Ten Puzzles and Paradoxes about Knowledge and Belief

ESSLLI 2013, Düsseldorf

Wes Holliday **Eric Pacuit**

August 13, 2013

Outline

- **Higher-Order Knowledge**: the Margin of Error Paradox
- **Knowability**: Fitch's Paradox
- **The Dynamics of Knowledge**: the Puzzle of the Gifts

The KK Principle

In our study of the prediction paradox, we spotted the principle:

$$4_i^< \quad K_i\varphi \to K_iK_j\varphi \quad (j > i).$$

The KK Principle

In our study of the prediction paradox, we spotted the principle:

$$4_i^< \quad K_i\varphi \to K_i K_j \varphi \quad (j > i).$$

More famous is the "KK principle" (or "positive introspection"):

$$4_i \quad K_i\varphi \to K_i K_i \varphi.$$

The KK Principle

In our study of the prediction paradox, we spotted the principle:

$$4_i^{<} \quad K_i\varphi \to K_iK_j\varphi \quad (j > i).$$

More famous is the "KK principle" (or "positive introspection"):

$$4_i \quad K_i\varphi \to K_iK_i\varphi.$$

Hintikka, one of the inventors of epistemic logic, endorsed the 4 axiom—at least for what he considered a strong notion of knowledge, found in philosophy from Aristotle to Schopenhauer.

J. Hintikka. 1962. *Knowledge and Belief*, Cornell University Press.

The KK Principle

More famous is the "KK principle" (or "positive introspection"):

$$4_i \quad K_i\varphi \to K_iK_i\varphi.$$

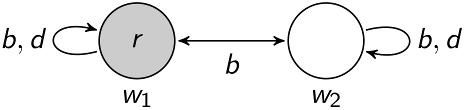
Hintikka, one of the inventors of epistemic logic, endorsed the 4 axiom—at least for what he considered a strong notion of knowledge, found in philosophy from Aristotle to Schopenhauer.

J. Hintikka. 1962. *Knowledge and Belief*, Cornell University Press.

Hintikka rejected arguments for 4 based on claims about agents introspective powers, or what he called "the myth of the self-illumination of certain mental activities" (67). Instead, his claim was that for a strong notion of knowledge, *knowing that one knows* "differs only in words" from *knowing* (§2.1-2.2).

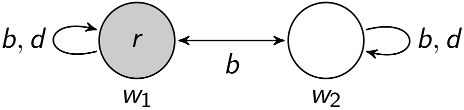
Recall that the relational semantics for normal epistemic logics uses models $\mathcal{M} = \{W, \{R_i\}_{i\in\mathbb{N}}, V\rangle$ where each R_i is a binary "epistemic accessibility" relation on W:

$$\mathcal{M}, w \vDash K_i\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_iv \text{ then } \mathcal{M}, v \vDash \varphi.$$



Recall that the relational semantics for normal epistemic logics uses models $\mathcal{M} = \{W, \{R_i\}_{i \in \mathbb{N}}, V\rangle$ where each R_i is a binary "epistemic accessibility" relation on W:

$$\mathcal{M}, w \vDash K_i\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_iv \text{ then } \mathcal{M}, v \vDash \varphi.$$

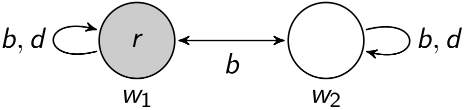


Where $\hat{K}\varphi$ is defined as $\neg K \neg \varphi$, its derived truth clause is:

$$\mathcal{M}, w \vDash \hat{K}_i\varphi \quad \text{iff} \quad \exists v \in W\colon wR_iv \text{ and } \mathcal{M}, v \vDash \varphi.$$

Recall that the relational semantics for normal epistemic logics uses models $\mathcal{M} = \{W, \{R_i\}_{i\in\mathbb{N}}, V\rangle$ where each R_i is a binary "epistemic accessibility" relation on W:

$$\mathcal{M}, w \vDash K_i\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_iv \text{ then } \mathcal{M}, v \vDash \varphi.$$



Where $\hat{K}\varphi$ is defined as $\neg K\neg\varphi$, its derived truth clause is:

$$\mathcal{M}, w \vDash \hat{K}_i\varphi \quad \text{iff} \quad \exists v \in W\colon wR_iv \text{ and } \mathcal{M}, v \vDash \varphi.$$

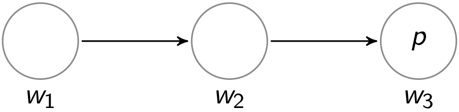
We take wR_iv (an arrow pointing from w to v) to mean that the possibility v is compatible with what the agent knows in w.

The KK Principle and Transitive Accessibility

Now let's return to the KK principle:

$$Kp \to KKp, \text{ or equivalently, } \hat{K}\hat{K}p \to \hat{K}p.$$

This corresponds to the transitivity of the accessibility relation:

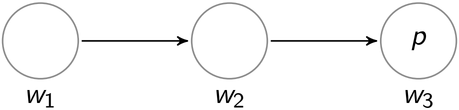


The KK Principle and Transitive Accessibility

Now let's return to the KK principle:

$$Kp \to KKp, \text{ or equivalently, } \hat{K}\hat{K}p \to \hat{K}p.$$

This corresponds to the transitivity of the accessibility relation:



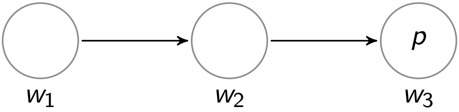
$\mathcal{M}, w_1 \vDash \hat{K}\hat{K}p$

The KK Principle and Transitive Accessibility

Now let's return to the KK principle:

$$Kp \to KKp, \text{ or equivalently, } \hat{K}\hat{K}p \to \hat{K}p.$$

This corresponds to the transitivity of the accessibility relation:



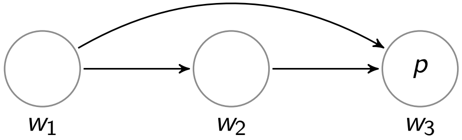
$\mathcal{M}, w_1 \vDash \hat{K}\hat{K}p \quad \mathcal{M}, w_2 \vDash \hat{K}p$

The KK Principle and Transitive Accessibility

Now let's return to the KK principle:

$$Kp \to KKp, \text{ or equivalently, } \hat{K}\hat{K}p \to \hat{K}p.$$

This corresponds to the transitivity of the accessibility relation:



$\mathcal{M}, w_1 \vDash \hat{K}p$ $\mathcal{M}, w_2 \vDash \hat{K}p$

Williamson's Margin of Error Puzzle

We will now consider an argument, due to Williamson, that purports to be a *reductio ad absurdum* of the KK principle.

T. Williamson. 2000. *Knowlege and Its Limits*, Oxford University Press

T. Williamson. 2007. "Rational Failures of the KK Principle." *The Logic of Strategy*, eds. C. Bicchieri, R. Jeffrey, and B. Skyrms, OUP.

Williamson's Margin of Error Puzzle

Suppose an agent is estimating the height of a faraway tree, which is in fact k inches. While the agent's rationality is perfect, his eyesight is not. As Williamson (2000) explains, "anyone who can tell by looking at the tree that it is not i inches tall, when in fact it is $i+1$ inches tall, has much better eyesight and a much greater ability to judge heights" than this agent (115).

Williamson's Margin of Error Puzzle

Suppose an agent is estimating the height of a faraway tree, which is in fact k inches. While the agent's rationality is perfect, his eyesight is not. As Williamson (2000) explains, "anyone who can tell by looking at the tree that it is not i inches tall, when in fact it is $i + 1$ inches tall, has much better eyesight and a much greater ability to judge heights" than this agent (115).

Let h_i stand for *the height of the tree is i inches*, so h_k is true.

Williamson's Margin of Error Puzzle

Suppose an agent is estimating the height of a faraway tree, which is in fact k inches. While the agent's rationality is perfect, his eyesight is not. As Williamson (2000) explains, "anyone who can tell by looking at the tree that it is not i inches tall, when in fact it is $i+1$ inches tall, has much better eyesight and a much greater ability to judge heights" than this agent (115).

Let h_i stand for *the height of the tree is i inches*, so h_k is true.

Given the limited visual discrimination of the agent, we have:

(0) $\forall i : h_{i+1} \to \neg K \neg h_i$.

Williamson's Margin of Error Puzzle

Suppose an agent is estimating the height of a faraway tree, which is in fact k inches. While the agent's rationality is perfect, his eyesight is not. As Williamson (2000) explains, "anyone who can tell by looking at the tree that it is not i inches tall, when in fact it is $i+1$ inches tall, has much better eyesight and a much greater ability to judge heights" than this agent (115).

Let h_i stand for *the height of the tree is i inches*, so h_k is true.

Given the limited visual discrimination of the agent, we have:

(0) $\forall i : h_{i+1} \to \neg K \neg h_i$.

Taking the contrapositive, we have:

(1) $\forall i : K\neg h_i \to \neg h_{i+1}$

Williamson's Margin of Error Puzzle

(1) $\forall i : K\neg h_i \rightarrow \neg h_{i+1}$

Suppose that the agent reflects on the limitations of his visual discrimination and comes to know every instance of (1):

(2) $\forall i : K(K\neg h_i \rightarrow \neg h_{i+1})$.

Williamson's Margin of Error Puzzle

(1) $\forall i : K\neg h_i \rightarrow \neg h_{i+1}$

Suppose that the agent reflects on the limitations of his visual discrimination and comes to know every instance of (1):

(2) $\forall i : K(K\neg h_i \rightarrow \neg h_{i+1})$.

Given these assumptions, it follows that for any j, if the agent knows that the height of the tree is not j inches, then he also knows that the height of the tree is not $j+1$ inches:

Williamson's Margin of Error Puzzle

(1) $\forall i : K\neg h_i \rightarrow \neg h_{i+1}$

Suppose that the agent reflects on the limitations of his visual discrimination and comes to know every instance of (1):

(2) $\forall i : K(K\neg h_i \rightarrow \neg h_{i+1})$.

Given these assumptions, it follows that for any j, if the agent knows that the height of the tree is not j inches, then he also knows that the height of the tree is not $j+1$ inches:

(3) $K\neg h_j$ assumption;

Williamson's Margin of Error Puzzle

(1) $\forall i : K\neg h_i \rightarrow \neg h_{i+1}$

Suppose that the agent reflects on the limitations of his visual discrimination and comes to know every instance of (1):

(2) $\forall i : K(K\neg h_i \rightarrow \neg h_{i+1})$.

Given these assumptions, it follows that for any j, if the agent knows that the height of the tree is not j inches, then he also knows that the height of the tree is not $j + 1$ inches:

(3) $K\neg h_j$ assumption;

(4) $KK\neg h_j$ from (3) using 4 and PL;

Williamson's Margin of Error Puzzle

(1) $\forall i : K\neg h_i \rightarrow \neg h_{i+1}$

Suppose that the agent reflects on the limitations of his visual discrimination and comes to know every instance of (1):

(2) $\forall i : K(K\neg h_i \rightarrow \neg h_{i+1})$.

Given these assumptions, it follows that for any j, if the agent knows that the height of the tree is not j inches, then he also knows that the height of the tree is not $j+1$ inches:

(3) $K\neg h_j$ assumption;

(4) $KK\neg h_j$ from (3) using 4 and PL;

(5) $K(K\neg h_j \rightarrow \neg h_{j+1})$ instance of (2);

Williamson's Margin of Error Puzzle

(1) $\forall i : K\neg h_i \to \neg h_{i+1}$

Suppose that the agent reflects on the limitations of his visual discrimination and comes to know every instance of (1):

(2) $\forall i : K(K\neg h_i \to \neg h_{i+1})$.

Given these assumptions, it follows that for any j, if the agent knows that the height of the tree is not j inches, then he also knows that the height of the tree is not $j + 1$ inches:

(3) $K\neg h_j$ assumption;

(4) $KK\neg h_j$ from (3) using 4 and PL;

(5) $K(K\neg h_j \to \neg h_{j+1})$ instance of (2);

(6) $K\neg h_{j+1}$ from (4) and (5) using RK and PL.

Williamson's Margin of Error Puzzle

(2) $\forall i : K(K\neg h_i \rightarrow \neg h_{i+1})$.

Given these assumptions, it follows that for any j, if the agent knows that the height of the tree is not j inches, then he also knows that the height of the tree is not $j+1$ inches:

(3) $K\neg h_j$ assumption;

(4) $KK\neg h_j$ from (3) using 4 and PL;

(5) $K(K\neg h_j \rightarrow \neg h_{j+1})$ instance of (2);

(6) $K\neg h_{j+1}$ from (4) and (5) using RK and PL.

Assuming $K\neg h_0$ holds, by repeating the steps of (3) - (6), we reach the conclusion $K\neg h_k$ by induction. Finally, by T, $K\neg h_k$ implies $\neg h_k$, contradicting our initial assumption of h_k.

Williamson's Margin of Error Puzzle

Formally, Williamson's observation is that for all $i, j \in \mathbb{N}$ with $j > i$:

$$\{K(K\neg h_i \to \neg h_{i+1}) \mid i \in \mathbb{N}\} \vdash_{\mathbf{K4}} K\neg h_i \to K\neg h_j.$$

This gives us the absurd result that $K\neg h_0 \to K\neg h_k$.

Williamson's Margin of Error Puzzle

Formally, Williamson's observation is that for all $i, j \in \mathbb{N}$ with $j > i$:

$$\{K(K\neg h_i \rightarrow \neg h_{i+1}) \mid i \in \mathbb{N}\} \vdash_{\mathbf{K4}} K\neg h_i \rightarrow K\neg h_j.$$

This gives us the absurd result that $K\neg h_0 \rightarrow K\neg h_k$.

Since Williamson defends the principles of the form $K(K\neg h_i \rightarrow \neg h_{i+1})$, he argues that we should reject 4.

Williamson's Margin of Error Puzzle

Formally, Williamson's observation is that for all $i, j \in \mathbb{N}$ with $j > i$:

$$\{K(K\neg h_i \to \neg h_{i+1}) \mid i \in \mathbb{N}\} \vdash_{\mathbf{K4}} K\neg h_i \to K\neg h_j.$$

This gives us the absurd result that $K\neg h_0 \to K\neg h_k$.

Since Williamson defends the principles of the form $K(K\neg h_i \to \neg h_{i+1})$, he argues that we should reject 4.

To model agents with limited discrimination, Williamson proposes epistemic models with non-transitive accessibility relations.

Non-transitive Models for Limited Discrimination

Suppose the agent has a fixed margin of error ϵ for judging the heights of the tree: so if the tree is height i, it is compatible with the agent's knowledge that its height is between $i - \epsilon$ and $i + \epsilon$.

Non-transitive Models for Limited Discrimination

Suppose the agent has a fixed margin of error ϵ for judging the heights of the tree: so if the tree is height i, it is compatible with the agent's knowledge that its height is between $i - \epsilon$ and $i + \epsilon$.

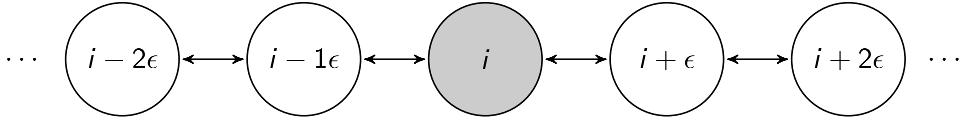
According to Williamson, part of the epistemic model for the agent should look like this (ignoring heights between i and $i \pm \epsilon$):

$$\cdots \quad i - 2\epsilon \longleftrightarrow i - 1\epsilon \longleftrightarrow i \longleftrightarrow i + \epsilon \longleftrightarrow i + 2\epsilon \quad \cdots$$

Non-transitive Models for Limited Discrimination

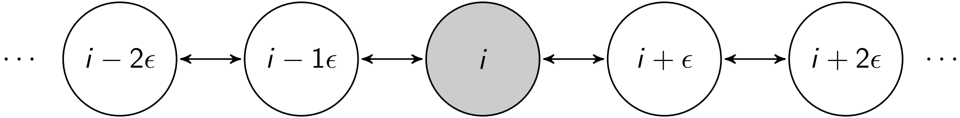
Suppose the agent has a fixed margin of error ϵ for judging the heights of the tree: so if the tree is height i, it is compatible with the agent's knowledge that its height is between $i - \epsilon$ and $i + \epsilon$.

According to Williamson, part of the epistemic model for the agent should look like this (ignoring heights between i and $i \pm \epsilon$):



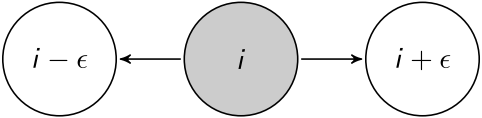
Note: at the shaded world, $K\neg i + 2\epsilon \wedge \neg KK\neg i + 2\epsilon$ is true.

Non-transitive Models for Limited Discrimination



Note: at the shaded world, $K\neg i + 2\epsilon \wedge \neg KK\neg i + 2\epsilon$ is true.

Compare the non-transitive model above with the transitive model:



Now $K\neg i + 2\epsilon \wedge KK\neg i + 2\epsilon$ is true at the shaded world.

Non-transitive Models for Limited Discrimination

$$\cdots \quad i-2\epsilon \longleftrightarrow i-1\epsilon \longleftrightarrow i \longleftrightarrow i+\epsilon \longleftrightarrow i+2\epsilon \quad \cdots$$

Note: at the shaded world, $K^l\neg 0$ (for some $l \in \mathbb{N}$) is *false*.

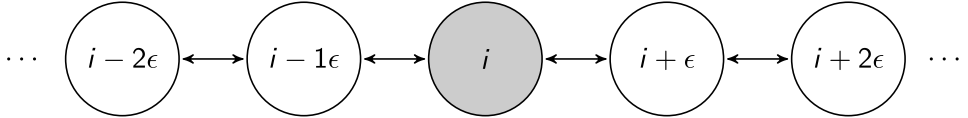
M. Gómez-Torrente. 1997.
"Two Problems for an Epistemicist View of Vagueness," *Philosophical Issues*.

Compare the non-transitive model above with the transitive model:

$$i-\epsilon \longleftarrow i \longrightarrow i+\epsilon$$

In this model, $K^l\neg 0$ is *true* at the shaded world.

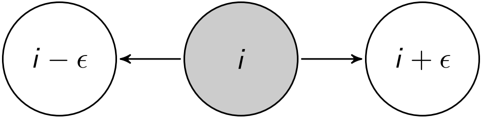
Non-transitive Models for Limited Discrimination



Note: at the shaded world, $K^l \neg 0$ (for some $l \in \mathbb{N}$) is *false*.

What is preventing the agent from knowing that he knows that he knows ... (l times) ... that the tree is not 0 inches?

Compare the non-transitive model above with the transitive model:



In this model, $K^l \neg 0$ is *true* at the shaded world.

Fitch's Paradox

Fitch (1963) derived an unexpected consequence from the thesis, advocated by some anti-realists, that *every truth is knowable*:

Fitch's Paradox

Fitch (1963) derived an unexpected consequence from the thesis, advocated by some anti-realists, that *every truth is knowable*:

(VT) $q \rightarrow \Diamond Kq$,

where $\Diamond$ is a *possibility* operator (more on this later).

Fitch's Paradox

Fitch (1963) derived an unexpected consequence from the thesis, advocated by some anti-realists, that *every truth is knowable*:

(VT) $q \to \Diamond Kq$,

where $\Diamond$ is a *possibility* operator (more on this later).

Fitch make two modest assumptions for K, $K\varphi \to \varphi$ (T) and $K(\varphi \wedge \psi) \to (K\varphi \wedge K\psi)$ (M), and two modest assumptions for $\Diamond$:

Fitch's Paradox

Fitch (1963) derived an unexpected consequence from the thesis, advocated by some anti-realists, that *every truth is knowable*:

(VT) $q \to \Diamond Kq$,

where $\Diamond$ is a *possibility* operator (more on this later).

Fitch make two modest assumptions for K, $K\varphi \to \varphi$ (T) and $K(\varphi \wedge \psi) \to (K\varphi \wedge K\psi)$ (M), and two modest assumptions for $\Diamond$:

- $\Diamond$ is the dual of $\Box$ for *necessity*, so $\neg\Diamond\varphi$ follows from $\Box\neg\varphi$.

Fitch's Paradox

Fitch (1963) derived an unexpected consequence from the thesis, advocated by some anti-realists, that *every truth is knowable*:

(VT) $q \rightarrow \Diamond Kq$,

where $\Diamond$ is a *possibility* operator (more on this later).

Fitch make two modest assumptions for K, $K\varphi \rightarrow \varphi$ (T) and $K(\varphi \wedge \psi) \rightarrow (K\varphi \wedge K\psi)$ (M), and two modest assumptions for $\Diamond$:

- $\Diamond$ is the dual of $\Box$ for *necessity*, so $\neg\Diamond\varphi$ follows from $\Box\neg\varphi$.
- $\Box$ obeys the rule of Necessitation: if φ is a theorem, so is $\Box\varphi$.

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \land \neg Kp) \rightarrow \Diamond K(p \land \neg Kp)$

Here is the proof for Fitch's Paradox:

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \rightarrow \neg Kp$ instance of T axiom

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \rightarrow \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge \neg Kp)$ from (1) and (2) by PL

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \to \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \to (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \to \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \to (Kp \wedge \neg Kp)$ from (1) and (2) by PL

(4) $\neg K(p \wedge \neg Kp)$ from (3) by PL

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \to \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \to (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \to \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \to (Kp \wedge \neg Kp)$ from (1) and (2) by PL

(4) $\neg K(p \wedge \neg Kp)$ from (3) by PL

(5) $\Box\neg K(p \wedge \neg Kp)$ from (4) by $\Box$-Necessitation

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \rightarrow \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge \neg Kp)$ from (1) and (2) by PL

(4) $\neg K(p \wedge \neg Kp)$ from (3) by PL

(5) $\Box\neg K(p \wedge \neg Kp)$ from (4) by $\Box$-Necessitation

(6) $\neg\Diamond K(p \wedge \neg Kp)$ from (5) by $\Box$ - $\Diamond$ Duality

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \rightarrow \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge \neg Kp)$ from (1) and (2) by PL

(4) $\neg K(p \wedge \neg Kp)$ from (3) by PL

(5) $\Box\neg K(p \wedge \neg Kp)$ from (4) by $\Box$-Necessitation

(6) $\neg\Diamond K(p \wedge \neg Kp)$ from (5) by $\Box$ - $\Diamond$ Duality

(7) $\neg(p \wedge \neg Kp)$ from (0) by PL

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \rightarrow \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge \neg Kp)$ from (1) and (2) by PL

(4) $\neg K(p \wedge \neg Kp)$ from (3) by PL

(5) $\Box\neg K(p \wedge \neg Kp)$ from (4) by $\Box$-Necessitation

(6) $\neg\Diamond K(p \wedge \neg Kp)$ from (5) by $\Box$ - $\Diamond$ Duality

(7) $\neg(p \wedge \neg Kp)$ from (0) by PL

(8) $p \rightarrow Kp$ from (7) by classical PL

Fitch's Paradox

For an arbitrary p, consider the following instance of (VT):

(0) $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$

Here is the proof for Fitch's Paradox:

(1) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge K\neg Kp)$ instance of M axiom

(2) $K\neg Kp \rightarrow \neg Kp$ instance of T axiom

(3) $K(p \wedge \neg Kp) \rightarrow (Kp \wedge \neg Kp)$ from (1) and (2) by PL

(4) $\neg K(p \wedge \neg Kp)$ from (3) by PL

(5) $\Box\neg K(p \wedge \neg Kp)$ from (4) by $\Box$-Necessitation

(6) $\neg\Diamond K(p \wedge \neg Kp)$ from (5) by $\Box$ - $\Diamond$ Duality

(7) $\neg(p \wedge \neg Kp)$ from (0) by PL

(8) $p \rightarrow Kp$ from (7) by classical PL

Since p was arbitrary, we have shown that *every truth is known*.

The Question

Fitch's Paradox leaves us with **the question**: what must we require in addition to the truth of φ to ensure the knowability of φ?

The Question

Fitch's Paradox leaves us with **the question**: what must we require in addition to the truth of φ to ensure the knowability of φ?

There is a fairly large literature on knowability and related issues. See, e.g.:

J. Salerno. 2009. *New Essays on the Knowability Paradox*, OUP

J. van Benthem. 2004. "What One May Come to Know," *Analysis*.

P. Balbiani et al. 2008. "'Knowable' as 'Known after an Announcement,"' *Review of Symbolic Logic*.

Dynamic Epistemic Logic

The key idea of dynamic epistemic logic is that we can represent changes in agents' epistemic states by *transforming models*.

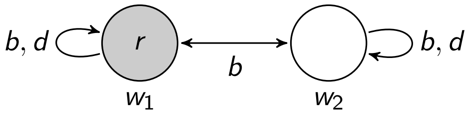
Dynamic Epistemic Logic

The key idea of dynamic epistemic logic is that we can represent changes in agents' epistemic states by *transforming models*.

In the simplest case, we model an agent's acquisition of knowledge by the elimination of possibilities from an initial epistemic model.

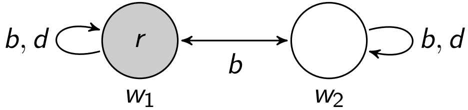
Example (Berkeley and Düsseldorf)

Recall the Berkeley agent who doesn't know whether it's raining in Düsseldorf, whose epistemic state is represented by the model:



Example (Berkeley and Düsseldorf)

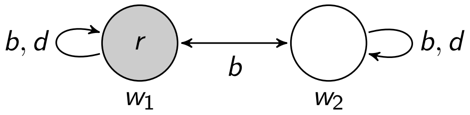
Recall the Berkeley agent who doesn't know whether it's raining in Düsseldorf, whose epistemic state is represented by the model:



What happens when the Düsseldorf agent calls the Berkeley agent on the phone and says, "It's raining in Düsseldorf"?

Example (Berkeley and Düsseldorf)

Recall the Berkeley agent who doesn't know whether it's raining in Düsseldorf, whose epistemic state is represented by the model:

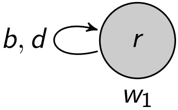


What happens when the Düsseldorf agent calls the Berkeley agent on the phone and says, "It's raining in Düsseldorf"?

We model the change in b's epistemic state by eliminating all epistemic possibilities in which it's *not* raining in Düsseldorf.

Example (Berkeley and Düsseldorf)

Recall the Berkeley agent who doesn't know whether it's raining in Düsseldorf, whose epistemic state is represented by the model:



What happens when the Düsseldorf agent calls the Berkeley agent on the phone and says, "It's raining in Düsseldorf"?

We model the change in b's epistemic state by eliminating all epistemic possibilities in which it's *not* raining in Düsseldorf.

Model Update

We can easily give a formal definition that captures the idea of knowledge acquisition as the elimination of possibilities.

Model Update

We can easily give a formal definition that captures the idea of knowledge acquisition as the elimination of possibilities.

Given $\mathcal{M} = \langle W, \{R_a \mid a \in \mathrm{Agt}\}, V\rangle$, the *updated model* $\mathcal{M}_{|\varphi}$ is obtained by deleting from $\mathcal{M}$ all worlds in which φ was false.

Model Update

We can easily give a formal definition that captures the idea of knowledge acquisition as the elimination of possibilities.

Given $\mathcal{M} = \langle W, \{R_a \mid a \in \text{Agt}\}, V\rangle$, the *updated model* $\mathcal{M}_{|\varphi}$ is obtained by deleting from $\mathcal{M}$ all worlds in which φ was false.

Formally, $\mathcal{M}_{|\varphi} = \langle W_{|\varphi}, \{R_{a_{|\varphi}} \mid a \in \text{Agt}\}, V_{|\varphi}\rangle$ is the model s.th.:

$$W_{|\varphi} = \{v \in W \mid \mathcal{M}, v \models \varphi\};$$

Model Update

We can easily give a formal definition that captures the idea of knowledge acquisition as the elimination of possibilities.

Given $\mathcal{M} = \langle W, \{R_a \mid a \in \text{Agt}\}, V\rangle$, the *updated model* $\mathcal{M}_{|\varphi}$ is obtained by deleting from $\mathcal{M}$ all worlds in which φ was false.

Formally, $\mathcal{M}_{|\varphi} = \langle W_{|\varphi}, \{R_{a_{|\varphi}} \mid a \in \text{Agt}\}, V_{|\varphi}\rangle$ is the model s.th.:

$$W_{|\varphi} = \{v \in W \mid \mathcal{M}, v \vDash \varphi\};$$

$R_{a_{|\varphi}}$ is the restriction of R_a to $W_{|\varphi}$;

Model Update

We can easily give a formal definition that captures the idea of knowledge acquisition as the elimination of possibilities.

Given $\mathcal{M} = \langle W, \{R_a \mid a \in \text{Agt}\}, V\rangle$, the *updated model* $\mathcal{M}_{|\varphi}$ is obtained by deleting from $\mathcal{M}$ all worlds in which φ was false.

Formally, $\mathcal{M}_{|\varphi} = \langle W_{|\varphi}, \{R_{a_{|\varphi}} \mid a \in \text{Agt}\}, V_{|\varphi}\rangle$ is the model s.th.:

$$W_{|\varphi} = \{v \in W \mid \mathcal{M}, v \vDash \varphi\};$$

$R_{a_{|\varphi}}$ is the restriction of R_a to $W_{|\varphi}$;

$V_{|\varphi}(p)$ is the intersection of $V(p)$ and $W_{|\varphi}$.

Model Update

We can easily give a formal definition that captures the idea of knowledge acquisition as the elimination of possibilities.

Given $\mathcal{M} = \langle W, \{R_a \mid a \in \text{Agt}\}, V\rangle$, the *updated model* $\mathcal{M}_{|\varphi}$ is obtained by deleting from $\mathcal{M}$ all worlds in which φ was false.

Formally, $\mathcal{M}_{|\varphi} = \langle W_{|\varphi}, \{R_{a_{|\varphi}} \mid a \in \text{Agt}\}, V_{|\varphi}\rangle$ is the model s.th.:

$W_{|\varphi} = \{v \in W \mid \mathcal{M}, v \vDash \varphi\}$;

$R_{a_{|\varphi}}$ is the restriction of R_a to $W_{|\varphi}$;

$V_{|\varphi}(p)$ is the intersection of $V(p)$ and $W_{|\varphi}$.

In the single-agent case, this models the agent learning φ. In the multi-agent case, this models all agents *publicly* learning φ.

Public Announcement Logic

One of the **big ideas** of dynamic epistemic logic is to add to our formal language operators that can describe the kinds of model updates that we just saw for the Berkeley and Düsseldorf example.

Public Announcement Logic

One of the **big ideas** of dynamic epistemic logic is to add to our formal language operators that can describe the kinds of model updates that we just saw for the Berkeley and Düsseldorf example.

The language of Public Announcement Logic (PAL) is given by:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [!\varphi]\varphi$$

Public Announcement Logic

One of the **big ideas** of dynamic epistemic logic is to add to our formal language operators that can describe the kinds of model updates that we just saw for the Berkeley and Düsseldorf example.

The language of Public Announcement Logic (PAL) is given by:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [!\varphi]\varphi$$

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Public Announcement Logic

One of the **big ideas** of dynamic epistemic logic is to add to our formal language operators that can describe the kinds of model updates that we just saw for the Berkeley and Düsseldorf example.

The language of Public Announcement Logic (PAL) is given by:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [!\varphi]\varphi$$

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi\rangle\psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

Public Announcement Logic

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi\rangle\psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

The truth clause for the dynamic operator $[!\varphi]$ is:

Public Announcement Logic

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi\rangle\psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

The truth clause for the dynamic operator $[!\varphi]$ is:

- $\mathcal{M}, w \vDash [!\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

Public Announcement Logic

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi \rangle \psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

The truth clause for the dynamic operator $[!\varphi]$ is:

- $\mathcal{M}, w \models [!\varphi]\psi$ iff $\mathcal{M}, w \models \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \models \psi$.

So if φ is false, $[!\varphi]\psi$ is vacuously true.

Public Announcement Logic

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi\rangle\psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

The truth clause for the dynamic operator $[!\varphi]$ is:

- $\mathcal{M}, w \vDash [!\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

So if φ is false, $[!\varphi]\psi$ is vacuously true. Here is the $\langle !\varphi\rangle$ clause:

Public Announcement Logic

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi\rangle\psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

The truth clause for the dynamic operator $[!\varphi]$ is:

- $\mathcal{M}, w \vDash [!\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

So if φ is false, $[!\varphi]\psi$ is vacuously true. Here is the $\langle !\varphi\rangle$ clause:

- $\mathcal{M}, w \vDash \langle !\varphi\rangle\psi$ iff $\mathcal{M}, w \vDash \varphi$ *and* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

Public Announcement Logic

Read $[!\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle !\varphi\rangle\psi := \neg[!\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

The truth clause for the dynamic operator $[!\varphi]$ is:

- $\mathcal{M}, w \vDash [!\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

So if φ is false, $[!\varphi]\psi$ is vacuously true. Here is the $\langle !\varphi\rangle$ clause:

- $\mathcal{M}, w \vDash \langle !\varphi\rangle\psi$ iff $\mathcal{M}, w \vDash \varphi$ *and* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

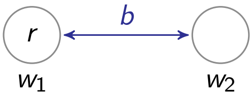
Big Idea: we evaluate $[!\varphi]\psi$ and $\langle !\varphi\rangle\psi$ not by looking at *other worlds in the same model*, but rather by looking at a new model.

Self-Refuting Announcements

Suppose that in the Berkeley and Düsseldorf example, the Düsseldorf agent (a perfectly trustworthy source of weather information) tells the Berkeley agent over the phone, "You don't know it, but it's raining in Düsseldorf": $\neg K_b r \wedge r$.

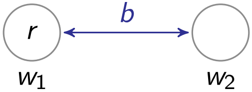
Self-Refuting Announcements

Suppose that in the Berkeley and Düsseldorf example, the Düsseldorf agent (a perfectly trustworthy source of weather information) tells the Berkeley agent over the phone, "You don't know it, but it's raining in Düsseldorf": $\neg K_b r \wedge r$.



Self-Refuting Announcements

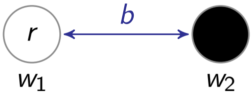
Suppose that in the Berkeley and Düsseldorf example, the Düsseldorf agent (a perfectly trustworthy source of weather information) tells the Berkeley agent over the phone, “You don’t know it, but it’s raining in Düsseldorf”: $\neg K_b r \wedge r$.



Observe that $\mathcal{M}, w_1 \vDash \langle ! \neg K_b r \wedge r \rangle \neg(\neg K_b r \wedge r)$.

Self-Refuting Announcements

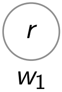
Suppose that in the Berkeley and Düsseldorf example, the Düsseldorf agent (a perfectly trustworthy source of weather information) tells the Berkeley agent over the phone, "You don't know it, but it's raining in Düsseldorf": $\neg K_b r \wedge r$.



Observe that $\mathcal{M}, w_1 \models \langle ! \neg K_b r \wedge r \rangle \neg(\neg K_b r \wedge r)$.
Delete the world w_2 where $\neg K_b r \wedge r$ is false.

Self-Refuting Announcements

Suppose that in the Berkeley and Düsseldorf example, the Düsseldorf agent (a perfectly trustworthy source of weather information) tells the Berkeley agent over the phone, “You don’t know it, but it’s raining in Düsseldorf”: $\neg K_b r \wedge r$.



Observe that $\mathcal{M}, w_1 \vDash \langle ! \neg K_b r \wedge r \rangle \neg(\neg K_b r \wedge r)$.
Observe that $\mathcal{M}_{|\neg K_b r \wedge r}, w_1 \vDash \neg(\neg K_b r \wedge r)$.

Self-Refuting Announcements

Not only is the update with $\neg K_b r \wedge r$ unsuccessful in this specific case, but in general $\neg K_b r \wedge r$ is self-refuting. Let $\alpha := \neg K_b r \wedge r$.

Self-Refuting Announcements

Not only is the update with $\neg K_b r \wedge r$ unsuccessful in this specific case, but in general $\neg K_b r \wedge r$ is self-refuting. Let $\alpha := \neg K_b r \wedge r$.

Proof. Suppose $\mathcal{M}, w \vDash \alpha$. In $\mathcal{M}_{|\alpha}$, there are no worlds where r is false. Hence $\mathcal{M}_{|\alpha}, w \vDash K_b r$, which means $\mathcal{M}_{|\alpha}, w \vDash \neg\alpha$. Thus, $\mathcal{M}, w \vDash [!\alpha]\neg\alpha$. Since $\mathcal{M}, w$ was arbitrary, $[!\alpha]\neg\alpha$ is valid.

Self-Refuting Announcements

Not only is the update with $\neg K_b r \wedge r$ unsuccessful in this specific case, but in general $\neg K_b r \wedge r$ is self-refuting. Let $\alpha := \neg K_b r \wedge r$.

Proof. Suppose $\mathcal{M}, w \vDash \alpha$. In $\mathcal{M}_{|\alpha}$, there are no worlds where r is false. Hence $\mathcal{M}_{|\alpha}, w \vDash K_b r$, which means $\mathcal{M}_{|\alpha}, w \vDash \neg\alpha$. Thus, $\mathcal{M}, w \vDash [!\alpha]\neg\alpha$. Since $\mathcal{M}, w$ was arbitrary, $[!\alpha]\neg\alpha$ is valid.

Question: is $\neg K_b \varphi \wedge \varphi$ self-refuting for all φ?

Self-Refuting Announcements

Not only is the update with $\neg K_b r \wedge r$ unsuccessful in this specific case, but in general $\neg K_b r \wedge r$ is self-refuting. Let $\alpha := \neg K_b r \wedge r$.

Proof. Suppose $\mathcal{M}, w \vDash \alpha$. In $\mathcal{M}_{|\alpha}$, there are no worlds where r is false. Hence $\mathcal{M}_{|\alpha}, w \vDash K_b r$, which means $\mathcal{M}_{|\alpha}, w \vDash \neg\alpha$. Thus, $\mathcal{M}, w \vDash [!\alpha]\neg\alpha$. Since $\mathcal{M}, w$ was arbitrary, $[!\alpha]\neg\alpha$ is valid.

Question: is $\neg K_b \varphi \wedge \varphi$ self-refuting for all φ?

Or is there a φ such that if you receive the true information (from a source you know to be infallible) that "you don't know it, but φ," it can *remain true* afterward that you don't know it, but φ?

What's Wrong with Moore Sentences?

Is there a φ such that if you receive the true information (from a source you know to be infallible) that "you don't know it, but φ," it can *remain true* afterward that you don't know it, but φ?

What's Wrong with Moore Sentences?

Is there a φ such that if you receive the true information (from a source you know to be infallible) that "you don't know it, but φ," it can *remain true* afterward that you don't know it, but φ?

> If you know that I am well informed and if I address the words . . . to you, these words have a curious effect which may perhaps be called anti-performatory. You may come to know that what I say *was* true, but saying it in so many words has the effect of making what is being said false. (68-69)

J. Hintikka 1962. *Knowledge and Belief*.

What's Wrong with Moore Sentences?

Is there a φ such that if you receive the true information (from a source you know to be infallible) that "you don't know it, but φ," it can *remain true* afterward that you don't know it, but φ?

> If you know that I am well informed and if I address the words . . . to you, these words have a curious effect which may perhaps be called anti-performatory. You may come to know that what I say *was* true, but saying it in so many words has the effect of making what is being said false. (68-69)

J. Hintikka 1962. *Knowledge and Belief*.

Surprisingly, this is not always the case, as we will now show...

What's Wrong with Moore Sentences?

> If you know that I am well informed and if I address the words . . . to you, these words have a curious effect which may perhaps be called anti-performatory. You may come to know that what I say *was* true, but saying it in so many words has the effect of making what is being said false. (68-69)

J. Hintikka 1962. *Knowledge and Belief*.

Surprisingly, this is not always the case, as we will now show...

We will show this with the **Puzzle of the Gifts** from

W. Holliday, T. Hoshi, and T. Icard. 2013

"Information Dynamics and Uniform Substitution," *Synthese*.

The Puzzle of the Gifts

With my hands behind my back, **I** walk into a room where a friend **F** is sitting. **F** did not see what if anything **I** put in my hands, and **I** know this. In fact, **I** have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," **I** truthfully announce:

The Puzzle of the Gifts

With my hands behind my back, **I** walk into a room where a friend **F** is sitting. **F** did not see what if anything **I** put in my hands, and **I** know this. In fact, **I** have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," **I** truthfully announce:

(G) Either **I** have a gift in my *right* hand and you don't know it, or **I** have gifts in *both* hands and you don't know **I** have one in my *left* hand.

The Puzzle of the Gifts

With my hands behind my back, I walk into a room where a friend **F** is sitting. **F** did not see what if anything I put in my hands, and I know this. In fact, I have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," I truthfully announce:

(G) Either I have a gift in my *right* hand and you don't know it, or I have gifts in *both* hands and you don't know I have one in my *left* hand.

F takes me to be an infallible source and therefore accepts G.

The Puzzle of the Gifts

With my hands behind my back, I walk into a room where a friend **F** is sitting. **F** did not see what if anything I put in my hands, and I know this. In fact, I have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," I truthfully announce:

(G) Either I have a gift in my *right* hand and you don't know it, or I have gifts in *both* hands and you don't know I have one in my *left* hand.

F takes me to be an infallible source and therefore accepts G.

1. After my announcement, does **F** know if I have a gift in my left/right/both hand(s)?

The Puzzle of the Gifts

With my hands behind my back, I walk into a room where a friend **F** is sitting. **F** did not see what if anything I put in my hands, and I know this. In fact, I have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," I truthfully announce:

(G) Either I have a gift in my *right* hand and you don't know it, or I have gifts in *both* hands and you don't know I have one in my *left* hand.

F takes me to be an infallible source and therefore accepts G.

1. After my announcement, does **F** know if I have a gift in my left/right/both hand(s)?
2. After my announcement, is G *true*?

The Puzzle of the Gifts

With my hands behind my back, **I** walk into a room where a friend **F** is sitting. **F** did not see what if anything **I** put in my hands, and **I** know this. In fact, **I** have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," **I** truthfully announce:

(G) Either **I** have a gift in my *right* hand and you don't know it, or **I** have gifts in *both* hands and you don't know **I** have one in my *left* hand.

F takes me to be an infallible source and therefore accepts G.

1. After my announcement, does **F** know if **I** have a gift in my left/right/both hand(s)?
2. After my announcement, is G *true*?
3. After my announcement, does **F** *know* G?

The Puzzle of the Gifts

With my hands behind my back, I walk into a room where a friend **F** is sitting. **F** did not see what if anything I put in my hands, and I know this. In fact, I have gifts for **F** in both hands. Instead of asking **F** to "pick a hand, any hand," I truthfully announce:

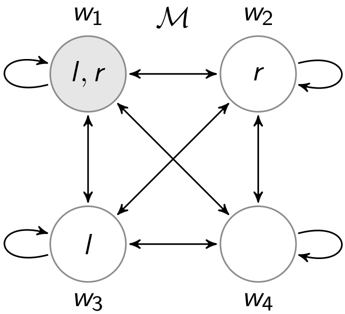
(G) Either I have a gift in my *right* hand and you don't know it, or I have gifts in *both* hands and you don't know I have one in my *left* hand.

F takes me to be an infallible source and therefore accepts G.

1. After my announcement, does **F** know if I have a gift in my left/right/both hand(s)?
2. After my announcement, is G *true*?
3. After my announcement, does **F** *know* G?
4. If 'yes' to 2, what happens if I announce G again?

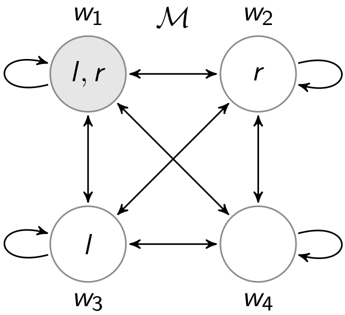
Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.

Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.



We can translate G into the language of epistemic logic as

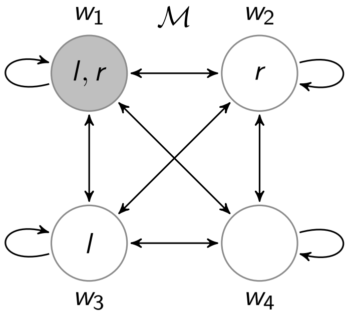
Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.



We can translate G into the language of epistemic logic as

(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.

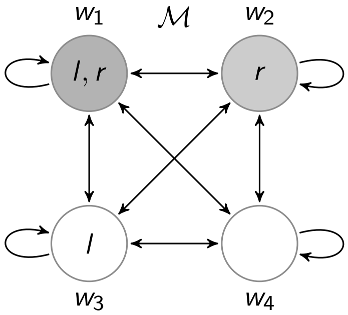


We can translate G into the language of epistemic logic as

(*G*) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Note: $\mathcal{M}, w_1 \vDash G$

Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.

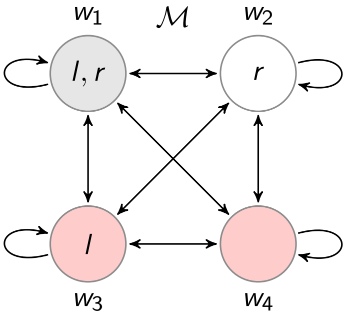


We can translate G into the language of epistemic logic as

(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Note: $\mathcal{M}, w_1 \models G$ and $\mathcal{M}, w_2 \models G$.

Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.

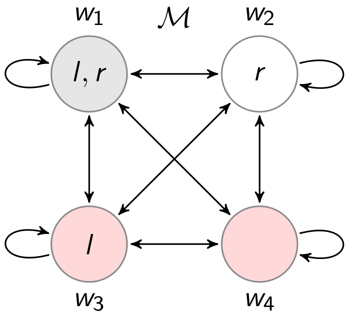


We can translate G into the language of epistemic logic as

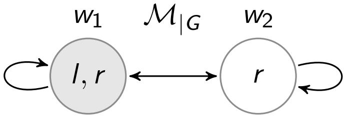
(*G*) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Note: $\mathcal{M}, w_1 \vDash G$, $\mathcal{M}, w_2 \vDash G$, but $\mathcal{M}, w_3 \nvDash G$, $\mathcal{M}, w_4 \nvDash G$.

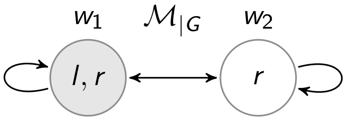
What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?



What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?



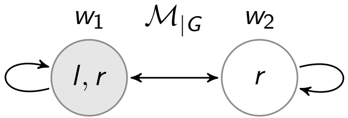
What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?



(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

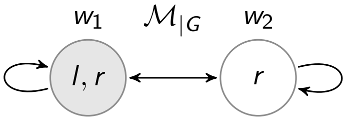


(*G*) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G . . .

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

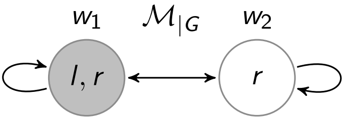


(*G*) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true?

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

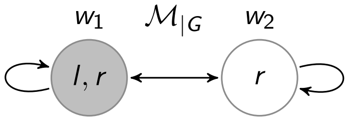


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes.

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

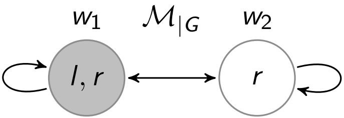


(*G*) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes.

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

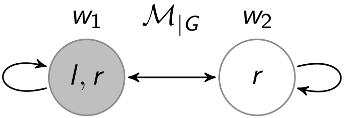


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}_{|G}, w_1 \vDash G$.

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

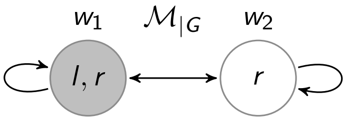


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}, w_1 \models \langle !G \rangle G$.

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

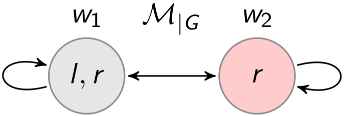


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G?

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

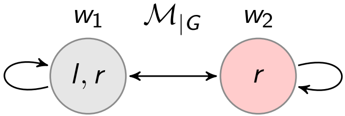


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}, w_1 \models \langle !G \rangle G$.
3. Does **F** now know G? No!

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

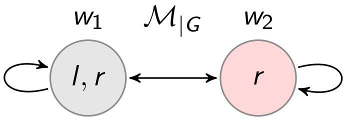


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}, w_1 \models \langle ! G \rangle G$.
3. Does **F** now know G? No! $\mathcal{M}_{|G}, w_1 \models \neg K_{\mathbf{F}} G$.

What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?

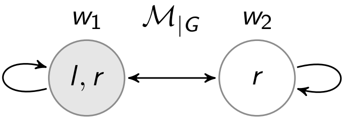


(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Questions. After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No! $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

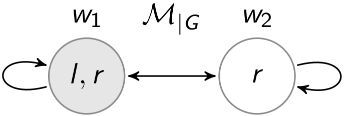
What happens if **I** truthfully announce G, and **F** knows that **I** am an infallible source of information?



(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

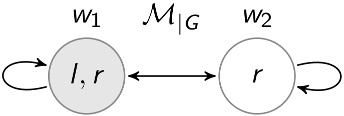
After my announcement of G ...

1. Does **F** know if **I** have a gift in my left/right/both hand(s)?
2. Is G still true? Yes. $\mathcal{M}, w_1 \models \langle ! G \rangle G$.
3. Does **F** now know G? No! $\mathcal{M}, w_1 \models \langle ! G \rangle \neg K_{\mathbf{F}} G$.



Questions. After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

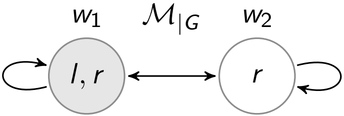


Questions. After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

Given 2 and 3, the following is *not valid*:

$$[!\varphi]\varphi \rightarrow [!\varphi]K\varphi$$



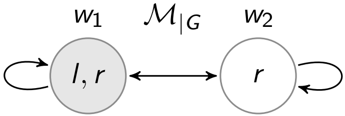
Questions. After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

Given 2 and 3, the following is *not valid*:

$$[!\varphi]\varphi \rightarrow [!\varphi]K\varphi$$

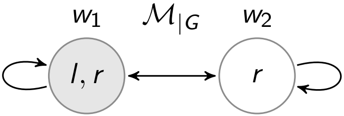
There are formulas φ such that even if φ remains true after being truly announced by a source whom you know to be infallible, you can fail to know that φ is still true.



Questions. After my announcement of G . . .

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

It follows from the answers to 2 and 3 that
$\mathcal{M}, w_1 \vDash \langle !G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.



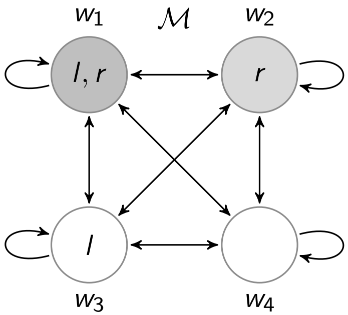
Questions. After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

It follows from the answers to 2 and 3 that
$\mathcal{M}, w_1 \vDash \langle !G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

Let's check that G and $(G \wedge \neg K_{\mathbf{F}} G)$ are true at the same states in our *original* model $\mathcal{M}$, namely w_1 and w_2.

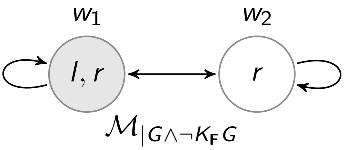
Let l be 'a gift is in the left hand' and r be 'a gift is in the right'.



We can translate G into the language of epistemic logic as

(G) $(r \wedge \neg K_{\mathbf{F}} r) \vee (l \wedge r \wedge \neg K_{\mathbf{F}} l)$.

Note: $\mathcal{M}, w_1 \vDash G \wedge \neg K_{\mathbf{F}} G$ and $\mathcal{M}, w_2 \vDash G \wedge \neg K_{\mathbf{F}} G$.

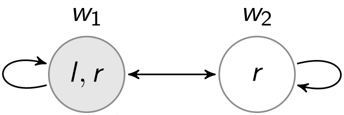


After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

It follows from the answers to 2 and 3 that $\mathcal{M}, w_1 \vDash \langle !G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

We've seen that G and $(G \wedge \neg K_{\mathbf{F}} G)$ are true at the same states in $\mathcal{M}$: w_1 and w_2.



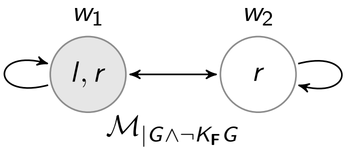
$\mathcal{M}_{|G \wedge \neg K_{\mathbf{F}} G}$

After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \models \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \models \langle !G \rangle \neg K_{\mathbf{F}} G$.

It follows from the answers to 2 and 3 that $\mathcal{M}, w_1 \models \langle !G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

We've seen that G and $(G \wedge \neg K_{\mathbf{F}} G)$ are true at the same states in $\mathcal{M}$: w_1 and w_2. Hence $\mathcal{M}, w_1 \models \langle !G \wedge \neg K_{\mathbf{F}} G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.



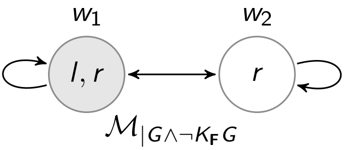
After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

It follows from the answers to 2 and 3 that
$\mathcal{M}, w_1 \vDash \langle !G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

We've seen that G and $(G \wedge \neg K_{\mathbf{F}} G)$ are true at the same states in $\mathcal{M}$: w_1 and w_2. Hence $\mathcal{M}, w_1 \vDash \langle !G \wedge \neg K_{\mathbf{F}} G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

$[!\varphi \wedge \neg K\varphi]\neg(\varphi \wedge \neg K\varphi)$ is not valid for all φ.



After my announcement of G ...

2. Is G still true? Yes. $\mathcal{M}, w_1 \vDash \langle !G \rangle G$.
3. Does **F** now know G? No. $\mathcal{M}, w_1 \vDash \langle !G \rangle \neg K_{\mathbf{F}} G$.

It follows from the answers to 2 and 3 that $\mathcal{M}, w_1 \vDash \langle !G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

We've seen that G and $(G \wedge \neg K_{\mathbf{F}} G)$ are true at the same states in $\mathcal{M}$: w_1 and w_2. Hence $\mathcal{M}, w_1 \vDash \langle !G \wedge \neg K_{\mathbf{F}} G \rangle (G \wedge \neg K_{\mathbf{F}} G)$.

$[!\varphi \wedge \neg K \varphi] \neg (\varphi \wedge \neg K \varphi)$ is not valid for all φ.

Moorean utterances are not always self-refuting.

What's Wrong with Moore Sentences?

Is there a φ such that if you receive the true information (from a source you know to be infallible) that "you don't know it, but φ," it can *remain true* afterward that you don't know it, but φ?

> If you know that I am well informed and if I address the words ... to you, these words have a curious effect which may perhaps be called anti-performatory. You may come to know that what I say *was* true, but saying it in so many words has the effect of making what is being said false. (68-69)

J. Hintikka 1962. *Knowledge and Belief*.

Surprisingly, this is not always the case, as we just showed.