# COMPUTATIONAL LEARNING IN DYNAMIC LOGICS DAY 2: DYNAMIC EPISTEMIC LOGIC AND BELIEF REVISION

Nina Gierasimczuk and Caleb Schultz Kisby

@NASSLLI, June 2025

## Course Homepage:

https://sites.google.com/view/nasslli25-learning-in-del

## PLAN FOR TODAY

Muddy Children Puzzle

2 Dynamic Epistemic Logic

3 Doxastic Logic and Belief Revision

## PLAN FOR TODAY

Muddy Children Puzzle

2 Dynamic Epistemic Logic

3 Doxastic Logic and Belief Revision

## MUDDY CHILDREN REVISITED

Imagine *n* children are playing outside together. Now it happens that during their play some of them, say *k* get mud on their foreheads. Each can see mud on others but not on his own forehead.

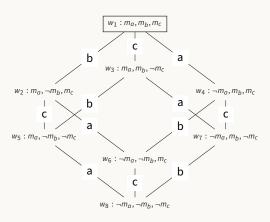
Along comes the father, who says, "At least one of you have mud on your forehead". The father then asks the following question, over and over: "Does any of you know whether you have mud on your own forehead?" Assuming that all the children are perceptive, intelligent, truthful, and they answer simultaneously, what will happen?

Surprisingly, after the father asks the question for the  $k^{th}$  time all muddy children will say "yes". How come?

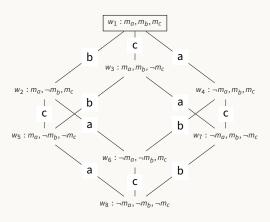


## MUDDY CHILDREN: THE UNDERLYING ASSUMPTIONS

- Common knowledge that the father is truthful,
- that all the children hear the father,
- that all the children see each other,
- that none of them can see their own forehead,
- and that all the children are truthful and intelligent.

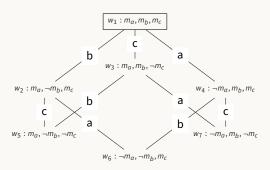


Before the announcement.



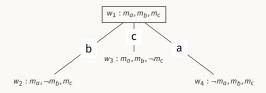
# Father says:

"
$$m_a \vee m_b \vee m_c$$
"



- Does any of you know?

- No!



- Does any of you know?

- No!

 $w_1:m_a,m_b,m_c$ 

- Does any of you know?

– Yes!

## PLAN FOR TODAY

1 Muddy Children Puzzle

2 Dynamic Epistemic Logic

3 Doxastic Logic and Belief Revision

#### LOGICS OF PUBLIC ANNOUNCEMENTS

- PAL (Public Announcement Logic) was proposed by Jan Plaza in 1989
- PAC (Public Announcement logic with common knowledge) is PAL + C
- PAL and PAC are examples of dynamic epistemic logic.

Dynamic Epistemic Logics formalize informational changes: the dynamics of knowledge/belief.

## DYNAMIC MODALITIES

To express informational changes, dynamic epistemic logics use a new kind of operators, called **dynamic modalities**:

$$[\alpha]\phi$$
,

where  $\boldsymbol{\alpha}$  is the name of some action involving communication.

## DYNAMIC MODALITIES

To express informational changes, dynamic epistemic logics use a new kind of operators, called **dynamic modalities**:

$$[\alpha]\phi$$
,

where  $\alpha$  is the name of some action involving communication.

Such actions are called **epistemic actions** (as opposed to ontic actions) since they affect only the knowledge/beliefs of the agents.

## DYNAMIC MODALITIES

To express informational changes, dynamic epistemic logics use a new kind of operators, called **dynamic modalities**:

$$[\alpha]\phi$$
,

where  $\alpha$  is the name of some action involving communication.

Such actions are called **epistemic actions** (as opposed to ontic actions) since they affect only the knowledge/beliefs of the agents.

The intended meaning of  $[\alpha]\phi$  is: if action  $\alpha$  is performed, then  $\phi$  will become true.

## EXAMPLE: THE PUBLIC ANNOUNCEMENT MODALITY

An example is the **truthful public announcement** of some sentence  $\boldsymbol{\phi} \colon$ 

 $[!\phi]$ 

#### **EXAMPLE: THE PUBLIC ANNOUNCEMENT MODALITY**

An example is the **truthful public announcement** of some sentence  $\phi$ :

 $[!\phi]$ 

The intended meaning of  $[!\phi]\psi$  is:

if a truthful public announcement of  $\phi$  is performed, then  $\psi$  will become true.

# WHAT HAPPENS IF THE ANNOUNCEMENT IS FALSE?

 $!\phi \text{ can only be performed if } \phi \text{ is } \textbf{true},$  so  $[!\phi]\psi$  is by definition true in worlds where  $\phi$  is false.

# WHAT HAPPENS IF THE ANNOUNCEMENT IS FALSE?

 $!\phi \text{ can only be performed if } \phi \text{ is } \textbf{true},$  so  $[!\phi]\psi$  is by definition true in worlds where  $\phi$  is false.

In particular, if a false sentence is "truthfully announced", then everything is true after that (including contradictions):

$$\neg \phi \rightarrow [!\phi] \bot$$
,

where  $\bot$  is any sentence that is always false (contradictory).

## LANGUAGE OF PUBLIC ANNOUNCEMENT LOGIC

# **Definition (Syntax)**

 $\Phi$  is a set of propositions, with  $p \in \Phi$ , and  $A = \{1, ..., n\}$  is a set of agents.

$$\varphi := \top \mid p \mid \dots \mid K_i \varphi \mid [!\varphi] \varphi$$

where  $\top$  abbreviates a tautology and  $i \in A$  is the name of some agent.

## LANGUAGE OF PUBLIC ANNOUNCEMENT LOGIC

# **Definition (Syntax)**

 $\Phi$  is a set of propositions, with  $p \in \Phi$ , and  $A = \{1, ..., n\}$  is a set of agents.

$$\varphi := \top \mid p \mid \dots \mid K_i \varphi \mid [!\varphi] \varphi$$

where  $\top$  abbreviates a tautology and  $i \in A$  is the name of some agent.

As before, these formulas are interpreted in possible world models.

## PUBLIC ANNOUNCEMENT AS JOINT UPDATE

How can we model the effect of a public announcement?

#### PUBLIC ANNOUNCEMENT AS JOINT UPDATE

How can we model the effect of a public announcement?

As we saw in many examples, this can be done by **deleting worlds**.

LEARNING = ELIMINATING POSSIBILITIES

## PUBLIC ANNOUNCEMENT AS JOINT UPDATE

How can we model the effect of a public announcement?

As we saw in many examples, this can be done by **deleting worlds**.

LEARNING = ELIMINATING POSSIBILITIES

From now on, we denote by  $!\phi$  the operation of deleting the non- $\phi$  worlds, and call it **public announcement with**  $\phi$ , or **joint update with**  $\phi$ .

## SEMANTICS OF PUBLIC ANNOUNCEMENT LOGIC

## **Definition**

Let 
$$M = (S, \pi, \mathcal{K}_1, \dots, \mathcal{K}_n)$$
.  
 $(M, s) \vDash p$  iff  $\pi(s)(p) = 1$   
 $\dots$   
 $(M, s) \vDash \mathcal{K}_i \varphi$  iff for all  $v$  with  $(s, v) \in \mathcal{K}_i$ ,  $(M, v) \vDash \varphi$   
 $(M, s) \vDash [!\varphi] \psi$  iff if  $(M, s) \vDash \varphi$  then  $(M|\varphi, s) \vDash \psi$ 

where  $M|\phi = (S', \pi', \mathcal{K}'_1, \dots, \mathcal{K}'_n)$  is defined as follows:

- $S' := \{ s \in S \mid (M, s) \models \phi \}$
- $\pi' := \pi$  restricted to S'
- for each  $i \in \{1, ..., n\}$ ,  $\mathcal{K}'_i := \mathcal{K}_i \cap (S' \times S')$

# ARE SENTENCES KNOWN AFTER TRUTHFULLY ANNOUNCED?

Intuitively, it may seem that: every sentence becomes known after it is truthfully publicly announced.

(?)  $[!\phi]K_a\phi$ , for any sentence  $\phi$  and any agent a.

# ARE SENTENCES KNOWN AFTER TRUTHFULLY ANNOUNCED?

Intuitively, it may seem that: every sentence becomes known after it is truthfully publicly announced.

(?)  $[!\phi]K_a\phi$ , for any sentence  $\phi$  and any agent a.

If the above were true then, assuming truthfulness, we would also get:

(?)  $[!\phi]\phi$ , for any sentence  $\phi$ .

#### **MOORE SENTENCES**

Two stockbrokers Alice and Bob are lunching in a Wall Street café. A messenger comes in and delivers a letter to Alice. On the envelope it is written 'urgently requested data on United Agents'. Alice opens and reads the letter, which informs her of the fact that United Agents is doing well, such that she intends to buy a portfolio of stocks of that company, immediately.



Alice says to Bob:

'UA is doing well.'

#### **MOORE SENTENCES**

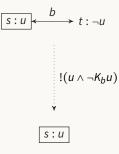
Two stockbrokers Alice and Bob are lunching in a Wall Street café. A messenger comes in and delivers a letter to Alice. On the envelope it is written 'urgently requested data on United Agents'. Alice opens and reads the letter, which informs her of the fact that United Agents is doing well, such that she intends to buy a portfolio of stocks of that company, immediately.



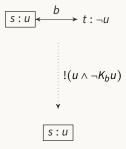
Alice says to Bob:

'Guess you don't know it, but UA is doing well.'

# UNSUCCESSFUL UPDATE

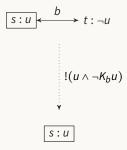


## UNSUCCESSFUL UPDATE



Alice truthfully announced something that became false after the announcement!

## UNSUCCESSFUL UPDATE



Alice truthfully announced something that became false after the announcement!

$$[!(u \wedge \neg K_b u)] \neg (u \wedge \neg K_b u)$$

In PAL we can iterate and combine announcements.

In PAL we can iterate and combine announcements.

We can announce not only facts: [!u]

In PAL we can iterate and combine announcements.

We can announce not only facts: [!u]

or combinations of facts (Boolean formulas):  $[!(u \lor \neg q)]$ 

In PAL we can iterate and combine announcements.

We can announce not only facts: [!u]

or combinations of facts (Boolean formulas):  $[!(u \lor \neg q)]$ 

but also epistemic formulas:  $[!(\neg K_b u)]$ 

### ANNOUNCEMENTS ABOUT ANNOUNCEMENTS

In PAL we can iterate and combine announcements.

We can announce not only facts: [!u]

or combinations of facts (Boolean formulas):  $[!(u \lor \neg q)]$ 

but also epistemic formulas:  $[!(\neg K_b u)]$ 

and make announcements about other announcements:  $[!([!u]K_bu)].$ 

# CLOSURE OF PUBLIC ANNOUNCEMENT UNDER COMPOSITION

$$\vDash [!\phi][!\psi]\theta \leftrightarrow [!(\phi \land [!\phi]\psi)]\theta.$$

## CLOSURE OF PUBLIC ANNOUNCEMENT UNDER COMPOSITION

$$\models [!\phi][!\psi]\theta \leftrightarrow [!(\phi \land [!\phi]\psi)]\theta.$$

this expresses closure of public announcements under sequential composition:

performing successively two public announcements:  $!\phi$ ;  $!\psi$  is equivalent to performing one:  $!(\phi \land [!\phi]\psi)$ .

The following formulas are valid in  $\mathfrak{M}_n$ :

The following formulas are valid in  $\mathcal{M}_n$ :

**Atomic Permanence**  $[!\phi]p \leftrightarrow (\phi \rightarrow p)$ , for atomic propositions p

The following formulas are valid in  $\mathcal{M}_n$ :

**Atomic Permanence**  $[!\phi]p \leftrightarrow (\phi \rightarrow p)$ , for atomic propositions p

Announcement-Negation  $[!\phi]\neg\psi\leftrightarrow(\phi\rightarrow\neg[!\phi]\psi)$ 

The following formulas are valid in  $\mathcal{M}_n$ :

**Atomic Permanence**  $[!\phi]p \leftrightarrow (\phi \rightarrow p)$ , for atomic propositions p

Announcement-Negation  $[!\phi]\neg\psi\leftrightarrow (\phi\rightarrow\neg[!\phi]\psi)$ 

Announcement-Conjunction  $[!\phi](\psi \wedge \theta) \leftrightarrow ([!\phi]\psi \wedge [!\phi]\theta)$ 

The following formulas are valid in  $\mathcal{M}_n$ :

**Atomic Permanence**  $[!\phi]p \leftrightarrow (\phi \rightarrow p)$ , for atomic propositions p

Announcement-Negation  $[!\phi]\neg\psi\leftrightarrow (\phi\rightarrow\neg[!\phi]\psi)$ 

Announcement-Conjunction  $[!\phi](\psi \wedge \theta) \leftrightarrow ([!\phi]\psi \wedge [!\phi]\theta)$ 

Announcement-Knowledge  $[!\phi]K_i\psi \leftrightarrow (\phi \rightarrow K_i[!\phi]\psi)$ 

The following formulas are valid in  $\mathcal{M}_n$ :

**Atomic Permanence**  $[!\phi]p \leftrightarrow (\phi \rightarrow p)$ , for atomic propositions p

Announcement-Negation  $[!\phi]\neg\psi\leftrightarrow (\phi\rightarrow\neg[!\phi]\psi)$ 

Announcement-Conjunction  $[!\phi](\psi \wedge \theta) \leftrightarrow ([!\phi]\psi \wedge [!\phi]\theta)$ 

Announcement-Knowledge  $[!\phi]K_i\psi \leftrightarrow (\phi \rightarrow K_i[!\phi]\psi)$ 

Using those reduction axioms, one can translate formulas with announcement modalities into ones without.

This shows PAL can be reduced to Epistemic Logic.

### DYNAMIC EPISTEMIC LOGIC IN GENERAL

- DEL comprises a family of logics.
- Each has syntax and semantics.
- DEL concerns explicit informational actions.
- Corresponding knowledge and belief changes in agents.
- Often uses special action models.

We use the action models of DEL with postconditions (ontic actions).

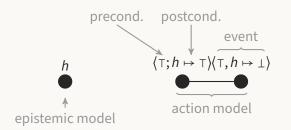
h:="the coin faces heads up"



Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

We use the action models of DEL with postconditions (ontic actions).

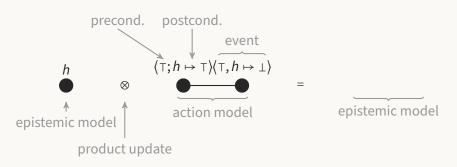
h:="the coin faces heads up"

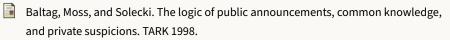


- Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.
- van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

We use the action models of DEL with postconditions (ontic actions).

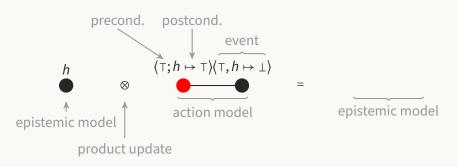
h:="the coin faces heads up"

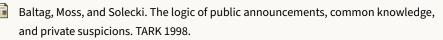




We use the action models of DEL with postconditions (ontic actions).

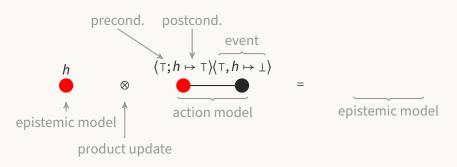
h:="the coin faces heads up"

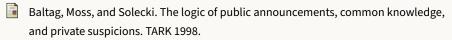




We use the action models of DEL with postconditions (ontic actions).

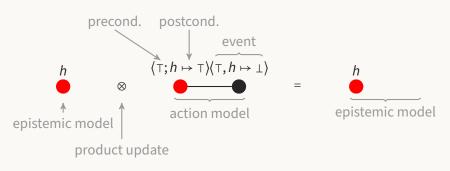
h:="the coin faces heads up"

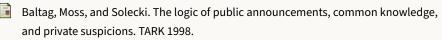




We use the action models of DEL with postconditions (ontic actions).

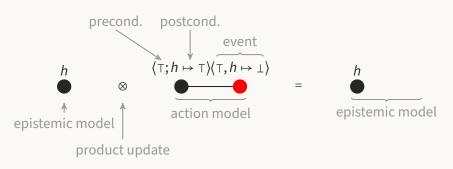
h:="the coin faces heads up"

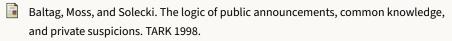




We use the action models of DEL with postconditions (ontic actions).

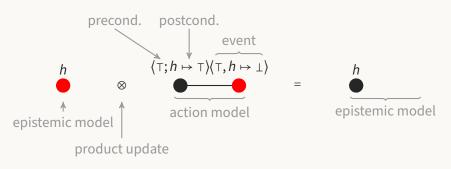
h:="the coin faces heads up"

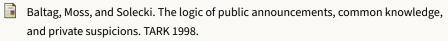




We use the action models of DEL with postconditions (ontic actions).

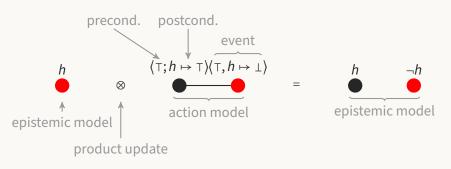
h:="the coin faces heads up"

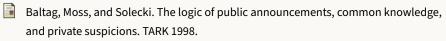




We use the action models of DEL with postconditions (ontic actions).

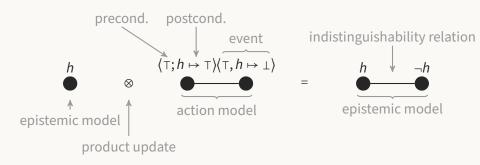
h:="the coin faces heads up"

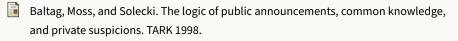




We use the action models of DEL with postconditions (ontic actions).

h:="the coin faces heads up"





## PLAN FOR TODAY

1 Muddy Children Puzzle

2 Dynamic Epistemic Logic

3 Doxastic Logic and Belief Revision

# DOXASTIC LOGIC AND BELIEF REVISION

• So far, we have only talked about learning in terms of knowledge update

# DOXASTIC LOGIC AND BELIEF REVISION

- So far, we have only talked about learning in terms of knowledge update
- **Knowledge is monotonic:** Once  $\varphi$  is *known*, it cannot be taken back or revised. (*hard information*)

# DOXASTIC LOGIC AND BELIEF REVISION

- $\bullet~$  So far, we have only talked about learning in terms of  ${\it knowledge}~{\it update}$
- **Knowledge is monotonic:** Once  $\varphi$  is *known*, it cannot be taken back or revised. (*hard information*)
- Belief is not monotonic: Our beliefs are tentative convictions, which allow for exceptions and future revision (soft information)
  - An agent can believe false things
  - Beliefs can allow for exceptions, e.g., the belief that birds fly
  - An agent can revise or retract her beliefs

# THE LANGUAGE OF CONDITIONAL BELIEF

## **Definition**

Take a countable set of propositions Prop.

$$\varphi, \psi := \top \mid p \mid \neg \varphi \mid \varphi \wedge \psi \mid A\varphi \mid \mathbf{B}^{\psi} \varphi$$

for all  $p \in \mathsf{Prop}$ . The usual abbreviations are  $\lor$ ,  $\to$ , and E (dual to A)

# THE LANGUAGE OF CONDITIONAL BELIEF

# **Definition**

Take a countable set of propositions Prop.

$$\varphi, \psi := \top \mid p \mid \neg \varphi \mid \varphi \wedge \psi \mid A\varphi \mid \mathbf{B}^{\psi} \varphi$$

for all  $p \in Prop$ . The usual abbreviations are  $\lor$ ,  $\rightarrow$ , and E (dual to A)

• We read  $\mathbf{B}^{\psi}\varphi$  as saying:

"The agent believes that, in the most plausible (or normal) scenarios where  $\psi$  holds,  $\varphi$  also holds."

## THE LANGUAGE OF CONDITIONAL BELIEF

## **Definition**

Take a countable set of propositions Prop.

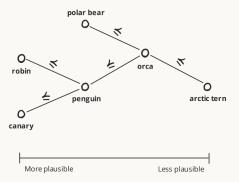
$$\varphi, \psi := \top \mid p \mid \neg \varphi \mid \varphi \wedge \psi \mid A\varphi \mid \mathbf{B}^{\psi} \varphi$$

for all  $p \in Prop$ . The usual abbreviations are  $\lor$ ,  $\rightarrow$ , and E (dual to A)

- We read  ${\bf B}^{\psi}\varphi$  as saying:
  - "The agent believes that, in the most plausible (or normal) scenarios where  $\psi$  holds,  $\varphi$  also holds."
- Example: B<sup>bird</sup>(fly) means "the agent believes that normal birds fly"
  - taking some liberties: "normally, birds fly"
  - o allows for exceptional birds that do not fly, e.g. penguins, dodos

• In order to model the agent's concept of "most normal", we use a plausibility order on worlds  $\leq: W \times W$ :

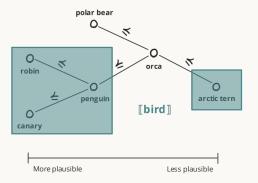
 $u \leq w$  if the agent considers u at least as plausible as w



• Let  $\mathsf{best}_{\preceq}(S) = \{ w \in S \mid \mathsf{for\ all}\ u \in S, \mathsf{not}\ u \preceq w \}$ , the set of most plausible worlds over S. In cognitive science, this is known as the *prototype* of S.

• In order to model the agent's concept of "most normal", we use a plausibility order on worlds  $\leq: W \times W$ :

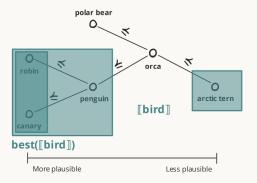
 $u \leq w$  if the agent considers u at least as plausible as w



• Let  $\mathsf{best}_{\preceq}(S) = \{ w \in S \mid \mathsf{for\ all}\ u \in S, \mathsf{not}\ u \preceq w \}$ , the set of most plausible worlds over S. In cognitive science, this is known as the *prototype* of S.

• In order to model the agent's concept of "most normal", we use a plausibility order on worlds  $\leq: W \times W$ :

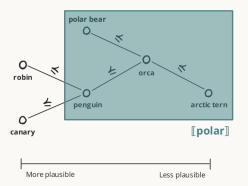
 $u \leq w$  if the agent considers u at least as plausible as w



• Let  $\mathsf{best}_{\preceq}(S) = \{w \in S \mid \mathsf{for\ all}\ u \in S, \mathsf{not}\ u \preceq w\}$ , the set of most plausible worlds over S. In cognitive science, this is known as the *prototype* of S.

• In order to model the agent's concept of "most normal", we use a plausibility order on worlds  $\leq: W \times W$ :

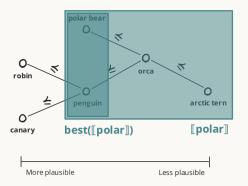
 $u \leq w$  if the agent considers u at least as plausible as w



• Let  $\mathsf{best}_{\preceq}(S) = \{w \in S \mid \mathsf{for\ all}\ u \in S, \mathsf{not}\ u \preceq w\}$ , the set of most plausible worlds over S. In cognitive science, this is known as the *prototype* of S.

• In order to model the agent's concept of "most normal", we use a plausibility order on worlds  $\leq: W \times W$ :

 $u \leq w$  if the agent considers u at least as plausible as w



• Let  $\mathsf{best}_{\preceq}(S) = \{ w \in S \mid \mathsf{for\ all}\ u \in S, \mathsf{not}\ u \preceq w \}$ , the set of most plausible worlds over S. In cognitive science, this is known as the *prototype* of S.

## SEMANTICS FOR CONDITIONAL BELIEF

# **Definition**

Given a plausibility model  $M = (W, \preceq, V)$  and a state  $s \in W$ :

$$\begin{array}{ll} M,s \vDash p & \text{iff} \quad s \in V(p) \text{ for each } p \in \mathsf{Prop} \\ M,s \vDash \neg \varphi & \text{iff} \quad \mathsf{not} \ M,s \vDash \varphi \\ M,s \vDash \varphi \wedge \psi & \text{iff} \quad M,s \vDash \varphi \text{ and } M,s \vDash \psi \\ M,s \vDash A\varphi & \text{iff} \quad M,u \vDash \varphi \text{ for all } u \in W \text{ whatsoever} \\ M,s \vDash \mathbf{B}^{\psi}\varphi & \text{iff} \quad \mathsf{best}_{\preceq}(\llbracket \psi \rrbracket) \subseteq \llbracket \varphi \rrbracket \\ & \text{"the most normal } \psi\text{-worlds are } \varphi\text{-worlds"} \end{array}$$

Again,  $\llbracket \varphi \rrbracket = \{u \mid M, u \vDash \varphi\}$  is the set of  $\varphi$ -states.

# ADDITIONAL COMMENTS ON CONDITIONAL BELIEF

- The semantics for  ${\bf B}^{\psi}\varphi$  doesn't depend on the state s at all!
  - $\circ$  We can fix this by having a different plausibility order  $\preceq_s$  per state.

$$M, s \models \mathbf{B}^{\psi} \varphi$$
 iff  $\operatorname{best}_{\prec_s}(\llbracket \psi \rrbracket) \subseteq \llbracket \varphi \rrbracket$ 

# ADDITIONAL COMMENTS ON CONDITIONAL BELIEF

- The semantics for  $\mathbf{B}^{\psi}\varphi$  doesn't depend on the state s at all!
  - $\circ$  We can fix this by having a different plausibility order  $\preceq_s$  per state.

$$M,s \vDash \mathbf{B}^{\psi} \varphi \qquad \text{iff} \qquad \mathsf{best}_{\preceq_s}(\llbracket \psi \rrbracket) \subseteq \llbracket \varphi \rrbracket$$

individual's learning policy here.

We've dropped the subscripts i — we're more interested in an

- $\circ$  We could have a different plausibility order  $\preceq_{i,s}$  per agent, per state
- This gets messy!

## ADDITIONAL COMMENTS ON CONDITIONAL BELIEF

- The semantics for  ${\bf B}^\psi \varphi$  doesn't depend on the state s at all!
  - We can fix this by having a different plausibility order  $\leq_s$  per state.

$$M,s \vDash \mathbf{B}^{\psi} \varphi \qquad \text{iff} \qquad \text{best}_{\preceq_s}(\llbracket \psi \rrbracket) \subseteq \llbracket \varphi \rrbracket$$

- ullet We've dropped the subscripts i we're more interested in an individual's learning policy here.
  - $\circ~$  We could have a different plausibility order  $\preceq_{i,s}$  per agent, per state
  - o This gets messy!
- We can also define ordinary (nonconditional) belief as

$$M,s \vDash \mathbf{B} \varphi \qquad \text{iff} \qquad \text{For all } t \in \mathsf{best}_{\preceq}(\llbracket \varphi \rrbracket), M,t \vDash \varphi$$

o But conditional belief is more expressive!

$$\models\! \mathbf{B} \varphi \! \leftrightarrow \! \mathbf{B}^{\varphi} \varphi, \text{ since } M, s \! \models \! \mathbf{B} \varphi \text{ iff best}_{\preceq}([\![\varphi]\!]) \! \subseteq [\![\varphi]\!]$$

## CHANGING OUR BELIEFS

• **First stab:** Try public announcement  $[!\varphi]$ , but for belief revision

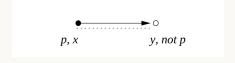
# CHANGING OUR BELIEFS

- **First stab:** Try public announcement  $[!\varphi]$ , but for belief revision
- Problem: Public announcement of p does not allow us to later retract, revise, or undo our agent's belief in p.



## CHANGING OUR BELIEFS

- **First stab:** Try public announcement  $[!\varphi]$ , but for belief revision
- Problem: Public announcement of p does not allow us to later retract,
   revise, or undo our agent's belief in p.



• We need a less destructive idea than world elimination...

# BELIEF CHANGE AS PLAUSIBILITY RE-ORDERING

• New idea: Re-order (rearrange) the plausibility relation

## BELIEF CHANGE AS PLAUSIBILITY RE-ORDERING

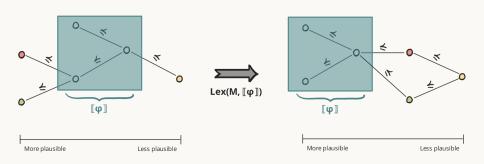
- New idea: Re-order (rearrange) the plausibility relation
- There are many, many different policies we could use to re-order!
  - Hans Rott, Shifting priorities: Simple representations for 27 iterated theory change operators (2006)

## BELIEF CHANGE AS PLAUSIBILITY RE-ORDERING

- New idea: Re-order (rearrange) the plausibility relation
- There are many, many different policies we could use to re-order!
  - Hans Rott, Shifting priorities: Simple representations for 27 iterated theory change operators (2006)
- Each revision policy represents the agent's "style" of response to incoming information (hard vs soft, radical vs minimal)

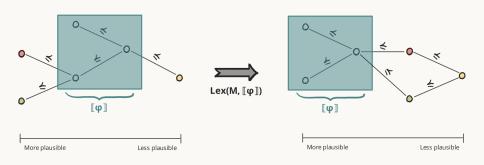
# LEXICOGRAPHIC UPGRADE

# *Make all* $\varphi$ -worlds more plausible than $\neg \varphi$ worlds



#### LEXICOGRAPHIC UPGRADE

# *Make all* $\varphi$ -worlds more plausible than $\neg \varphi$ worlds

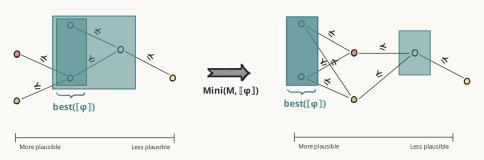


• Formally,  $\text{Lex}((W, \preceq, V), \llbracket \varphi \rrbracket) = (W, \preceq', V)$ , where the plausibility order is replaced with the following:

All  $[\![\varphi]\!]$ -worlds are  $\preceq'$ -better than all  $[\![\neg\varphi]\!]$ -worlds, but within those two groups the old ordering  $\preceq$  remains

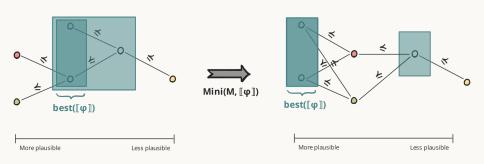
# MINIMAL UPGRADE

# Make only the **best** $\varphi$ -worlds more plausible than the rest



#### MINIMAL UPGRADE

Make only the **best**  $\varphi$ -worlds more plausible than the rest



• Formally,  $\operatorname{Mini}((W, \preceq, V), \llbracket \varphi \rrbracket) = (W, \preceq', V)$ , where the plausibility order is replaced with the following:

The best $\leq (\llbracket \varphi \rrbracket)$  worlds come on top, but otherwise the old order remains

#### DYNAMIC DOXASTIC LOGIC

## **Definition**

We extend our language from before with two new upgrade operators:

$$\varphi, \psi := \top \mid p \mid \neg \varphi \mid \varphi \wedge \psi \mid A\varphi \mid \mathbf{B}^{\psi} \varphi \mid [\!\!\uparrow \varphi] \psi \mid [\!\!\uparrow \varphi] \psi$$

These operators are functional, so no duals are necessary

#### DYNAMIC DOXASTIC LOGIC

#### **Definition**

We extend our language from before with two new upgrade operators:

$$\varphi, \psi := \top \mid p \mid \neg \varphi \mid \varphi \wedge \psi \mid A\varphi \mid \mathbf{B}^{\psi} \varphi \mid [\!\!\uparrow \varphi] \psi \mid [\!\!\uparrow \varphi] \psi$$

These operators are functional, so no duals are necessary

• Given a plausibility model  $M = (W, \preceq, V)$  and a state  $s \in W$ :

$$\begin{split} M,s &\vDash [\!\!\!\uparrow \varphi] \psi \quad \text{iff} \quad \mathsf{Lex}(M,[\![\varphi]\!]),s \vDash \psi \\ M,s &\vDash [\!\!\!\uparrow \varphi] \psi \quad \text{iff} \quad \mathsf{Mini}(M,[\![\varphi]\!]),s \vDash \psi \end{split}$$

#### DYNAMIC DOXASTIC LOGIC

#### **Definition**

We extend our language from before with two new upgrade operators:

$$\varphi, \psi := \top \mid p \mid \neg \varphi \mid \varphi \wedge \psi \mid A\varphi \mid \mathbf{B}^{\psi}\varphi \mid [\!\! \uparrow \varphi \!\! ]\psi \mid [\!\! \uparrow \varphi \!\! ]\psi$$

These operators are functional, so no duals are necessary

• Given a plausibility model  $M = (W, \preceq, V)$  and a state  $s \in W$ :

$$\begin{split} M,s &\models [\!\!\!\upharpoonright \varphi] \psi \quad \text{iff} \quad \mathsf{Lex}(M,[\![\varphi]\!]), s \models \psi \\ M,s &\models [\!\!\!\upharpoonright \varphi] \psi \quad \text{iff} \quad \mathsf{Mini}(M,[\![\varphi]\!]), s \models \psi \end{split}$$

We read them as follows:

"After re-ordering the agent's plausibility order via Lex (or Mini) on input  $\|\varphi\|$ ,  $\psi$  holds."

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\!\!\!\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Lex-Negation.  $[\!\!\Uparrow \varphi] \neg \psi \leftrightarrow \neg [\!\!\Uparrow \varphi] \psi$ 

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Lex-Negation.  $[\!\!\Uparrow \varphi] \neg \psi \leftrightarrow \neg [\!\!\Uparrow \varphi] \psi$ 

**Lex-Conjunction.**  $[\!\!\!\uparrow \varphi](\psi \wedge \theta) \leftrightarrow ([\!\!\!\uparrow \varphi]\psi \wedge [\!\!\!\uparrow \varphi]\theta)$ 

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

**Lex-Negation.**  $[\!\!\!\uparrow \varphi] \neg \psi \leftrightarrow \neg [\!\!\!\uparrow \varphi] \psi$ 

**Lex-Conjunction.**  $[\!\!\!\uparrow \varphi](\psi \wedge \theta) \leftrightarrow ([\!\!\!\uparrow \varphi]\psi \wedge [\!\!\!\uparrow \varphi]\theta)$ 

**Lex-Universal.**  $[\!\!\!\uparrow \varphi] A \psi \leftrightarrow A [\!\!\!\uparrow \varphi] \psi$ 

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

**Lex-Negation.**  $[\!\!\!\uparrow \varphi] \neg \psi \leftrightarrow \neg [\!\!\!\uparrow \varphi] \psi$ 

**Lex-Conjunction.**  $[\!\!\!\!\uparrow\varphi](\psi\wedge\theta) \leftrightarrow ([\!\!\!\!\uparrow\varphi]\psi\wedge[\!\!\!\!\uparrow\varphi]\theta)$ 

Lex-Universal.  $[\!\!\!\uparrow \varphi] A \, \psi \leftrightarrow A [\!\!\!\uparrow \varphi] \, \psi$ 

$$\vee (\neg E (\varphi \wedge [\uparrow \varphi] \theta) \wedge \mathbf{B}^{[\uparrow \varphi] \theta} ([\uparrow \varphi] \psi))$$

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Lex-Negation.  $[\!\!\Uparrow \!\!\varphi] \neg \psi \! \leftrightarrow \! \neg [\!\!\Uparrow \!\!\varphi] \psi$ 

**Lex-Universal.**  $[\!\!\!\uparrow \varphi] A \psi \leftrightarrow A [\!\!\!\uparrow \varphi] \psi$ 

 $\textbf{Lex-Belief.} \ [\!\! \Uparrow \varphi] \mathbf{B}^{\theta} \psi \leftrightarrow (E \ (\varphi \wedge [\!\! \Uparrow \varphi] \theta) \wedge \mathbf{B}^{\varphi \wedge [\!\! \Uparrow \varphi] \theta} ([\!\! \Uparrow \varphi] \psi))$ 

$$\vee (\neg E (\varphi \wedge [\uparrow \varphi] \theta) \wedge \mathbf{B}^{[\uparrow \varphi] \theta} ([\uparrow \varphi] \psi))$$

• Just like with  $[!\varphi]$ ,  $[\Uparrow \varphi]$  can be reduced to Doxastic Logic.

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

**Lex-Negation.**  $[\!\!\Uparrow \varphi] \neg \psi \leftrightarrow \neg [\!\!\Uparrow \varphi] \psi$ 

**Lex-Conjunction.**  $[\!\!\!\uparrow \varphi](\psi \wedge \theta) \leftrightarrow ([\!\!\!\uparrow \varphi]\psi \wedge [\!\!\!\uparrow \varphi]\theta)$ 

**Lex-Universal.**  $[\!\!\!\uparrow \varphi] A \psi \leftrightarrow A [\!\!\!\uparrow \varphi] \psi$ 

 $\textbf{Lex-Belief.} \ [\!\! \Uparrow \varphi] \mathbf{B}^{\theta} \psi \leftrightarrow (E \ (\varphi \wedge [\!\! \Uparrow \varphi] \theta) \wedge \mathbf{B}^{\varphi \wedge [\!\! \Uparrow \varphi] \theta} ([\!\! \Uparrow \varphi] \psi))$ 

$$\vee (\neg E (\varphi \wedge [\Uparrow \varphi] \theta) \wedge \mathbf{B}^{[\Uparrow \varphi] \theta} ([\Uparrow \varphi] \psi))$$

- Just like with  $[!\varphi]$ ,  $[\!\!\!\uparrow\varphi]$  can be reduced to Doxastic Logic.
- This gives us a complete axiomatization of  $[\uparrow \varphi]!$

The following formulas are valid for Lex:

**Lex-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

**Lex-Negation.**  $[\!\!\Uparrow \varphi] \neg \psi \leftrightarrow \neg [\!\!\Uparrow \varphi] \psi$ 

**Lex-Conjunction.**  $[\!\!\!\uparrow \varphi](\psi \wedge \theta) \leftrightarrow ([\!\!\!\uparrow \varphi]\psi \wedge [\!\!\!\uparrow \varphi]\theta)$ 

**Lex-Universal.**  $[\!\!\!\uparrow \varphi] A \psi \leftrightarrow A [\!\!\!\uparrow \varphi] \psi$ 

$$\textbf{Lex-Belief.} \ [\!\!\!\Uparrow\varphi] \textbf{B}^{\theta} \psi \leftrightarrow (E \ (\varphi \wedge [\!\!\!\Uparrow\varphi] \theta) \wedge \textbf{B}^{\varphi \wedge [\!\!\!\Uparrow\varphi] \theta} ([\!\!\!\Uparrow\varphi] \psi))$$

$$\vee (\neg E (\varphi \wedge [\Uparrow \varphi] \theta) \wedge \mathbf{B}^{[\Uparrow \varphi] \theta} ([\Uparrow \varphi] \psi))$$

- Just like with  $[!\varphi]$ ,  $[\!\!\!\uparrow\varphi]$  can be reduced to Doxastic Logic.
- This gives us a complete axiomatization of  $[\!\!\! \uparrow \varphi]\!\!\! !$
- **Note:**  $E\varphi$  here just says "there is a world (at all) where  $\varphi$  holds"

This last law is a little mysterious:

what's going on? Why does this hold?

This last law is a little mysterious:

what's going on? Why does this hold?

Let's look at what this means for the propositional case:

This last law is a little mysterious:

what's going on? Why does this hold?

# Let's look at what this means for the propositional case:

$$[\!\! \uparrow \!\! p] \mathbf{B}^q r \! \leftrightarrow \! (E (p \wedge q) \wedge \mathbf{B}^{p \wedge q} r) \vee (\neg E (p \wedge q) \wedge \mathbf{B}^q r)$$

$$[\uparrow p]\mathbf{B}^q r \leftrightarrow (E(p \land q) \land \mathbf{B}^{p \land q} r) \lor (\neg E(p \land q) \land \mathbf{B}^q r)$$

$$[\uparrow p]\mathbf{B}^q r \leftrightarrow (E(p \land q) \land \mathbf{B}^{p \land q} r) \lor (\neg E(p \land q) \land \mathbf{B}^q r)$$

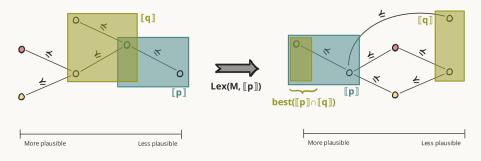
This actually encodes in our logic a complete description of the effect
 Lex has on the plausibility ordering. Here is that description:

$$\mathsf{best}_{\mathsf{Lex}(\preceq, \llbracket p \rrbracket)}(\llbracket q \rrbracket) = \left\{ \begin{array}{ll} \mathsf{best}_{\preceq}(\llbracket p \rrbracket \cap \llbracket q \rrbracket) & \mathsf{if} \; \llbracket p \rrbracket \cap \llbracket q \rrbracket \neq \emptyset \\ \mathsf{best}_{\preceq}(\llbracket q \rrbracket) & \mathsf{if} \; \llbracket p \rrbracket \cap \llbracket q \rrbracket = \emptyset \end{array} \right.$$

$$[\uparrow p]\mathbf{B}^q r \leftrightarrow (E(p \land q) \land \mathbf{B}^{p \land q} r) \lor (\neg E(p \land q) \land \mathbf{B}^q r)$$

This actually encodes in our logic a complete description of the effect
 Lex has on the plausibility ordering. Here is that description:

$$\mathsf{best}_{\mathsf{Lex}(\preceq, \llbracket p \rrbracket)}(\llbracket q \rrbracket) = \left\{ \begin{array}{ll} \mathsf{best}_{\preceq}(\llbracket p \rrbracket \cap \llbracket q \rrbracket) & \mathsf{if} \; \llbracket p \rrbracket \cap \llbracket q \rrbracket \neq \emptyset \\ \mathsf{best}_{\preceq}(\llbracket q \rrbracket) & \mathsf{if} \; \llbracket p \rrbracket \cap \llbracket q \rrbracket = \emptyset \end{array} \right.$$



The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Mini-Negation.  $[\uparrow \varphi] \neg \psi \leftrightarrow \neg [\uparrow \varphi] \psi$ 

The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Mini-Negation.  $[\uparrow \varphi] \neg \psi \leftrightarrow \neg [\uparrow \varphi] \psi$ 

 $\textbf{Mini-Conjunction.} \ \ [\uparrow\varphi](\psi\wedge\theta) \leftrightarrow ([\uparrow\varphi]\psi\wedge[\uparrow\varphi]\theta)$ 

The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Mini-Negation.  $[\uparrow \varphi] \neg \psi \leftrightarrow \neg [\uparrow \varphi] \psi$ 

 $\textbf{Mini-Conjunction.} \ \ [\uparrow\varphi](\psi\wedge\theta) \leftrightarrow ([\uparrow\varphi]\psi\wedge[\uparrow\varphi]\theta)$ 

Mini-Universal.  $[\uparrow \varphi] A \psi \leftrightarrow A[\uparrow \varphi] \psi$ 

The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Mini-Negation.  $[\uparrow \varphi] \neg \psi \leftrightarrow \neg [\uparrow \varphi] \psi$ 

 $\textbf{Mini-Conjunction.} \ \ [\uparrow\varphi](\psi\wedge\theta) \leftrightarrow ([\uparrow\varphi]\psi\wedge[\uparrow\varphi]\theta)$ 

Mini-Universal.  $[\uparrow \varphi] A \psi \leftrightarrow A [\uparrow \varphi] \psi$ 

 $\textbf{Mini-Belief.} \ [\!\!\uparrow\varphi\!\!] \mathbf{B}^{\theta} \psi \leftrightarrow \mathbf{B}^{\varphi} (\neg [\!\!\uparrow\varphi\!\!] \theta) \wedge \mathbf{B}^{[\!\!\uparrow\varphi\!\!] \theta} ([\!\!\uparrow\varphi\!\!] \psi)$ 

$$\vee \neg \mathbf{B}^{\varphi}(\neg[\uparrow\varphi]\theta) \wedge \mathbf{B}^{\varphi \wedge [\uparrow\varphi]\theta}([\uparrow\varphi]\psi)$$

The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Mini-Negation.  $[\uparrow \varphi] \neg \psi \leftrightarrow \neg [\uparrow \varphi] \psi$ 

Mini-Conjunction.  $[\uparrow \varphi](\psi \land \theta) \leftrightarrow ([\uparrow \varphi]\psi \land [\uparrow \varphi]\theta)$ 

Mini-Universal.  $[\uparrow \varphi] A \psi \leftrightarrow A [\uparrow \varphi] \psi$ 

 $\textbf{Mini-Belief.} \ [\!\uparrow\!\varphi] \mathbf{B}^\theta \psi \! \leftrightarrow \mathbf{B}^\varphi (\neg [\!\uparrow\!\varphi] \theta) \wedge \mathbf{B}^{[\!\uparrow\!\varphi]} \theta ([\!\uparrow\!\varphi] \psi)$ 

$$\vee \ \neg \mathbf{B}^{\varphi}(\neg[\uparrow\varphi]\theta) \wedge \mathbf{B}^{\varphi \wedge [\uparrow\varphi]\theta}([\uparrow\varphi]\psi)$$

•  $[\uparrow \varphi]$  can also be reduced to Doxastic Logic.

The following formulas are valid for Mini:

**Mini-Atomic.**  $[\uparrow \varphi] p \leftrightarrow p$ , for atomic propositions p

Mini-Negation.  $[\uparrow \varphi] \neg \psi \leftrightarrow \neg [\uparrow \varphi] \psi$ 

Mini-Conjunction.  $[\uparrow \varphi](\psi \land \theta) \leftrightarrow ([\uparrow \varphi]\psi \land [\uparrow \varphi]\theta)$ 

Mini-Universal.  $[\uparrow \varphi] A \psi \leftrightarrow A [\uparrow \varphi] \psi$ 

$$\textbf{Mini-Belief.} \ [\!\!\uparrow\varphi] \mathbf{B}^{\theta} \psi \leftrightarrow \mathbf{B}^{\varphi} (\neg [\!\!\uparrow\varphi] \theta) \wedge \mathbf{B}^{[\!\!\uparrow\varphi] \theta} ([\!\!\uparrow\varphi] \psi)$$

$$\vee \neg \mathbf{B}^{\varphi}(\neg[\uparrow\varphi]\theta) \wedge \mathbf{B}^{\varphi \wedge [\uparrow\varphi]\theta}([\uparrow\varphi]\psi)$$

- $[\uparrow \varphi]$  can also be reduced to Doxastic Logic.
- This gives us a complete axiomatization of  $[\uparrow \varphi]!$

Again, the last law seems mysterious.

Mini-Belief. 
$$[\uparrow \varphi] \mathbf{B}^{\theta} \psi \leftrightarrow \mathbf{B}^{\varphi} (\neg [\uparrow \varphi] \theta) \wedge \mathbf{B}^{[\uparrow \varphi] \theta} ([\uparrow \varphi] \psi)$$
  
$$\lor \neg \mathbf{B}^{\varphi} (\neg [\uparrow \varphi] \theta) \wedge \mathbf{B}^{\varphi \wedge [\uparrow \varphi] \theta} ([\uparrow \varphi] \psi)$$

Let's demystify it!

Again, the last law seems mysterious.

Mini-Belief. 
$$[\uparrow \varphi] \mathbf{B}^{\theta} \psi \leftrightarrow \mathbf{B}^{\varphi} (\neg [\uparrow \varphi] \theta) \wedge \mathbf{B}^{[\uparrow \varphi] \theta} ([\uparrow \varphi] \psi)$$

$$\lor \neg \mathbf{B}^{\varphi} (\neg [\uparrow \varphi] \theta) \wedge \mathbf{B}^{\varphi \wedge [\uparrow \varphi] \theta} ([\uparrow \varphi] \psi)$$

Let's demystify it!

Again, look at what this means for the propositional case:

Again, the last law seems mysterious.

$$\begin{split} \textbf{Mini-Belief.} \ [\uparrow\varphi] \mathbf{B}^{\theta} \psi &\leftrightarrow \mathbf{B}^{\varphi} (\neg [\uparrow\varphi] \theta) \wedge \mathbf{B}^{[\uparrow\varphi] \theta} ([\uparrow\varphi] \psi) \\ & \lor \neg \mathbf{B}^{\varphi} (\neg [\uparrow\varphi] \theta) \wedge \mathbf{B}^{\varphi \wedge [\uparrow\varphi] \theta} ([\uparrow\varphi] \psi) \end{split}$$

Let's demystify it!

# Again, look at what this means for the propositional case:

$$[\uparrow p]\mathbf{B}^q r \leftrightarrow (\mathbf{B}^p \neg q \wedge \mathbf{B}^q r) \vee (\neg \mathbf{B}^p \neg q \wedge \mathbf{B}^{p \wedge q} r)$$

$$[\uparrow p] \mathbf{B}^q r \leftrightarrow (\mathbf{B}^p \neg q \wedge \mathbf{B}^q r) \vee (\neg \mathbf{B}^p \neg q \wedge \mathbf{B}^{p \wedge q} r)$$

$$[\uparrow p]\mathbf{B}^q r \leftrightarrow (\mathbf{B}^p \neg q \wedge \mathbf{B}^q r) \vee (\neg \mathbf{B}^p \neg q \wedge \mathbf{B}^{p \wedge q} r)$$

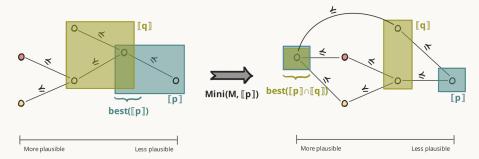
 This encodes in our logic a complete description of the effect Mini has on the plausibility ordering. This is the description:

$$\mathsf{best}_{\mathsf{Mini}(\underline{\prec}, \llbracket p \rrbracket)}(\llbracket q \rrbracket) = \left\{ \begin{array}{ll} \mathsf{best}_{\underline{\prec}}(\llbracket q \rrbracket) & \mathsf{if} \; \mathsf{best}_{\underline{\prec}}(\llbracket p \rrbracket) \subseteq \llbracket q \rrbracket^{\complement} \\ \mathsf{best}_{\underline{\prec}}(\llbracket p \rrbracket \cap \llbracket q \rrbracket) & \mathsf{if} \; \mathsf{best}_{\underline{\prec}}(\llbracket p \rrbracket) \not\subseteq \llbracket q \rrbracket^{\complement} \end{array} \right.$$

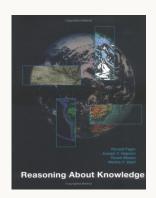
$$[\uparrow p]\mathbf{B}^q r \leftrightarrow (\mathbf{B}^p \neg q \wedge \mathbf{B}^q r) \vee (\neg \mathbf{B}^p \neg q \wedge \mathbf{B}^{p \wedge q} r)$$

 This encodes in our logic a complete description of the effect Mini has on the plausibility ordering. This is the description:

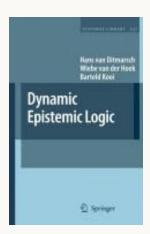
$$\mathsf{best}_{\mathsf{Mini}(\preceq, \llbracket p \rrbracket)}(\llbracket q \rrbracket) = \left\{ \begin{array}{ll} \mathsf{best}_{\preceq}(\llbracket q \rrbracket) & \mathsf{if} \; \mathsf{best}_{\preceq}(\llbracket p \rrbracket) \subseteq \llbracket q \rrbracket^{\complement} \\ \mathsf{best}_{\preceq}(\llbracket p \rrbracket \cap \llbracket q \rrbracket) & \mathsf{if} \; \mathsf{best}_{\preceq}(\llbracket p \rrbracket) \not \subseteq \llbracket q \rrbracket^{\complement} \end{array} \right.$$



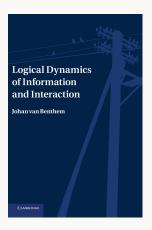
## TO LEARN MORE CHECK OUT...



(a) <u>Reasoning about</u>
<u>Knowledge</u> by Halpern,
Vardi, Fagin, & Moses



(b) <u>Dynamic Epistemic</u> <u>Logic</u> by van Ditmarsch, van der Hoek, & Kooi



(c) Logical Dynamics of Information and Interaction by van Benthem

