Reasoning about Knowledge and Beliefs

Lecture 13

Eric Pacuit

University of Maryland, College Park

pacuit.org
epacuit@umd.edu

October 23, 2013

The Theory of Belief Revision

C. Alchourrón, P. Gärdenfors and D. Makinson. *On the logic of theory change: Partial meet contraction and revision functions*. Journal of Symbolic Logic, 50, 510 - 530, 1985.

Hans Rott. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*. Oxford University Press, 2001.

A.P. Pedersen and H. Arló-Costa. *"Belief Revision."*. In Continuum Companion to Philosophical Logic. Continuum Press, 2011.

$$K * \varphi$$

$$K * \varphi$$

Initial set of beliefs

$$K * \varphi$$

Initial set of beliefs

New evidence φ

Revision operator: $* : \mathcal{B} \times \mathcal{L} \to \mathcal{B}$

$$K * \varphi$$

Initial set of beliefs

New evidence φ

Minimal Change

When accepting a new piece of information, an agent should aim at a minimal change of his old beliefs.

"The concept of contraction leads us to the concept of minimal change of belief, or briey, revision" (Makinson 1985, p. 352).

"The criterion of informational economy demands that as few beliefs as possible be given up so that the change is in some sense a minimal change of K to accommodate for A" (Gardenfors 1988, p. 53).

"The amount of information lost in a belief change should be kept minimal" (Gardenfors and Rott 1995, p. 38).

Minimal Change

"At the center of the AGM theory [of theory change] are a number of approaches to giving formal substance to the maxim [of minimal mutilation: keep incisions into theories as small as possible!]" (Fuhrmann 1997, p. 17).

"The hallmark of the AGM postulates is the principle of minimal belief change, that is, the need to preserve as much of earlier beliefs as possible and to add only those beliefs which are absolutely compelled by the revision specified" (Darwiche and Pearl 1997, p. 2).

Keep the *Most Entrenched* Beliefs

If there are different ways to effect a belief change, the agent should give up those beliefs which are least entrenched.

"When a belief set K is contracted (or revised), the sentences in K that are given up are those with the lowest epistemic entrenchment" (Gardenfors 1988, p. 87).

"The guiding idea for the construction is that when a knowledge system K is revised or contracted, the sentences in K which are given up are those having the lowest degrees of epistemic entrenchment" (Gardenfors and Makinson 1988, p. 88).

Keep the *Most Entrenched* Beliefs

"In so far as some beliefs are considered more important or entrenched than others, one should retract the least important ones" (Gardenfors and Rott 1995, p. 38).

"If there are different ways to effect a belief change, the agent should give up those beliefs which are least entrenched...when it comes to choosing between candidates for removal, the least entrenched ones ought to be given up" (Fuhrmann 1997, p. 24).

"A hallmark of the AGM theory is its commitment to the principle of informational economy: beliefs are only given up when there are no less entrenched candidates.... If one of two beliefs must be retracted in order to accommodate some new fact, the less entrenched belief will be relinquished, while the more entrenched persists" (Boutilier 1996, pp. 264-265).

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM 5: $K * \varphi$ is inconsistent only if φ is inconsistent

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM 5: $K * \varphi$ is inconsistent only if φ is inconsistent

AGM 6: If φ and ψ are logically equivalent then $K * \varphi = K * \psi$

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM 5: $K * \varphi$ is inconsistent only if φ is inconsistent

AGM 6: If φ and ψ are logically equivalent then $K * \varphi = K * \psi$

AGM 7: $K * (\varphi \wedge \psi) \subseteq Cn(K * \varphi \cup \{\psi\})$

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM 5: $K * \varphi$ is inconsistent only if φ is inconsistent

AGM 6: If φ and ψ are logically equivalent then $K * \varphi = K * \psi$

AGM 7: $K * (\varphi \wedge \psi) \subseteq Cn(K * \varphi \cup \{\psi\})$

AGM 8: if $\neg\psi \notin K * \varphi$ then $Cn(K * \varphi \cup \{\psi\}) \subseteq K * (\varphi \wedge \psi)$

AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM 5: $K * \varphi$ is inconsistent only if φ is inconsistent

AGM 6: If φ and ψ are logically equivalent then $K * \varphi = K * \psi$

AGM 7: $K * (\varphi \wedge \psi) \subseteq Cn(K * \varphi \cup \{\psi\})$

AGM 8: if $\neg\psi \notin K * \varphi$ then $Cn(K * \varphi \cup \{\psi\}) \subseteq K * (\varphi \wedge \psi)$

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

You are walking down a street and see someone holding a sign reading "The World will End Tomorrow", but you don't add this add this to your beliefs.

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

You are walking down a street and see someone holding a sign reading "The World will End Tomorrow", but you don't add this add this to your beliefs.

Two people, Ann and Bob, are reliable sources of information on whether The Netherlands will win the world cup. They are equally reliable.

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

You are walking down a street and see someone holding a sign reading "The World will End Tomorrow", but you don't add this add this to your beliefs.

Two people, Ann and Bob, are reliable sources of information on whether The Netherlands will win the world cup. They are equally reliable. AGM assumes that the most recent evidence that you received takes precedent. Ann says "yes" and a little bit later, Bob says "no".

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

You are walking down a street and see someone holding a sign reading "The World will End Tomorrow", but you don't add this add this to your beliefs.

Two people, Ann and Bob, are reliable sources of information on whether The Netherlands will win the world cup. They are equally reliable. AGM assumes that the most recent evidence that you received takes precedent. Ann says "yes" and a little bit later, Bob says "no". Why should the, possibly arbitrary, order in which you receive the information give more weight to Bob's announcement?

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

You are walking down a street and see someone holding a sign reading "The World will End Tomorrow", but you don't add this add this to your beliefs.

Two people, Ann and Bob, are reliable sources of information on whether The Netherlands will win the world cup. They are equally reliable. AGM assumes that the most recent evidence that you received takes precedent. Ann says "yes" and a little bit later, Bob says "no". Why should the, possibly arbitrary, order in which you receive the information give more weight to Bob's announcement? Is this a counterexample to AGM 2?

Counterexample to AGM 2

$$\varphi \in K * \varphi$$

You are walking down a street and see someone holding a sign reading "The World will End Tomorrow", but you don't add this add this to your beliefs.

Two people, Ann and Bob, are reliable sources of information on whether The Netherlands will win the world cup. They are equally reliable. AGM assumes that the most recent evidence that you received takes precedent. Ann says "yes" and a little bit later, Bob says "no". Why should the, possibly arbitrary, order in which you receive the information give more weight to Bob's announcement? Is this a counterexample to AGM 2? No (Why?)

Rott's Counterexample

AGM 7: $K * (\varphi \wedge \psi) \subseteq Cn(K * \varphi \cup \{\psi\})$

AGM 8: if $\neg\psi \notin K * \varphi$ then $Cn(K * \varphi \cup \{\psi\}) \subseteq K * (\varphi \wedge \psi)$

So, if $\psi \in Cn(\{\varphi\})$, then $K * \varphi = Cn(K * \varphi \cup \{\psi\})$

Rott's Counterexample

There is an appointment to be made in a philosophy department. The position is a metaphysics position, and there are three main candidates: Andrew, Becker and Cortez.

1. Andrew is clearly the best metaphysician, but is weak in logic.
2. Becker is a very good metaphysician, also good in logic.
3. Cortez is a brilliant logician, but weak in metaphysics.

Rott's Counterexample

There is an appointment to be made in a philosophy department. The position is a metaphysics position, and there are three main candidates: Andrew, Becker and Cortez.

1. Andrew is clearly the best metaphysician, but is weak in logic.
2. Becker is a very good metaphysician, also good in logic.
3. Cortez is a brilliant logician, but weak in metaphysics.

Scenario 1: Paul is told by the dean, that the chosen candidate is either Andrew or Becker. Since Andrew is clearly the better metaphysician of the two, Paul concludes that the winning candidate will be Andrew.

Rott's Counterexample

1. Andrew is clearly the best metaphysician, but is weak in logic.
2. Becker is a very good metaphysician, also good in logic.
3. Cortez is a brilliant logician, but weak in metaphysics.

Scenario 2: Paul is told by the dean that the chosen candidate is either Andrew, Becker or Cortez.

Rott's Counterexample

1. Andrew is clearly the best metaphysician, but is weak in logic.
2. Becker is a very good metaphysician, also good in logic.
3. Cortez is a brilliant logician, but weak in metaphysics.

Scenario 2: Paul is told by the dean that the chosen candidate is either Andrew, Becker or Cortez.

" This piece of information sets off a rather subtle line of reasoning. Knowing that Cortez is a splendid logician, but that he can hardly be called a metaphysician, Paul comes to realize that his background assumption that expertise in the field advertised is the decisive criterion for the appointment cannot be upheld. Apparently, competence in logic is regarded as a considerable asset by the selection committee." Paul concludes Becker will be hired.

"Meta-Information"

"...Rott seems to take the point about meta-information to explain why the example conflicts with the theoretical principles,

"Meta-Information"

"...Rott seems to take the point about meta-information to explain why the example conflicts with the theoretical principles, whereas I want to conclude that it shows why the example does not conflict with the theoretical principles, since I take the relevance of the meta-information to show that the conditions for applying the principles in question are not met by the example.

"Meta-Information"

"...Rott seems to take the point about meta-information to explain why the example conflicts with the theoretical principles, whereas I want to conclude that it shows why the example does not conflict with the theoretical principles, since I take the relevance of the meta-information to show that the conditions for applying the principles in question are not met by the example.... I think proper attention to the relation between concrete examples and the abstract models will allow us to reconcile some of the beautiful properties [of the abstract theory of belief revision] with the complexity of concrete reasoning." (Stalnaker, 204)

Three Epistemic Changes

1. In **expansion**, a sentence φ is added to a belief set K to obtain an expanded belief set $K + \varphi$.
2. In **revision**, a sentence φ is added to a belief set K to obtain a revised belief set $K * \varphi$ in a way that preserves logical consistency.
3. In **contraction**, a sentence φ is removed from K to obtain a contracted belief set $K - \varphi$ that does not include φ.

Three Epistemic Changes

1. In **expansion**, a sentence φ is added to a belief set K to obtain an expanded belief set $K + \varphi$.
2. In **revision**, a sentence φ is added to a belief set K to obtain a revised belief set $K * \varphi$ in a way that preserves logical consistency.
3. In **contraction**, a sentence φ is removed from K to obtain a contracted belief set $K - \varphi$ that does not include φ.

Levi Identity: Revision can be reduced to contraction via the so-called Levi identity,

$$K * \varphi = (K - \neg\varphi) + \varphi$$

Contraction Postulates

$(C1)$	$K - \alpha$ is deductively closed
$(C2)$	$K - \alpha \subseteq K$
$(C3)$	If $\alpha \notin K$ or $\vdash \alpha$ then $K - \alpha = K$
$(C4)$	If $\not\vdash \alpha$, then $\alpha \notin K - \alpha$
$(C5)$	If $\vdash \alpha \leftrightarrow \beta$, then $K - \alpha = K - \beta$
$(C6)$	$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

While reading a book about Cleopatra I learned that she had both a son and a daughter. I therefore believe both that Cleopatra had a son (s) and Cleopatra had a daughter (d).

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

While reading a book about Cleopatra I learned that she had both a son and a daughter. I therefore believe both that Cleopatra had a son (s) and Cleopatra had a daughter (d). Later I learn from a well-informed friend that the book in question is just a historical novel. I accordingly contract my belief that Cleopatra had a child ($s \vee d$).

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

While reading a book about Cleopatra I learned that she had both a son and a daughter. I therefore believe both that Cleopatra had a son (s) and Cleopatra had a daughter (d). Later I learn from a well-informed friend that the book in question is just a historical novel. I accordingly contract my belief that Cleopatra had a child ($s \vee d$). However, shortly thereafter I learn from a reliable source that in fact Cleopatra had a child. I thereby reintroduce $s \vee d$ to my collection of beliefs without also returning either s or d.

(Hansson, 1991)

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

I believed both that George is a criminal (c) and George is a mass murderer (m).

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

I believed both that George is a criminal (c) and George is a mass murderer (m). Upon receiving certain information I am induced to retract my belief set K by my belief that George is a criminal (c). Of course, I therefore retract my belief set by my belief that George is a mass murderer (m).

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

I believed both that George is a criminal (c) and George is a mass murderer (m). Upon receiving certain information I am induced to retract my belief set K by my belief that George is a criminal (c). Of course, I therefore retract my belief set by my belief that George is a mass murderer (m). Later I learn that in fact George is a shoplifter (s), so I expand my contracted belief set $K - c$ by s to obtain $(K - C) + s$.

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

I believed both that George is a criminal (c) and George is a mass murderer (m). Upon receiving certain information I am induced to retract my belief set K by my belief that George is a criminal (c). Of course, I therefore retract my belief set by my belief that George is a mass murderer (m). Later I learn that in fact George is a shoplifter (s), so I expand my contracted belief set $K - c$ by s to obtain $(K - C) + s$. As Georges being a shoplifter (s) entails his being a criminal (c), $(K - c) + c$ is a subset of $(K - c) + s$. Yet by Recovery it follows that $K \subseteq (K - c) + c$, so m is a member of the expanded belief, so m is a member of the expanded belief set $(K - c) + s$.

Counterexamples to Recovery

$$K \subseteq Cn((K - \alpha) \cup \{\alpha\})$$

I believed both that George is a criminal (c) and George is a mass murderer (m). Upon receiving certain information I am induced to retract my belief set K by my belief that George is a criminal (c). Of course, I therefore retract my belief set by my belief that George is a mass murderer (m). Later I learn that in fact George is a shoplifter (s), so I expand my contracted belief set $K - c$ by s to obtain $(K - C) + s$. As Georges being a shoplifter (s) entails his being a criminal (c), $(K - c) + c$ is a subset of $(K - c) + s$. Yet by Recovery it follows that $K \subseteq (K - c) + c$, so m is a member of the expanded belief, so m is a member of the expanded belief set $(K - c) + s$. But I do not believe that George is a mass murdered (m). (Hansson, 1996)

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Complete vs. incomplete belief sets:
$K = Cn(\{p \vee q\})$ vs. $K = Cn(\{p \vee q, p, q\})$

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Complete vs. incomplete belief sets:
$K = Cn(\{p \vee q\})$ vs. $K = Cn(\{p \vee q, p, q\})$

Revising by $\neg p$ ($K * \neg p$) vs. Updating by $\neg p$ ($K \diamond \neg p$)

H. Katsuno and A. O. Mendelzon. *Propositional knowledge base revision and minimal change*. Artificial Intelligence, 52, pp. 263 - 294 (1991).

KM Postulates

KM 1: $K \diamond \varphi = Cn(K \diamond \varphi)$

KM 2: $\varphi \in K \diamond \varphi$

KM 3: If $\varphi \in K$ then $K \diamond \varphi = K$

KM 4: $K \diamond \varphi$ is inconsistent iff φ is inconsistent

KM 5: If φ and ψ are logically equivalent then $K \diamond \varphi = K \diamond \psi$

KM 6: $K \diamond (\varphi \wedge \psi) \subseteq Cn(K \diamond \varphi \cup \{\psi\})$

KM 7: If $\psi \in K \diamond \varphi$ and $\varphi \in K \diamond \psi$ then $K \diamond \varphi = K \diamond \psi$

KM 8: If K is complete then $K \diamond (\varphi \wedge \psi) \subseteq K \diamond \varphi \cap K \diamond \psi$

KM 9: $K \diamond \varphi = \bigcap_{M \in Comp(K)} M \diamond \varphi$, where $Comp(K)$ is the class of all complete theories containing K.

Updating and Revising

$$K \diamond \varphi = \bigcap_{M \in Comp(K)} M * \varphi$$

H. Katsuno and A. O. Mendelzon. *On the difference between updating a knowledge base and revising it.* *Belief Revision*, P. Gärdenfors (ed.), pp 182 - 203 (1992).