# Logic and Artificial Intelligence

## Lecture 17

Eric Pacuit

Currently Visiting the Center for Formal Epistemology, CMU

Center for Logic and Philosophy of Science
Tilburg University

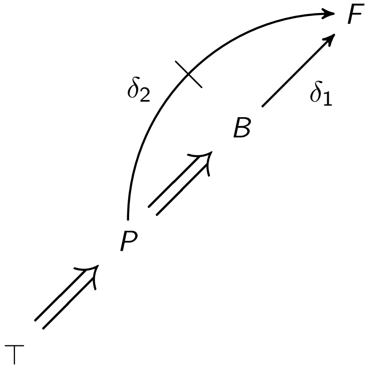
ai.stanford.edu/∼epacuit
e.j.pacuit@uvt.nl

November 3, 2011

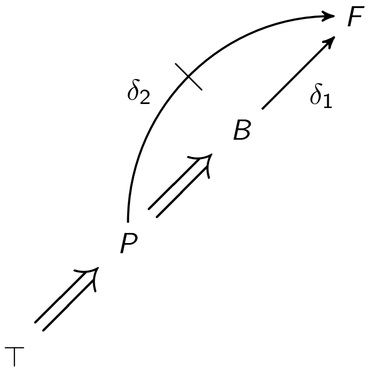
# A Distinction

Rebutting vs. undercutting

# Tweety Triangle: Rebutting



# Tweety Triangle: Rebutting




$\delta_1$ is the default: $B \longrightarrow F$
$\delta_2$ is the default $B \longrightarrow \neg F$
There is a *conflict* in that $Conclusion(\delta_1)$ is inconsistent with $Conclusion(\delta_2)$

# Undercutting, I

Suppose an object in front of me looks red. Then it is reasonable for me to conclude that it is red, through an application of a general default according to which things that look red tend to be red.

# Undercutting, I

Suppose an object in front of me looks red. Then it is reasonable for me to conclude that it is red, through an application of a general default according to which things that look red tend to be red.

$$L \longrightarrow R$$

# Undercutting, I

Suppose an object in front of me looks red. Then it is reasonable for me to conclude that it is red, through an application of a general default according to which things that look red tend to be red.

$$L \longrightarrow R$$

Suppose that a reliable source — so reliable that I trust this source more than my own sense perception — might inform me that the object is not, in fact, red. (We can suppose that, as I am well aware, the reliable source may know things that I do not: the object could be illuminated by red lights, a red screen could have been inserted in front of the object, and so on.)

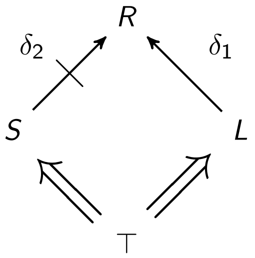
# Undercutting, I

Suppose an object in front of me looks red. Then it is reasonable for me to conclude that it is red, through an application of a general default according to which things that look red tend to be red.

$$L \longrightarrow R$$

Suppose that a reliable source — so reliable that I trust this source more than my own sense perception — might inform me that the object is not, in fact, red. (We can suppose that, as I am well aware, the reliable source may know things that I do not: the object could be illuminated by red lights, a red screen could have been inserted in front of the object, and so on.)

$$S \longrightarrow \neg R$$




$\delta_1 < \delta_2$

$\delta_2 < \delta_1$

# Undercutting, II

If the object looks red but the reliable source tells me otherwise, then it is natural to appeal to another default, with the conclusion that the object is not red, since what the reliable source says tends to be true and the reliable source has told me that it is not red. And because, by hypothesis, the reliable source is more reliable than perception, this new default would have to be stronger than the original, that whatever looks red tends to be red, and so would defeat this original default in the sense we have considered so far, by providing a stronger reason for a conflicting conclusion.

# Undercutting, II

If the object looks red but the reliable source tells me otherwise, then it is natural to appeal to another default, with the conclusion that the object is not red, since what the reliable source says tends to be true and the reliable source has told me that it is not red. And because, by hypothesis, the reliable source is more reliable than perception, this new default would have to be stronger than the original, that whatever looks red tends to be red, and so would defeat this original default in the sense we have considered so far, by providing a stronger reason for a conflicting conclusion.

$\delta_2 < \delta_1$, so the agent believes the object is not red.

# Undercutting, III

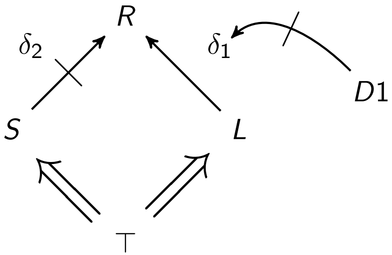
Suppose that I have taken a drug — let us call it Drug 1 — that makes everything looks red. If the object looks red but I have taken Drug 1, on the other hand, then it seems again that I am no longer entitled to the conclusion that the object is red. But in this case, the original default is not defeated in the same way. There is no stronger reason for concluding that the object is not red; instead, it is as if the favoring relation represented by the original default is itself severed, so that what was once a reason no longer provides any support for its conclusion.

# Undercutting, III

Suppose that I have taken a drug — let us call it Drug 1 — that makes everything looks red. If the object looks red but I have taken Drug 1, on the other hand, then it seems again that I am no longer entitled to the conclusion that the object is red. But in this case, the original default is not defeated in the same way. There is no stronger reason for concluding that the object is not red; instead, it is as if the favoring relation represented by the original default is itself severed, so that what was once a reason no longer provides any support for its conclusion.

$$D1 \longrightarrow Out(\delta_1)$$

# Undercutting, IV



When should a rational agent believe/accept/assert and *conditional*? (If $A$, then $B$: $A > B$)

Note: $A > B$ is NOT the same as $A \Rightarrow B$

# Ramsey Test

If two people are arguing 'If $p$, then $q$?' and are both in doubt as to $p$, they are adding $p$ hypothetically to their stock of knowledge and arguing on that basis about $q$; so that in a sense 'If $p$, $q$' and 'If $p$, $\neg q$' are contradictories. We can say that they are fixing their degree of belief in $q$ given $p$. If $p$ turns out false, these degrees of belief are rendered void. If either party believes not $p$ for certain, the question ceases to mean anything to him except as a question about what follows from certain laws or hypotheses.

F. P. Ramsey. *General Propositions and Causality*. 1929.

# AGM and the Ramsey Test

Recall that $\mathcal{L}$ is a propositional language. Let $\mathcal{L} \subseteq \mathcal{L}_{<}$ extends $\mathcal{L}$ with a conditional operator '$>$': $A > B$ means 'If $A$, $B$'

# AGM and the Ramsey Test

Recall that $\mathcal{L}$ is a propositional language. Let $\mathcal{L} \subseteq \mathcal{L}_{<}$ extends $\mathcal{L}$ with a conditional operator '$>$': $A > B$ means 'If $A$, $B$'

$$(R) \qquad A > B \in K \text{ iff } B \in K * A$$

# AGM and the Ramsey Test

Recall that $\mathcal{L}$ is a propositional language. Let $\mathcal{L} \subseteq \mathcal{L}_<$ extends $\mathcal{L}$ with a conditional operator '$>$': $A > B$ means 'If $A$, $B$'

$$(R) \qquad A > B \in K \text{ iff } B \in K * A$$

Belief revision model: $\langle \mathbf{K}, * \rangle$ where $* : \mathbf{K} \times \mathcal{L} \to \mathbf{K}$.

## Gärdenfors Impossibility Result, I

For each $A \in \mathcal{L}$, write $K_A$ for $K * A$.

# Gärdenfors Impossibility Result, I

For each $A \in \mathcal{L}$, write $K_A$ for $K * A$.

**Assumptions**

# Gärdenfors Impossibility Result, I

For each $A \in \mathcal{L}$, write $K_A$ for $K * A$.

**Assumptions**

(1) $A \in K_A$

# Gärdenfors Impossibility Result, I

For each $A \in \mathcal{L}$, write $K_A$ for $K * A$.

**Assumptions**

(1) $A \in K_A$

(2) If $K \neq K_\perp$ and $K_A = K_\perp$ then $\vdash \neg A$

# Gärdenfors Impossibility Result, I

For each $A \in \mathcal{L}$, write $K_A$ for $K * A$.

**Assumptions**

(1) $A \in K_A$

(2) If $K \neq K_\perp$ and $K_A = K_\perp$ then $\vdash \neg A$

(R) $A > B \in K$ iff $B \in K * A$

# Gärdenfors Impossibility Result, II

From (R) we can conclude

(M) For all belief sets $K$, $K'$ and propositions $A$, $K \subseteq K'$ implies $K_A \subseteq K'_A$

## Gärdenfors Impossibility Result, II

From (R) we can conclude

(M) For all belief sets $K$, $K'$ and propositions $A$, $K \subseteq K'$ implies $K_A \subseteq K'_A$

**Preservation Criterion**

(P) If $\neg A \not\in K$ and $B \in K$ then $B \in K_A$

# Gärdenfors Impossibility Result, III

Let $K/A = Cn(K \cup \{A\})$, We have

# Gärdenfors Impossibility Result, III

Let $K/A = Cn(K \cup \{A\})$, We have

(3) $(K/A)/B = K/(A \wedge B)$
(4) $K/(A \vee B) \subseteq K/A$

# Gärdenfors Impossibility Result, III

Let $K/A = Cn(K \cup \{A\})$, We have

(3) $(K/A)/B = K/(A \wedge B)$
(4) $K/(A \vee B) \subseteq K/A$

From (1) and (P)

(I) If $\neg A \notin K$ then $K/A \subseteq K_A$

# Gärdenfors Impossibility Result, IV

Two propositions $A$ and $B$ are *disjoint* provided $\vdash \neg(A \wedge B)$

A belief revision model $\langle \mathbf{K}, \mathbf{F} \rangle$ is **nontrivial** provided there is a $K \in \mathbf{K}$ and *pairwise disjoint* propositions $A, B$ and $C$ such that $\neg A \notin K$, $\neg B \notin K$ and $\neg C \notin K$.

**Theorem** There is no nontrivial belief revision model that satisfies all the conditions (1), (2), (M) and (P).

P. Gärdenfors. *Belief Revision and the Ramsey Test*. Philosophical Review, Vol. 95, pp. 81 - 93, 1986.

# Gärdenfors Impossibility Result: Proof

**Nontrivial** There are $A$, $B$, and $C$ such that $\vdash \neg(A \wedge B)$, $\vdash \neg(B \wedge C)$, $\vdash \neg(A \wedge C)$ and belief set $K$ such that $\neg A \notin K$, $\neg B \notin K$ and $\neg C \notin K$.

(1) $A \in K_A$

(2) If $K \neq K_\perp$ and $K_A = K_\perp$ then $\vdash \neg A$

(R) $A > B \in K$ iff $B \in K * A$

(M) For all $K$, $K'$ and $A$, $K \subseteq K'$ implies $K_A \subseteq K'_A$

(P) If $\neg A \notin K$ and $B \in K$ then $B \in K_A$

(3) $(K/A)/B = K/(A \wedge B)$

(4) $K/(A \vee B) \subseteq K/A$

(I) If $\neg A \notin K$ then $K/A \subseteq K_A$

H. Leitgeb. *On the Ramsey Test without Triviality*. Notre Dame Journal of Formal Logic, 51:1, 2010, pp. 21 - 54.

H. Leitgeb. *Beliefs in Conditionals Vs. Conditional Beliefs*. Topoi, 26 (1), 2007, pp. 115 - 132.

I. Levi. *Iteration of Conditionals and the Ramsey Test*. Synthese, 76, 1988, pp. 49 - 81.

# Ingredients of a Logical Analysis of Rational Agency

# Ingredients of a Logical Analysis of Rational Agency

- What are the basic building blocks?

# Ingredients of a Logical Analysis of Rational Agency

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)

# Ingredients of a Logical Analysis of Rational Agency

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)
- Single agent vs. many agents.

# Ingredients of a Logical Analysis of Rational Agency

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)

- Single agent vs. many agents.

- What the the primitive operators?
  - Informational attitudes
  - Motivational attitudes
  - Normative attitudes

# Ingredients of a Logical Analysis of Rational Agency

- What are the basic building blocks? (the nature of time (continuous or discrete/branching or linear), how (primitive) *events* or *actions* are represented, how *causal* relationships are represented and what constitutes a *state of affairs*.)
- Single agent vs. many agents.
- What the the primitive operators?
  - Informational attitudes
  - Motivational attitudes
  - Normative attitudes
- Static vs. dynamic

# Ingredients of a Logical Analysis of Rational Agency

- $\Rightarrow$ informational attitudes (eg., knowledge, belief, certainty)
- $\Rightarrow$ time, actions and ability
- $\Rightarrow$ motivational attitudes (eg., preferences)
- $\Rightarrow$ group notions (eg., common knowledge and coalitional ability)
- $\Rightarrow$ normative attitudes (eg., obligations)

# Ingredients of a Logical Analysis of Rational Agency

- ✓ informational attitudes (eg., knowledge, belief, certainty)
- ⇒ time, actions and ability
- ⇒ motivational attitudes (eg., preferences)
- ⇒ group notions (eg., common knowledge and coalitional ability)
- ⇒ normative attitudes (eg., obligations)