Failures of Contingent Thinking

Evan Piermont
Royal Holloway – University of London

Peio Zuazo-Garin
ICEF – Higher School of Economics

Bounded Rationality in Choice, June 2022

We are motivated by two facts:

(1) There is a disconnect between the way uncertainty is modeled and how it is perceived by agents:

- Modeled by semantic state spaces; each state represents a complete description
  - State-space Ω, and probability μ over Ω.
- Real world uncertainty often doled out as a set of interconnected statements
  - “It is raining” or “The S&P500 went up today.”

We are motivated by two facts:

(1) There is a disconnect between the way uncertainty is modeled and how it is perceived by agents:
- Modeled by semantic state spaces; each state represents a complete description
  - State-space Ω, and probability μ over Ω.
- Real world uncertainty often doled out as a set of interconnected statements
  - “It is raining” or “The S&P500 went up today.”

(2) While rationality assumptions allow translating back and forth, humans are, unfortunately, not perfect reasoners.

Tversky and Kahneman (1983) provided subjects with the following vignette:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

When asked to rank the following statements in order of likelihood:

$\text{F} =$ "Linda is active in the feminist movement."

$\text{T} =$ "Linda is a bank teller."

$\text{T} \wedge \text{F} =$ "Linda is a bank teller and is active in the feminist movement."

- 85% of subjects ranked $\text{F} > \text{T} \wedge \text{F} > \text{T}$
- $\text{T} \wedge \text{F}$ implies T, any (classical) state space must rank the later weakly more likely.

1. Contingent thinking *is* the ability to recognize implications,

   ‘`if` … `then` …’

   - In the example, subjects do not recognize that $\mathrm{T} \wedge \mathrm{F}$ implies T
   - The opposite is also possible — perceive an implication that does not exist
   - We provide a precise behavioral criterion for perceiving implication

2. A state-space implicitly determines implication relations:
   - we must start with a more primitive *syntactic* objects
   - representation connects a subjective state-space to perceived implication

Failures of contingent thinking abound:

- Voting (pivotality)
  - Feddersen (JEP, 2004); Esponda and Vespa (AEJ Micro, 2014)
- Auctions (winning)
  - Thaler (JEP, 1988); Eyster and Rabin, (ETCA, 2005); Li (AER, 2017)
- Disclosure (no news)
  - Jin, Luca and Martin (WP, 2015), Enke (QJE, 2020)
- Information acquisition (generation process)
  - Enke and Zimmermann (ReStud, 2019), Enke (QJE, 2020)

Our goal is to provide:

- an abstract *definition* of contingent thinking that is context independent
- a methodology for *identifying* what the decision maker (DM) understands
- the limits of the *choice-data* necessary for such identification

Related (decision theoretic) Literature:

- Syntactic Decision Theory:
  - Tversky and Kahneman (1983); Mukerji (1997); Blume, Easley, Halpern (2021) and Bjorndahl and Halpern (2021)
- Incomplete of Inconsistent State-Spaces:
  - Inconsistent: Lipman (1999); Sadler (2020)
  - Incomplete: Fagin and Halpern (1988); Modica and Rustichini (1999); Heifetz, Meier, Schipper (2008)
- Misspecified Models:
  - Acemoglu et al. (2016); Mailath and Samuelson (2020); Frick et al. (2020); Eliaz et al. (2020); Ellis and Thysen (2021),

Decision environment

$\mathbb{P}$ is a set of **primitive statements**

- In the example, "Linda is bank teller," etc.

$\mathcal{L}$ is the **language** induced by $\mathbb{P}$ via negation, conjunction and disjunction

- If φ is in $\mathbb{P}$ then it is in $\mathcal{L}$ too
- If φ is in $\mathcal{L}$ then $\neg\varphi$ ("not φ") is in $\mathcal{L}$
- If φ and ψ are in $\mathcal{L}$ then $\varphi \wedge \psi$ ("φ and ψ") is in $\mathcal{L}$
- If φ and ψ are in $\mathcal{L}$ then $\varphi \vee \psi$ ("φ or ψ") is in $\mathcal{L}$

Let '$\varphi \to \psi$' mean that ϕ implies ψ according to the rules of classical logic.

- For example $\phi \Rightarrow \phi \vee \psi$

Decision environment

An **act** is a function $f : \Phi \to [0, \infty)$ with finite domain $\Phi \subseteq \mathcal{L}$

- f yields (vNM) utility $f(\varphi)$ when $\varphi \in \Phi$ is *true*, and is *called off* if no $\varphi \in \Phi$ is true
- x_Φ denotes the constant act $\Phi \mapsto x$
- $\mathcal{F}$ denotes the set of all acts

The DM has a **strict** preference $\succ$ over the set of acts

- Let $f \approx g$ if f and g satisfy the same $\succ$ relations.

Interpretations of uncertainty

An **interpretation** (**of uncertainty**) is a list (Ω, t, μ) consisting of:

1. A set of *states* Ω
2. A *truth-valuation map* $t : \mathcal{L} \to 2^\Omega$
3. A *likelihood assessment* $\mu : \mathcal{A} \to [0, \infty)$, where $\mathcal{A}$ is the algebra generated by the image of t

Interpretations of uncertainty

- $t(\varphi) \subseteq \Omega$ is the set of states in which φ is considered to true
  - t may not obey the usual logical dictates; e.g., $t(\varphi \wedge \psi) \not\subseteq t(\varphi)$

Interpretations of uncertainty

- $t(\varphi) \subseteq \Omega$ is the set of states in which φ is considered to true
  - t may not obey the usual logical dictates; e.g., $t(\varphi \wedge \psi) \not\subseteq t(\varphi)$
- $\mu(t(\varphi))$ quantifies the likelihood that statement φ is true
  - In the paper: μ may not be additive—not even monotone (e.g., $\mu(E) > \mu(F)$ for $E \subseteq F$)
  - For now: μ is a probability

Properties of t

- **Exact**: $\varphi \leftrightarrow \psi$ implies $t(\varphi) = t(\psi)$
- **Monotone**: $\varphi \to \psi$ implies $t(\varphi) \subseteq t(\psi)$
- **Distributive**: $t(\varphi \wedge \psi) = t(\varphi) \cap t(\psi)$
- **Symmetric**: $t(\neg\varphi) = \Omega \setminus t(\psi)$
- **Sound**: all of the above

Interpretations of uncertainty allow for evaluating an act in two steps:

interpretation and **aggregation**

Evaluation of an act - Interpretation

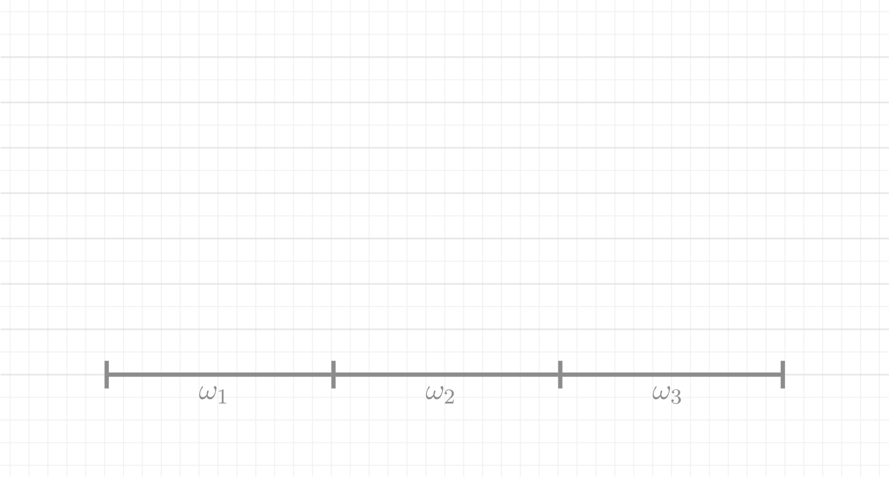
- An act $f : \Phi \to [0, \infty)$ is defined over linguistic statements $\varphi \in \Phi$
- Thus, the first step will consist in *interpreting* f as a map $\boldsymbol{f} : \Omega \to [0, \infty)$
- This interpretation is subjective and hinges on (Ω, t)
- There is an obvious ambiguity involved: Where to map states $\omega \in t(\varphi) \cap t(\psi)$ (if any) for $\varphi, \psi \in \Phi$?

Evaluation of an act - Interpretation

- $\Phi = \{\varphi, \neg\varphi\}$ and $f : \varphi \mapsto 2, f : \neg\varphi \mapsto 3$

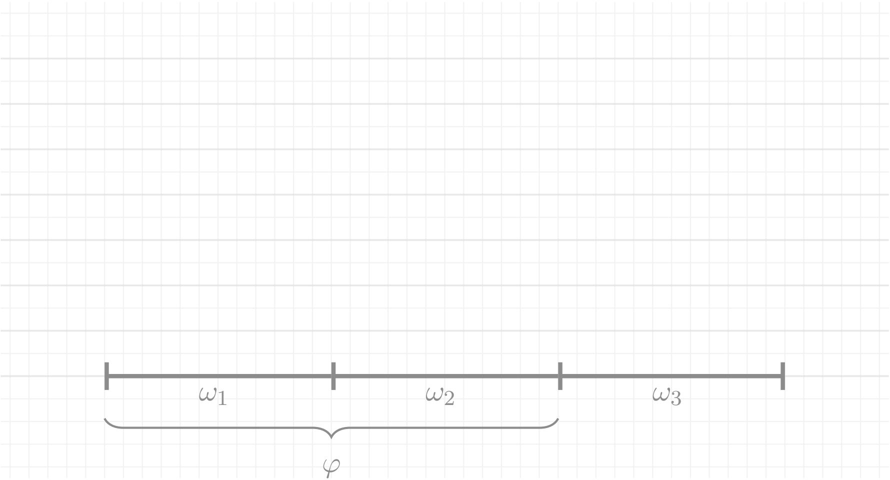
Evaluation of an act - Interpretation

- $\Omega = \{\omega_1, \omega_2, \omega_3\}$



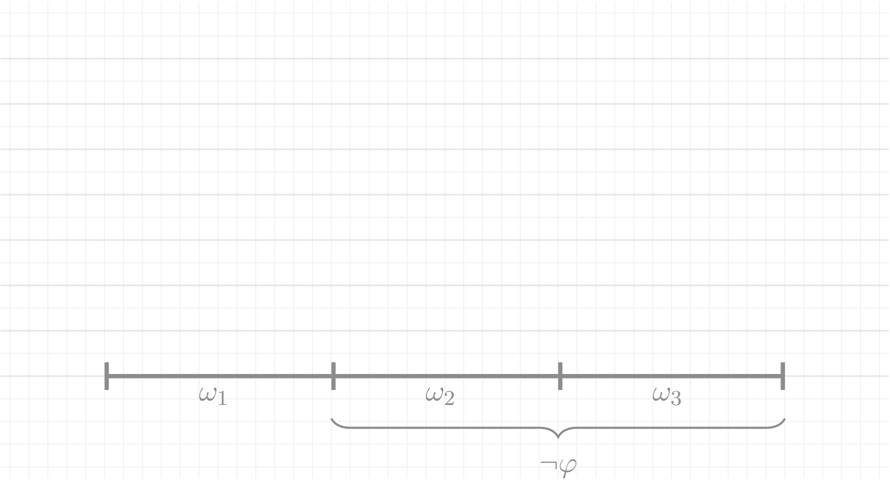
Evaluation of an act - Interpretation

- $t : \varphi \mapsto \{\omega_1, \omega_2\}$



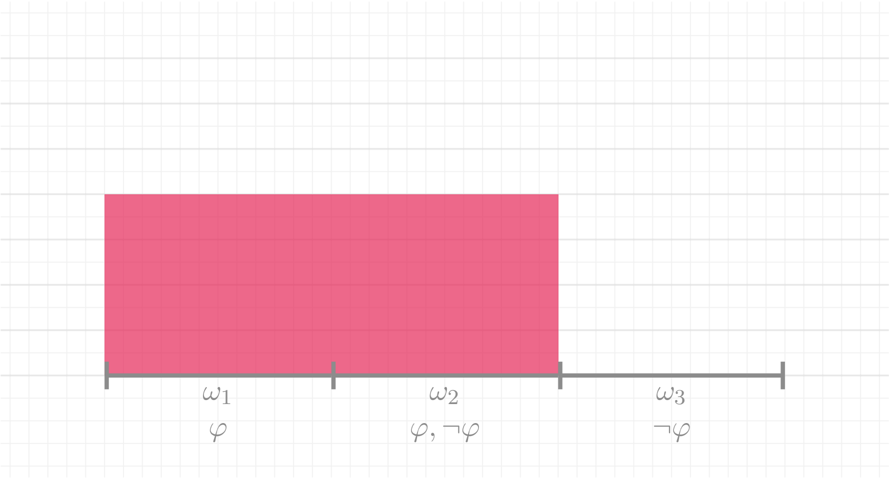
Evaluation of an act - Interpretation

- $t : \neg\varphi \mapsto \{\omega_2, \omega_3\}$



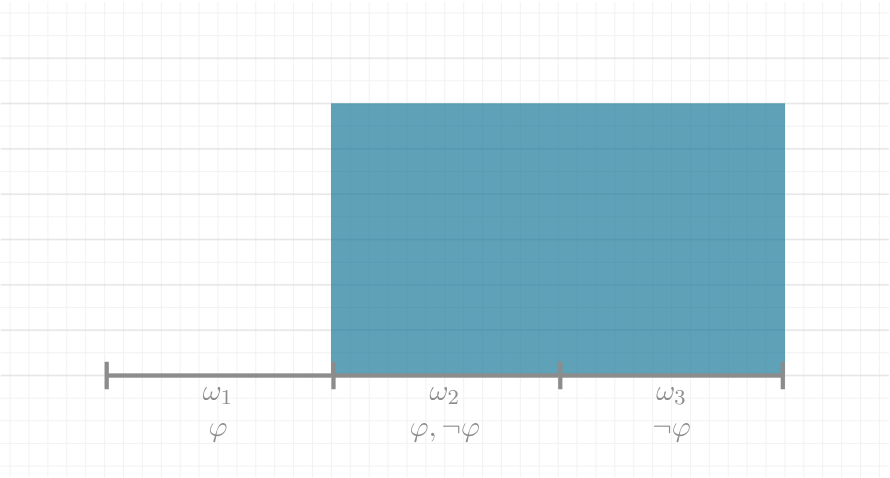
Evaluation of an act - Interpretation

- $f : \varphi \mapsto 2$



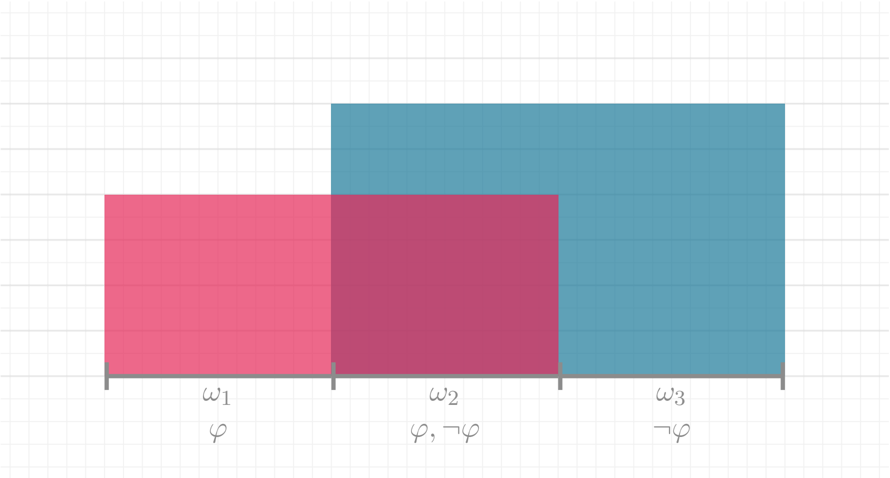
Evaluation of an act - Interpretation

- $f : \neg\varphi \mapsto 3$



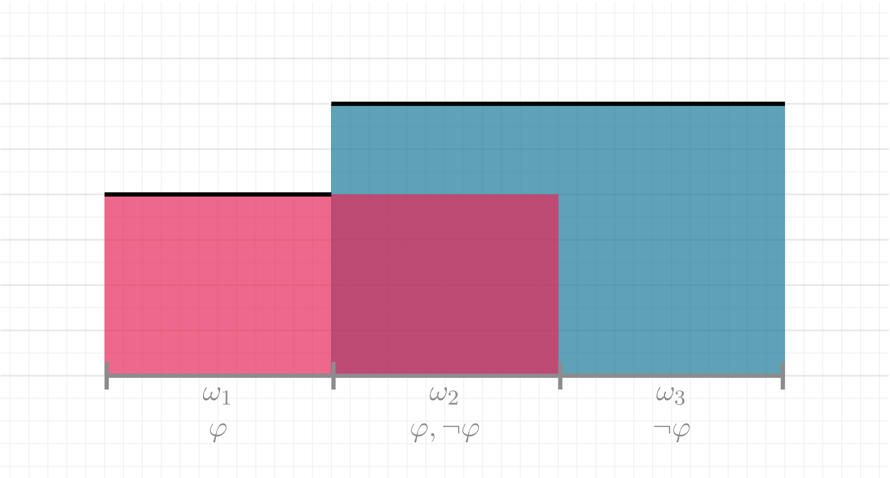
Evaluation of an act - Interpretation

- Payoff in state ω_2 is undefined



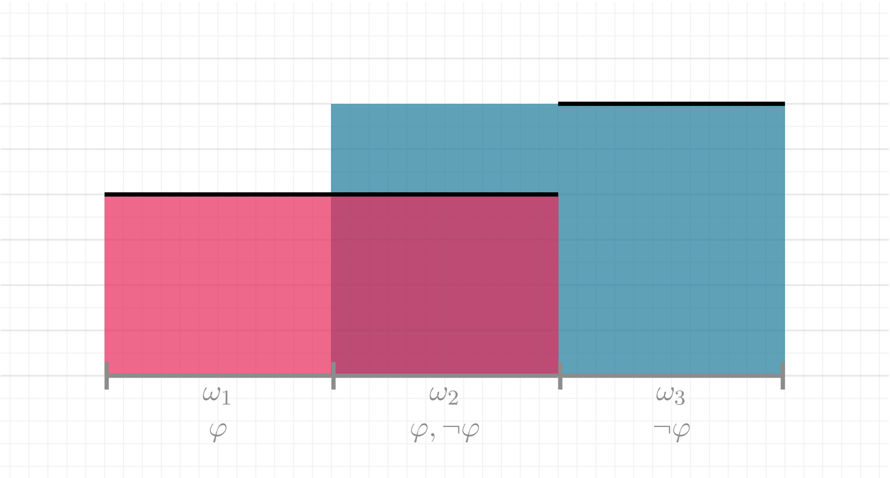
Evaluation of an act - Interpretation

- Maximal consistent act $\bar{\boldsymbol{f}}(\omega) = \sup\{f(\varphi) \mid \omega \in t(\varphi)\}$



Evaluation of an act - Interpretation

- Minimal consistent act $\underline{\boldsymbol{f}}(\omega) = \inf\{f(\varphi) \mid \omega \in t(\varphi)\}$



Evaluation of an act - Interpretation

Fix act f with domain Φ

- $\boldsymbol{f} : t(\Phi) \to [0, \infty)$ is **consistent** with f if:

$$\boldsymbol{f}(\omega) \in \{ f(\varphi) \mid \omega \in t(\varphi) \}$$

  for every $\omega \in t(\Phi)$

- Let $[\![f]\!]$ collect all the maps $\boldsymbol{f} : t(\Phi) \to [0, \infty)$ consistent with f
  - The multiplicity of $[\![f]\!]$ represents the ambiguity arising in the interpretation of f

Evaluation of an act - Aggregation

- Fix act f with domain Φ
- For each $\boldsymbol{f} \in [\![f]\!]$ we have a well-defined **Expected Utility**: $\int \boldsymbol{f} \, \mathrm{d}\mu$
- Thus, each syntactic act is associated to a subset of $[0, \infty)$:

$$\left\{ \int \boldsymbol{f} \, \mathrm{d}\mu \,\middle|\, \boldsymbol{f} \in [\![f]\!] \right\}$$

Interpretation-Dependent Expected Utility

$\succ$ is an **interpretation-dependent expected utility** (**IDEU**) **preference** if there exists some interpretation (Ω, t, μ) that represents $\succ$; i.e., such that, for every pair of acts f and g,

$$f \succ g \Longleftrightarrow \inf\left\{ \int \boldsymbol{f}\, \mathrm{d}\mu \,\middle|\, \boldsymbol{f} \in [\![f]\!] \right\} > \sup\left\{ \int \boldsymbol{g}\, \mathrm{d}\mu \,\middle|\, \boldsymbol{g} \in [\![g]\!] \right\}$$

- The ambiguity that arises when interpreting an act manifest as incompleteness of the preference
- We provide an axiomatization of these preferences in the paper.

The interpretation t formally captures the DM's contingent thinking:

- Benchmarks of rationality
- Understanding of contingencies via implication
- Relation to objective knowledge / models
- Belief updating

- Exact: $\varphi \leftrightarrow \psi$ implies $t(\varphi) = t(\psi)$
  - If $\varphi \leftrightarrow \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$

- Exact: $\varphi \leftrightarrow \psi$ implies $t(\varphi) = t(\psi)$
    - If $\varphi \leftrightarrow \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$
- Monotone: $\varphi \to \psi$ implies $t(\varphi) \subseteq t(\psi)$
    - If $\varphi \to \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$

- Exact: $\varphi \leftrightarrow \psi$ implies $t(\varphi) = t(\psi)$
  - If $\varphi \leftrightarrow \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$
- Monotone: $\varphi \to \psi$ implies $t(\varphi) \subseteq t(\psi)$
  - If $\varphi \to \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$
- Distributive: $t(\varphi \wedge \psi) = t(\varphi) \cap t(\psi)$
  - $\begin{cases} \varphi \mapsto x \\ \psi \mapsto 0 \end{cases} \approx \begin{cases} \varphi \mapsto x \\ \psi \wedge \varphi \mapsto 0 \end{cases}$

- Exact: $\varphi \leftrightarrow \psi$ implies $t(\varphi) = t(\psi)$
  - If $\varphi \leftrightarrow \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$
- Monotone: $\varphi \to \psi$ implies $t(\varphi) \subseteq t(\psi)$
  - If $\varphi \to \psi$ and $\varphi \in \Phi$ then $x_\Phi \approx x_{\Phi \cup \psi}$
- Distributive: $t(\varphi \wedge \psi) = t(\varphi) \cap t(\psi)$
  - $\begin{cases} \varphi \mapsto x \\ \psi \mapsto 0 \end{cases} \approx \begin{cases} \varphi \mapsto x \\ \psi \wedge \varphi \mapsto 0 \end{cases}$
- Symmetric: $t(\neg\varphi) = \Omega \setminus t(\psi)$
  - $x_{\{\varphi, \neg\varphi\}} \approx x_{\{\varphi \vee \neg\varphi\}}$

Perceived implications

For any $\varphi, \psi \in \mathcal{L}$, a DM with preference $\succ$ **perceives** that φ implies ψ, denoted by $\varphi \mathrel{=\!\succ} \psi$ if,

$$1_{\{\psi\}} \approx 1_{\{\varphi,\psi\}}.$$

- There is never any benefit to betting on φ given a bet on ψ
- We show $\mathrel{=\!\succ}$ is a partial order over propositions
  - Not true if μ is not a measure

- $\varphi \not\succ\!\!\!\!= \psi$ necessarily implies that $t(\varphi) \not\subseteq t(\psi)$ (otherwise $t(\{\psi\}) = t(\{\varphi\}) \cup t(\{\psi\})$

- Thus, there exists some $\omega \in t(\varphi) \setminus t(\psi)$, i.e., it is conceivable for the DM that φ holds and ψ does not

Faithful Representation

An interpretation (Ω, t, μ) is **faithful** (for $\succ$)if it represents $\succ$ and for every $\varphi, \psi \in \mathcal{L}$,

$$\varphi \Rightarrow_{\succ} \psi \text{ if and only if } t(\varphi) \subseteq t(\psi)$$

- Subjective implication exactly corresponds to set containment

Theorem

If $\succ$ has an IDEU representation it has a faithful IDEU representation.

- Moreover this is unique in some meaningful sense.

Identification of theories

Typically, elicitation experiments rely on the analyst assuming some (non-logical) relation between statements.

- ‘It is raining’ implies ‘The ground is wet‘
- This implications is true but not logically necessary—it requires a **theory**

Question

- Could seemingly erratic behavior be explained as not understanding that some statement is true?
- Can we disentangle irrationality from differing beliefs?

Identification of theories

Theories

A set $\mathcal{T} \subseteq \mathcal{L}$ is a **theory** if, for any $\varphi \in \mathcal{T}$ and any $\psi \in \mathcal{L}$ the following two hold:

1. If $\varphi \rightarrow \psi$ then $\psi \in \mathcal{T}$
2. If $\varphi \rightarrow \neg\psi$ then $\psi \notin \mathcal{T}$

A theory $\mathcal{T}$ allows for additional implications, denoted by $\xrightarrow{\mathcal{T}}$ ($\mathcal{T}$-implications)

E.g., if $\neg\varphi \vee \psi \in \mathcal{T}$ then $\varphi \xrightarrow{\mathcal{T}} \psi$ regardless of whether $\varphi \rightarrow \psi$ or not

An IDEU preference perceives $\mathcal{T}$-implications if $\varphi \Rightarrow\!\succ \psi$ for every $\varphi, \psi \in \mathcal{L}$ such that $\varphi \xrightarrow{\mathcal{T}} \psi$

Identification of theories

Theorem

Let $\succ$ be an IDEU preference that perceives implications, and let $\mathcal{T}$ be a theory. Then, there exists a unique theory $\mathcal{T}^{\succ} \subseteq \mathcal{T}$ such that:

1. $\succ$ perceives $\mathcal{T}^{\succ}$-implications
2. For any theory $\mathcal{T}'$ such that $\mathcal{T}^{\succ} \subseteq \mathcal{T}' \subseteq \mathcal{T}$, $\succ$ does not perceive all $\mathcal{T}'$-implications

We can identify the largest theory that rationalizes a rational DM's choices—even if inconsistent with the analyst's theory

What if we only observed preferences over simple bets of the form:

$$x_\varphi \mapsto \begin{cases} x & \text{if } \varphi \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

- In many environments (for example SEU) this suffices for identification.
- $x_\varphi \succ x_\psi$ implies φ is assessed as more likely than ψ.

- Given simple bets, we cannot distinguish failures of *logical* thinking from failures of *probabilistic* thinking.
- Non-rational t imparts the same preferences (over simple bets!) as weakening the conditions on μ

μ additive and	t sound and:
any t	N/A
t exact	any μ
t monotone	μ monotone
t $\wedge$-distributive	μ totally monotone
t sound	μ additive

Thank You!