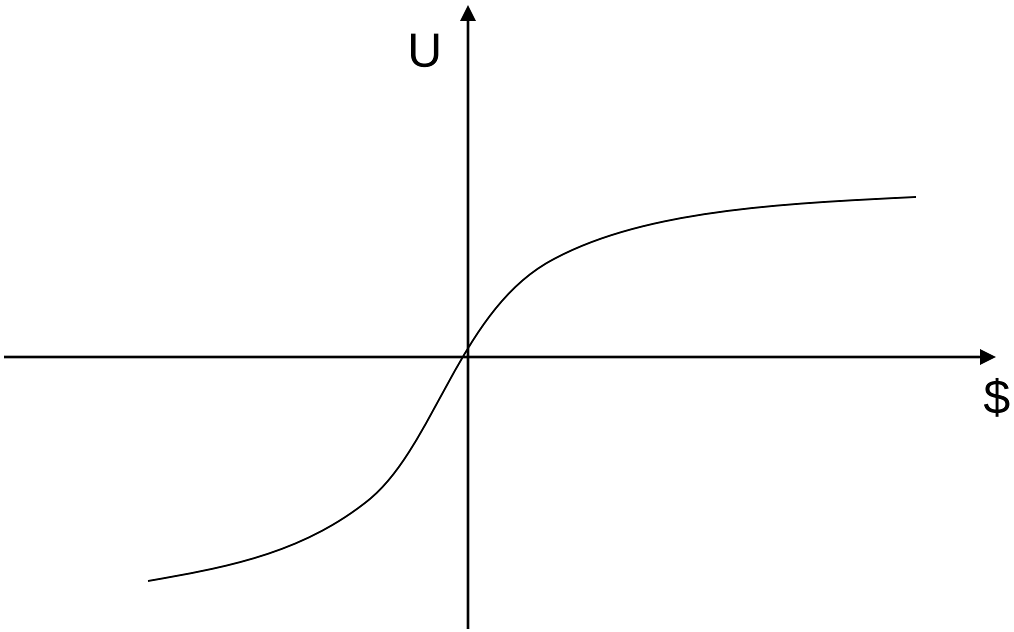


Risk Averse



Risk Seeking

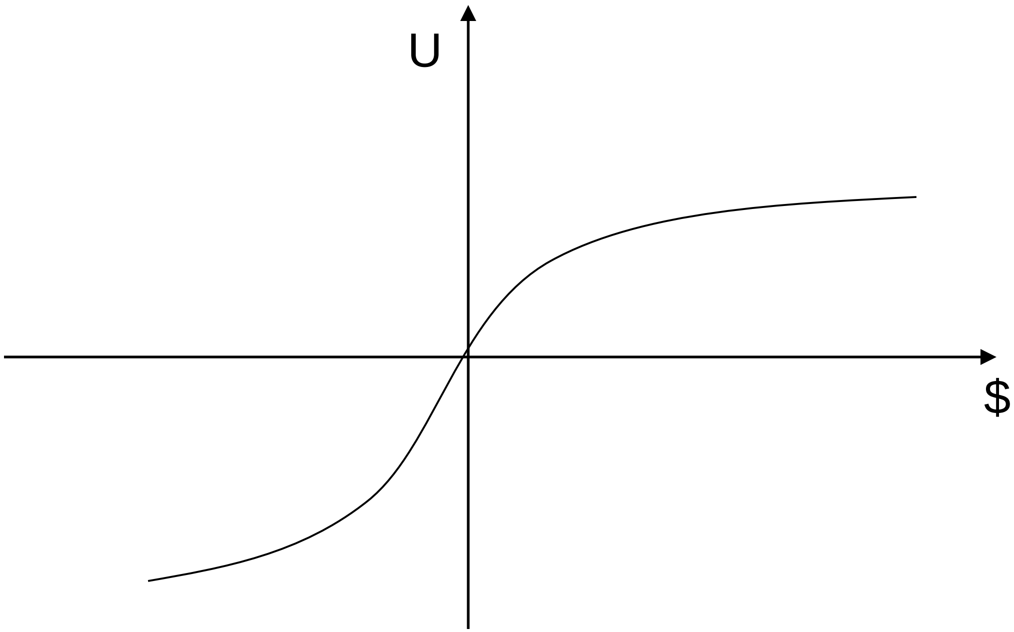
The utility of money



- Decision-theory is **normative**: describes how **rational** agents should act.

  => Useful to define an optimization criteria for AI agents.

The utility of money



- Decision-theory is **normative**: describes how **rational** agents should act.
  => Useful to define an optimization criteria for AI agents.

- People systematically violate the axioms of utility theory!
  => Or maybe we don't understand their utility function?

Poll: Let’s Play a Game

- You are forced to play the following pair of lotteries concurrently. For each, indicate the option that you prefer.

Decision (i)

A. a sure gain of $240
B. 25% chance to gain $1000, 75% chance to gain nothing

Decision (ii)

C. a sure loss of $750
D. 75% chance to lose $1000, 25% chance to lose nothing

Framing Effects

- Experiment by Tversky and Kahneman (1981)
- They found that 84% of respondents chose A, and 87% of respondents chose D, so the majority chose A&D [N=150].

- Let's compare A&D vs B&C:

A&D.	25% chance to win $240, 75% chance to lose $760
B&C.	25% chance to win $250, 75% chance to lose $750

- If presented this way, everybody chose B&C.
- The way in which decisions are presented matters! This is called a **framing effect**.

Utility at a societal level

- Government-run health insurance: how to decide what treatments are covered, given limited resources?
- NHS (UK): utility based on **quality-adjusted of life years** (**QALY**) – maximize years of good health
- But non-intuitive outcomes are possible:
  - Being completely blind is *much* worse than being blind in one eye
  - Being blind in one eye is *somewhat* worse than being fully sighted
  - **Conclusion**: lower priority to prevent vision loss in just one eye
  - So, expensive treatment for macular degeneration only covered if you are already blind in one eye: http://www.telegraph.co.uk/news/uknews/4182723/You-must-lose-sight-in-one-eye-before-NHS-will-treat-you.html

Acting Under Uncertainty

- **MEU principle**: Choose the action that maximizes expected utility.
  - Most widely accepted as a standard for rational behavior.

- Note than an agent can be entirely rational, i.e. consistent with MEU, without ever representing or manipulating utilities and probabilities. **Example??**

Maximizing expected utility (MEU)

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944):

Given preferences that satisfy these axioms, there exists a real-valued function U such that:

$$A \succsim B \text{ iff } U(A) \geq U(B)$$

where

$$U([p_1, C_1; \ldots; p_n, C_n]) = \sum_i p_i \, U(C_i)$$

Example: single-stage decision-making

- One random variable, *X*: does the child have an ear infection or not?
- One decision, *d*: give antibiotic (yes) or not (no)
- **Utility function**: associates a real value to the possible states of the world and possible decisions.

	$X =$ no	$X =$ yes
$d =$ no	0	-50
$d =$ yes	-100	10

- Unfortunately *X* is not directly observable!
- But we know *Pr(X=yes) = 0.1* and *Pr(X=no)=0.9*.
- According to MEU what is the best action?

Maximizing expected utility

- Compute:

$$\begin{aligned} EU(d = \text{no}) &= 0.9 \times 0 + 0.1 \times (-50) = -5 \\ EU(d = \text{yes}) &= 0.9 \times (-100) + 0.1 \times 10 = -89 \end{aligned}$$

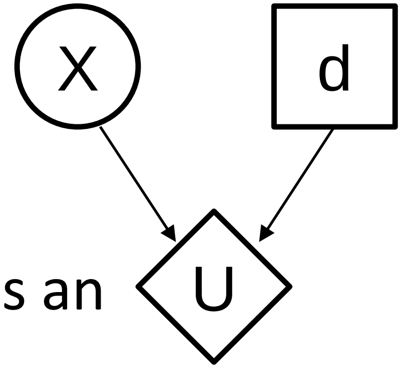
- Best action give this utility function and probability is *d* = no.

Useful definitions for utility theory

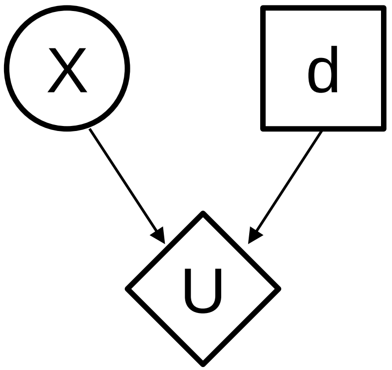
- **Utility function** $U(x)$
  - Numerical expression of the desirability of a state
- **Expected Utility** $EU(\, a \mid x \,)$
  $= \sum_i Pr(\, Effect(a) \mid x \,)\, U(\, Effect(a) \,)$
  - Utility of an action weighted by the expected outcome of that action
- **Maximum Expected Utility** $max_a\, EU(\, a \mid x \,)$
  - Best average payoff that can be achieved in situation *x*.
- **Optimal Action** $argmax_a\, EU(\, a \mid x \,)$
  - Action chosen according to the MEU principle.
- **Policy** $\pi(x)\text{:}\;\; X \text{->} A$
  - A strategy for picking actions in all states.

Decision graphs

- Represent decision models graphically:
  - **Random variables** are represented as **oval** nodes.
    - Parameters associated with such nodes are *probabilities*.
  - **Decisions (actions)** are represented as **rectangles**.
  - **Utilities** are represented as **diamonds**.
    - Parameters associated with such nodes are *utility values* for all possible values of the parents.
- Restrictions on nodes:
  - Utility nodes have no out-going arcs.
  - Decision nodes have no incoming arcs.
- Computing the optimal action can be viewed as an **inference** task.



Example



1. Suppose we had evidence that *X=yes*:
   - We can set *d* to each possible value (*yes/no*).
   - For each value, ask the utility node to give the utility of that situation, then pick *d* according to MEU.

2. If there is no evidence at *X*: we will have to *sum out* (*marginalize*) over all possible values of *X*, like in Bayes net inference.

- *X* can be a set of variables (incl. partially observable, e.g. HMM.)
- This gives the expected utility at node *U*, for each choice of action *d*.

Example

Buying oil drilling rights:

- Two blocks, A and B, exactly one has oil, worth *k*.
- Prior probability 0.5 for each block, mutually exclusive.
- Current price of each block is *k/2*.

- What does the decision network look like?

Information gathering

- In an environment with hidden information, an agent can choose to perform **information-gathering actions**.

  e.g., taking the child to the doctor.

  e.g., scouting the price of a product at different companies.

- Sometimes, such actions take time, or have associated costs (e.g., medical tests.) *When are they worth pursuing?*

- The **value of information** specifies the utility of every piece of evidence that can be acquired.

Example: Value of information

Buying oil drilling rights:

- Two blocks, A and B, exactly one has oil, worth *k*.
- Prior probability 0.5 for each block, mutually exclusive.
- Current price of each block is *k/2*.
- **Consultant offers accurate survey of *A*.**

What is a fair price for the survey?

Solution for the example

- Compute:

Expected value of information =

expected value of best action **given** the information

- expected value of best action **without** information.

Survey may say "oil in A" or "no oil in A" with Pr=0.5 each.

Value = [0.5 * value of "buy A" given "oil in A"

+ 0.5 * value of "buy B" given "no oil in A"]

- [expected return of "buy A" - cost of "buy A"]

= [0.5 * $k/2$ + 0.5 * $k/2$] - [$k/2$ - $k/2$]

= $k/2$

Value of Perfect Information (VPI)

- Suppose you have current evidence E, current best action a^*, with possible outcomes c_i. Then the expected value of a^* is:

$$EU(a^*|E) = \max_a U(a) = \max_a \sum_i U(c_i)P(c_i|E,a)$$

- Suppose you could gather further evidence about a variable X, should you do it?

Value of Perfect Information

- Suppose we knew *X=x*, then we would choose a_x* such that:

$$EU(a_x^*|E, X = x) = \max_a \sum_i U(c_i)P(c_i|E, a, X = x)$$

- *X* is a random variable whose value is unknown, so we must compute expected gain over all possible values:

$$VPI_E(X) = \left(\sum_x P(X = x|E)EU(a_x^*|E, X = x) \right) - EU(a^*|E)$$

This is the value of knowing *X* exactly!

Exercise

- The final exam for a course will test topic A or topic B.
  - Your prior: Pr(Test A = 0.4), Pr(Test B = 0.6)
- You can choose to focus your studying on topic A or topic B, with the following utilities:

	Test A	Test B
Study A	100	50
Study B	60	100

- At the final review tutorial, the TA will tell you which topic will be tested. *How much is this information worth to you?* (After all, it takes time to attend the tutorial!)

Properties of VPI

- **Non-negative**:

  $$\forall X, E \quad \mathrm{VPI}_E(X) \geq 0$$

  Note that VPI is an *expectation*. Depending on the actual value we find for X, there can actually be a loss post-hoc.

- **Non-additive**: E.g. consider obtaining X twice.

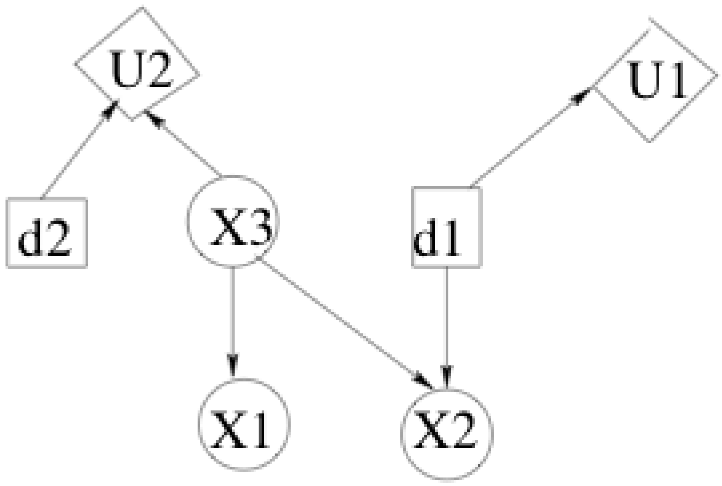
  $$\mathrm{VPI}_E(X, X) \neq \mathrm{VPI}_E(X) + \mathrm{VPI}_E(X)$$

- **Order-independent**:

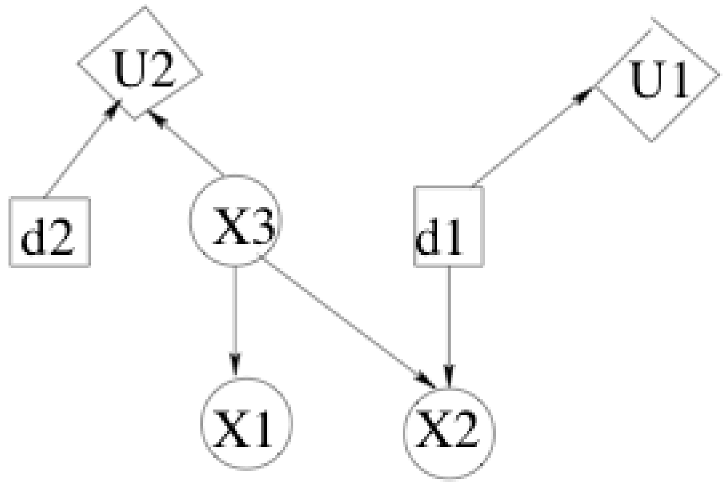
  $$\mathrm{VPI}_E(X, Y) = \mathrm{VPI}_E(X) + \mathrm{VPI}_{E,X}(Y) = \mathrm{VPI}_E(Y) + \mathrm{VPI}_{E,Y}(X)$$

A more complex example

- X1: Symptoms present
- X2: Infection observed
- X3: Infection present
- d1: Go see the doctor?
- d2: Take antibiotics?



A more complex example (cont'd)



- Total utility is U1+U2
- X2 is only observed if
- X3 is never observed

Now we have to optimize d1 and d2 together!

Acting under uncertainty: Other options

- Sometimes it can be advantageous to not always choose actions according to MEU, e.g. if the environment may change, or it is not fully known to the agent.
- **Random choice models**: choose the action with the highest expected utility most of the time, but keep non-zero probabilities for other actions as well.
  - Avoids being too predictable.
  - If utilities are not perfect, allows for exploration.
- **Minimizing regret**: consider the loss between the current behaviour and some "gold standard" behaviour, and try to minimize this loss.

Preference Elicitation

- Applications often require recommending something to a user or making a decision for them:
  - Online movie or book recommendation systems
  - Deciding which cancer treatment to give to a patient (has to take into account chance of survival, cost, side effects, etc.)
  - Deciding which ads to show on a dynamic web page
- For this, we need to know the utility that the user associates to different items.
- People are very bad at specifying utility values!
- **Preference elicitation** refers to finding out their preferences and translating them into utilities. Hard problem!

Summary

- To make decisions under uncertainty, we need to know likelihood (probability) of different outcomes, and have preferences among outcomes:

  Decision Theory = Probability Theory + Utility Theory

- An agent with consistent preferences has a utility function, which associates a real number to each possible state.
  - Rational agents try to maximize their (expected) utility.
  - Utility theory allows us to tell if gathering more information is valuable.
- Decision graphs can be used to represent decision problems.
  - An algorithm similar to variable elimination is useful to compute optimal decision, but this is very expensive in general.