

Logic, Interaction and Collective Agency

Lecture 5

ESSLLI'10, Copenhagen

Eric Pacuit

Olivier Roy

TiLPS, Tilburg University
`ai.stanford.edu/~epacuit`

University of Groningen
`philos.rug.nl/~olivier`

August 20, 2010

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		
\bullet, M, MH		
\vdots		
\bullet, H, MM		
\bullet, H, MH		
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		
\bullet, M, MH		
\vdots		
\bullet, H, MM		
\bullet, H, MH		
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		
\bullet, M, MH		
\vdots		
\bullet, H, MM		
\bullet, H, MH		
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		0
\bullet, M, MH		
\vdots		
\bullet, H, MM		
\bullet, H, MH		
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		0
\bullet, M, MH		0
\vdots		
\bullet, H, MM		
\bullet, H, MH		
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		0
\bullet, M, MH		0
\vdots		
\bullet, H, MM	0	
\bullet, H, MH		
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O
\bullet, M, MM		0
\bullet, M, MH		0
\vdots		
\bullet, H, MM	0	
\bullet, H, MH	0	
\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O	$\lambda_E(I_E)(C_O)$	M_O	H_O
\bullet, M, MM		0	\bullet, M, MM		
\bullet, M, MH		0	\bullet, M, MH		
\vdots			\vdots		
\bullet, H, MM	0		\bullet, H, MM		
\bullet, H, MH	0		\bullet, H, MH		
\vdots			\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O	$\lambda_E(I_E)(C_O)$	M_O	H_O
\bullet, M, MM		0	\bullet, M, MM		0
\bullet, M, MH		0	\bullet, M, MH		
\vdots			\vdots		
\bullet, H, MM	0		\bullet, H, MM		
\bullet, H, MH	0		\bullet, H, MH		
\vdots			\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O	$\lambda_E(I_E)(C_O)$	M_O	H_O
\bullet, M, MM		0	\bullet, M, MM		0
\bullet, M, MH		0	\bullet, M, MH	0	
\vdots			\vdots		
\bullet, H, MM	0		\bullet, H, MM		
\bullet, H, MH	0		\bullet, H, MH		
\vdots			\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O	$\lambda_E(I_E)(C_O)$	M_O	H_O
\bullet, M, MM		0	\bullet, M, MM		0
\bullet, M, MH		0	\bullet, M, MH	0	
\vdots			\vdots		
\bullet, H, MM	0		\bullet, H, MM		0
\bullet, H, MH	0		\bullet, H, MH		
\vdots			\vdots		

Correlations in UTI

Correlations arise from the **Team Authority** requirement:

$$\text{If } t_i = k \text{ then } \alpha_i = \alpha_i^k(s)$$

(Part of) an ex-interim type space for this TI:

$\lambda_E(I_E)(I_O)$	M_O	H_O	$\lambda_E(I_E)(C_O)$	M_O	H_O
\bullet, M, MM		0	\bullet, M, MM		0
\bullet, M, MH		0	\bullet, M, MH	0	
\vdots			\vdots		
\bullet, H, MM	0		\bullet, H, MM		0
\bullet, H, MH	0		\bullet, H, MH	0	
\vdots			\vdots		

Plan

1. Prior Beliefs and the Common Prior Assumption.
2. Correlations: strategies, beliefs, quantitatively, qualitatively.
3. General Discussion on Commitments and Intentions.
4. Wrap up.

Prior Beliefs and the Common Prior Assumption

Prior and posterior beliefs

- ▶ Different stages of information disclosure:

Prior and posterior beliefs

- ▶ Different stages of information disclosure:
 1. *Ex ante*: *A priori*. No information. Possibly background beliefs about who you are playing against (background beliefs about the “context of interaction”).

Prior and posterior beliefs

► Different stages of information disclosure:

1. *Ex ante*: *A priori*. No information. Possibly background beliefs about who you are playing against (background beliefs about the “context of interaction”).
2. *Ex interim*: the players receive a *piece of (private) information*. They know more.

Prior and posterior beliefs

► Different stages of information disclosure:

1. *Ex ante*: *A priori*. No information. Possibly background beliefs about who you are playing against (background beliefs about the “context of interaction”).

Prior beliefs.

2. *Ex interim*: the players receive a piece of (private) information. They know more.

Posterior beliefs.

- ▶ We encountered priors many time already:
 - In type spaces, in UTIs, in Bayesian games...

- ▶ We encountered priors many time already:
 - In **type spaces**, in UTIs, in Bayesian games...

$$\lambda_i(t_i)(s, t) = \frac{\Omega((s, t) \ \& \ t_i)}{\Omega(t_i)}$$

- ▶ We encountered priors many time already:
 - In type spaces, in UTIs, in Bayesian games...
 - We mostly used **common** priors:
 - ▶ $\Omega_i = \Omega_j$ for all $i, j \in \mathcal{A}$.
 - ▶ $\preceq_i = \preceq_j$ for all $i, j \in \mathcal{A}$ (in epistemic-plausibility models)

Common priors

- ▶ \nRightarrow same posteriors!

Common priors

- ▶ \nRightarrow same posteriors! For the simple reason that agents can receive **different private information**.

Common priors

- ▶ \nrightarrow same posteriors! For the simple reason that agents can receive **different private information**.
- ▶ Olivier and I play card together.



Common priors

- ▶ \nrightarrow same posteriors! For the simple reason that agents can receive **different private information**.
- ▶ Olivier and I play card together.
Before the cards are dealt, our common prior belief that the other end up with a Joker is $0.37 = 2/54$.



Common priors

- ▶ \nrightarrow same posteriors! For the simple reason that agents can receive **different private information**.
- ▶ Olivier and I play card together. Before the cards are dealt, our common prior belief that the other end up with a Joker is $0.37 = 2/54$.
- ▶ We get 5 card each (and don't show them to each other). I end up with the 2 Jokers.
 - My posterior belief that Olivier has a Joker is 0.
 - Olivier's posterior belief that I have a Joker is $0.4 = 2/49$.



Common Prior Assumption

- Differences in posterior beliefs should be seen as coming from different information, not from different priors.

Common Prior Assumption

- ▶ Differences in posterior beliefs should be seen as coming from different information, not from different priors.
- ▶ Alternative formulation: “...people with different information may legitimately entertain different probabilities, but there is no rational basis for people who have always been fed precisely the same information to do so.” (...Given that we are Bayesians)

Common Prior Assumption

- ▶ Differences in posterior beliefs should be seen as coming from different information, not from different priors.
- ▶ Alternative formulation: “...people with different information may legitimately entertain different probabilities, but there is no rational basis for people who have always been fed precisely the same information to do so.” (...Given that we are Bayesians)

“Harsanyi doctrine” [Aumann, 1976].

R. Aumann. *Agreeing to Disagree*. *Annals of Statistics*, Vol.4, No.6, 1976.

J.C. Harsanyi. *Games with incomplete informations played by Bayesian players*. *Management Science* 14:159, pgs. 182, 320 - 334, 486 -502, 1967-68.

Common Prior Assumption

- ▶ Differences in posterior beliefs should be seen as coming from different information, not from different priors.
- ▶ Alternative formulation: “...people with different information may legitimately entertain different probabilities, but there is no rational basis for people who have always been fed precisely the same information to do so.” (...Given that we are Bayesians)

“Harsanyi doctrine” [Aumann, 1976].

- ▶ CPA not an innocuous assumption!

▶ Explanation

R. Aumann. *Agreeing to Disagree*. *Annals of Statistics*, Vol.4, No.6, 1976.

J.C. Harsanyi. *Games with incomplete informations played by Bayesian players*. *Management Science* 14:159, pgs. 182, 320 - 334, 486 -502, 1967-68.

Why CPA?

Why CPA?

- ▶ A controversial question!
 - See Morris (1995) for a thorough discussion.

Why CPA?

- ▶ A controversial question!
 - See Morris (1995) for a thorough discussion. One **methodological** observation:

Why CPA?

- ▶ A controversial question!
 - See Morris (1995) for a thorough discussion. One **methodological** observation:
 - ▶ Better to explain differences in posterior on the basis of **identifiable differences** in information or **plausible errors** in information processing.

Why CPA?

- ▶ A controversial question!
 - See Morris (1995) for a thorough discussion. One **methodological** observation:
 - ▶ Better to explain differences in posterior on the basis of **identifiable differences** in information or **plausible errors** in information processing.
 - ▶ Resorting on differences in priors often appears **ad hoc** (the resulting theory is “too permissive”).

S. Morris. *The Common Prior Assumption in Economic Theory*. *Economics and Philosophy*, 11(2): pgs. 227- 253, 1995.

Common Prior and Shared Perspective

- ▶ Common prior is already a form of **shared background** within a group, although a **relatively weak one**.
 - RCBR characterizes non-dominance; but RCBR + (non-correlated) CP characterizes Nash equilibrium.

Common Prior and Shared Perspective

- ▶ Common prior is already a form of **shared background** within a group, although a **relatively weak one**.
 - RCBR characterizes non-dominance; but RCBR + (non-correlated) CP characterizes Nash equilibrium.
- ▶ But the connection between group members can be strengthened if the priors are **correlated**.

Correlations

Correlated Strategies

	H	M
H	3, 3	0, 4
M	4, 0	1, 1

Correlated Strategies

	H	M
H	3, 3	0, 4
M	4, 0	1, 1

We each can make our choice dependent on a flip of the coin:

Correlated Strategies

	H	M
H	3, 3	0, 4
M	4, 0	1, 1

We each can make our choice dependent on a flip of the coin:

- ▶ A mixed strategy is a random variable assigning actions to the outcome of a coin toss.
- ▶ The coins are *independent*: (H, H) is never rational.

Correlated Strategies

	H	M
H	3, 3	0, 4
M	4, 0	1, 1

Suppose we base our choices on the *same coin*:

Correlated Strategies

	H	M
H	3, 3	0, 4
M	4, 0	1, 1

Suppose we base our choices on the *same coin*:

- ▶ $\rho_E = (\text{Head} \Rightarrow H, \text{Tail} \Rightarrow M)$
- ▶ $\rho_O = (\text{Head} \Rightarrow H, \text{Tail} \Rightarrow M)$

Our choice are **correlated**, i.e. **not independent**

(In general, the players need not observe the *same* random event.)

Correlated Strategies

	H	M
H	3, 3	0, 4
M	4, 0	1, 1

	H	M
H	$1/2$	0
M	0	$1/2$

Convenient representation of correlated strategies in terms of the distribution of the external signal.

Correlated Equilibrium

Let \mathbb{G} be a game in strategic form and W a (measurable) set of **outside signals** and $\Omega(S)$ a probability distribution over S .

- ▶ A **correlated strategy** for player i is a function $\rho_i : W \rightarrow S_i$ assigning to each signal an action in S_i .
- ▶ The **expected value** for player i of the profile ρ :

$$EV_i(\rho) = \sum_{w \in W} \Omega(w) v_i(\rho(w))$$

- ▶ A **correlated equilibrium** is a profile ρ_i of correlated strategies such that, for all i :

$$\rho_i \in \operatorname{argmax}_{\rho'_i} (v_i(\rho'_i, \rho_{-i}))$$

R.J. Aumann. *Correlated equilibrium as an expression of bayesian rationality*. *Econometrica*, 55(1-18), 1987.

Nash equilibrium \subsetneq Correlated Equilibrium

	H	M
H	3, 3	0, 0
M	0, 0	1, 1

Nash equilibrium \subsetneq Correlated Equilibrium

▶

	H	M
H	6, 6	2, 7
M	7, 2	0, 0

Nash equilibrium \subsetneq Correlated Equilibrium

►

	H	M
H	6, 6	2, 7
M	7, 2	0, 0

- Three Nash equilibria: (M, H), (H, M) and $((2/3 \text{ M}, 1/3 \text{ H}), (2/3 \text{ M}, 1/3 \text{ H}))$.

Nash equilibrium \subsetneq Correlated Equilibrium

►

	H	M
H	6, 6	2, 7
M	7, 2	0, 0

	H	M
H	Blue Card	Red card
M	Green Card	

- Now suppose we make our choice dependent on drawing one of three cards.

Nash equilibrium \subsetneq Correlated Equilibrium

►

	H	M
H	6, 6	2, 7
M	7, 2	0, 0

	H	M
H	1/3	1/3
M	1/3	

- Now suppose we make our choice dependent on drawing one of three cards.

Nash equilibrium \subsetneq Correlated Equilibrium

▶

	H	M
H	6, 6	2, 7
M	7, 2	0, 0

	H	M
H	1/3	1/3
M	1/3	

- ▶ Now suppose we make our choice dependent on drawing one of three cards.
- ▶ This is a **correlated equilibrium**, with **expected value of 5** for each of us.

Nash equilibrium \subsetneq Correlated Equilibrium

▶

	H	M
H	6, 6	2, 7
M	7, 2	0, 0

	H	M
H	1/3	1/3
M	1/3	

- ▶ Now suppose we make our choice dependent on drawing one of three cards.
- ▶ This is a **correlated equilibrium**, with **expected value of 5** for each of us.
- ▶ No Nash equilibrium can give you that.

Nash equilibrium \subsetneq Correlated Equilibrium

	H	M		H	M
H	6, 6	2, 7	H	1/3	1/3
M	7, 2	0, 0	M	1/3	

- ▶ Now suppose we make our choice dependent on drawing one of three cards.
- ▶ This is a **correlated equilibrium**, with **expected value of 5** for each of us.
- ▶ **No mixing** can give you the probability distribution ρ on $\Pi_i S_i$.

Nash equilibrium \subsetneq Correlated Equilibrium

	H	M		H	M
H	6, 6	2, 7	H	1/3	1/3
M	7, 2	0, 0	M	1/3	

- ▶ Now suppose we make our choice dependent on drawing one of three cards.
- ▶ This is a **correlated equilibrium**, with **expected value of 5** for each of us.
- ▶ The probability distribution ρ on $\Pi_i S_i$ **cannot be factorized** in independent probability distributions on S_i 's.

Correlations in Belief

- Small step from correlated strategies to correlated priors.

	H	M
H	3, 3	0, 0
M	0, 0	1, 1

Correlations in Belief

- Small step from correlated strategies to correlated priors.

	H	M	Ω	M	H
H	3, 3	0, 0	H	1/2	0
M	0, 0	1, 1	M	0	1/2

- $\mathcal{T} = \{t_O, t_E\}$.

Correlations in Belief

- Small step from correlated strategies to correlated priors.

	H	M	Ω	M	H
H	3, 3	0, 0	H	1/2	0
M	0, 0	1, 1	M	0	1/2

- $\mathcal{T} = \{t_O, t_E\}$.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U		1, 1, 0	1, 0, 0
			D		0, 1, 3	0, 0, 3

- Three players: Ann, Bob and Charlie.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U	1, 1, 0	1, 0, 0	
			D	0, 1, 3	0, 0, 3	

- Ann chooses the column.

Beliefs about correlations

X	L	R	Y	L	R
U	1, 1, 3	1, 0, 3	U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0	D	0, 0, 0	1, 1, 2

Z	L	R
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

- Ann chooses the **column**.

Beliefs about correlations

X	L	R	Y	L	R
U	1, 1, 3	1, 0, 3	U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0	D	0, 0, 0	1, 1, 2

Z	L	R
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

- Bob chooses the row.

Beliefs about correlations

X	L	R
U	1, 1, 3	1, 0, 3
D	0, 1, 0	0, 0, 0

Y	L	R
U	1, 1, 2	0, 0, 0
D	0, 0, 0	1, 1, 2

Z	L	R
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

- Bob chooses the row.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U		1, 1, 0	1, 0, 0
			D		0, 1, 3	0, 0, 3

- Charlie chooses the **matrix**.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U			
			D			

- Charlie chooses the **matrix**.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z	L	R	
			U	1, 1, 0	1, 0, 0	
			D	0, 1, 3	0, 0, 3	

- Charlie chooses the **matrix**.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U		1, 1, 0	1, 0, 0
			D		0, 1, 3	0, 0, 3

- Strategy **Y** is **rationalizable** for Charlie.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z		L	R
			U		1, 1, 0	1, 0, 0
			D		0, 1, 3	0, 0, 3

- Strategy **Y** is **rationalizable** for Charlie.
- Take: $\lambda(t_C)(U, L) = 1/2$, $\lambda(t_C)(D, R) = 1/2$.

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U		1, 1, 0	1, 0, 0
			D		0, 1, 3	0, 0, 3

- Strategy **Y** is **rationalizable** for Charlie.
- Take: $\lambda(t_C)(U, L) = 1/2$, $\lambda(t_C)(D, R) = 1/2$. Then:

$$EV_C(X, t_C) = 1/2(3) + 1/2(0) = 1.5$$

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2

Z	L	R
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

- Strategy **Y** is **rationalizable** for Charlie.
- Take: $\lambda(t_C)(U, L) = 1/2$, $\lambda(t_C)(D, R) = 1/2$. Then:

$$EV_C(X, t_C) = 1/2(3) + 1/2(0) = 1.5$$

$$EV_C(Y, t_C) = 1/2(2) + 1/2(2) = 2$$

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z	L	R	
			U	1, 1, 0	1, 0, 0	
			D	0, 1, 3	0, 0, 3	

- Strategy **Y** is **rationalizable** for Charlie.
- Take: $\lambda(t_C)(U, L) = 1/2$, $\lambda(t_C)(D, R) = 1/2$. Then:

$$EV_C(X, t_C) = 1/2(3) + 1/2(0) = 1.5$$

$$EV_C(Y, t_C) = 1/2(2) + 1/2(2) = 2$$

$$EV_C(Z, t_C) = 1/2(0) + 1/2(3) = 1.5$$

Beliefs about correlations

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z			
			L			
			R			
			U		1, 1, 0	1, 0, 0
			D		0, 1, 3	0, 0, 3

- ▶ Strategy **Y** is **rationalizable** for Charlie.
- ▶ Take: $\lambda(t_C)(U, L) = 1/2$, $\lambda(t_C)(D, R) = 1/2$.
- ▶ **Y** is not rationalizable if Charlie doesn't believe that Ann and Bob correlate their choices. (Left as an exercise).

The Importance of Correlations

The Importance of Correlations

- ▶ RCBR \Leftrightarrow Playing iteratively rationalizable strategies.

The Importance of Correlations

- ▶ RCBR \Leftrightarrow Playing iteratively rationalizable strategies.
 - For more than two agents: some (beliefs in) correlations are bound to arise.
 - We just saw an example of that.

The Importance of Correlations

- ▶ RCBR \Leftrightarrow Playing iteratively rationalizable strategies.
 - For more than two agents: some (beliefs in) correlations are bound to arise.
 - We just saw an example of that.
 - Team Interactions generate correlations.

- ▶ Question: where do (beliefs in) correlations come from?

The Importance of Correlations

- ▶ RCBR \Leftrightarrow Playing iteratively rationalizable strategies.
 - For more than two agents: some (beliefs in) correlations are bound to arise.
 - We just saw an example of that.
 - Team Interactions generate correlations.

- ▶ Question: where do (beliefs in) correlations come from?
Answer: A player can think that other players' strategy choices are correlated, because he thinks what they believe about the game is correlated.

Correlations through types

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2

Z	L	R
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

$$\lambda(t_C)(U, L) = 1/2$$

$$\lambda(t_C)(D, R) = 1/2$$

Correlations through types

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2
			Z	L	R	
			U	1, 1, 0	1, 0, 0	
			D	0, 1, 3	0, 0, 3	

- Instead: make strategy choices conditional on types/beliefs/hierarchies.

Correlations through types

X	L	R		Y	L	R
U	1, 1, 3	1, 0, 3		U	1, 1, 2	0, 0, 0
D	0, 1, 0	0, 0, 0		D	0, 0, 0	1, 1, 2

Z	L	R
U	1, 1, 0	1, 0, 0
D	0, 1, 3	0, 0, 3

- Instead: make strategy choices conditional on types/beliefs/hierarchies.

- Suppose $T_A = \{t_A, u_A\}$ and $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

Correlations through types

t_C	UL	UR	DL	DR
$t_A t_B$	1/2	0	0	0
$t_A u_B$	0	0	0	0
$u_A t_B$	0	0	0	0
$u_A u_B$	0	0	0	1/2

- Instead: make strategy choices conditional on types/beliefs/hierarchies.

- Suppose $T_A = \{t_A, u_A\}$ and $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

- We get the same result: Charlie is rational at state

$$(t_C, \bullet), (s_A, s_B, Y)$$

Correlations through types

- ▶ Strategy choices conditional on signals.
 - $\rho_A = \{Head \Rightarrow U, Tail \Rightarrow D\}$, $\rho_B = \{Head \Rightarrow L, Tail \Rightarrow R\}$.
- ▶ Strategy choices conditional on types/beliefs/hierarchies.
 - $T_A = \{t_A, u_A\}$, $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

Correlations through types

- ▶ Strategy choices conditional on **signals**.
 - $\rho_A = \{\text{Head} \Rightarrow U, \text{Tail} \Rightarrow D\}, \rho_B = \{\text{Head} \Rightarrow L, \text{Tail} \Rightarrow R\}.$
- ▶ Strategy choices conditional on types/beliefs/hierarchies.
 - $T_A = \{t_A, u_A\}, T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

Correlations through types

- ▶ Strategy choices conditional on **signals**.
 - $\rho_A = \{Head \Rightarrow U, Tail \Rightarrow D\}$, $\rho_B = \{Head \Rightarrow L, Tail \Rightarrow R\}$.
- ▶ Strategy choices conditional on types/beliefs/hierarchies.
 - $T_A = \{t_A, u_A\}$, $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

Correlations through types

- ▶ Strategy choices conditional on signals.
 - $\rho_A = \{Head \Rightarrow U, Tail \Rightarrow D\}$, $\rho_B = \{Head \Rightarrow L, Tail \Rightarrow R\}$.
- ▶ Strategy choices conditional on **types/beliefs/hierarchies**.
 - $T_A = \{t_A, u_A\}$, $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(\mathbf{t}_A, \mathbf{t}_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

Correlations through types

- ▶ Strategy choices conditional on signals.
 - $\rho_A = \{Head \Rightarrow U, Tail \Rightarrow D\}$, $\rho_B = \{Head \Rightarrow L, Tail \Rightarrow R\}$.
- ▶ Strategy choices conditional on **types/beliefs/hierarchies**.
 - $T_A = \{t_A, u_A\}$, $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

Correlations through types

- ▶ Strategy choices conditional on signals.
 - $\rho_A = \{Head \Rightarrow U, Tail \Rightarrow D\}$, $\rho_B = \{Head \Rightarrow L, Tail \Rightarrow R\}$.
- ▶ Strategy choices conditional on types/beliefs/hierarchies.
 - $T_A = \{t_A, u_A\}$, $T_B = \{t_B, u_B\}$ and:

$$\lambda(t_C)(t_A, t_B, U, L) = 1/2 = \lambda(t_C)(u_A, u_B, D, R)$$

- ▶ How much explanatory power?

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any **type** t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are **independent (i.e. non-correlated)**, conditional on the type of the other players.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), **conditional on the type of the other players**.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

- In the example:

$$\begin{aligned} \lambda(t_C)(U, L|t_A, t_B) &= \lambda(t_C)(U|t_A, t_B) \times \lambda(t_C)(L|t_A, t_B) \\ &= \times \end{aligned}$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

- In the example:

$$\begin{aligned} \lambda(t_C)(U, L|t_A, t_B) &= \lambda(t_C)(U|t_A, t_B) \times \lambda(t_C)(L|t_A, t_B) \\ &= \quad \times \end{aligned}$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

- In the example:

$$\begin{aligned} \lambda(t_C)(U, L|t_A, t_B) &= \lambda(t_C)(U|t_A, t_B) \times \lambda(t_C)(L|t_A, t_B) \\ &= \times \end{aligned}$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

- In the example:

$$\begin{array}{rclcl} \lambda(t_C)(U, L|t_A, t_B) & = & \lambda(t_C)(U|t_A, t_B) & \times & \lambda(t_C)(L|t_A, t_B) \\ 1 & = & & \times & \end{array}$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

- In the example:

$$\frac{\lambda(t_C)(U, L|t_A, t_B)}{1} = \frac{\lambda(t_C)(U|t_A, t_B)}{1} \times \frac{\lambda(t_C)(L|t_A, t_B)}{1}$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **CI**: For any type t_i for agent i , the choices (s_j, s_k, \dots) for the other agents ($i \neq j, \neq k, \dots$) are independent (i.e. non-correlated), conditional on the type of the other players.

- In the example:

$$\frac{\lambda(t_C)(U, L|t_A, t_B)}{1} = \frac{\lambda(t_C)(U|t_A, t_B)}{1} \times \frac{\lambda(t_C)(L|t_A, t_B)}{1}$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is **sufficient**, to determine t_i 's beliefs about j 's choice.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- ▶ **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.
- ▶ In the example:

$$\lambda(t_C)(U|t_A, t_B) = \lambda(t_C)(U|t_A)$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.
- In the example:

$$\lambda(t_C)(U|t_A, t_B) = \lambda(t_C)(U|t_A)$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- ▶ **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.
- ▶ In the example:

$$\lambda(t_C)(U|t_A, t_B) = \lambda(t_C)(U|t_A)$$

$$1 =$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

- **SUFF**: For any type t_i for agent i , a type t_j of agent $(j \neq i)$ is sufficient, to determine t_i 's beliefs about j 's choice.
- In the example:

$$\lambda(t_C)(U|t_A, t_B) = \lambda(t_C)(U|t_A)$$

$$1 = 1$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations through types

Theorem If Conditional Independence (CI) and Sufficient (SUFF) hold and type t_i has independent assessments about the types of players j and k , then t_i must have an independent assessment about j 's and k 's strategy choice.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

There are some conceptual issues...

► Explanation

Correlations through types

Theorem If Conditional Independence (CI) and Sufficient (SUFF) hold and type t_i has independent assessments about the types of players j and k , then t_i must have an independent assessment about j 's and k 's strategy choice.

CI & SUFF \Rightarrow

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

There are some conceptual issues...

► Explanation

Correlations through types

Theorem If Conditional Independence (CI) and Sufficient (SUFF) hold and type t_i has independent assessments about the types of players j and k , then t_i must have an independent assessment about j 's and k 's strategy choice.

$$\text{CI \& SUFF} \Rightarrow \left(\text{Ind Bel about types} \Rightarrow \right.$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

There are some conceptual issues...

► Explanation

Correlations through types

Theorem If Conditional Independence (CI) and Sufficient (SUFF) hold and type t_i has independent assessments about the types of players j and k , then t_i must have an independent assessment about j 's and k 's strategy choice.

$$\text{CI \& SUFF} \Rightarrow \left(\text{Ind Bel about types} \Rightarrow \text{Ind Bel about choices} \right)$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

There are some conceptual issues...

► Explanation

Correlations through types

Theorem If Conditional Independence (CI) and Sufficient (SUFF) hold and type t_i has independent assessments about the types of players j and k , then t_i must have an independent assessment about j 's and k 's strategy choice.

$$\text{CI \& SUFF} \Rightarrow \left(\text{Corr Bel about choices} \Rightarrow \text{Corr Bel about types} \right)$$

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

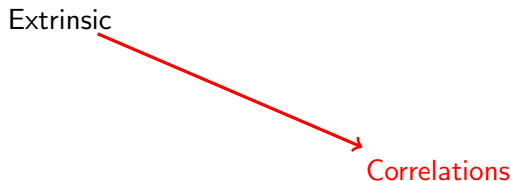
There are some conceptual issues...

► Explanation

Two routes to explain correlations

Correlations

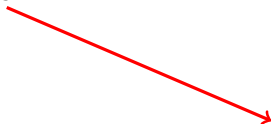
Two routes to explain correlations



R.J. Aumann. *Correlated equilibrium as an expression of bayesian rationality.* *Econometrica*, 55(1-18), 1987.

Two routes to explain correlations

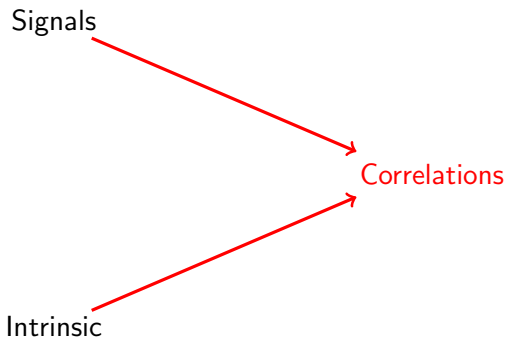
Signals



Correlations

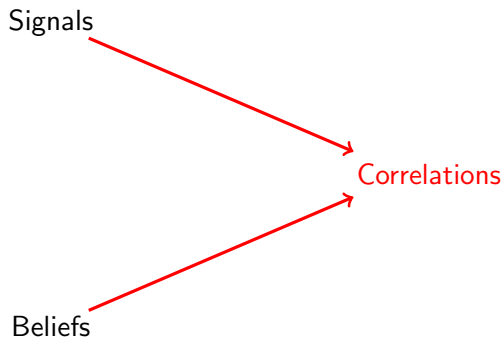
R.J. Aumann. *Correlated equilibrium as an expression of bayesian rationality.* *Econometrica*, 55(1-18), 1987.

Two routes to explain correlations



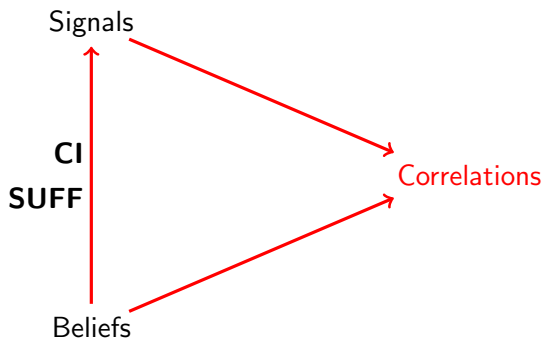
A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Two routes to explain correlations



A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

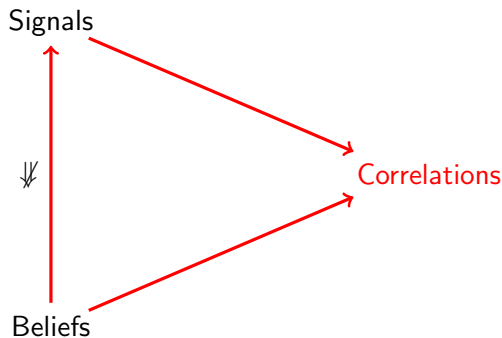
Two routes to explain correlations



R.J. Aumann. *Correlated equilibrium as an expression of bayesian rationality*. *Econometrica*, 55(1-18), 1987.

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Two routes to explain correlations



A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

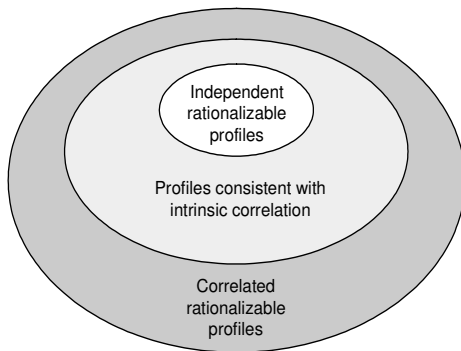


Figure 2.2

A. Brandenburger and A. Friedenberg. *Intrinsic Correlations in Games*. J.E.T., vol. 141, 2008.

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

- ▶ Is there a **systematic way in which correlations arises** in *ex interim* UTIs?

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

- ▶ Is there a systematic way in which correlations arises in *ex interim* UTIs?
- ▶ **Conditionals on protocols**, are **types** and/or **strategy** choices **independent**? (Protocols as signals?)

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

- ▶ Is there a systematic way in which correlations arises in *ex interim* UTIs?
- ▶ Conditionals on protocols, are types and/or strategy choices independent? (Protocols as signals?)
 - If yes: can we capture this trough (a larger set of types in) UTIs?

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

- ▶ Is there a systematic way in which correlations arises in *ex interim* UTIs?
- ▶ Conditionals on protocols, are types and/or strategy choices independent? (Protocols as signals?)
 - If yes: can we capture this trough (a larger set of types in) UTIs?

More general questions:

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

- ▶ Is there a systematic way in which correlations arises in *ex interim* UTIs?
- ▶ Conditionals on protocols, are types and/or strategy choices independent? (Protocols as signals?)
 - If yes: can we capture this trough (a larger set of types in) UTIs?

More general questions:

- What does it means to say that a **team** is **rational** at a given state?

Correlations in UTI

The kind of correlations that arise in *UTI* are akin to correlations induced by types. **Questions:**

- ▶ Is there a systematic way in which correlations arises in *ex interim* UTIs?
- ▶ Conditionals on protocols, are types and/or strategy choices independent? (Protocols as signals?)
 - If yes: can we capture this trough (a larger set of types in) UTIs?

More general questions:

- What does it means to say that a team is rational at a given state?
- And what about **correlations in non-probabilistic models of beliefs?**

Additional requirements on team work.

What is a team?

1. Group identification.
 - Information about who's in and who's out.
 - Reasoning as group members.
 - Shared goal.
 - ▶ Group preference / utilities.
2. Shared commitments.
 - Shared intentions.
 - Sanctions for lapsing?
 - Shared praise[blame] for success[failure]?
3. Common knowledge (beliefs?) of the above?

What is a team?

1. Group identification.

- Information about who's in and who's out.
- Reasoning as group members.
- Shared goal.
 - ▶ Group preference / utilities.

2. Shared commitments.

- Shared intentions.
- Sanctions for lapsing?
- Shared praise[blame] for success[failure]?

3. Common knowledge (beliefs?) of the above?

Intentions

Motivational attitudes which:

- ▶ Are relatively stable.
- ▶ Are conduct-controlling, i.e. commit to action.
- ▶ Constraint further practical reasoning.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions

Motivational attitudes which:

- ▶ Are relatively stable.
- ▶ Are conduct-controlling, i.e. commit to action.
- ▶ Constraint further practical reasoning.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions

Motivational attitudes which:

- ▶ Are **relatively stable**.
- ▶ Are conduct-controlling, i.e. commit to action.
- ▶ Constraint further practical reasoning.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions

Motivational attitudes which:

- ▶ Are relatively stable.
- ▶ Are **conduct-controlling**, i.e. commit to action.
- ▶ Constraint further practical reasoning.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions

Motivational attitudes which:

- ▶ Are relatively stable.
- ▶ Are conduct-controlling, i.e. commit to action.
- ▶ **Constraint** further **practical reasoning**.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions

Motivational attitudes which:

- ▶ Are relatively stable.
- ▶ Are conduct-controlling, i.e. commit to action.
- ▶ Constraint further practical reasoning.

Intentions **anchor inter-temporal** and interpersonal coordination.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions

Motivational attitudes which:

- ▶ Are relatively stable.
- ▶ Are conduct-controlling, i.e. commit to action.
- ▶ Constraint further practical reasoning.

Intentions **anchor** inter-temporal and **interpersonal** coordination.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Commitments and Intentions

Key Philosophical Work:

Commitments and Intentions

Key Philosophical Work:

G. Harmann. *Practical Reasoning. Review of Metaphysics*, 1976.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Commitments and Intentions

Key Philosophical Work:

G. Harmann. *Practical Reasoning. Review of Metaphysics*, 1976.

M. Bratman. *Intention, Plans, Practical Reason*. Harvard UP, 1987.

Intentions and Teamwork:

M. Gilbert. *On Social Facts*. Princeton UP, 1989.

J. Searle. *The Construction of Social Reality*. Free Press, 1995.

M. Bratman. *Faces of Intentions*. Cambridge UP, 1999.

R. Tuomela. *The Philosophy of Sociality*. Oxford UP, 2010.

Commitments and Intentions

Formally - but mostly single-agent:

Commitments and Intentions

Formally - but mostly single-agent:

M. Bratman, D. Israel and M. Pollack. *Plans and Ressource-Bounded Reasoning*. *Computational Intelligence*, 4(3), 1988.

P.R. Cohen and HJ Levesque. *Intention is Choice with Commitment*. *Artificial Intelligence*, 42(3), 1990.

F Gul, W Pesendorfer. *Temptation and Self-Control*. *Econometrica*, 2001.

⇒ See also (T. Icard, EP and Y. Shoham 2010).

Commitments and Intentions

Formally - but mostly single-agent:

M. Bratman, D. Israel and M. Pollack. *Plans and Ressource-Bounded Reasoning. Computational Intelligence*, 4(3), 1988.

P.R. Cohen and HJ Levesque. *Intention is Choice with Commitment. Artificial Intelligence*, 42(3), 1990.

F Gul, W Pesendorfer. *Temptation and Self-Control. Econometrica*, 2001.

⇒ See also (T. Icard, EP and Y. Shoham 2010).

Formally - Multi-agent:

B. Dunin-Keplicz and R. Verbrugge. *Teamwork in Multi-Agent Systems*. Wiley, 2010.

⇒ See also (Lorini and Herzig:

<http://www.irit.fr/~Andreas.Herzig/Esslli09course/>)

Teamwork as Shared cooperative activity:



Teamwork as Shared cooperative activity:

A The Intention part:

1. Me:

1.1 I intend that we J.

1.2 I intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

M. Bratman. *Faces of Intentions*. Cambridge UP, 1999.

Teamwork as Shared cooperative activity:

A The Intention part:

1. Me:

1.1 I intend that we J.

1.2 I intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

2. You:

2.1 You intend that we J.

2.2 You intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

M. Bratman. *Faces of Intentions*. Cambridge UP, 1999.

Teamwork as Shared cooperative activity:

A The Intention part:

1. Me:

1.1 I intend that we J.

1.2 I intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

2. You:

2.1 You intend that we J.

2.2 You intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

3. Additional requirements:

3.1 The intentions in (1) and in (2) are not coerced by the other participant.

3.2 The intentions in (1) and (2) are minimally cooperatively stable.

M. Bratman. *Faces of Intentions*. Cambridge UP, 1999.

Teamwork as Shared cooperative activity:

A The Intention part:

1. Me:

1.1 I intend that we J.

1.2 I intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

2. You:

2.1 You intend that we J.

2.2 You intend that we J in accordance with and because of meshing subplans of (1.1) and (2.1).

3. Additional requirements:

3.1 The intentions in (1) and in (2) are not coerced by the other participant.

3.2 The intentions in (1) and (2) are minimally cooperatively stable.

B: The epistemic part:

1. It is common knowledge between us that (A).

M. Bratman. *Faces of Intentions*. Cambridge UP, 1999.

Conclusions: Teams in Interaction

1. Group identification.
 - Information about who's in and who's out.
 - Reasoning as group members.
 - Shared goal.
 - ▶ Group preference / utilities.
2. Shared commitments.
 - Shared intentions.
 - Sanctions for lapsing?
 - Shared praise[blame] for success[failure]?
3. Common knowledge (beliefs?) of the above?

Thanks and Merci (and Bedankt)

CPA not an innocuous assumption!

- ▶ Agreement Theorem:

*If two people have the **common priors**, and their posteriors for an event A are **common knowledge**, then these posterior have to be **the same**.*

CPA not an innocuous assumption!

► Agreement Theorem:

The Annals of Statistics
1976, Vol. 4, No. 6, 1236-1239

AGREEING TO DISAGREE¹

BY ROBERT J. AUMANN

Stanford University and the Hebrew University of Jerusalem

Two people, 1 and 2, are said to have *common knowledge* of an event E if both know it, 1 knows that 2 knows it, 2 knows that 1 knows it, 1 knows that 2 knows that 1 knows it, and so on.

THEOREM. *If two people have the same priors, and their posteriors for an event A are common knowledge, then these posteriors are equal.*

If two people have the same priors, and their posteriors for a given event A are common knowledge, then these posteriors must be equal. This is so even though they may base their posteriors on quite different information. In brief, people with the same priors *cannot agree to disagree*.

We publish this observation with some diffidence, since once one has the appropriate framework, it is mathematically trivial. Intuitively, though, it is not

CPA not an innocuous assumption!

► Agreement Theorem:

The Annals of Statistics
1976, Vol. 4, No. 6, 1236-1239

AGREEING TO DISAGREE¹

BY ROBERT J. AUMANN

Stanford University and the Hebrew University of Jerusalem

Two people, 1 and 2, are said to have *common knowledge* of an event E if both know it, 1 knows that 2 knows it, 2 knows that 1 knows it, 1 knows that 2 knows that 1 knows it, and so on.

THEOREM. *If two people have the same priors, and their posteriors for an event A are common knowledge, then these posteriors are equal.*

If two people have the same priors, and their posteriors for a given event A are common knowledge, then these posteriors must be equal. This is so even though they may base their posteriors on quite different information. In brief, people with the same priors *cannot agree to disagree*.

We publish this observation with some diffidence, since once one has the appropriate framework, it is mathematically trivial. Intuitively, though, it is not

CPA not an innocuous assumption!

► Agreement Theorem:

The Annals of Statistics
1976, Vol. 4, No. 6, 1236-1239

AGREEING TO DISAGREE¹

BY ROBERT J. AUMANN

Stanford University and the Hebrew University of Jerusalem

Two people, 1 and 2, are said to have *common knowledge* of an event E if both know it, 1 knows that 2 knows it, 2 knows that 1 knows it, 1 knows that 2 knows that 1 knows it, and so on.

THEOREM. *If two people have the same priors, and their posteriors for an event A are common knowledge, then these posteriors are equal.*

If two people have the same priors, and their posteriors for a given event A are common knowledge, then these posteriors must be equal. This is so even though they may base their posteriors on quite different information. In brief, people with the same priors *cannot agree to disagree*.

We publish this observation with some diffidence, since once one has the appropriate framework, it is mathematically trivial. Intuitively, though, it is not

CPA not an innocuous assumption!

- ▶ Agreement Theorem: [▶ Back to Lecture](#)

An digression on CI and SUFF

Suppose a being in whose power to predict your choices you have enormous confidence. [...]

R. Nozick. *Newcomb's Problem and Two Principles of Choice*. 1969.

An digression on CI and SUFF

Suppose a being in whose power to predict your choices you have enormous confidence. [...] There are two boxes, (*A*) and (*B*). (*B*1) contains 1000 [euros]. (*B*2) contains either 1 million [euros] (*M*), or nothing.

An digression on CI and SUFF

Suppose a being in whose power to predict your choices you have enormous confidence. [...] There are two boxes, (*A*) and (*B*). (*B*1) contains 1000 [euros]. (*B*2) contains either 1 million [euros] (*M*), or nothing. You have a choice between two actions:

1. taking what is in both boxes
2. taking only what is in the second box.

An digression on CI and SUFF

Suppose a being in whose power to predict your choices you have enormous confidence. [...] There are two boxes, (A) and (B). (B1) contains 1000 [euros]. (B2) contains either 1 million [euros] (M), or nothing. You have a choice between two actions:

1. taking what is in both boxes
2. taking only what is in the second box.

Furthermore, [the following is common knowledge]:

- ▶ If the being predicts you will take what is in both boxes, he does not put the 1 M in the second box.
- ▶ If the being predicts you will take only what is in the second box, he does put the 1M in the second box.

An digression on CI and SUFF

Suppose a being in whose power to predict your choices you have enormous confidence. [...] There are two boxes, (A) and (B). (B1) contains 1000 [euros]. (B2) contains either 1 million [euros] (M), or nothing. You have a choice between two actions:

1. taking what is in both boxes
2. taking only what is in the second box.

Furthermore, [the following is common knowledge]:

- ▶ If the being predicts you will take what is in both boxes, he does not put the 1 M in the second box.
- ▶ If the being predicts you will take only what is in the second box, he does put the 1M in the second box.

The situation is as follows. First the being makes its prediction. Then it puts the 1M in the second box, or does not, depending upon what it has predicted. Then you make your choice. What do you do?

An digression on CI and SUFF

	$A = 1M$	$A = 0$
1 Box	1M	0
2 Boxes	$1M + 100$	100



An digression on CI and SUFF

	$A = 1M$	$A = 0$
1 Box	1M	0
2 Boxes	$1M + 100$	100

	$A = 1M$	$A = 0$
1 Box	$1/2$	0
2 Boxes	0	$1/2$



An digression on CI and SUFF

	$A = 1M$	$A = 0$
1 Box	1M	0
2 Boxes	$1M + 100$	100

	$A = 1M$	$A = 0$
1 Box	$1/2$	0
2 Boxes	0	$1/2$

► Back to Lecture

An digression on CI and SUFF

	$A = 1M$	$A = 0$
1 Box	1M	0
2 Boxes	$1M + 100$	100

	$A = 1M$	$A = 0$
1 Box	$1/2$	0
2 Boxes	0	$1/2$

- No uncertainty on types here.

► Back to Lecture

An digression on CI and SUFF

	$A = 1M$	$A = 0$
1 Box	1M	0
2 Boxes	$1M + 100$	100

	$A = 1M$	$A = 0$
1 Box	$1/2$	0
2 Boxes	0	$1/2$

- ▶ No uncertainty on types here.
- ▶ SUFF is trivially satisfied (always for two agents).

▶ Back to Lecture

An digression on CI and SUFF

	$A = 1M$	$A = 0$
1 Box	1M	0
2 Boxes	$1M + 100$	100

	$A = 1M$	$A = 0$
1 Box	$1/2$	0
2 Boxes	0	$1/2$

- ▶ No uncertainty on types here.
- ▶ SUFF is trivially satisfied (always for two agents).
- ▶ CI is violated.

▶ Back to Lecture