Epistemic Game Theory

Lecture 2

ESSLLI'12, Opole

Eric Pacuit Olivier Roy

TiLPS, Tilburg University MCMP, LMU Munich
`ai.stanford.edu/~epacuit`
`http://olivier.amonbofis.net`

August 7, 2012

Plan for the week

1. **Monday** Basic Concepts.
2. **Tuesday** Epistemics.
   - Relating dominance reasoning with maximizing expected utility
   - Probabilistic/graded models of beliefs, knowledge and higher-order attitudes.
   - Logical/qualitative models of beliefs, knowledge and higher-order attitudes.
3. **Wednesday** Fundamentals of Epistemic Game Theory.
4. **Thursday** Puzzles and Paradoxes.
5. **Friday** Extensions and New Directions.

Dominance vs MEU

		Bob	
		L	*R*
Ann	*U*	3,3	0,0
	D	0,0	1,1

Dominance vs MEU

		Bob	
		L	R
Ann	U	3,3	0,0
	D	0,0	1,1

- **Ann's beliefs**: $p_A \in \Delta(\{L, R\})$ with $p_A(L) = 1/6$
  Bob's beliefs: $p_B \in \Delta(\{U, D\})$ with $p_B(U) = 3/4$.

Dominance vs MEU

		Bob	
		L	R
Ann	U	3,3	0,0
	D	0,0	1,1

- **Ann's beliefs**: $p_A \in \Delta(\{L, R\})$ with $p_A(L) = 1/6$
  Bob's beliefs: $p_B \in \Delta(\{U, D\})$ with $p_B(U) = 3/4$.

  $EU(U, p_A) = p_A(L) \cdot u_A(U, L) + p_A(R) \cdot u_A(U, R)$
  $EU(D, p_A) = p_A(L) \cdot u_A(D, L) + p_A(R) \cdot u_A(D, R)$

Dominance vs MEU

		Bob	
		L	R
Ann	U	3,3	0,0
	D	0,0	1,1

- **Ann's beliefs**: $p_A \in \Delta(\{L, R\})$ with $p_A(L) = 1/6$
  Bob's beliefs: $p_B \in \Delta(\{U, D\})$ with $p_B(U) = 3/4$.

  $EU(U, p_A) = 1/6 \cdot 3 + 5/6 \cdot 0 = 0.5$
  $EU(D, p_A) = 1/6 \cdot 0 + 5/6 \cdot 1 = 0.833$

Dominance vs MEU

		Bob	
		L	R
Ann	U	3,3	0,0
	D	0,0	1,1

- **Ann's beliefs**: $p_A \in \Delta(\{L, R\})$ with $p_A(L) = 1/6$
  Bob's beliefs: $p_B \in \Delta(\{U, D\})$ with $p_B(U) = 3/4$.

  $$EU(U, p_A) = 1/6 \cdot 3 + 5/6 \cdot 0 = 0.5$$
  $$EU(D, p_A) = 1/6 \cdot 0 + 5/6 \cdot 1 = 0.833$$
  $$EU(L, p_B) = p_B(U) \cdot u_B(U, L) + p_B(D) \cdot u_B(D, R)$$
  $$EU(R, p_B) = p_B(U) \cdot u_B(U, R) + p_B(D) \cdot u_B(D, R)$$

Dominance vs MEU

		Bob	
		L	R
Ann	U	3,3	0,0
	D	0,0	1,1

- **Ann's beliefs**: $p_A \in \Delta(\{L, R\})$ with $p_A(L) = 1/6$
  Bob's beliefs: $p_B \in \Delta(\{U, D\})$ with $p_B(U) = 3/4$.

  $$EU(U, p_A) = 1/6 \cdot 3 + 5/6 \cdot 0 = 0.5$$
  $$EU(D, p_A) = 1/6 \cdot 0 + 5/6 \cdot 1 = 0.833$$
  $$EU(L, p_B) = 3/4 \cdot 3 + 1/4 \cdot 0 = 2.25$$
  $$EU(R, p_B) = 3/4 \cdot 0 + 1/4 \cdot 1 = 0.25$$

Dominance vs MEU

		Bob	
		L	R
Ann	U	3,3	0,0
	D	0,0	1,1

- **Ann's beliefs**: $p_A \in \Delta(\{L, R\})$ with $p_A(L) = 1/6$
  Bob's beliefs: $p_B \in \Delta(\{U, D\})$ with $p_B(U) = 3/4$.

  $EU(U, p_A) = 1/6 \cdot 3 + 5/6 \cdot 0 = 0.5$
  $EU(D, p_A) = 1/6 \cdot 0 + 5/6 \cdot 1 = 0.833$
  $EU(L, p_B) = 3/4 \cdot 3 + 1/4 \cdot 0 = 2.25$
  $EU(R, p_B) = 3/4 \cdot 0 + 1/4 \cdot 1 = 0.25$

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$

$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$
$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

$s, s' \in S_i$, s **strictly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \quad u_i(s, s_{-i}) > u_i(s', s_{-i})$$

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$
$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

$s, s' \in S_i$, s **strictly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \quad u_i(s, s_{-i}) > u_i(s', s_{-i})$$

$s, s' \in S_i$, s **weakly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \;\; u_i(s, s_{-i}) \geq u_i(s', s_{-i}) \text{ and } \exists s_{-i} \in X, \;\; u_i(s, s_{-i}) > u_i(s', s_{-i})$$

Comparing Dominance Reasoning and MEU

$G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$
$X \subseteq S_{-i}$ (a set of strategy profiles for all players except i)

$s, s' \in S_i$, s **strictly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \quad u_i(s, s_{-i}) > u_i(s', s_{-i})$$

$s, s' \in S_i$, s **weakly dominates** s' with respect to X provided

$$\forall s_{-i} \in X, \ \ u_i(s, s_{-i}) \geq u_i(s', s_{-i}) \text{ and } \exists s_{-i} \in X, \ \ u_i(s, s_{-i}) > u_i(s', s_{-i})$$

$p \in \Delta(X)$, s is a **best response** to p with respect to X provided

$$\forall s' \in S_i, \quad EU(s, p) \geq EU(s', p)$$

Strict Dominance and MEU

Fact. Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a strategic game and $X \subseteq S_{-i}$. A strategy $s_i \in S_i$ is strictly dominated (possibly by a mixed strategy) with respect to X iff there is no probability measure $p \in \Delta(X)$ such that s_i is a best response to p.

Dominance vs MEU

Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a finite strategic game.

Dominance vs MEU

Suppose that $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Dominance vs MEU

Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Let $p \in \Delta(X)$ be any probability measure. Then,

$$\forall s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) \geq p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

$$\exists s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) > p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Suppose that $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Let $p \in \Delta(X)$ be any probability measure. Then,

$$\forall s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) \geq p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

$$\exists s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) > p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Hence,

$$\sum_{s_{-i} \in S_{-i}} p(s_{-i}) \cdot u_i(s_i', s_{-i}) > \sum_{s_{-i} \in S_{-i}} p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Dominance vs MEU

Suppose that $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ is a finite strategic game.

Suppose that $s_i \in S_i$ is strictly dominated with respect to X:

$$\exists s_i' \in S_i, \forall s_{-i} \in X, \quad u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

Let $p \in \Delta(X)$ be any probability measure. Then,

$$\forall s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) \geq p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

$$\exists s_{-i} \in X, \quad p(s_{-i}) \cdot u_i(s_i', s_{-i}) > p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

Hence,

$$\sum_{s_{-i} \in S_{-i}} p(s_{-i}) \cdot u_i(s_i', s_{-i}) > \sum_{s_{-i} \in S_{-i}} p(s_{-i}) \cdot u_i(s_i, s_{-i})$$

So, $EU(s_i', p) > EU(s_i, p)$: s_i is not a best response to p.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two-player game.
(Let $U_i : \Delta(S_1) \times \Delta(S_2) \to \mathbb{R}$ be the expected utility for i)

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two-player game.
(Let $U_i : \Delta(S_1) \times \Delta(S_2) \to \mathbb{R}$ be the expected utility for i)

Suppose that $\alpha \in \Delta(S_1)$ is not a best response to any $p \in \Delta(S_2)$.

$$\forall p \in \Delta(S_2) \;\; \exists q \in \Delta(S_1), \quad U_1(q, p) > U_1(\alpha, p)$$

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

Dominance vs MEU

For the converse direction, we sketch the proof for two player games and where $X = S_{-i}$. [1]

Let $G = \langle S_1, S_2, u_1, u_2 \rangle$ be a two-player game.
(Let $U_i : \Delta(S_1) \times \Delta(S_2) \rightarrow \mathbb{R}$ be the expected utility for i)

Suppose that $\alpha \in \Delta(S_1)$ is not a best response to any $p \in \Delta(S_2)$.

$$\forall p \in \Delta(S_2) \;\; \exists q \in \Delta(S_1), \quad U_1(q, p) > U_1(\alpha, p)$$

We can define a function $b : \Delta(S_2) \rightarrow \Delta(S_1)$ where, for each $p \in \Delta(S_2)$, $U_1(b(p), p) > U_1(\alpha, p)$.

---

[1] The proof of the more general statement uses the *supporting hyperplane theorem* from convex analysis.

Dominance vs MEU

Consider the game $G' = \langle S_1, S_2, \overline{u}_1, \overline{u}_2 \rangle$ where

$$\overline{u}_1(s_1, s_2) = u_1(s_1, s_2) - U_1(\alpha, s_2) \text{ and } \overline{u}_2(s_1, s_2) = -\overline{u}_1(s_1, s_2)$$

Dominance vs MEU

Consider the game $G' = \langle S_1, S_2, \overline{u}_1, \overline{u}_2 \rangle$ where

$$\overline{u}_1(s_1, s_2) = u_1(s_1, s_2) - U_1(\alpha, s_2) \text{ and } \overline{u}_2(s_1, s_2) = -\overline{u}_1(s_1, s_2)$$

By the minimax theorem, there is a Nash equilibrium (p_1^*, p_2^*) such that for all $m \in \Delta(S_2)$,

$$\overline{U}(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*)$$

Consider the game $G' = \langle S_1, S_2, \overline{u}_1, \overline{u}_2 \rangle$ where

$$\overline{u}_1(s_1, s_2) = u_1(s_1, s_2) - U_1(\alpha, s_2) \text{ and } \overline{u}_2(s_1, s_2) = -\overline{u}_1(s_1, s_2)$$

By the minimax theorem, there is a Nash equilibrium (p_1^*, p_2^*) such that for all $m \in \Delta(S_2)$,

$$\overline{U}(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*)$$

We now prove that $\overline{U}_1(b(p_2^*), p_2^*) > 0$:

Dominance vs MEU

$$\overline{U}_1(b(p_2^*), p_2^*) \quad = \quad \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) \overline{u}_1(x, y)$$

Dominance vs MEU

$$\begin{aligned}\overline{U}_1(b(p_2^*), p_2^*) &= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y)\\ &= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha, y)]\end{aligned}$$

Dominance vs MEU

$$\begin{aligned}
\overline{U}_1(b(p_2^*), p_2^*) &= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y)\\
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)]\\
&= \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y)\\
&- \textstyle\sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)
\end{aligned}$$

Dominance vs MEU

$$
\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\[2ex]
 & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)] \\[2ex]
 & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
 & - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\[2ex]
 & = & U_1(b(p_2^*), p_2^*) \\
 & - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y)
\end{array}
$$

Dominance vs MEU

$$
\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) \overline{u}_1(x, y) \\
& = & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) [u_1(x, y) - U_1(\alpha, y)] \\
& = & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) u_1(x, y) \\
& - & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) U_1(\alpha, y) \\
& = & U_1(b(p_2^*), p_2^*) \\
& - & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) U_1(\alpha, y) \\
& > & U_1(\alpha, p_2^*) - \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) U_1(\alpha, y)
\end{array}
$$

Dominance vs MEU

$$\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)] \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(b(p_2^*), p_2^*) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& > & U_1(\alpha, p_2^*) - \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(\alpha, p_2^*) - \sum_{x\in S_1} b(p_2^*)(x)\sum_{y\in S_2} p_2^*(y)U_1(\alpha,y)
\end{array}$$

Dominance vs MEU

$$\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)\overline{u}_1(x,y) \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)[u_1(x,y) - U_1(\alpha,y)] \\
& = & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)u_1(x,y) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(b(p_2^*), p_2^*) \\
& - & \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& > & U_1(\alpha, p_2^*) - \sum_{x\in S_1}\sum_{y\in S_2} b(p_2^*)(x)p_2^*(y)U_1(\alpha,y) \\
& = & U_1(\alpha, p_2^*) - \sum_{x\in S_1} b(p_2^*)(x)\sum_{y\in S_2} p_2^*(y)U_1(\alpha,y) \\
& = & U_1(\alpha, p_2^*) - U_1(\alpha, p_2^*)\cdot\sum_{x\in S_1} b(p_2^*)(x)
\end{array}$$

Dominance vs MEU

$$\begin{array}{rcl}
\overline{U}_1(b(p_2^*), p_2^*) & = & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) \overline{u}_1(x,y) \\
& = & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) [u_1(x,y) - U_1(\alpha, y)] \\
& = & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) u_1(x,y) \\
& - & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) U_1(\alpha, y) \\
& = & U_1(b(p_2^*), p_2^*) \\
& - & \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) U_1(\alpha, y) \\
& > & U_1(\alpha, p_2^*) - \sum_{x \in S_1} \sum_{y \in S_2} b(p_2^*)(x) p_2^*(y) U_1(\alpha, y) \\
& = & U_1(\alpha, p_2^*) - \sum_{x \in S_1} b(p_2^*)(x) \sum_{y \in S_2} p_2^*(y) U_1(\alpha, y) \\
& = & U_1(\alpha, p_2^*) - U_1(\alpha, p_2^*) \cdot \sum_{x \in S_1} b(p_2^*)(x) \\
& = & U_1(\alpha, p_2^*) - U_1(\alpha, p_2^*) = 0
\end{array}$$

Dominance vs MEU

Hence, for all $m \in \Delta(S_2)$ we have

$$\overline{U}(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*) > 0$$

Hence, for all $m \in \Delta(S_2)$ we have

$$\overline{U}(p_1^*, m) \geq \overline{U}_1(p_1^*, p_2^*) \geq \overline{U}_1(b(p_2^*), p_2^*) > 0$$

which implies for all $m \in \Delta(S_2)$, $U_1(p_1^*, m) > U_1(\alpha, m)$, and so α is strictly dominated by p_1^*.

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- Note that y is not strictly dominated for Charles.

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- Note that y is not strictly dominated for Charles.
- It is easy to find a probability measure $p \in \Delta(S_A \times S_B)$ such that y is a best response to p. Suppose that $p(u,l) = p(d,r) = \frac{1}{2}$. Then, $EU(x,p) = EU(z,p) = 1.5$ while $EU(y,p) = 2$.

Important Issue: Correlated Beliefs

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- Note that y is not strictly dominated for Charles.
- It is easy to find a probability measure $p \in \Delta(S_A \times S_B)$ such that y is a best response to p. Suppose that $p(u, l) = p(d, r) = \frac{1}{2}$. Then, $EU(x, p) = EU(z, p) = 1.5$ while $EU(y, p) = 2$.
- However, there is no probability measure $p \in \Delta(S_A \times S_B)$ such that y is a best response to p and $p(u, l) = p(u) \cdot p(l)$.

Dominance vs MEU

x	l	r
u	1,1,3	1,0,3
d	0,1,0	0,0,0

y	l	r
u	1,1,2	1,0,0
d	0,1,0	1,1,2

z	l	r
u	1,1,0	1,0,0
d	0,1,3	0,0,3

- To see this, suppose that a is the probability assigned to u and b is the probability assigned to l. Then, we have:
  - The expected utility of y is $2ab + 2(1-a)(1-b)$;
  - The expected utility of x is $3ab + 3a(1-b) = 3a(b + (1-b)) = 3a$; and
  - The expected utility of z is $3(1-a)b + 3(1-a)(1-b) = 3(1-a)(b + (1-b)) = 3(1-a)$.

Weak Dominance and MEU

Fact. Suppose that $G = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$ is a strategic game and $X \subseteq S_{-i}$. A strategy $s_i \in S_i$ is weakly dominated (possibly by a mixed strategy) with respect to X iff there is **no full support probability measure** $p \in \Delta^{>0}(X)$ such that s_i is a best response to p.

Dominance vs MEU

Some preliminary remarks

Propositional Attitudes

- We will talk about so-called **propositional attitudes**. These are attitudes (like knowledge, beliefs, desires, intentions, etc...) that take propositions as objects.

Propositional Attitudes

- We will talk about so-called **propositional attitudes**. These are attitudes (like knowledge, beliefs, desires, intentions, etc...) that take propositions as objects.
- Proposition will be taken to be element of a given algebra. I.e. measurable subsets of a state space (sigma- and/or power-set algebra), formulas in a given language (abstract Boolean algebra)...

All-out vs graded attitudes

- A propositional attitude A is **all-out** when, for any proposition p, the agent can only be in **three states** of that attitude regarding p:
    1. Ap: the agent "believes" that p.
    2. $A\neg p$: the agent "disbelieve" that p.
    3. $\neg Ap \wedge \neg A\neg p$: the agent "suspends judgment" about p.

All-out vs graded attitudes

- A propositional attitude A is **all-out** when, for any proposition p, the agent can only be in **three states** of that attitude regarding p:
  1. Ap: the agent “believes” that p.
  2. $A\neg p$: the agent “disbelieve” that p.
  3. $\neg Ap \wedge \neg A\neg p$: the agent “suspends judgment” about p.

- A propositional attitude A is **graded** when, for any proposition p, the states of that attitude that the agent be in w.r.t. a proposition p can be compared according to their **strength** on a given scale.

p_i	P	$\neg P$
A	1/8	3/8

Hard and Soft Attitudes

- Hard attitudes:
  - Truthful.
  - Unrevisable.
  - Fully introspective.
- Soft attitudes:
  - Can be false / mistaken.
  - Revisable / can be reversed.
  - Not fully introspective.

Models of graded beliefs

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

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

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

- A **type** is everything a player knows privately at the beginning of the game which could affect his beliefs about payoffs and about all other players' possible types.
- Each **type** is assigned a joint probability over the space of types and actions

$$\lambda_i : T_i \rightarrow \Delta(T_{-i} \times S_{-i})$$

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

- A **type** is everything a player knows privately at the beginning of the game which could affect his beliefs about payoffs and about all other players' possible types.
- Each **type** is assigned a joint probability over the space of types and actions

$$\lambda_i : T_i \to \Delta(T_{-i} \times S_{-i})$$

Player i's types

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

- A **type** is everything a player knows privately at the beginning of the game which could affect his beliefs about payoffs and about all other players' possible types.
- Each **type** is assigned a joint probability over the space of types and actions

$$\lambda_i : T_i \to \Delta(T_{-i} \times S_{-i})$$

The set of all probability distributions

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

- A **type** is everything a player knows privately at the beginning of the game which could affect his beliefs about payoffs and about all other players' possible types.
- Each **type** is assigned a joint probability over the space of types and actions

$$\lambda_i : T_i \to \Delta(T_{-i} \times S_{-i})$$

The other players' types

Harsanyi Type Space

Based on the work of John Harsanyi on games with *incomplete information*, game theorists have developed an elegant formalism that makes precise talk about beliefs, knowledge and rationality:

- A **type** is everything a player knows privately at the beginning of the game which could affect his beliefs about payoffs and about all other players' possible types.
- Each **type** is assigned a joint probability over the space of types and actions

$$\lambda_i : T_i \to \Delta(T_{-i} \times S_{-i})$$

The other players' choices

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

Bob

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

- One type for Ann (t_A) and two types for Bob (t_B, u_B)

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

- One type for Ann (t_A) and two types for Bob (t_B, u_B)
- A **state** is a tuple of choices and types: (M, M, t_A, t_B)

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

- One type for Ann (t_A) and two types for Bob (t_B, u_B)
- A **state** is a tuple of choices and types: (M, t_A, M, u_B)
- Calculate **expected utility** in the usual way...

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

Bob

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

- M is **rational** for Ann (t_A)
  $0 \cdot 0.2 + 1 \cdot 0.8 \geq 3 \cdot 0.2 + 0 \cdot 0.8$

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

- M is **rational** for Ann (t_A)
  $0 \cdot 0.2 + 1 \cdot 0.8 \geq 3 \cdot 0.2 + 0 \cdot 0.8$
- M is **rational** for Bob (t_B)
  $0 \cdot 0 + 1 \cdot 1 \geq 3 \cdot 0 + 0 \cdot 1$

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

- M is **rational** for Ann (t_A)
  $0 \cdot 0.2 + 1 \cdot 0.8 \geq 3 \cdot 0.2 + 0 \cdot 0.8$
- M is **rational** for Bob (t_B)
  $0 \cdot 0 + 1 \cdot 1 \geq 3 \cdot 0 + 0 \cdot 1$
- Ann thinks Bob may be irrational

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Returning to the Example: A Game Model

		Bob	
		H	M
Ann	H	3,3	0,0
	M	0,0	1,1

- M is **rational** for Ann (t_A)
  $0 \cdot 0.2 + 1 \cdot 0.8 \geq 3 \cdot 0.2 + 0 \cdot 0.8$
- M is **rational** for Bob (t_B)
  $0 \cdot 0 + 1 \cdot 1 \geq 3 \cdot 0 + 0 \cdot 1$
- Ann thinks Bob may be irrational
  $P_A(Irrat[B]) = 0.3$, $P_A(Rat[B]) = 0.7$

	H	M
t_B	0	0.5
u_B	0.2	0.3

t_A

	H	M
t_A	0	1

t_B

	H	M
t_A	0.4	0.6

u_B

Rationality

Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game and $\mathcal{T} = \langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ a type space for G.

Rationality

Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game and $\mathcal{T} = \langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ a type space for G.

For each $t_i \in T_i$, we can define a probability measure $p_{t_i} \in \Delta(S_{-i})$:

$$p_{t_i}(s_{-i}) = \sum_{t_{-i} \in T_{-i}} \lambda_i(t_i)(s_{-i}, t_{-i})$$

Rationality

Let $G = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$ be a strategic game and $\mathcal{T} = \langle \{T_i\}_{i \in N}, \{\lambda_i\}_{i \in N}, S \rangle$ a type space for G.

For each $t_i \in T_i$, we can define a probability measure $p_{t_i} \in \Delta(S_{-i})$:

$$p_{t_i}(s_{-i}) = \sum_{t_{-i} \in T_{-i}} \lambda_i(t_i)(s_{-i}, t_{-i})$$

The set of states (pairs of strategy profiles and type profiles) where player i chooses **rationally** is:

$$\mathsf{Rat_i} \;:=\; \{(s_i, t_i) \mid s_i \text{ is a best response to } p_{t_i}\}$$

The event that all players are *rational* is
$\mathsf{Rat} = \{(s, t) \mid \text{ for all } i, \; (s_i, t_i) \in \mathsf{Rat_i}\}$.

Common "knowledge" of rationality

In much of this literature, "full belief" or sometimes "knowledge" is identified with probability 1.

(This is not a philosophical commitment, but rather a term of art!)

Common knowledge of rationality

Define R_i^n by induction on n:

Common knowledge of rationality

Define R_i^n by induction on n:

Let $R_i^1 = \mathsf{Rat_i}$.

Common knowledge of rationality

Define R_i^n by induction on n:

Let $R_i^1 = \mathsf{Rat_i}$.

Suppose that for each i, R_i^n has been defined.

Define R^n_{-i} as follows:

$$R^n_{-i} = \{(s,t) \mid s \in S_{-i},\ t \in T_{-j}, \text{ and for each } j \neq i,\ (s_j, t_j) \in R_j^n\}.$$

Common knowledge of rationality

Define R_i^n by induction on n:

Let $R_i^1 = \mathsf{Rat_i}$.

Suppose that for each i, R_i^n has been defined.

Define R^n_{-i} as follows:

$$R^n_{-i} = \{(s,t) \mid s \in S_{-i},\ t \in T_{-j}, \text{ and for each } j \neq i,\ (s_j, t_j) \in R_j^n\}.$$

For each $n > 1$,

$$R_i^{n+1} = \{(s,t) \mid (s,t) \in R_i^n \text{ and } \lambda_i(t) \text{ assigns probability 1 to } R^n_{-i}\}$$

Common knowledge of rationality

Define R_i^n by induction on n:

Let $R_i^1 = \mathsf{Rat_i}$.

Suppose that for each i, R_i^n has been defined.

Define R_{-i}^n as follows:

$$R_{-i}^n = \{(s,t) \mid s \in S_{-i},\ t \in T_{-j}, \text{ and for each } j \neq i,\ (s_j, t_j) \in R_j^n\}.$$

For each $n > 1$,

$$R_i^{n+1} = \{(s,t) \mid (s,t) \in R_i^n \text{ and } \lambda_i(t) \text{ assigns probability 1 to } R_{-i}^n\}$$

Common knowledge of rationality is:

$$\bigcap_{n \geq 1} R_1^n \ \times \ \bigcap_{n \geq 1} R_2^n \times \cdots \times \bigcap_{n \geq 1} R_N^n$$

Models of graded beliefs

		B	
		L	R
A	U	2,2	0,0
	D	0,0	1,1

- Consider the state (d, r, a_3, b_3). Both a_3 and b_3 correctly believe that (i.e., assign probability 1 to) the outcome is (d, r)

$\lambda_A(a_1)$	L	R
b_1	0.5	0.5
b_2	0	0
b_3	0	0

$\lambda_A(a_2)$	L	R
b_1	0.5	0
b_2	0	0
b_3	0	0.5

$\lambda_A(a_3)$	L	R
b_1	0	0
b_2	0	0.5
b_3	0	0.5

$\lambda_B(b_1)$	U	D
a_1	0.5	0
a_2	0	0.5
a_3	0	0

$\lambda_B(b_2)$	U	D
a_1	0.5	0
a_2	0	0
a_3	0	0.5

$\lambda_B(b_3)$	U	D
a_1	0	0
a_2	0	0.5
a_3	0	0.5

Models of graded beliefs

		B	
		L	R
A	U	2,2	0,0
	D	0,0	1,1

- This fact is not common knowledge: a_3 assigns a 0.5 probability to Bob being of type b_2, and type b_2 assigns a 0.5 probability to Ann playing l. *Ann does not know that Bob knows that she is playing r*

$\lambda_A(a_1)$	L	R
b_1	0.5	0.5
b_2	0	0
b_3	0	0

$\lambda_A(a_2)$	L	R
b_1	0.5	0
b_2	0	0
b_3	0	0.5

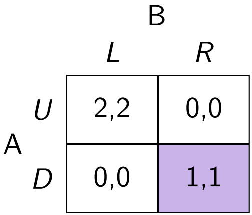
$\lambda_A(a_3)$	L	R
b_1	0	0
b_2	0	0.5
b_3	0	0.5

$\lambda_B(b_1)$	U	D
a_1	0.5	0
a_2	0	0.5
a_3	0	0

$\lambda_B(b_2)$	U	D
a_1	0.5	0
a_2	0	0
a_3	0	0.5

$\lambda_B(b_3)$	U	D
a_1	0	0
a_2	0	0.5
a_3	0	0.5

Models of graded beliefs



		B	
		L	R
A	U	2,2	0,0
	D	0,0	1,1

- Furthermore, while it is true that both Ann and Bob are rational, it is not common knowledge that they are rational.

$\lambda_A(a_1)$	L	R
b_1	0.5	0.5
b_2	0	0
b_3	0	0

$\lambda_A(a_2)$	L	R
b_1	0.5	0
b_2	0	0
b_3	0	0.5

$\lambda_A(a_3)$	L	R
b_1	0	0
b_2	0	0.5
b_3	0	0.5

$\lambda_B(b_1)$	U	D
a_1	0.5	0
a_2	0	0.5
a_3	0	0

$\lambda_B(b_2)$	U	D
a_1	0.5	0
a_2	0	0
a_3	0	0.5

$\lambda_B(b_3)$	U	D
a_1	0	0
a_2	0	0.5
a_3	0	0.5

General Comments

- Suppressed mathematical details about probabilities (σ-algebra, etc.)

General Comments

- Suppressed mathematical details about probabilities (σ-algebra, etc.)

- "Impossibility" is identified with probability 0, but it is an important distinction (especially for infinite games)

General Comments

- Suppressed mathematical details about probabilities (σ-algebra, etc.)
- "Impossibility" is identified with probability 0, but it is an important distinction (especially for infinite games)
- We can model "soft" information using conditional probability systems, lexicographic probabilities, nonstandard probabilities (more on this later).

Models of all-out attitudes.

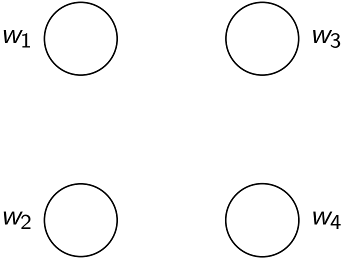
Hard Information

Example

		Bob	
		r	l
Ann	u	1, -1	-1, 1
	d	-1,1	1, -1

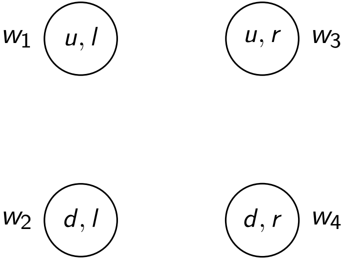
Example

		Bob	
		r	l
Ann	u	1, -1	-1, 1
	d	-1,1	1, -1



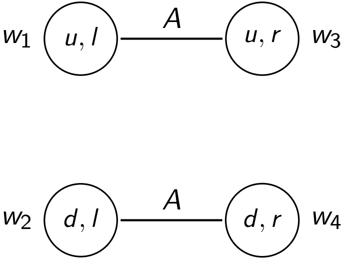
Example

		Bob	
		r	l
Ann	u	1, -1	-1, 1
	d	-1,1	1, -1



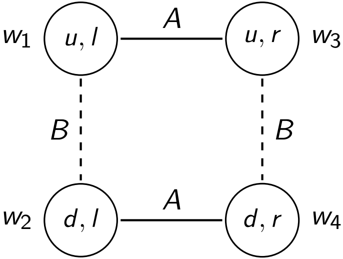
Example

		Bob	
		r	l
Ann	u	1, -1	-1, 1
	d	-1,1	1, -1



Example

		Bob	
		r	l
Ann	u	1, -1	-1, 1
	d	-1,1	1, -1



Epistemic Model

Suppose that G is a strategic game, S is the set of strategy profiles of G, and Ag is the set of players. An **epistemic model based on S and Ag** is a triple $\langle W, \{\Pi_i\}_{i\in \mathrm{Ag}}, \sigma\rangle$, where W is a nonempty set, for each $i \in \mathrm{Ag}$, Π_i is a partition[2] over W and $\sigma : W \to S$ is a strategy function.

---

[2] A partition of W is a pairwise disjoint collection of subsets of W whose union is all of W. Elements of a partition Π on W are called **cells**, and for $w \in W$, let $\Pi(w)$ denote the cell of Π containing w.

Epistemic Model

Suppose that G is a strategic game, S is the set of strategy profiles of G, and Ag is the set of players. An **epistemic model based on S and Ag** is a triple $\langle W, \{\Pi_i\}_{i \in \mathrm{Ag}}, \sigma \rangle$, where W is a nonempty set, for each $i \in \mathrm{Ag}$, Π_i is a partition[2] over W and $\sigma : W \to S$ is a strategy function.

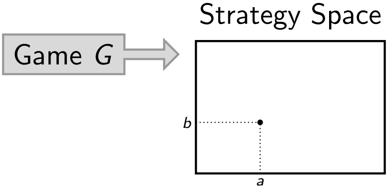
---

[2] A partition of W is a pairwise disjoint collection of subsets of W whose union is all of W. Elements of a partition Π on W are called **cells**, and for $w \in W$, let $\Pi(w)$ denote the cell of Π containing w.

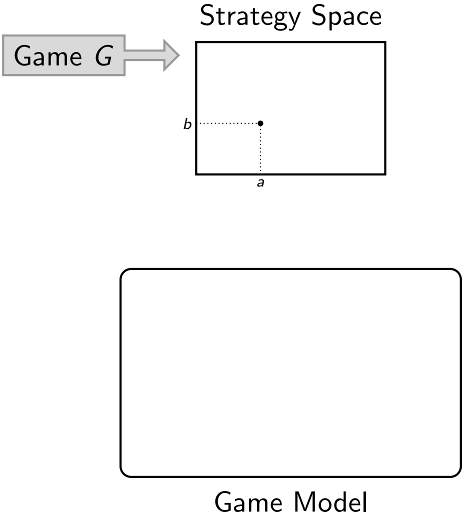
Game models

Game G

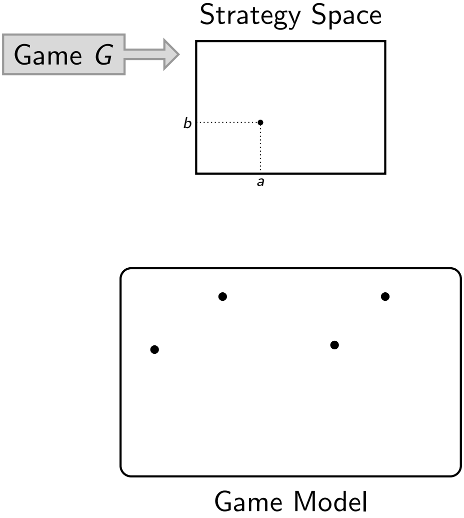
Game models



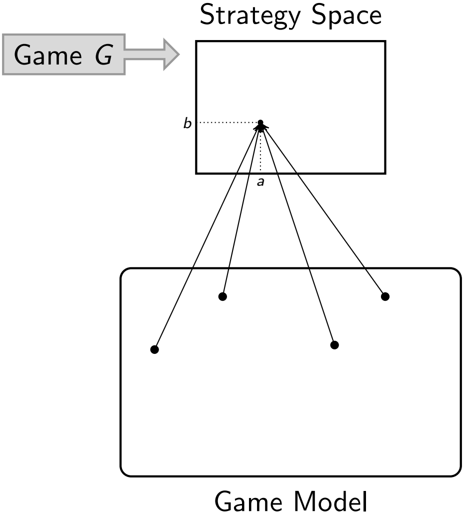
Game models



Game models



Game models

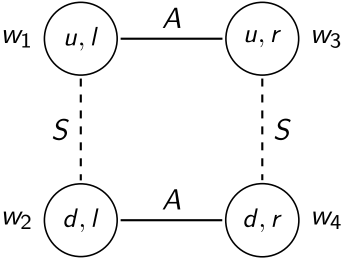


Kripke Model for S5

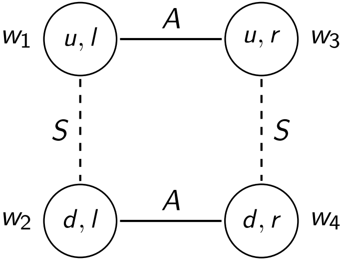
Prop is a given set of atomic propositions and Ag is a set of agents. An **epistemic model based on *Prop* and** Ag is a triple $\langle W, \{\Pi_i\}_{i\in\mathrm{Ag}}, V\rangle$, where W is a nonempty set, for each $i \in \mathrm{Ag}$, Π_i is a partition over W and $V : W \to \mathcal{P}(\mathit{Prop})$ is a valuation function.

Example

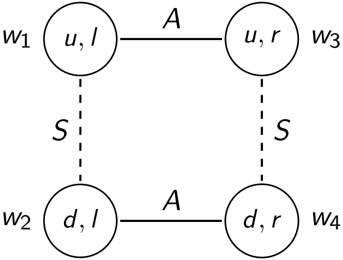
		A	
		r	l
S	u	1, -1	-1, 1
	d	-1,1	1, -1



Example

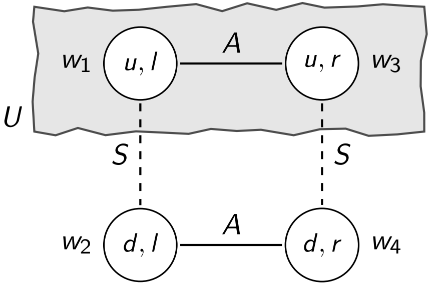


Example



- $\mathcal{M}, w \models K_i\varphi$ iff for all $w' \in \pi_i(w)$, $\mathcal{M} w' \models \varphi$.

Example



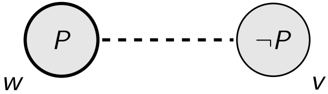
- $\mathcal{M}, w \models K_i\varphi$ iff for all $w' \in \pi_i(w)$, $\mathcal{M} w' \models \varphi$.
- One assumption: Ex-interim condition.
    - If $w' \in \pi_i(w)$ then $\sigma(w)_i = \sigma(w')_i$.

Hard Information, Axiomatically

1. Closed under known implication (K): $K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi)$
2. Logical truth are known (NEC): If $\models \varphi$ then $\models K_i\varphi$
3. Truthful, (T): $K_i\varphi \rightarrow \varphi$
4. Positive introspection (4): $K_i\varphi \rightarrow K_iK_i\varphi$
5. Negative introspection (5): $\neg K_i\varphi \rightarrow K_i\neg K_i\varphi$

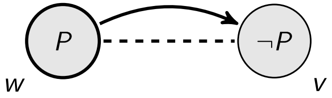
Soft Information

Modeling soft attitudes



Ann does not know that P

Modeling soft attitudes



Ann does not know that P, but she believes that $\neg P$

Plausibility Models

Let *Prop* be a countable set of propositions and Ag a set of agents.
A **plausibility model** $\mathcal{M}$ is a tuple $\langle W, \{\preceq_i\}_{i\in\mathrm{Ag}}, V\rangle$ where:

Plausibility Models

Let *Prop* be a countable set of propositions and Ag a set of agents.
A **plausibility model** $\mathcal{M}$ is a tuple $\langle W, \{\preceq_i\}_{i \in \mathrm{Ag}}, V \rangle$ where:

- W is a non-empty set of states.

Plausibility Models

Let *Prop* be a countable set of propositions and Ag a set of agents.
A **plausibility model** $\mathcal{M}$ is a tuple $\langle W, \{\preceq_i\}_{i\in\mathrm{Ag}}, V\rangle$ where:

- W is a non-empty set of states.
- for each $i \in \mathrm{Ag}$, $\preceq_i$ is a well-founded pre-order on W.
- $V : W \longrightarrow \mathcal{P}(Prop)$ is a valuation function.

For all φ, write $||\varphi||$ for $\{w|\mathcal{M}, w \models \varphi\}$

Plausibility Models

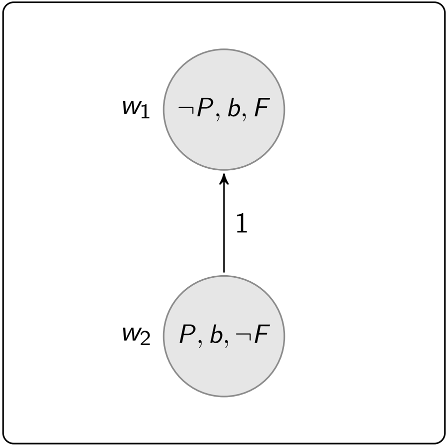
Let *Prop* be a countable set of propositions and Ag a set of agents. A **plausibility model** $\mathcal{M}$ is a tuple $\langle W, \{\preceq_i\}_{i\in\mathrm{Ag}}, V\rangle$ where:

- W is a non-empty set of states.
- for each $i \in \mathrm{Ag}$, $\preceq_i$ is a well-founded pre-order on W.
- $V : W \longrightarrow \mathcal{P}(Prop)$ is a valuation function.

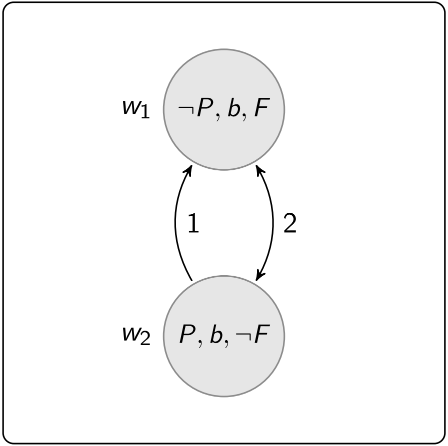
For all φ, write $||\varphi||$ for $\{w|\mathcal{M}, w \models \varphi\}$

- Maximum plausibility in a given set X:
  - $max_{\preceq_i}(X) = \{w \in X : \text{for all } w' \in X, w' \preceq_i w\}$
- Hard information defined:
  - $w \sim_i w'$ iff either $w' \preceq_i w$ or $w \preceq_i w'$.
  - Let $\pi_i(w) = \{w' : w \sim_i w'\}$. Then $\{\pi_i(w) : w \in W\}$ is a partition of W.

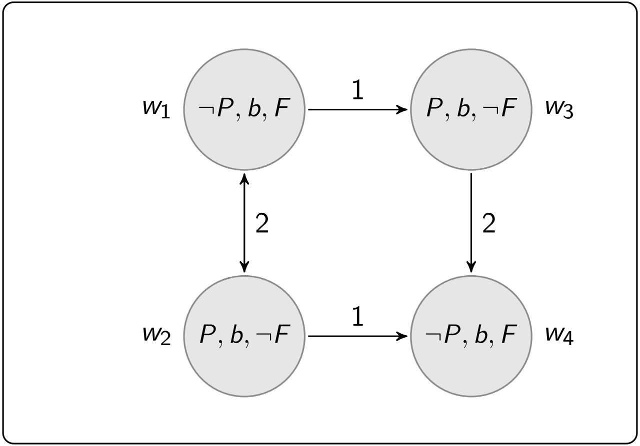
Example - Tweety is a penguin



Example - Tweety is a penguin



Example - Tweety is a penguin



Final Remarks

- Two broad families of models of higher-order information:
  - Probabilistic/graded.
  - Logical/qualitative.
- This is not meant to be a sharp distinction! (See SEP entry).

Final Remarks

- Two broad families of models of higher-order information:
  - Probabilistic/graded.
  - Logical/qualitative.
- This is not meant to be a sharp distinction! (See SEP entry).
- In both the notion of a state is crucial. A state encodes:
  1. The "non-epistemic facts". Here, mostly: what the agents are playing.
  2. What the agents know and/or believe about 1.
  3. What the agents know and/or believe about 2.
  4. ...

Final Remarks

- Two broad families of models of higher-order information:
  - Probabilistic/graded.
  - Logical/qualitative.
- This is not meant to be a sharp distinction! (See SEP entry).
- In both the notion of a state is crucial. A state encodes:
  1. The "non-epistemic facts". Here, mostly: what the agents are playing.
  2. What the agents know and/or believe about 1.
  3. What the agents know and/or believe about 2.
  4. ...
- Tomorrow: we put all this machinery to work in the context of games.
- Tonight: don't miss the evening lecture.