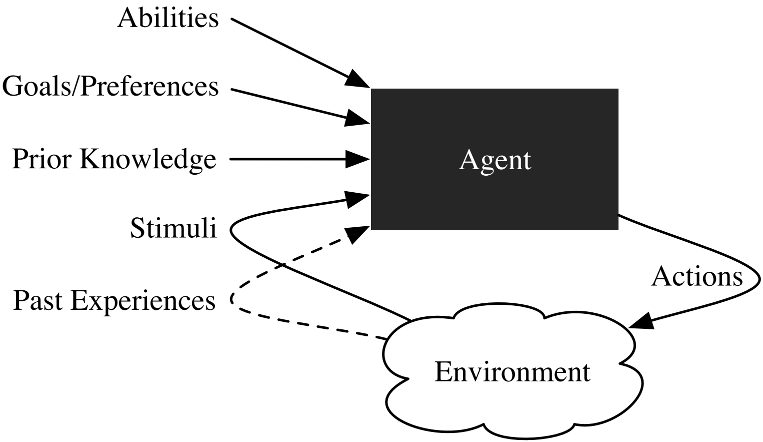
Agents acting in an environment: inputs and output



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

Learning Objectives

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

- justify the use and semantics of utility
- know the assumptions behind measures of preference
- estimate the utility of an outcome

Preferences

- Actions result in outcomes
- Agents have preferences over outcomes

Preferences

- Actions result in outcomes
- Agents have preferences over outcomes
- A rational agent will do the action that has the best outcome for them

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.

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 $o_2 \succeq o_1$.

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 its actions, but only have a probability distribution of the outcomes.

Lotteries

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

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

where the o_i are outcomes and $p_i \geq 0$ such that

$$\sum_i p_i = 1$$

The lottery specifies that outcome o_i occurs with probability p_i.

Lotteries

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

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

  where the o_i are outcomes and $p_i \geq 0$ 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:

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

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:

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

(Similarly for other mixtures of $\succ$ and $\succeq$.)

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:

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

(Similarly for other mixtures of $\succ$ and $\succeq$.)
Rationale: otherwise $o_1 \succeq o_2$ and $o_2 \succ o_3$ and $o_3 \succeq o_1$.
If they are prepared to pay to get o_2 instead of o_3,
and are happy to have o_1 instead of o_2,
and are happy to have o_3 instead of o_1
$\longrightarrow$ money pump.

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 $p \in [0, 1]$.
- 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:

$$[p : o_1, 1-p : [q : o_2, 1-q : o_3]]$$
$$\sim \quad [p : o_1, (1-p)q : o_2, (1-p)(1-q) : o_3]$$

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

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

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

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

What we would like

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

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

- 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:

$$utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i * utility(o_i)$$

Proof

- If all outcomes are equally preferred,

Proof

- If all outcomes are equally preferred, set $utility(o_i) = 0$ for all outcomes o_i.

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*.

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

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,

$$[u_1 : best, 1 - u_1 : worst]$$
$$\succeq$$

Proof (cont.)

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

$$[u_1 : best, 1 - u_1 : worst] \succeq [u_2 : best, 1 - u_2 : worst]$$

Which, by completeness and monotonicity implies

Proof (cont.)

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

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

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

Proof (cont.)

- To prove utilities are linear with probabilities

Proof (cont.)

- To prove utilities are linear with probabilities
- Suppose $u = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.

Proof (cont.)

- To prove utilities are linear with probabilities
- Suppose $u = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.
- Suppose $utility(o_i) = u_i$. We know:

Proof (cont.)

- To prove utilities are linear with probabilities
- Suppose $u = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.
- Suppose $utility(o_i) = u_i$. We know:

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

Proof (cont.)

- To prove utilities are linear with probabilities
- Suppose $u = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.
- 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

Proof (cont.)

- To prove utilities are linear with probabilities
- Suppose $u = utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k])$.
- 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

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

Proof (cont.)

- By decomposability, this is equivalent to:

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

- Thus, by definition of utility:

Proof (cont.)

- By decomposability, this is equivalent to:

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

- Thus, by definition of utility:

$$u = p_1 * u_1 + \cdots + p_k * u_k$$

Utilities

Two conditions of utility:

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

$$utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i * utility(o_i)$$

Utilities

Two conditions of utility:

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

$$utility([p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]) = \sum_{i=1}^{k} p_i * utility(o_i)$$

- Proved: probability of indifference satisfies the conditions.

Utilities

Two conditions of utility:

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

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

- Proved: probability of indifference satisfies the conditions.
- A (positive) linear function – multiplying by positive constant and/or adding a constant – of a utility function also satisfies the conditions.

Utilities

Two conditions of utility:

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

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

- Proved: probability of indifference satisfies the conditions.
- A (positive) linear function – multiplying by positive constant and/or adding a constant – of a utility function also satisfies the conditions.
- Often a different scale, such as $[0, 100]$, is used for utility.

Utilities

Two conditions of utility:

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

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

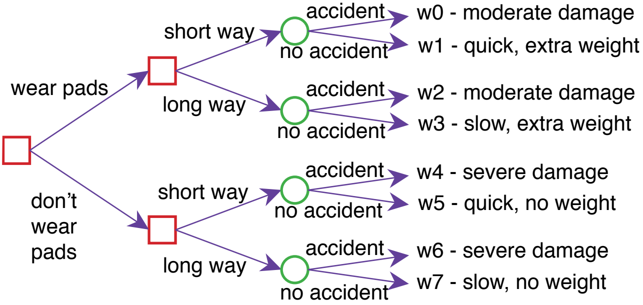
- Proved: probability of indifference satisfies the conditions.
- A (positive) linear function – multiplying by positive constant and/or adding a constant – of a utility function also satisfies the conditions.
- Often a different scale, such as $[0, 100]$, is used for utility.
- Sometimes negative values – costs – are used.

Delivery Robot Decision

- The robot can choose to wear pads to protect itself or not.
- The robot can choose to go the short way past the stairs or a long way that reduces the chance of an accident.
- There uncertainty about whether there will be an accident.

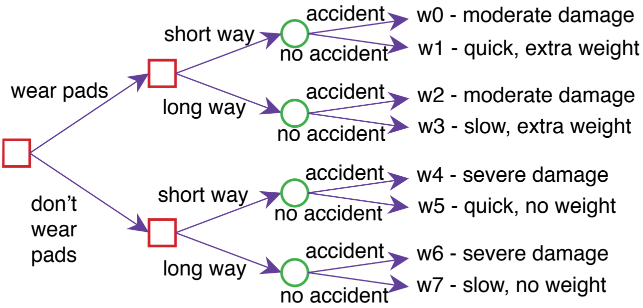
Delivery Robot Decision

- The robot can choose to wear pads to protect itself or not.
- The robot can choose to go the short way past the stairs or a long way that reduces the chance of an accident.
- There uncertainty about whether there will be an accident.



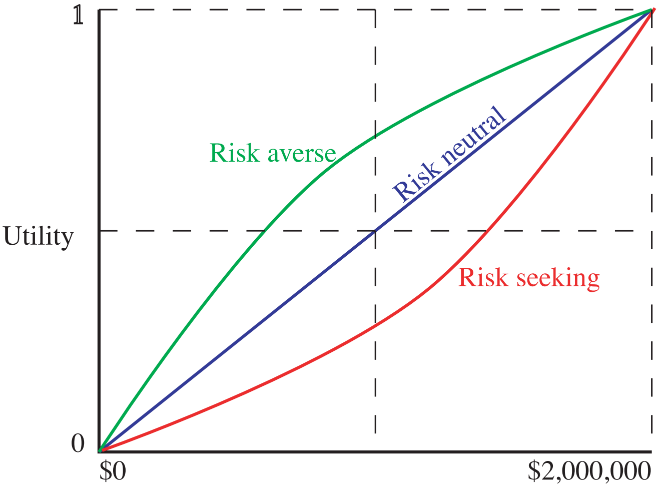
Delivery Robot Decision

- The robot can choose to wear pads to protect itself or not.
- The robot can choose to go the short way past the stairs or a long way that reduces the chance of an accident.
- There uncertainty about whether there will be an accident.



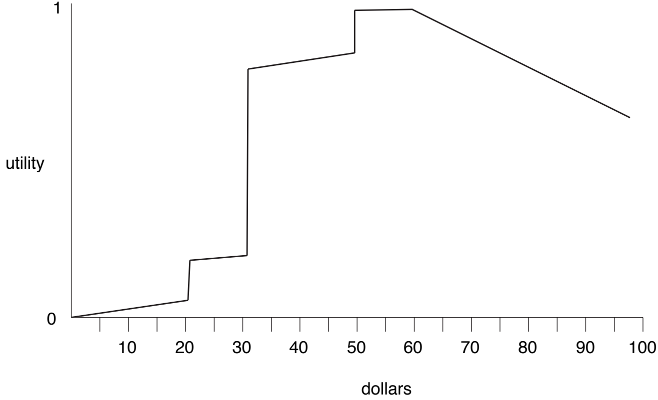
- What are reasonable utilities for the 8 outcomes $w0, \ldots, w7$? (suppose range $[0, 100]$)

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:



Factored Representation of Utility

- Under strong assumptions (see later), utility can be decomposed into a sum of factors:

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

This is called additive utility.

Factored Representation of Utility

- Under strong assumptions (see later), utility can be decomposed into a sum of factors:

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

  This is called additive utility.
- Many ways to represent the same utility:

Factored Representation of Utility

- Under strong assumptions (see later), utility can be decomposed into a sum of factors:

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

  This is called additive utility.
- 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:

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

- 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.

Additive Utility

- An additive utility has a canonical representation:

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

- 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:

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

- 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 $X_2, \ldots, X_n$.

Additive Utility

- An additive utility has a canonical representation:

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

- 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 $X_2, \ldots, X_n$.
- Assumption behind additive utility:

Additive Utility

- An additive utility has a canonical representation:

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

- 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 $X_2, \ldots, X_n$.
- Assumption behind additive utility: for all x_1, x_1',
  $u(x_1, x_2, \ldots, x_n) - u(x_1', x_2, \ldots, x_n)$ is the same for all
  $x_2, \ldots, x_n$, and similarly for other positions.

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
  - An excursion for 6 hours North on day 3.
  - An 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
  - An excursion for 6 hours North on day 3.
  - An 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.

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.
- The canonical representation of utility allows weights for conjunctions of feature values. For Boolean $\{0, 1\}$ features:

$$
\begin{aligned}
u(x_1, \ldots, x_n) = {} & w_0 + w_1 * x_1 + w_2 * x_2 \cdots + w_n * x_n \\
& + w_{12} * x_1 * x_2 + w_{13} * x_1 * x_3 + \ldots \\
& + w_{123} * x_1 * x_2 * x_3 + \ldots \\
& \ldots
\end{aligned}
$$

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.
- The canonical representation of utility allows weights for conjunctions of feature values. For Boolean $\{0, 1\}$ features:

$$
\begin{aligned}
u(x_1, \ldots, x_n) = {} & w_0 + w_1 * x_1 + w_2 * x_2 \cdots + w_n * x_n \\
& + w_{12} * x_1 * x_2 + w_{13} * x_1 * x_3 + \ldots \\
& + w_{123} * x_1 * x_2 * x_3 + \ldots \\
& \ldots
\end{aligned}
$$

- 2^n weights can represent any utility on n Boolean features. Most weights can be 0 (and omitted).

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.
- The canonical representation of utility allows weights for conjunctions of feature values. For Boolean $\{0,1\}$ features:

$$
\begin{aligned}
u(x_1,\dots,x_n) = {} & w_0 + w_1 * x_1 + w_2 * x_2 \cdots + w_n * x_n \\
& + w_{12} * x_1 * x_2 + w_{13} * x_1 * x_3 + \dots \\
& + w_{123} * x_1 * x_2 * x_3 + \dots \\
& \dots
\end{aligned}
$$

- 2^n weights can represent any utility on n Boolean features. Most weights can be 0 (and omitted).
- x_i and x_j are complements iff

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.
- The canonical representation of utility allows weights for conjunctions of feature values. For Boolean $\{0,1\}$ features:

$$
\begin{aligned}
u(x_1,\ldots,x_n) = {} & w_0 + w_1 * x_1 + w_2 * x_2 \cdots + w_n * x_n \\
& + w_{12} * x_1 * x_2 + w_{13} * x_1 * x_3 + \ldots \\
& + w_{123} * x_1 * x_2 * x_3 + \ldots \\
& \ldots
\end{aligned}
$$

- 2^n weights can represent any utility on n Boolean features. Most weights can be 0 (and omitted).
- x_i and x_j are complements iff $w_{ij} > 0$

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.
- The canonical representation of utility allows weights for conjunctions of feature values. For Boolean $\{0, 1\}$ features:

$$
\begin{aligned}
u(x_1, \ldots, x_n) = {} & w_0 + w_1 * x_1 + w_2 * x_2 \cdots + w_n * x_n \\
& + w_{12} * x_1 * x_2 + w_{13} * x_1 * x_3 + \ldots \\
& + w_{123} * x_1 * x_2 * x_3 + \ldots \\
& \ldots
\end{aligned}
$$

- 2^n weights can represent any utility on n Boolean features. Most weights can be 0 (and omitted).
- x_i and x_j are complements iff $w_{ij} > 0$
- x_i and x_j are substitutes iff

Canonical Representation of Utility

- Generalized additive utility defines utility in terms of sum of factors, where a factor is a function on some of the variables.
- The canonical representation of utility allows weights for conjunctions of feature values. For Boolean $\{0, 1\}$ features:

$$\begin{aligned} u(x_1, \ldots, x_n) = & \ w_0 + w_1 * x_1 + w_2 * x_2 \cdots + w_n * x_n \\ & + w_{12} * x_1 * x_2 + w_{13} * x_1 * x_3 + \ldots \\ & + w_{123} * x_1 * x_2 * x_3 + \ldots \\ & \ldots \end{aligned}$$

- 2^n weights can represent any utility on n Boolean features. Most weights can be 0 (and omitted).
- x_i and x_j are complements iff $w_{ij} > 0$
- x_i and x_j are substitutes iff $w_{ij} < 0$

Utility and time

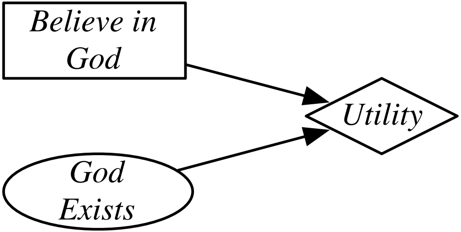
- Would you prefer \$1000 today or \$1000 next year?
- What price would you pay now to have an eternity of happiness?
- How can you trade off pleasures today with pleasures in the future?

Pascal's Wager (1670)

Decide whether to believe in God.

Pascal's Wager (1670)

Decide whether to believe in God.



Utility and time

- How would you compare the following sequences of rewards (per week):

  A: \$1000000, \$0, \$0, \$0, \$0, \$0,...
  B: \$1000, \$1000, \$1000, \$1000, \$1000,...
  C: \$1000, \$0, \$0, \$0, \$0,...
  D: \$1, \$1, \$1, \$1, \$1,...
  E: \$1, \$2, \$3, \$4, \$5,...

Rewards and Values

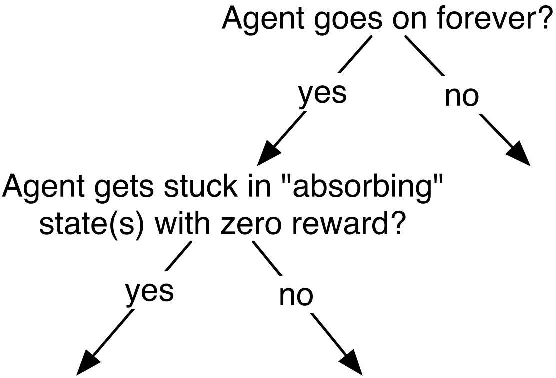
Suppose the agent receives a sequence of rewards $r_1, r_2, r_3, r_4, \ldots$ in time. What utility should be assigned? "Return" or "value"

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? "Return" or "value"

- total reward $V = \sum_{i=1}^{\infty} r_i$
- average reward $V = \lim_{n\to\infty} (r_1 + \cdots + r_n)/n$

Average vs Accumulated Rewards



Rewards and Values

Suppose the agent receives a sequence of rewards $r_1, r_2, r_3, r_4, \ldots$ in time.

- discounted return $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
  γ is the discount factor $0 \leq \gamma \leq 1$.

Properties of the Discounted Rewards

- The discounted return for rewards $r_1, r_2, r_3, r_4, \ldots$ is

$$\begin{aligned} V &= r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots \\ &= \end{aligned}$$

Properties of the Discounted Rewards

- The discounted return for 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

$$V_t =$$

Properties of the Discounted Rewards

- The discounted return for 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

$$V_t = r_t + \gamma V_{t+1}$$

Properties of the Discounted Rewards

- The discounted return for 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

$$V_t = r_t + \gamma V_{t+1}$$

- How is the infinite future valued compared to immediate rewards?

Properties of the Discounted Rewards

- The discounted return for 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
$$V_t = r_t + \gamma V_{t+1}$$
- How is the infinite future valued compared to immediate rewards?
$1 + \gamma + \gamma^2 + \gamma^3 + \cdots =$

Properties of the Discounted Rewards

- The discounted return for 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

$$V_t = r_t + \gamma V_{t+1}$$

- How is the infinite future valued compared to immediate rewards?
  $1 + \gamma + \gamma^2 + \gamma^3 + \cdots = 1/(1-\gamma)$
  Therefore $\dfrac{\text{minimum reward}}{1-\gamma} \leq V_t \leq \dfrac{\text{maximum reward}}{1-\gamma}$

Properties of the Discounted Rewards

- The discounted return for 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
$$V_t = r_t + \gamma V_{t+1}$$
- How is the infinite future valued compared to immediate rewards?
$1 + \gamma + \gamma^2 + \gamma^3 + \cdots = 1/(1-\gamma)$
Therefore $\dfrac{\text{minimum reward}}{1-\gamma} \leq V_t \leq \dfrac{\text{maximum reward}}{1-\gamma}$
- You can approximate V with the first k terms, with error:
$$\begin{aligned} V - (r_1 + \gamma r_2 + \cdots + \gamma^{k-1} r_k) &= \gamma^k V_{k+1} \\ &\propto \gamma^k/(1-\gamma) \end{aligned}$$

Properties of the Discounted Rewards

- $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
- At each time:
  - with probability γ, agent keeps going
  - otherwise agent stops

  with return is

Properties of the Discounted Rewards

- $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
- At each time:
  - with probability γ, agent keeps going
  - otherwise agent stops

  with return is total reward is equivalent to discounting.

Properties of the Discounted Rewards

- $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
- At each time:
  - with probability γ, agent keeps going
  - otherwise agent stops

  with return is total reward is equivalent to discounting.
- With an interest rate of i, a dollar now is worth $1 + i$ in a year. So a dollar in a year is worth

Properties of the Discounted Rewards

- $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
- At each time:
  - with probability γ, agent keeps going
  - otherwise agent stops

  with return is total reward is equivalent to discounting.
- With an interest rate of i, a dollar now is worth $1 + i$ in a year. So a dollar in a year is worth $1/(1 + i)$ now. γ can be seen as $1/(1 + i)$ where i is interest rate.

Properties of the Discounted Rewards

- $V = r_1 + \gamma r_2 + \gamma^2 r_3 + \gamma^3 r_4 + \cdots$
- At each time:
  - with probability γ, agent keeps going
  - otherwise agent stops

  with return is total reward is equivalent to discounting.
- With an interest rate of i, a dollar now is worth $1 + i$ in a year. So a dollar in a year is worth $1/(1 + i)$ now. γ can be seen as $1/(1 + i)$ where i is interest rate.
- γ should reflect an agent's utility.

Why discounting? [Koopmans 1972]

With an infinite sequence of outcomes $\langle o_1, o_2, o_3, \ldots \rangle$ if

- the first time period matters, so $\exists\ o_1, o_2, o_3, \ldots$ and o_1' where

$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1', o_2, o_3, \ldots \rangle$$

Why discounting? [Koopmans 1972]

With an infinite sequence of outcomes $\langle o_1, o_2, o_3, \ldots \rangle$ if

- the first time period matters, so $\exists\ o_1, o_2, o_3, \ldots$ and o_1' where

$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1', o_2, o_3, \ldots \rangle$$

- preferences on first two times do not depend on the future:

$$\langle x_1, x_2, o_3, o_4 \ldots \rangle \succ \langle y_1, y_2, o_3, o_4 \ldots \rangle$$

if and only if $\langle x_1, x_2, o_3', o_4' \ldots \rangle \succ \langle y_1, y_2, o_3', o_4' \ldots \rangle$

Why discounting? [Koopmans 1972]

With an infinite sequence of outcomes $\langle o_1, o_2, o_3, \ldots \rangle$ if

- the first time period matters, so $\exists\ o_1, o_2, o_3, \ldots$ and o_1' where

$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1', o_2, o_3, \ldots \rangle$$

- preferences on first two times do not depend on the future:

$$\langle x_1, x_2, o_3, o_4 \ldots \rangle \succ \langle y_1, y_2, o_3, o_4 \ldots \rangle$$

if and only if $\langle x_1, x_2, o_3', o_4' \ldots \rangle \succ \langle y_1, y_2, o_3', o_4' \ldots \rangle$

- stationarity:

$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1, o_2', o_3', \ldots \rangle$$

if and only if $\langle o_2, o_3, \ldots \rangle \succ \langle o_2', o_3', \ldots \rangle$

Why discounting? [Koopmans 1972]

With an infinite sequence of outcomes $\langle o_1, o_2, o_3, \ldots \rangle$ if

- the first time period matters, so $\exists\ o_1, o_2, o_3, \ldots$ and o_1' where
$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1', o_2, o_3, \ldots \rangle$$
- preferences on first two times do not depend on the future:
$$\langle x_1, x_2, o_3, o_4 \ldots \rangle \succ \langle y_1, y_2, o_3, o_4 \ldots \rangle$$
if and only if $\langle x_1, x_2, o_3', o_4' \ldots \rangle \succ \langle y_1, y_2, o_3', o_4' \ldots \rangle$
- stationarity:
$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1, o_2', o_3', \ldots \rangle$$
if and only if $\langle o_2, o_3, \ldots \rangle \succ \langle o_2', o_3', \ldots \rangle$
- the agent only cares about finite subspaces of infinite time

Why discounting? [Koopmans 1972]

With an infinite sequence of outcomes $\langle o_1, o_2, o_3, \ldots \rangle$ if

- the first time period matters, so $\exists\ o_1, o_2, o_3, \ldots$ and o_1' where

$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1', o_2, o_3, \ldots \rangle$$

- preferences on first two times do not depend on the future:

$$\langle x_1, x_2, o_3, o_4 \ldots \rangle \succ \langle y_1, y_2, o_3, o_4 \ldots \rangle$$
$$\text{if and only if } \langle x_1, x_2, o_3', o_4' \ldots \rangle \succ \langle y_1, y_2, o_3', o_4' \ldots \rangle$$

- stationarity:

$$\langle o_1, o_2, o_3, \ldots \rangle \succ \langle o_1, o_2', o_3', \ldots \rangle$$
$$\text{if and only if } \langle o_2, o_3, \ldots \rangle \succ \langle o_2', o_3', \ldots \rangle$$

- the agent only cares about finite subspaces of infinite time

then there exists a discount factor γ and function r such that

$$utility(\langle o_1, o_2, o_3, \ldots \rangle) = \sum_i \gamma^{i-1} r(o_i)$$

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

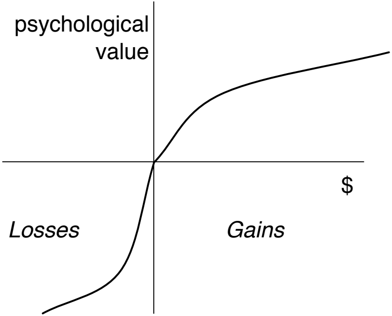
D: lottery $[0.10 : \$2.5m, 0.9 : \$0]$

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

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

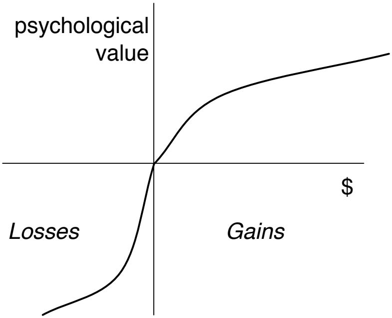
B,D: lottery $[0.10 : \$2.5m, 0.01 : \$0, 0.89 : X]$

Prospect Theory



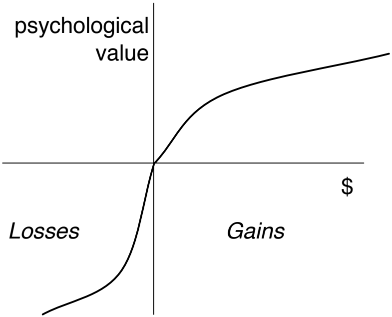
- Preferences depend on the agent's reference point: current wealth.

Prospect Theory



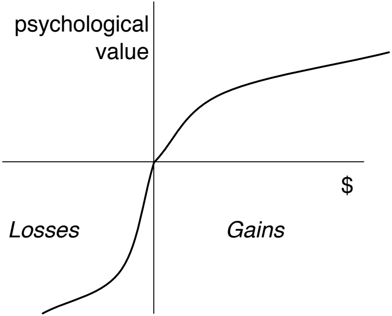
- Preferences depend on the agent's reference point: current wealth.
- For gains,

Prospect Theory



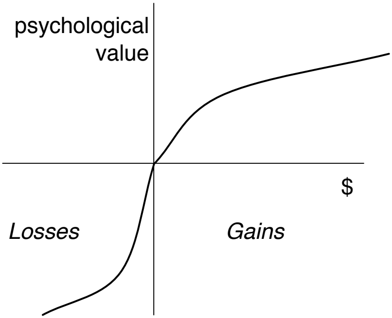
- Preferences depend on the agent's reference point: current wealth.
- For gains, they are risk averse.

Prospect Theory



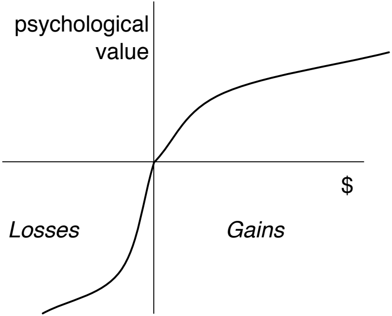
- Preferences depend on the agent's reference point: current wealth.
- For gains, they are risk averse.
- For losses,

Prospect Theory



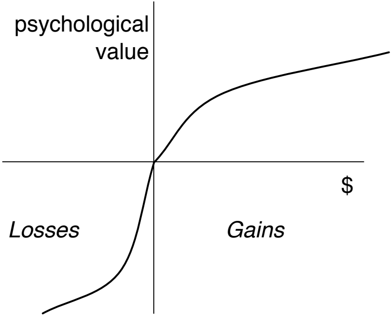
- Preferences depend on the agent's reference point: current wealth.
- For gains, they are risk averse.
- For losses, they are risk seeking

Prospect Theory



- Preferences depend on the agent's reference point: current wealth.
- For gains, they are risk averse.
- For losses, they are risk seeking
- Losses are (about) twice as bad as gains.

Prospect Theory



- Preferences depend on the agent's reference point: current wealth.
- For gains, they are risk averse.
- For losses, they are risk seeking
- Losses are (about) twice as bad as gains.

This better fits with human preferences.

Reference Points

Consider Anthony and Betty who (for argument) are essentially the same except:

- Anthony's current wealth is \$1 million.
- Betty's current wealth is \$4 million.

They are both offered the choice between a gamble and a sure thing:

- Gamble: equal chance to end up owning \$1 million or \$4 million.
- Sure Thing: own \$2 million

What does expected utility theory predict?

Reference Points

Consider Anthony and Betty who (for argument) are essentially the same except:

- Anthony's current wealth is $1 million.
- Betty's current wealth is $4 million.

They are both offered the choice between a gamble and a sure thing:

- Gamble: equal chance to end up owning $1 million or $4 million.
- Sure Thing: own $2 million

What does expected utility theory predict? They make same choice as preference only depends on the outcomes.

Reference Points

Consider Anthony and Betty who (for argument) are essentially the same except:

- Anthony's current wealth is \$1 million.
- Betty's current wealth is \$4 million.

They are both offered the choice between a gamble and a sure thing:

- Gamble: equal chance to end up owning \$1 million or \$4 million.
- Sure Thing: own \$2 million

What does expected utility theory predict? They make same choice as preference only depends on the outcomes.

What does prospect theory predict?

Reference Points

Consider Anthony and Betty who (for argument) are essentially the same except:

- Anthony's current wealth is \$1 million.
- Betty's current wealth is \$4 million.

They are both offered the choice between a gamble and a sure thing:

- Gamble: equal chance to end up owning \$1 million or \$4 million.
- Sure Thing: own \$2 million

What does expected utility theory predict? They make same choice as preference only depends on the outcomes.

What does prospect theory predict?
Anthony is making a gain so will will be risk averse and take the sure thing.
Better is making a loss and so will will be risk seeking and gamble.

Reference Points [Kahneman 2011]

Twins Andy and Bobbie, have identical tastes and identical starting jobs. There are two jobs that are identical, except that

- job A gives a raise of \$10000
- job B gives an extra day of vacation per month.

They are each indifferent to the outcomes and toss a coin. Andy takes job A, and Bobbie takes job B.
Now the company suggests they swap jobs with a \$500 bonus.
Will they swap?

Reference Points [Kahneman 2011]

Twins Andy and Bobbie, have identical tastes and identical starting jobs. There are two jobs that are identical, except that

- job A gives a raise of $10000
- job B gives an extra day of vacation per month.

They are each indifferent to the outcomes and toss a coin. Andy takes job A, and Bobbie takes job B.

Now the company suggests they swap jobs with a $500 bonus.

Will they swap?

What does utility theory predict?

Reference Points [Kahneman 2011]

Twins Andy and Bobbie, have identical tastes and identical starting jobs. There are two jobs that are identical, except that

- job A gives a raise of \$10000
- job B gives an extra day of vacation per month.

They are each indifferent to the outcomes and toss a coin. Andy takes job A, and Bobbie takes job B.

Now the company suggests they swap jobs with a \$500 bonus.

Will they swap?

What does utility theory predict?

What does prospect theory predict?

Reference Points [Kahneman 2011]

Twins Andy and Bobbie, have identical tastes and identical starting jobs. There are two jobs that are identical, except that

- job *A* gives a raise of $10000
- job *B* gives an extra day of vacation per month.

They are each indifferent to the outcomes and toss a coin. Andy takes job A, and Bobbie takes job B.
Now the company suggests they swap jobs with a $500 bonus.
Will they swap?
What does utility theory predict?
What does prospect theory predict?
Utility theory predicts they swap. Prospect theory predicts they do not swap.
[From D. Kahneman, *Thinking, Fast and Slow*, 2011, p. 291.]

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

What do you think of Alan and Ben:

- Alan: intelligent—industrious—impulsive—critical—stubborn—envious

Framing Effects

What do you think of Alan and Ben:

- Ben: envious—stubborn—critical—impulsive—industrious—intelligent

Framing Effects

What do you think of Alan and Ben:

- Alan: intelligent—industrious—impulsive—critical—stubborn—envious
- Ben: envious—stubborn—critical—impulsive—industrious—intelligent

[From D. Kahneman, Thinking Fast and Slow, 2011, p. 82]

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?

[From R.M. Dawes, Rational Choice in an Uncertain World, 1988.]

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 utilities are unbounded.
- Then for any outcome o_i there is an outcome o_{i+1} such that $u(o_{i+1}) > 2u(o_i)$.

St. Petersburg Paradox

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

- Suppose utilities are unbounded.
- Then for any outcome o_i there is an outcome o_{i+1} such that $u(o_{i+1}) > 2u(o_i)$.
- Would the agent prefer o_1 or the lottery $[0.5 : o_2, 0.5 : 0]$ where 0 is the worst outcome?

St. Petersburg Paradox

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

- Suppose utilities are unbounded.
- Then for any outcome o_i there is an outcome o_{i+1} such that $u(o_{i+1}) > 2u(o_i)$.
- Would the agent prefer o_1 or the lottery $[0.5 : o_2, 0.5 : 0]$ where 0 is the worst outcome?
- Is it rational to gamble o_1 to on a coin toss to get o_2?
- Is it rational to gamble o_2 to on a coin toss to get o_3?
- Is it 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 utilities are unbounded.
- Then for any outcome o_i there is an outcome o_{i+1} such that $u(o_{i+1}) > 2u(o_i)$.
- Would the agent prefer o_1 or the lottery $[0.5 : o_2, 0.5 : 0]$ where 0 is the worst outcome?
- Is it rational to gamble o_1 to on a coin toss to get o_2?
- Is it rational to gamble o_2 to on a coin toss to get o_3?
- Is it rational to gamble o_3 to on a coin toss to get o_4?
- What will eventually happen?

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?