Puzzles

Eric Pacuit

Department of Philosophy
University of Maryland

pacuit.org/esslli2019/puzzles

August 9, 2019

Coordinated Attack/Email Game

J. Gray. *Notes on database operating systems*. in *Operating Systems: An Advanced Course*, Lecture Notes in Computer Science, Vol. 66, 1978.

J. Y. Halpern and Y. Moses. *Knowledge and common knowledge in a distributed environment*. Journal of the ACM, 37(3):549 - 587, 1990.

A. Rubinstein. *The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge'*. American Economic Review, 79, 385 - 391, 1989.

Coordinated Attack/Email Game

- Two generals, each commanding a division of an army, want to attack a common enemy.
- They will win the battle only if they attack the enemy simultaneously; if only one division attacks, it will be defeated.
- Thus, the generals want to coordinate their attack.

Coordinated Attack/Email Game

- Two generals, each commanding a division of an army, want to attack a common enemy.
- They will win the battle only if they attack the enemy simultaneously; if only one division attacks, it will be defeated.
- Thus, the generals want to coordinate their attack.
- Unfortunately, the only way they have of communicating is by means of messengers, who might get lost or captured by the enemy.

Coordinated Attack/Email Game

- Two generals, each commanding a division of an army, want to attack a common enemy.
- They will win the battle only if they attack the enemy simultaneously; if only one division attacks, it will be defeated.
- Thus, the generals want to coordinate their attack.
- Unfortunately, the only way they have of communicating is by means of messengers, who might get lost or captured by the enemy.
- How many messages are needed for the generals to coordinate their attack?

The Guessing Game

Guess a number between 1 & 100.

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

app.pacuit.io/games/avg

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess?

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? 100

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, 99

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, ~~99~~, . . . , 67

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, ~~99~~, . . . , ~~67~~, . . . , 2, 1

The Guessing Game

Guess a number between 1 & 100.
The closest to 2/3 of the average wins.

What number should you guess? ~~100~~, ~~99~~, . . . , ~~67~~, . . . , ~~2~~, (1)

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.
4. The person that wrote the smaller number will receive that amount plus 2 EUR (as a reward), and the person that wrote the larger number will receive the smaller number minus 2 EUR (as a punishment).

Traveler's Dilemma

1. You and your friend write down an integer between 2 and 100 (without discussing).
2. If both of you write down the same number, then both will receive that amount in Euros from the airline in compensation.
3. If the numbers are different, then the airline assumes that the smaller number is the actual price of the luggage.
4. The person that wrote the smaller number will receive that amount plus 2 EUR (as a reward), and the person that wrote the larger number will receive the smaller number minus 2 EUR (as a punishment).

Suppose that you are randomly paired with another person here at ESSLLI. What number would you write down?

app.pacuit.io/games/td

- Belief paradoxes: The knower paradox, Buriden-Burge sentences, Anti-expert sentences, Prior’s sentence
- The Brandenburger-Keisler (BK) paradox

Buridan-Burge: $p \leftrightarrow \neg B_a p$

Anti-Expert: $p \leftrightarrow B_a \neg p$

Tyler Burge. *Epistemic paradox*. Journal of Philosophy, 81(1), pgs. 5 - 29, 1984.

Michael Caie. *Belief and indeterminacy*. The Philosophical Review, 121(1), pgs. 1 - 54, 2012.

Buridan-Burge: $p \leftrightarrow \neg B_a p$
$B_a(p \leftrightarrow \neg B_a p)$

Anti-Expert: $p \leftrightarrow B_a \neg p$
$B_a(p \leftrightarrow B_a \neg p)$

Tyler Burge. *Epistemic paradox*. Journal of Philosophy, 81(1), pgs. 5 - 29, 1984.

Michael Caie. *Belief and indeterminacy*. The Philosophical Review, 121(1), pgs. 1 - 54, 2012.

Prior's Theorem

$$Q(\forall p(Qp \to \neg p)) \to (\exists p(Qp \wedge p) \wedge \exists p(Qp \wedge \neg p))$$

is a derivable using Universal Instantiation and propositional reasoning.

A. N. Prior. *On a family of paradoxes*. Notre Dame Journal of Formal Logic, 2(1), pgs. 16 - 32, 1961.

Prior's Theorem

T1. *C(UpCdpNp) C(dUpCdpNp)(NUpCdpNp)* – from *CUpdpdq* by substitution.

T2. *C(dUpCdpNp) C(UpCdpNp) (NUpCdpNp)* – from *T1* and *CCpCqrCqCpr*.

T3. *C(dUpCdpNp) (NUpCdpNp)* – from *T2* and *CCpCqNqCpNq*.

T4. *C(dUpCdpNp)(EpKdpp)* – from *T3* and equivalence of 'not-none' and 'some', i.e. of 'not-all-not' and 'some'.

T5. *C(dUpCdpNp) K(dUpCdpNp) (NUpCdpNp)* – from *T3* and *CCpqCpKpq*.

T6. *CK(dUpCdpNp)(NUpCdpNp)(EpKdpNp)* – substitution in *CdqEpdp*.

T7. *C(dUpCdpNp)(EpKdpNp)* – syllogistically from *T5* and *T6*.

T8. *C(dUpCdpNp) K(EpKdpp)(EpKdpNp)* – from *T4*, *T7* and *CCpqCCprCpKqr*.

Prior's Theorem

1. $\forall p (Qp \to \neg p) \to (Q(\forall p(Qp \to \neg p)) \to \neg \forall p(Qp \to \neg p))$

$(\forall p\, \varphi(p) \to \varphi[p/q])$

Prior's Theorem

1. $\forall p(Qp \rightarrow \neg p) \rightarrow (Q(\forall p(Qp \rightarrow \neg p)) \rightarrow \neg\forall p(Qp \rightarrow \neg p))$

   $(\forall p\varphi(p) \rightarrow \varphi[p/q])$

2. $Q(\forall p(Qp \rightarrow \neg p)) \rightarrow (\forall p(Qp \rightarrow \neg p) \rightarrow \neg\forall p(Qp \rightarrow \neg p))$

   $((a \rightarrow (b \rightarrow c)) \rightarrow (b \rightarrow (a \rightarrow c)))$

Prior's Theorem

1. $\forall p\,(Qp \to \neg p) \;\to (Q(\,\forall p(Qp \to \neg p)\,) \to \neg\,\forall p(Qp \to \neg p)\,)$

   $(\forall p \varphi(p) \to \varphi[p/q])$

2. $Q(\forall p(Qp \to \neg p)) \to (\,\forall p(Qp \to \neg p) \to \neg\forall p(Qp \to \neg p)\,)$

   $((\,a \;\to (b \to c)) \to (b \to (\,a \;\to c)))$

3. $Q(\forall p(Qp \to \neg p)) \to \neg\forall p(Qp \to \neg p)$

   $((\,a \to (\,b \to \neg b\,)) \to (\,a \to \neg b\,))$

Prior's Theorem

1. $\forall p\,(Qp \rightarrow \neg p) \rightarrow (Q(\forall p(Qp \rightarrow \neg p)) \rightarrow \neg\forall p(Qp \rightarrow \neg p))$

   $(\forall p\varphi(p) \rightarrow \varphi[p/q])$

2. $Q(\forall p(Qp \rightarrow \neg p)) \rightarrow (\forall p(Qp \rightarrow \neg p) \rightarrow \neg\forall p(Qp \rightarrow \neg p))$

   $((a \rightarrow (b \rightarrow c)) \rightarrow (b \rightarrow (a \rightarrow c)))$

3. $Q(\forall p(Qp \rightarrow \neg p)) \rightarrow \neg\forall p(Qp \rightarrow \neg p)$

   $((a \rightarrow (b \rightarrow \neg b)) \rightarrow (a \rightarrow \neg b))$

4. $Q(\forall p(Qp \rightarrow \neg p)) \rightarrow \exists p(Qp \wedge p)$

   ($\neg\forall p\varphi \leftrightarrow \exists p\neg\varphi$ and $\neg(a \rightarrow \neg b) \leftrightarrow (a \wedge b)$)

Prior's Theorem

1. $\forall p\,(Qp \to \neg p) \to (Q(\forall p(Qp \to \neg p)) \to \neg\forall p(Qp \to \neg p))$

   $(\forall p\varphi(p) \to \varphi[p/q])$

2. $Q(\forall p(Qp \to \neg p)) \to (\forall p(Qp \to \neg p) \to \neg\forall p(Qp \to \neg p))$

   $((a \to (b \to c)) \to (b \to (a \to c)))$

3. $Q(\forall p(Qp \to \neg p)) \to \neg\forall p(Qp \to \neg p)$

   $((a \to (b \to \neg b)) \to (a \to \neg b))$

4. $Q(\forall p(Qp \to \neg p)) \to \exists p(Qp \wedge p)$

   ($\neg\forall p\varphi \leftrightarrow \exists p\neg\varphi$ and $\neg(a \to \neg b) \leftrightarrow (a \wedge b)$)

5. $Q(\forall p(Qp \to \neg p)) \to (Q(\forall p(Qp \to \neg p)) \wedge \neg\forall p(Qp \to \neg p))$

   $((a \to b) \to (a \to (a \wedge b)))$

Prior's Theorem

1. $\forall p\,(Qp \to \neg p) \to (Q(\,\forall p(Qp \to \neg p)\,) \to \neg\,\forall p(Qp \to \neg p)\,)$
   $(\forall p\varphi(p) \to \varphi[p/q])$

2. $Q(\forall p(Qp \to \neg p)) \to (\,\forall p(Qp \to \neg p) \to \neg\forall p(Qp \to \neg p)\,)$
   $((\,a \to (b \to c)) \to (b \to (\,a \to c)))$

3. $Q(\forall p(Qp \to \neg p)) \to \neg\forall p(Qp \to \neg p)$
   $((a \to (b \to \neg b)) \to (a \to \neg b))$

4. $Q(\forall p(Qp \to \neg p)) \to \exists p(Qp \wedge p)$
   $(\,\neg\forall p\varphi \leftrightarrow \exists p\neg\varphi$ and $\neg(a \to \neg b) \leftrightarrow (a \wedge b))$

5. $Q(\forall p(Qp \to \neg p)) \to (Q(\forall p(Qp \to \neg p)) \wedge \neg\forall p(Qp \to \neg p))$
   $(a \to b) \to (a \to (a \wedge b))$

6. $Q(\forall p(Qp \to \neg p)) \to \exists p(Qp \wedge \neg p)$
   $((Q\varphi \wedge \neg\varphi) \to \exists p(Qp \wedge \neg p))$

Prior's Theorem

1. $\forall p\,(Qp \to \neg p) \to (Q(\,\forall p(Qp \to \neg p)\,) \to \neg\,\forall p(Qp \to \neg p)\,)$

   $(\forall p\varphi(p) \to \varphi[p/q])$

2. $Q(\forall p(Qp \to \neg p)) \to (\,\forall p(Qp \to \neg p) \to \neg\forall p(Qp \to \neg p)\,)$

   $((\,a \to (b \to c)) \to (b \to (\,a \to c)))$

3. $Q(\forall p(Qp \to \neg p)) \to \neg\forall p(Qp \to \neg p)$

   $((a \to (b \to \neg b)) \to (a \to \neg b))$

4. $Q(\forall p(Qp \to \neg p)) \to \exists p(Qp \wedge p)$

   ($\neg\forall p\varphi \leftrightarrow \exists p\neg\varphi$ and $\neg(a \to \neg b) \leftrightarrow (a \wedge b)$)

5. $Q(\forall p(Qp \to \neg p)) \to (\,Q(\forall p(Qp \to \neg p)) \wedge \neg\forall p(Qp \to \neg p)\,)$

   $(a \to b) \to (a \to (a \wedge b))$

6. $Q(\forall p(Qp \to \neg p)) \to \exists p(Qp \wedge \neg p)$

   $((Q\varphi \wedge \neg\varphi) \to \exists p(Qp \wedge \neg p))$

7. $Q(\forall p(Qp \to \neg p)) \to (\exists p(Qp \wedge p) \wedge \exists p(Qp \wedge \neg p))$

   $(((a \to b) \wedge (a \to c)) \to (a \to (b \wedge c)))$

Prior's Theorem

1. $\forall p\,(Qp \to \neg p) \;\to (Q(\;\forall p(Qp \to \neg p)\;) \to \neg\;\forall p(Qp \to \neg p)\;)$

   $(\forall p\varphi(p) \to \varphi[p/q])$

2. $Q(\forall p(Qp \to \neg p)) \;\to (\;\forall p(Qp \to \neg p) \;\to\; \neg\forall p(Qp \to \neg p)\;)$

   $((\;a \;\to (b \to c)) \to (b \to (\;a \;\to c)))$

3. $Q(\forall p(Qp \to \neg p)) \;\to\; \neg\forall p(Qp \to \neg p)$

   $((a \to (b \to \neg b)) \to (a \to \neg b))$

4. $Q(\forall p(Qp \to \neg p)) \to \;\exists p(Qp \wedge p)$

   $(\;\neg\forall p\varphi \leftrightarrow \exists p\neg\varphi$ and $\neg(a \to \neg b) \leftrightarrow (a \wedge b))$

5. $Q(\forall p(Qp \to \neg p)) \;\to (\;Q(\forall p(Qp \to \neg p)) \;\wedge\; \neg\forall p(Qp \to \neg p)\;)$

   $(a \to b) \to (a \to (a \wedge b))$

6. $Q(\forall p(Qp \to \neg p)) \to \;\exists p(Qp \wedge \neg p)$

   $((Q\varphi \wedge \neg\varphi) \to \exists p(Qp \wedge \neg p))$

7. $Q(\forall p(Qp \to \neg p)) \to (\;\exists p(Qp \wedge p) \;\wedge\; \exists p(Qp \wedge \neg p)\;)$

   $(((a \to b) \wedge (a \to c)) \to (a \to (b \wedge c)))$

$$Q(\forall p(Qp \to \neg p)) \to (\exists p(Qp \wedge p) \wedge \exists p(Qp \wedge \neg p))$$

$$Q(\forall p(Qp \rightarrow \neg p)) \rightarrow (\exists p(Qp \wedge p) \wedge \exists p(Qp \wedge \neg p))$$

- $Q\varphi :=$ Ann believes that φ

  If Ann believes that everything that Ann believes is wrong, then Ann believes something true and Ann believes something wrong.

$$Q(\forall p(Qp \to \neg p)) \to (\exists p(Qp \wedge p) \wedge \exists p(Qp \wedge \neg p))$$

- $Q\varphi :=$ Ann believes that φ

  If Ann believes that everything that Ann believes is wrong, then Ann believes something true and Ann believes something wrong.

- $Q\varphi :=$ Ann says that φ

  If Ann says that everything that Ann says is wrong, then Ann says something true and Ann says something wrong.

$$Q(\forall p(Qp \to \neg p)) \to (\exists p(Qp \wedge p) \wedge \exists p(Qp \wedge \neg p))$$

- $Q\varphi :=$ Ann believes that φ

  If Ann believes that everything that Ann believes is wrong, then Ann believes something true and Ann believes something wrong.

- $Q\varphi :=$ Ann says that φ

  If Ann says that everything that Ann says is wrong, then Ann says something true and Ann says something wrong.

- $Q\varphi :=$ Ann wrote on the board at midnight that φ

  If Ann wrote on the board at midnight that everything that Ann wrote on the board at midnight is wrong, then Ann wrote a true thing on the board at midnight and Ann wrote a false thing on the board at midnight.

A. Bacon, J. Hawthorne and G. Uzquiano. *Higher-Order Free Logic and the Prior-Kaplan Paradox*. Forthcoming in *Williamson on Modality*.

A. Bacon and G. Uzquiano. *Some results on the limits of thought*. Journal of Philosophical Logic, 2018.

R. H. Thomason and D. Tucker. *Paradoxes of Intensionality*. Review of Symbolic Logic, 4, pgs. 394 - 411, 2011.

The Epistemic Program in Game Theory

...the analysis constitutes a fleshing-out of the textbook interpretation of equilibrium as 'rationality' plus correct beliefs.

E. Dekel and M. Siniscalchi. *Epistemic Game Theory*. Handbook of Game Theory with Economic Applications, Volume 4, 2015, pgs. 619 - 702.

A. Brandenburger. *The Language of Game Theory*. World Scientific Press, 2014.

A. Perea. *Epistemic Game Theory: Reasoning and Choice*. Cambridge University, 2012.

EP and O. Roy. *Epistemic Foundations of Game Theory*. Stanford Encyclopedia of Philosophy, 2015.

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).

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).

EP. *Understanding the Brandenburger-Keisler Paradox*. Studia Logica, 86(3), pgs. 435 - 454, 2007.

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

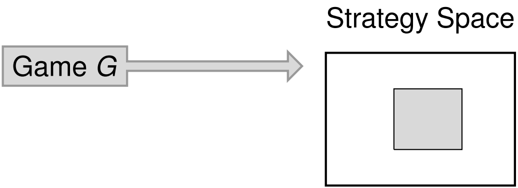
C. Baskent. *Some non-classical approaches to the Brandenburger¢Keisler paradox* . Logic Journal of the IGPL, 23(4): 533-552, 2015.

The Epistemic Program in Game Theory

Game G

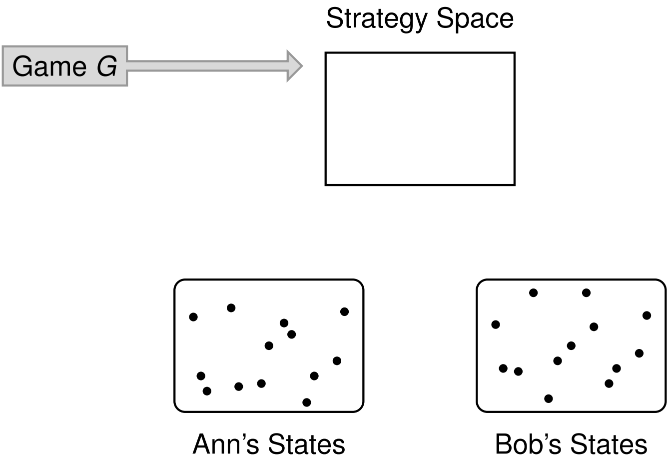
G: available actions, payoffs, structure of the decision problem

The Epistemic Program in Game Theory



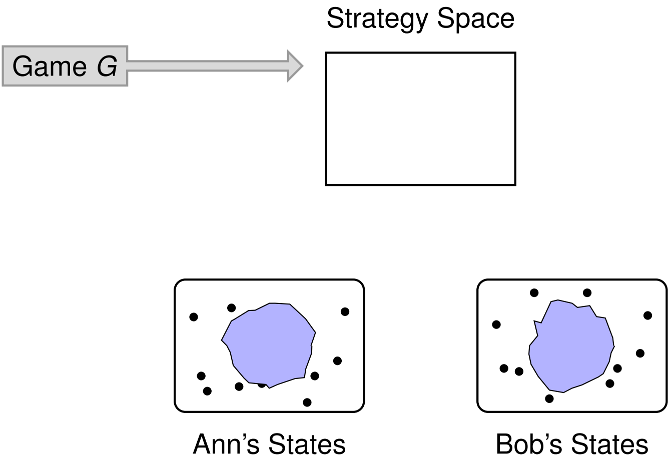
solution concepts are systematic descriptions of what players *do*

The Epistemic Program in Game Theory



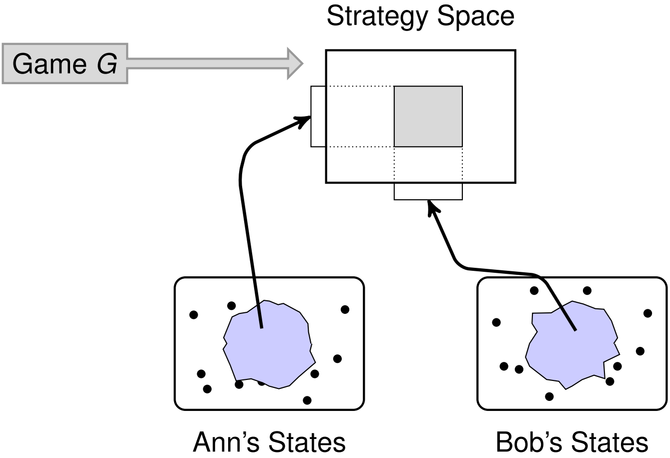
The game model includes *information states* of the players

The Epistemic Program in Game Theory

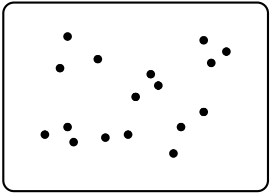


Restrict to information states satisfying some rationality condition

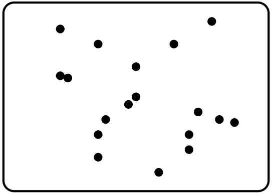
The Epistemic Program in Game Theory



Project onto the strategy space



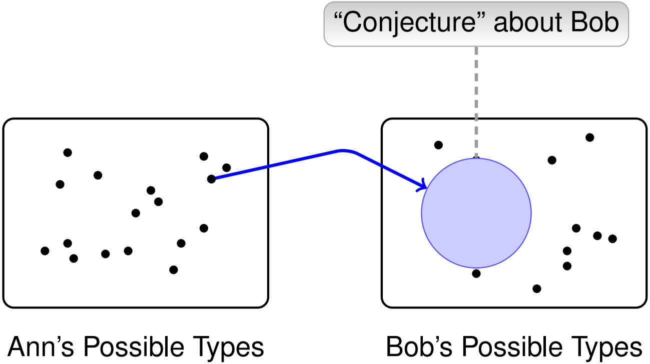
Ann's Possible Types

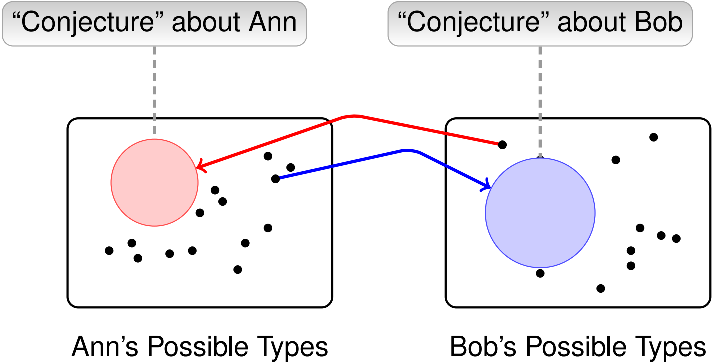


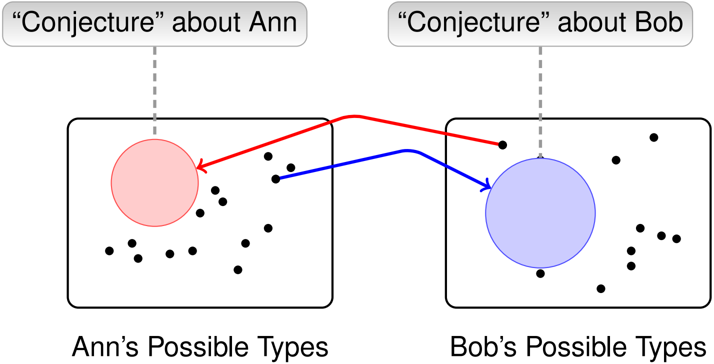
Bob's Possible Types

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.

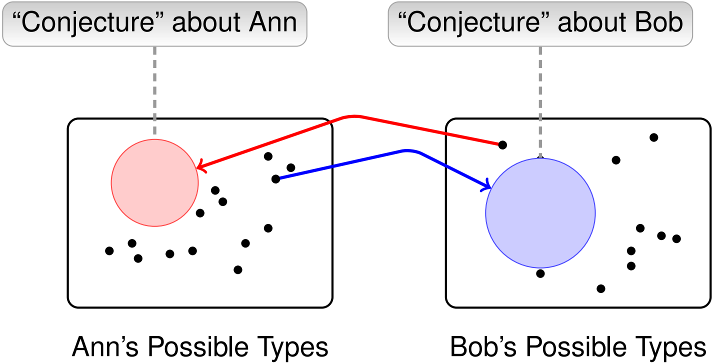
(Harsanyi argued that all uncertainty in a game can be equivalently modeled as uncertainty about payoff functions.)


"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...

Results

Language for i: A set $\mathcal{L}_i \subseteq \wp(T_{-i})$.

Results

Language for i: A set $\mathcal{L}_i \subseteq \wp(T_{-i})$.

Richness Property for $\mathcal{L}_i$: For all $X \in \mathcal{L}_i$, if $X \neq \emptyset$, then there is some type $t \in T_i$ of player i such that X describes t's conjecture about $-i$.

Results

Language for i: A set $\mathcal{L}_i \subseteq \wp(T_{-i})$.

Richness Property for $\mathcal{L}_i$: For all $X \in \mathcal{L}_i$, if $X \neq \emptyset$, then there is some type $t \in T_i$ of player i such that X describes t's conjecture about $-i$.

- $\mathcal{L}_i$ **can't** be the set of all non-empty subsets. (Brandenburger, 2003)
- $\mathcal{L}_i$ **can't** be the set of sets that are definable in first-order logic. (Brandenburger and Keisler, 2006)
- $\mathcal{L}_i$ **can't** be the set of sets definable in a propositional modal logic with an assumption modality. (Brandenburger and Keisler, 2006)
- $\mathcal{L}_i$ **can** be the set of compact subsets of a topological space (Mariotti, Meier and Piccione, 2005)
- $\cdots$

The BK Paradox

Ann believes that Bob's assumption is that Ann believes that Bob's assumption is wrong.

Does Ann believe that Bob's assumption is wrong?

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

The BK Paradox

Ann believes that the strongest proposition that Bob believes is that Ann believes that the strongest proposition that Bob believes is false.

Does Ann believe that the strongest proposition that Bob believes is false?

The BK Paradox

> *Ann believes that the strongest proposition that Bob believes is that Ann believes that the strongest proposition that Bob believes is false.*

Does Ann believe that the strongest proposition that Bob believes is false?

The BK Paradox

> *Ann believes that the strangest proposition that Bob believes is that Ann believes that the strangest proposition that Bob believes is false.*

Does Ann believe that the strangest proposition that Bob believes is false?

The BK Paradox

Ann believes that the most interesting proposition that Bob believes is that Ann believes that the most interesting proposition that Bob believes is false.

Does Ann believe that the most interesting proposition that Bob believes is false?

Suppose that At is a set of atomic propositions, $\mathcal{A}$ is a set of agents, and Lab is a set of "labels" for propositions. The language $\mathcal{L}$:

$$\varphi \quad ::= \quad p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid B_i\varphi \mid \gamma \text{ is } \varphi$$

where $p \in \text{At}$, $i \in \mathcal{A}$, and $\gamma \in \text{Lab}$.

- $B_i\varphi$: "agent i believes that φ"
- γ is φ: "the definite description γ denotes the proposition expressed by φ" (or "the γ-proposition is φ").

The BK Paradox

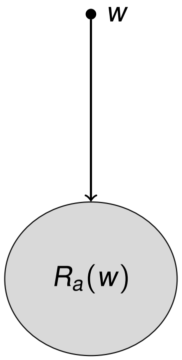
Ann believes that the strongest proposition that Bob believes is that Ann believes that the strongest proposition that Bob believes is wrong.

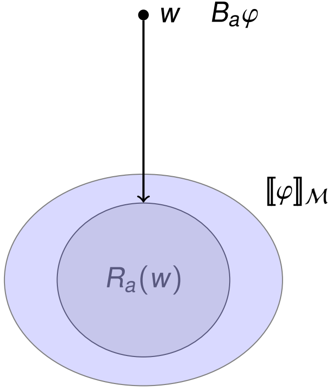
Does Ann believe that the strongest proposition that Bob believes is wrong?

The BK Paradox

Ann believes that the γ-proposition is that Ann believes that the γ-proposition is wrong.

- $B_a(\gamma \text{ is } B_a^{dicto}\mathsf{F}(\gamma))$
- $B_a(\gamma \text{ is } B_a^{re}\mathsf{F}(\gamma))$


w
$R_a(w)$

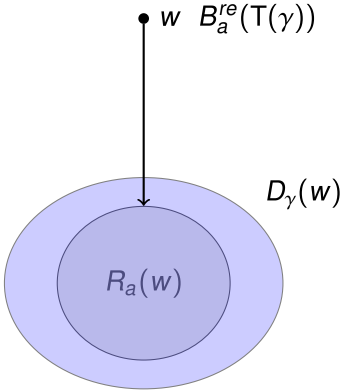

w
$B_a\varphi$
$[\![\varphi]\!]_{\mathcal{M}}$
$R_a(w)$

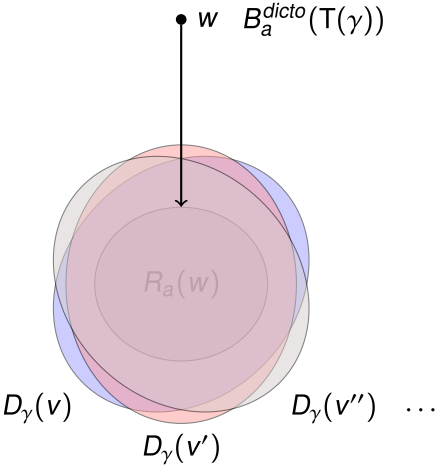
$$B_a(\gamma \text{ is } \varphi)$$

$$B_a(\gamma)$$

$$B_a(\gamma \text{ is } \varphi)$$

$$\cancel{B_a(\gamma)} \qquad B_a(\mathsf{T}(\gamma)), \quad B_a(\mathsf{F}(\gamma))$$


w $B_a^{re}(\mathsf{T}(\gamma))$
$D_\gamma(w)$
$R_a(w)$


w
$B_a^{dicto}(\mathsf{T}(\gamma))$
$R_a(w)$
$D_\gamma(v)$
$D_\gamma(v'')$
$\cdots$
$D_\gamma(v')$

The BK Paradox

Ann believes that the strongest proposition that Bob believes is that Ann believes that the strongest proposition that Bob believes is false.

Does Ann believe that the strongest proposition that Bob believes is false?

The BK Paradox

Ann believes that the γ-proposition is that Ann believes that the γ-proposition is false.

Claim 1. $\{B_a(\gamma \text{ is } B_a^{dicto}\mathsf{F}(\gamma))\}$ is inconsistent in any modal logic containing K, Nec, Cor, I, $S2^{dicto}$

Claim 2. $\{B_a(\gamma \text{ is } B_a^{re}\mathsf{F}(\gamma))\}$ is inconsistent in any modal logic containing K, Nec, Cor, I, $S2^{re}$

The Knower Paradox

Let T be a theory in the language of arithmetic that can prove the Gödel-Carnap fixed-point theorem and K a (perhaps complex) unary predicate in the language of T, such that, for every sentence φ in the language of T, T satisfies:

- $K\varphi \rightarrow \varphi$
- If $T \vdash \varphi$, then $T \vdash K\varphi$

Then, T is inconsistent.

D. Kaplan and R. Montague. *A Paradox Regained*. Notre Dame Journal of Formal Logic, 1, 1960, pp. 79-90.

The Knower Paradox

P. Egré. *The Knower Paradox in the Light of Provability Interpretations of Modal Logic*. Journal of Logic, Language and Information, 14, pgs. 13-48, 2005.

The Knower Paradox

1. $\gamma \leftrightarrow \neg K\ulcorner\gamma\urcorner$ Gödel-Carnap Fixed-Point Lemma

The Knower Paradox

1. $\gamma \leftrightarrow \neg K \ulcorner \gamma \urcorner$ Gödel-Carnap Fixed-Point Lemma
1. $\gamma \leftrightarrow \neg K \gamma$ (Forget Gödel numbering)

The Knower Paradox

1.	$\gamma \leftrightarrow \neg K \ulcorner \gamma \urcorner$	Gödel-Carnap Fixed-Point Lemma
1.	$\gamma \leftrightarrow \neg K\gamma$	(Forget Gödel numbering)
2.	$\gamma \rightarrow \neg K\gamma$	Prop. Reasoning
3.	$K\gamma \rightarrow K\neg K\gamma$	Modal Reasoning
4.	$K\neg K\gamma \rightarrow \neg K\gamma$	T
5.	$K\gamma \rightarrow \neg K\gamma$	Prop. Reasoning
6.	$\neg K\gamma$	Prop. Reasoning

The Knower Paradox

1.	$\gamma \leftrightarrow \neg K \ulcorner \gamma \urcorner$	Gödel-Carnap Fixed-Point Lemma
1.	$\gamma \leftrightarrow \neg K\gamma$	(Forget Gödel numbering)
2.	$\gamma \rightarrow \neg K\gamma$	Prop. Reasoning
3.	$K\gamma \rightarrow K\neg K\gamma$	Modal Reasoning
4.	$K\neg K\gamma \rightarrow \neg K\gamma$	T
5.	$K\gamma \rightarrow \neg K\gamma$	Prop. Reasoning
6.	$\neg K\gamma$	Prop. Reasoning
7.	$\neg K\gamma \rightarrow \gamma$	Prop. Reasoning
8.	γ	Prop. Reasoning
9.	$K\gamma$	Nec
	$\lightning$	

Let T be a theory in the language of arithmetic that can prove the Gödel-Carnap fixed-point theorem and B a (perhaps complex) unary predicate in the language of T, such that, for every sentence φ and ψ in the language of T, T satisfies:

- $B\neg B\varphi \rightarrow \neg B\varphi$
- If $T \vdash \varphi$, then $T \vdash B\varphi$
- If $T \vdash \varphi \leftrightarrow \psi$, then $T \vdash B\varphi \leftrightarrow B\psi$

then T is inconsistent.

R. Thomason. *A note on syntactical treatments of modality*. Synthese 44, pgs. 391 - 395, 1980.

1.	$\gamma \leftrightarrow \neg K \ulcorner \gamma \urcorner$	Gödel-Carnap Fixed-Point Lemma
1.	$\gamma \leftrightarrow \neg K\gamma$	(Forget Gödel numbering)
2.	$\gamma \rightarrow \neg K\gamma$	Prop. Reasoning
3.	$K\gamma \rightarrow K\neg K\gamma$	Modal Reasoning
4.	$K\neg K\gamma \rightarrow \neg K\gamma$	T
5.	$K\gamma \rightarrow \neg K\gamma$	Prop. Reasoning
6.	$\neg K\gamma$	Prop. Reasoning
7.	$\neg K\gamma \rightarrow \gamma$	Prop. Reasoning
8.	γ	Prop. Reasoning
9.	$K\gamma$	Nec
	$\lightning$	

1.	$\gamma \leftrightarrow \neg B\ulcorner\gamma\urcorner$	Gödel-Carnap Fixed-Point Lemma
1.	$\gamma \leftrightarrow \neg B\gamma$	(Forget Gödel numbering)
2.	$\gamma \rightarrow \neg B\gamma$	Prop. Reasoning
3.	$B\gamma \rightarrow B\neg B\gamma$	Modal Reasoning
4.	$B\neg B\gamma \rightarrow \neg B\gamma$	T
5.	$B\gamma \rightarrow \neg B\gamma$	Prop. Reasoning
6.	$\neg B\gamma$	Prop. Reasoning
7.	$\neg B\gamma \rightarrow \gamma$	Prop. Reasoning
8.	γ	Prop. Reasoning
9.	$B\gamma$	Nec
	$\lightning$	

1.	$\gamma \leftrightarrow \neg B\ulcorner\gamma\urcorner$	Gödel-Carnap Fixed-Point Lemma
1.	$\gamma \leftrightarrow \neg B\gamma$	(Forget Gödel numbering)
2.	$\gamma \rightarrow \neg B\gamma$	Prop. Reasoning
3.	$B\gamma \rightarrow B\neg B\gamma$	Modal Reasoning
4.	$B\neg B\gamma \rightarrow \neg B\gamma$	T
5.	$B\gamma \rightarrow \neg B\gamma$	Prop. Reasoning
6.	$\neg B\gamma$	Prop. Reasoning
7.	$\neg B\gamma \rightarrow \gamma$	Prop. Reasoning
8.	γ	Prop. Reasoning
9.	$B\gamma$	Nec
	$\lightning$	

1.	$\gamma \leftrightarrow \neg B\ulcorner\gamma\urcorner$	Gödel-Carnap Fixed-Point Lemma
1.	$\gamma \leftrightarrow \neg B\gamma$	(Forget Gödel numbering)
2.	$\gamma \rightarrow \neg B\gamma$	Prop. Reasoning
3.	$B\gamma \rightarrow B\neg B\gamma$	Modal Reasoning
4.	$B\neg B\gamma \rightarrow \neg B\gamma$	Corr_N
5.	$B\gamma \rightarrow \neg B\gamma$	Prop. Reasoning
6.	$\neg B\gamma$	Prop. Reasoning
7.	$\neg B\gamma \rightarrow \gamma$	Prop. Reasoning
8.	γ	Prop. Reasoning
9.	$B\gamma$	Nec
	$\lightning$	

Proposition. $\{B_a(\gamma \text{ is } B_a^{dicto}\mathsf{F}(\gamma))\}$ is inconsistent in any modal logic containing K, Nec, Cor_N, I_N, $S2^{dicto}$.

1.	$B_i(\gamma \text{ is } B_i^{dicto}\mathsf{F}(\gamma))$	(assumption)
2.	$B_i(\gamma \text{ is } B_i^{dicto}\mathsf{F}(\gamma)) \rightarrow$ $(B_i^{dicto}\mathsf{F}(\gamma) \leftrightarrow B_i(\neg B_i^{dicto}\mathsf{F}(\gamma)))$	($S2^{dicto}$)
3.	$B_i^{dicto}\mathsf{F}(\gamma) \leftrightarrow B_i(\neg B_i^{dicto}\mathsf{F}(\gamma))$	(MP, 1, 2)
	$\vdots$	
10.	Contradiction	

Proposition. $\{B_a(\gamma \text{ is } B_a^{re}\mathsf{F}(\gamma))\}$ is inconsistent in any modal logic containing *K*, *Nec*, *Cor*, *I*, $S2^{re}$.

1.	$B_i(\gamma \text{ is } B_i^{re}\mathsf{F}(\gamma))$	(assumption)
2.	$(\gamma \text{ is } B_i^{re}\mathsf{F}(\gamma)) \rightarrow$ $(B_i^{re}\mathsf{F}(\gamma) \leftrightarrow B_i(\neg B_i^{re}\mathsf{F}(\gamma)))$	($S2^{re}$)
3.	$B_i(\gamma \text{ is } B_i^{re}\mathsf{F}(\gamma)) \rightarrow$ $B_i(B_i^{re}\mathsf{F}(\gamma) \leftrightarrow B_i(\neg B_i^{re}\mathsf{F}(\gamma)))$	(Mon, 2)
	$\vdots$	
22.	Contradiction	

Taking Stock

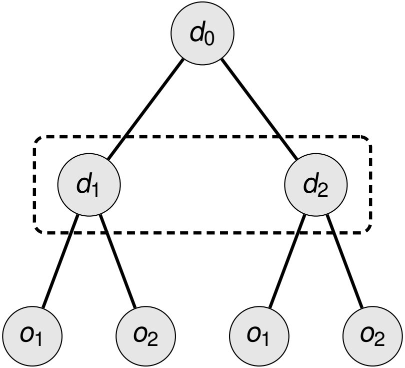
- Propositional modal logic with definition descriptions for propositions.
- On this analysis, the BK Paradox is **not** a paradox of *interactive* beliefs.
- The proof of the BK Paradox is similar to the proof of the Knower Paradox.

Knower	BB/AE	BK
γ is $\neg B_i\mathrm{T}(\gamma)$	$p \leftrightarrow \neg B_i p$	$B_i(\gamma$ is $\neg B_i\mathrm{T}(\gamma))$
γ is $B_i\mathrm{F}(\gamma)$	$p \leftrightarrow B_i \neg p$	$B_i(\gamma$ is $B_i\mathrm{F}(\gamma))$

γ is $B_i\mathsf{F}(\gamma)$	$\lightning$	$B_i(\gamma$ is $B_i\mathsf{F}(\gamma))$	$\lightning$
$p \leftrightarrow B_i\neg p$		$B_i(p \leftrightarrow B_i\neg p)$	$\lightning$

The Absent-Minded Driver

Games of Imperfect Information



The Absent-Minded Driver

An individual is sitting late at night in a bar planning his midnight trip home. In order to get home he has to take the highway and get off at the second exit.

The Absent-Minded Driver

An individual is sitting late at night in a bar planning his midnight trip home. In order to get home he has to take the highway and get off at the second exit. Turning at the first exit leads into a disastrous area (payoff 0). Turning at the second exit yields the highest reward (payoff 4).

The Absent-Minded Driver

An individual is sitting late at night in a bar planning his midnight trip home. In order to get home he has to take the highway and get off at the second exit. Turning at the first exit leads into a disastrous area (payoff 0). Turning at the second exit yields the highest reward (payoff 4). If he continues beyond the second exit, he cannot go back and at the end of the highway he will find a motel where he can spend the night (payoff 1).

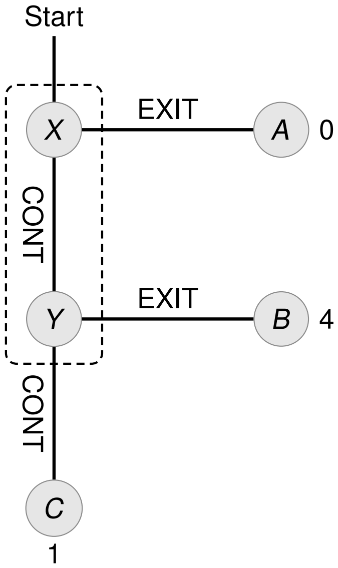
The Absent-Minded Driver

The driver is absentminded and is aware of this fact. At an intersection, he cannot tell whether it is the first or the second intersection and he cannot remember how many he has passed (one can make the situation more realistic by referring to the 17th intersection).

The Absent-Minded Driver

The driver is absentminded and is aware of this fact. At an intersection, he cannot tell whether it is the first or the second intersection and he cannot remember how many he has passed (one can make the situation more realistic by referring to the 17th intersection). While sitting at the bar, all he can do is to decide whether or not to exit at an intersection. (pg. 7)

M. Piccione and A. Rubinstein. *On the Interpretation of Decision Problems with Imperfect Recall*. Games and Econ Behavior, 20, pgs. 3- 24, 1997.


Start
X
EXIT
A
0
CONT
Y
EXIT
B
4
CONT
C
1

Planning stage: While planning his trip home at the bar, the decision maker is faced with a choice between “Continue; Continue” and “Exit”. Since he cannot distinguish between the two intersections, he cannot plan to “Exit” at the second intersection (he must plan the same behavior at both X and Y). Since “Exit” will lead to the worst outcome (with a payoff of 0), the optimal strategy is “Continue; Continue” with a guaranteed payoff of 1.

Action stage: When arriving at an intersection, the decision maker is faced with a local choice of either “Exit” or “Continue” (possibly followed by another decision). Now the decision maker knows that since he committed to the plan of choosing “Continue” at each intersection, it is possible that he is at the second intersection. Indeed, the decision maker concludes that he is at the first intersection with probability 1/2. But then, his expected payoff for “Exit” is 2, which is greater than the payoff guaranteed by following the strategy he previously committed to. Thus, he chooses to “Exit”.

W. Schwarz. *Lost Memories and Useless Coins: Revisiting the Absentminded Driver*. Synthese 192, 3011-3036, 2015.

Thank you!