Reasoning about Knowledge and Beliefs

Lecture 7

Eric Pacuit

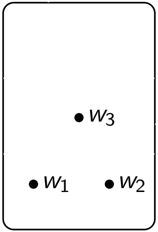
University of Maryland, College Park

pacuit.org
epacuit@umd.edu

October 2, 2013

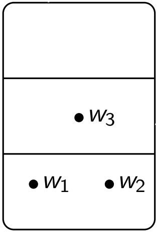
Beliefs via Plausibility

- $W = \{w_1, w_2, w_3\}$



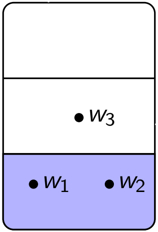
Beliefs via Plausibility

- $W = \{w_1, w_2, w_3\}$
- $w_1 \preceq w_2$ and $w_2 \preceq w_1$ (w_1 and w_2 are equi-plausbile)
- $w_1 \prec w_3$ ($w_1 \preceq w_3$ and $w_3 \not\preceq w_1$)
- $w_2 \prec w_3$ ($w_2 \preceq w_3$ and $w_3 \not\preceq w_2$)

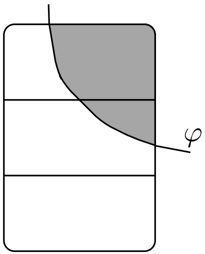


Beliefs via Plausibility

- $W = \{w_1, w_2, w_3\}$
- $w_1 \preceq w_2$ and $w_2 \preceq w_1$ (w_1 and w_2 are equi-plausbile)
- $w_1 \prec w_3$ ($w_1 \preceq w_3$ and $w_3 \not\preceq w_1$)
- $w_2 \prec w_3$ ($w_2 \preceq w_3$ and $w_3 \not\preceq w_2$)
- $\{w_1, w_2\} \subseteq Min_{\preceq}([w_i])$

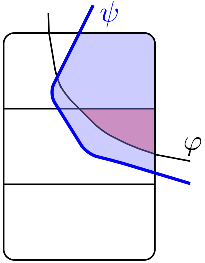


Beliefs via Plausibility



Conditional Belief: $B^{\varphi}\psi$

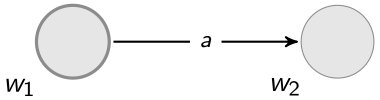
Beliefs via Plausibility



Conditional Belief: $B^{\varphi}\psi$

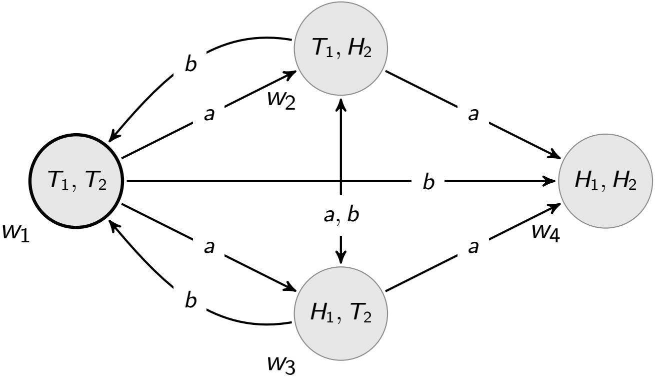
$$Min_{\preceq}([\![\varphi]\!]_{\mathcal{M}}) \subseteq [\![\psi]\!]_{\mathcal{M}}$$

Example

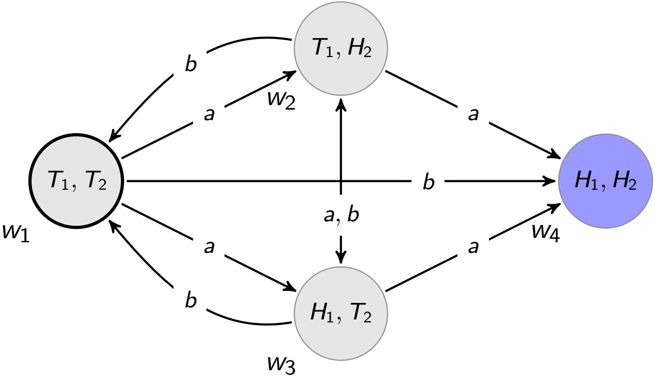


$$w_2 \preceq_b w_1$$

Example

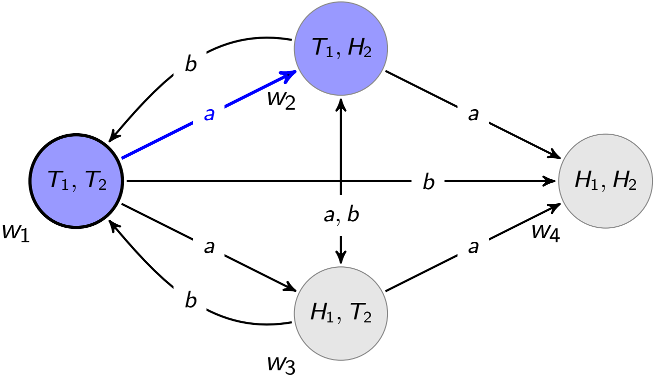


Example



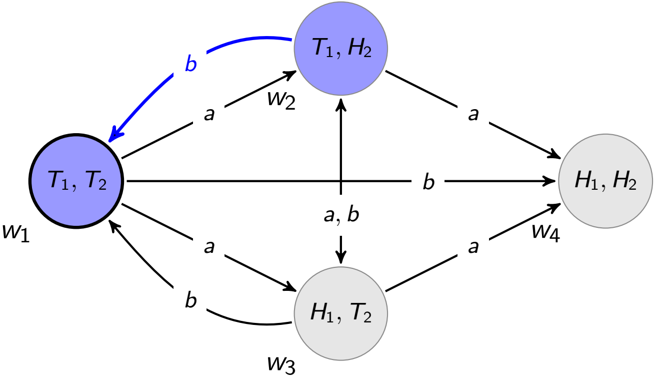
- $w_1 \models B_a(H_1 \wedge H_2) \wedge B_b(H_1 \wedge H_2)$

Example



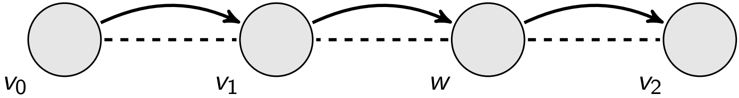
- $w_1 \models B_a(H_1 \wedge H_2) \wedge B_b(H_1 \wedge H_2)$
- $w_1 \models B_a^{T_1} H_2$

Example

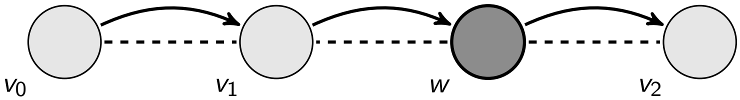


- $w_1 \models B_a(H_1 \wedge H_2) \wedge B_b(H_1 \wedge H_2)$
- $w_1 \models B_a^{T_1} H_2$
- $w_1 \models B_b^{T_1} T_2$

Grades of Doxastic Strength

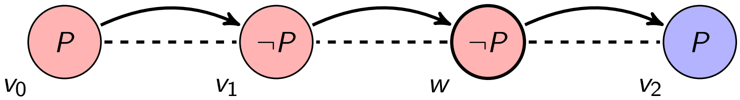


Grades of Doxastic Strength



Suppose that w is the current state.

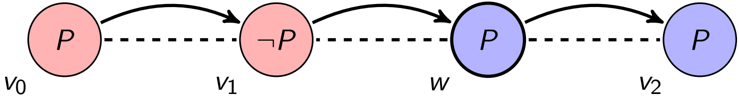
Grades of Doxastic Strength



Suppose that w is the current state.

- **Belief** (BP)

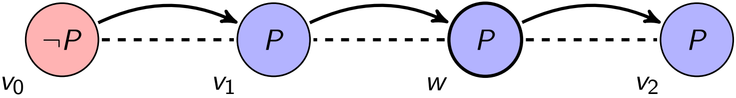
Grades of Doxastic Strength



Suppose that w is the current state.

- **Belief** (BP)
- **Robust Belief** ($[\preceq]P$)

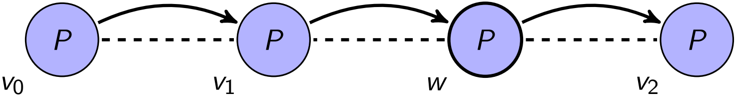
Grades of Doxastic Strength



Suppose that w is the current state.

- **Belief** (BP)
- **Robust Belief** ($[\preceq]P$)
- **Strong Belief** ($B^s P$)

Grades of Doxastic Strength



Suppose that w is the current state.

- **Belief** (BP)
- **Robust Belief** ($[\preceq]P$)
- **Strong Belief** ($B^s P$)
- **Knowledge** (KP)

Is $B\varphi \to B^{\psi}\varphi$ valid?

Is $B\varphi \to B^{\psi}\varphi$ valid?

Is $B^{\alpha}\varphi \to B^{\alpha\wedge\beta}\varphi$ valid?

Is $B\varphi \to B^{\psi}\varphi$ valid?

Is $B^{\alpha}\varphi \to B^{\alpha \wedge \beta}\varphi$ valid?

Is $B\varphi \to B^{\psi}\varphi \vee B^{\neg\psi}\varphi$ valid?

Is $B\varphi \rightarrow B^{\psi}\varphi$ valid?

Is $B^{\alpha}\varphi \rightarrow B^{\alpha\wedge\beta}\varphi$ valid?

Is $B\varphi \rightarrow B^{\psi}\varphi \vee B^{\neg\psi}\varphi$ valid?

Exercise: Prove that B, B^{φ} and B^{s} are definable in the language with K and $[\preceq]$ modalities.

$\mathcal{M}, w \models B^{\varphi}\psi$ if for each $v \in Min_{\preceq}([w] \cap [\![\varphi]\!])$, $\mathcal{M}, v \models \varphi$
where $[\![\varphi]\!] = \{w \mid \mathcal{M}, w \models \varphi\}$ and $[w] = \{v \mid w \sim v\}$

$\mathcal{M}, w \models B^{\varphi}\psi$ if for each $v \in Min_{\preceq}([w] \cap [\![\varphi]\!])$, $\mathcal{M}, v \models \varphi$
where $[\![\varphi]\!] = \{w \mid \mathcal{M}, w \models \varphi\}$ and $[w] = \{v \mid w \sim v\}$

Core Logical Principles:

1. $B^{\varphi}\varphi$
2. $B^{\varphi}\psi \rightarrow B^{\varphi}(\psi \vee \chi)$
3. $(B^{\varphi}\psi_1 \wedge B^{\varphi}\psi_2) \rightarrow B^{\varphi}(\psi_1 \wedge \psi_2)$
4. $(B^{\varphi_1}\psi \wedge B^{\varphi_2}\psi) \rightarrow B^{\varphi_1 \vee \varphi_2}\psi$
5. $(B^{\varphi}\psi \wedge B^{\psi}\varphi) \rightarrow (B^{\varphi}\chi \leftrightarrow B^{\psi}\chi)$

J. Burgess. *Quick completeness proofs for some logics of conditionals*. *Notre Dame Journal of Formal Logic 22*, 76 – 84, 1981.

Types of Beliefs: Logical Characterizations

- $\mathcal{M}, w \models K_i\varphi$ iff $\mathcal{M}, w \models B_i^{\psi}\varphi$ for all ψ
  i knows φ iff i continues to believe φ given any new information

Types of Beliefs: Logical Characterizations

- $\mathcal{M}, w \models K_i\varphi$ iff $\mathcal{M}, w \models B_i^{\psi}\varphi$ for all ψ
  i knows φ iff *i* continues to believe φ given any new information

- $\mathcal{M}, w \models [\preceq_i]\varphi$ iff $\mathcal{M}, w \models B_i^{\psi}\varphi$ for all ψ with $\mathcal{M}, w \models \psi$.
  i robustly believes φ iff *i* continues to believe φ given any true formula.

Types of Beliefs: Logical Characterizations

- $\mathcal{M}, w \models K_i\varphi$ iff $\mathcal{M}, w \models B_i^{\psi}\varphi$ for all ψ
  i knows φ iff *i* continues to believe φ given any new information

- $\mathcal{M}, w \models [\preceq_i]\varphi$ iff $\mathcal{M}, w \models B_i^{\psi}\varphi$ for all ψ with $\mathcal{M}, w \models \psi$.
  i robustly believes φ iff *i* continues to believe φ given any true formula.

- $\mathcal{M}, w \models B_i^{s}\varphi$ iff $\mathcal{M}, w \models B_i\varphi$ and $\mathcal{M}, w \models B_i^{\psi}\varphi$ for all ψ with $\mathcal{M}, w \models \neg K_i(\psi \rightarrow \neg\varphi)$.
  i strongly believes φ iff *i* believes φ and continues to believe φ given any evidence (truthful or not) that is not known to contradict φ.

Additional Axioms

Success: $B_i^{\varphi}\varphi$

Additional Axioms

Success:	$B_i^{\varphi}\varphi$
Knowledge entails belief	$K_i\varphi \rightarrow B_i^{\psi}\varphi$

Additional Axioms

Success:	$B_i^{\varphi}\varphi$
Knowledge entails belief	$K_i\varphi \to B_i^{\psi}\varphi$
Full introspection:	$B_i^{\varphi}\psi \to K_i B_i^{\varphi}\psi$ and $\neg B_i^{\varphi}\psi \to K_i \neg B_i^{\varphi}\psi$

Additional Axioms

Success:	$B_i^{\varphi}\varphi$
Knowledge entails belief	$K_i\varphi \to B_i^{\psi}\varphi$
Full introspection:	$B_i^{\varphi}\psi \to K_i B_i^{\varphi}\psi$ and $\neg B_i^{\varphi}\psi \to K_i \neg B_i^{\varphi}\psi$
Cautious Monotonicity:	$(B_i^{\varphi}\alpha \wedge B_i^{\varphi}\beta) \to B_i^{\varphi\wedge\beta}\alpha$

Additional Axioms

Success:	$B_i^{\varphi}\varphi$
Knowledge entails belief	$K_i\varphi \to B_i^{\psi}\varphi$
Full introspection:	$B_i^{\varphi}\psi \to K_i B_i^{\varphi}\psi$ and $\neg B_i^{\varphi}\psi \to K_i \neg B_i^{\varphi}\psi$
Cautious Monotonicity:	$(B_i^{\varphi}\alpha \wedge B_i^{\varphi}\beta) \to B_i^{\varphi\wedge\beta}\alpha$
Rational Monotonicity:	$(B_i^{\varphi}\alpha \wedge \neg B_i^{\varphi}\neg\beta) \to B_i^{\varphi\wedge\beta}\alpha$

Additional Axioms

Success:	$B_i^{\varphi}\varphi$
Knowledge entails belief	$K_i\varphi \rightarrow B_i^{\psi}\varphi$
Full introspection:	$B_i^{\varphi}\psi \rightarrow K_iB_i^{\varphi}\psi$ and $\neg B_i^{\varphi}\psi \rightarrow K_i\neg B_i^{\varphi}\psi$
Cautious Monotonicity:	$(B_i^{\varphi}\alpha \wedge B_i^{\varphi}\beta) \rightarrow B_i^{\varphi\wedge\beta}\alpha$
Rational Monotonicity:	$(B_i^{\varphi}\alpha \wedge \neg B_i^{\varphi}\neg\beta) \rightarrow B_i^{\varphi\wedge\beta}\alpha$

Rational Monotonicity, I

Rational Monotonicity: $(B_i^{\varphi}\alpha \wedge \neg B_i^{\varphi}\neg\beta) \rightarrow B_i^{\varphi\wedge\beta}\alpha$

R. Stalnaker. *Nonmonotonic consequence relations*. Fundamenta Informaticae, 21: 721, 1994.

Rational Monotonicity, I

Rational Monotonicity: $(B_i^{\varphi}\alpha \wedge \neg B_i^{\varphi}\neg\beta) \rightarrow B_i^{\varphi\wedge\beta}\alpha$

R. Stalnaker. *Nonmonotonic consequence relations*. Fundamenta Informaticae, 21: 721, 1994.

Consider the three composers: Verdi, Bizet, and Satie, and suppose that we initially accept (correctly but defeasibly) that Verdi is Italian $I(v)$, while Bizet and Satie are French $(F(b) \wedge F(s))$.

Rational Monotonicity, II

Suppose now that we are told by a reliable (but not infallible!) source of information that that Verdi and Bizet are compatriots ($C(v,b)$). This leads us no longer to endorse either the proposition that Verdi is Italian (because he could be French), or that Bizet is French (because he could be Italian); but we would still draw the defeasible consequence that Satie is French, since nothing that we have learned conflicts with it.

$$B^{C(v,b)}F(s)$$

Rational Monotonicity, III

Now consider the proposition $C(v, s)$ that Verdi and Satie are compatriots. Before learning that $C(v, b)$ we would be inclined to reject the proposition $C(v, s)$ because we accept $I(v)$ and $F(s)$, but after learning that Verdi and Bizet are compatriots, we can no longer endorse $I(v)$, and therefore no longer reject $C(v, s)$.

$$\neg B^{C(v,b)} \neg C(v, s)$$

Rational Monotonicity, IV

However, if we added $C(v, s)$ to our stock of beliefs, we would lose the inference to $F(s)$: in the context of $C(v, b)$, the proposition $C(v, s)$ is equivalent to the statement that all three composers have the same nationality. This leads us to suspend our belief in the proposition $F(s)$.

$$\neg B^{C(v,b) \wedge C(v,s)} F(s)$$

Rational Monotonicity, IV

However, if we added $C(v,s)$ to our stock of beliefs, we would lose the inference to $F(s)$: in the context of $C(v,b)$, the proposition $C(v,s)$ is equivalent to the statement that all three composers have the same nationality. This leads us to suspend our belief in the proposition $F(s)$.

$$\neg B^{C(v,b)\wedge C(v,s)}F(s)$$

$B^{C(v,b)}F(s)$ and $\neg B^{C(v,b)}\neg C(v,s)$ but $\neg B^{C(v,b)\wedge C(v,s)}F(s)$