Measure Semantics and Qualitative Semantics for Epistemic Modals

Perspectives on Modality

Wes Holliday and Thomas Icard
Berkeley and Stanford

April 12, 2013

Outline

- ‘probably’ and ‘at least as likely as’
- Previous Proposals
- Is Probability Necessary?
  - Fuzzy Measure Semantics
  - Qualitative Semantics
- Methodological Issues

Consider the English locution 'at least as likely as', as in

(1) It is at least as likely that our visitor is coming in on American Airlines as it is that he is coming on Continental Airlines.

Consider the English locution ‘at least as likely as’, as in

(1) It is at least as likely that our visitor is coming in on American Airlines as it is that he is coming on Continental Airlines.

What does this mean? Specifically, what is its logic?

Consider the English locution 'at least as likely as', as in

(1) It is at least as likely that our visitor is coming in on American Airlines as it is that he is coming on Continental Airlines.

What does this mean? Specifically, what is its logic?

Some entailments are clear. For instance, (1) follows from (2):

(2) American is at least as likely as Continental or Delta.

Consider the English locution ‘at least as likely as’, as in

(1) It is at least as likely that our visitor is coming in on American Airlines as it is that he is coming on Continental Airlines.

What does this mean? Specifically, what is its logic?

Some entailments are clear. For instance, (1) follows from (2):

(2) American is at least as likely as Continental or Delta.

What else? How might we interpret such talk model-theoretically?

What is the relation between ordinary talk using ‘probably’ and ‘at least as likely as’ and the mathematical theory of probability?

Is Kolmogorovian probability implicated in their semantics?

What is the relation between ordinary talk using 'probably' and 'at least as likely as' and the mathematical theory of probability?

Is Kolmogorovian probability implicated in their semantics?

Hamblin (1959, 234): "Metrical probability theory is well-established, scientifically important and, in essentials, beyond logical reproof. But when, for example, we say 'It's probably going to rain', or 'I shall probably be in the library this afternoon', are we, even vaguely, using the metrical probability concept?"

What is the relation between ordinary talk using 'probably' and 'at least as likely as' and the mathematical theory of probability?

Is Kolmogorovian probability implicated in their semantics?

Hamblin (1959, 234): "Metrical probability theory is well-established, scientifically important and, in essentials, beyond logical reproof. But when, for example, we say 'It's probably going to rain', or 'I shall probably be in the library this afternoon', are we, even vaguely, using the metrical probability concept?"

Kratzer (2012, 25): "Our semantic knowledge alone does not give us the precise quantitative notions of probability and desirability that mathematicians and scientists work with."

Formal Language

Given a set $\mathrm{At} = \{p, q, r, \dots\}$ of atomic sentence symbols, the language $\mathcal{L}(\Diamond, \geqslant)$ is generated by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Diamond\varphi \mid (\varphi \geqslant \varphi),$$

Formal Language

Given a set $\mathsf{At} = \{p, q, r, \dots\}$ of atomic sentence symbols, the language $\mathcal{L}(\Diamond, \geqslant)$ is generated by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Diamond\varphi \mid (\varphi \geqslant \varphi),$$

with the following intuitive readings:

$\Diamond\varphi$	"it might be that φ";
$\varphi \geqslant \psi$	"φ is at least as likely as ψ";

Formal Language

Given a set $\mathsf{At} = \{p, q, r, \dots\}$ of atomic sentence symbols, the language $\mathcal{L}(\Diamond, \succcurlyeq)$ is generated by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Diamond\varphi \mid (\varphi \succcurlyeq \varphi),$$

with the following intuitive readings:

$\Diamond\varphi$	"it might be that φ";
$\varphi \succcurlyeq \psi$	"φ is at least as likely as ψ";

We take $\vee$, $\rightarrow$, and $\leftrightarrow$ to be abbreviations, as well as the following:

$\Box\varphi := \neg\Diamond\neg\varphi$	"it must be that φ";
$\varphi \succ \psi := (\varphi \succcurlyeq \psi) \wedge \neg(\psi \succcurlyeq \varphi)$	"φ is more likely than ψ";
$\triangle\varphi := \varphi \succ \neg\varphi$	"probably φ".

Kratzer's Semantics

Definition (World-Ordering Model)

A (total) world-ordering model is a tuple
$\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V\rangle$:

- W is a non-empty set;

Kratzer's Semantics

Definition (World-Ordering Model)

A (total) world-ordering model is a tuple
$\mathbf{M} = \langle W, R, \{\succeq_w | \; w \in W\}, V\rangle$:

- W is a non-empty set;
- R is a (serial) binary relation on W;

Kratzer's Semantics

Definition (World-Ordering Model)

A (total) world-ordering model is a tuple
$\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V\rangle$:

- W is a non-empty set;
- R is a (serial) binary relation on W; $R(w) = \{v \in W \mid wRv\}$;

Kratzer's Semantics

Definition (World-Ordering Model)

A (total) world-ordering model is a tuple
$\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V \rangle$:

- W is a non-empty set;
- R is a (serial) binary relation on W; $R(w) = \{v \in W \mid wRv\}$;
- For each $w \in W$, $\succeq_w$ is a (total) **preorder** on $R(w)$;

Kratzer's Semantics

Definition (World-Ordering Model)

A (total) world-ordering model is a tuple $\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V \rangle$:

- W is a non-empty set;
- R is a (serial) binary relation on W; $R(w) = \{v \in W \mid wRv\}$;
- For each $w \in W$, $\succeq_w$ is a (total) **preorder** on $R(w)$;
- $V : \text{At} \to \wp(W)$ is a valuation function.

Kratzer's Semantics

Definition (World-Ordering Model)

A (total) world-ordering model is a tuple
$\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V\rangle$:

- W is a non-empty set;
- R is a (serial) binary relation on W; $R(w) = \{v \in W \mid wRv\}$;
- For each $w \in W$, $\succeq_w$ is a (total) **preorder** on $R(w)$;
- V : At $\to \wp(W)$ is a valuation function.

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \forall b \in B_w\, \exists a \in A_w \colon a \succeq_w b.$$

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq^l_w$ on $\wp(W)$:

$$A \succeq^l_w B \text{ iff } \forall b \in B_w \, \exists a \in A_w \colon a \succeq_w b.$$

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \forall b \in B_w \, \exists a \in A_w \colon a \succeq_w b.$$

Definition (Truth)

Given a pointed model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash \varphi$ and $[\![\varphi]\!]^{\mathbf{M}} = \{v \in W \mid \mathbf{M}, v \vDash \varphi\}$ as follows:

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \forall b \in B_w \, \exists a \in A_w \colon a \succeq_w b.$$

Definition (Truth)

Given a pointed model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash \varphi$ and $[\![\varphi]\!]^{\mathbf{M}} = \{v \in W \mid \mathbf{M}, v \vDash \varphi\}$ as follows:

$$\begin{array}{lll}
\mathbf{M}, w \vDash p & \text{iff} & w \in V(p); \\
\mathbf{M}, w \vDash \neg\varphi & \text{iff} & \mathbf{M}, w \nvDash \varphi; \\
\mathbf{M}, w \vDash \varphi \wedge \psi & \text{iff} & \mathbf{M}, w \vDash \varphi \text{ and } \mathbf{M}, w \vDash \psi;
\end{array}$$

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq^l_w$ on $\wp(W)$:

$$A \succeq^l_w B \text{ iff } \forall b \in B_w \, \exists a \in A_w : a \succeq_w b.$$

Definition (Truth)

Given a pointed model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash \varphi$ and $[\![\varphi]\!]^{\mathbf{M}} = \{v \in W \mid \mathbf{M}, v \vDash \varphi\}$ as follows:

$$\begin{array}{lll} \mathbf{M}, w \vDash p & \text{iff} & w \in V(p); \\ \mathbf{M}, w \vDash \neg\varphi & \text{iff} & \mathbf{M}, w \nvDash \varphi; \\ \mathbf{M}, w \vDash \varphi \wedge \psi & \text{iff} & \mathbf{M}, w \vDash \varphi \text{ and } \mathbf{M}, w \vDash \psi; \\ \mathbf{M}, w \vDash \Diamond\varphi & \text{iff} & \exists v \in R(w) : \mathbf{M}, v \vDash \varphi; \end{array}$$

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \forall b \in B_w\, \exists a \in A_w : a \succeq_w b.$$

Definition (Truth)

Given a pointed model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash \varphi$ and $[\![\varphi]\!]^{\mathbf{M}} = \{v \in W \mid \mathbf{M}, v \vDash \varphi\}$ as follows:

$\mathbf{M}, w \vDash p$	iff	$w \in V(p)$;
$\mathbf{M}, w \vDash \neg\varphi$	iff	$\mathbf{M}, w \nvDash \varphi$;
$\mathbf{M}, w \vDash \varphi \wedge \psi$	iff	$\mathbf{M}, w \vDash \varphi$ and $\mathbf{M}, w \vDash \psi$;
$\mathbf{M}, w \vDash \Diamond\varphi$	iff	$\exists v \in R(w) : \mathbf{M}, v \vDash \varphi$;
$\mathbf{M}, w \vDash \varphi \geqslant \psi$	iff	$[\![\varphi]\!]^{\mathbf{M}} \succeq_w^l [\![\psi]\!]^{\mathbf{M}}$.

As pointed out by Yalcin (2010) and Lassiter (2010), Kratzer's approach validates some rather dubious patterns. For instance, it predicts that (3) should follow from (1) and (2):

(1) American is at least as likely as Continental.

(2) American is at least as likely as Delta.

(3) American is at least as likely as Continental or Delta.

As pointed out by Yalcin (2010) and Lassiter (2010), Kratzer's approach validates some rather dubious patterns. For instance, it predicts that (3) should follow from (1) and (2):

(1) American is at least as likely as Continental.

(2) American is at least as likely as Delta.

(3) American is at least as likely as Continental or Delta.

It also fails to validate some intuitively obvious patterns.

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

V6 $\Box\varphi \to \triangle\varphi$

V7 $\triangle\varphi \to \Diamond\varphi$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

V6 $\Box\varphi \to \triangle\varphi$

V7 $\triangle\varphi \to \Diamond\varphi$

V11 $(\psi \geqslant \varphi) \to (\triangle\varphi \to \triangle\psi)$

V12 $(\psi \geqslant \varphi) \to ((\varphi \geqslant \neg\varphi) \to (\psi \geqslant \neg\psi))$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

V6 $\Box\varphi \to \triangle\varphi$

V7 $\triangle\varphi \to \Diamond\varphi$

V11 $(\psi \geqslant \varphi) \to (\triangle\varphi \to \triangle\psi)$

V12 $(\psi \geqslant \varphi) \to ((\varphi \geqslant \neg\varphi) \to (\psi \geqslant \neg\psi))$

I1 $((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi))$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

V6 $\Box\varphi \to \triangle\varphi$

V7 $\triangle\varphi \to \Diamond\varphi$

V11 $(\psi \geqslant \varphi) \to (\triangle\varphi \to \triangle\psi)$

V12 $(\psi \geqslant \varphi) \to ((\varphi \geqslant \neg\varphi) \to (\psi \geqslant \neg\psi))$

I1 $((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi))$

I2 $(\varphi \geqslant \neg\varphi) \to (\varphi \geqslant \psi)$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

V6 $\Box\varphi \to \triangle\varphi$

V7 $\triangle\varphi \to \Diamond\varphi$

V11 $(\psi \geqslant \varphi) \to (\triangle\varphi \to \triangle\psi)$

V12 $(\psi \geqslant \varphi) \to ((\varphi \geqslant \neg\varphi) \to (\psi \geqslant \neg\psi))$

I1 $((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi))$

I2 $(\varphi \geqslant \neg\varphi) \to (\varphi \geqslant \psi)$

I3 $\triangle\varphi \to (\varphi \geqslant \psi)$

Yalcin's List of Intuitively Valid and Invalid Patterns

V1 $\triangle\varphi \to \neg\triangle\neg\varphi$

V2 $\triangle(\varphi \wedge \psi) \to (\triangle\varphi \wedge \triangle\psi)$

V3 $\triangle\varphi \to \triangle(\varphi \vee \psi)$

V4 $\varphi \geqslant \bot$

V5 $\top \geqslant \varphi$

V6 $\Box\varphi \to \triangle\varphi$

V7 $\triangle\varphi \to \Diamond\varphi$

V11 $(\psi \geqslant \varphi) \to (\triangle\varphi \to \triangle\psi)$

V12 $(\psi \geqslant \varphi) \to ((\varphi \geqslant \neg\varphi) \to (\psi \geqslant \neg\psi))$

I1 $((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi))$

I2 $(\varphi \geqslant \neg\varphi) \to (\varphi \geqslant \psi)$

I3 $\triangle\varphi \to (\varphi \geqslant \psi)$

E1 $(\triangle\varphi \wedge \triangle\psi) \to \triangle(\varphi \wedge \psi)$

Set-Function Models

Definition (Relational Set-Function Model)

Consider models $\mathcal{M} = \langle W, R, \{\nu_w \mid w \in W\}, V\rangle$ such that

- $\nu_w : \wp(W) \to [0, 1]$ is a normalized set-function:
  - $\nu(\emptyset) = 0$;
  - $\nu(R(w)) = 1$;

Set-Function Models

Definition (Relational Set-Function Model)

Consider models $\mathcal{M} = \langle W, R, \{\nu_w \mid w \in W\}, V\rangle$ such that

- $\nu_w : \wp(W) \to [0, 1]$ is a normalized set-function:
  - $\nu(\emptyset) = 0$;
  - $\nu(R(w)) = 1$;

Definition (Truth)

Truth in a model is defined in the same way, except for the following clause:

$$\mathcal{M}, w \vDash \varphi \geqslant \psi \quad \text{iff} \quad \nu_w([\![\varphi]\!]^{\mathcal{M}}) \geq \nu_w([\![\psi]\!]^{\mathcal{M}}).$$

Probability Measures

Definition (Probability Measure)

A probability measure on a set W is a normalized set-function $\nu \colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- $A \cap B = \emptyset$, then $\nu(A \cup B) = \nu(A) + \nu(B)$.

Probability Measures

Definition (Probability Measure)

A probability measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- $A \cap B = \varnothing$, then $\nu(A \cup B) = \nu(A) + \nu(B)$.

Fact

V1-V12 are valid over the class of all probability measure models, while I1-I3 and E1 are not valid. ✓

Probability Measures

Definition (Probability Measure)

A probability measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- $A \cap B = \emptyset$, then $\nu(A \cup B) = \nu(A) + \nu(B)$.

Fact

V1-V12 are valid over the class of all probability measure models, while I1-I3 and E1 are not valid. ✓

What about axiomatization?

System FP

System FP

Taut all tautologies

System FP

Taut all tautologies

MP $\dfrac{\varphi \to \psi \qquad \varphi}{\psi}$

System FP

Taut all tautologies

MP $\dfrac{\varphi \to \psi \qquad \varphi}{\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

System FP

Taut all tautologies

MP $\dfrac{\varphi \to \psi \qquad \varphi}{\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

Ex $(\Box(\varphi \leftrightarrow \varphi') \wedge \Box(\psi \leftrightarrow \psi')) \to ((\varphi \geqslant \psi) \leftrightarrow (\varphi' \geqslant \psi'))$

System FP

Taut all tautologies

MP $\dfrac{\varphi \to \psi \qquad \varphi}{\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

Ex $(\Box(\varphi \leftrightarrow \varphi') \wedge \Box(\psi \leftrightarrow \psi')) \to ((\varphi \geqslant \psi) \leftrightarrow (\varphi' \geqslant \psi'))$

Bot $\varphi \geqslant \bot$

BT $\neg(\bot \geqslant \top)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

Scott $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m \to ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Scott $\varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m \rightarrow ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \rightarrow (\psi_m \geqslant \varphi_m))$

Scott $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m \to ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Here $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ abbreviates a $\mathcal{L}(\Diamond)$ formula such that:

- $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ iff for all $v \in R(w)$: $|\{\varphi_i \mid i \leq m, \mathcal{M}, v \vDash \varphi_i\}| = |\{\psi_i \mid i \leq m, \mathcal{M}, v \vDash \psi_i\}|$.

Scott $\varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m \to ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Here $\varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$ abbreviates a $\mathcal{L}(\diamond)$ formula such that:

- $\mathcal{M}, w \vDash \varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$ iff for all $v \in R(w)$: $|\{\varphi_i \mid i \leq m, \mathcal{M}, v \vDash \varphi_i\}| = |\{\psi_i \mid i \leq m, \mathcal{M}, v \vDash \psi_i\}|$.

We claim that if $\mathcal{M}, w \vDash \varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$, then

$$\sum_{i \leq m} \nu_w([\![\varphi_i]\!]^{\mathcal{M}}) = \sum_{i \leq m} \nu_w([\![\psi_i]\!]^{\mathcal{M}}). \tag{1}$$

Scott $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m \to ((\bigwedge_{i \le m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Here $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ abbreviates a $\mathcal{L}(\diamond)$ formula such that:

- $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ iff for all $v \in R(w)$:
  $|\{\varphi_i \mid i \le m, \mathcal{M}, v \vDash \varphi_i\}| = |\{\psi_i \mid i \le m, \mathcal{M}, v \vDash \psi_i\}|$.

We claim that if $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$, then

$$\sum_{i \le m} \nu_w([\![\varphi_i]\!]^{\mathcal{M}}) = \sum_{i \le m} \nu_w([\![\psi_i]\!]^{\mathcal{M}}). \tag{1}$$

If the model is finite, then to show (1) it suffices to show

$$\sum_{i \le m} \sum_{x \in [\![\varphi_i]\!]^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}) = \sum_{i \le m} \sum_{x \in [\![\psi_i]\!]^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}),$$

Scott $\varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m \to ((\bigwedge_{i \le m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Here $\varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$ abbreviates a $\mathcal{L}(\Diamond)$ formula such that:

- $\mathcal{M}, w \vDash \varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$ iff for all $v \in R(w)$: $|\{\varphi_i \mid i \le m,\, \mathcal{M}, v \vDash \varphi_i\}| = |\{\psi_i \mid i \le m,\, \mathcal{M}, v \vDash \psi_i\}|$.

We claim that if $\mathcal{M}, w \vDash \varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$, then

$$\sum_{i \le m} \nu_w(\llbracket \varphi_i \rrbracket^{\mathcal{M}}) = \sum_{i \le m} \nu_w(\llbracket \psi_i \rrbracket^{\mathcal{M}}). \tag{1}$$

If the model is finite, then to show (1) it suffices to show

$$\sum_{i \le m} \sum_{x \in \llbracket \varphi_i \rrbracket^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}) = \sum_{i \le m} \sum_{x \in \llbracket \psi_i \rrbracket^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}), \tag{2}$$

which follows from $\mathcal{M}, w \vDash \varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m$.

Scott $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m \to ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Here $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ abbreviates a $\mathcal{L}(\Diamond)$ formula such that:

- $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ iff for all $v \in R(w)$:
  $|\{\varphi_i \mid i \leq m, \mathcal{M}, v \vDash \varphi_i\}| = |\{\psi_i \mid i \leq m, \mathcal{M}, v \vDash \psi_i\}|$.

We claim that if $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$, then

$$\sum_{i \leq m} \nu_w(\llbracket \varphi_i \rrbracket^{\mathcal{M}}) = \sum_{i \leq m} \nu_w(\llbracket \psi_i \rrbracket^{\mathcal{M}}). \tag{1}$$

If the model is finite, then to show (1) it suffices to show

$$\sum_{i \leq m} \sum_{x \in \llbracket \varphi_i \rrbracket^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}) = \sum_{i \leq m} \sum_{x \in \llbracket \psi_i \rrbracket^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}), \tag{2}$$

which follows from $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$. Given (1), $\mathcal{M}, w \vDash (\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m)$.

Scott $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m \rightarrow ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \rightarrow (\psi_m \geqslant \varphi_m))$

Here $\varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ abbreviates a $\mathcal{L}(\Diamond)$ formula such that:

- $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$ iff for all $v \in R(w)$:
  $|\{\varphi_i \mid i \leq m, \mathcal{M}, v \vDash \varphi_i\}| = |\{\psi_i \mid i \leq m, \mathcal{M}, v \vDash \psi_i\}|$.

We claim that if $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$, then

$$\sum_{i \leq m} \nu_w([\![\varphi_i]\!]^{\mathcal{M}}) = \sum_{i \leq m} \nu_w([\![\psi_i]\!]^{\mathcal{M}}). \tag{1}$$

If the model is finite, then to show (1) it suffices to show

$$\sum_{i \leq m} \sum_{x \in [\![\varphi_i]\!]^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}) = \sum_{i \leq m} \sum_{x \in [\![\psi_i]\!]^{\mathcal{M}} \cap R(w)} \nu_w(\{x\}), \tag{2}$$

which follows from $\mathcal{M}, w \vDash \varphi_1 \dots \varphi_m \mathbb{E} \psi_1 \dots \psi_m$. Given (1), $\mathcal{M}, w \vDash (\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \rightarrow (\psi_m \geqslant \varphi_m)$. Holds in infinite too.

System FP

Bot $\varphi \geqslant \bot$

BT $\neg(\bot \geqslant \top)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

Scott $\varphi_1 \ldots \varphi_m \mathbb{E} \psi_1 \ldots \psi_m \to ((\bigwedge_{i \leq m-1} (\varphi_i \geqslant \psi_i)) \to (\psi_m \geqslant \varphi_m))$

Theorem (Scott 1964, Segerberg 1971, Gärdenfors 1975)

FP is sound/complete with respect to probability measure models.

Is Probability Necessary?

Having seen that a probability-based semantics is sufficient for validating V1-V12 and invalidating I1-I3 and E1, let us now consider whether such a semantics is necessary.

Is Probability Necessary?

Having seen that a probability-based semantics is sufficient for validating V1-V12 and invalidating I1-I3 and E1, let us now consider whether such a semantics is necessary.

> [I]t may be questioned whether probability spaces really are appropriate to the semantics of (what superficially appears to be) natural language probability talk. Hamblin 1959, an impressive early investigation into this question, seems to favour a plausibility measure approach; and Kratzer 1991 gives a semantics for probability operators in terms of nonnumerical qualitative orderings of possibilities. It would be desirable to demonstrate, in so far as possible, that the resources of probability theory are in fact needed. (Yalcin 2007, 1019)

Is Probability Necessary?

While Yalcin (2010) shows that the semantics of Kratzer and Hamblin validate too much and yet not enough, and Lassiter (2011) gives additional arguments for a probability-based semantics, there are other options.

Is Probability Necessary?

While Yalcin (2010) shows that the semantics of Kratzer and Hamblin validate too much and yet not enough, and Lassiter (2011) gives additional arguments for a probability-based semantics, there are other options. We will show that semantics based on fuzzy measures solve the entailment problems raised for Kratzer and Hamblin, as do some purely qualitative semantics.

Alternative Systems

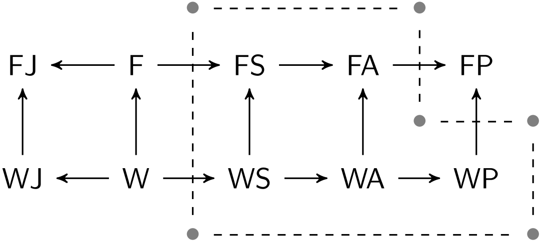

Figure : Logical Landscape

Hamblin's Semantics

Definition (Possibility Measure)

A possibility measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- $\nu(A \cup B) = \max(\nu(A), \nu(B))$.

Hamblin's Semantics

Definition (Possibility Measure)

A possibility measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- $\nu(A \cup B) = \max(\nu(A), \nu(B))$.

Fact

V1-V10 and V12 are all valid over possibility measure models; V11 is not valid; I1-13 and E1 are all valid. X

Alternative 1

Definition (Fuzzy Measure)

A (normalized) fuzzy measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- if $A \subseteq B$, then $\nu(A) \leq \nu(B)$.

Alternative 1

Definition (Fuzzy Measure)

A (normalized) fuzzy measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- if $A \subseteq B$, then $\nu(A) \leq \nu(B)$.

A fuzzy measure is self-dual iff for all $A \subseteq W$:

- $\nu(A) + \nu(A^c) = 1$, where $A^c = \{w \in W \mid w \notin A\}$.

Alternative 1

Definition (Fuzzy Measure)

A (normalized) fuzzy measure on a set W is a normalized set-function $\nu\colon \wp(W) \to [0,1]$ such that for all $A, B \subseteq W$:

- if $A \subseteq B$, then $\nu(A) \leq \nu(B)$.

A fuzzy measure is self-dual iff for all $A \subseteq W$:

- $\nu(A) + \nu(A^c) = 1$, where $A^c = \{w \in W \mid w \notin A\}$.

Fact

V1-V12 are all valid over self-dual fuzzy measure models, while none of I1-I3 or E1 are valid. ✓

Systems **F**, **FS**, and **FJ**

System **F** is **K** plus:

Mon $\Box(\varphi \to \psi) \to (\psi \geqslant \varphi)$

Systems **F**, **FS**, and **FJ**

System **F** is **K** plus:

Mon $\Box(\varphi \to \psi) \to (\psi \geqslant \varphi)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

BT $\neg(\bot \geqslant \top)$

Tran $(\varphi \geqslant \psi) \to ((\psi \geqslant \chi) \to (\varphi \geqslant \chi))$

Systems **F**, **FS**, and **FJ**

System **F** is **K** plus:

Mon $\Box(\varphi \to \psi) \to (\psi \geqslant \varphi)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

BT $\neg(\bot \geqslant \top)$

Tran $(\varphi \geqslant \psi) \to ((\psi \geqslant \chi) \to (\varphi \geqslant \chi))$

System **FS** is **F** plus: S $(\varphi \geqslant \psi) \to (\neg\psi \geqslant \neg\varphi)$

Systems **F**, **FS**, and **FJ**

System **F** is **K** plus:

Mon $\Box(\varphi \to \psi) \to (\psi \geqslant \varphi)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

BT $\neg(\bot \geqslant \top)$

Tran $(\varphi \geqslant \psi) \to ((\psi \geqslant \chi) \to (\varphi \geqslant \chi))$

System **FS** is **F** plus: S $(\varphi \geqslant \psi) \to (\neg\psi \geqslant \neg\varphi)$

System **FJ** is **F** plus: J $((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi))$

Systems **F**, **FS**, and **FJ**

System **F** is **K** plus:

Mon $\Box(\varphi \to \psi) \to (\psi \geqslant \varphi)$ Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

BT $\neg(\bot \geqslant \top)$ Tran $(\varphi \geqslant \psi) \to ((\psi \geqslant \chi) \to (\varphi \geqslant \chi))$

System **FS** is **F** plus: S $(\varphi \geqslant \psi) \to (\neg\psi \geqslant \neg\varphi)$

System **FJ** is **F** plus: J $((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi))$

Theorem (Fuzzy Measure Axiomatizations)

1. **F** is sound/complete for the class of fuzzy measure models.
2. **FS** is sound/complete for self-dual fuzzy measure models.
3. **FJ** is sound/complete for possibility measure models.

Systems **F**, **FS**, **FJ**

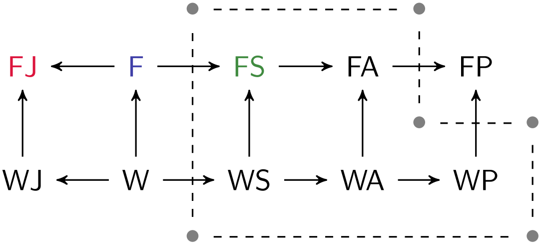

Figure : Logical Landscape

Stronger Systems

Although self-dual fuzzy measure semantics solves the entailment problems raised for Kratzer and Hamblin's semantics, one may still argue in favor of moving to a semantics with a stronger logic, if not as strong as **FP**, to capture reasoning that depends on some form of additivity. In the following slides, we will put additional constraints on fuzzy measures to obtain such semantics.

Alternative 2

Definition (Quasi-Additive Measures)

A quasi-additive measure on a set W is a normalized set-function $\nu : \wp(W) \to [0,1]$ such that for all $A, B, C \subseteq W$:

- $A \cap (B \cup C) = \varnothing \Rightarrow [\nu(B) \leq \nu(C) \text{ iff } \nu(A \cup B) \leq \nu(A \cup C)]$

Alternative 2

Definition (Quasi-Additive Measures)

A quasi-additive measure on a set W is a normalized set-function $\nu : \wp(W) \rightarrow [0, 1]$ such that for all $A, B, C \subseteq W$:

- $A \cap (B \cup C) = \emptyset \Rightarrow [\nu(B) \leq \nu(C) \text{ iff } \nu(A \cup B) \leq \nu(A \cup C)]$

A quasi-additive measure is self-dual iff for all $A \subseteq W$:

- $\nu(A) + \nu(A^c) = 1.$

Alternative 2

Definition (Quasi-Additive Measures)

A quasi-additive measure on a set W is a normalized set-function $\nu : \wp(W) \to [0,1]$ such that for all $A, B, C \subseteq W$:

- $A \cap (B \cup C) = \emptyset \Rightarrow [\nu(B) \leq \nu(C) \text{ iff } \nu(A \cup B) \leq \nu(A \cup C)]$

A quasi-additive measure is self-dual iff for all $A \subseteq W$:

- $\nu(A) + \nu(A^c) = 1.$

Fact

V1-V12 are all valid over quasi-additive measure models, while none of I1-I3 or E1 are valid over these (self-dual) models. ✓

de Finetti's System **FA**

System **FA** is **K** plus Ex and:

Bot $\varphi \geqslant \bot$

BT $\neg(\bot \geqslant \top)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

Tran $(\varphi \geqslant \psi) \rightarrow ((\psi \geqslant \chi) \rightarrow (\varphi \geqslant \chi))$

A $\neg\Diamond(\chi \wedge (\varphi \vee \psi)) \rightarrow (\varphi \geqslant \psi \leftrightarrow ((\chi \vee \varphi) \geqslant (\chi \vee \psi)))$

de Finetti's System **FA**

System **FA** is **K** plus Ex and:

Bot $\varphi \geqslant \bot$

BT $\neg(\bot \geqslant \top)$

Tot $(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi)$

Tran $(\varphi \geqslant \psi) \to ((\psi \geqslant \chi) \to (\varphi \geqslant \chi))$

A $\neg\Diamond(\chi \wedge (\varphi \vee \psi)) \to (\varphi \geqslant \psi \leftrightarrow ((\chi \vee \varphi) \geqslant (\chi \vee \psi)))$

Theorem

FA is sound and complete for the class of quasi-additive measure models and the class of self-dual quasi-additive measure models.

System **FA**

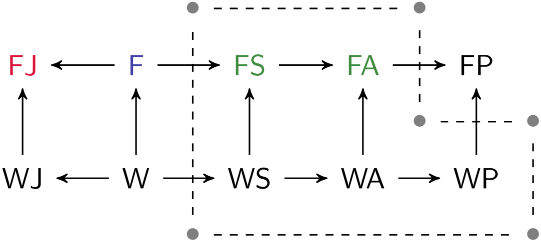

Figure : Logical Landscape

Alternative 3

Definition (Qualitative Probability Orderings)

Given a set W, a weak qualitative probability ordering $\succsim$ is a binary relation on $\wp(W)$ such that for all $A, B, C \subseteq W$:

not $\emptyset \succsim W$; if $A \succsim B$ and $B \succsim C$, then $A \succsim C$;

if $A \supseteq B$, then $A \succsim B$.

Alternative 3

Definition (Qualitative Probability Orderings)

Given a set W, a weak qualitative probability ordering $\succsim$ is a binary relation on $\wp(W)$ such that for all $A, B, C \subseteq W$:

not $\emptyset \succsim W$; if $A \succsim B$ and $B \succsim C$, then $A \succsim C$;

if $A \supseteq B$, then $A \succsim B$.

$\succsim$ is complementary iff all $A, B \subseteq W$: $A \succsim B$ iff $B^c \succsim A^c$.

Alternative 3

Definition (Qualitative Probability Orderings)

Given a set W, a weak qualitative probability ordering $\succsim$ is a binary relation on $\wp(W)$ such that for all $A, B, C \subseteq W$:

not $\emptyset \succsim W$; if $A \succsim B$ and $B \succsim C$, then $A \succsim C$;

if $A \supseteq B$, then $A \succsim B$.

$\succsim$ is complementary iff all $A, B \subseteq W$: $A \succsim B$ iff $B^c \succsim A^c$.

Quasi-additive QP orderings replace the last two by $A \succsim \emptyset$ and

if $A \cap (B \cup C) = \emptyset$, then $B \succsim C$ iff $A \cup B \succsim A \cup C$.

Alternative 3

Definition (Qualitative Probability Orderings)

Given a set W, a weak qualitative probability ordering $\succsim$ is a binary relation on $\wp(W)$ such that for all $A, B, C \subseteq W$:

not $\emptyset \succsim W$; if $A \succsim B$ and $B \succsim C$, then $A \succsim C$;

if $A \supseteq B$, then $A \succsim B$.

$\succsim$ is complementary iff all $A, B \subseteq W$: $A \succsim B$ iff $B^c \succsim A^c$.

Quasi-additive QP orderings replace the last two by $A \succsim \emptyset$ and

if $A \cap (B \cup C) = \emptyset$, then $B \succsim C$ iff $A \cup B \succsim A \cup C$.

Finally, $\succsim$ is total iff for all $A, B \subseteq W$: $A \succsim B$ or $B \succsim A$.

Alternative 3

A weak qualitative probability model is a tuple $\mathcal{M} = \langle W, R, \{\succsim_w \mid w \in W\}, V\rangle$, where $\succsim_w$ is a weak qualitative probability ordering such that $R(w) \succsim_w W$.

Alternative 3

A weak qualitative probability model is a tuple $\mathcal{M} = \langle W, R, \{\succsim_w \mid w \in W\}, V\rangle$, where $\succsim_w$ is a weak qualitative probability ordering such that $R(w) \succsim_w W$.

Definition (Truth)

Given a pointed model $\mathbb{M}, w$ and φ in $\mathcal{L}(\Diamond, \geqslant)$, we define $\mathbb{M}, w \vDash \varphi$ as follows (with other cases as before):

$$\mathbb{M}, w \vDash \varphi \geqslant \psi \quad \text{iff} \quad [\![\varphi]\!]^{\mathbb{M}} \succsim_w [\![\psi]\!]^{\mathbb{M}}.$$

Alternative 3

A weak qualitative probability model is a tuple $\mathcal{M} = \langle W, R, \{\succsim_w \mid w \in W\}, V\rangle$, where $\succsim_w$ is a weak qualitative probability ordering such that $R(w) \succsim_w W$.

Definition (Truth)

Given a pointed model $\mathbb{M}, w$ and φ in $\mathcal{L}(\Diamond, \geqslant)$, we define $\mathbb{M}, w \vDash \varphi$ as follows (with other cases as before):

$$\mathbb{M}, w \vDash \varphi \geqslant \psi \quad \text{iff} \quad [\![\varphi]\!]^{\mathbb{M}} \succsim_w [\![\psi]\!]^{\mathbb{M}}.$$

Fact

V1-V12 are all valid over complementary weak qualitative probability models, while none of I1-I3 or E1 are valid. ✓

Systems **W**, **WS**, **WA**

System **W** is **F** minus Tot. System **WS** is **FS** minus Tot.

S $(\varphi \geqslant \psi) \rightarrow (\neg\psi \geqslant \neg\varphi)$.

System **WA** is **FA** minus Tot.

A $\neg\Diamond(\chi \wedge (\varphi \vee \psi)) \rightarrow (\varphi \geqslant \psi \leftrightarrow ((\chi \vee \varphi) \geqslant (\chi \vee \psi)))$.

Systems **W**, **WS**, **WA**

System **W** is **F** minus Tot. System **WS** is **FS** minus Tot.

S $(\varphi \geqslant \psi) \rightarrow (\neg\psi \geqslant \neg\varphi)$.

System **WA** is **FA** minus Tot.

A $\neg\Diamond(\chi \wedge (\varphi \vee \psi)) \rightarrow (\varphi \geqslant \psi \leftrightarrow ((\chi \vee \varphi) \geqslant (\chi \vee \psi)))$.

Theorem (Qualitative Probability Axiomatizations)

1. **W** is sound/complete for weak QP models.
2. **WS** is sound/complete for complementary weak QP models.
3. **F** is sound/complete for total weak QP models.
4. **FS** is sound/complete for complementary total weak QP models.
5. **WA** is sound/complete for quasi-additive QP models.
6. **FA** is sound/complete for total quasi-additive QP models.

Systems **WJ**, **W**, **WS**, **WA**

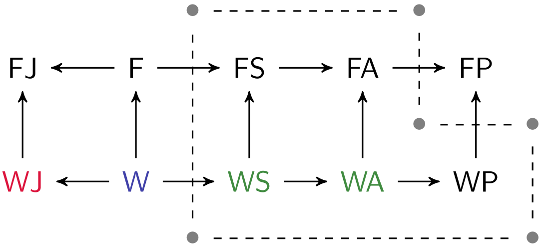

Figure : Logical Landscape

Kratzer Revisited

Definition (World-Ordering Model)

A (total) world-ordering model $\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V \rangle$ has for each $w \in W$ a (total) **preorder** $\succeq_w$ on $R(w)$.

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \forall b \in B_w \, \exists a \in A_w \colon a \succeq_w b.$$

Kratzer gives the truth clause for $\geqslant$ using the lifted relation $\succeq_w^l$.

Definition (Truth)

Given a pointed world-ordering model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash_l \varphi$ as follows (with the other clauses as before):

$$\mathbf{M}, w \vDash_l \varphi \geqslant \psi \quad \text{iff} \quad [\![\varphi]\!]^{\mathbf{M}} \succeq_w^l [\![\psi]\!]^{\mathbf{M}}.$$

Kratzer Revisited

Fact

V1-V10 and V12 are all valid over world-ordering models according to Kratzer's semantics; V11 is not valid; I1-13 are all valid. *X*

Kratzer Revisited

Fact

V1-V10 and V12 are all valid over world-ordering models according to Kratzer's semantics; V11 is not valid; I1-13 are all valid. *X*

Theorem (Axiomatization of Kratzer's Semantics)

1. **WJ** is sound and complete with respect to the class of world-ordering models with Lewis's lifting.
2. **FJ** is sound and complete with respect to the class of total world-ordering models with Lewis's lifting.

Kratzer Revisited

Fact

V1-V10 and V12 are all valid over world-ordering models according to Kratzer's semantics; V11 is not valid; I1-13 are all valid. X

Theorem (Axiomatization of Kratzer's Semantics)

1. **WJ** is sound and complete with respect to the class of world-ordering models with Lewis's lifting.
2. **FJ** is sound and complete with respect to the class of total world-ordering models with Lewis's lifting.

Recall that **FJ** was the complete logic for Hamblin's semantics

Kratzer and Hamblin

We can think of Hamblin's semantics as almost the quantitative version of Kratzer's semantics, given this representation result:

Proposition

Given a set X, consider a relation $\succsim$ on $\wp(X)$.

1. If $\succsim = \succeq^l$ for a total preorder $\succeq$ on X, then there is a possibility measure ν on $\wp(X)$ such that

$$A \succsim B \text{ iff } \nu(A) \geq \nu(B).$$

2. If $\succsim = \succeq^l$ for a preorder $\succeq$ on X, then there is a possibility measure ν on $\wp(X)$ such that

$$A \succsim B \text{ implies } \nu(A) \geq \nu(B).$$

Kratzer and Hamblin

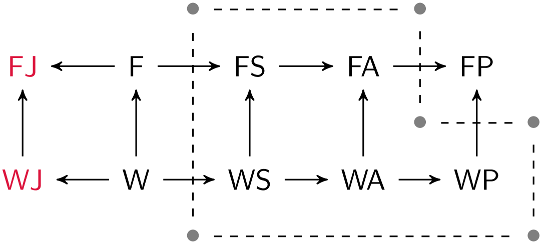

Figure : Logical Landscape

Kratzer Revisited

Definition (World-Ordering Model)

A (total) world-ordering model $\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V \rangle$ has for each $w \in W$ a (total) **preorder** $\succeq_w$ on $R(w)$.

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \forall b \in B_w \, \exists a \in A_w : a \succeq_w b.$$

Kratzer gives the truth clause for $\geqslant$ using the lifted relation $\succeq_w^l$.

Definition (Truth)

Given a pointed world-ordering model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash_l \varphi$ as follows (with the other clauses as before):

$$\mathbf{M}, w \vDash_l \varphi \geqslant \psi \quad \text{iff} \quad [\![\varphi]\!]^{\mathbf{M}} \succeq_w^l [\![\psi]\!]^{\mathbf{M}}.$$

Kratzer Revisited

Definition (World-Ordering Model)

A (total) world-ordering model $\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V \rangle$ has for each $w \in W$ a (total) **preorder** $\succeq_w$ on $R(w)$.

Following Lewis, we can lift $\succeq_w$ to a relation $\succeq_w^l$ on $\wp(W)$:

$$A \succeq_w^l B \text{ iff } \exists \text{ function } f\colon B_w \to A_w \text{ s.th. } \forall x \in B : f(x) \succeq_w x.$$

Kratzer gives the truth clause for $\geqslant$ using the lifted relation $\succeq_w^l$.

Definition (Truth)

Given a pointed world-ordering model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash_l \varphi$ as follows (with the other clauses as before):

$$\mathbf{M}, w \vDash_l \varphi \geqslant \psi \quad \text{iff} \quad [\![\varphi]\!]^{\mathbf{M}} \succeq_w^l [\![\psi]\!]^{\mathbf{M}}.$$

Alternative 4: A Better Lifting

Definition (World-Ordering Model)

A (total) world-ordering model $\mathbf{M} = \langle W, R, \{\succeq_w \mid w \in W\}, V\rangle$ has for each $w \in W$ a (total) preorder $\succeq_w$ on $R(w)$.

Here is a **better way to lift** $\succeq_w$ to a relation $\succeq_w^{\uparrow}$ on $\wp(W)$:

$$A \succeq_w^{\uparrow} B \text{ iff } \exists \text{ injection } f\colon B_w \to A_w \text{ s.th. } \forall x \in B : f(x) \succeq_w x.$$

Definition (Truth)

Given a pointed world-ordering model $\mathbf{M}, w$ and formula φ, we define $\mathbf{M}, w \vDash_{\uparrow} \varphi$ as follows (with the other clauses as before):

$$\mathbf{M}, w \vDash_{\uparrow} \varphi \geqslant \psi \quad \text{iff} \quad [\![\varphi]\!]^{\mathbf{M}} \succeq_w^{\uparrow} [\![\psi]\!]^{\mathbf{M}}.$$

Alternative 4: A Better Lifting

Given $a \succ b \succ c \succ d$, consider the liftings:

$$\begin{array}{ccccccccccccccccc}
abcd & \simeq^l & abc & \simeq^l & abd & \simeq^l & acd & \simeq^l & ab & \simeq^l & ac & \simeq^l & ad & \simeq^l & a & \succ^l \\
bcd & \simeq^l & bc & \simeq^l & bd & \simeq^l & b & \succ^l & cd & \simeq^l & c & \succ^l & d & \succ^l & \emptyset &
\end{array}$$

$$\begin{array}{ccccccccccccccccc}
abcd & \succ^{\uparrow} & abc & \succ^{\uparrow} & abd & \succ^{\uparrow} & acd & \succ^{\uparrow} & bcd & & & & & & \\
 & & & & \underset{\rightarrow}{\curlyvee} & & \underset{\rightarrow}{\curlyvee} & & \underset{\rightarrow}{\curlyvee} & & & & & & \\
 & & & & ab & \succ^{\uparrow} & ac & \succ^{\uparrow} & bc & & & & & & \\
 & & & & & & \underset{\rightarrow}{\curlyvee} & & \underset{\rightarrow}{\curlyvee} & & & & & & \\
 & & & & & & ad & \succ^{\uparrow} & bd & \succ^{\uparrow} & cd & & & & \\
 & & & & & & \underset{\rightarrow}{\curlyvee} & & \underset{\rightarrow}{\curlyvee} & & \underset{\rightarrow}{\curlyvee} & & & & \\
 & & & & & & a & \succ^{\uparrow} & b & \succ^{\uparrow} & c & \succ^{\uparrow} & d & \succ^{\uparrow} & \emptyset
\end{array}$$

Figure : Comparison of Lewis's lifting $\succeq^l$ and the new lifting $\succ^{\uparrow}$

Alternative 4: A Better Lifting

Here is a **better way to lift** $\succeq_w$ to a relation $\succeq_w^{\uparrow}$ on $\wp(W)$:

$$A \succeq^{\uparrow} B \text{ iff } \exists \text{ injective } f\colon B \to A \text{ s.th. } \forall x \in B : f(x) \succeq x$$

Proposition (Soundness)

WP *is sound with respect to the class of* path-finite[1] *world-ordering models with the* $\uparrow$ *lifting.*

[1] I.e., there is no infinite path $x_1 \preceq_w x_2 \preceq_w x_3 \ldots$ with $x_i \neq x_j$ for $i \neq j$.

Alternative 4: A Better Lifting

Here is a **better way to lift** $\succeq_w$ to a relation $\succeq_w^{\uparrow}$ on $\wp(W)$:

$$A \succeq^{\uparrow} B \text{ iff } \exists \text{ injective } f\colon B \to A \text{ s.th. } \forall x \in B : f(x) \succeq x$$

Proposition (Soundness)

WP *is sound with respect to the class of* path-finite[1] *world-ordering models with the* $\uparrow$ *lifting.*

Moral: simply changing Kratzer's semantics by requiring that the function be *injective* yields a logic of 'at least as likely as' that validates everything that the logic **FP** of full probability does, except the (controversial) totality axiom.

[1] I.e., there is no infinite path $x_1 \preceq_w x_2 \preceq_w x_3 \ldots$ with $x_i \neq x_j$ for $i \neq j$.

Alternative 4: A Better Lifting

Here is a **better way to lift** $\succeq_w$ to a relation $\succeq_w^{\uparrow}$ on $\wp(W)$:

$$A \succeq^{\uparrow} B \text{ iff } \exists \text{ injective } f\colon B \to A \text{ s.th. } \forall x \in B : f(x) \succeq x$$

Proposition (Soundness)

WP *is sound with respect to the class of* path-finite[2] *world-ordering models with the* $\uparrow$ *lifting.*

Trying to prove completeness is on our agenda. We know the complete logic for path-finite world-ordering models is below **FP**.

[2] I.e., there is no infinite path $x_1 \preceq_w x_2 \preceq_w x_3 \ldots$ with $x_i \neq x_j$ for $i \neq j$.

Alternative 4: A Better Lifting

Fact

Given any probability function μ on a set X, define a relation $\succeq$ on X by $x \succeq y$ iff $\mu(\{x\}) \geq \mu(\{y\})$. Then for any $A, B \subseteq X$,

$$A \succeq^{\uparrow} B \text{ implies } \mu(A) \geq \mu(B).$$

Alternative 4: A Better Lifting

Fact

Given any probability function μ on a set X, define a relation $\succeq$ on X by $x \succeq y$ iff $\mu(\{x\}) \geq \mu(\{y\})$. Then for any $A, B \subseteq X$,

$$A \succeq^{\uparrow} B \text{ implies } \mu(A) \geq \mu(B).$$

It is straightforward to construct orderings on worlds such that the lifted ordering $\succeq^{\uparrow}$ does not satisfy the problematic principles I1-I3 and E1. This shows that a semantics based on world-ordering models with a truth clause for $\geqslant$ stated in terms of $\succeq^{\uparrow}$ avoids the entailment problems raised for Kratzer's semantics.

System **WP**

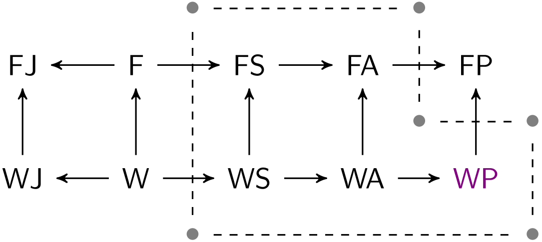

Figure : Logical Landscape

Summary of Results

We have seen four different kinds of semantics that yield the same results as the probability-based semantics with respect to Yalcin's list of intuitive validities and invalidities:

Summary of Results

We have seen four different kinds of semantics that yield the same results as the probability-based semantics with respect to Yalcin's list of intuitive validities and invalidities:

- self-dual fuzzy measure semantics;
- quasi-additive measure semantics;
- qualitative probability semantics;
- the semantics based on the lifting $\uparrow$.

Summary of Results

We have seen four different kinds of semantics that yield the same results as the probability-based semantics with respect to Yalcin's list of intuitive validities and invalidities:

- self-dual fuzzy measure semantics;
- quasi-additive measure semantics;
- qualitative probability semantics;
- the semantics based on the lifting $\uparrow$.

How do we decide between these semantics and the probability-based semantics?

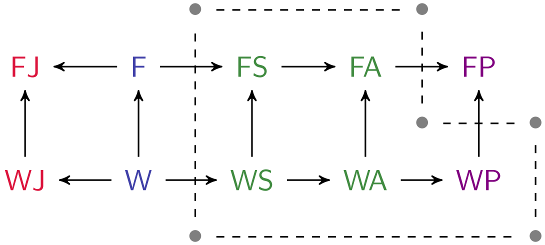

Figure : Logical Landscape

The diagram suggests the following way of thinking about the semantics for ‘at least as likely as’ and ‘probably’ that have been proposed: earlier proposals took off from **W** in the wrong direction. The new proposals head in the right direction, but the question is whether going all the way to **FP** is going too far.

Semantic Intuitions as Data

The standard data for semantic theory have traditionally been speakers' intuitions about entailment, implication, contradiction, validity, and other paradigmatic "semantic properties".

Semantic Intuitions as Data

The standard data for semantic theory have traditionally been speakers' intuitions about entailment, implication, contradiction, validity, and other paradigmatic "semantic properties".

This quotation from Chierchia & McConnell-Ginet's (2001) popular semantics textbook is characteristic:

> We are capable of assessing certain semantic properties of expressions and how two expressions are semantically related. These properties and relationships and the capacity that underlies our recognition of them constitute the empirical base of semantics. (52)

Semantic Intuitions as Data

Unfortunately, if we naïvely test speakers' intuitions about epistemic modals, we may not make it very far:

Semantic Intuitions as Data

Unfortunately, if we naïvely test speakers' intuitions about epistemic modals, we may not make it very far:

1. Keynes (1921) and many since, e.g., Gaifman (2009), have argued that **Tot** is not generally satisfied, nor should it be.

Semantic Intuitions as Data

Unfortunately, if we naïvely test speakers' intuitions about epistemic modals, we may not make it very far:

1. Keynes (1921) and many since, e.g., Gaifman (2009), have argued that **Tot** is not generally satisfied, nor should it be.
2. Tversky (1969), Fishburn (1983), and others have argued **Tran** is not always obeyed.

Semantic Intuitions as Data

Unfortunately, if we naïvely test speakers' intuitions about epistemic modals, we may not make it very far:

1. Keynes (1921) and many since, e.g., Gaifman (2009), have argued that **Tot** is not generally satisfied, nor should it be.
2. Tversky (1969), Fishburn (1983), and others have argued **Tran** is not always obeyed.
3. Tversky and Kahneman (1983) have famously argued people do not even obey **Mon**.

Semantic Intuitions as Data

Unfortunately, if we naïvely test speakers' intuitions about epistemic modals, we may not make it very far:

1. Keynes (1921) and many since, e.g., Gaifman (2009), have argued that **Tot** is not generally satisfied, nor should it be.
2. Tversky (1969), Fishburn (1983), and others have argued **Tran** is not always obeyed.
3. Tversky and Kahneman (1983) have famously argued people do not even obey **Mon**.

Are any non-trivial principles universally satisfied?

Semantic Intuitions as Data

Suppose for the moment that these experimental results can be explained away or otherwise dismissed, and we can justify, e.g., **WA**, on the basis of what should intuitively follow from what.

Semantic Intuitions as Data

Suppose for the moment that these experimental results can be explained away or otherwise dismissed, and we can justify, e.g., **WA**, on the basis of what should intuitively follow from what.

Question: What is the status of **FP**, and in particular the strong **Scott** axiom, with respect to ordinary semantic intuitions?

Semantic Intuitions as Data

Suppose for the moment that these experimental results can be explained away or otherwise dismissed, and we can justify, e.g., **WA**, on the basis of what should intuitively follow from what.

Question: What is the status of **FP**, and in particular the strong **Scott** axiom, with respect to ordinary semantic intuitions?

Many theorists have searched for the most intuitive principles that would guarantee an agreeing probability measure. Some theorists, e.g., Fine (1973), have argued that there are systems of inequalities that do not admit of an agreeing probability measure, but are in fact quite reasonable (c.f. Kraft, Pratt, and Seidenberg 1959).

Kraft et al.'s (counter)example

Where $\Omega = \{a, b, c, d, e\}$:

$$d \succ ac \qquad bc \succ ad \qquad ae \succ cd \qquad acd \succ be$$

Kraft et al.'s (counter)example

Where $\Omega = \{a, b, c, d, e\}$:

$$d \succ ac \qquad bc \succ ad \qquad ae \succ cd \qquad acd \succ be$$

1. X is more likely to be on Delta than American or Continental;

Kraft et al.'s (counter)example

Where $\Omega = \{a, b, c, d, e\}$:

$$d \succ ac \qquad bc \succ ad \qquad ae \succ cd \qquad acd \succ be$$

1. X is more likely to be on Delta than American or Continental;
2. British or Continental is more likely than American or Delta;

Kraft et al.'s (counter)example

Where $\Omega = \{a, b, c, d, e\}$:

$$d \succ ac \qquad bc \succ ad \qquad ae \succ cd \qquad acd \succ be$$

1. X is more likely to be on Delta than American or Continental;
2. British or Continental is more likely than American or Delta;
3. American or Emirates is more likely than Continental or Delta;

Kraft et al.'s (counter)example

Where $\Omega = \{a, b, c, d, e\}$:

$$d \succ ac \qquad bc \succ ad \qquad ae \succ cd \qquad acd \succ be$$

1. X is more likely to be on Delta than American or Continental;
2. British or Continental is more likely than American or Delta;
3. American or Emirates is more likely than Continental or Delta;
4. American, Continental, or Delta is more likely than British Air or United Arab Emirates.

Kraft et al.'s (counter)example

Where $\Omega = \{a, b, c, d, e\}$:

$$d \succ ac \qquad bc \succ ad \qquad ae \succ cd \qquad acd \succ be$$

1. X is more likely to be on Delta than American or Continental;
2. British or Continental is more likely than American or Delta;
3. American or Emirates is more likely than Continental or Delta;
4. American, Continental, or Delta is more likely than British Air or United Arab Emirates.

Fact

There is no probability measure that agrees with 1-4. In particular, this system of inequalities is inconsistent with **FP**.

Beyond Semantic Intuitions

Consider the following argument for **FP**:

Beyond Semantic Intuitions

Consider the following argument for **FP**:

1. When we use the words ‘at least as likely as’, ‘probable’, and so on, it is clear we are talking about chance and probability.

Beyond Semantic Intuitions

Consider the following argument for **FP**:

1. When we use the words ‘at least as likely as’, ‘probable’, and so on, it is clear we are talking about chance and probability.
2. The best theory of chance and probability is that given by the standard Kolmogorov axioms.

Beyond Semantic Intuitions

Consider the following argument for **FP**:

1. When we use the words ‘at least as likely as’, ‘probable’, and so on, it is clear we are talking about chance and probability.
2. The best theory of chance and probability is that given by the standard Kolmogorov axioms.
3. Therefore, Kolmogorovian probability captures what we mean by these words.

Beyond Semantic Intuitions

Consider the following argument for **FP**:

1. When we use the words 'at least as likely as', 'probable', and so on, it is clear we are talking about chance and probability.
2. The best theory of chance and probability is that given by the standard Kolmogorov axioms.
3. Therefore, Kolmogorovian probability captures what we mean by these words.

Bracketing the disagreement about additivity mentioned previously, what could be wrong with this rather commonsensical argument?

Quotation from Portner (2009):

> *Must* and *may* are widely attested in human language, and obviously existed before the development of a mathematical understanding of probability; in contrast, *there is a 60 percent probability that* expresses a meaning that had to be invented (or discovered) through the advancement of mathematical knowledge . . . . [I]t could be that *must* and *may* should be analyzed in terms of a non-mathematical theory, while *there is a 60 percent probability that* is to be understood in terms of a separate theory presupposing an additional modern mathematical apparatus. (73-74)

Quotation from Portner (2009):

> *Must* and *may* are widely attested in human language, and obviously existed before the development of a mathematical understanding of probability; in contrast, *there is a 60 percent probability that* expresses a meaning that had to be invented (or discovered) through the advancement of mathematical knowledge . . . . [I]t could be that *must* and *may* should be analyzed in terms of a non-mathematical theory, while *there is a 60 percent probability that* is to be understood in terms of a separate theory presupposing an additional modern mathematical apparatus. (73-74)

Background issue: Where does linguistic semantics stop and science, mathematics, philosophy, or other types of inquiry begin?

These issues are of course not unique to the logic of epistemic modality; nor are these questions new in semantics.

These issues are of course not unique to the logic of epistemic modality; nor are these questions new in semantics.

It may be instructive to consider related domains of discourse—for instance, talk about extensive properties like height, and time—and compare what considerations have motivated theorists in these areas to observe, or disregard, analogous assumptions.

Height

Since semanticists first started studying adjectives like 'tall', it has been assumed that the meaning is somehow related to points on a scale, such as a normalized measure function (Lewis 1970).

Height

Since semanticists first started studying adjectives like ‘tall’, it has been assumed that the meaning is somehow related to points on a scale, such as a normalized measure function (Lewis 1970).

Some early studies, such as Bartsch & Vennemann (1972), sought a *general* treatment of gradable adjectives, capable of explaining, e.g., how the positive form ‘tall’ and the comparative ‘taller than’ are related. They assumed this should extend to other gradable adjectives like ‘beautiful’, ‘intelligent’, and the like, which do not have obvious scales associated with them.

Height

Since semanticists first started studying adjectives like ‘tall’, it has been assumed that the meaning is somehow related to points on a scale, such as a normalized measure function (Lewis 1970).

Some early studies, such as Bartsch & Vennemann (1972), sought a *general* treatment of gradable adjectives, capable of explaining, e.g., how the positive form ‘tall’ and the comparative ‘taller than’ are related. They assumed this should extend to other gradable adjectives like ‘beautiful’, ‘intelligent’, and the like, which do not have obvious scales associated with them.

Cresswell (1977) addressed the issue explicitly:

> Whether [$\succ$] should be strict or not or total or not seems unimportant, and perhaps we should be liberal enough not to insist on transitivity or antisymmetry. (266)

An Extreme View

In one of the earliest discussions, Wheeler (1972) went further:

> Semantics, as we see it, is solely concerned with finding out what the forms of sentences in English are. When we have found where the predicates are, semantics is finished. It is certainly a worthwhile project, when semantics is done, to state some truths using the predicates the semantics has arrived at, but this is to do science, not semantics. . . . The tendency we oppose is the tendency to turn high-level truths into *analytic* truths; to build information into a theory of a language; to treat languages as first-order theories rather than as first-order languages. (319)

Scale types

In the mean time, the situation has become more complicated, and purely grammatical considerations have motivated linguists to posit more structure on the underlying domains.

Scale types

In the mean time, the situation has become more complicated, and purely grammatical considerations have motivated linguists to posit more structure on the underlying domains.

For instance, Kennedy (2007) has argued that many adjectives can be classified on the basis of whether they form grammatical expressions when combined with modifiers like ‘perfectly’, ‘slightly’, or ‘completely’. This leads to a classification of scale types, specifying such properties as closed, open, and bounded.

Scale types

In the mean time, the situation has become more complicated, and purely grammatical considerations have motivated linguists to posit more structure on the underlying domains.

For instance, Kennedy (2007) has argued that many adjectives can be classified on the basis of whether they form grammatical expressions when combined with modifiers like ‘perfectly’, ‘slightly’, or ‘completely’. This leads to a classification of scale types, specifying such properties as closed, open, and bounded.

Classic work in the Theory of Measurement, as explicated in Krantz, Luce, Suppes, and Tversky (1971), has collected a number of representation theorems for extensive measurement. It remains to be seen whether purely linguistic, or semantic, considerations motivate the need for real number scales, say, for height.

Time

There is a long, distinguished history of work in the logic of time, and the references to time in language are many and varied.

Time

There is a long, distinguished history of work in the logic of time, and the references to time in language are many and varied.

If we focus on a simple language with F and P, interpreted with respect to a temporal precedence order $\prec$, we can play a game very much like we did for epistemic modals.

Time

There is a long, distinguished history of work in the logic of time, and the references to time in language are many and varied.

If we focus on a simple language with F and P, interpreted with respect to a temporal precedence order $\prec$, we can play a game very much like we did for epistemic modals.

Which "intuitive" principles involving the future and the past can we state, to ensure certain properties of $\prec$?

Time

There is a long, distinguished history of work in the logic of time, and the references to time in language are many and varied.

If we focus on a simple language with F and P, interpreted with respect to a temporal precedence order $\prec$, we can play a game very much like we did for epistemic modals.

Which "intuitive" principles involving the future and the past can we state, to ensure certain properties of $\prec$?

For instance, the statement $P\top$ corresponds to "having no beginning point", while $F\top$ corresponds to "having no end point".

Time

There is a long, distinguished history of work in the logic of time, and the references to time in language are many and varied.

If we focus on a simple language with F and P, interpreted with respect to a temporal precedence order $\prec$, we can play a game very much like we did for epistemic modals.

Which "intuitive" principles involving the future and the past can we state, to ensure certain properties of $\prec$?

For instance, the statement $P\top$ corresponds to "having no beginning point", while $F\top$ corresponds to "having no end point".

(Bach (1986): "Are questions about the Big Bang *linguistic* questions?")

Temporal Ontology

Some authors have been rather insistent that natural language semantics is independent of considerations about how time really is. Mark Steedman (1997), for instance, says:

> As in any epistemological domain, neither the ontology nor the relations should be confused with the corresponding descriptors that we use to define the physics and mechanics of the real world. The notion of time that is reflected in linguistic categories is only indirectly related to common-sense physics of clock-time and the related Newtonian representation of it as a dimension comprising an infinite number of instants corresponding to the real numbers, still less to the more abstruse representation of time in modern physics. (925)

Conclusion

- We have seen several classes of semantic models, and their associated logics, which overcome the entailment problems for previous accounts of epistemic modals.

Conclusion

- We have seen several classes of semantic models, and their associated logics, which overcome the entailment problems for previous accounts of epistemic modals.
- In light of this, the question naturally arises: why might we prefer one system over another? In particular, do we have reason to prefer **FP** over its weaker fragments?

Conclusion

- There are strategies that might lead one to **FP**. However, these go beyond what we called the "standard" methodology in linguistic semantics of relying on ordinary speakers' intuitions about what follows from what.

Conclusion

- There are strategies that might lead one to **FP**. However, these go beyond what we called the "standard" methodology in linguistic semantics of relying on ordinary speakers' intuitions about what follows from what.
- In the analogous domains of height and time, there has been resistance to go too far beyond what seems necessary for systematizing semantic or grammatical intuitions. It is an interesting to ask how epistemic modality might be different.

- One might want a semantic account:
  - to provide a reasonable approximation to what we have in our heads, or of what underlies our communicative behavior;
  - to capture the range of claims we can make about the world, in science or otherwise;
  - to endow an automated agent with the ability to use and process and natural language;
  - ...

- One might want a semantic account:
  - to provide a reasonable approximation to what we have in our heads, or of what underlies our communicative behavior;
  - to capture the range of claims we can make about the world, in science or otherwise;
  - to endow an automated agent with the ability to use and process and natural language;
  - ...

- To that extent, there may be no substantive disagreement between semanticists who prefer stronger or weaker systems.

- One might want a semantic account:
  - to provide a reasonable approximation to what we have in our heads, or of what underlies our communicative behavior;
  - to capture the range of claims we can make about the world, in science or otherwise;
  - to endow an automated agent with the ability to use and process and natural language;
  - ...
- To that extent, there may be no substantive disagreement between semanticists who prefer stronger or weaker systems. It may be misleading to speak of "the" logic of epistemic modals, since different projects call for different methodology, which may lead to different conclusions about validity.

- One might want a semantic account:
  - to provide a reasonable approximation to what we have in our heads, or of what underlies our communicative behavior;
  - to capture the range of claims we can make about the world, in science or otherwise;
  - to endow an automated agent with the ability to use and process and natural language;
  - ...

- To that extent, there may be no substantive disagreement between semanticists who prefer stronger or weaker systems. It may be misleading to speak of "the" logic of epistemic modals, since different projects call for different methodology, which may lead to different conclusions about validity.

- At any rate, we hope to have made clear the landscape of options for the semanticist.

Thank you!

Theorem (Representation Theorem, Scott 1964)

If $(W, \succeq)$ *satisfies the axioms of FP, then there is a probability measure* $\nu : \wp(W) \to [0,1]$ *such that:*

if $A \succ B$ *then* $\nu(A) > \nu(B)$*, and if* $A \sim B$ *then* $\nu(A) = \nu(B)$.

Theorem (Representation Theorem, Scott 1964)

If $(W, \succeq)$ satisfies the axioms of FP, then there is a probability measure $\nu : \wp(W) \to [0, 1]$ such that:

if $A \succ B$ then $\nu(A) > \nu(B)$, and if $A \sim B$ then $\nu(A) = \nu(B)$.

Proof.

Finite case. Each $A \in \wp(W)$ can be associated with a vector $\overline{A} \in \{0, 1\}^n$, with $|W| = n$, the "characteristic function" of A.

Theorem (Representation Theorem, Scott 1964)

If $(W, \succeq)$ satisfies the axioms of FP, then there is a probability measure $\nu : \wp(W) \to [0, 1]$ such that:

$$\text{if } A \succ B \text{ then } \nu(A) > \nu(B), \text{ and if } A \sim B \text{ then } \nu(A) = \nu(B).$$

Proof.

Finite case. Each $A \in \wp(W)$ can be associated with a vector $\overline{A} \in \{0, 1\}^n$, with $|W| = n$, the "characteristic function" of A.

Let Γ be the set of strict inequalities $A \succ B$, and Σ the set of equivalences $A \sim B$. For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

□

Proof.

For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

Proof.

For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

Lemma

There exists $c \in \mathbb{R}^n$ such that $c \cdot \overline{\gamma} > 0$ for all $\gamma \in \Gamma$, and $c \cdot \overline{\sigma} = 0$ for all $\sigma \in \Sigma$.

Proof.

For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

Lemma

There exists $c \in \mathbb{R}^n$ such that $c \cdot \overline{\gamma} > 0$ for all $\gamma \in \Gamma$, and $c \cdot \overline{\sigma} = 0$ for all $\sigma \in \Sigma$.

Given this lemma, we set:

$$\nu(A) = \frac{c \cdot \overline{A}}{c \cdot \overline{W}}.$$

Proof.

For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

Lemma

There exists $c \in \mathbb{R}^n$ such that $c \cdot \overline{\gamma} > 0$ for all $\gamma \in \Gamma$, and $c \cdot \overline{\sigma} = 0$ for all $\sigma \in \Sigma$.

Given this lemma, we set:

$$\nu(A) = \frac{c \cdot \overline{A}}{c \cdot \overline{W}}.$$

Then:

- If $A \succ B$, then by the lemma, $\nu(A) > \nu(B)$;

Proof.

For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

Lemma

There exists $c \in \mathbb{R}^n$ such that $c \cdot \overline{\gamma} > 0$ for all $\gamma \in \Gamma$, and $c \cdot \overline{\sigma} = 0$ for all $\sigma \in \Sigma$.

Given this lemma, we set:

$$\nu(A) = \frac{c \cdot \overline{A}}{c \cdot \overline{W}}.$$

Then:

- If $A \succ B$, then by the lemma, $\nu(A) > \nu(B)$;
- If $A \sim B$, then again by the lemma, $\nu(A) = \nu(B)$.

Proof.

For $\gamma = A \succ B$, and $\sigma = A \sim B$, let

$$\overline{\gamma} = \overline{A} - \overline{B}, \text{ and } \overline{\sigma} = \overline{A} - \overline{B}.$$

Lemma

There exists $c \in \mathbb{R}^n$ such that $c \cdot \overline{\gamma} > 0$ for all $\gamma \in \Gamma$, and $c \cdot \overline{\sigma} = 0$ for all $\sigma \in \Sigma$.

Given this lemma, we set:

$$\nu(A) = \frac{c \cdot \overline{A}}{c \cdot \overline{W}}.$$

Then:

- If $A \succ B$, then by the lemma, $\nu(A) > \nu(B)$;
- If $A \sim B$, then again by the lemma, $\nu(A) = \nu(B)$.
- Showing ν is a probability measure is easy.