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

Preferences

- Actions result in outcomes
- Agents have preferences over outcomes
- A rational agent will do the action that has the best outcome for them
- Sometimes agents don't know the outcomes of the actions, but they still need to compare actions
- Agents have to act (doing nothing is (often) an action).

Preferences Over Outcomes

If o_1 and o_2 are outcomes

- $o_1 \succeq o_2$ means o_1 is at least as desirable as o_2.
- $o_1 \sim o_2$ means $o_1 \succeq o_2$ and $o_2 \succeq o_1$.
- $o_1 \succ o_2$ means $o_1 \succeq o_2$ and $o_2 \not\succeq o_1$

Lotteries

- An agent may not know the outcomes of their actions, but only have a probability distribution of the outcomes.
- A lottery is a probability distribution over outcomes. It is written

$$[p_1 : o_1, p_2 : o_2, \dots, p_k : o_k]$$

where the o_i are outcomes and $p_i > 0$ such that

$$\sum_i p_i = 1$$

The lottery specifies that outcome o_i occurs with probability p_i.

- When we talk about outcomes, we will include lotteries.

Properties of Preferences

- **Completeness:** Agents have to act, so they must have preferences:

$$\forall o_1 \forall o_2 \; o_1 \succeq o_2 \text{ or } o_2 \succeq o_1$$

- **Transitivity:** Preferences must be transitive:

$$\text{if } o_1 \succeq o_2 \text{ and } o_2 \succeq o_3 \text{ then } o_1 \succeq o_3$$

otherwise $o_1 \succeq o_2$ and $o_2 \succeq o_3$ and $o_3 \succ o_1$. If they are prepared to pay to get from o_1 to o_3 $\longrightarrow$ money pump. (Similarly for mixtures of $\succ$ and $\succeq$.)

Properties of Preferences (cont.)

Monotonicity: An agent prefers a larger chance of getting a better outcome than a smaller chance:

- If $o_1 \succ o_2$ and $p > q$ then

$$[p : o_1, 1 - p : o_2] \succ [q : o_1, 1 - q : o_2]$$

Consequence of axioms

- Suppose $o_1 \succ o_2$ and $o_2 \succ o_3$. Consider whether the agent would prefer
  - o_2
  - the lottery $[p : o_1, 1 - p : o_3]$

  for different values of $p \in [0, 1]$.
- You can plot which one is preferred as a function of p:

Properties of Preferences (cont.)

Continuity: Suppose $o_1 \succ o_2$ and $o_2 \succ o_3$, then there exists a $p \in [0, 1]$ such that

$$o_2 \sim [p : o_1, 1 - p : o_3]$$

Properties of Preferences (cont.)

Decomposability: (no fun in gambling). An agent is indifferent between lotteries that have same probabilities and outcomes. This includes lotteries over lotteries. For example:

$$
\begin{aligned}
& [p : o_1, 1-p : [q : o_2, 1-q : o_3]] \\
& \quad \sim \quad [p : o_1, (1-p)q : o_2, (1-p)(1-q) : o_3]
\end{aligned}
$$

Properties of Preferences (cont.)

Substitutability: if $o_1 \sim o_2$ then the agent is indifferent between lotteries that only differ by o_1 and o_2:

$$[p : o_1, 1 - p : o_3] \sim [p : o_2, 1 - p : o_3]$$

Alternative Axiom for Substitutability

Substitutability: if $o_1 \succeq o_2$ then the agent weakly prefers lotteries that contain o_1 instead of o_2, everything else being equal.

That is, for any number p and outcome o_3:

$$[p : o_1, (1-p) : o_3] \succeq [p : o_2, (1-p) : o_3]$$

What we would like

- We would like a measure of preference that can be combined with probabilities. So that

$$\begin{aligned} & value([p : o_1, 1 - p : o_2]) \\ = \;\; & p \times value(o_1) + (1 - p) \times value(o_2) \end{aligned}$$

- Money does not act like this.
  What would you prefer

$$\$1,000,000 \text{ or } [0.5 : \$0, 0.5 : \$2,000,000]?$$

- It may seem that preferences are too complex and muti-faceted to be represented by single numbers.

Theorem

If preferences follow the preceding properties, then preferences can be measured by a function

$$utility : outcomes \rightarrow [0, 1]$$

such that

- $o_1 \succeq o_2$ if and only if $utility(o_1) \geq utility(o2)$.
- Utilities are linear with probabilities:

$$\begin{aligned} & utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) \\ = \; & \sum_{i=1}^{k} p_i \times utility(o_i) \end{aligned}$$

Proof

- If all outcomes are equally preferred, set $utility(o_i) = 0$ for all outcomes o_i.
- Otherwise, suppose the best outcome is *best* and the worst outcome is *worst*.
- For any outcome o_i, define $utility(o_i)$ to be the number u_i such that

$$o_i \sim [u_i : best, 1 - u_i : worst]$$

This exists by the Continuity property.

Proof (cont.)

- Suppose $o_1 \succeq o_2$ and $utility(o_i) = u_i$, then by Substitutability,

$$\begin{aligned} & [u_1 : best, 1 - u_1 : worst] \\ & \quad \succeq \quad [u_2 : best, 1 - u_2 : worst] \end{aligned}$$

  Which, by completeness and monotonicity implies $u_1 \geq u_2$.

Proof (cont.)

- Suppose $p = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.
- Suppose $utility(o_i) = u_i$. We know:

$$o_i \sim [u_i : best, 1 - u_i : worst]$$

- By substitutability, we can replace each o_i by $[u_i : best, 1 - u_i : worst]$, so

$$\begin{aligned} p = utility(\quad [\quad & p_1 : [u_1 : best, 1 - u_1 : worst] \\ & \ldots \\ & p_k : [u_k : best, 1 - u_k : worst]]) \end{aligned}$$

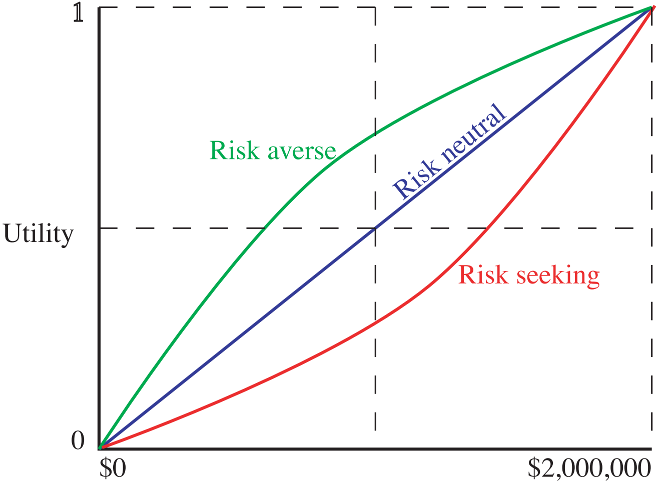
- By decomposability, this is equivalent to:

$$p = utility(\ [\ p_1 u_1 + \cdots + p_k u_k : best,\ p_1(1-u_1) + \cdots + p_k(1-u_k) : worst]])$$

- Thus, by definition of utility,

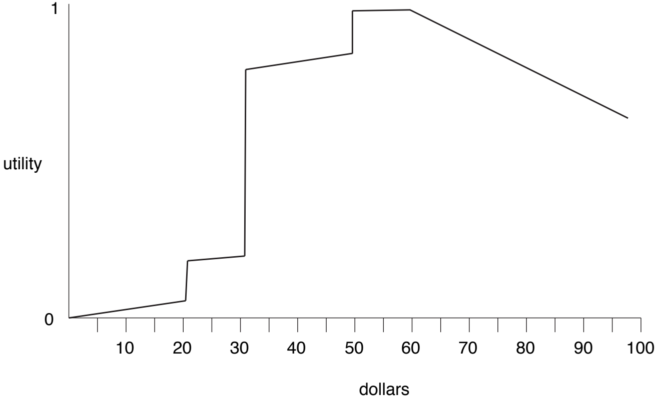
$$p = p_1 \times u_1 + \cdots + p_k \times u_k$$

Utility as a function of money



Possible utility as a function of money

Someone who really wants a toy worth \$30, but who would also like one worth \$20:



Allais Paradox (1953)

What would you prefer:

A: 1m$ — one million dollars

B: lottery $[0.10 : \$2.5m, 0.89 : \$1m, 0.01 : \$0]$

Allais Paradox (1953)

What would you prefer:

- A: $\$1m$ — one million dollars
- B: lottery $[0.10 : \$2.5m, 0.89 : \$1m, 0.01 : \$0]$

What would you prefer:

- C: lottery $[0.11 : \$1m, 0.89 : \$0]$
- D: lottery $[0.10 : \$2.5m, 0.9 : \$0]$

Allais Paradox (1953)

What would you prefer:

A: 1m$ — one million dollars

B: lottery $[0.10 : \$2.5m, 0.89 : \$1m, 0.01 : \$0]$

What would you prefer:

C: lottery $[0.11 : \$1m, 0.89 : \$0]$

D: lottery $[0.10 : \$2.5m, 0.9 : \$0]$

It is inconsistent with the axioms of preferences to have $A \succ B$ and $D \succ C$.

Allais Paradox (1953)

What would you prefer:

A: \$1m — one million dollars

B: lottery $[0.10 : \$2.5m, 0.89 : \$1m, 0.01 : \$0]$

What would you prefer:

C: lottery $[0.11 : \$1m, 0.89 : \$0]$

D: lottery $[0.10 : \$2.5m, 0.9 : \$0]$

It is inconsistent with the axioms of preferences to have $A \succ B$ and $D \succ C$.

A,C: lottery $[0.11 : \$1m, 0.89 : X]$

B,D: lottery $[0.10 : \$2.5m, 0.01 : \$0, 0.89 : X]$

The Ellsberg Paradox

Two bags:

- **Bag 1** 40 white chips, 30 yellow chips, 30 green chips
- **Bag 2** 40 white chips, 60 chips that are yellow or green

What do you prefer:

- A: Receive \$1m if a white or yellow chip is drawn from bag 1
- B: Receive \$1m if a white or yellow chip is drawn from bag 2
- C: Receive \$1m if a white or green chip is drawn from bag 2

The Ellsberg Paradox

Two bags:

- **Bag 1** 40 white chips, 30 yellow chips, 30 green chips
- **Bag 2** 40 white chips, 60 chips that are yellow or green

What do you prefer:

- A: Receive \$1m if a white or yellow chip is drawn from bag 1
- B: Receive \$1m if a white or yellow chip is drawn from bag 2
- C: Receive \$1m if a white or green chip is drawn from bag 2

What about

- D: Lottery $[0.5 : B, 0.5 : C]$

The Ellsberg Paradox

Two bags:

- **Bag 1** 40 white chips, 30 yellow chips, 30 green chips
- **Bag 2** 40 white chips, 60 chips that are yellow or green

What do you prefer:

- A: Receive \$1m if a white or yellow chip is drawn from bag 1
- B: Receive \$1m if a white or yellow chip is drawn from bag 2
- C: Receive \$1m if a white or green chip is drawn from bag 2

What about

- D: Lottery $[0.5 : B, 0.5 : C]$

However A and D should give same outcome, no matter what the proportion in Bag 2.

St. Petersburg Paradox

What if there is no "best" outcome?
Are utilities unbounded?

St. Petersburg Paradox

What if there is no "best" outcome?
Are utilities unbounded?

- Suppose they are unbounded.
- Then for any outcome o_i there is an outcome o_{i+1} such that $u(o_{i+1}) > 2u(o_i)$.
- It is rational to give up o_1 to play the lottery $[0.5 : o_2, 0.5 : 0]$.
- It is then rational to gamble o_2 to on a coin toss to get o_3.
- It is then rational to gamble o_3 to on a coin toss to get o_4.

St. Petersburg Paradox

What if there is no "best" outcome?
Are utilities unbounded?

- Suppose they are unbounded.
- Then for any outcome o_i there is an outcome o_{i+1} such that $u(o_{i+1}) > 2u(o_i)$.
- It is rational to give up o_1 to play the lottery $[0.5 : o_2, 0.5 : 0]$.
- It is then rational to gamble o_2 to on a coin toss to get o_3.
- It is then rational to gamble o_3 to on a coin toss to get o_4.
- In this infinite sequence of bets you are guaranteed to lose everything.

Predictor Paradox

Two boxes:

Box 1: contains \$10,000

Box 2: contains either \$0 or \$1m

- You can either choose both boxes or just box 2.

Predictor Paradox

Two boxes:

Box 1: contains \$10,000

Box 2: contains either \$0 or \$1m

- You can either choose both boxes or just box 2.
- The "predictor" has put \$1m in box 2 if he thinks you will take box 2 and \$0 in box 2 if he thinks you will take both.
- The predictor has been correct in previous predictions.
- Do you take both boxes or just box 2?

Framing Effects [Tversky and Kahneman]

- A disease is expected to kill 600 people. Two alternative programs have been proposed:

  Program A: 200 people will be saved
  Program B: probability $1/3$: 600 people will be saved
  probability $2/3$: no one will be saved

  Which Program Would you favor?

Framing Effects [Tversky and Kahneman]

- A disease is expected to kill 600 people. Two alternative programs have been proposed:

  Program C: 400 people will die

  Program D: probability $1/3$: no one will die
  probability $2/3$: 600 will die

  Which Program Would you favor?

Framing Effects [Tversky and Kahneman]

- A disease is expected to kill 600 people. Two alternative programs have been proposed:

  Program A: 200 people will be saved
  Program B: probability 1/3: 600 people will be saved
  probability 2/3: no one will be saved

  Which Program Would you favor?

- A disease is expected to kill 600 people. Two alternative programs have been proposed:

  Program C: 400 people will die
  Program D: probability 1/3: no one will die
  probability 2/3: 600 will die

  Which Program Would you favor?

Tversky and Kahneman: 72% chose A over B.
22% chose C over D.

Framing Effects

- Suppose you had bought tickets for the theatre for \$50. When you got to the theatre, you had lost the tickets. You have your credit card and can buy equivalent tickets for \$50. Do you buy the replacement tickets on your credit card?

Framing Effects

- Suppose you had bought tickets for the theatre for \$50. When you got to the theatre, you had lost the tickets. You have your credit card and can buy equivalent tickets for \$50. Do you buy the replacement tickets on your credit card?
- Suppose you had \$50 in your pocket to buy tickets. When you got to the theatre, you had lost the \$50. You have your credit card and can buy equivalent tickets for \$50. Do you buy the tickets on your credit card?

Factored Representation of Utility

- Suppose the outcomes can be described in terms of features $X_1, \ldots, X_n$.
- An additive utility is one that can be decomposed into set of factors:

$$u(X_1, \ldots, X_n) = f_1(X_1) + \cdots + f_n(X_n).$$

  This assumes additive independence.
- Strong assumption: each feature has a best and a worst value that doesn't depend on other features.
- Many ways to represent the same utility:
  — a number can be added to one factor as long as it is subtracted from others.

Additive Utility

- An additive utility has a canonical representation:

$$u(X_1, \ldots, X_n) = w_1 \times u_1(X_1) + \cdots + w_n \times u_n(X_n).$$

- If $best_i$ is the best value of X_i, $u_i(X_i{=}best_i) = 1$.
  If $worst_i$ is the worst value of X_i, $u_i(X_i{=}worst_i) = 0$.
- w_i are weights, $\sum_i w_i = 1$.
  The weights reflect the relative importance of features.
- We can determine weights by comparing outcomes.

$$w_1 =$$

Additive Utility

- An additive utility has a canonical representation:

$$u(X_1, \ldots, X_n) = w_1 \times u_1(X_1) + \cdots + w_n \times u_n(X_n).$$

- If $best_i$ is the best value of X_i, $u_i(X_i{=}best_i) = 1$.
  If $worst_i$ is the worst value of X_i, $u_i(X_i{=}worst_i) = 0$.
- w_i are weights, $\sum_i w_i = 1$.
  The weights reflect the relative importance of features.
- We can determine weights by comparing outcomes.

$$w_1 = u(best_1, x_2, \ldots, x_n) - u(worst_1, x_2, \ldots, x_n).$$

for any values $x_2, \ldots, x_n$ of $X_2, \ldots, X_n$.

Complements and Substitutes

- Often additive independence is not a good assumption.
- Values x_1 of feature X_1 and x_2 of feature X_2 are **complements** if having both is better than the sum of the two.
- Values x_1 of feature X_1 and x_2 of feature X_2 are **substitutes** if having both is worse than the sum of the two.

Complements and Substitutes

- Often additive independence is not a good assumption.
- Values x_1 of feature X_1 and x_2 of feature X_2 are complements if having both is better than the sum of the two.
- Values x_1 of feature X_1 and x_2 of feature X_2 are substitutes if having both is worse than the sum of the two.
- Example: on a holiday
  - A excursion for 6 hours North on day 3.
  - A excursion for 6 hours South on day 3.

Complements and Substitutes

- Often additive independence is not a good assumption.
- Values x_1 of feature X_1 and x_2 of feature X_2 are complements if having both is better than the sum of the two.
- Values x_1 of feature X_1 and x_2 of feature X_2 are substitutes if having both is worse than the sum of the two.
- Example: on a holiday
  - A excursion for 6 hours North on day 3.
  - A excursion for 6 hours South on day 3.
- Example: on a holiday
  - A trip to a location 3 hours North on day 3
  - The return trip for the same day.

Generalized Additive Utility

- A generalized additive utility can be written as a sum of factors:

$$u(X_1, \ldots, X_n) = f_1(\overline{X_1}) + \cdots + f_k(\overline{X_k})$$

  where $\overline{X_i} \subseteq \{X_1, \ldots, X_n\}$.
- An intuitive canonical representation is difficult to find.
- It can represent complements and substitutes.

Utility and time

- Would you prefer \$1000 today or \$1000 next year?
- What price would you pay now to have an eternity of happiness?
- How can you trade off pleasures today with pleasures in the future?
- How would you compare the following sequences of rewards (per week):
  - \$1000000, \$0, \$0, \$0, \$0, \$0,...
  - \$1000, \$1000, \$1000, \$1000, \$1000,...
  - \$1000, \$0, \$0, \$0, \$0,...
  - \$1, \$1, \$1, \$1, \$1,...
  - \$1, \$2, \$3, \$4, \$5,...

Rewards and Values

Suppose the agent receives a sequence of rewards $r_1, r_2, r_3, r_4, \ldots$ in time. What utility should be assigned?

- total reward $V = \sum_{i=1}^{\infty} r_i$
- average reward $V = \lim_{n \to \infty} (r_1 + \cdots + r_n)/n$
- discounted reward $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
  γ is the discount factor $0 \le \gamma \le 1$.

Properties of the Discounted Reward

- The discounted value of rewards $r_1, r_2, r_3, r_4, \ldots$ is
$$\begin{aligned} V &= r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots \\ &= r_1 + \gamma(r_2 + \gamma(r_3 + \gamma(r_4 + \ldots))) \end{aligned}$$
- If $V(t)$ is the value obtained from time step t
$$V(t) = r_t + \gamma V(t+1)$$
- $1 + \gamma + \gamma^2 + \gamma^3 + \cdots = 1/(1-\gamma)$
Therefore $\dfrac{\text{minimum reward}}{1-\gamma} \leq V(t) \leq \dfrac{\text{maximum reward}}{1-\gamma}$
- We can approximate V with the first k terms, with error:
$$V - (r_1 + \gamma r_2 + \cdots + \gamma^{k-1} r_k) = \gamma^k V(k+1)$$