Puzzles and Paradoxes from Decision and Game Theory

Eric Pacuit
University of Maryland
pacuit.org

July 17, 2017

Information

Email:	epacuit@umd.edu
Website:	pacuit.org/esslli2017/paradoxes_games_dec/
Reading:	See website
Game Theory:	www.game-theory-class.org

Plan

- Day 1: Rational Choice Theory, Expected Utility
- Day 2: Decision-Theoretic Paradoxes, Absent-Minded Driver
- Day 3: Game Theory
- Day 4: Common Knowledge, Backward Induction and Epistemic Game Theory
- Day 5: Paradoxes of Interactive Epistemology, Framing in Games and Decisions

Topics

- Paradoxes of expected utility: St. Petersberg paradox, Pasadena game, The Two-envelop paradox
- Allais and Ellsberg paradox
- Newcomb's paradox and the psychopath button problem
- Puzzling games: the Prisoner's Dilemma and the Traveler's Dilemma
- The absent-minded driver problem
- Rubinstein's email game and the general's problem
- Backward induction and common knowledge of rationality
- The Brandenburger-Keisler paradox
- Framing in decision and game theory: language-dependent decisions and games, coordination problems and the theory of focal points.

Simple Choice Model

Menu

Simple Choice Model

Choice

Simple Choice Model

Rational Choice?

Simple Choice Model

Rational Choice?

Preference

 >  > 

Simple Choice Model

Rational Choice

Preference

 >  > 

Simple Choice Model

Irrational Choice

Preference

- **Option uncertainty**: What type of wine is it? Is the red wine sweet or dry? Is the white wine spoiled? Is the lemonade very sugary? …

- **Option uncertainty**: What type of wine is it? Is the red wine sweet or dry? Is the white wine spoiled? Is the lemonade very sugary? …

  [Options vs. **Prospects**]

- **Option uncertainty**: What type of wine is it? Is the red wine sweet or dry? Is the white wine spoiled? Is the lemonade very sugary? …

  [Options vs. **Prospects**]

- **Menu uncertainty**: Are there other drink choices that are available (e.g., a beer or a soda)? …

- **Option uncertainty**: What type of wine is it? Is the red wine sweet or dry? Is the white wine spoiled? Is the lemonade very sugary? …

  [Options vs. **Prospects**]

- **Menu uncertainty**: Are there other drink choices that are available (e.g., a beer or a soda)? …

- **Context**: What are we having to eat? What time of day is it? How many drinks have you had? Are you driving home? …

Decision Problems

Decision Problems

Individual decision-making (against nature)

- E.g., Gambling

Decision Problems

Individual decision-making (against nature)

- E.g., Gambling

Individual decision making in interaction

- E.g., Playing chess

Decision Problems

Individual decision-making (against nature)

- E.g., Gambling

Individual decision making in interaction

- E.g., Playing chess

Collective decision making

- E.g., Carrying a piano

Decision Problems

Individual decision-making (against nature)

- E.g., Gambling

Individual decision making in interaction

- E.g., Playing chess

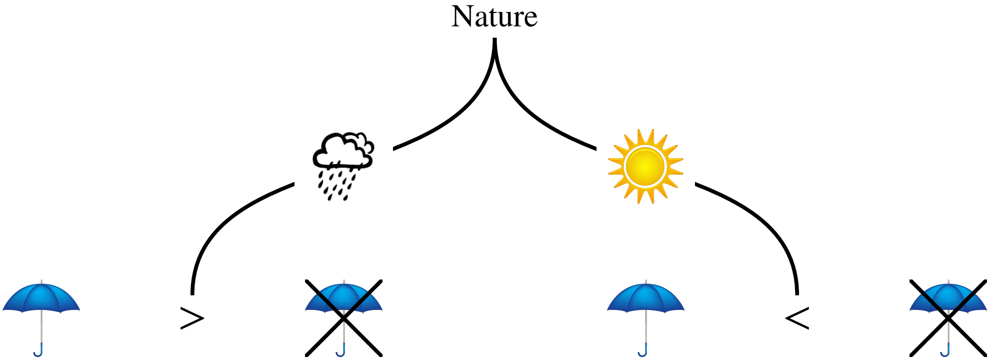
Collective decision making

- E.g., Carrying a piano
- E.g., Voting in an election

Individual decision-making (against nature)

🌂 > ~~🌂~~

🌂 < ~~🌂~~


Nature
>
<

		
	encumbered, dry	encumbered, dry
	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

		
	encumbered, dry	encumbered, dry
	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

		
	encumbered, dry	encumbered, dry
	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

		
	encumbered, dry	encumbered, dry
	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

		
	encumbered, dry	encumbered, dry
	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

		
	encumbered, dry	encumbered, dry
	wet	free, dry

States: it rains; it does not rain

Outcomes: encumbered, dry; wet; free, dry

Actions: take umbrella; leave umbrella

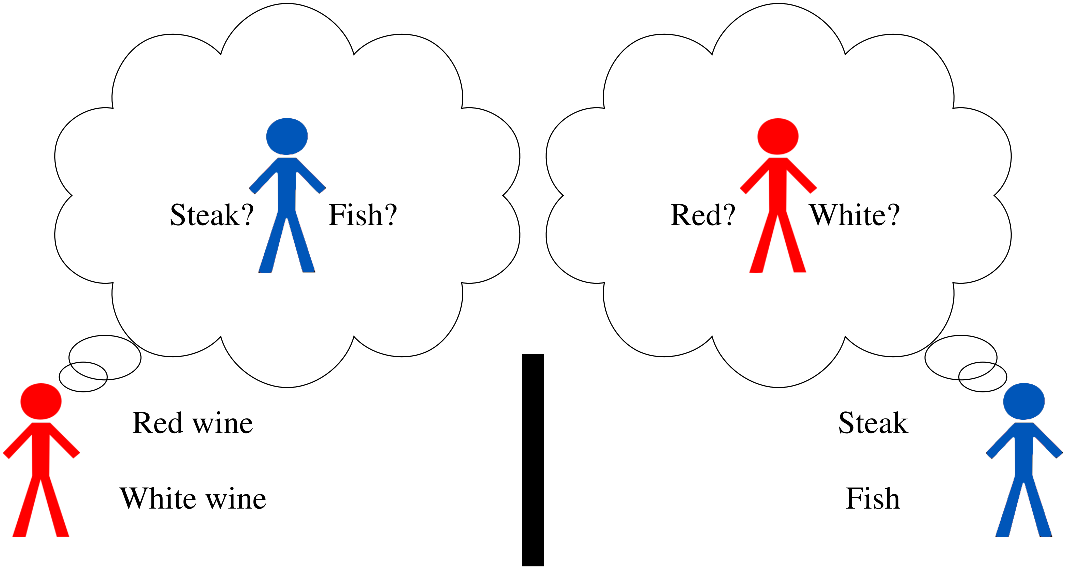
Decision making in interaction

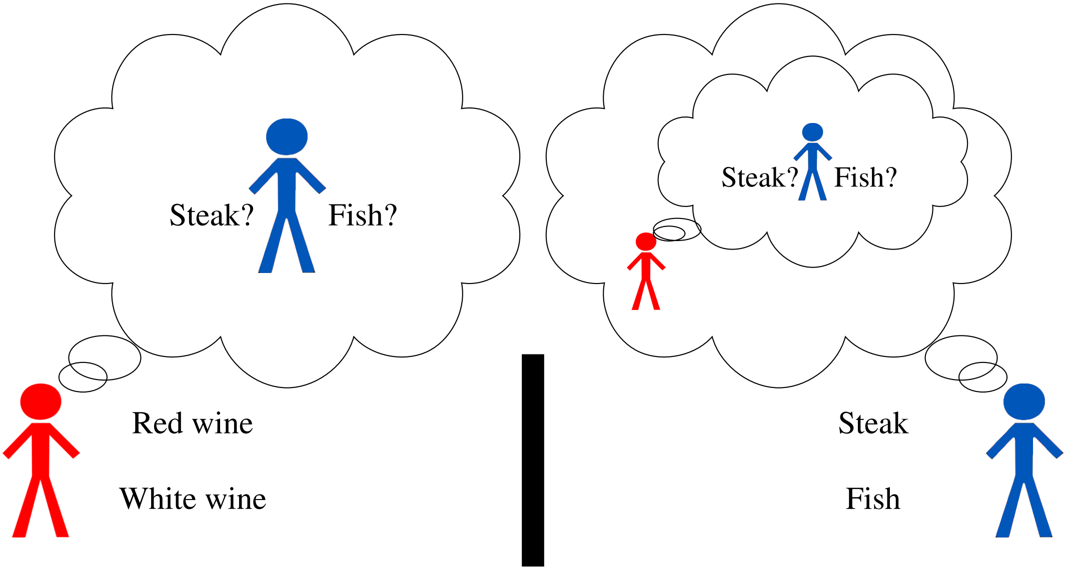
Red wine

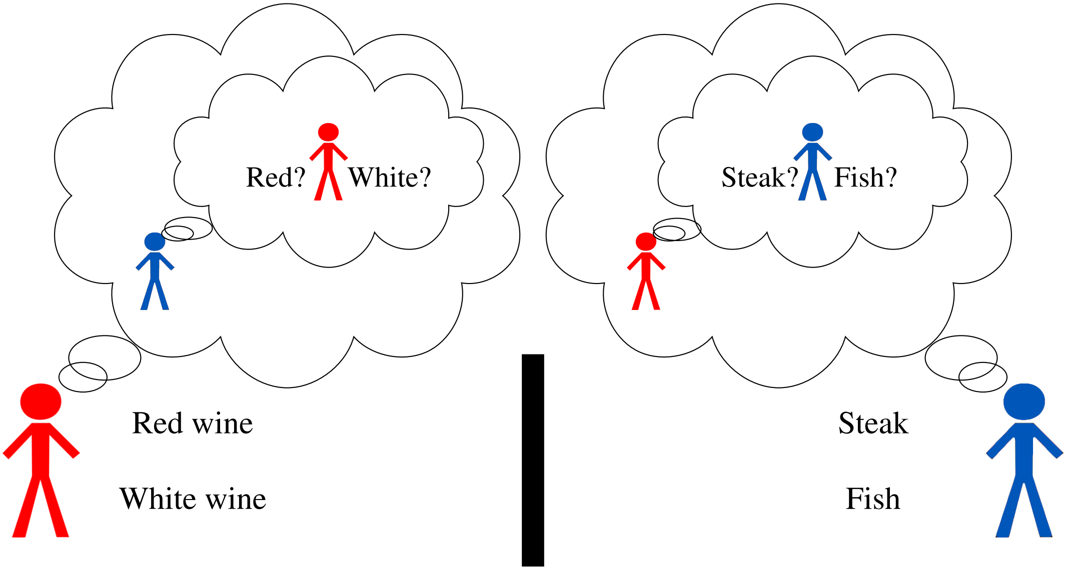
White wine

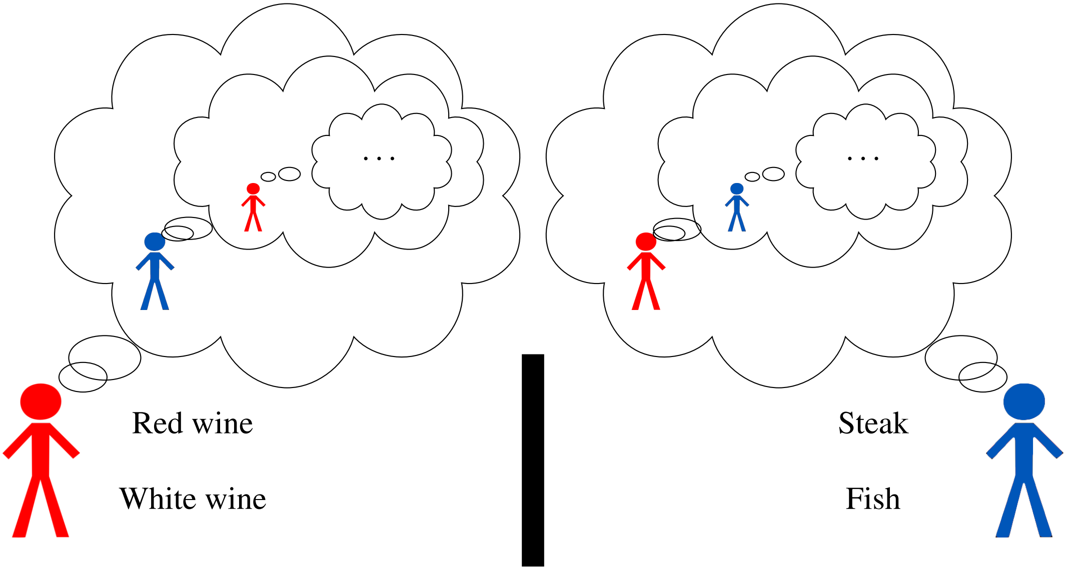
Steak

Fish


Steak?
Fish?
Red?
White?
Red wine
White wine
Steak
Fish


Steak?
Fish?
Red wine
White wine
Steak?
Fish?
Steak
Fish


Red? White?
Steak? Fish?
Red wine
White wine
Steak
Fish


. . .
. . .
Red wine
White wine
Steak
Fish

Preferences

Preferring or choosing x is different that “liking” x or “having a taste for x”: one can prefer x to y but *dislike* both options

Preferences are always understood as comparative: “preference” is more like “bigger” than “big”

Concepts of *preference*

1. *Enjoyment comparison*: I prefer red wine to white wine means that I *enjoy* red wine more than white wine

Concepts of *preference*

1. *Enjoyment comparison*: I prefer red wine to white wine means that I *enjoy* red wine more than white wine

2. *Comparative evaluation*: I prefer candidate *A* over candidate *B* means "I judge *A* to be *superior* to *B*". This can be *partial* (ranking with respect to some criterion) or *total* (with respect to every relevant consideration).

Concepts of *preference*

1. *Enjoyment comparison*: I prefer red wine to white wine means that I *enjoy* red wine more than white wine

2. *Comparative evaluation*: I prefer candidate *A* over candidate *B* means "I judge *A* to be *superior* to *B*". This can be *partial* (ranking with respect to some criterion) or *total* (with respect to every relevant consideration).

3. *Favoring*: Affirmative action calls for racial/gender preferences in hiring.

Concepts of *preference*

1. *Enjoyment comparison*: I prefer red wine to white wine means that I *enjoy* red wine more than white wine

2. *Comparative evaluation*: I prefer candidate *A* over candidate *B* means "I judge *A* to be *superior* to *B*". This can be *partial* (ranking with respect to some criterion) or *total* (with respect to every relevant consideration).

3. *Favoring*: Affirmative action calls for racial/gender preferences in hiring.

4. *Choice ranking*: In a restaurant, when asked "do you prefer red wine or white wine", the waiter wants to know which option I choose.

Concepts of *preference*

1. *Enjoyment comparison*: I prefer red wine to white wine means that I *enjoy* red wine more than white wine

2. *Comparative evaluation*: I prefer candidate *A* over candidate *B* means “I judge *A* to be *superior* to *B*”. This can be *partial* (ranking with respect to some criterion) or *total* (with respect to every relevant consideration).

3. *Favoring*: Affirmative action calls for racial/gender preferences in hiring.

4. *Choice ranking*: In a restaurant, when asked “do you prefer red wine or white wine”, the waiter wants to know which option I choose.

Preferences

Preferences are understood as *mental rankings* of alternatives “all things considered”.

Preferences

Preferences are understood as *mental rankings* of alternatives “all things considered”.

What exactly constitutes an “all things considered preference”?

Preferences

Preferences are understood as *mental rankings* of alternatives “all things considered”.

What exactly constitutes an “all things considered preference”?

1. Lauren drank water rather than wine with dinner, despite preferring to drink wine, because she promised her husband she would stay sober.

2. Lauren drank water with dinner because she preferred to do so. But for the promise she made her husband to stay sober, she would have preferred to drink wine rather than water with dinner.

Invoking someone's preferences will suffice to explain why some choices were not made (i.e. in terms of rational impermissibility) but not typically why some particular choice was made. To take up the slack, explanations must draw on factors other than preference: psychological one such as the framing of the choice problem or the saliency of particular options, or sociological ones such as the existence of norms or conventions governing choices of the relevant kind.

Mathematically describing preferences

Mathematical background: Relations

Suppose that X is a set. A **relation** on X is a set of **ordered pairs** from X: $R \subseteq X \times X$.

Mathematical background: Relations

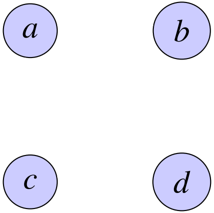
Suppose that X is a set. A **relation** on X is a set of **ordered pairs** from X: $R \subseteq X \times X$.

E.g., $X = \{a, b, c, d\}$, $R = \{(a, a), (b, a), (c, d), (a, c), (d, d)\}$

Mathematical background: Relations

Suppose that X is a set. A **relation** on X is a set of **ordered pairs** from X: $R \subseteq X \times X$.

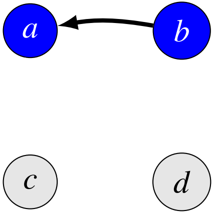
E.g., $X = \{a, b, c, d\}$, $R = \{(a, a), (b, a), (c, d), (a, c), (d, d)\}$



Mathematical background: Relations

Suppose that X is a set. A **relation** on X is a set of **ordered pairs** from X: $R \subseteq X \times X$.

E.g., $X = \{a, b, c, d\}$, $R = \{(a, a), (b, a), (c, d), (a, c), (d, d)\}$

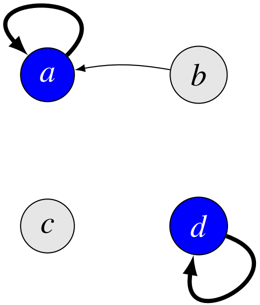


$b \; R \; a$

Mathematical background: Relations

Suppose that X is a set. A **relation** on X is a set of **ordered pairs** from X: $R \subseteq X \times X$.

E.g., $X = \{a, b, c, d\}$, $R = \{(a, a), (b, a), (c, d), (a, c), (d, d)\}$



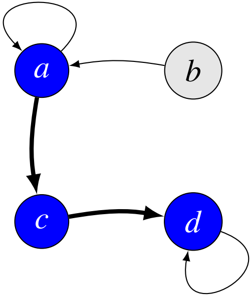
$a\ R\ a$
$b\ R\ a$

$d\ R\ d$

Mathematical background: Relations

Suppose that X is a set. A **relation** on X is a set of **ordered pairs** from X: $R \subseteq X \times X$.

E.g., $X = \{a, b, c, d\}$, $R = \{(a, a), (b, a), (c, d), (a, c), (d, d)\}$



$a \; R \; a$
$b \; R \; a$
$c \; R \; d$
$a \; R \; c$
$d \; R \; d$

Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

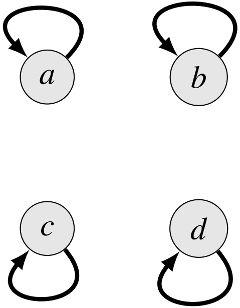
Reflexive relation: for all $x \in X$, $x \; R \; x$

Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Reflexive relation: for all $x \in X$, $x \; R \; x$

E.g., $X = \{a, b, c, d\}$



Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

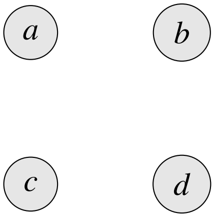
Irreflexive relation: for all $x \in X$, $x \not R\, x$ (i.e., $(x, x) \notin R$)

Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Irreflexive relation: for all $x \in X$, $x \not R\, x$ (i.e., $(x, x) \notin R$)

E.g., $X = \{a, b, c, d\}$



Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

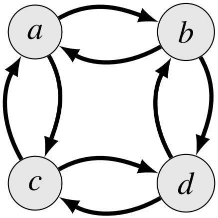
Symmetric relation: for all $x, y \in X$, if $x \; R \; y$, then $y \; R \; x$

Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Symmetric relation: for all $x, y \in X$, if $x \; R \; y$, then $y \; R \; x$

E.g., $X = \{a, b, c, d\}$



Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

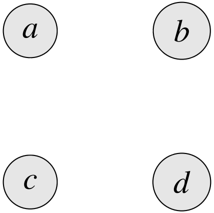
Complete relation: for all $x, y \in X$, either $x \; R \; y$ or $y \; R \; x$

Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Complete relation: for all $x, y \in X$, either $x \; R \; y$ or $y \; R \; x$

E.g., $X = \{a, b, c, d\}$

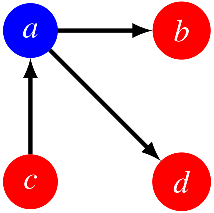


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Complete relation: for all $x, y \in X$, either $x\ R\ y$ or $y\ R\ x$

E.g., $X = \{a, b, c, d\}$

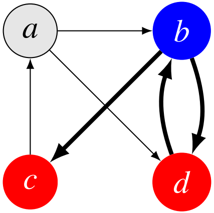


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Complete relation: for all $x, y \in X$, either $x\ R\ y$ or $y\ R\ x$

E.g., $X = \{a, b, c, d\}$

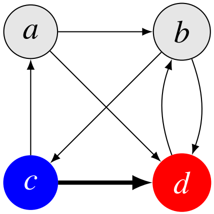


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Complete relation: for all $x, y \in X$, either $x \; R \; y$ or $y \; R \; x$

E.g., $X = \{a, b, c, d\}$

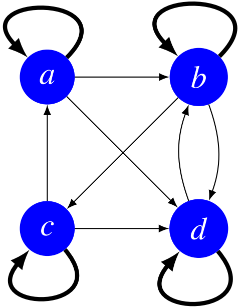


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Complete relation: for all $x, y \in X$, either $x \; R \; y$ or $y \; R \; x$

E.g., $X = \{a, b, c, d\}$



Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

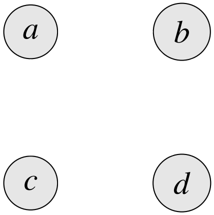
Transitive relation: for all $x, y, z \in X$, if $x \; R \; y$ and $y \; R \; z$, then $x \; R \; z$

Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Transitive relation: for all $x, y, z \in X$, if $x \; R \; y$ and $y \; R \; z$, then $x \; R \; z$

E.g., $X = \{a, b, c, d\}$

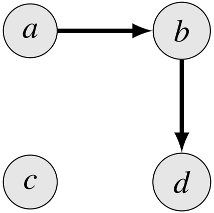


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Transitive relation: for all $x, y, z \in X$, if $x \; R \; y$ and $y \; R \; z$, then $x \; R \; z$

E.g., $X = \{a, b, c, d\}$

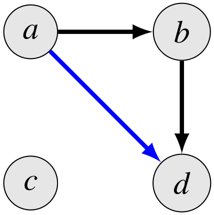


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Transitive relation: for all $x, y, z \in X$, if $x \; R \; y$ and $y \; R \; z$, then $x \; R \; z$

E.g., $X = \{a, b, c, d\}$

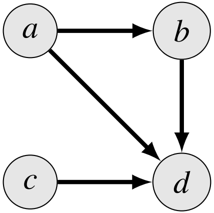


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Transitive relation: for all $x, y, z \in X$, if $x\ R\ y$ and $y\ R\ z$, then $x\ R\ z$

E.g., $X = \{a, b, c, d\}$

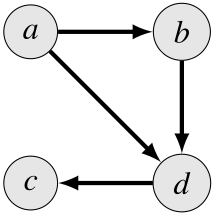


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Transitive relation: for all $x, y, z \in X$, if $x\ R\ y$ and $y\ R\ z$, then $x\ R\ z$

E.g., $X = \{a, b, c, d\}$

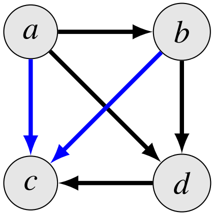


Mathematical background: Relations

Suppose that X is a set and $R \subseteq X \times X$ is a relation.

Transitive relation: for all $x, y, z \in X$, if $x\ R\ y$ and $y\ R\ z$, then $x\ R\ z$

E.g., $X = \{a, b, c, d\}$



Maximal elements, Cycles

Suppose that $R \subseteq X \times X$ is a relation.

$x \in X$ is **maximal** with respect to R provided there is no $y \in X$ such that $y\ R\ x$.

For $Y \subseteq X$, let $\max_R(Y) = \{x \in Y \mid \text{ there is no } y \in Y \text{ such that } y\ R\ x\}$

Maximal elements, Cycles

Suppose that $R \subseteq X \times X$ is a relation.

$x \in X$ is **maximal** with respect to R provided there is no $y \in X$ such that $y \, R \, x$.

For $Y \subseteq X$, let $\max_R(Y) = \{x \in Y \mid \text{ there is no } y \in Y \text{ such that } y \, R \, x\}$

A **cycle** is a set of distinct elements $x_1, \ldots, x_n$ such that

$$x_1 \; R \; x_2 \cdots x_{n-1} \; R \; x_n \; R \; x_1$$

R is **acyclic** if it does not contain any cycles.

Representing Preferences

Let X be a set of options/outcomes. A decision maker's *preference* over X is represented by a *relation* $\succeq \subseteq X \times X$.

Representing Preferences

Given $x, y \in X$, there are four possibilities:

Representing Preferences

Given $x, y \in X$, there are four possibilities:

1. $x \succsim y$ and $y \not\succsim x$: *The decision maker ranks x above y* (the decision maker strictly prefers x to y).

Representing Preferences

Given $x, y \in X$, there are four possibilities:

1. $x \succeq y$ and $y \not\succeq x$: *The decision maker ranks x above y* (the decision maker strictly prefers x to y).

2. $y \succeq x$ and $x \not\succeq y$: *The decision maker ranks y above x* (the decision maker strictly prefers y to x).

Representing Preferences

Given $x, y \in X$, there are four possibilities:

1. $x \succeq y$ and $y \not\succeq x$: *The decision maker ranks x above y* (the decision maker strictly prefers x to y).

2. $y \succeq x$ and $x \not\succeq y$: *The decision maker ranks y above x* (the decision maker strictly prefers y to x).

3. $x \succeq y$ and $y \succeq x$: The agent is *indifferent* between x and y.

Representing Preferences

Given $x, y \in X$, there are four possibilities:

1. $x \succeq y$ and $y \not\succeq x$: *The decision maker ranks x above y* (the decision maker strictly prefers x to y).

2. $y \succeq x$ and $x \not\succeq y$: *The decision maker ranks y above x* (the decision maker strictly prefers y to x).

3. $x \succeq y$ and $y \succeq x$: The agent is *indifferent* between x and y.

4. $x \not\succeq y$ and $y \not\succeq x$: The agent *cannot compare* x and y

Representing Preferences

Given $x, y \in X$, there are four possibilities:

1. $x \succeq y$ and $y \not\succeq x$: *The decision maker ranks x above y* (the decision maker strictly prefers x to y).

2. $y \succeq x$ and $x \not\succeq y$: *The decision maker ranks y above x* (the decision maker strictly prefers y to x).

3. $x \succeq y$ and $y \succeq x$: The agent is ***indifferent*** between x and y.

4. $x \not\succeq y$ and $y \not\succeq x$: The agent ***cannot compare*** x and y

Representing Preferences

Suppose that $\succeq$ is a preference relation. Then,

- **Strict preference**: $x \succ y$ iff $x \succeq y$ and $y \not\succeq x$
- **Indifference**: $x \sim y$ iff $x \succeq y$ and $y \succeq x$

Rational preferences

A relation $\succeq \subseteq X \times X$ is a **rational preference relation** (for a decision maker) provided that

1. $\succeq$ is complete (and hence reflexive)
2. $\succeq$ is transitive

- What is the relationship between choice and preference?
- What makes a preference *rational*?
- *Should* a decision maker's preference be complete and transitive?
- *Are* people's preferences complete and transitive?

Revealed Preference Theory

Standard economics focuses on revealed preference because economic data comes in this form. Economic data can—at best—reveal what the agent wants (or has chosen) in a particular situation. Such data do not enable the economist to distinguish between what the agent intended to choose and what he ended up choosing; what he chose and what he ought to have chosen. (Gul and Pesendorfer, 2008)

Given some choices of a decision maker, in what circumtances can we understand those choices as being made by a *rational* decision maker?

Sen's α Condition

R: red wine

W: white wine

L: lemonade

Sen's α Condition

R: red wine

W: white wine

L: lemonade

Sen's α Condition

R: red wine

W: white wine

Sen's α Condition

R: red wine

W: white wine

Sen's α Condition

R: red wine

W: white wine

L: lemonade

R: red wine

W: white wine

Sen's α Condition

R: red wine
W: white wine
L: lemonade

R: red wine
W: white wine

If the world champion is American, then she must be a US champion too.

Observations of actual choices will only partially constrain preference attribution. That someone chooses red wine when white wine is available does not allow one to conclude that the choice of an white wine was ruled out by her preferences, only that her preferences ruled the red wine in.

Sen's β Condition

R: red wine

W: white wine

Sen's β Condition

R: red wine

W: white wine

Sen's β Condition

R: red wine

W: white wine

L: lemonade

Sen's β Condition

R: red wine

W: white wine

L: lemonade

Sen's β Condition

R: red wine	R: red wine
W: white wine	W: white wine
	L: lemonade

If some American is a world champion, then all champions of America must be world champions.

Revealed Preference Theory

A decision maker's choices over a set of alternatives X are **rationalizable** iff there is a (rational) preference relation on X such that the decision maker's choices *maximize* the preference relation.

Revealed Preference Theory

A decision maker's choices over a set of alternatives X are **rationalizable** iff there is a (rational) preference relation on X such that the decision maker's choices *maximize* the preference relation.

Revelation Theorem. A decision maker's choices satisfy Sen's α and β if and only if the decision maker's choices are **rationalizable**.

Choice Functions

Suppose X is a set of options. And consider $B \subseteq X$ as a choice problem. A **choice function** is any function where $C(B) \subseteq B$. B is sometimes called a menu and $C(B)$ the set of "rational" or "desired" choices.

Choice Functions

Suppose X is a set of options. And consider $B \subseteq X$ as a choice problem. A **choice function** is any function where $C(B) \subseteq B$. B is sometimes called a menu and $C(B)$ the set of "rational" or "desired" choices.

A relation R on X **rationalizes a choice function** C if for all B
$C(B) = \{x \in B \mid \text{for all } y \in B \;\; xRy\}$.

Choice Functions

Suppose X is a set of options. And consider $B \subseteq X$ as a choice problem. A **choice function** is any function where $C(B) \subseteq B$. B is sometimes called a menu and $C(B)$ the set of "rational" or "desired" choices.

A relation R on X **rationalizes a choice function** C if for all B
$C(B) = \{x \in B \mid \text{for all } y \in B \;\; xRy\}$.

Sen's α: If $x \in C(A)$ and $B \subseteq A$ and $x \in B$ then $x \in C(B)$

Choice Functions

Suppose X is a set of options. And consider $B \subseteq X$ as a choice problem. A **choice function** is any function where $C(B) \subseteq B$. B is sometimes called a menu and $C(B)$ the set of "rational" or "desired" choices.

A relation R on X **rationalizes a choice function** C if for all B
$C(B) = \{x \in B \mid \text{for all } y \in B \;\; xRy\}$.

Sen's α: If $x \in C(A)$ and $B \subseteq A$ and $x \in B$ then $x \in C(B)$

Sen's β: If $x, y \in C(A)$, $A \subseteq B$ and $y \in C(B)$ then $x \in C(B)$.

Maximizing

A. Sen. *Maximization and the Act of Choice*. Econometrica, Vol. 65, No. 4, 1997, 745 - 779.

“The formulation of maximizing behavior in economics has often paralleled the modeling of maximization in physics an related disciplines.

Maximizing

A. Sen. *Maximization and the Act of Choice*. Econometrica, Vol. 65, No. 4, 1997, 745 - 779.

“The formulation of maximizing behavior in economics has often paralleled the modeling of maximization in physics an related disciplines. But maximizing *behavior* differs from nonvolitional *maximization* because of the fundamental relevance of the choice act, which has to be placed in a central position in analyzing maximizing behavior.

Maximizing

A. Sen. *Maximization and the Act of Choice*. Econometrica, Vol. 65, No. 4, 1997, 745 - 779.

"The formulation of maximizing behavior in economics has often paralleled the modeling of maximization in physics an related disciplines. But maximizing *behavior* differs from nonvolitional *maximization* because of the fundamental relevance of the choice act, which has to be placed in a central position in analyzing maximizing behavior. A person's preferences over *comprehensive* outcomes (including the choice process) have to be distinguished form the conditional preferences over *culmination* outcomes *given* the act of choice." (pg. 745)

Maximizing

You arrive at a garden party and can readily identify the most comfortable chair. You would be delighted if an imperious host were to assign you that chair. However, if the matter is left to your own choice, you may refuse to rush to it.

Maximizing

You arrive at a garden party and can readily identify the most comfortable chair. You would be delighted if an imperious host were to assign you that chair. However, if the matter is left to your own choice, you may refuse to rush to it. You select a “less preferred” chair.

Maximizing

You arrive at a garden party and can readily identify the most comfortable chair. You would be delighted if an imperious host were to assign you that chair. However, if the matter is left to your own choice, you may refuse to rush to it. You select a “less preferred” chair. Are you still a maximizer?

Maximizing

You arrive at a garden party and can readily identify the most comfortable chair. You would be delighted if an imperious host were to assign you that chair. However, if the matter is left to your own choice, you may refuse to rush to it. You select a "less preferred" chair. Are you still a maximizer? Quite possibly you are, since your preference ranking for choice behavior may well be defined over "comprehensive outcomes", including choice processes (in particular, who does the choosing) as well as the outcomes at culmination (the distribution of chairs). (Sen, pg. 747)

Invoking someone's preferences will suffice to explain why some choices were not made (i.e. in terms of rational impermissibility) but not typically why some particular choice was made. To take up the slack, explanations must draw on factors other than preference: psychological one such as the framing of the choice problem or the saliency of particular options, or sociological ones such as the existence of norms or conventions governing choices of the relevant kind.

Ordinal Utility Theory

Utility Function

A **utility function** on a set X is a function $u : X \to \mathbb{R}$

Utility Function

A **utility function** on a set X is a function $u : X \to \mathbb{R}$

A preference ordering is **represented** by a utility function iff x is (weakly) preferred to y provided $u(x) \geq u(y)$

Utility Function

A **utility function** on a set X is a function $u : X \to \mathbb{R}$

A preference ordering is **represented** by a utility function iff x is (weakly) preferred to y provided $u(x) \geq u(y)$

What properties does such a preference ordering have?

Ordinal Utility Theory

Fact. Suppose that X is finite and $\geq$ is a complete and transitive ordering over X, then there is a utility function $u : X \to \Re$ that represents $\geq$

(i.e., $x \geq y$ iff $u(x) \geq u(y)$)

Ordinal Utility Theory

Fact. Suppose that X is finite and $\geq$ is a complete and transitive ordering over X, then there is a utility function $u : X \to \Re$ that represents $\geq$

(i.e., $x \geq y$ iff $u(x) \geq u(y)$)

Utility is *defined* in terms of preference (so it is an error to say that the agent prefers x to y *because* she assigns a higher utility to x than to y).

Important

All three of the utility functions represent the preference $x \succ y \succ z$

Item	u_1	u_2	u_3
x	3	10	1000
y	2	5	99
z	1	0	1

$x \succ y \succ z$ is represented by both $(3, 2, 1)$ and $(1000, 999, 1)$, so one cannot say that y is "closer" to x than to z.

- What is the relationship between choice and preference?
- Why *should* preferences be complete and transitive?
- *Are* people's preferences complete and transitive?

- Transitivity: Money-pump argument
- Completeness: Incommensurable options

Transitivity

For all $x, y, z \in X$, if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

Transitivity

For all $x, y, z \in X$, if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

Indifference: For all $x, y, z \in X$, if $x \sim y$ and $y \sim z$, then $x \sim z$.

Transitivity

For all $x, y, z \in X$, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.

Indifference: For all $x, y, z \in X$, if $x \sim y$ and $y \sim z$, then $x \sim z$.

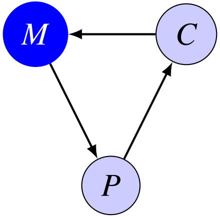
Strict preference: For all $x, y, z \in X$, if $x \succ y$ and $y \succ z$, then $x \succ z$.

Indifference is not transitive: $x_1 \sim x_2 \sim \cdots \sim x_n$, yet $x_1 \succ x_n$

Indifference is not transitive: $x_1 \sim x_2 \sim \cdots \sim x_n$, yet $x_1 \succ x_n$

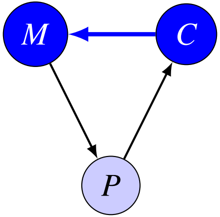
Cycle: $x_1 \succ x_2 \cdots \succ x_n$, yet $x_n \succ x_1$

Money-Pump Argument



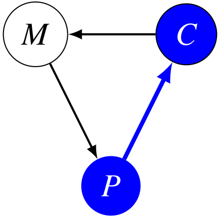
(M)

Money-Pump Argument



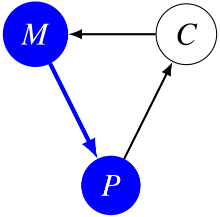
$$(M) \implies (C, -1)$$

Money-Pump Argument



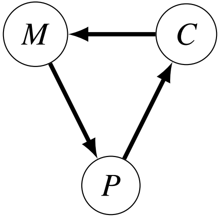
$$(M) \implies (C, -1) \implies (P, -2)$$

Money-Pump Argument



$$(M) \implies (C, -1) \implies (P, -2) \implies (M, -3)$$

Money-Pump Argument



$$(M) \implies (C, -1) \implies (P, -2) \implies (M, -3) \implies (C, -4) \implies \cdots$$

Completeness

For all $x, y \in X$, one of the following obtains:

1. the decision maker strictly prefers x over y ($x \succ y$);
2. the decision maker strictly prefers y over x ($y \succ x$); or
3. the decision maker is indifferent between x over y ($y \sim x$)

[O]f all the axioms of utility theory, the completeness axiom is perhaps the most questionable. Like others, it is inaccurate as a description of real life; but unlike them we find it hard to accept even from the normative viewpoint.

(Aumann, 1962)

Rather than trying to provide instrumental or pragmatic justifications for the axioms of ordinal utility, it is better...to see them as constitutive of our conception of a fully rational agent....those disposed to blatantly ignore transitivity are unintelligible to us: we can't understand their pattern of actions as sensible. [Gaus], pg. 39

Decision Problems

In many circumstances the decision maker doesn't get to choose outcomes directly, but rather chooses an instrument that affects what outcome actually occurs.

Decision Problems

In many circumstances the decision maker doesn't get to choose outcomes directly, but rather chooses an instrument that affects what outcome actually occurs.

Choice under

- *certainty*: highly confident about the relationship between actions and outcomes
- *risk*: clear sense of possibilities and their likelihoods
- *uncertainty*: the relationship between actions and outcomes is so imprecise that it is not possible to assign likelihoods

Decision Problems

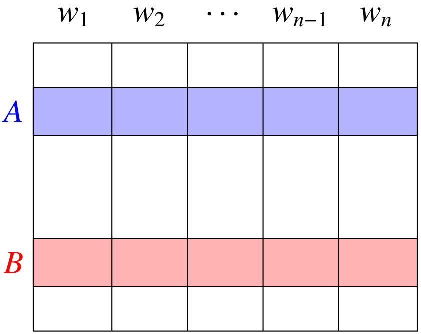
A

B

Decision Problems

	w_1	w_2	$\cdots$	w_{n-1}	w_n
A					
B					

Decision Problems



An **act** is a function $F : W \to O$

Making an omelet

States: {the sixth egg is good, the sixth egg is rotten}

Consequences: { six-egg omelet, no omelet and five good eggs destroyed, six-egg omelet and a cup to wash....}

Acts: { break egg into bowl, break egg into a cup, throw egg away}

Making an omelet

	Good egg (s_1)	Bad egg (s_2)
Break into a bowl (A_1)	six egg omelet (o_1)	no omelet and five good eggs destroyed (o_2)
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3)	five egg omelet and a cup to wash (o_4)
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5)	five egg omelet (o_6)

Making an omelet

	Good egg (s_1)	Bad egg (s_2)
Break into a bowl (A_1)	six egg omelet (o_1)	no omelet and five good eggs destroyed (o_2)
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3)	five egg omelet and a cup to wash (o_4)
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5)	five egg omelet (o_6)

$A_1(s_1) = o_1$ $\quad\quad$ $A_1(s_2) = o_2$

Making an omelet

	Good egg (s_1)	Bad egg (s_2)
Break into a bowl (A_1)	six egg omelet (o_1)	no omelet and five good eggs destroyed (o_2)
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3)	five egg omelet and a cup to wash (o_4)
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5)	five egg omelet (o_6)

$A_1(s_1) = o_1$ $\quad$ $A_1(s_2) = o_2$

Making an omelet

	Good egg (s_1)	Bad egg (s_2)
Break into a bowl (A_1)	six egg omelet (o_1)	no omelet and five good eggs destroyed (o_2)
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3)	five egg omelet and a cup to wash (o_4)
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5)	five egg omelet (o_6)

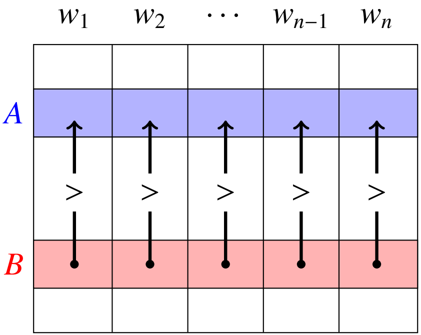
$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2$$

Making an omelet

	Good egg (s_1)	Bad egg (s_2)
Break into a bowl (A_1)	six egg omelet (o_1)	no omelet and five good eggs destroyed (o_2)
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3)	five egg omelet and a cup to wash (o_4)
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5)	five egg omelet (o_6)

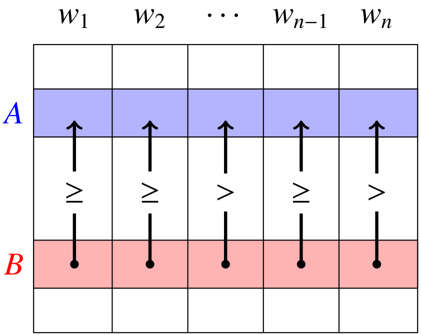
$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2$ How should A_1, A_2 and A_3 be ranked?

Strict Dominance



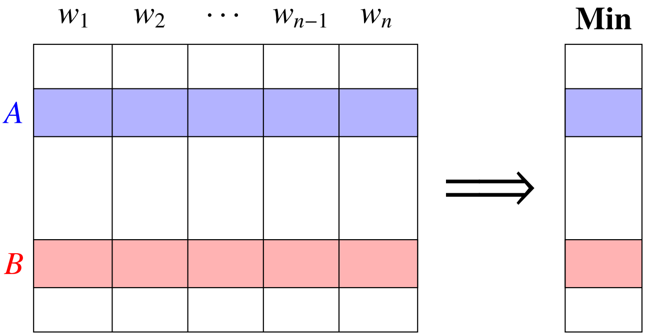
$$\forall\, w \in W,\ u(A(w)) > u(B(w))$$

Weak Dominance



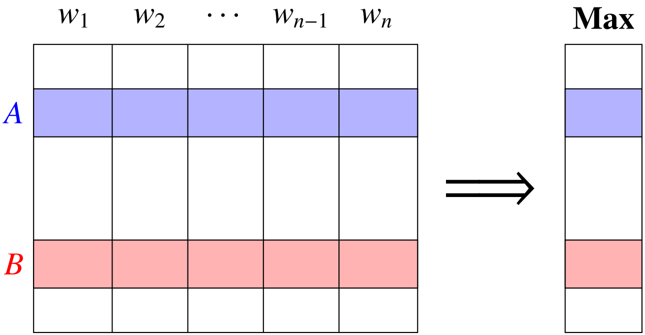
$$\forall\, w \in W,\, u(A(w)) \geq u(B(w)) \text{ and } \exists\, w \in W,\, u(A(w)) > u(B(w))$$

MaxMin (Security)



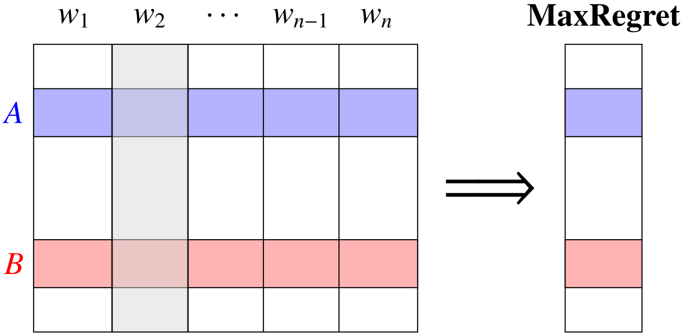
$$\min(\{u(A(w)) \mid w \in W\}) > \min(\{u(B(w)) \mid w \in W\})$$

MaxMax



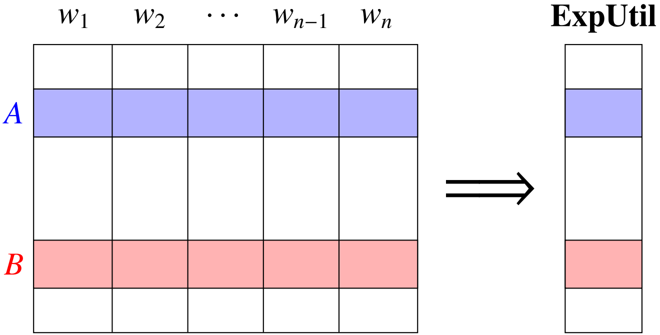
$$\max(\{u(A(w)) \mid w \in W\}) > \max(\{u(B(w)) \mid w \in W\})$$

MinMax Regret



$$\max(\{u(A(w_i)) - \max(\{u(A_i(w_i)) \mid A_i \in \mathsf{Act}\})$$

Maximize (Subjective) Expected Utility



$$\sum_{w \in W} P_A(w) * u(A(w)) > \sum_{w \in W} P_A(w) * u(B(w))$$

Subjective Expected Utility

Probability: Suppose that $W = \{w_1, \ldots, w_n\}$ is a finite set of states. A probability function on W is a function $P : W \to [0, 1]$ where $\sum_{w \in W} P(w) = 1$ (i.e., $P(w_1) + P(w_2) + \cdots + P(w_n) = 1$).

Suppose that A is an act for a set of outcomes O (i.e., $A : W \to O$). The **expected utility** of A is:

$$\sum_{w \in W} P(w) * u(A(w))$$

Making an omelet

	Good egg (s_1)	Bad egg (s_2)
Break into a bowl (A_1)	six egg omelet (o_1)	no omelet and five good eggs destroyed (o_2)
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3)	five egg omelet and a cup to wash (o_4)
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5)	five egg omelet (o_6)

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **6**	no omelet and five good eggs destroyed (o_2) **1**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **4**	five egg omelet and a cup to wash (o_4) **3**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **2**	five egg omelet (o_6) **5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$u(o_1) = 6, u(o_6) = 5, u(o_3) = 4, u(o_4) = 3, u(o_5) = 2, u(o_2) = 1$$

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **6**	no omelet and five good eggs destroyed (o_2) **1**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **4**	five egg omelet and a cup to wash (o_4) **3**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **2**	five egg omelet (o_6) **5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$EU(A_1) = P(s_1) * u(A_1(s_1)) + P(s_2) * u(A_1(s_2)) = 0.8 * 6 + 0.2 * 1 = 5.0$$

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **6**	no omelet and five good eggs destroyed (o_2) **1**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **4**	five egg omelet and a cup to wash (o_4) **3**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **2**	five egg omelet (o_6) **5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$EU(A_2) = P(s_1) * u(A_2(s_1)) + P(s_2) * u(A_2(s_2)) = 0.8 * 4 + 0.2 * 3 = 3.8$$

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **6**	no omelet and five good eggs destroyed (o_2) **1**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **4**	five egg omelet and a cup to wash (o_4) **3**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **2**	five egg omelet (o_6) **5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$EU(A_3) = P(s_1) * u(A_3(s_1)) + P(s_2) * u(A_3(s_2)) = 0.8 * 2 + 0.2 * 5 = 2.6$$

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **6**	no omelet and five good eggs destroyed (o_2) **1**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **4**	five egg omelet and a cup to wash (o_4) **3**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **2**	five egg omelet (o_6) **5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$EU(A_1) = 5 > EU(A_2) = 3.8 > EU(A_3) = 2.6$$

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **9**	no omelet and five good eggs destroyed (o_2) **0**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **8**	five egg omelet and a cup to wash (o_4) **7**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **1**	five egg omelet (o_6) **8.5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$u(o_1) = 9, u(o_6) = 8.5, u(o_3) = 8, u(o_4) = 7, u(o_5) = 1, u(o_2) = 0$$

Making an omelet

	Good egg (s_1) **0.8**	Bad egg (s_2) **0.2**
Break into a bowl (A_1)	six egg omelet (o_1) **9**	no omelet and five good eggs destroyed (o_2) **0**
Break into a cup (A_2)	six egg omelet and a cup to wash (o_3) **8**	five egg omelet and a cup to wash (o_4) **7**
Throw away (A_3)	five egg omelet and one good egg destroyed (o_5) **1**	five egg omelet (o_6) **8.5**

$$o_1 \succ o_6 \succ o_3 \succ o_4 \succ o_5 \succ o_2 \qquad P(s_1) = 0.8, P(s_2) = 0.2$$

$$EU(A_2) = 7.8 > EU(A_1) = 7.2 > EU(A_3) = 2.7$$

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

Expected utility of action A

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

Expected utility of action A

Utility of outcome o

$$EU(A) = \sum_{o \in O} P_A(o) \times U(o)$$

Expected utility of action *A*

Utility of outcome *o*

Probability of outcome *o* conditional on *A*

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

Evidential: $P_A(o) = P(o \mid A) = \frac{P(o \;\&\; A)}{P(A)}$

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

Evidential: $P_A(o) = P(o \mid A) = \frac{P(o \,\&\, A)}{P(A)}$

Classical: $P_A(o) = \sum_{s \in S} P(s) f_{A,s}(o)$, where

$$f_{A,s}(o) = \begin{cases} 1 & A(s) = o \\ 0 & A(s) \neq o \end{cases}$$

$P_A(o)$: probability of o conditional on A — how likely it is that outcome o will occur, on the supposition that the agent chooses act A.

Evidential: $P_A(o) = P(o \mid A) = \frac{P(o \,\&\, A)}{P(A)}$

Classical: $P_A(o) = \sum_{s \in S} P(s) f_{A,s}(o)$, where

$$f_{A,s}(o) = \begin{cases} 1 & A(s) = o \\ 0 & A(s) \neq o \end{cases}$$

Causal: $P_A(o) = P(A \mathrel{\Box\!\!\rightarrow} o)$

P("if A were performed, outcome o would ensue")

(Lewis, 1981)