Reasoning about Knowledge and Beliefs

Lecture 18

Eric Pacuit

University of Maryland, College Park

pacuit.org
epacuit@umd.edu

November 19, 2013

Puzzles about *interactive* knowledge and beliefs

$K_i E$: "i knows that E"

$K_i K_j E$: "i knows that j knows that E"

Alternative history...

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III. Management Science Theory* **14***: 159-182, 1967-68.*

Robert Aumann. *Agreeing to Disagree*. Annals of Statistics **4** (1976).

Alternative history...

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III. Management Science Theory* **14***: 159-182, 1967-68.*

Robert Aumann. *Agreeing to Disagree*. Annals of Statistics **4** (1976).

R. Aumann. *Interactive Epistemology I & II*. International Journal of Game Theory (1999).

P. Battigalli and G. Bonanno. *Recent results on belief, knowledge and the epistemic foundations of game theory*. Research in Economics (1999).

R. Myerson. *Harsanyi's Games with Incomplete Information*. Special 50th anniversary issue of *Management Science*, 2004.

Harsanyi Type Space

John C. Harsanyi, nobel prize winner in economics, developed a theory of games with **incomplete information**.

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III. Management Science Theory* **14***: 159-182, 1967-68.*

Harsanyi Type Space

John C. Harsanyi, nobel prize winner in economics, developed a theory of games with **incomplete information**.

1. incomplete information: uncertainty about the *structure* of the game (outcomes, payoffs, strategy space)
2. imperfect information: uncertainty *within the game* about the previous moves of the players

J. Harsanyi. *Games with incomplete information played by "Bayesian" players I-III. Management Science Theory* **14**: *159-182, 1967-68.*

Harsanyi's Problem

A natural question following any game-theoretic analysis is

Harsanyi's Problem

A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*?

Harsanyi's Problem

A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

Harsanyi's Problem

A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know

Harsanyi's Problem

A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)

Harsanyi's Problem

A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)

Harsanyi's Problem

A natural question following any game-theoretic analysis is *how would the players react if some parameters of the model are not known to the players*? How do we completely specify such a model?

1. Suppose there is a parameter that some player i does not know
2. i's uncertainty about the parameter must be included in the model (first-order beliefs)
3. this is a new parameter that the other players may not know, so we must specify the players beliefs about this parameter (second-order beliefs)
4. but this is a new parameter, and so on....

Harsanyi's Problem

A (game-theoretic) **type** of a player summarizes everything the player knows privately at the beginning of the game which could affect his beliefs about payoffs in the game and about all other players' types.

(Harsanyi argued that all uncertainty in a game can be equivalently modeled as uncertainty about payoff functions.)

Information in games situations

- imperfect information about the play of the game
- incomplete information about the structure of the game

Information in games situations

- imperfect information about the play of the game
- incomplete information about the structure of the game
- strategic information (what will the other players do?)
- higher-order information (what are the other players thinking?)

Epistemic Game Theory

Formally, a game is described by its strategy sets and payoff functions.

Epistemic Game Theory

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game.

Epistemic Game Theory

Formally, a game is described by its strategy sets and payoff functions. But in real life, may other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game.... The difference lies in the attitudes of the players, in their expectations about each other, in custom, and in history, though the rules of the game do not distinguish between the two situations. *(pg. 72)*

R. Aumann and J. H. Dreze. *Rational Expectations in Games*. American Economic Review 98 (2008), pp. 72-86.

The Epistemic Program in Game Theory

"...the analysis constitutes a fleshing-out of the textbook interpretation of equilibrium as 'rationality plus correct beliefs.'...this suggests that equilibrium behavior cannot arise out of strategic reasoning alone. "

E. Dekel and M. Siniscalchi. *Epistemic Game Theory*. manuscript, 2013.

A. Brandenburger. *The Power of Paradox*. International Journal of Game Theory, 35, pgs. 465 - 492, 2007.

EP and O. Roy. *Epistemic Game Theory*. Stanford Encyclopedia of Philosophy, forthcoming, 2013.

Doesn't such talk of what Ann believes Bob believes about her, and so on, suggest that some kind of self-reference arises in games, similar to the well-known examples of self-reference in mathematical logic.

A. Brandenburger and H. J. Keisler. *An Impossibility Theorem on Beliefs in Games*. *Studia Logica* (2006).

A Paradox

Ann believes that Bob's strongest belief is
that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false?

$*$ A **strongest belief** is a belief that implies all other beliefs.

A. Brandenburger and H. J. Keisler. *An Impossibility Theorem on Beliefs in Games*. *Studia Logica* (2006).

A Paradox

Ann believes that Bob's strongest belief is
that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose Yes.

A Paradox

Ann believes that Bob's strongest belief is
that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose Yes.

Then, Ann believes that it's not the case that Ann believes that Bob's strongest belief is false.

A Paradox

Ann believes that Bob's strongest belief is
that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose Yes.

Then, Ann believes that it's not the case that Ann believes that Bob's strongest belief is false.

So, it's not the case that Ann believes that Bob's strongest belief is false. ($B\neg B\varphi \to \neg B\varphi$)

A Paradox

Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose Yes.

Then, Ann believes that it's not the case that Ann believes that Bob's strongest belief is false.

So, it's not the case that Ann believes that Bob's strongest belief is false. ($B\neg B\varphi \rightarrow \neg B\varphi$)

So, the answer is no.

A Paradox

Ann believes that Bob's strongest belief is
that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose No.

A Paradox

Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose No.

Then, it's not the case that Ann believes it's not the case that Ann believes that Bob's strongest belief is false.

A Paradox

Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose No.

Then, it's not the case that Ann believes it's not the case that Ann believes that Bob's strongest belief is false.

So, Ann believes that Bob's strongest belief is false.
$(\neg B \neg B\varphi \to B\varphi)$

A Paradox

Ann believes that Bob's strongest belief is
that Ann believes that Bob's strongest belief is false.

Does Ann believe that Bob's strongest belief is false? Suppose No.

Then, it's not the case that Ann believes it's not the case that Ann believes that Bob's strongest belief is false.

So, Ann believes that Bob's strongest belief is false.
$(\neg B \neg B \varphi \to B \varphi)$

So, the answer must be yes.

- strongest belief

- strongest belief
- weakest belief

- strongest belief
- weakest belief
- craziest belief

- strongest belief
- weakest belief
- craziest belief
- all of Bob's belief

Is there a space of all possible interactive beliefs of a game?

Is there a space of all possible interactive beliefs of a game?

Two questions

Is there a space of all possible interactive beliefs of a game?

Two questions

- What exactly does “all possible” mean?

Is there a space of all possible interactive beliefs of a game?

Two questions

- What exactly does “all possible” mean?
  (Complete, Canonical, Universal)

Is there a space of all possible interactive beliefs of a game?

Two questions

- What exactly does "all possible" mean?
  (Complete, Canonical, Universal)
- Who cares?

Who Cares?

A. Brandenburger and E. Dekel. *Hierarchies of Beliefs and Common Knowledge*. Journal of Economic Theory (1993).

A. Heifetz and D. Samet. *Knoweldge Spaces with Arbitrarily High Rank*. Games and Economic Behavior (1998).

L. Moss and I. Viglizzo. *Harsanyi type spaces and final coalgebras constructed from satisfied theories*. EN in Theoretical Computer Science (2004).

A. Friendenberg. *When do type structures contain all hierarchies of beliefs?*. working paper (2007).

Who cares?

We think of a particular incomplete structure as giving the "context" in which the game is played.

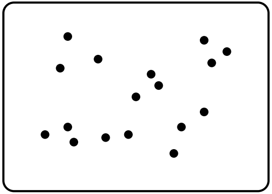
Who cares?

We think of a particular incomplete structure as giving the "context" in which the game is played. In line with Savage's Small-Worlds idea in decision theory [...], who the players are in the given game can be seen as a shorthand for their experiences before the game.

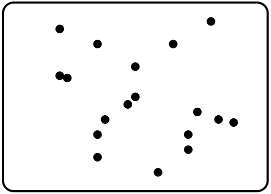
Who cares?

We think of a particular incomplete structure as giving the "context" in which the game is played. In line with Savage's Small-Worlds idea in decision theory [...], who the players are in the given game can be seen as a shorthand for their experiences before the game. The players' possible characteristics — including their possible types — then reflect the prior history or context. (Seen in this light, complete structures represent a special "context-free" case, in which there has been no narrowing down of types.) (pg. 319)

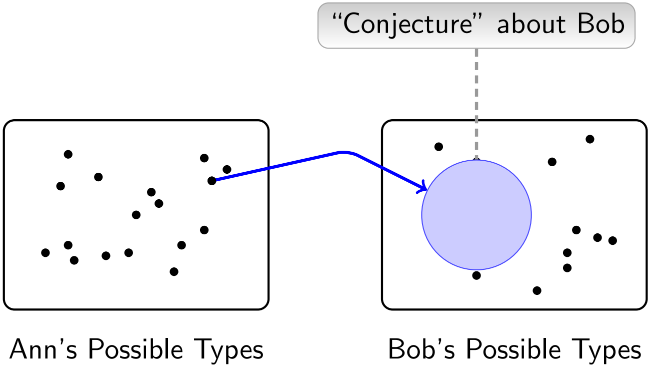
A. Brandenburger, A. Friedenberg, H. J. Keisler. *Admissibility in Games*. Econometrica (2008).

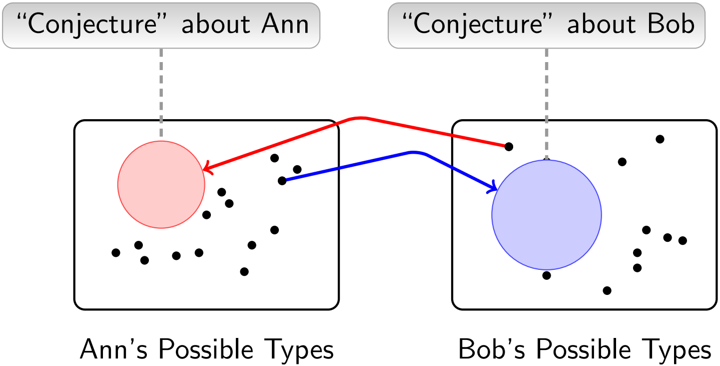


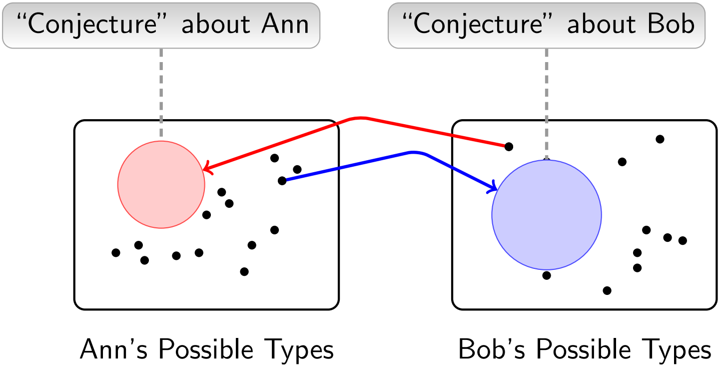
Ann's Possible Types



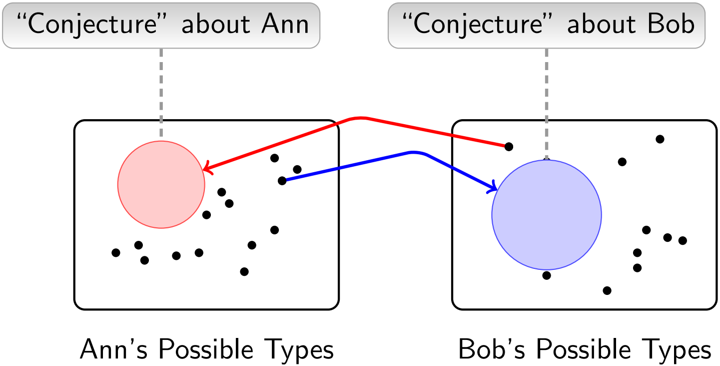
Bob's Possible Types


"Conjecture" about Bob
Ann's Possible Types
Bob's Possible Types


"Conjecture" about Ann
"Conjecture" about Bob
Ann's Possible Types
Bob's Possible Types



Is there a space where every *possible* conjecture is considered by *some* type?



Is there a space where every *possible* conjecture is considered by *some* type? It depends...

S. Abramsky and J. Zvesper. *From Lawvere to Brandenburger-Keisler: interactive forms of diagonalization and self-reference*. Proceedings of LOFT 2010.

EP. *Understanding the Brandenburger Keisler Pardox*. Studia Logica (2007).

Impossibility Results

Language: the (formal) language used by the players to formulate conjectures about their opponents.

Impossibility Results

Language: the (formal) language used by the players to formulate conjectures about their opponents.

Completeness: A model is **complete for a language** if every (consistent) statement in a player's language about an opponent is *considered* by some type.

Qualitative Type Spaces: $\langle T_a, T_b, \lambda_a, \lambda_b \rangle$

$\lambda_a : T_a \rightarrow \wp(T_b)$
$\lambda_b : T_b \rightarrow \wp(T_a)$

Qualitative Type Spaces: $\langle T_a, T_b, \lambda_a, \lambda_b \rangle$

$\lambda_a : T_a \rightarrow \wp(T_b)$
$\lambda_b : T_b \rightarrow \wp(T_a)$

x **believes** a set $Y \subseteq T_b$ if $\lambda_a(x) \subseteq Y$

x **assumes** a set $Y \subseteq T_b$ if $\lambda_a(x) = Y$

Impossibility Results

Impossibility 1 There is no complete interactive belief structure for the *powerset language*.

Proof. Cantor: there is no onto map from X to the nonempty subsets of X.

Impossibility Results

Impossibility 1 There is no complete interactive belief structure for the *powerset language*.

Proof. Cantor: there is no onto map from X to the nonempty subsets of X.

Impossibility 2 (Brandenburger and Keisler) There is no complete interactive belief structure for *first-order logic*.

Suppose that $\mathcal{C}_A \subseteq \wp(T_A)$ is a set of *conjectures* about Ann and $\mathcal{C}_B \subseteq \wp(T_B)$ a set of conjectures about Bob states.

Assume For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$: "in state x_0, Ann has consistent beliefs"
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$: "in state x_0, Ann believes that Bob's strongest belief is that X"

Lemma. Under the above assumption, for each $X \in \mathcal{C}_A$ there is an x_0 such that

$$x_0 \in X \text{ iff there is a } y \in T_B \text{ such that } y \in \lambda_A(x_0) \text{ and } x_0 \in \lambda_B(y)$$

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$. By 1., $\lambda_A(x_0) \neq \emptyset$ so there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$. By 1., $\lambda_A(x_0) \neq \emptyset$ so there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$. We show that $x_0 \in \lambda_B(y_0)$.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$. By 1., $\lambda_A(x_0) \neq \emptyset$ so there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$. We show that $x_0 \in \lambda_B(y_0)$. By 2., we have $y_0 \in \lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$. Hence, $x_0 \in X = \lambda_B(y_0)$.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$. By 1., $\lambda_A(x_0) \neq \emptyset$ so there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$. We show that $x_0 \in \lambda_B(y_0)$. By 2., we have $y_0 \in \lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$. Hence, $x_0 \in X = \lambda_B(y_0)$.

Suppose that there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y_0)$.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$. By 1., $\lambda_A(x_0) \neq \emptyset$ so there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$. We show that $x_0 \in \lambda_B(y_0)$. By 2., we have $y_0 \in \lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$. Hence, $x_0 \in X = \lambda_B(y_0)$.

Suppose that there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y_0)$. By 2., $y_0 \in \lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$.

Claim. $x_0 \in X$ iff $\exists y \in T_B$, $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Assumption: For all $X \in \mathcal{C}_A$ there is a $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$

Suppose that $X \in \mathcal{C}_A$. Then there is an $x_0 \in T_A$ satisfying 1 and 2.

Suppose that $x_0 \in X$. By 1., $\lambda_A(x_0) \neq \emptyset$ so there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$. We show that $x_0 \in \lambda_B(y_0)$. By 2., we have $y_0 \in \lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$. Hence, $x_0 \in X = \lambda_B(y_0)$.

Suppose that there is a $y_0 \in T_B$ such that $y_0 \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y_0)$. By 2., $y_0 \in \lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$. Hence, $x_0 \in \lambda_B(y_0) = X$.

Consider a first-order language $\mathcal{L}$ containing binary relational symbols $R_A(x, y)$ and $R_B(x, y)$ defining λ_A and λ_B, respectively.

Consider a first-order language $\mathcal{L}$ containing binary relational symbols $R_A(x, y)$ and $R_B(x, y)$ defining λ_A and λ_B, respectively.

$\mathcal{L}$ is interpreted over qualitative type structures where the interpretation of R_A is $\{(t, s) \mid t \in T_A, s \in T_B, \text{ and } s \in \lambda_A(t)\}$.

Consider a first-order language $\mathcal{L}$ containing binary relational symbols $R_A(x,y)$ and $R_B(x,y)$ defining λ_A and λ_B, respectively.

$\mathcal{L}$ is interpreted over qualitative type structures where the interpretation of R_A is $\{(t,s) \mid t \in T_A, s \in T_B, \text{ and } s \in \lambda_A(t)\}$.

Consider the formula φ in $\mathcal{L}$:

$$\varphi(x) := \exists y(R_A(x,y) \wedge R_B(y,x))$$

Consider a first-order language $\mathcal{L}$ containing binary relational symbols $R_A(x, y)$ and $R_B(x, y)$ defining λ_A and λ_B, respectively.

$\mathcal{L}$ is interpreted over qualitative type structures where the interpretation of R_A is $\{(t, s) \mid t \in T_A, s \in T_B, \text{ and } s \in \lambda_A(t)\}$.

Consider the formula φ in $\mathcal{L}$:

$$\varphi(x) := \exists y(R_A(x, y) \wedge R_B(y, x))$$

$\neg\varphi(x) := \forall y(R_A(x, y) \to \neg R_B(y, x))$: "Ann believes that Bob's strongest belief is *false*."

Proof of the Theorem

Suppose that $X \in \mathcal{C}_A$ is defined by the formula
$\neg\varphi(x) := \neg\exists y(R_A(x,y) \wedge R_B(y,x))$.

Proof of the Theorem

Suppose that $X \in \mathcal{C}_A$ is defined by the formula $\neg\varphi(x) := \neg\exists y(R_A(x,y) \wedge R_B(y,x))$.

There is an $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$: Ann's beliefs at x_0 are consistent.
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$: At x_0, Ann believes that Bob's strongest belief is that $X = \{x \mid \neg\varphi(x)\}$ (i.e., Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.)

Proof of the Theorem

Suppose that $X \in \mathcal{C}_A$ is defined by the formula
$\neg\varphi(x) := \neg\exists y(R_A(x,y) \wedge R_B(y,x))$.

There is an $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$: Ann's beliefs at x_0 are consistent.
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$: At x_0, Ann believes that Bob's strongest belief is that $X = \{x \mid \neg\varphi(x)\}$ (i.e., Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.)

$\neg\varphi(x_0)$ is true iff (def. of X) $x_0 \in X$

Proof of the Theorem

Suppose that $X \in \mathcal{C}_A$ is defined by the formula $\neg\varphi(x) := \neg\exists y(R_A(x,y) \wedge R_B(y,x))$.

There is an $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$: Ann's beliefs at x_0 are consistent.
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$: At x_0, Ann believes that Bob's strongest belief is that $X = \{x \mid \neg\varphi(x)\}$ (i.e., Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.)

$\neg\varphi(x_0)$ is true	iff (def. of X)	$x_0 \in X$
	iff (Lemma)	there is a $y \in T_B$ with $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$

Proof of the Theorem

Suppose that $X \in \mathcal{C}_A$ is defined by the formula
$\neg\varphi(x) := \neg\exists y(R_A(x,y) \wedge R_B(y,x))$.

There is an $x_0 \in T_A$ such that

1. $\lambda_A(x_0) \neq \emptyset$: Ann's beliefs at x_0 are consistent.
2. $\lambda_A(x_0) \subseteq \{y \mid \lambda_B(y) = X\}$: At x_0, Ann believes that Bob's strongest belief is that $X = \{x \mid \neg\varphi(x)\}$ (i.e., Ann believes that Bob's strongest belief is that Ann believes that Bob's strongest belief is false.)

$\neg\varphi(x_0)$ is true	iff (def. of X)	$x_0 \in X$
	iff (Lemma)	there is a $y \in T_B$ with $y \in \lambda_A(x_0)$ and $x_0 \in \lambda_B(y)$
	iff (def. of $\varphi(x)$)	$\varphi(x_0)$ is true.