Dynamic Epistemic Logic — an overview

Hans van Ditmarsch
Logic, University of Sevilla, Spain, hvd@us.es

personal.us.es/hvd/

- the logic of knowledge and belief
- public announcement logic
- non-public actions
- factual change
- Moore-sentences
- knowability
- epistemic protocol synthesis

Sevilla

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2

<u>012</u>

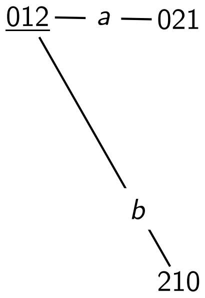
The state 012 represents the actual deal of cards.

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2

$\underline{012}$ — a — 021

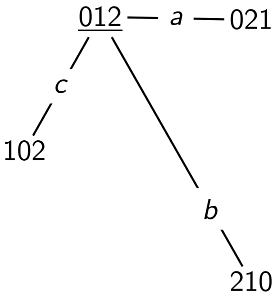
The state 012 represents the actual deal of cards.
Anne is uncertain between card deals 012 and 021.

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2



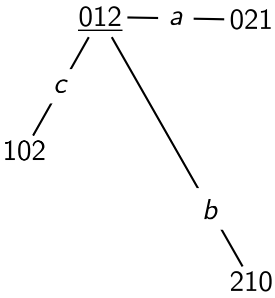
The state 012 represents the actual deal of cards.
Anne is uncertain between card deals 012 and 021.
Bill is uncertain between card deals 012 and 210.

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2



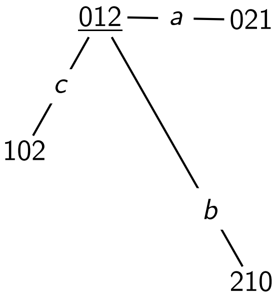
The state 012 represents the actual deal of cards.
Anne is uncertain between card deals 012 and 021.
Bill is uncertain between card deals 012 and 210.
Cath is uncertain between card deals 012 and 102.

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2



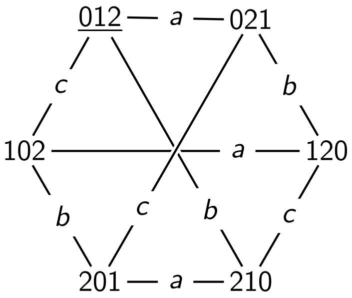
The state 012 represents the actual deal of cards.
Anne is uncertain between card deals 012 and 021.
Bill is uncertain between card deals 012 and 210.
Cath is uncertain between card deals 012 and 102.
But...

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2



The state 012 represents the actual deal of cards.
Anne is uncertain between card deals 012 and 021.
Bill is uncertain between card deals 012 and 210.
Cath is uncertain between card deals 012 and 102.
But... Anne holds 0, but Anne considers is possible that Bill considers it possible that Cath holds 0, namely in deal 210...

Three agents: Anne, Bill, Cath draw cards 0, 1, and 2



- Anne knows that Bill knows that Cath knows her own card: $K_a K_b (K_c 0_c \vee K_c 1_c \vee K_c 2_c)$
- Anne has card 0, but she considers it possible that Bill considers it possible that Cath knows that Anne does not have card 0: $0_a \wedge \hat{K}_a \hat{K}_b K_c \neg 0_a$

Multi-agent Epistemic Logic – Syntax & Semantics

Language $\varphi \quad ::= \quad p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi$

Structures

A *Kripke model* is a structure $M = \langle S, R, V \rangle$, where

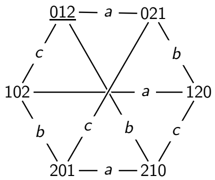
- *domain* S is a nonempty set of states;
- R yields an *accessibility relation* $R_a \subseteq S \times S$ for every $a \in A$;
- V is a *valuation* (function) $V : P \to \mathcal{P}(S)$.

If all R_a are equivalence relations $\sim_a$, M is an *epistemic model*.
A pointed epistemic model is an *epistemic state* (M, s).

Semantics

$$\begin{array}{lll} M, s \models p & \text{iff} & s \in V(p) \\ M, s \models (\varphi \wedge \psi) & \text{iff} & M, s \models \varphi \text{ and } M, s \models \psi \\ M, s \models \neg\varphi & \text{iff} & \text{not } (M, s \models \varphi) \\ M, s \models K_a\varphi & \text{iff} & \text{for all } t \text{ such that } s \sim_a t \text{ it holds that } M, t \models \varphi \end{array}$$

Example



$$Hexa, 012 \models \hat{K}_a \hat{K}_b K_c \neg 0_a$$
$$\Leftarrow$$
$$012 \sim_a 021 \text{ and } Hexa, 021 \models \hat{K}_b K_c \neg 0_a$$
$$\Leftarrow$$
$$021 \sim_b 120 \text{ and } Hexa, 120 \models K_c \neg 0_a$$
$$\Leftarrow$$
$$\sim_c (120) = \{120, 210\},\ Hexa, 120 \models \neg 0_a \text{ and } Hexa, 210 \models \neg 0_a$$
$$\Leftarrow$$
$$Hexa, 120 \not\models 0_a \text{ and } Hexa, 210 \not\models 0_a$$
$$\Leftarrow$$
$$120, 210 \not\in V(0_a) = \{012, 021\}$$

Axiomatization

all instantiations of propositional tautologies
$K_a(\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$
$K_a\varphi \rightarrow \varphi$
$K_a\varphi \rightarrow K_aK_a\varphi$
$\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$
From φ and $\varphi \rightarrow \psi$, infer ψ
From φ, infer $K_a\varphi$

Intermezzo — Common knowledge

- language: $\varphi \quad ::= \quad p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_B\varphi$
- accessibility: $\sim_B := (\bigcup_{a \in B} \sim_a)^*$
- semantics:
  $M, s \models C_B\varphi$ iff for all $t : s \sim_B t$ implies $M, t \models \varphi$

Common knowledge has the properties of individual knowledge, and the axiomatization can be extended, e.g., with *induction*:

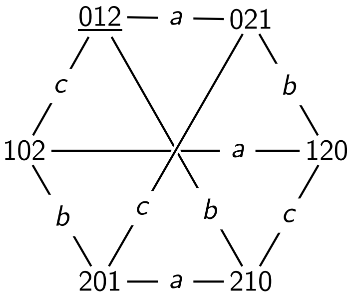
$$C_B(\varphi \rightarrow \bigwedge_{a \in B} K_a\varphi) \rightarrow (\varphi \rightarrow C_B\varphi)$$

Recent technical innovation: *conditional common knowledge* $C_B^{\psi}\varphi$

'along all the B-paths satisfying ψ it holds that φ.'

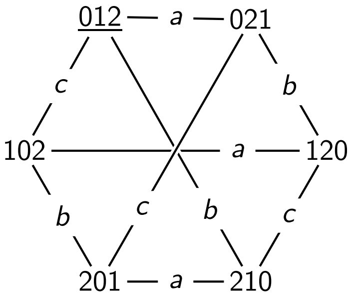
We have that $C_B^{\top}\varphi$ iff $C_B\varphi$.

Public announcements: Example



- After Anne says that she does not have card 1, Cath knows that Bill has card 1.
- After Anne says that she does not have card 1, Cath knows Anne's card.
- Bill still doesn't know Anne's card after that.

Public announcements: Example



- After Anne says that she does not have card 1, Cath knows that Bill has card 1.
  $[\neg 1_a]K_c 1_b$
- After Anne says that she does not have card 1, Cath knows Anne's card.
  $[\neg 1_a](K_c 0_a \vee K_c 1_a \vee K_c 2_a)$
- Bill still doesn't know Anne's card after that:
  $[\neg 1_a]\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$

Public Announcement Logic: language

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

Write $\langle\varphi\rangle\psi$ for $\neg[\varphi]\neg\psi$

For $[\varphi]\psi$ read "after the announcement of φ, ψ (is true)."

For $\langle\varphi\rangle\psi$ read "φ is true and after the announcement of φ, ψ."

Public Announcement Logic: semantics

The effect of the public announcement of φ is the restriction of the epistemic state to all states where φ holds. So, 'announce φ' can be seen as an epistemic state transformer, with a corresponding dynamic modal operator $[\varphi]$.

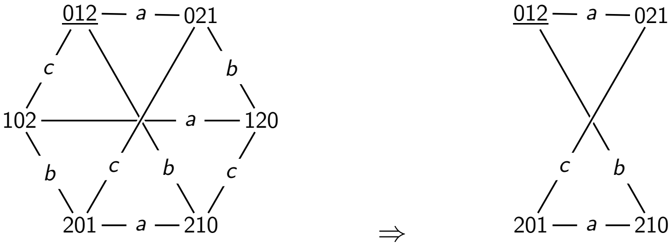
'φ is the announcement'
means
'φ is publicly and truthfully announced'.

$$M, s \models [\varphi]\psi \text{ iff } (M, s \models \varphi \text{ implies } M|\varphi, s \models \psi)$$

$M|\varphi := \langle S', \sim', V' \rangle$:

$$\begin{array}{lcll} S' & := & [\![\varphi]\!]_M & := \{s \in S \mid M, s \models \varphi\} \\ \sim'_a & := & \sim_a \cap ([\![\varphi]\!]_M \times [\![\varphi]\!]_M) & \\ V'(p) & := & V(p) \cap [\![\varphi]\!]_M & \end{array}$$

Example announcement in Hexa



$$Hexa, 012 \models \langle \neg 1_a \rangle K_c 0_a$$

$$\Leftarrow$$

$$Hexa, 012 \models \neg 1_a \text{ and } Hexa|\neg 1_a, 012 \models K_c 0_a$$

$$\Leftarrow$$

$$Hexa, 012 \models \neg 1_a \ \& \ (Hexa|\neg 1_a, 012 \models 0_a \ \& \ \sim_c(012) = \{012\})$$

$$\Leftarrow$$

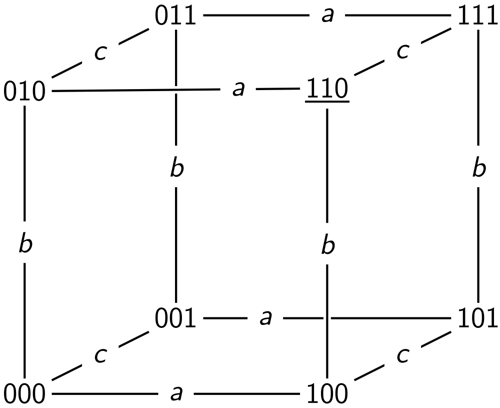
$$012 \neq V(1_a) \text{ and } 012 \in V'(0_a)$$

A dynamic epistemic logic classic

Muddy Children

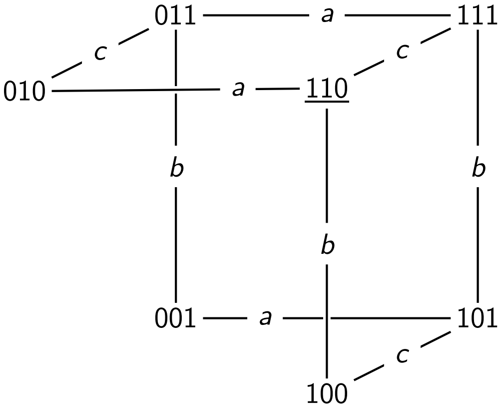
A group of children has been playing outside and are called back into the house by their father. The children gather round him. As one may imagine, some of them have become dirty from the play and in particular: they may have mud on their forehead. Children can only see whether other children are muddy, and not if there is any mud on their own forehead. All this is commonly known, and the children are, obviously, perfect logicians. Father now says: "At least one of you has mud on his or her forehead." And then: "Will those who know whether they are muddy step forward." If nobody steps forward, father keeps repeating the request. What happens?

Muddy Children



Given: The children can see each other

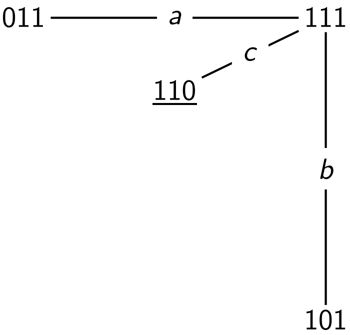
Muddy Children



After: At least one of you has mud on his or her forehead.

$m_a \vee m_b \vee m_c$

Muddy Children



After: Will those who know whether they are muddy step forward?

$\neg(\ (K_a m_a \vee K_a \neg m_a) \vee (K_b m_b \vee K_b \neg m_b) \vee (K_c m_c \vee K_c \neg m_c)\ )$

Muddy Children

<u>110</u>

After: Will those who know whether they are muddy step forward?
$(K_a m_a \vee K_a \neg m_a) \vee (K_b m_b \vee K_b \neg m_b) \vee (K_c m_c \vee K_c \neg m_c)$

On the origin of Muddy Children

On the origin of Muddy Children

German translation of Rabelais' Gargantua et Pantagruel:
Gottlob Regis, *Meister Franz Rabelais der Arzeney Doctoren Gargantua und Pantagruel, usw.*, Barth, Leipzig, 1832.

Ungelacht pfetz ich dich. Gesellschaftsspiel. Jeder zwickt seinen rechten Nachbar an Kinn oder Nase; wenn er lacht, giebt er ein Pfand. Zwei von der Gesellschaft sind nämlich im Complot und haben einen verkohlten Korkstöpsel, woran sie sich die Finger, und mithin denen, die sie zupfen, die Gesichter schwärzen. Diese werden nun um so lächerlicher, weil jeder glaubt, man lache über den anderen.

I pinch you without laughing. Parlour game. Everybody pinches his right neighbour into chin or nose; if one laughs, one must give a pledge. Two in the round have secretly blackened their fingers on a charred piece of cork, and hence will blacken the faces of their neighbours. These neighbours make a fool of themselves, since they both think that everybody is laughing about the other one.

Axiomatization of public announcement logic

$[\varphi]p \leftrightarrow (\varphi \rightarrow p)$
$[\varphi]\neg\psi \leftrightarrow (\varphi \rightarrow \neg[\varphi]\psi)$
$[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$
$[\varphi]K_a\psi \leftrightarrow (\varphi \rightarrow K_a[\varphi]\psi)$
$[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$
From φ, infer $[\psi]\varphi$
From $\chi \rightarrow [\varphi]\psi$ and $\chi \wedge \varphi \rightarrow E_B\chi$, infer $\chi \rightarrow [\varphi]C_B\psi$

Expressivity (Plaza, Gerbrandy): *Every formula in the language of public announcement logic without common knowledge is equivalent to a formula in the language of epistemic logic.*

Axiomatization of public announcement logic

$[\varphi]p \leftrightarrow (\varphi \rightarrow p)$
$[\varphi]\neg\psi \leftrightarrow (\varphi \rightarrow \neg[\varphi]\psi)$
$[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$
$[\varphi]K_a\psi \leftrightarrow (\varphi \rightarrow K_a[\varphi]\psi)$
$[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$
From φ, infer $[\psi]\varphi$
From $\chi \rightarrow [\varphi]\psi$ and $\chi \wedge \varphi \rightarrow E_B\chi$, infer $\chi \rightarrow [\varphi]C_B\psi$

Expressivity (Plaza, Gerbrandy): *Every formula in the language of public announcement logic without common knowledge is equivalent to a formula in the language of epistemic logic.*

Announcement and relativized common knowledge

$[\varphi]C_B^{\chi}\psi \leftrightarrow C_B^{\varphi \wedge [\varphi]\chi}[\varphi]\psi$

Sequence of announcements

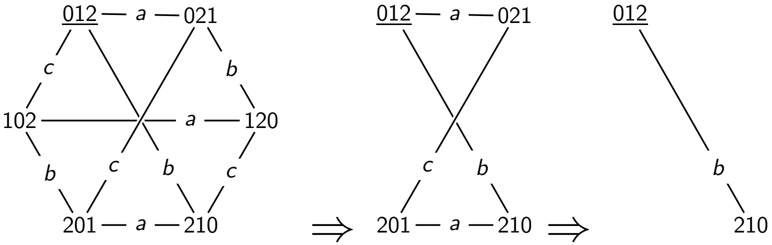
$[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$

Anne does not have card 1, and Cath now knows Anne's card.
Sequence of two announcements:

$$\neg 1_a \; ; \; (K_c 0_a \vee K_c 1_a \vee K_c 2_a)$$

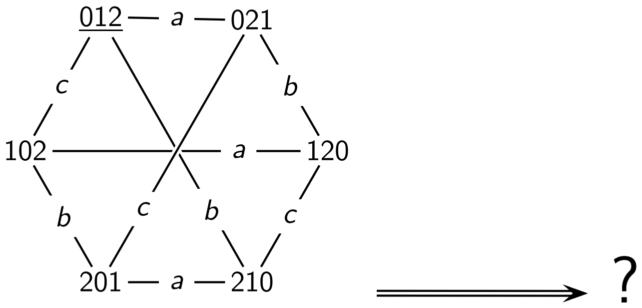
Single announcement:

$$\neg 1_a \wedge [\neg 1_a](K_c 0_a \vee K_c 1_a \vee K_c 2_a)$$



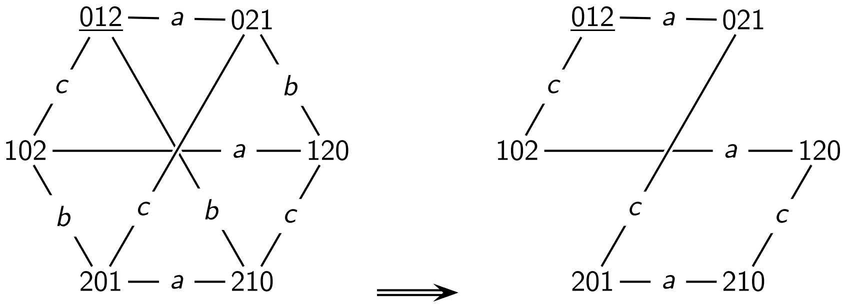
Intermezzo — More complex dynamics (= non-public)

(Anne holds 0, Bill holds 1, and Cath holds 2.) Anne shows (only) Bill her card. (She shows card 0.) Cath cannot see the face of the shown card, but notices that a card is being shown.

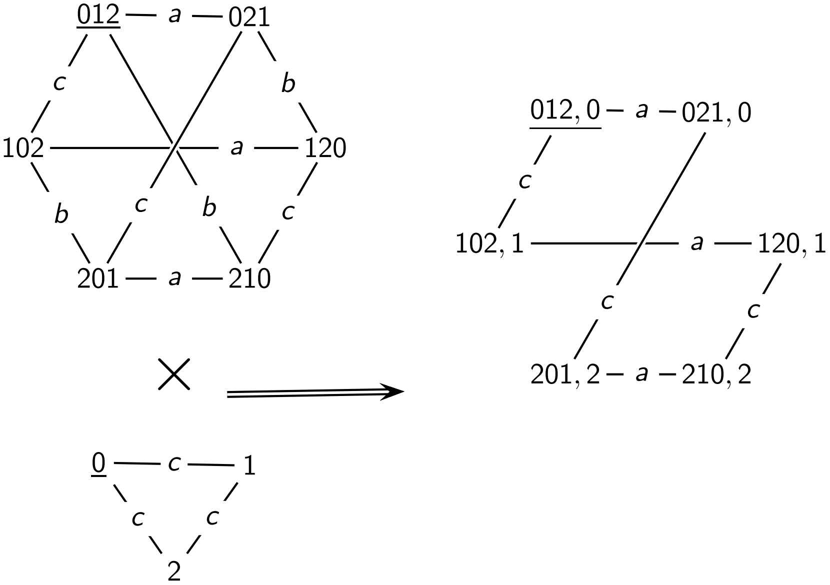


Intermezzo — More complex dynamics (= non-public)

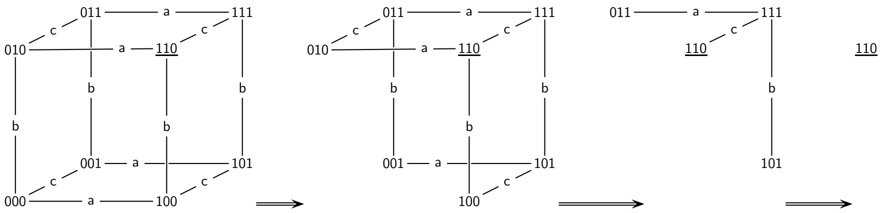
(Anne holds 0, Bill holds 1, and Cath holds 2.) Anne shows (only) Bill her card. (She shows card 0.) Cath cannot see the face of the shown card, but notices that a card is being shown.



Intermezzo — Anne shows card 0 to Bill



Muddy Children — No, not again!!

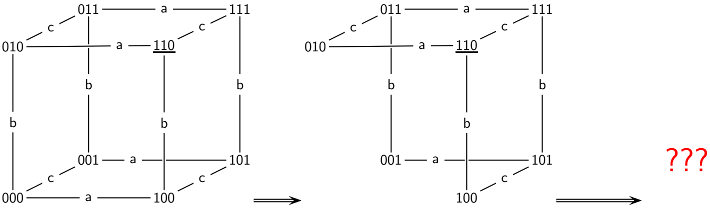


- At least one of you has mud on his or her forehead.
- Will those who know whether they are muddy step forward?
- Will those who know whether they are muddy step forward?

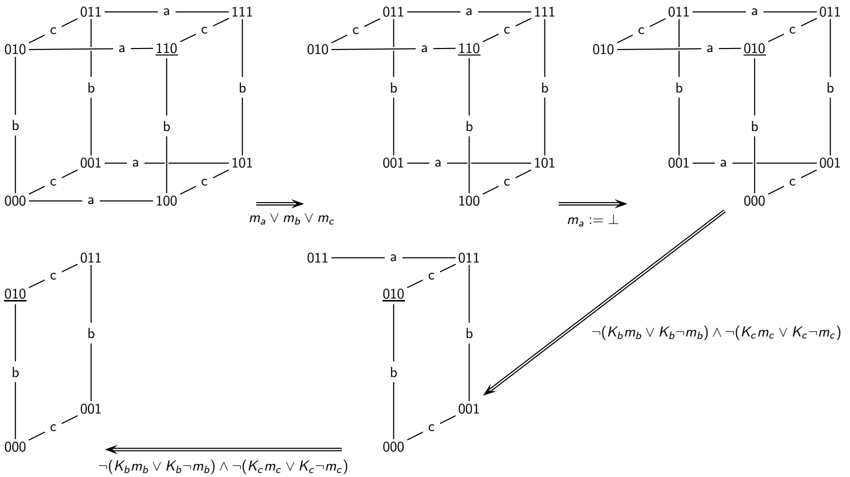
Muddy Children — with cleaning!!

Suppose that after telling the children that at least one of them is muddy... father empties a bucket of water over Anne (splash!).

Muddy Children — with cleaning



- At least one of you has mud on his or her forehead.
- Father empties a bucket of water over Anne (splash!). ?
- Will those who know whether they are muddy step forward? ?
- Will those who know whether they are muddy step forward? ?



- Last step: Cath learns that Anne knows that she **was** muddy.
- Assignment reduction principle: $[p := \varphi]p \leftrightarrow \varphi$
- Applications to the computational frame problem.

Unsuccessful updates and Moore sentences

Unsuccessful updates and Moore sentences

You do not know that this is my first visit to Brazil.

Unsuccessful updates and Moore sentences

You do not know that this is my first visit to Brazil.

So what?

Unsuccessful updates and Moore sentences

You do not know that this is my first visit to Brazil.

So what?

You do not know that this is my first visit to Brazil.

Unsuccessful updates and Moore sentences

You do not know that this is my first visit to Brazil.

So what?

You do not know that this is my first visit to Brazil.

Liar!

Unsuccessful updates and Moore sentences

You do not know that this is my first visit to Brazil.

So what?

You do not know that this is my first visit to Brazil.

Liar!

- You do not know that this is my first visit to Brazil: $p \wedge \neg Kp$.
- After announcing this you know that p: $[p \wedge \neg Kp]Kp$ is valid.
- And therefore also $[p \wedge \neg Kp](\neg p \vee Kp)$.
- And therefore also $[p \wedge \neg Kp]\neg(p \wedge \neg Kp)$.
- I cannot truthfully announce $p \wedge \neg Kp$ twice.

Unsuccessful updates and Moore sentences

Postulate of success:

$$\varphi \to \langle \varphi \rangle K\varphi$$

Announcement of a *fact* always makes it public:

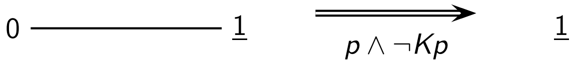
$$\models [p]Kp$$

Announcements of non-facts do not have to make them public:

$$\not\models [\varphi]K\varphi$$

It can be even worse, as we have seen:

$$\models [p \wedge \neg Kp]\neg(p \wedge \neg Kp)$$
$$\models [p \wedge \neg Kp]\neg K(p \wedge \neg Kp)$$



Fitch paradox of knowability

Fitch's paradox is that:
there is an unknown truth is incons. with all truths are knowable.
$\exists p(p \wedge \neg Kp)$ is inconsistent with $\forall q(q \rightarrow \Diamond Kq)$.

Substitute $p \wedge \neg Kp$ for q and you get: $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$.
For any 'reasonable' interpretation of $\Diamond$, this can't be.

Fitch paradox of knowability

Fitch's paradox is that:
there is an unknown truth is incons. with all truths are knowable.
$\exists p(p \wedge \neg Kp)$ is inconsistent with $\forall q(q \rightarrow \Diamond Kq)$.

Substitute $p \wedge \neg Kp$ for q and you get: $(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp)$.
For any 'reasonable' interpretation of $\Diamond$, this can't be.

But what is reasonable? Consider the following:

Interpret
$\varphi \rightarrow \Diamond K\varphi$
as
If φ is true, then there is an announcement after which φ is known.

This is false when φ is $p \wedge \neg Kp$.

Arbitrary announcement logic – $\Diamond(Kp \vee K\neg p)$ is valid

$\Diamond\varphi$ is true in a model, iff
there is an epistemic ψ such that $\langle\psi\rangle\varphi$ is true, iff
there is a ... model restriction such that φ is true in the restriction.

$\underline{1}$——0 $\overset{p}{\Rightarrow}$ $\underline{1}$

$\Diamond(Kp \vee K\neg p), \langle p\rangle(Kp \vee K\neg p)$ — p, Kp

1——$\underline{0}$ $\overset{\neg p}{\Rightarrow}$ $\underline{0}$

$\Diamond(Kp \vee K\neg p), \langle\neg p\rangle(Kp \vee K\neg p)$ — $\neg p, K\neg p$

Moore-sentence:
$p \wedge \neg Kp$ $\overset{p \wedge \neg Kp}{\Rightarrow}$ $Kp, \neg p \vee Kp, \neg(p \wedge \neg Kp)$

Further issues in dynamic epistemic logic

- AGM-type belief revision
  (from belief in $\neg p$ to belief in p).
- Logics with quantification over information change
  (as in $\varphi \to \Diamond K\varphi$)
- Dynamic epistemic logic and temporal epistemic logic
  (replace $[\varphi]\psi$ by $X\psi$)
- Public announcement protocols for three agents
  (sender, receiver, eavesdropper)
- Protocol synthesis and knowledge progression.

Feria de Sevilla