Awareness in Games, Awareness in Logic

Joseph Halpern   Leandro Rêgo

Cornell University

Game Theory

- Standard game theory models assume that the structure of the game is common knowledge among the players.
  - This includes the possible moves and the set of players
- **Problem:** Not always a reasonable assumption; for example:
  - war settings
    - one side may not be aware of weapons the other side has
  - financial markets
    - an investor may not be aware of new innovations
  - auctions in large networks,
    - you may not be aware of who the bidders are
  - . . .

This talk:

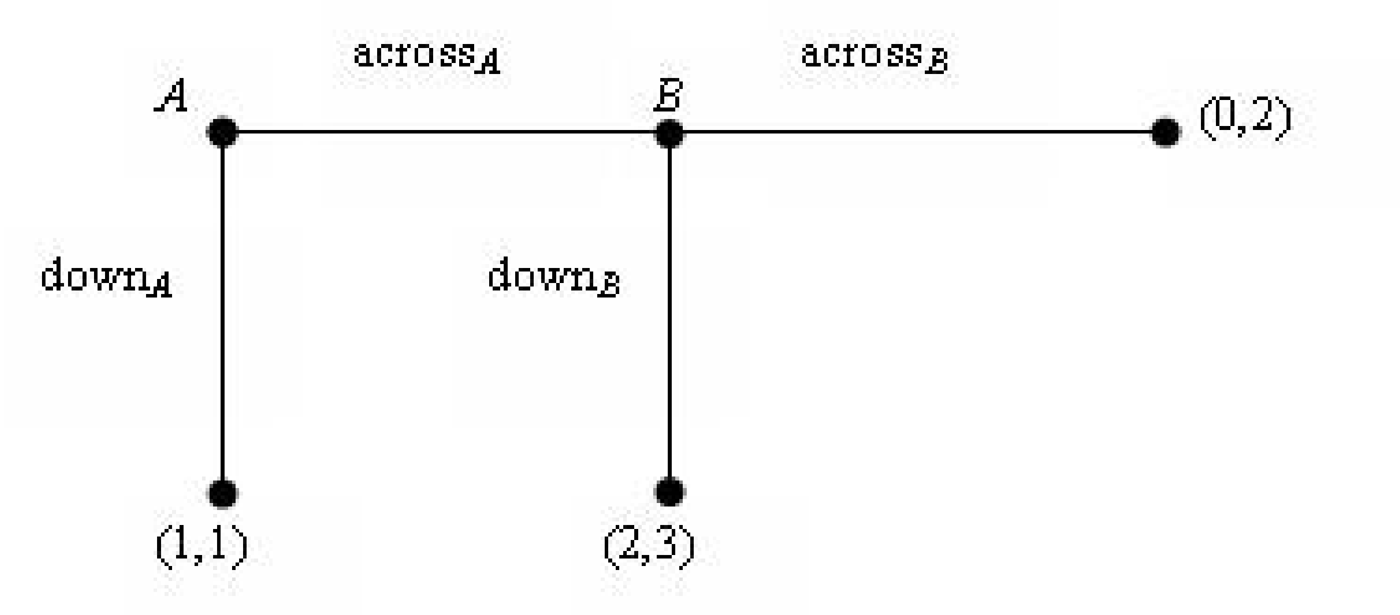
- Discuss how awareness can be added in games
  - Halpern and Rego: Extensive games with possibly unaware players
  - Lots of other work in the game theory community
- Discuss how awareness can be added to epistemic logic
  - Halpern and Rego: Reasoning About Knowledge of Unawareness Revisited
  - Work goes back to Fagin and Halpern (1985)
  - Now lots of work in the game theory community!

Nash Equilibrium

- Arguably, the major goal is to define appropriate *solution concepts*
  - how a game is/should be played
- The standard solution concept in game theory is *Nash equilibrium* (NE)
  - No player can gain by unilaterally changing his strategy

But Nash equilibrium does not always make sense if players are not aware of all moves . . .

A Simple Game



- One Nash equilibrium of this game
  - A plays across$_A$, B plays down$_B$ (not unique).
- But if A is not aware that B can play down$_B$, A will play down$_A$.

We need a solution concept that takes awareness into account!

- First step: represent games where players may be unaware

Representing lack of awareness

- Γ: an underlying standard extensive game.
  - Γ describes the moves actually available to players
- An *augmented game based on* Γ is essentially a standard game that also determines for each history h an *awareness level*,
  - the set of runs in the underlying game that the player who moves at h is aware of
  - Intuition: an augmented game describes the game from the point of view of an omniscient modeler or one of the players.

Augmented Games

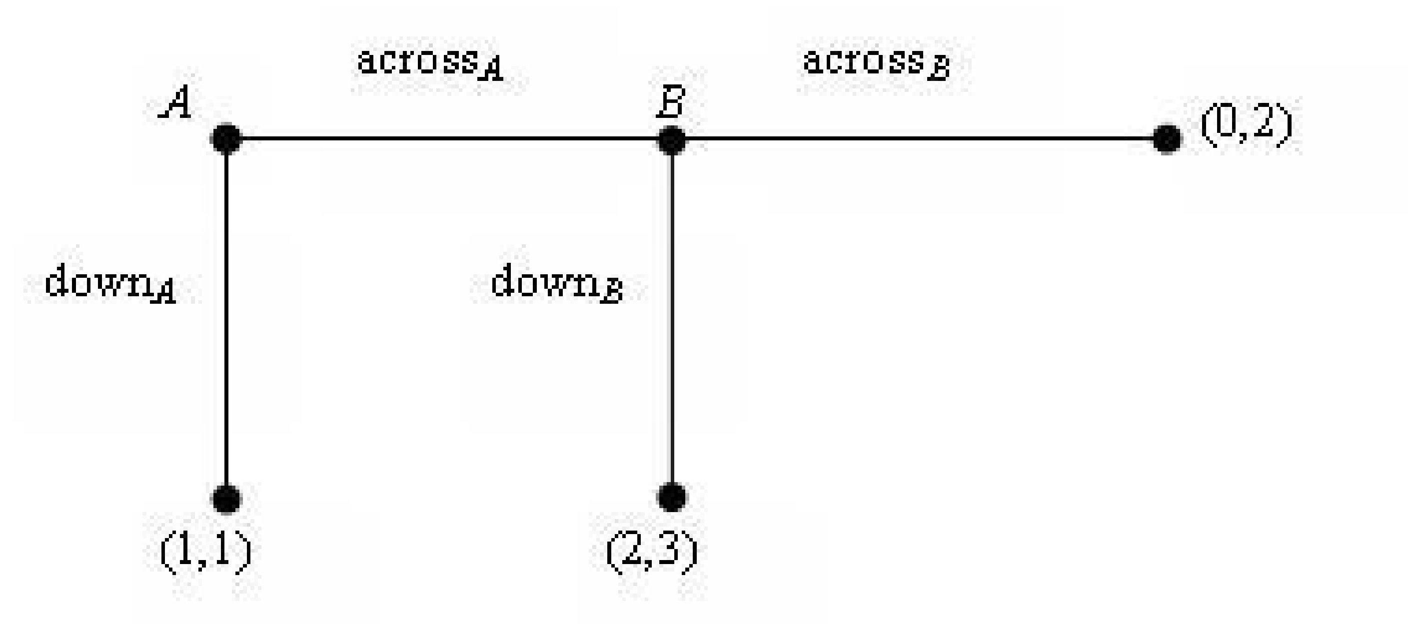
Consider the earlier game. Suppose that

- players A and B are aware of all histories of the game;
- player A is uncertain as to whether player B is aware of run $\langle \text{across}_A, \text{down}_B \rangle$ and believes that B is unaware of it with probability p; and
- the type of player B that is aware of the run $\langle \text{across}_A, \text{down}_B \rangle$ is aware that player A is aware of all histories, and he knows A is uncertain about B's awareness level and knows the probability p.

To represent this, we need three augmented games.

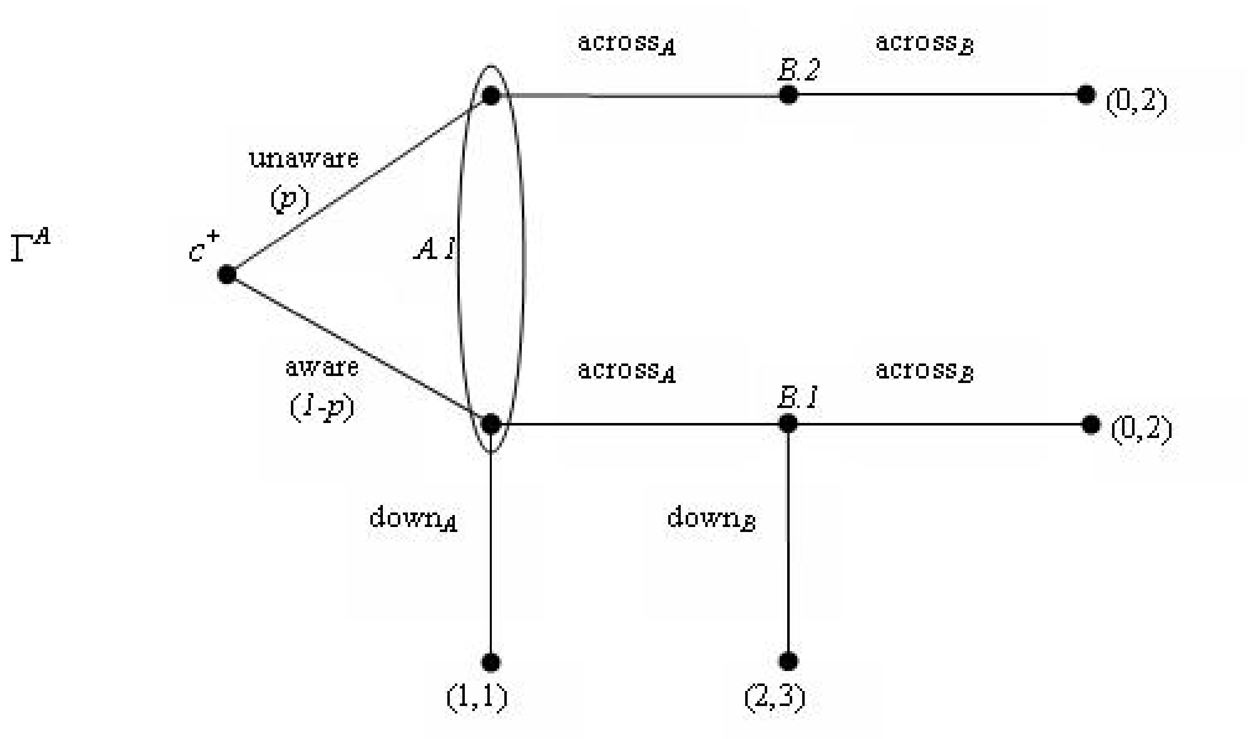
Modeler's Game

Γ^m



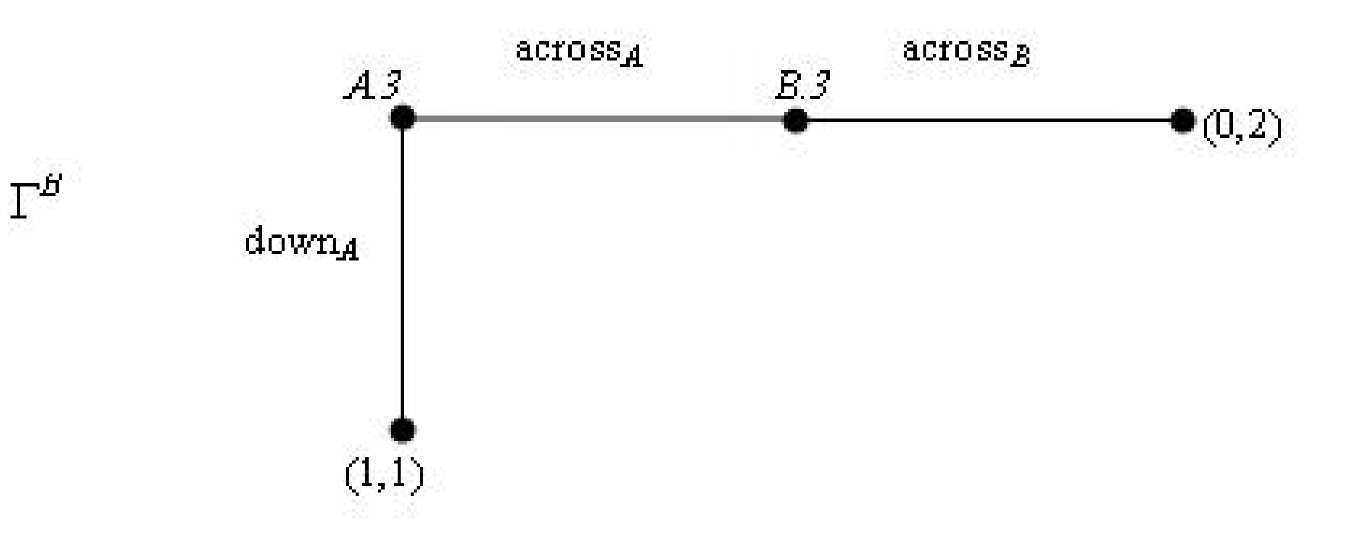
- Both A and B are aware of all histories of the underlying game.
- But A considers it possible that B is unaware.
  - To represent A's viewpoint, we need another augmented game.

A's View of the Game



- At node $B.2$, B is not aware of the run $\langle$across$_A$, down$_B\rangle$.
  - We need yet another augmented game to represent this.

(A's view of) B's view



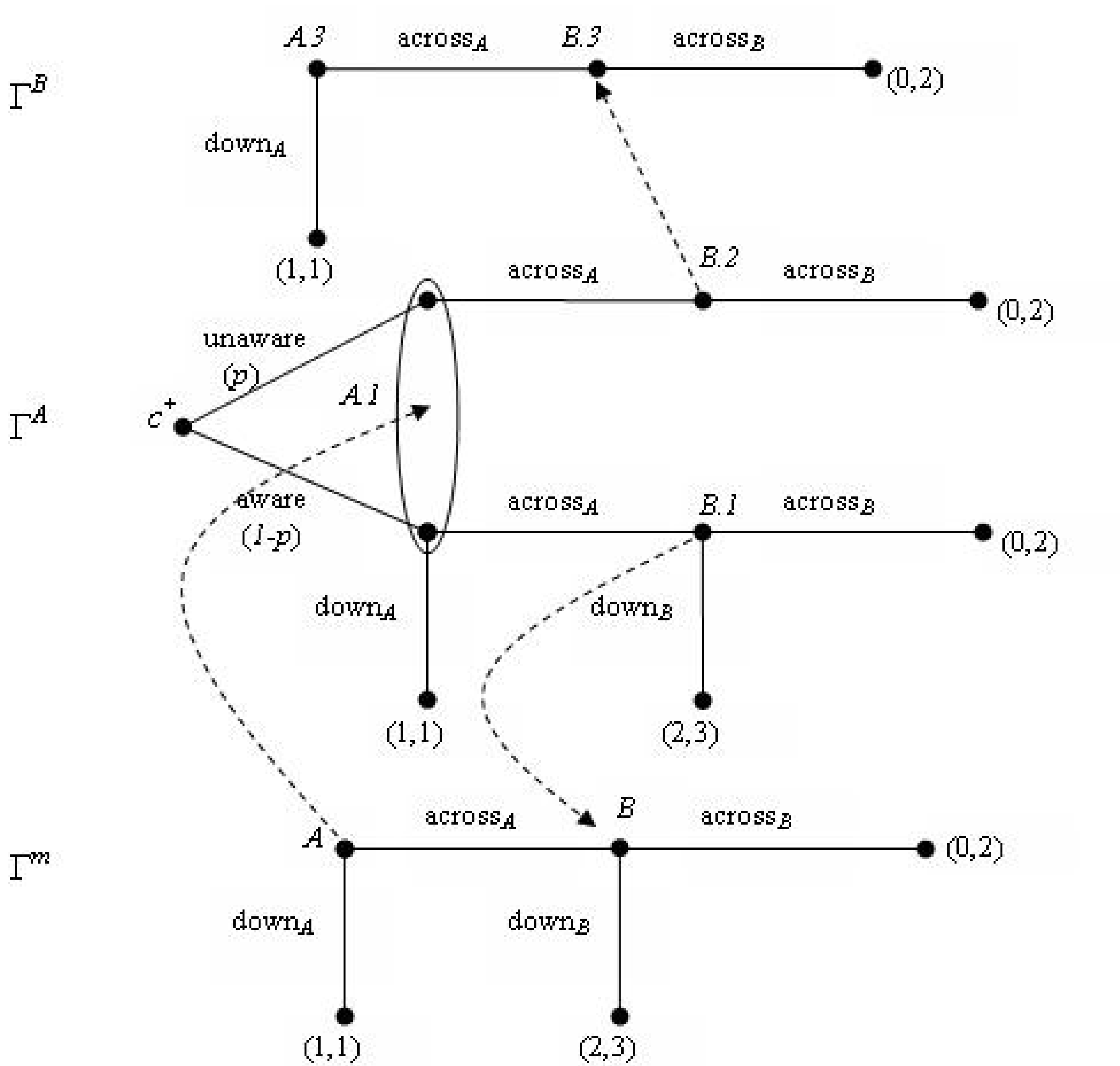
- At node $A.3$, A is not aware of $\langle$across$_A$, down$_B\rangle$;
  - neither is B at $B.3$.
- **Moral:** to fully represent a *game with awareness* we need a set of augmented games.
  - Like a set of possible worlds in Kripke structures

Game with Awareness

A *game with awareness based on* Γ is a tuple $\Gamma^* = (\mathcal{G}, \Gamma^m, \mathcal{F})$, where

- $\mathcal{G}$ is a countable set of augmented games based on Γ;
- $\Gamma^m \in \mathcal{G}$ is an omniscient modeler's view of the game
- If $\mathcal{F}(\Gamma^+, h, i) = (\Gamma^h, I)$ then
  - in history h of game Γ^+, player i believes that the actual game is Γ^h and he is in information set I
    - I (i's *information set*) describes the histories where i thinks might be in Γ^h
- There are some consistency conditions that Γ^* must satisfy
  - See paper

Example Continued



Adding Awareness to Games

- There are many games of awareness based on an underlying game Γ, that vary depending on
  - how players' awareness levels change over time;
  - players' beliefs about other players' awareness level.
- **Example:** If in the simple game Γ, we have considered so far, both players were indeed unaware of run $\langle \text{across}_A, \text{down}_B \rangle$, and this was common knowledge, then in the modeler's game of this example, players' awareness level would not include this run, and we would only need game Γ^B to model players' view of the game.

Canonical Representation

A standard extensive game Γ can be identified with the game $(\{\Gamma^m\}, \Gamma^m, \mathcal{F})$, where $\mathcal{F}(\Gamma^m, h) = (\Gamma^m, I)$ for $h \in I$ and $\Gamma^m = (\Gamma, \{A_i : i \in N\})$, where $\forall i$ and $\forall h \in H_i^m$, $A_i(h) = H$.

- This is the *canonical representation of* Γ as a game with awareness.

Intuition:

- In the canonical representation of Γ as a game with awareness, all players understand the structure of the underlying game Γ and this is common knowledge among players.
- A standard game can be viewed as a special case of a game with awareness, where the underlying game is common knowledge.

Strategies

- In a standard game, a *behavioral strategy* for player i is a function from i-information sets to a distribution over moves.
  - i must do the same thing at histories i cannot tell apart.
  - A strategy is a *universal plan*, describing what i will do in every possible circumstance.
  - In games with awareness, this does not make sense!
    - A player cannot plan in advance what he will do when he becomes aware of new moves.

Local Strategies

- In a game $\Gamma^* = (\mathcal{G}, \Gamma^m, \mathcal{F})$ with awareness, we consider a collection of *local strategies*, one for each augmented game an agent may consider to be the true one in some situation.
  - Intuitively, a local strategy $\sigma_{i,\Gamma'}$ for game Γ' is the strategy that i would use if i were called upon to play and i thought that the true game was Γ'.
- There may be no relationship between the strategies $\sigma_{i,\Gamma'}$ for different games Γ'.

Generalized Nash Equilibrium

- Intuition: $\vec{\sigma}$ is a generalized Nash equilibrium if for every player i, if i believes he is playing game Γ', then his local strategy $\sigma_{i,\Gamma'}$ is a best response to the local strategies of other players in Γ'.
  - The local strategies of the other players are part of $\vec{\sigma}$.
- $\vec{\sigma}$ is a Nash equilibrium of a standard game Γ iff $\vec{\sigma}$ is a (generalized) Nash equilibrium of the canonical representation of Γ as a game with awareness.

Theorem: Every game with awareness has at least one generalized Nash equilibrium.

Awareness of Unawareness

Sometimes players may be aware that they are unaware of relevant moves:

- War settings:
  - you know that an enemy may have new technologies of which you are not aware
- Delaying a decision
  - you may become aware of new issues tomorrow
- Chess
  - “lack of awareness” $\leftrightarrow$ “inability to compute”

Modeling Awareness of Unawareness

- If i is aware that j can make a move at h that i is not aware of, then j can make a "virtual move" at h in i's subjective representation of the game
  - The payoffs after a virtual move reflect i's beliefs about the outcome after the move.
    - Just like associating a value to a board position in chess
- Again, there is guaranteed to be a generalized Nash equilibrium.

Reasoning About Games

- Game theorists reason about games using knowledge
  - Do you know/believe your opponent(s) are rational?
- They essentially model knowledge using Kripke structures of the form $M = (S, \mathcal{K}_1, \ldots, \mathcal{K}_n, \pi)$
  - S is a set of worlds (states)
    - In game theory: S describes, e.g., the strategies used by the players, the game being played
  - $\mathcal{K}_i$ is a binary relation on S:
    - $(s, s') \in \mathcal{K}_i$ if, in world s, agent i considers s' possible
  - π gives meaning to primitive propositions
  - $(M, s) \models K_i\varphi$ if $(M, s') \models \varphi$ for all $(s, s') \in \mathcal{K}_i$.

Adding Awareness: A Biased History

[Hintikka, 1962]: The standard semantics for epistemic logic suffers from the *logical omniscience* problem:

- agents know all tautologies and know all the logical consequences of their knowledge.

One approach for dealing with logical omniscience:

- model agent's lack of awareness [Fagin and Halpern, 1985/88] (FH from now on)
- This allows us to model, e.g., agents who are not aware of all moves in a game

Capturing Awareness

- FH model awareness using a syntactic awareness operator.
  - Awareness described explicitly
    - by listing the formulas an agent is aware of at each state
- Add operators A_i and X_i for each agent i to the language
  - $A_i\varphi$: agent i is aware of φ
  - $X_i\varphi$: agent i explicitly knows φ
- $\mathcal{A}_i(s)$: the formulas that agent i is aware of at state s
  - $(M, s) \models A_i\varphi$ if $\varphi \in \mathcal{A}_i(s)$
- $X_i\varphi$ is true if i implicitly knows φ (it's true at all worlds the agent considers possible) and is aware of it
  - $X_i\varphi \Leftrightarrow K_i\varphi \wedge A_i\varphi$

The MR-HMS approach

Modica and Rustichini [1994,1999] (MR) took a different approach:

- A possibly different set $\mathcal{L}(s)$ of primitive propositions is associated with each world s.
- $(M, s) \models \varphi$ only if every primitive proposition in φ is in $\mathcal{L}(s)$
- $K\varphi$ is defined as usual (truth in all possible worlds)
- $A\varphi$ is an abbreviation for $K\varphi \vee K\neg K\varphi$
  - Ap holds iff $p \in \mathcal{L}(s)$

Heifetz, Meier, and Schipper [2003, 2008] (HMS) extend the MR approach to multiple agents.

- The extension is nontrivial: requires lattices of state spaces, with projection functions between them.

Constraints on FH Awareness

Can impose constraints on $\mathcal{A}_i$:

- awareness is *generated by primitive propositions (agpp)* if an agent is aware of φ iff he is aware of all of the primitive in φ: $\varphi \in \mathcal{A}_i(s)$ iff $\Phi_\varphi \subseteq \mathcal{A}_i(s)$
  - $\Phi_\varphi =$ primitive propositions in φ
- *Agents know what they are aware of (ka)*: if $(s,t) \in \mathcal{K}_i$ then $\mathcal{A}_i(s) = \mathcal{A}_i(t)$

Halpern [2001] and Halpern and Rêgo [2007] show that MR-HMS models are special cases of the FH awareness model where awareness is generated by primitive propositions and agents know what they are aware of.

Capturing Knowledge of Unawareness

If awareness is generated by primitive propositions (as in the MR-HMS approach), then an agent *cannot* know that he is unaware of a fact φ.

So how do we model knowledge and awareness of unawareness?

- In [Halpern and Rêgo, 2006/09] (HR), we model knowledge of unawareness by allowing quantification over formulas
  - Can say "agent i knows $\exists$ a formula of which he is unaware.

  Formally, allow formulas of the form $\forall x \varphi$, $\exists x \varphi$
  - The quantification is over quantifier-free formulas
    - $(M, s) \models \forall x \varphi$ iff $(M, s) \models \varphi[x/\psi]$ for all quantifier-free formulas ψ
    - Restriction is necessary to make semantics well defined

The Good News

- We can capture knowledge of lack of awareness:
  - $K_i \exists x(\neg A_i x \wedge K_j x)$: i knows that there is some formula that j knows to be true that he (i) is not aware of
- There is an elegant complete axiomatization

. . . and the Bad News

(Under standard assumptions) it is impossible for an agent to be uncertain about whether he is aware of all formulas.

- Consider $\psi = \neg X_i \neg \forall x A_i x \wedge \neg X_i \forall x A_i x$.
  - agent i considers it possible that she is aware of all formulas and also considers it possible that she is not aware of all formulas.
- ψ is unsatisfiable!
  - Follows from assumption that in all worlds you consider possible you are aware of the same primitive propositions + KD45 axioms

. . . and the Bad News

(Under standard assumptions) it is impossible for an agent to be uncertain about whether he is aware of all formulas.

- Consider $\psi = \neg X_i \neg \forall x A_i x \wedge \neg X_i \forall x A_i x$.
  - agent i considers it possible that she is aware of all formulas and also considers it possible that she is not aware of all formulas.
- ψ is unsatisfiable!
  - Follows from assumption that in all worlds you consider possible you are aware of the same primitive propositions
    + KD45 axioms
    + assumption that the language is the same at all worlds!

A Better Approach

We combine a number of ideas:

- from FH: the basic framework (syntactic awareness)
- from HR: quantification to express knowledge of lack of awareness
- from MR/HMS: idea of allowing a different language at each state.

We get

- an elegant complete axiomatization (a variant of the HR axioms)
    - Extends the HMS axioms to allow knowledge of unawareness
- a model that satisfies ψ
- don't need a "syntactic" awareness function in the model:
    - $A_i\varphi$, $K_i(\varphi \vee \neg\varphi)$, and $K_i\varphi \vee K_i\neg K_i\varphi$ are equivalent.

Syntax and Semantics

Syntax: The syntax is identical to that of HR:

- start with a set of primitive propositions, close off under X_i, A_i, K_i, and quantification ($\forall x$, $\exists x$)

Semantics: Extend FH awareness structures to allow a different language at each state. An *extended awareness structure* is a tuple $M = (S, \mathcal{L}, \mathcal{K}_1, \ldots, \mathcal{K}_n, \mathcal{A}_1, \ldots, \mathcal{A}_n, \pi)$

- S is a set of states
- $\mathcal{L}(s)$ is a set of primitive propositions—the language at state s
- $\mathcal{A}_i(s)$ the set of formulas i is aware of at s
  - At s, i can be aware only of formulas $\varphi \in \Phi(\mathcal{L}(s))$
    - formulas all of whose primitive propositions are in $\mathcal{L}(s)$.

The Truth Relation

We give semantics to formulas just as in HR, except that for a formula φ to be true at a world s, we must have $\varphi \in \Phi(\mathcal{L}(s))$. E.g.:

- $(M, s) \models p$ if $p \in \mathcal{L}(s)$ and $\pi(s, p) = \mathbf{true}$
- $(M, s) \models \neg\varphi$ if $\varphi \in \Phi(\mathcal{L}(s))$ and $(M, s) \not\models \varphi$
- $(M, s) \models \forall x \varphi$ if $\varphi \in \Phi(\mathcal{L}(s))$ and $(M, s) \models \varphi[x/\psi]$ for all quantifier-free sentences $\psi \in \Phi(\mathcal{L}(s))$
- $(M, s) \models K_i\varphi$ if $\varphi \in \Phi(\mathcal{L}(s))$ and $(M, s') \models \varphi$ for all $(s, s') \in \mathcal{K}_i$.

$\neg X_i \neg \forall x A_i x \wedge \neg X_i \forall x A_i x$ is satisfiable in a state s where i considers possible a world t_1 where $\mathcal{L}(t_1) \subseteq \mathcal{A}_i(s)$ and another world t_2 where $\mathcal{L}(t_2) \not\subseteq \mathcal{A}_i(s)$.

Axiomatization

A formula φ is *valid* in a class $\mathcal{N}$ of extended awareness structures if, for all extended awareness structures $M \in \mathcal{N}$ and worlds s such that $\varphi \in \Phi(\mathcal{L}(s))$, $(M, s) \models \varphi$.

In the full paper, we give an axiom system $\mathrm{S5}_e^{K,X,A,A^*,\forall}$ that is sound and complete with respect to structures where $\mathcal{K}_i$ is an equivalence relation and both *agpp* and *ka* hold.

- These are the standard assumptions in the literature.
- *Sound*: all provable formulas are valid

  Complete: all valid formulas are provable
  - Need the notion of validity above.

Axiomatizing K_i Using A_i^*

The standard axioms for K_i do not quite hold.

- $\neg K_i\varphi \Rightarrow K_i\neg K_i\varphi$ is not sound
  - E.g.: if $\varphi \notin \mathcal{L}(s)$, will have $(M, s) \models \neg K_i\varphi \wedge \neg K_i\neg K_i\varphi$.

Let $A_i^*\varphi$ be an abbreviation for $K_i(\varphi \vee \neg\varphi)$.

Proposition: If $\mathcal{K}_i$ is Euclidean, then

$$(A_i^*\varphi \wedge \neg K_i\varphi) \Rightarrow K_i\neg K_i\varphi \text{ is valid.}$$

Well known: If $\mathcal{K}_i$ is Euclidean and *ka* holds, then

$$(A_i\varphi \wedge \neg X_i\varphi) \Rightarrow X_i\neg X_i\varphi \text{ is valid.}$$

Get analogous axioms for K_i and X_i, with A_i^* playing role of A_i.

A, A', and A^*

- Recall that $A_i^*\varphi$ is an abbreviation for $K_i(\varphi \vee \neg\varphi)$
- Let $A_i'\varphi$ be an abbreviation for $K_i\varphi \vee K_i\neg K_i\varphi$
  - Recall that this is how MR and HMS define awareness.

Theorem:

(a) $A_i'\varphi \Rightarrow A_i^*\varphi$ is valid.

(b) If $\mathcal{K}_i$ is Euclidean, then $A_i^*\varphi \Rightarrow A_i'\varphi$ is valid.

(c) If agents know what they are aware of, then $A_i\varphi \Rightarrow A_i^*\varphi$ is valid.

(d) $A_i^*\varphi \Rightarrow A_i\varphi$ is (trivially) valid under the following assumption:

- If $p \notin \mathcal{A}_i(s)$, then $p \notin \mathcal{L}(t)$ for some t such that $(s,t) \in \mathcal{K}_i$.

Under minimal assumptions, A_i, A_i', and A_i^* are equivalent.

Language and Awareness

If $(s,t) \in \mathcal{K}_i$, $\mathcal{L}(t) - \mathcal{A}_i(s)$ may be nonempty

- i may consider possible formulas of which he is unaware.
- i can “label” formulas that he does not “understand”
- This is a feature: we want to allow agents to have some partial information about formulas that they are unaware of.
  - E.g., want $X_1(\exists x(\neg A_1(x) \wedge K_2(x)))$ to be consistent.
- An agent may have enough partial information about a formula he is unaware of that he can describe it sufficiently well to communicate about it.
  - When this happens in natural language, people will come up with a name for a concept and add it to their language.

Summary

- We have
  - a flexible framework for reasoning about (lack of) awareness of moves in games
    - an arguably reasonable extension of Nash equilibrium that takes (lack of) awareness into account
  - a logic for reasoning about knowledge and (lack of) awareness
    - an elegant complete axiomatization for the logic
    - connections to other approaches
    - importance of language

(Some) Open Problems

- Game Theory:
  - What is the right solution concept in games of awareness?
  - Where do your beliefs come from when you become aware of something new
    - Ozbay (2007): that should be part of the solution concept
  - Connect to work defining solution concepts that take computation into account (joint with Rafael Pass)
- Logic:
  - adding probability
  - dynamics of language change