Implicit Knowledge in Unawareness Structures

Gaia Belardinelli
Center for Information and Bubble Studies, University of Copenhagen

Zoom mini-workshop
02 December 2022

Joint work with Burkhard C. Schipper
University of California, Davis

UNIVERSITY OF COPENHAGEN

UCDAVIS UNIVERSITY OF CALIFORNIA

Example

Ann is a violin maker. She needs to buy wood for her next creations, so she goes into a wood shop, where she is shown several pieces of wood that they need to choose from.

Together with Ann there is Carol, Ann's student, who learning the job and joined Ann at the shop for the first time. "There is only one way to choose the wood for our violins" says Ann to Carol "**Tap on it** and hear which sound it produces." Carol nods, but is puzzled as she has noticed **something else**: "I thought we mainly had to look at the **color** of the wood. In fact, I noticed that all the wood pieces you used in the past have some dark brown shades."

Ann is a bit **surprised** to hear that. She never noticed. But after thinking a bit about it, she realises that color **is actually crucial as wel**l, and replies "That's actually correct. Now that I think about it, I wouldn't even consider a piece of wood that doesn't have these dark shades."

→ Ann's choices are affected by **more information than she is currently aware of**.

→ Two different kinds of knowledge:

1. Ann knows that sound is an indicator of quality. — Explicit knowledge
2. Ann knows that color shades are an indicator of quality. — **Implicit knowledge**

Implicit knowledge

We are interested in that: knowledge **without awareness.**

- All phenomena in which agents have some **true information**, but they are currently **unaware they have it** (and unaware they use it).

- Knowledge that can only be ascribed from the **outside**, from the **modeler's** perspective.
- It can only be displayed in behavior, choices, or data.

- Does it really exist? **Implicit measures:** originally invented to measure information that subjects are unable or unwilling to report (cannot be measured with self-report measures)

- In general, every decision or interaction between agents or with the environment may be affected by more information than we are currently aware of. Some classical examples include:
  - implicit bias, e.g., decisions in recruiting processes;
  - expertise and knowledge how. Many skills are acquired without awareness, e.g., chicken sexers.

Our Goal

→ Introduce **implicit knowledge** in unawareness structures by Heifetz, Meier, and Schipper (2006), (2008)

Why in unawareness structures (HMS models)?

- "overt" levels of awareness;
- it is easy to "plug into" decision theory and game theory, and thus develop applications in economics & social sciences;
- feature explicit knowledge, but lack any notion of implicit knowledge.

Implicit Knowledge in Logic, Computer Science

Fagin & Halpern (1988), answer to **logical omniscience problem**:

"Explicit knowledge = Implicit Knowledge & Awareness"

Primitive unawareness structures by Heifetz et al. (2006, 2008)

Primitives of awareness structures by Fagin & Halpern (FH models)

→Fagin Halpern (FH) implicit knowledge notion:

- standard **S5** properties
- implicit knowledge is the same of explicit knowledge modulo awareness

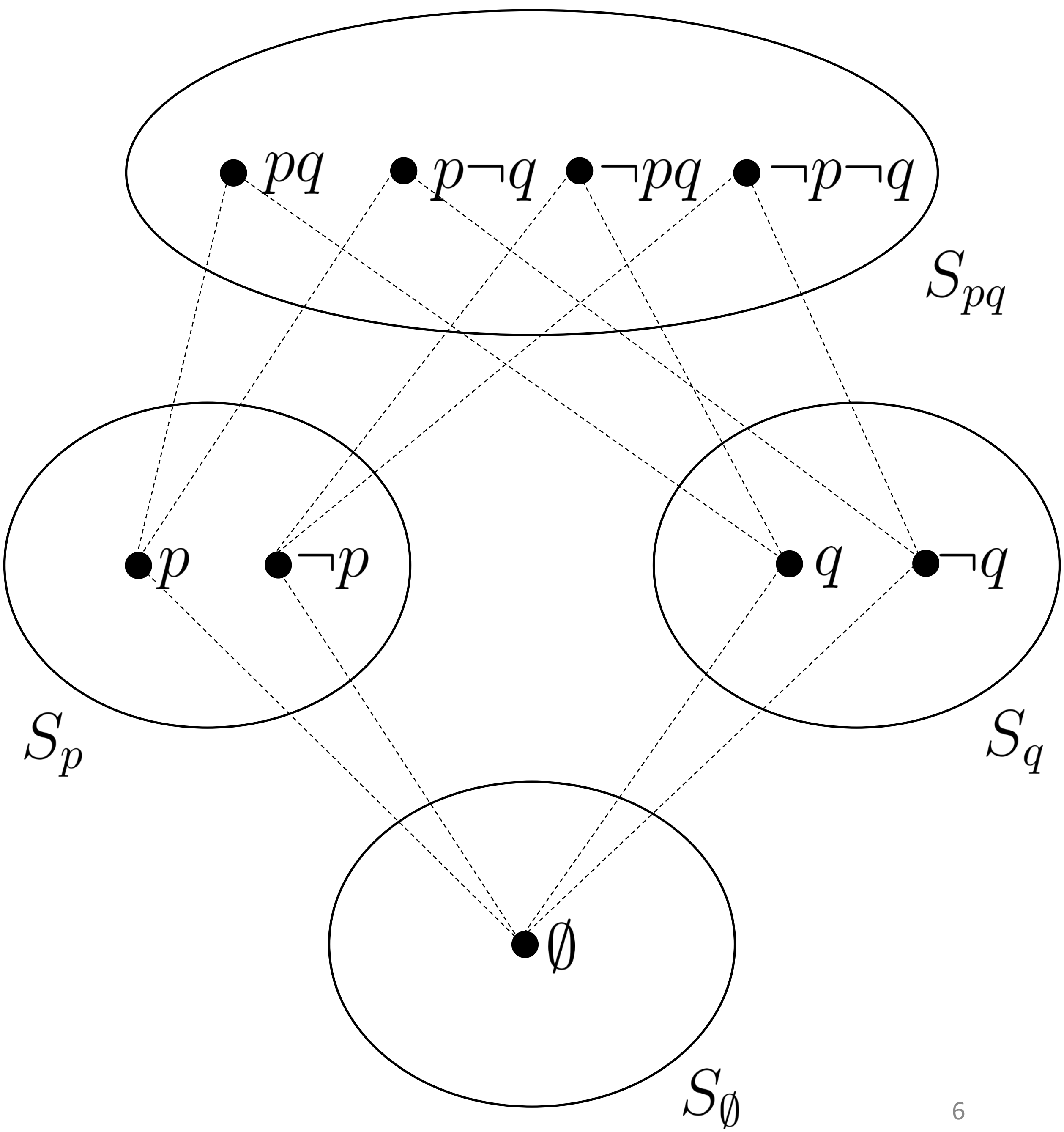
→Equivalence of HMS models with impl. K with FH models:

- it answers the theoretical question: are the two implicit knowledge notions the same?
- we use the equivalence to **obtain soundness and completeness**.
- one of the constructions used to show equivalence is informative about the **nature of HMS models** and relations with FH models.

Unawareness structure:

p means “the sound of the wood is an indicator of quality”

q means “the color shades of the wood are indicators of quality”



Unawareness structure:

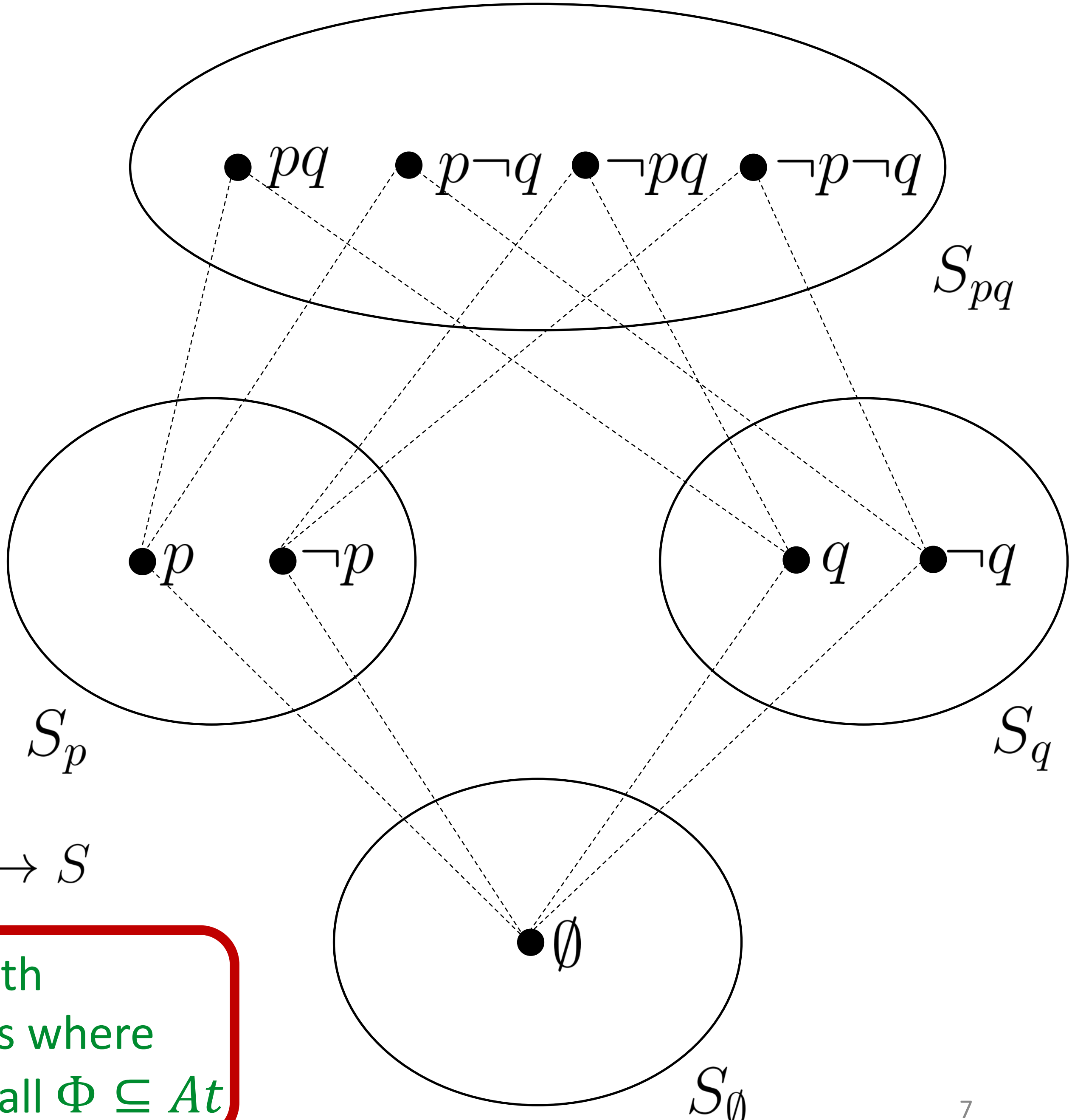
Complete lattice of state spaces: $\mathcal{S} = \{S, S', S'' \ldots\}$

Partially ordered by expressiveness: $S' \succeq S$

$\Omega := S \cup S' \cup \ldots$

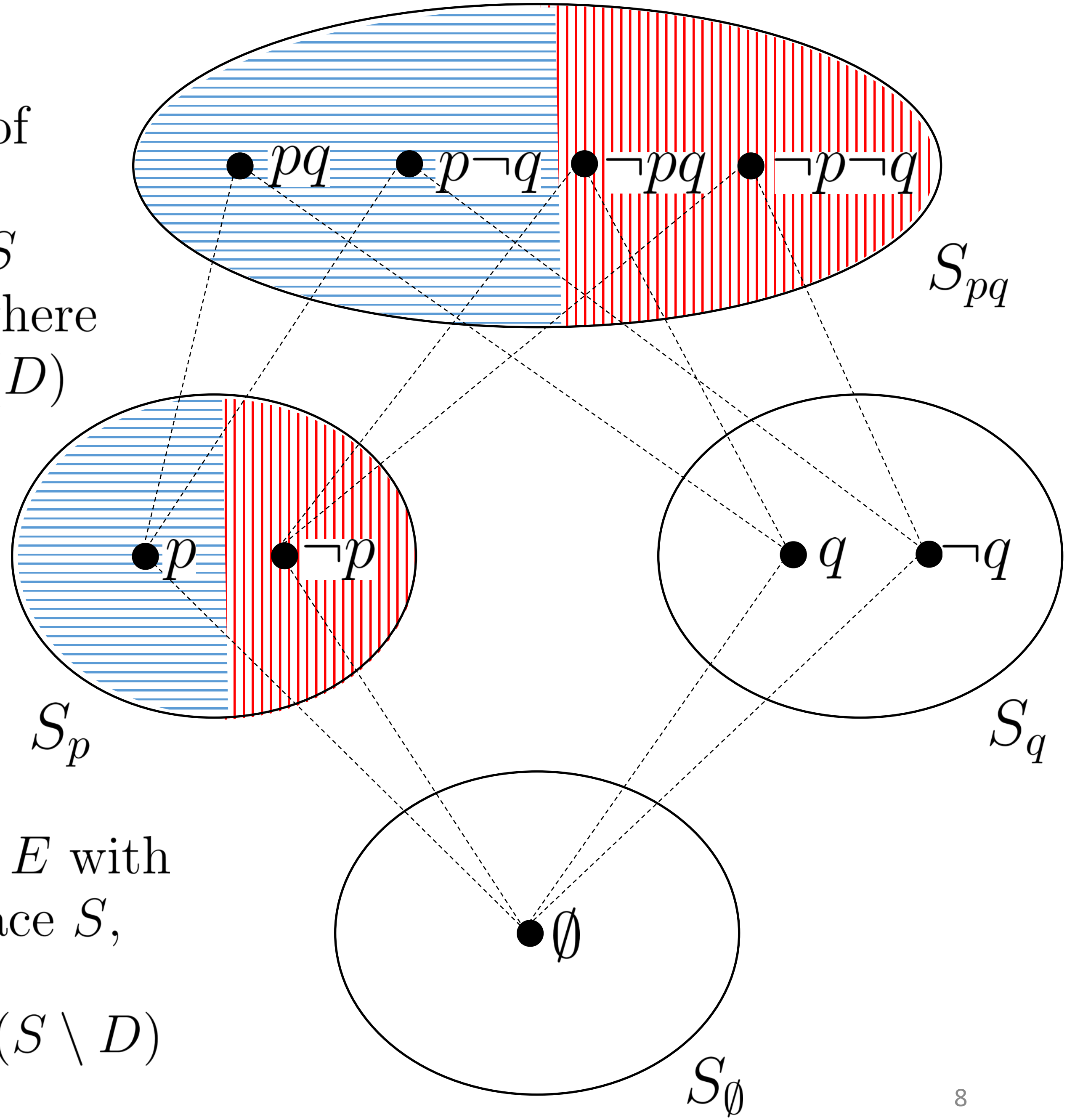
Projections:
For $S' \succeq S$, $r_S^{S'} : S' \to S$

We will mainly work with unawareness structures where there is a space S_Φ for all $\Phi \subseteq At$



Events

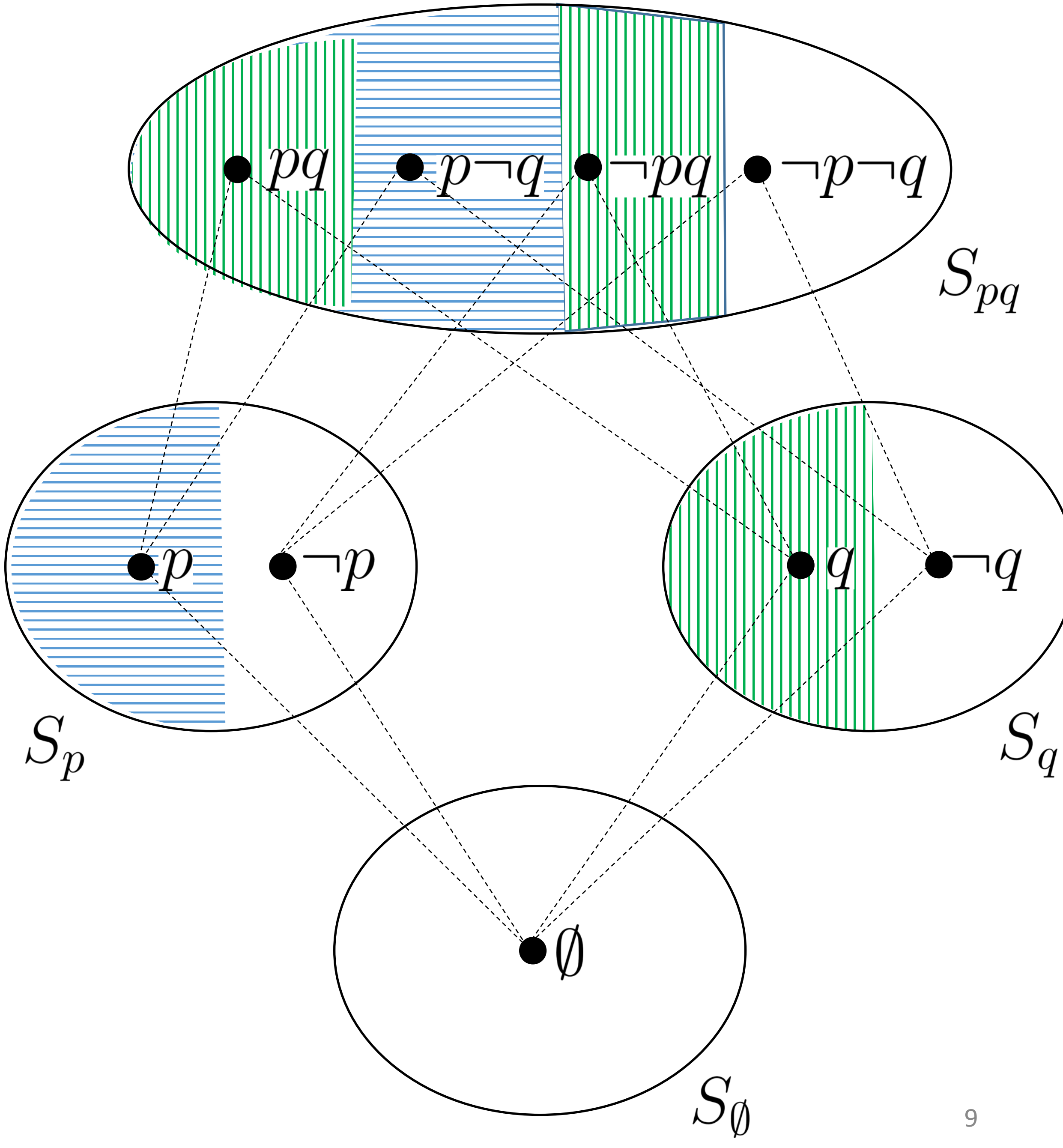
An *event* $E \subseteq \Omega$ is of the form $E = D^{\uparrow}$ for some base $D \subseteq S$ and base space S, where $D^{\uparrow} := \bigcup_{S' \succeq S} (r_S^{S'})^{-1}(D)$



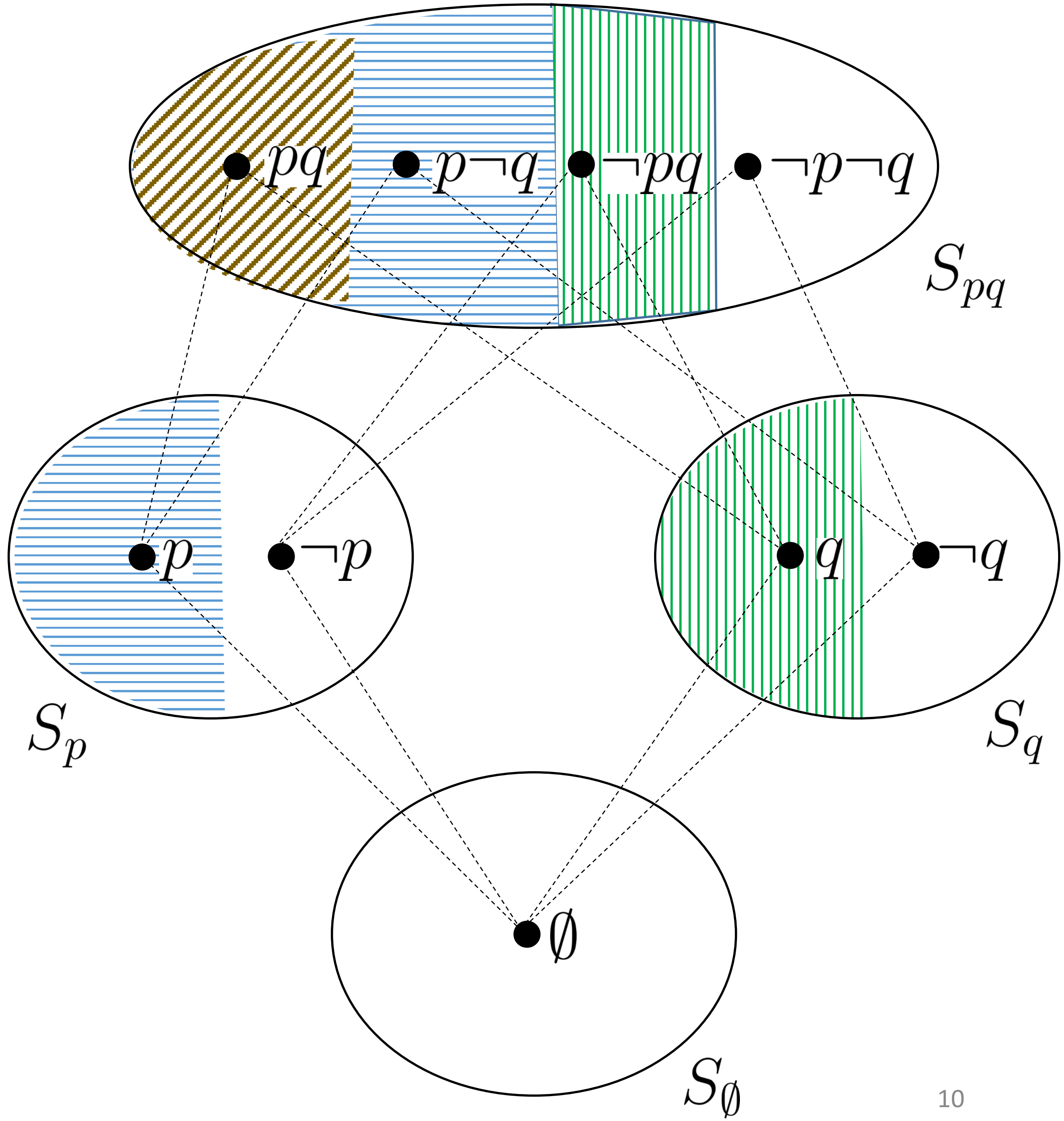
Negation: For event E with base D and base-space S,

$$\neg E := \bigcup_{S' \succeq S} (r_S^{S'})^{-1}(S \setminus D)$$

Conjunction: $E \cap F$

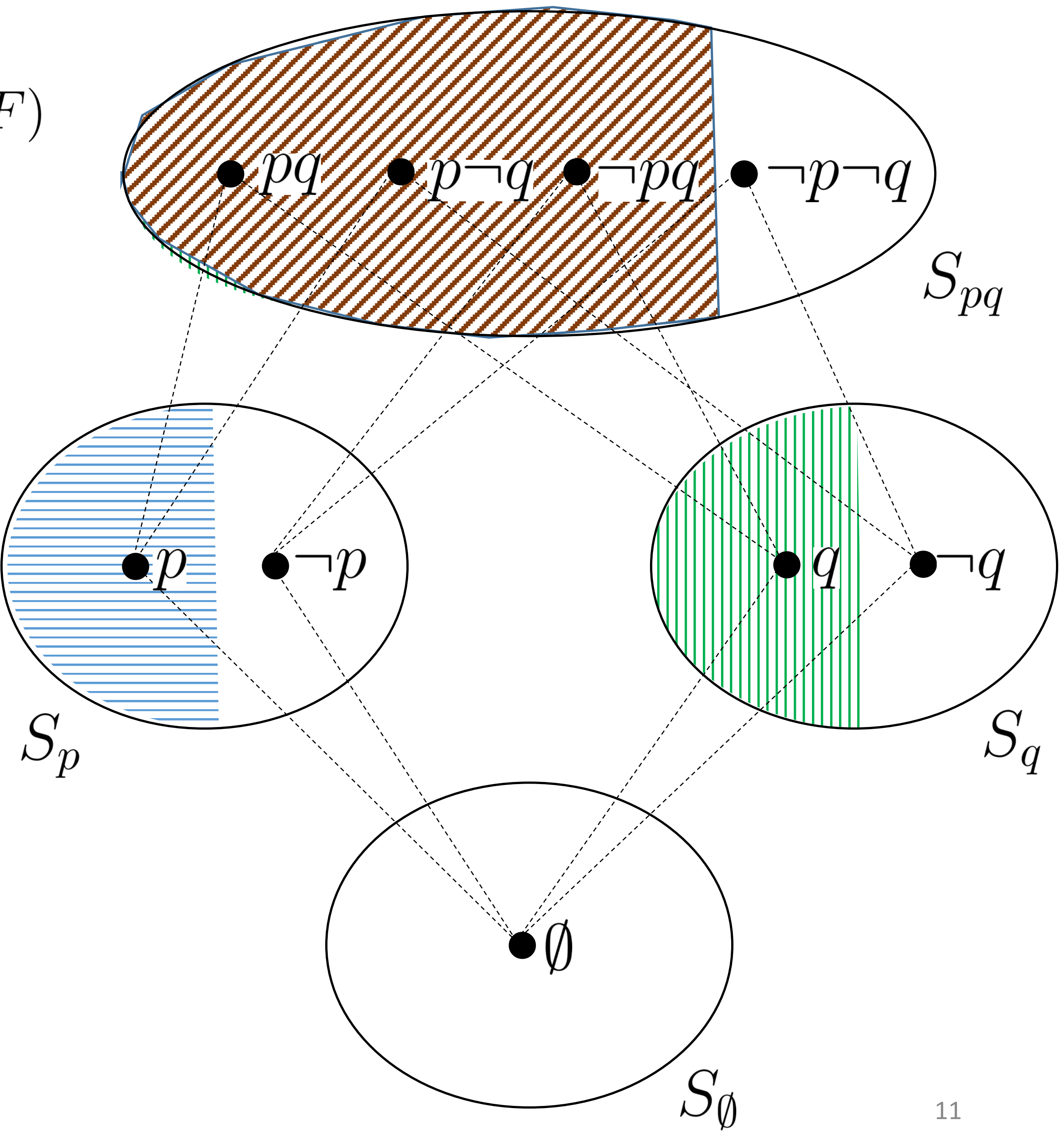


Conjunction: $E \cap F$



Disjunction:

$E \vee F := \neg(\neg E \cap \neg F)$



For each individual $i \in I$, there is a possibility correspondence $\Pi_i : \Omega \longrightarrow 2^{\Omega}$ such that

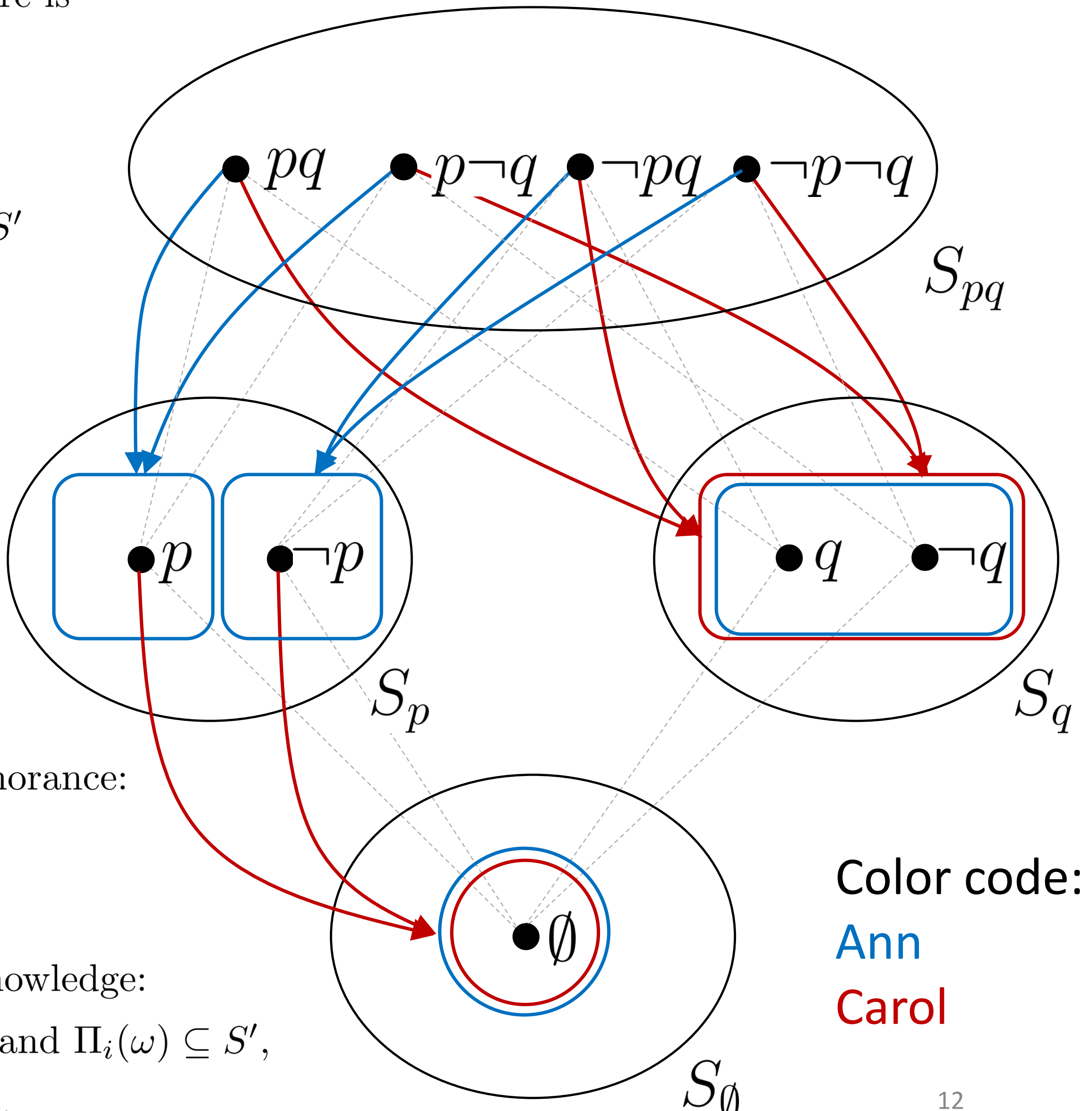
(0) Confinement:
If $\omega \in S$ then $\Pi_i(\omega) \subseteq S'$ for some $S' \preceq S$.

(i) Generalized Reflexivity:
$\omega \in \Pi_i^{\uparrow}(\omega)$ for every $\omega \in \Omega$.

(ii) Stationarity:
$\omega' \in \Pi_i(\omega)$ implies $\Pi_i(\omega') = \Pi_i(\omega)$.

(iii) Projections Preserve Ignorance:
If $\omega \in S'$ and $S \preceq S'$ then $\Pi_i^{\uparrow}(\omega) \subseteq \Pi_i^{\uparrow}(\omega_S)$.

(iv) Projections Preserve Knowledge:
If $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$, then $(\Pi_i(\omega))_S = \Pi_i(\omega_S)$.

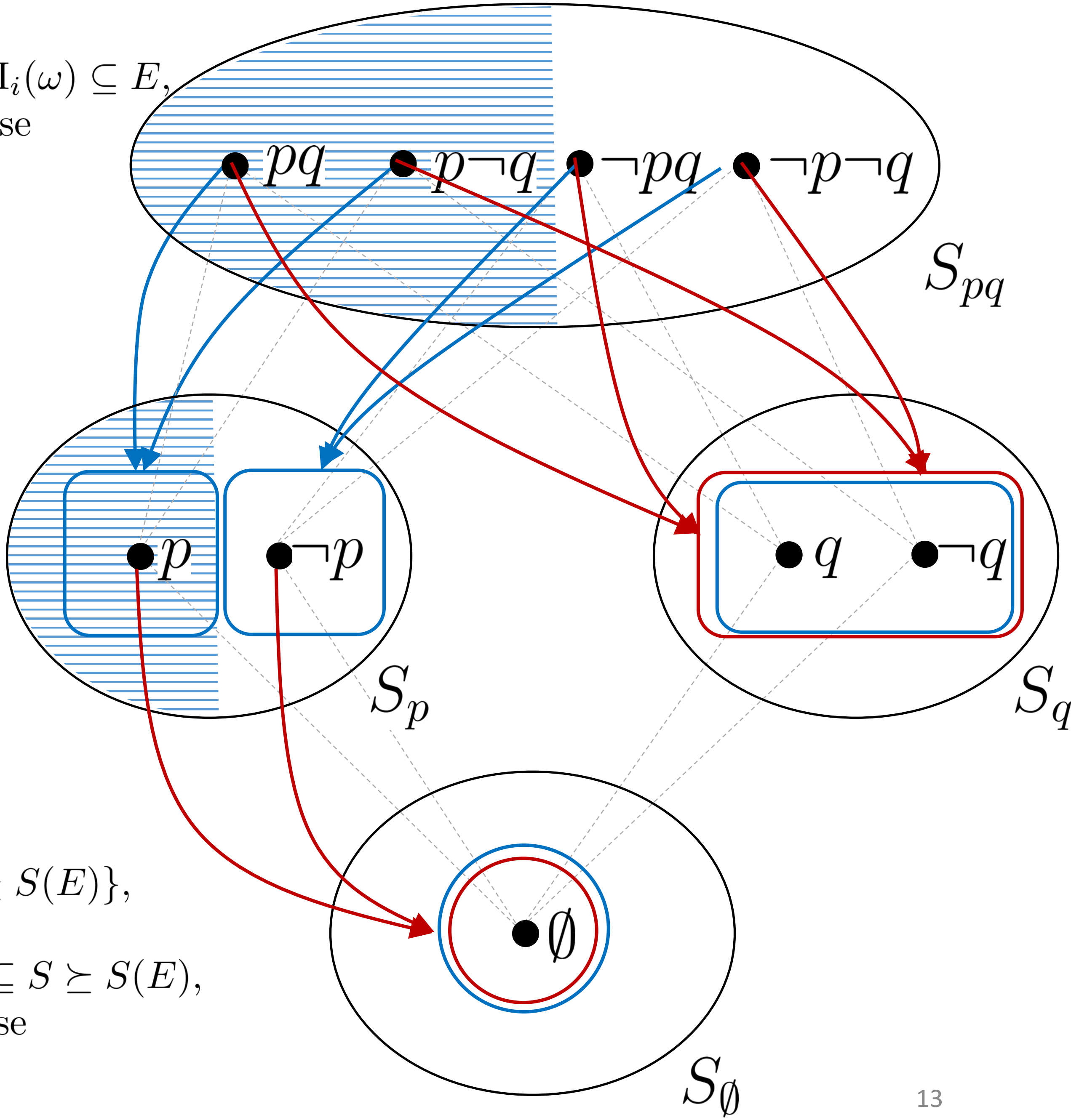


$K_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq E\}$,

if there is a state ω such that $\Pi_i(\omega) \subseteq E$, and by $K_i(E) := \emptyset^{S(E)}$ otherwise

It holds that:

1. At *pq*, Ann (explicitly) knows *p;*
2. At *pq,* Ann thinks that Carol is unaware of *p*;



$A_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq S \succeq S(E)\}$,

if there is a state ω s.t. $\Pi_i(\omega) \subseteq S \succeq S(E)$, and by $A_i(E) := \emptyset^{S(E)}$ otherwise

$U_i(E) = \neg A_i(E)$

$$K_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq E\},$$

if there is a state ω such that $\Pi_i(\omega) \subseteq E$, and by $K_i(E) := \emptyset^{S(E)}$ otherwise

Proposition K_i satisfies

(i) $K_i(\Omega) = \Omega$,

(ii) $K_i\left(\bigcap_{\lambda \in L} E_\lambda\right) = \bigcap_{\lambda \in L} K_i(E_\lambda)$,

(iii) $E \subseteq F$ implies $K_i(E) \subseteq K_i(F)$,

(iv) $K_i(E) \subseteq E$,

(v) $K_i(E) \subseteq K_i K_i(E)$,

(vi) $\neg K_i(E) \cap \neg K_i \neg K_i(E) \subseteq \neg K_i \neg K_i \neg K_i(E)$.

Explicit knowledge

$$A_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq S \succeq S(E)\},$$

if there is a state ω s.t. $\Pi_i(\omega) \subseteq S \succeq S(E)$, and by $A_i(E) := \emptyset^{S(E)}$ otherwise

$$U_i(E) = \neg A_i(E)$$

Proposition K_i and A_i satisfy

1. $K_i U_i(E) = \emptyset^{S(E)}$,
2. $U_i(E) = U_i U_i(E)$,
3. $A_i(E) = K_i\left(S(E)^{\uparrow}\right)$,
4. $U_i(E) = \bigcap_{n=1}^{\infty} (\neg K_i)^n (E)$,
5. $\neg K_i(E) \cap A_i \neg K_i(E) = K_i \neg K_i(E)$,
6. $A_i(E) = A_i(\neg E)$,
7. $\bigcap_{\lambda \in L} A_i(E_\lambda) = A_i\left(\bigcap_{\lambda \in L} E_\lambda\right)$,
8. $A_i(E) = A_i K_i(E)$,
9. $A_i(E) = A_i A_i(E)$,
10. $A_i(E) = K_i A_i(E)$.

$$K_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq E\},$$

if there is a state ω such that $\Pi_i(\omega) \subseteq E$, and by $K_i(E) := \emptyset^{S(E)}$ otherwise

Proposition K_i satisfies

(i) $K_i(\Omega) = \Omega$,

(ii) $K_i\left(\bigcap_{\lambda \in L} E_\lambda\right) = \bigcap_{\lambda \in L} K_i\left(E_\lambda\right)$,

(iii) $E \subseteq F$ implies $K_i(E) \subseteq K_i(F)$,

(iv) $K_i(E) \subseteq E$,

(v) $K_i(E) \subseteq K_i K_i(E)$,

(vi) $\neg K_i(E) \cap \neg K_i \neg K_i(E) \subseteq \neg K_i \neg K_i \neg K_i(E)$.

Explicit knowledge

$$A_i(E) := \{\omega \in \Omega : \Pi_i(\omega) \subseteq S \succeq S(E)\},$$

if there is a state ω s.t. $\Pi_i(\omega) \subseteq S \succeq S(E)$, and by $A_i(E) := \emptyset^{S(E)}$ otherwise

$$U_i(E) = \neg A_i(E)$$

Proposition K_i and A_i satisfy

1. $K_i U_i(E) = \emptyset^{S(E)}$,
2. $U_i(E) = U_i U_i(E)$,
3. $A_i(E) = K_i\left(S\left(E\right)^{\uparrow}\right)$,
4. $U_i(E) = \bigcap_{n=1}^{\infty} \left(\neg K_i\right)^n \left(E\right)$,
5. $\neg K_i(E) \cap A_i \neg K_i(E) = K_i \neg K_i(E)$,
6. $A_i(E) = A_i(\neg E)$,
7. $\bigcap_{\lambda \in L} A_i\left(E_\lambda\right) = A_i\left(\bigcap_{\lambda \in L} E_\lambda\right)$,
8. $A_i(E) = A_i K_i(E)$,
9. $A_i(E) = A_i A_i(E)$,
10. $A_i(E) = K_i A_i(E)$.

Implicit Knowledge

Fagin & Halpern (1988):

“Explicit knowledge = Implicit Knowledge & Awareness”

1) Can we derive this notion of implicit knowledge from explicit knowledge?
2) Can we take implicit knowledge and awareness as a primitive in unawareness structures and derive explicit knowledge?

Implicit Knowledge

Fagin & Halpern (1988):

“Explicit knowledge = Implicit Knowledge & Awareness”

1) Can we derive this notion of implicit knowledge from explicit knowledge?
2) Can we take implicit knowledge and awareness as a primitive in unawareness structures and derive explicit knowledge?

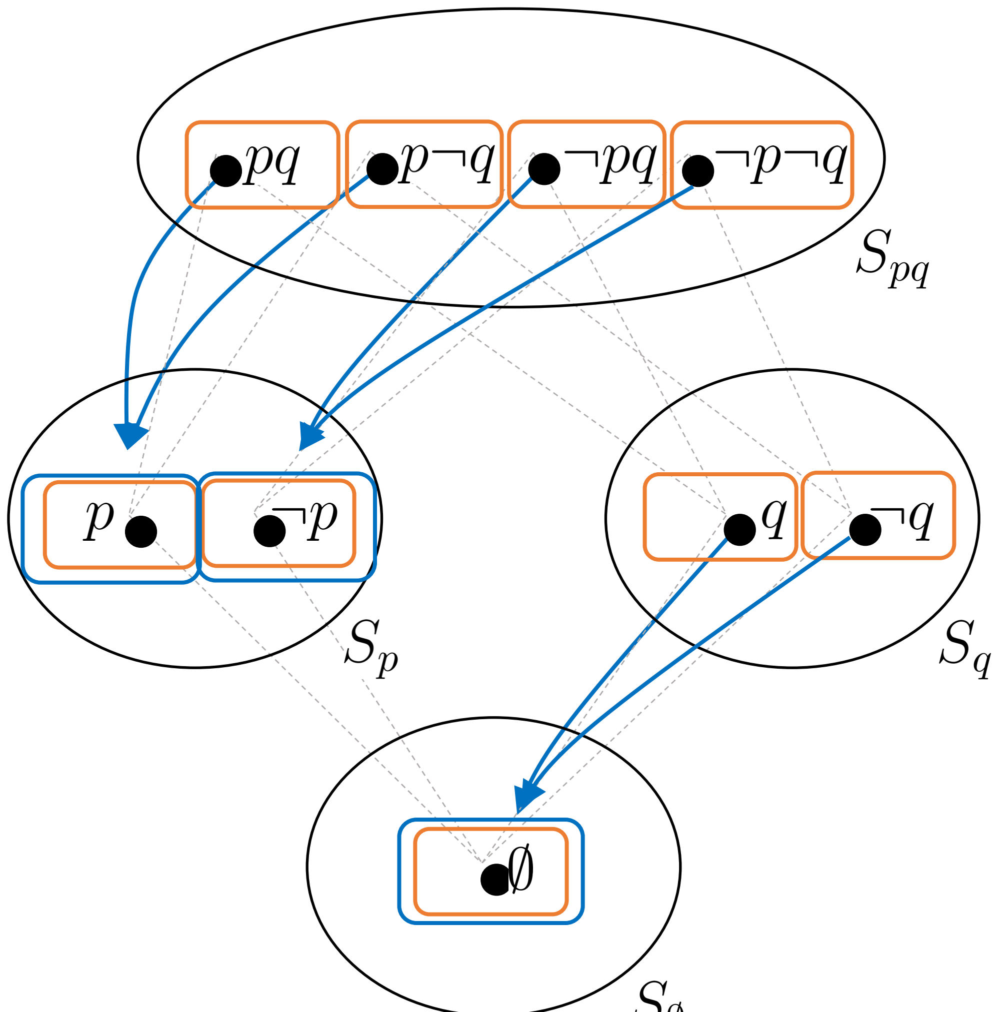
Derived Implicit Knowledge

Let's focus on Ann only.

1. Ann explicitly knows ***p*** and implicitly knows ***q***.
2. Ann's implicit knowledge is the same of explicit knowledge *modulo awareness*.

→ Introduce **another possibility correspondence.**
→ Tie it to the explicit possibility correspondence: they are the same at the agents' awareness level.

Color code:
Explicit knowledge
Implicit knowledge



→ Which **properties** give us this notion of implicit knowledge?

Derived Implicit Knowledge

Given an explicit possibility correspondence Π_i for individual $i \in I$, the *implicit possibility correspondence* $\Lambda_i : \Omega \longrightarrow 2^{\Omega}$ of individual i satisfies

(i) *Strong Confinement:* For any $\Phi \subseteq At$ and $\omega \in S_{\Phi}$, $\Lambda_i(\omega) \subseteq S_{\Phi}$.

(ii) *Reflexivity:* For any $\omega \in \Omega$, $\omega \in \Lambda_i(\omega)$.

(iii) *Stationarity*: $\omega' \in \Lambda_i(\omega)$ implies $\Lambda_i(\omega') = \Lambda_i(\omega)$.

(iv) *Projections Preserve Knowledge:* For any $\Phi \subseteq At$, if $\omega \in S_{\Phi}$, then $\Lambda_i(\omega)_{\Psi} = \Lambda_i(\omega_{\Psi})$ for all $\Psi \subseteq \Phi$.

(v) *Explicit Measurability:* $\omega' \in \Lambda_i(\omega)$ implies $\Pi_i(\omega') = \Pi_i(\omega)$.

(vi) *Implicit Measurability:* $\omega' \in \Pi_i(\omega)$ implies $\Lambda_i(\omega') = \Lambda_i(\omega)_{S_{\Pi_i(\omega)}}$.

Derived Implicit Knowledge

Implicit Knowledge Operator:

$L_i(E) := \{\omega \in \Omega : \Lambda_i(\omega) \subseteq E\}$

if there is a state $\omega \in S$ such that $\Lambda_i(\omega) \subseteq E$,
and by $L_i(E) := \emptyset^S$ otherwise

Proposition L_i satisfies

(i) *For* $\Phi \subseteq At$, $L_i(S^{\uparrow}_{\Phi}) = S_{\Phi}$

(ii) $L_i(\bigcap_{E \in \mathcal{E}} E) = \bigcap_{E \in \mathcal{E}} L_i(E)$.

(iii) $E \subseteq F$ implies $L_i(E) \subseteq L_i(F)$.

(iv) $L_i(E) \subseteq E$.

(v) $L_i(E) \subseteq L_i L_i(E)$.

(vi) $\neg L_i(E) \subseteq L_i \neg L_i(E)$.

Proposition K_i, A_i, and L_i satisfy

1. $K_i(E) = L_i(E) \cap A_i(E)$,
2. $U_i(E) = L_i(U_i(E))$,
3. $A_i(E) = L_i(A_i(E))$,
4. $A_i L_i(E) = A_i(E)$.

NEXT STEPS

- Transform HMS models with implicit knowledge into FH models, and viceversa HMS into FH models.

- Show that the two satisfy the same formulas from a language with explicit, implicit knowledge and awareness.

- Derive soundness and completeness of HMS models with implicit knowledge wrt a logic proposed by FH '88.

As we talk about formulas…

We need to move to a **syntax-based** framework:

$$\mathsf{M} = \langle I, \{S_\alpha\}_{\alpha \subseteq At}, (r^\alpha_\beta)_{\beta \subseteq \alpha}, (\Pi_i)_{i \in I}, (\Lambda_i)_{i \in I}, v \rangle$$

The valuation function maps propositions to events:

$$v : At \longrightarrow \Sigma, \text{ where } \Sigma \text{ is the set of events.}$$

(Recall: events are sets of states, upward closed, where certain propositions are true)

As the valuation function takes the set At as input, we say that M is an HMS model with implicit information, *defined for* At.

Language for Explicit Knowledge, Implicit Knowledge, and Awareness

With $i \in I$ and $p \in At$, define the language $\mathcal{L}$ by

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid k_i\varphi \mid a_i\varphi \mid \ell_i\varphi$$

Let $\mathcal{L}_\alpha = \{\varphi \in \mathcal{L} : At(\varphi) \subseteq \alpha\}$ be the sublanguage of $\mathcal{L}$ built on propositional variables in $\alpha \subseteq At$.

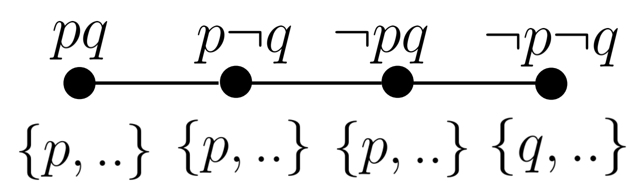
We will use this sublanguage definition in the construction of HMS lattice, to define the awareness in the subspaces.

Fagin Halpern (1998) defines

$$k_i\varphi = a_i\varphi \wedge \ell_i\varphi, \text{ for any } \varphi \in \mathcal{L}$$

Fagin Halpern '88 Awareness Model

An *FH model* is a tuple $\mathsf{K} = (I, W, R, V, \mathcal{A})$ consisting of



- a non-empty set of individuals I,
- a non-empty set of states W,
- an accessibility relation $R_i \subseteq W^2$ for all $i \in I$,
- a valuation $V : At \longrightarrow 2^W$,
- an awareness function $\mathcal{A}_i : W \longrightarrow 2^{\mathcal{L}}$, for all $i \in I$.

Let $At(\varphi) = \{p \in At\colon p \text{ is a subformula of } \varphi\}$, for all $\varphi \in \mathcal{L}$. The function $\mathcal{A}$ satisfies

PP (*Awareness is Generated by Primitive Propositions*) if for all $i \in I$ and $\varphi \in \mathcal{L}$, $\varphi \in \mathcal{A}_i(w)$ iff for all $p \in At(\varphi)$, $p \in \mathcal{A}_i(w)$.

KA (*Agents Know What They are Aware of*) if for all $i \in I$, $(w, v) \in R_i$ implies $\mathcal{A}_i(w) = \mathcal{A}_i(v)$.

→ We say that an FH model is *defined for* At, if the valuation function takes At as input.

FH Models as Semantics

Let $\mathsf{K} = (I, W, R, V, \mathcal{A})$ be an FH model for At and let $w \in W$. Satisfaction of $\mathcal{L}$ formulas in K is given by

$$
\begin{array}{lll}
\mathsf{K}, w \Vdash \top & \text{for all } w \in W; & \\
\mathsf{K}, w \Vdash p & \text{iff} & w \in V(p); \\
\mathsf{K}, w \Vdash \neg\varphi & \text{iff} & \mathsf{K}, w \nVdash \varphi;
\end{array}
\qquad
\begin{array}{lll}
\mathsf{K}, w \Vdash \varphi \wedge \psi & \text{iff} & \mathsf{K}, w \Vdash \varphi \text{ and } \mathsf{K}, w \Vdash \psi; \\
\mathsf{K}, w \Vdash \ell_i\varphi & \text{iff} & \mathsf{K}, t \Vdash \varphi \text{ for all } (w,t) \in R_i; \\
\mathsf{K}, w \Vdash a_i\varphi & \text{iff} & \varphi \in \mathcal{A}_i(w). \\
\mathsf{K}, w \Vdash k_i\varphi & \text{iff} & \varphi \in \mathcal{A}_i(w) \text{ and} \\
 & & \mathsf{K}, t \Vdash \varphi \text{ for all } (w,t) \in R_i.
\end{array}
$$

Unawareness Models as Semantics

Let $\mathsf{M} = \langle I, \mathcal{S}, \mathcal{R}, \Pi, \Lambda, v \rangle$ be an HMS model for At with implicit information, let $\omega \in \Omega$. Satisfaction of $\mathcal{L}$ formulas in M is given by

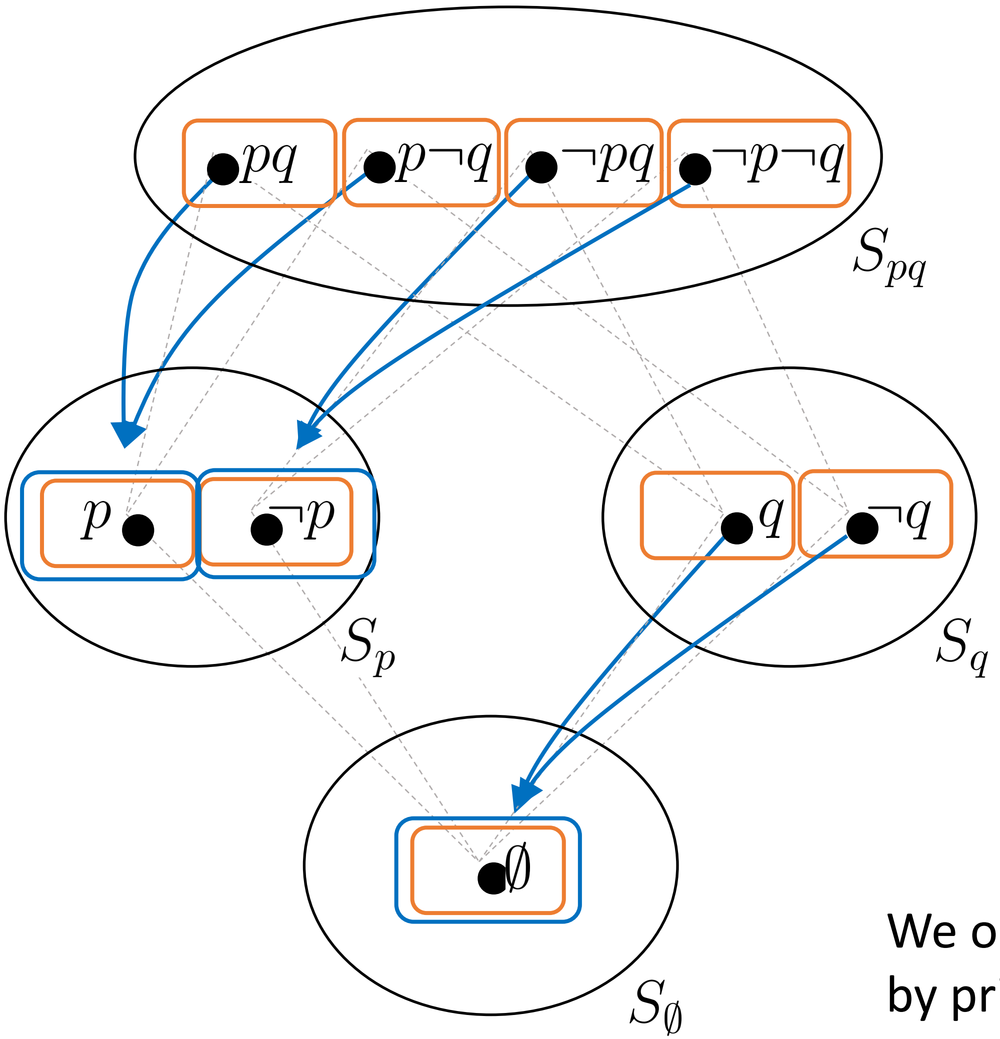
$$
\begin{array}{lll}
\mathsf{M}, \omega \vDash \top & \text{for all } \omega \in \Omega & \\
\mathsf{M}, \omega \vDash p & \text{iff} & \omega \in v(p) \\
\mathsf{M}, \omega \vDash \neg\varphi & \text{iff} & \omega \in \neg[\![\varphi]\!] \\
\mathsf{M}, \omega \vDash \varphi \wedge \psi & \text{iff} & \omega \in [\![\varphi]\!] \cap [\![\psi]\!]
\end{array}
\qquad
\begin{array}{lll}
 & & \\
\mathsf{M}, \omega \vDash \ell_i\varphi & \text{iff} & \Lambda_i(\omega) \subseteq [\![\varphi]\!] \\
\mathsf{M}, \omega \vDash a_i\varphi & \text{iff} & S_{\Pi_i(\omega)} \succeq S_{[\![\varphi]\!]} \\
\mathsf{M}, \omega \vDash k_i\varphi & \text{iff} & \Pi_i(\omega) \subseteq [\![\varphi]\!]
\end{array}
$$

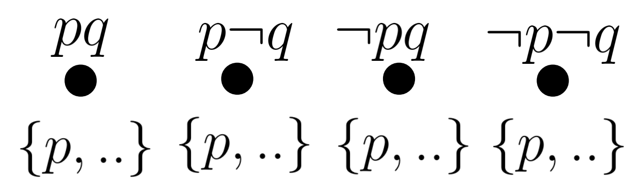
where $[\![\varphi]\!] = \{\omega' \in \Omega \colon \mathsf{M}, \omega' \vDash \varphi\}$ for all $\varphi \in \mathcal{L}$.

FH-Transform: From Unawareness to FH Models

FH-Transform: From Unawareness to FH Models

Easy! The supremum of the lattice is already a Kripke model. We only need to derive awareness:





To extract awareness info:

1. For each state, consider the space where the explicit possibility set lies;
2. Build all the formulas that can be built from the atoms defined at that space.

We obtain a notion of awareness generated by primitive propositions.

FH-Transform: From Unawareness to FH Models

Formal definition:

Let $\mathsf{M} = \langle I, \mathcal{S}, \mathcal{R}, \Pi, \Lambda, v\rangle$ be an HMS model for At with implicit information where the supremum is S_{At}. The *FH-transform model of* M is $FH(\mathsf{M}) = (I, W, R, V, \mathcal{A})$ where

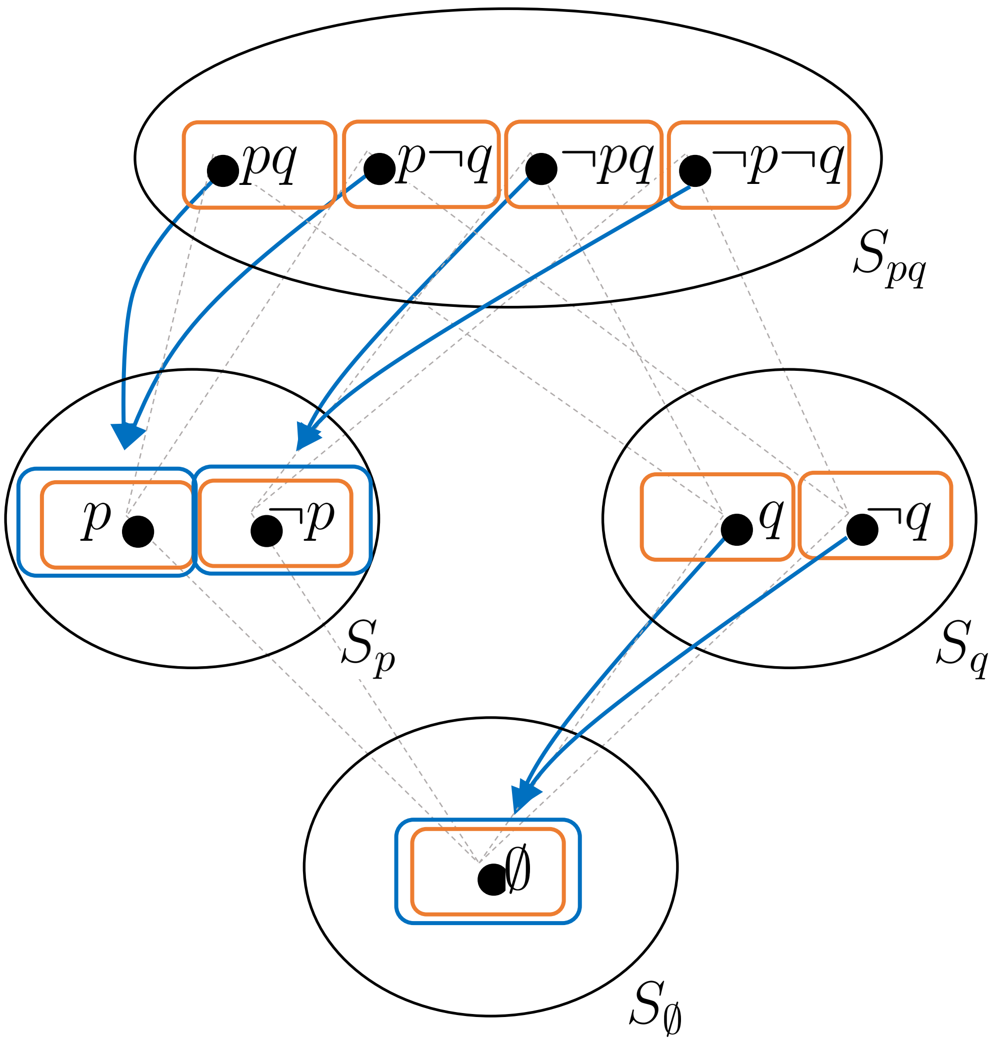
- $W = S_{At}$,
- $R_i \subseteq W^2$ is such that $(\omega, \omega') \in R_i$ iff $\omega' \in \Lambda_i(\omega)$,
- $V : At \longrightarrow \mathcal{P}(W)$ is such that $V(p) = \{\omega \in W : \omega \in v(p)\}$, for every $p \in At$,
- $\mathcal{A}_i : W \longrightarrow 2^{\mathcal{L}}$ is such that $\mathcal{A}_i(\omega) = \{\varphi \in \mathcal{L} : At(\varphi) \subseteq \Phi$ where $\Pi_i(\omega) \subseteq S_\Phi\}$.

Proposition *For any unawareness model* M*, the* FH*-transform* $FH(\mathsf{M}) = (W, R, V, \mathcal{A})$ *is an FH model, where* R *is an equivalence relation.*

Proposition *For any unawareness model* M*, the* FH*-transform* $FH(\mathsf{M}) = (W, R, V, \mathcal{A})$ *is propositionally determined, as* $\mathcal{A}$ *is generated by primitive propositions and agents know what they are aware of.*

U-Transform: From FH to Unawareness Models

U-Transform: From FH to Unawareness Models



Notice:

Spaces in an unawareness models are nothing but the **bisimulation contraction of the supremum**, for a restricted bisimulation notion (defined for some $\mathbf{\Phi} \subseteq At$).

Ex: Take the pq and the $p{\sim}q$ states. They are such that:

(atom) they contain the same p-information;

(aware) agents are aware of the same formulas from $\mathcal{L}_{\{p\}}$;

(zig and zag) they only "see" bisimilar states.

U-Transform: From FH to Unawareness Models

Strategy:

1. Take an FH model K defined for At.
2. Consider a notion of restricted bisimulation ($\mathbf{\Phi}$ -bisimulation) and define the $\mathbf{\Phi}$-bisimulation contraction $\boldsymbol{K}_{\mathbf{\Phi}}$ of the FH model K for all $\mathbf{\Phi} \subseteq At$.
3. Order the contracted models $\{\boldsymbol{K}_{\mathbf{\Phi}}\}_{\mathbf{\Phi} \subseteq At}$ by subset-inclusion of the atomic sets $\mathbf{\Phi}$. This gives a complete lattice of FH models.
4. Extract knowledge and unawareness out of it and define the unawareness model (*U*-transform).

U-Transform: From FH to Unawareness Models

Strategy:

1. Take an FH model K defined for At.
2. Consider a notion of restricted bisimulation ($\mathbf{\Phi}$ -bisimulation) and define the $\mathbf{\Phi}$-bisimulation contraction $\boldsymbol{K}_{\mathbf{\Phi}}$ of the FH model K for all $\mathbf{\Phi} \subseteq At$.
3. Order the contracted models $\{\boldsymbol{K}_{\mathbf{\Phi}}\}_{\mathbf{\Phi} \subseteq At}$ by subset-inclusion of the atomic sets $\mathbf{\Phi}$. This gives a complete lattice of FH models.
4. Extract knowledge and unawareness out of it and define the unawareness model (*U*-transform).

U-Transform: α-Bisimulation Example

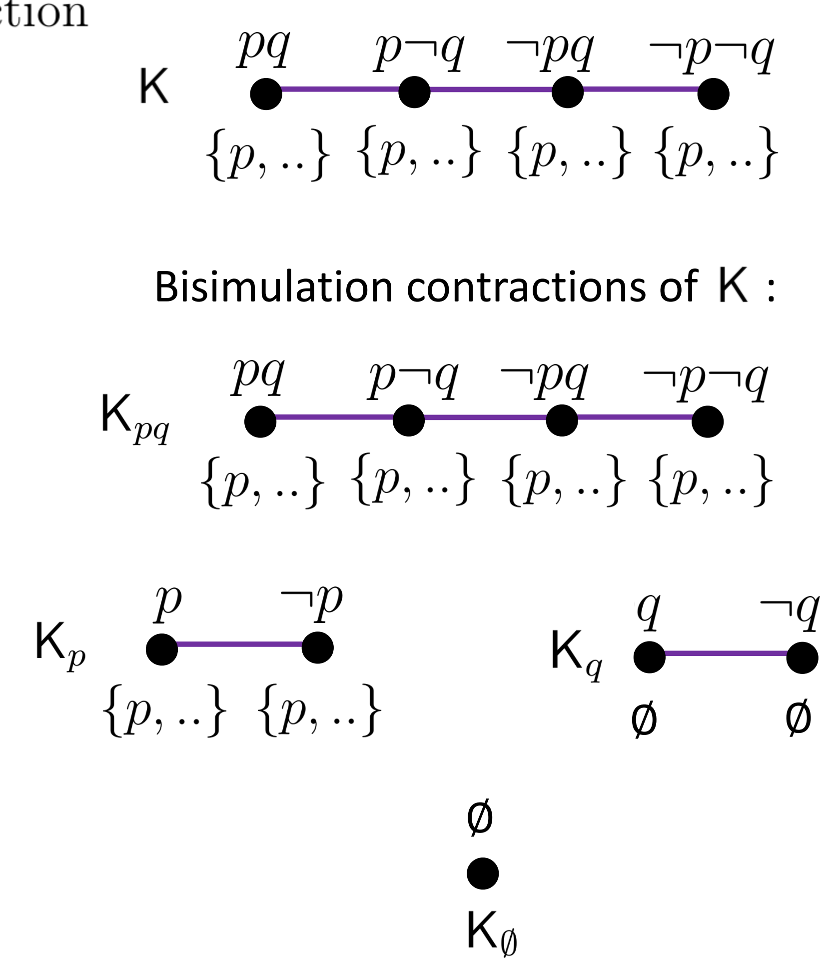
A Φ-*bisimulation* between two FH models $\mathsf{K} = (I, W, R, V, \mathcal{A})$ and $\mathsf{K}' = (I', W', R', V', \mathcal{A}')$ for At is a relation $\mathcal{Z}[\Phi] \subseteq W \times W'$ such that, for every $(w, w') \in \mathcal{Z}[\Phi]$, every agent $i \in I$, and every $p \in \Phi$:

- *atom:* $w \in V(p)$ iff $w' \in V'(p)$.
- *aware:* $\mathcal{L}_\Phi \cap \mathcal{A}_i(w) = \mathcal{L}_\Phi \cap \mathcal{A}'_i(w')$.
- *forth:* if $(w, t) \in R_i$ then there is a $t' \in W'$ such that $(w', t') \in R'_i$ and such that $(t, t') \in \mathcal{Z}[\Phi]$.
- *back:* if $(w', t') \in R'_i$ then there is a $t \in W$ such that $(w, t) \in R_i$ and such that $(t, t') \in \mathcal{Z}[\Phi]$.

U-Transform: FH-lattice

For all $\Phi \subseteq At$ take the Φ-bisimulation contraction $\mathsf{K}_\Phi = (I, W_\Phi, R_\Phi, V_\Phi, \mathcal{A}_\Phi)$ of an initial FH model $\mathsf{K} = (I, W, R, V, \mathcal{A})$, defined by

- $W_\Phi = \{[w]_\Phi \colon w \in W\}$ with $[w]_\Phi = \{\Phi\} \cup \{t \in M \colon (\mathsf{K}, w) \leftrightarrows_\Phi (\mathsf{K}, t)\}$;
- $R_{\Phi,i} = \{([w]_\Phi, [t]_\Phi) \colon \exists w' \in [w]_\Phi, \exists t' \in [t]_\Phi \text{ with } (w', t') \in R_i\}$;
- $V_\Phi : \Phi \longrightarrow 2^{W_\Phi}$ with $V_\Phi(p) = \{[w]_\Phi \in W_\Phi \colon w \in V(p)\}$ for all $p \in \Phi$;
- $\mathcal{A}_{\Phi,i}([w]_\Phi) = \mathcal{A}_i(w) \cap \mathcal{L}_\Phi$.



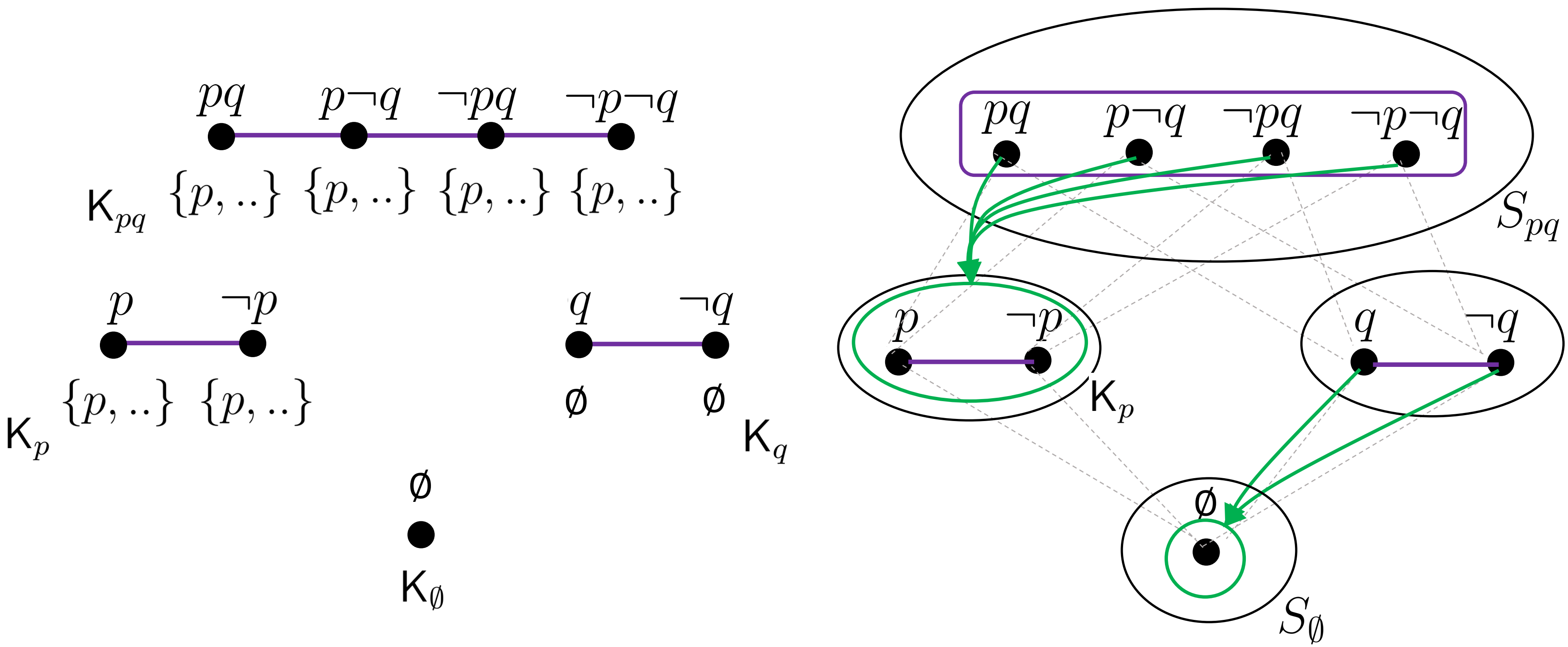
→ To construct the FH lattice then order the contracted models by subset inclusion.

U-Transform: From FH to Unawareness Models

Strategy:

1. Consider an FH model $\boldsymbol{S}$ defined for At. Define its bisimulation contraction for a notion of α-awareness bisimulation by van Ditmarch et al. (2018) (next slide).
2. Do it for all $\alpha \subseteq At$, so to obtain a set of α-bisimilar models $\boldsymbol{S}_{\alpha}$, for all $\alpha \subseteq At$.
3. Order the contracted models $\{\boldsymbol{S}_{\alpha}\}_{\alpha \subseteq At}$ by inclusion of the atomic sets α. This gives a complete lattice of FH models. Call it an FH-lattice.
4. Construct an unawareness model out of the FH-lattice (define the *U*-transform).

U-Transform: From FH to Unawareness Models



- Copy the frame (lattice structure and implicit information Λ_i);
- Each state $[w]_\Phi$ in the FH-lattice contains info about the awareness of agent i. We need that info to construct Π_i:

→ Map Π_i to the space defined for the set of atomic formulas Ψ that the agent is aware of at the considered state $[w]_\Phi$. Then let $\Pi_i([w]_\Phi) \subseteq \boldsymbol{S_\Psi}$ and take $[w']_\Psi$ that are related to $[w]_\Psi$.

U-Transform Model

Let $\mathsf{K} = (I, W, R, V, \mathcal{A})$ be an FH model for At and consider its Φ-bisimulation contractions $\mathsf{K}_\Phi = (I, W_\Phi, R_\Phi, V_\Phi, \mathcal{A}_\Phi)$, for all $\Phi \subseteq At$. Let $\Omega = \bigcup_{\Phi \subseteq At} W_\Phi$. The *$U$-transform model of* K is $U(\mathsf{K}) = \langle I, \mathcal{S}, \mathcal{R}, \Pi, \Lambda, v \rangle$, where

- $\mathcal{S} = \{W_\Phi\}_{\Phi \subseteq At}$ is a set of state-spaces W_Φ, defined for all $\Phi \subseteq At$;
- $\mathcal{R} = (r^\Phi_\Psi)_{\Psi \subseteq \Phi}$ is such that $r^\Phi_\Psi : W_\Phi \longrightarrow W_\Psi$ where $r^\Phi_\Psi(w_\Phi) = w_\Psi$, with $\Psi \subseteq \Phi \subseteq At$;
- $\Pi_i : \Omega \longrightarrow 2^\Omega$ is such that for all $\Phi \subseteq At$, $\Pi_i(w_\Phi) \ni w'_\Psi$ iff $(w_\Psi, w'_\Psi) \in R_{\Psi,i}$ and $\Psi = \{p \in At : p \in \bigcup_{\varphi \in \mathcal{A}_{\Phi,i}(w_\Phi)} At(\varphi)\}$;
- $\Lambda_i : \Omega \longrightarrow 2^\Omega$, such that $\Lambda_i(w_\Phi) \ni w'_\Phi$ iff $(w_\Phi, w'_\Phi) \in R_{\Phi,i}$;
- $v(p) = \{w_\Phi \in \Omega : \Phi \ni p \text{ and } w_\Phi \in V_\Phi(p)\}$ for all $p \in At$.

Proposition *For any partitional, propositionally determined FH model* K *for* At, *its U-transform* $U(\mathsf{K}) = \langle I, \mathcal{S}, \mathcal{R}, \Pi, \Lambda, v \rangle$ *is an HMS model for* At *with implicit information.*

Formula-equivalence follows

Proposition *For any* M *that is an HMS model for At with implicit information, where* S_{At} *is its supremum and where* $FH(\mathsf{M})$ *is its* FH*-transform, for all* $\varphi \in \mathcal{L}$ *and all* $\omega \in S_{At}$,

$$\mathsf{M}, \omega \vDash \varphi \text{ iff } FH(\mathsf{M}), \omega \Vdash \varphi.$$

Proposition *For any partitional, propositionally determined FH model* $\mathsf{K} = (I, W, R, V, \mathcal{A})$ *with* U*-transform* $U(\mathsf{K})$*, for all* $\varphi \in \mathcal{L}$*, all* $w \in W$*, and all* $w_\Phi \in \Omega$ *with* $At(\varphi) \subseteq \Phi$,

$$\mathsf{K}, w \vDash \varphi \text{ iff } U(\mathsf{K}), w_\Phi \Vdash \varphi.$$

Axiomatization

All substitution instances of propositional logic, including the formula $\top$	
$(\ell_i\varphi \wedge (\ell_i\varphi \to \ell_i\psi)) \to \ell_i\psi$	(K, Distribution)
$k_i\varphi \leftrightarrow (\ell_i\varphi \wedge a_i\varphi)$	(Explicit Knowledge)
$a_i(\varphi \wedge \psi) \leftrightarrow (a_i\varphi \wedge a_i\psi)$	(A1, Awareness Distribution)
$a_i\neg\varphi \leftrightarrow a_i\varphi$	(A2, Symmetry)
$a_ik_j\varphi \leftrightarrow a_i\varphi$	(A3, Awareness of Explicit Knowledge)
$a_ia_j\varphi \leftrightarrow a_i\varphi$	(A4, Awareness Reflection)
$a_i\ell_j\varphi \leftrightarrow a_i\varphi$	(A5, Awareness of Implicit Knowledge)
$a_i\varphi \to \ell_ia_i\varphi$	(A11, Awareness Introspection)
$\neg a_i\varphi \to \ell_i\neg a_i\varphi$	(A12, Unawareness Introspection)
From φ and $\varphi \to \psi$, infer ψ	(Modus Ponens)
From φ infer $\ell_i\varphi$	(K-Inference)
$\ell_i\varphi \to \varphi$	(T, Truth)
$\ell_i\varphi \to \ell_i\ell_i\varphi$	(4, Positive Introspection)
$\neg\ell_i\varphi \to \ell_i\neg\ell_i\varphi$	(5, Negative Introspection)

→ The logic given by rules and axioms in the table above is **sound and complete** with respect to unawareness models with implicit information.

Summary

- Introduced implicit knowledge in unawareness structures, namely knowledge the agent is not aware of.

- Unawareness structures are nothing but a lattice of bisimilar-spaces.

- Unawareness structures with implicit knowledge are formula equivalent to FH models, thus the logic for propositional awareness axiomatizes their model class.