18-0: **Making decisions**

- At this point, we know how to describe the probability of events occuring.
  - Or states being reached, in agent terms.
- Knowing the probabilities of events is only part of the battle.
- Agents are really interested in maximizing performance.
- This means making the correct decisions as to what to do.
- Often, performance can be captured by *utility*.
- Utility indicates the relative value of a state.

18-1: **Expected Utility**

- In episodic, stochastic worlds, we can use expected utility to select actions.
- An agent will know that an action can lead to one of a set S of states.
- The agent has a utility for each of these states.
- The agent also has a probability that these states will be reached.
- The *expected utility* of an action is:
- $\sum_{s \in S} P(s)U(s)$
- The principle of maximum expected utility says that an agent should choose the action that maximizes expected utility.

18-2: **Example**

- Let's say there are two levers.
  - Lever 1 costs \$1 to pull. With $p = 0.4$, you win \$2. With $p = 0.6$ you win nothing.
  - Lever 2 costs \$2 to pull. With $p = 0.1$ you win \$10. with $p = 0.9$ you lose \$1 (on top of the charge to pull).
- Should you a) pull lever 1 b) pull lever 2 c) pull neither?

18-3: **Example**

- EU(lever 1) = 0.4 * 1 + 0.6 * -1 = -0.2
- EU(lever 2) = 0.1 * 8 + 0.9 * -3 = -1.9
- EU(neither) = 0
- Lever 3 gives the maximum EU.
- TV digression - this is the choice contestants are faced with on "Deal or No Deal." The banker offers them a price slightly above the expected utility, and yet most contestants don't take it. Why?

18-4: **Example: Vegas**

- The typical roulette wheel has 38 numbers. (1-36, plus 0 and 00).

- 1-36 are either red or black.
- The payoff for betting on a single number is 35:1
  - In other words, if the number you picked comes up, you win $35. Otherwise, you lose $1.
- What is the expected utility of betting on a single number?

18-5: **Example: Vegas**

- The typical roulette wheel has 38 numbers. (1-36, plus 0 and 00).
- 1-36 are either red or black.
- The payoff for betting on a single number is 35:1
  - In other words, if the number you picked comes up, you win $35. Otherwise, you lose $1.
- What is the expected utility of betting on a single number?
- $\frac{1}{38} * 35 + \frac{37}{38} * -1 = -0.052$

18-6: **Example: Vegas**

- The typical roulette wheel has 38 numbers. (1-36, plus 0 and 00).
- 1-36 are either red or black.
- The payoff for betting on a single number is 35:1
  - In other words, if the number you picked comes up, you win $35. Otherwise, you lose $1.
- What if you decide to “spread the risk” and bet on two numbers?

18-7: **Example: Vegas**

- The typical roulette wheel has 38 numbers. (1-36, plus 0 and 00).
- 1-36 are either red or black.
- The payoff for betting on a single number is 35:1
  - In other words, if the number you picked comes up, you win $35. Otherwise, you lose $1.
- What if you decide to “spread the risk” and bet on two numbers?
- $\frac{2}{38} * 34 + \frac{36}{38} * -2 = -0.105$

18-8: **Example: Vegas**

- The typical roulette wheel has 38 numbers. (1-36, plus 0 and 00).
- 1-36 are either red or black.
- The payoff for betting on color (red or black) is 1:1
  - In other words, if you bet ’red’ and a red number comes up, you win $1. Otherwise, you lose $1.
- What is the expected utility of betting on ’red’?

18-9: **Example: Vegas**

- The typical roulette wheel has 38 numbers. (1-36, plus 0 and 00).
- 1-36 are either red or black.
- The payoff for betting on color (red or black) is 1:1
  - In other words, if you bet 'red' and a red number comes up, you win \$1. Otherwise, you lose \$1.
- What is the expected utility of betting on 'red'?
- $\frac{18}{38} * 1 + \frac{20}{38} * -1 = -0.052$

18-10: **Regarding Preferences**

- In order for MEU to make sense, we need to specify some (hopefully reasonable) constraints on agent preferences.
- Orderability. We must be able to say that A is preferred to B, B is preferred to A, or they are equally preferable. We cannot have the case where A and B are incomparable.
- Transitivity. If $A \prec B$ and $B \prec C$, then $A \prec C$.
- Continuity. If $A \prec B \prec C$, then there is a scenario where the agent is indifferent to getting B and having a probability p of getting A and $1 - p$ chance of getting C.

18-11: **Rational Preferences**

- Monotonicity. If two actions A and B have the same outcomes, and I prefer A to B, I should still prefer A if the probability of A increases.
- Decomposability. Utilities over a sequence of actions can be decomposed into utilities for atomic events.
- These preferences are (for the most part) quite reasonable, and allow an agent to avoid making foolish mistakes.

18-12: **Utility, Money, and Risk**

- Utility comes from economics
  - Money is often used as a substitute for utility.
- Preferences for money behave oddly when dealing with small or large amounts.

18-13: **Utility, Money, and Risk**

- You have two choices:
- Choice # 1
  - Win \$1 Million for sure
- Choice #2
  - Win \$2 Million with probability 0.51
  - Win nothing with probability 0.49
- Which do you pick?

18-14: **Utility, Money, and Risk**

- You have two choices:

- Choice # 1
  - Win $1 Million for sure
  - Expected winnings of $1 Million
- Choice #2
  - Win $2 Million with probability 0.51
  - Win nothing with probability 0.49
  - Expected winnings of > $1 Million
- Most people pick Choice #1
  - First $1 Million has huge impact, second $1 Million less so

18-15: **Utility, Money, and Risk**

- Flip a coin.
  - On heads, you win $4
  - On tails, you win nothing
- How much would you be willing to pay to play this game?

18-16: **Utility, Money, and Risk**

- Flip a coin until it comes up tails.
  - Count the number of heads h
  - Win $\$2^h$
- How much would you be willing to pay to play this game?

18-17: **Utility, Money, and Risk**

- Flip a coin until it comes up tails.
  - Count the number of heads h
  - Win $\$2^h$
- Expected return:

$$1/2 * 1 + 1/4 * 2 + 1/8 * 4 + 1/16 * 8 + \ldots$$

$$\sum_{i=1}^{\infty} 1/2 = \infty$$

18-18: **Utility, Money, and Risk**

- Does this mean utility theory is broken?

18-19: **Utility, Money, and Risk**

- Does this mean utility theory is broken?

  - No, just that the Utility of money is non-linear
  - Closer to logarithmic

18-20: **Utility, Money, and Risk**

- Utility comes from economics
  - Money is often used as a substitute for utility.
- Preferences for money behave oddly when dealing with small or large amounts.
- For example, you will often take more chance with small amounts, and be very conservative with very large amounts.
- This is called your *risk profile*
  - convex - risk-seeking (Vegas)
  - concave - risk-averse (Insurance)
- Typically, we say that we have a *quasilinear* utility function regarding money.

18-21: **Gathering information**

- When an agent has all the information it can get and just needs to select a single action, things are straightforward.
  - Find the action with the largest expected utility.
- What if an agent can choose to gather more information about the world?
- Now we have a sequential decision problem:
  - Should we just act, or gather information first?
  - What questions should we ask?
    - Agents should ask questions that give them useful information.
    - "Useful" means increasing expected utility.
  - Gathering information can cost time/money.

18-22: **Example**

- An agent can recommend to a prospecting company that they buy one of n plots of land.
- One of the plots has oil worth C dollars; the others are empty.
- Each block costs C/n dollars.
- Initially, agent is indifferent between buying and not buying. (why is that?)

18-23: **Example**

- Suppose the agent can perform a survey on a block that will indicate whether that block contains oil.
- How much should the agent pay to perform that survey?

18-24: **Example**

- How much should the agent pay to perform that survey?

- $P(oil) = 1/n$. $P(\neg oil) = (n-1)/n$
  - If oil found, buy for C/n. Profit = $C - C/n = (n-1)C/n$
  - If oil not found, buy a different block.
  - Probability of picking the oil block is now: $1/(n-1)$ Expected Profit: $C/(n-1) - C/n = C/(n*(n-1))$.
- So, the expected profit, given the information is:
- $\frac{1}{n}\frac{(n-1)C}{n} + \frac{n-1}{n}\frac{C}{n(n-1)} = \frac{C}{n}$

18-25: **Example**

- Without doing the survey, expected profit is 0
- With doing the survey, expected profit is $\frac{C}{n}$
- Company is willing to pay up to C/n (the expected value of the plot itself) to do the test.
  - This is the *value of that information*.

18-26: **Value of Perfect Information**

- Let's formalize this.
- We find the best action α in general with:
- $EU(\alpha) = max_{a \in actions} \sum_{i \in outcomes} U(i)P(i|a)$
- Let's say we acquire some new information E.
- Then we find the best action with:
- $EU(\alpha|E) = max_{a \in actions} \sum_{i \in outcomes} U(i)P(i|a, E)$
- The value of E is the difference between these two.

18-27: **Value of Perfect Information**

- However, before we do the test, we don't know what E will be.
- We must average over all possible values of E.
- $VPI(E) = (\sum_{j \in values_E} P(E = j)EU(\alpha|E = j)) - EU(\alpha)$
- In words, consider the possibility of each observation, along with the usefulness of that observation, to compute the expected information gain from this test.
- In general, information will be valuable to the extent that it changes the agent's decision.

18-28: **Example**

- Imagine that you are on a game show and are given a choice of three possible doors to open.
- If you open door number 1, you will win \$10.
- If you open door number 2, you have a 50% change of winning nothing, and a 50% chance of winning \$25.
- If you open door number 3, you have a 20% of winning \$20, and an 80% chance of winning \$10.
- Which door should you choose?

18-29: **Example**

- EU(door1) = 10
- EU(door2) = 0.5 * 0 + 0.5 * 25 = 12.5
- EU(door3) = 0.2 * 20 + 0.8 + 10 = 12
- Door 2 is best.

18-30: **Example**

- Now, suppose that the host offers to tell you honestly what you'll get if you open door number 2. How much would you pay for this information?
- Note: you can change your mind about which door to open after the host gives you this information.

18-31: **Example**

- There are two cases: either door 2 pays 0 or it pays 25.
- If it pays 25, we should choose it. This happens 50% of the time.
- If it pays 0, we should choose door3, which has an expected payment of 12. This happens 50% of the time.
- So, our utility will be: 0.5 * 25 + 0.5 * 12 = 18.5
- Our EU before we asked was 12.5, so the information is worth 6.

18-32: **Example II**

- Drive across canyon
  - Go around, takes 10 hours
  - Go through the canyon, takes 2 hours without traffic, and 15 hours with traffic. P(traffic) = 0.5
- Which route should you take, to minimize the expected travel time?

18-33: **Example II**

- Drive across canyon
  - Go around, takes 10 hours
    - Expected time = 10 hours
  - Go through the canyon, takes 2 hours without traffic, and 15 hours with traffic. P(traffic) = 0.5
    - Expected time = 2 * 0.5 + 15 * 0.5 = 8.5 hours

18-34: **Example II**

- Drive across canyon
  - Go around, takes 10 hours
  - Go through the canyon, takes 2 hours without traffic, and 15 hours with traffic. P(traffic) = 0.5
  - Drive to an overlook, see if there is traffic
- How long could you spend driving to / from the overlook, and have it still be worthwhile?

18-35: **Example II**

- Takes time n to drive to overlook and back
- If there is traffic, go around, for a total time of $10 + n$
- If there is no traffic, go through, for a total time of $2 + n$
- Expected time $= 0.5 * (10 + n) + 0.5 * (2 + n) = 6 + n$
- Without driving to overlook, best expected time = 8.5 hours
- As long as driving to the overlook takes less than 2.5 hours, worth it