Plan for the week

1. **Monday** Basic Concepts.
2. **Tuesday** Epistemics.
3. **Wednesday** Fundamentals of Epistemic Game Theory.
   - Models of all-out attitudes (cnt’d).
   - Common knowledge of Rationality and iterated strict dominance in the matrix.
   - (If time, o/w tomorrow.) Common knowledge of Rationality and backward induction (strict dominance in the tree).
4. **Thursday** Puzzles and Paradoxes.
5. **Friday** Extensions and New Directions.
A family of attitudes
A family of attitudes

- Conditional Beliefs: $\mathcal{M}, w \models B_i^\varphi \psi$ iff $\mathcal{M}, w' \models \psi$ for all $w' \in \max_{\leq i}(\pi_i(w) \cap \|\varphi\|)$. 

A family of attitudes

- **Conditional Beliefs**: \( M, w \models B^\varphi_i \psi \) iff \( M, w' \models \psi \) for all \( w' \in \max _{\leq i} (\pi_i(w) \cap \|\varphi\|) \).

- **Safe Belief**: \( M, w \models [\leq]_i \varphi \) iff \( M, w' \models \varphi \) for all \( w' \leq_i w \).
A family of attitudes

- **Conditional Beliefs**: \( \mathcal{M}, w \models B^\varphi_i \psi \) iff \( \mathcal{M}, w' \models \psi \) for all \( w' \in \max_{\preceq_i} (\pi_i(w) \cap ||\varphi||) \).

- **Safe Belief**: \( \mathcal{M}, w \models [\preceq_i] \varphi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \preceq_i w \).

- **Knowledge**: \( \mathcal{M}, w \models K_i \varphi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \) such that \( w' \sim_i w \).
A family of attitudes

- **Conditional Beliefs**: \( \mathcal{M}, w \models B_i^\varphi \psi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \in \max_{\preceq_i} (\pi_i(w) \cap \|\varphi\|) \).

- **Safe Belief**: \( \mathcal{M}, w \models [\preceq_i] \varphi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \preceq_i w \).

- **Knowledge**: \( \mathcal{M}, w \models K_i \varphi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \sim_i w \).

Plain beliefs defined:

\[
B_i^\psi \iff_{df} B_i^T \varphi
\]
A family of attitudes

- **Conditional Beliefs:** \( \mathcal{M}, w \models B_i^\varphi \psi \) iff \( \mathcal{M}, w' \models \psi \) for all \( w' \in \max_{\preceq_i}(\pi_i(w) \cap ||\varphi||) \).

- **Safe Belief:** \( \mathcal{M}, w \models [\preceq_i] \varphi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \preceq_i w \).

- **Knowledge:** \( \mathcal{M}, w \models K_i \varphi \) iff \( \mathcal{M}, w' \models \varphi \) for all \( w' \sim_i w \).

Plain beliefs defined:

\[
B_i \psi \iff_{df} B_i^T \varphi
\]

Conditional beliefs defined:

\[
B_i^\varphi \psi \iff_{df} \langle K \rangle_i \varphi \rightarrow \langle K \rangle_i (\varphi \land [\preceq_i] \varphi \rightarrow \psi)
\]
Suppose that $w$ is the current state.

Knowledge ($\text{KP}$)

Belief ($\text{BP}$)

Safe Belief ($\text{SP}$)

Strong Belief ($\text{BS}$)

Graded Beliefs ($\text{B}_{\varphi}$)

Eric Pacuit and Olivier Roy
Suppose that $w$ is the current state.
Suppose that $w$ is the current state.

- **Belief** ($B; p$)
Suppose that $w$ is the current state.

- **Belief** ($B; p$)
- **Safe Belief** ($[\leq]; p$)
Soft attitudes

Suppose that \( w \) is the current state.

- **Belief** \((B_i p)\)
- **Safe Belief** \((\preceq_i p)\)
- **Knowledge** \((K_i p)\)
Properties of Soft Attitudes

Beliefs and conditional beliefs can be mistaken.

\[ \not P \rightarrow \not P \not \rightarrow 0 \not \rightarrow \phi \rightarrow P \]

\[ \not B_i \phi \rightarrow \phi \]
Properties of Soft Attitudes

Beliefs and conditional beliefs are fully introspective.

\[ P \rightarrow \neg P \rightarrow \neg P \rightarrow P \]

\[ v_0 \rightarrow v_1 \rightarrow w \rightarrow v_2 \]

\[ \models B_i \varphi \rightarrow B_i B_i \varphi \]

\[ \models \neg B_i \varphi \rightarrow B_i \neg B_i \varphi \]
The whole problem with the world is that fools and fanatics are always so certain of themselves, and wiser people so full of doubts.

-Bertrand Russell
Properties of Soft Attitudes

Safe Belief is truthful and positively introspective.

|\[v_0\] \quad |\[P\] \quad |\[\neg P\] \quad |\[P\] \quad |\[P\] \quad |\[v_2\]|

\[\models [\preceq]_i \varphi \rightarrow \varphi\]

\[\models [\preceq]_i \varphi \rightarrow [\preceq]_i [\preceq]_i \varphi\]
Properties of Soft Attitudes

Safe Belief is **not** negatively introspective.

\[\neg P, Q \vdash P, Q\]

<table>
<thead>
<tr>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

\[\nRightarrow \neg[\leq]i\varphi \rightarrow [\leq]i\neg[\leq]i\varphi\]

but...

\[\models B_i\varphi \iff B_i[\leq]i\varphi\]
Higher-order attitudes and common knowledge.
“Common Knowledge” is informally described as what any fool would know, given a certain situation: It encompasses what is relevant, agreed upon, established by precedent, assumed, being attended to, salient, or in the conversational record.
“Common Knowledge” is informally described as what any fool would know, given a certain situation: It encompasses what is relevant, agreed upon, established by precedent, assumed, being attended to, salient, or in the conversational record.

It is not Common Knowledge who “defined” Common Knowledge!
The first formal definition of common knowledge?


The first formal definition of common knowledge?


<table>
<thead>
<tr>
<th>The first rigorous analysis of common knowledge</th>
</tr>
</thead>
</table>
The first formal definition of common knowledge?

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. Aumann</td>
<td><em>Agreeing to Disagree</em></td>
<td>Annals of Statistics (1976)</td>
</tr>
</tbody>
</table>

The first rigorous analysis of common knowledge

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. Lewis</td>
<td><em>Convention, A Philosophical Study</em></td>
<td>1969.</td>
</tr>
</tbody>
</table>

**Fixed-point definition:** \( \gamma := i \text{ and } j \text{ know that } (\varphi \text{ and } \gamma) \)

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Publication</th>
</tr>
</thead>
</table>
The first formal definition of common knowledge?

The first rigorous analysis of common knowledge

**Fixed-point definition:** \( \gamma := i \text{ and } j \text{ know that } (\varphi \text{ and } \gamma) \)

**Shared situation:** There is a *shared situation* \( s \) such that (1) \( s \) entails \( \varphi \), (2) \( s \) entails everyone knows \( \varphi \), plus other conditions
http://plato.stanford.edu/entries/common-knowledge/.
The “Standard” Account


The “Standard” Account

An event/proposition is any (definable) subset $E \subseteq W$
The “Standard” Account
The “Standard” Account

\[ w \models K_A(E) \text{ and } w \not\models K_B(E) \]
The “Standard” Account

The model also describes the agents’ higher-order knowledge/beliefs
The “Standard” Account

Everyone Knows: \( K(E) = \bigcap_{i \in \mathcal{A}} K_i(E), \ K^0(E) = E, \ K^m(E) = K(K^{m-1}(E)) \)
The “Standard” Account

Common Knowledge:

\[ C(E) = \bigcap_{m \geq 0} K^m(E) \]
The “Standard” Account

\[ w \models K(E) \quad w \nmodels C(E) \]
The “Standard” Account

\[ w \models C(E) \]
**Fact.** For all $i \in \mathcal{A}$ and $E \subseteq W$, $K_i C(E) = C(E)$. 
Fact. For all $i \in \mathcal{A}$ and $E \subseteq W$, $K_i C(E) = C(E)$.

Suppose you are told “Ann and Bob are going together,”’ and respond “sure, that’s common knowledge.” What you mean is not only that everyone knows this, but also that the announcement is pointless, occasions no surprise, reveals nothing new; pause in effect, that the situation after the announcement does not differ from that before. ... the event “Ann and Bob are going together” — call it $E$ — is common knowledge if and only if some event — call it $F$ — happened that entails $E$ and also entails all players’ knowing $F$ (like all players met Ann and Bob at an intimate party). (Aumann, 1999 pg. 271, footnote 8)
**Fact.** For all $i \in \mathcal{A}$ and $E \subseteq \mathcal{W}$, $K_i C(E) = C(E)$.

An event $F$ is **self-evident** if $K_i(F) = F$ for all $i \in \mathcal{A}$.

**Fact.** An event $E$ is commonly known iff some self-evident event that entails $E$ obtains.
Fact. For all $i \in A$ and $E \subseteq W$, $K_i C(E) = C(E)$.

An event $F$ is **self-evident** if $K_i(F) = F$ for all $i \in A$.

Fact. An event $E$ is commonly known iff some self-evident event that entails $E$ obtains.

Fact. $w \in C(E)$ if every finite path starting at $w$ ends in a state in $E$

The following axiomatize common knowledge:

- $C(\varphi \to \psi) \to (C\varphi \to C\psi)$
- $C\varphi \to (\varphi \land EC\varphi)$ (Fixed-Point)
- $C(\varphi \to E\varphi) \to (\varphi \to C\varphi)$ (Induction)

With $E\varphi := \bigwedge_{i \in A} K_i \varphi$. 
Some General Remarks

Two broad families of models of higher-order information:

• Type spaces. (probabilistic)
• Plausibility models. (all-out)

There’s also a natural notion of qualitative type spaces, just like a natural probabilistic version of plausibility models. No strict separation between the two ways of thinking about information in interaction.

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. ...
Some General Remarks

- Two broad families of models of higher-order information:
  - Type spaces. (probabilistic)
  - Plausibility models. (all-out)

- There’s also a natural notion of qualitative type spaces, just like a natural probabilistic version of plausibility models. No strict separation between the two ways of thinking about information in interaction.
Some General Remarks

- Two broad families of models of higher-order information:
  - Type spaces. (probabilistic)
  - Plausibility models. (all-out)

- There’s also a natural notion of qualitative type spaces, just like a natural probabilistic version of plausibility models. No strict separation between the two ways of thinking about information in interaction.

- 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. ...
Now let’s do epistemics in games…
The Epistemic or Bayesian View on Games

- Traditional game theory: Actions, outcomes, preferences, solution concepts.

- Epistemic game theory: Actions, outcomes, preferences, beliefs, choice rules.
Basics again

The Epistemic or Bayesian View on Games

- Traditional game theory:
  Actions, outcomes, preferences, solution concepts.

- Epistemic game theory:
  Actions, outcomes, preferences, beliefs, choice rules.
  \(\equiv\) (interactive) decision problem: choice rule and higher-order information.
The Epistemic or Bayesian View on Games

- Traditional game theory:
  Actions, outcomes, preferences, solution concepts.
- Decision theory:
  Actions, outcomes, preferences, beliefs, choice rules.
- Epistemic game theory:
  Actions, outcomes, preferences, beliefs, choice rules.
  \[=\] (interactive) decision problem: choice rule and higher-order information.
Beliefs, Choice Rules, Rationality

What do we mean when we say that a player chooses rationally? That she follows some given choice rules.

- Maximization of expected utility, (Strict) dominance reasoning, Admissibility, etc.
Beliefs, Choice Rules, Rationality

What do we mean when we say that a player chooses rationally? That she follows some given choice rules.

- Maximization of expected utility, (Strict) dominance reasoning, Admissibility, etc.

In game models:

- The model describes the choices and (higher-order) beliefs/attitudes at each state.
- It is the choice rules that determine whether the choice made at each state is ”rational” or not.
  - An agent can be rational at a state given one choice rule, but irrational given the other.
  - Rationality in this sense is not built in the models.
Rationality

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

$$\text{Rat}_i := \{(s_i, t_i) \mid s_i \text{ is a best response to } p_{t_i}\}$$

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

Let \( G = \langle N, \{ S_i \}_{i \in N}, \{ u_i \}_{i \in N} \rangle \) be a strategic game and \( T = \langle \{ T_i \}_{i \in N}, \{ \lambda_i \}_{i \in N}, S \rangle \) 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:

\[
\text{Rat}_i := \{ (s_i, t_i) \mid s_i \text{ is a best response to } p_{t_i} \}
\]

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

- **Types**, as opposed to players, are rational or not at a given state.
Rationality and common belief of rationality (RCBR) in the matrix
<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3,3</td>
<td>1,1</td>
<td>0,0</td>
</tr>
<tr>
<td>m</td>
<td>1,1</td>
<td>3,3</td>
<td>1,0</td>
</tr>
<tr>
<td>m</td>
<td>0,4</td>
<td>0,0</td>
<td>4,0</td>
</tr>
</tbody>
</table>
IESDS

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3,3</td>
<td>1,1</td>
</tr>
<tr>
<td>m</td>
<td>1,1</td>
<td>3,3</td>
</tr>
<tr>
<td>m</td>
<td>0,4</td>
<td>0,0</td>
</tr>
</tbody>
</table>

1 \rightarrow 1

<table>
<thead>
<tr>
<th></th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3,3</td>
</tr>
<tr>
<td>m</td>
<td>1,1</td>
</tr>
<tr>
<td>b</td>
<td>0,4</td>
</tr>
</tbody>
</table>
IESDS

Eric Pacuit and Olivier Roy
1’s types

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

For $t_2$:

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
2’s types

\[ \lambda_2(s_1) \]
\[
\begin{array}{ccc}
  t & m & b \\
  t_1 & 0.5 & 0.5 & 0 \\
  t_2 & 0 & 0 & 0 \\
\end{array}
\]

\[ \lambda_2(s_2) \]
\[
\begin{array}{ccc}
  t & m & b \\
  t_1 & 0.25 & 0.25 & 0 \\
  t_2 & 0.25 & 0.25 & 0 \\
\end{array}
\]

\[ \lambda_2(s_3) \]
\[
\begin{array}{ccc}
  t & m & b \\
  t_1 & 0.5 & 0 & 0 \\
  t_2 & 0 & 0 & 0.5 \\
\end{array}
\]
### RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3,3</td>
<td>1,1</td>
<td>0,0</td>
</tr>
<tr>
<td>m</td>
<td>1,1</td>
<td>3,3</td>
<td>1,0</td>
</tr>
<tr>
<td>b</td>
<td>0,4</td>
<td>0,0</td>
<td>4,0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_2(s_1)$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$\lambda_2(s_2)$</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>$\lambda_2(s_3)$</td>
<td>0.5</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Eric Pacuit and Olivier Roy
RCBR in the Matrix

\[
\begin{array}{c|ccc}
 & l & c & r \\
\hline
 t & 3,3 & 1,1 & 0,0 \\
m & 1,1 & 3,3 & 1,0 \\
b & 0,4 & 0,0 & 4,0 \\
\end{array}
\]

\[
\lambda_2(s_1)
\]

\[
\begin{array}{ccc}
t & m & b \\
\hline
t_1 & 0.5 & 0.5 & 0 \\
t_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\lambda_2(s_2)
\]

\[
\begin{array}{ccc}
t & m & b \\
\hline
t_1 & 0.25 & 0.25 & 0 \\
t_2 & 0.25 & 0.25 & 0 \\
\end{array}
\]

\[
\lambda_2(s_3)
\]

\[
\begin{array}{ccc}
t & m & b \\
\hline
t_1 & 0.5 & 0 & 0 \\
t_2 & 0 & 0 & 0.5 \\
\end{array}
\]

- l and c are rational for both \(s_1\) and \(s_2\).
Read the text below.

RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3, 3</td>
<td>1, 1</td>
<td>0, 0</td>
</tr>
<tr>
<td>m</td>
<td>1, 1</td>
<td>3, 3</td>
<td>1, 0</td>
</tr>
<tr>
<td>b</td>
<td>0, 4</td>
<td>0, 0</td>
<td>4, 0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_1) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>t2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_2) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>t2</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_3) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>t2</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- l and c are rational for both \( s_1 \) and \( s_2 \).

- l is the only rational action for \( s_3 \).

- Whatever her type, it is never rational to play r for 2.
RCBR in the Matrix

\[
\begin{array}{ccc}
  & l & c & r \\
 1 & 3,3 & 1,1 & 0,0 \\
  t & 1,1 & 3,3 & 1,0 \\
  m & 0,4 & 0,0 & 4,0 \\
  b & & & \\
\end{array}
\]

\[
\begin{array}{ccc}
  t & m & b \\
  t_1 & 0.5 & 0.5 & 0 \\
  t_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
  t & m & b \\
  t_1 & 0.25 & 0.25 & 0 \\
  t_2 & 0.25 & 0.25 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
  t & m & b \\
  t_1 & 0.5 & 0 & 0 \\
  t_2 & 0 & 0 & 0.5 \\
\end{array}
\]

- \( l \) and \( c \) are rational for both \( s_1 \) and \( s_2 \).
- \( l \) is the only rational action for \( s_3 \).
$\lambda_2(s_1)$

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$\lambda_2(s_2)$

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
</tbody>
</table>

$\lambda_2(s_3)$

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- $l$ and $c$ are rational for both $s_1$ and $s_2$.
- $l$ is the only rational action for $s_3$.
- Whatever her type, it is never rational to play $r$ for 2.
$\lambda_1(t_1)$

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$t$ and $m$ are rational for $t_1$.

$m$ and $b$ are rational for $t_2$.

$\lambda_1(t_1)$

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$\lambda_1(t_2)$

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>3,3</td>
<td>1,1</td>
<td>0,0</td>
</tr>
<tr>
<td>m</td>
<td>1,1</td>
<td>3,3</td>
<td>1,0</td>
</tr>
<tr>
<td>b</td>
<td>0,4</td>
<td>0,0</td>
<td>4,0</td>
</tr>
</tbody>
</table>

**$\lambda_1(t_1)$**

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**$\lambda_1(t_2)$**

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$t$ and $m$ are rational for $t_1$. 
RCBR in the Matrix

$t$ and $m$ are rational for $t_1$. 

\[ \lambda_1(t_1) \]

\[ \begin{array}{ccc}
    l & c & r \\
    s_1 & 0.5 & 0.5 & 0 \\
    s_2 & 0 & 0 & 0 \\
    s_3 & 0 & 0 & 0 \\
\end{array} \]

\[ \lambda_1(t_2) \]

\[ \begin{array}{ccc}
    l & c & r \\
    s_1 & 0 & 0.5 & 0 \\
    s_2 & 0 & 0 & 0.5 \\
    s_3 & 0 & 0 & 0 \\
\end{array} \]
RCBR in the Matrix

\[ \lambda_1(t_1) \]

<table>
<thead>
<tr>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>s_1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>s_2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>s_3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_1(t_2) \]

<table>
<thead>
<tr>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>s_1</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>s_2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>s_3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- \( t \) and \( m \) are rational for \( t_1 \).
- \( m \) and \( b \) are rational for \( t_2 \).
**RCBR in the Matrix**

<table>
<thead>
<tr>
<th>( \lambda_2(s_1) )</th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \lambda_2(s_2) )</th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \lambda_2(s_3) )</th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

All of 2's types believe that 1 is rational.
RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$\lambda_2(s_1)$

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
</tbody>
</table>

$\lambda_2(s_2)$

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

$\lambda_2(s_3)$

All of 2’s types believe that 1 is rational.
RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$\lambda_1(t_1)$</strong></td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$\lambda_1(t_2)$</strong></td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Type $t_1$ of 1 believes that 2 is rational.

But type $t_2$ doesn't! (1/2 probability that 2 is playing r.)
RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

▶ Type $t_1$ of 1 believes that 2 is rational.
Type $t_1$ of 1 believes that 2 is rational.
But type $t_2$ doesn’t! (1/2 probability that 2 is playing $r$.)

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
RCBR in the Matrix

\[ \lambda_2(s_1) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_2) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_3) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>
RCBR in the Matrix

\[ \lambda_2(s_1) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_2) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \lambda_2(s_3) \]

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>m</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- Only type \( s_1 \) of 2 believes that 1 is rational and that 1 believes that 2 is also rational.
RCBR in the Matrix

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
RCBR in the Matrix

$\lambda_1(t_1)$

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$\lambda_1(t_2)$

<table>
<thead>
<tr>
<th></th>
<th>l</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$s_2$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- Type $t_1$ of 1 believes that 2 is rational and that 2 believes that 1 believes that 2 is rational.
No further iteration of mutual belief in rationality eliminates some types or strategies. So at all the states in \( \{(t_1, s_1)\} \times \{(t, m)\} \times \{(l, c)\} \) we have rationality and common belief in rationality. But observe that \( \{(t, m)\} \times \{(l, c)\} \) is precisely the set of profiles that survive IESDS.

\[
\begin{array}{c|ccc}
\lambda_1(t_1) & \text{l} & \text{c} & \text{r} \\
\hline
s_1 & 0.5 & 0.5 & 0 \\
s_2 & 0 & 0 & 0 \\
s_3 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{c|ccc}
\lambda_2(s_1) & \text{t} & \text{m} & \text{b} \\
\hline
t_1 & 0.5 & 0.5 & 0 \\
t_2 & 0 & 0 & 0 \\
\end{array}
\]
No further iteration of mutual belief in rationality eliminate some types or strategies.
No further iteration of mutual belief in rationality eliminate some types or strategies.

So at all the states in \( \{(t_1, s_1)\} \times \{t, m\} \times \{l, c\} \) we have rationality and common belief in rationality.
No further iteration of mutual belief in rationality eliminate some types or strategies.

So at all the states in \( \{(t_1, s_1)\} \times \{t, m\} \times \{l, c\} \) we have rationality and common belief in rationality.

But observe that \( \{t, m\} \times \{l, c\} \) is precisely the set of profiles that survive IESDS.
The general result: RCBR ⇒ IESDS

Suppose that $G$ is a strategic game and $\mathcal{T}$ is any type space for $G$. If $(s, t)$ is a state in $\mathcal{T}$ in which all the players are rational and there is common belief of rationality, then $s$ is a strategy profile that survives iteratively removal of strictly dominated strategies.


Proof: RCBR $\Rightarrow$ IESDS

- We show by induction on $n$ that if the players have $n$-level of mutual belief in rationality then they do not play strategies that would be eliminated at the $n + 1^{th}$ round of IESDS.
Proof: $\text{RCBR} \Rightarrow \text{IESDS}$

- We show by induction on $n$ that if the players have $n$-level of mutual belief in rationality then they do not play strategies that would be eliminated at the $n + 1^{th}$ round of IESDS.

- Basic case, $n = 0$. All the players are rational. We know that a strictly dominated strategy, i.e. one that would be eliminated in the 1st round of IESDS, is never a best response. So no player is playing such a strategy.
Proof: RCBR $\Rightarrow$ IESDS

- We show by induction on $n$ that if the players have $n$-level of mutual belief in rationality then they do not play strategies that would be eliminated at the $n+1^{th}$ round of IESDS.

- Basic case, $n = 0$. All the players are rational. We know that a strictly dominated strategy, i.e. one that would be eliminated in the 1st round of IESDS, is never a best response. So no player is playing such a strategy.

- Inductive step. Suppose that it is mutual belief up to degree $n^{th}$ that all players are rational.
Proof: RCBR \Rightarrow IESDS

- We show by induction on $n$ that if the players have $n$-level of mutual belief in rationality then they do not play strategies that would be eliminated at the $n + 1^{th}$ round of IESDS.

- Basic case, $n = 0$. All the players are rational. We know that a strictly dominated strategy, i.e. one that would be eliminated in the 1st round of IESDS, is never a best response. So no player is playing such a strategy.

- Inductive step. Suppose that it is mutual belief up to degree $n^{th}$ that all players are rational. Take any strategy $s_i$ of an agent $i$ that would not survive $n + 1$ round of IESDS. This strategy is never a best response to a belief whose support is included in the set of states where the others play strategies that would not survive $n^{th}$ round of IESDS. But by our IH this is precisely the kind of belief that all $i$’s type have by IH, so $i$ is not playing $s_i$ either.
“Converse direction” From IESDS to RCBR

Given any strategy profile that survives IESDS, there is a model in and a state in that model where this profile RCBR holds at that state.
“Converse direction” From IESDS to RCBR

Given any strategy profile that survives IESDS, there is a model in and a state in that model where this profile RCBR holds at that state.

▶ Trivial? Mathematically, yes.
“Converse direction” From IESDS to RCBR

Given any strategy profile that survives IESDS, there is a model in
and a state in that model where this profile RCBR holds at that
state.

▶ Trivial? Mathematically, yes.
▶ ... but conceptually important. One can always view or
interpret the choice of a strategy profile that would survive the
iterative elimination procedure as one that results from RCBR.
“Converse direction” From IESDS to RCBR

Given any strategy profile that survives IESDS, there is a model in and a state in that model where this profile RCBR holds at that state.

▶ Trivial? Mathematically, yes.
▶ ... but conceptually important. One can always view or interpret the choice of a strategy profile that would survive the iterative elimination procedure as one that results from RCBR.

Is the entire set of strategy profiles that survive IESDS always consistent with rationality and common belief in rationality? Yes.

▶ For any game G, there is a type structure for that game in which the strategy profiles consistent with rationality and common belief in rationality is the set of strategies that survive iterative removal of strictly dominated strategies.

Subgames

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ be an arbitrary strategic game.
Subgames

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ be an arbitrary strategic game.

A restriction of $H$ is a sequence $G = (G_1, \ldots, G_n)$ such that $G_i \subseteq H_i$ for all $i \in \{1, \ldots, n\}$.

The set of all restrictions of a game $H$ ordered by componentwise set inclusion forms a complete lattice.
Relational models: \( \langle W, R_i \rangle \) where \( R_i \subseteq W \times W \). Write \( R_i(w) = \{ v \mid wR_iv \} \).

Events: \( E \subseteq W \)

Knowledge/Belief: \( \Box E = \{ w \mid R_i(w) \subseteq E \} \)

Common knowledge/belief:
\[
\begin{align*}
\Box^1 E &= \Box E \\
\Box^{k+1} E &= \Box \Box^k E \\
\Box^* E &= \bigcap_{k=1}^{\infty} \Box^k E
\end{align*}
\]

Fact. An event \( F \) is called evident provided \( F \subseteq \Box F \). \( w \in \Box^* E \) provided there is an evident event \( F \) such that \( w \in F \subseteq \Box E \).
Let $G = (G_1, \ldots, G_n)$ be a restriction of a game $H$.

A **knowledge/belief model of** $G$ is a tuple

$\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ where $\langle W, R_1, \ldots, R_n \rangle$ is a knowledge/belief model and $\sigma_i : W \rightarrow G_i$. 


Game Models

Let $G = (G_1, \ldots, G_n)$ be a restriction of a game $H$.

A knowledge/belief model of $G$ is a tuple $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ where $\langle W, R_1, \ldots, R_n \rangle$ is a knowledge/belief model and $\sigma_i : W \rightarrow G_i$.

Given a model $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ for a restriction $G$ and a sequence $\overline{E} = \{E_1, \ldots, E_n\}$ where $E_i \subseteq W$,

$$G_{\overline{E}} = (\sigma_1(E_1), \ldots, \sigma_n(E_n))$$
Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \rightarrow D$ an operator.
- $T$ is monotonic if for all $G, G'$, $G \subseteq G'$ implies $T(G) \subseteq T(G')$
Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \rightarrow D$ an operator.
- $T$ is monotonic if for all $G, G'$, $G \subseteq G'$ implies $T(G) \subseteq T(G')$
- $G$ is a fixed-point if $T(G) = G$

Tarski’s Fixed-Point Theorem: Every monotonic operator $T$ has a (least and largest) fixed point $T\infty = \nu_T = \bigcup \{G | G \subseteq T(G)\}$.
Some Lattice Theory

- $(D, \subseteq)$ is a lattice with largest element $\top$. $T : D \to D$ an operator.
- $T$ is monotonic if for all $G, G'$, $G \subseteq G'$ implies $T(G) \subseteq T(G')$.
- $G$ is a fixed-point if $T(G) = G$.
- $\nu T$ is the largest fixed point of $T$.
Some Lattice Theory

- \( (D, \subseteq) \) is a lattice with largest element \( \top \). \( T : D \to D \) an operator.
- \( T \) is monotonic if for all \( G, G' \), \( G \subseteq G' \) implies \( T(G) \subseteq T(G') \)
- \( G \) is a fixed-point if \( T(G) = G \)
- \( \nu T \) is the largest fixed point of \( T \)
- \( T^\infty \) is the “outcome of \( T \): \( T^0 = \top, T^{\alpha+1} = T(T^\alpha) \), \( T^\beta = \bigcap_{\alpha < \beta} T^\alpha \), The outcome of iterating \( T \) is the least \( \alpha \) such that \( T^{\alpha+1} = T^\alpha \), denoted \( T^\infty \)
Some Lattice Theory

- \((D, \subseteq)\) is a lattice with largest element \(\top\). \(T : D \rightarrow D\) an operator.
- \(T\) is monotonic if for all \(G, G', G \subseteq G'\) implies \(T(G) \subseteq T(G')\)
- \(G\) is a fixed-point if \(T(G) = G\)
- \(\nu T\) is the largest fixed point of \(T\)
- \(T^\infty\) is the “outcome of \(T\): \(T^0 = \top, T^{\alpha+1} = T(T^\alpha), T^\beta = \bigcap_{\alpha < \beta} T^\alpha\), The outcome of iterating \(T\) is the least \(\alpha\) such that \(T^{\alpha+1} = T^\alpha\), denoted \(T^\infty\)
- **Tarski’s Fixed-Point Theorem**: Every monotonic operator \(T\) has a (least and largest) fixed point \(T^\infty = \nu T = \bigcup\{G \mid G \subseteq T(G)\}\).
Some Lattice Theory

- \((D, \subseteq)\) is a lattice with largest element \(\top\). \(T : D \rightarrow D\) an operator.
- \(T\) is monotonic if for all \(G, G'\), \(G \subseteq G'\) implies \(T(G) \subseteq T(G')\)
- \(G\) is a fixed-point if \(T(G) = G\)
- \(\nu T\) is the largest fixed point of \(T\)
- \(T^\infty\) is the “outcome of \(T\): \(T^0 = \top, T^{\alpha+1} = T(T^\alpha)\), \(T^\beta = \bigcap_{\alpha < \beta} T^\alpha\), The outcome of iterating \(T\) is the least \(\alpha\) such that \(T^{\alpha+1} = T^\alpha\), denoted \(T^\infty\)
- **Tarski’s Fixed-Point Theorem**: Every monotonic operator \(T\) has a (least and largest) fixed point \(T^\infty = \nu T = \bigcup\{G \mid G \subseteq T(G)\}\).
- \(T\) is contracting if \(T(G) \subseteq G\). Every contracting operator has an outcome (\(T^\infty\) is well-defined)
Rationality Properties

$\varphi(s_i, G_i, G_{-i})$ holds between a strategy $s_i \in H_i$, a set of strategies $G_i$ for player $i$ and strategies $G_{-i}$ of the opponents. Intuitively $s_i$ is $\varphi$-optimal strategy for player $i$ in the restricted game $\langle G_i, G_{-i}, u_1, \ldots, u_n \rangle$ (where the payoffs are suitably restricted).
Rationality Properties

\( \varphi(s_i, G_i, G_{-i}) \) holds between a strategy \( s_i \in H_i \), a set of strategies \( G_i \) for player \( i \) and strategies \( G_{-i} \) of the opponents. Intuitively \( s_i \) is \( \varphi \)-optimal strategy for player \( i \) in the restricted game \( \langle G_i, G_{-i}, u_1, \ldots, u_n \rangle \) (where the payoffs are suitably restricted).

\( \varphi \) is **monotonic** if for all \( G_{-i}, G'_{-i} \subseteq H_{-i} \) and \( s_i \in H_i \)

\[ G_{-i} \subseteq G'_{-i} \] and \( \varphi(s_i, H_i, G_{-i}) \) implies \( \varphi(s_i, H_i, G'_{-i}) \)
Removing Strategies

If $\varphi = (\varphi_1, \ldots, \varphi_n)$, then define $T_\varphi(G) = G'$ where

- $G = (G_1, \ldots, G_n)$, $G' = (G'_1, \ldots, G'_n)$,
- for all $i \in \{1, \ldots, n\}$, $G'_i = \{s_i \in G_i \mid \varphi_i(s_i, H_i, G_{-i})\}$
Removing Strategies

If \( \varphi = (\varphi_1, \ldots, \varphi_n) \), then define \( T_{\varphi}(G) = G' \) where

- \( G = (G_1, \ldots, G_n) \), \( G' = (G'_1, \ldots, G'_n) \),
- for all \( i \in \{1, \ldots, n\} \), \( G'_i = \{ s_i \in G_i \mid \varphi_i(s_i, H_i, G_{-i}) \} \)

\( T_{\varphi} \) is contracting, so it has an outcome \( T_{\varphi}^\infty \)
Removing Strategies

If $\varphi = (\varphi_1, \ldots, \varphi_n)$, then define $T_\varphi(G) = G'$ where

$\triangleright$ $G = (G_1, \ldots, G_n)$, $G' = (G'_1, \ldots, G'_n)$,

$\triangleright$ for all $i \in \{1, \ldots, n\}$, $G'_i = \{s_i \in G_i \mid \varphi_i(s_i, H_i, G_{-i})\}$

$T_{\varphi}$ is contracting, so it has an outcome $T^\infty_{\varphi}$

If each $\varphi_i$ is monotonic, then $\nu T_{\varphi}$ exists and equals $T^\infty_{\varphi}$. 
Rational Play

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ a strategic game and
$\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ a model for $H$.

$\sigma_i(w)$ is the strategy player is using in state $w$.

$G_{R_i(w)}$ is a restriction of $H$ giving $i$’s view of the game.
Rational Play

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ a strategic game and $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ a model for $H$.

$\sigma_i(w)$ is the strategy player is using in state $w$.

$G_{R_i(w)}$ is a restriction of $H$ giving $i$’s view of the game.

Player $i$ is $\varphi_i$-rational in the state $w$ if $\varphi_i(\sigma_i(w), H_i, (G_{R_i(w)})_{-i})$ holds.
Rational Play

Let $H = \langle H_1, \ldots, H_n, u_1, \ldots, u_n \rangle$ a strategic game and $\langle W, R_1, \ldots, R_n, \sigma_1, \ldots, \sigma_n \rangle$ a model for $H$.

$\sigma_i(w)$ is the strategy player is using in state $w$.

$G_{R_i(w)}$ is a restriction of $H$ giving $i$’s view of the game.

Player $i$ is $\varphi_i$-rational in the state $w$ if $\varphi_i(\sigma_i(w), H_i, (G_{R_i(w)})_{-i})$ holds.

$\text{Rat}(\varphi) = \{ w \in W \mid \text{each player is } \varphi_i\text{-rational in } w \}$

$\Box \text{Rat}(\varphi)$
$\Box^* \text{Rat}(\varphi)$
Theorem (Apt and Zvesper).

- Suppose that each $\varphi_i$ is monotonic. Then for all belief models for $H$,
  $$G_{\text{Rat}}(\varphi) \cap B^*(\text{Rat}(\varphi)) \subseteq T^\infty_\varphi$$

- Suppose that each $\varphi_i$ is monotonic. Then for all knowledge models for $H$,
  $$G_{K^*}(\text{Rat}(\varphi)) \subseteq T^\infty_\varphi$$

- For some standard knowledge model for $H$,
  $$T^\infty_\varphi \subseteq G_{K^*}(\text{Rat}(\varphi))$$

**Claim** If each $\varphi_i$ is monotonic, then $G_{\text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)} \subseteq T^\infty$. 
**Claim** If each $\varphi_i$ is monotonic, then $G_{\text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)} \subseteq T^\infty$.

Let $s_i$ be an element of the $i$th component of $G_{\text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)}$: $s_i = \sigma_i(w)$ for some $w \in \text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)$. 
Claim If each $\varphi_i$ is monotonic, then $G_{\text{Rat}(\varphi) \cap \Box^* \text{Rat}(\varphi)} \subseteq T^\infty$.

Let $s_i$ be an element of the $i$th component of $G_{\text{Rat}(\varphi) \cap \Box^* \text{Rat}(\varphi)}$:

$s_i = \sigma_i(w)$ for some $w \in \text{Rat}(\varphi) \cap \Box^* \text{Rat}(\varphi)$

there is an $F$ such that $F \subseteq \Box F$ and

$$w \in F \subseteq \Box \text{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi)\}$$
**Claim** If each $\varphi_i$ is monotonic, then $G_{\text{Rat}(\varphi) \cap \Box^* \text{Rat}(\varphi)} \subseteq T^\infty \varphi$.

Let $s_i$ be an element of the $i$th component of $G_{\text{Rat}(\varphi) \cap \Box^* \text{Rat}(\varphi)}$:

$s_i = \sigma_i(w)$ for some $w \in \text{Rat}(\varphi) \cap \Box^* \text{Rat}(\varphi)$

there is an $F$ such that $F \subseteq \Box F$ and

$$w \in F \subseteq \Box \text{Rat}(\varphi) = \{ v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi) \}$$

**Claim.** $G_{F \cap \text{Rat}(\varphi)}$ is post-fixed point of $T \varphi$

$(G_{F \cap \text{Rat}(\varphi)} \subseteq T \varphi(G_{F \cap \text{Rat}(\varphi)})$.}
Claim If each $\varphi_i$ is monotonic, then $G_{\text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)} \subseteq T_\infty$.

Let $s_i$ be an element of the $i$th component of $G_{\text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)}$:

$s_i = \sigma_i(w)$ for some $w \in \text{Rat}(\varphi) \cap \square^* \text{Rat}(\varphi)$

there is an $F$ such that $F \subseteq \square F$ and

$$w \in F \subseteq \square \text{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi)\}$$

Claim. $G_{F \cap \text{Rat}(\varphi)}$ is post-fixed point of $T_\varphi$

$$(G_{F \cap \text{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \text{Rat}(\varphi)})).$$

Since each $\varphi_i$ is monotonic, $T_\varphi$ is monotonic and by Tarski’s fixed-point theorem, $G_{F \cap \text{Rat}(\varphi)} \subseteq T_\infty$. But $s_i = \sigma_i(w)$ and $w \in F \cap \text{Rat}(\varphi)$, so $s_i$ is the $i$th component in $T_\infty$. 

\[ F \subseteq \Box F \text{ and } w \in F \subseteq \Box \text{Rat}(\varphi) = \{ v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi) \} \]

**Claim.** \( G_{F \cap \text{Rat}(\varphi)} \) is post-fixed point of \( T_\varphi \)
\( (G_{F \cap \text{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \text{Rat}(\varphi)})). \)
$F \subseteq \Box F$ and $w \in F \subseteq \Box \text{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi)\}$

**Claim.** $G_{F \cap \text{Rat}(\varphi)}$ is post-fixed point of $T_\varphi$
$(G_{F \cap \text{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \text{Rat}(\varphi)}))$.

Let $w' \in F \cap \text{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$. 
$F \subseteq \square F$ and $w \in F \subseteq \square \text{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi)\}$

**Claim.** $G_{F \cap \text{Rat}(\varphi)}$ is post-fixed point of $T_\varphi$

($G_{F \cap \text{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \text{Rat}(\varphi)})$).

Let $w' \in F \cap \text{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \text{Rat}(\varphi)$, $\varphi_i(\sigma_i(w')), H_i, (G_{R_i(w)})_{-i})$ holds.
\[ F \subseteq \Box F \text{ and } w \in F \subseteq \Box \text{Rat}(\varphi) = \{ v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi) \} \]

**Claim.** \( G_{F \cap \text{Rat}(\varphi)} \) is post-fixed point of \( T_{\varphi} \)
\((G_{F \cap \text{Rat}(\varphi)} \subseteq T_{\varphi}(G_{F \cap \text{Rat}(\varphi)})\).

Let \( w' \in F \cap \text{Rat}(\varphi) \) and let \( i \in \{1, \ldots, n\} \).

Since \( w' \in \text{Rat}(\varphi), \varphi_i(\sigma_i(w')), H_i, (G_{R_i(w)})_{-i} \) holds.

\( F \) is evident, so \( R_i(w') \subseteq F \). We also have \( R_i(w') \subseteq \text{Rat}(\varphi) \).
\[ F \subseteq \Box F \text{ and } w \in F \subseteq \Box \text{Rat}(\varphi) = \{ v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi) \} \]

**Claim.** \( G_{F \cap \text{Rat}(\varphi)} \) is post-fixed point of \( T_{\varphi} \)
\((G_{F \cap \text{Rat}(\varphi)} \subseteq T_{\varphi}(G_{F \cap \text{Rat}(\varphi)})\).

Let \( w' \in F \cap \text{Rat}(\varphi) \) and let \( i \in \{1, \ldots, n\} \).

Since \( w' \in \text{Rat}(\varphi) \), \( \varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i}) \) holds.

\( F \) is evident, so \( R_i(w') \subseteq F \). We also have \( R_i(w') \subseteq \text{Rat}(\varphi) \).

Hence, \( R_i(w') \subseteq F \cap \text{Rat}(\varphi) \).
$$F \subseteq \Box F \text{ and } w \in F \subseteq \Box \text{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi)\}$$

**Claim.** $G_{F \cap \text{Rat}(\varphi)}$ is post-fixed point of $T_{\varphi}$

($G_{F \cap \text{Rat}(\varphi)} \subseteq T_{\varphi}(G_{F \cap \text{Rat}(\varphi)})$).

Let $w' \in F \cap \text{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \text{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

$F$ is evident, so $R_i(w') \subseteq F$. We also have $R_i(w') \subseteq \text{Rat}(\varphi)$.

Hence, $R_i(w') \subseteq F \cap \text{Rat}(\varphi)$.

This implies $(G_{R_i(w')}) \subseteq (G_{F \cap \text{Rat}(\varphi)})_{-i}$, and so by monotonicity of $\varphi_i$, $\varphi_i(s_i, H_i, (G_{F \cap \text{Rat}(\varphi)})_{-i})$ holds.
$F \subseteq \Box F$ and $w \in F \subseteq \Box \text{Rat}(\varphi) = \{v \in W \mid \forall i \ R_i(v) \subseteq \text{Rat}(\varphi)\}$

**Claim.** $G_{F \cap \text{Rat}(\varphi)}$ is post-fixed point of $T_\varphi$

$(G_{F \cap \text{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \text{Rat}(\varphi)}))$.

Let $w' \in F \cap \text{Rat}(\varphi)$ and let $i \in \{1, \ldots, n\}$.

Since $w' \in \text{Rat}(\varphi)$, $\varphi_i(\sigma_i(w'), H_i, (G_{R_i(w)})_{-i})$ holds.

$F$ is evident, so $R_i(w') \subseteq F$. We also have $R_i(w') \subseteq \text{Rat}(\varphi)$.

Hence, $R_i(w') \subseteq F \cap \text{Rat}(\varphi)$.

This implies $(G_{R_i(w')}) \subseteq (G_{F \cap \text{Rat}(\varphi)})_{-i}$, and so by monotonicity of $\varphi_i$, $\varphi_i(s_i, H_i, (G_{F \cap \text{Rat}(\varphi)})_{-i})$ holds.

This means $G_{F \cap \text{Rat}(\varphi)} \subseteq T_\varphi(G_{F \cap \text{Rat}(\varphi)})$. 
$sd_i(s_i, G_i, G_{-i})$ is $\neg \exists s'_i \in G_i, \forall s_{-i} \in G_{-i} u_i(s'_i, s_{-i}) > u_i(s_i, s_{-i})$
\[sd_i(s_i, G_i, G_{-i}) \text{ is } \neg \exists s'_i \in G_i, \forall s_{-i} \in G_{-i} u_i(s'_i, s_{-i}) > u_i(s_i, s_{-i})\]

\[br_i(s_i, G_i, G_{-i}) \text{ is } \exists \mu_i \in B_i(G_{-i}) \forall s'_i \in G_i, U_i(s_i, \mu_i) \geq U_i(s'_i, \mu_i).\]
\[sd_i(s_i, G_i, G_{-i}) \text{ is } \neg \exists s'_i \in G_i, \forall s_{-i} \in G_{-i} u_i(s'_i, s_{-i}) > u_i(s_i, s_{-i})\]

\[br_i(s_i, G_i, G_{-i}) \text{ is } \exists \mu_i \in B_i(G_{-i}) \forall s'_i \in G_i, U_i(s_i, \mu_i) \geq U_i(s'_i, \mu_i).\]

\[U_\varphi(G) = G' \text{ where } G'_i = \{s_i \in G_i \mid \varphi_i(s_i, G_i, G_{-i})\}.\]
\[ sd_i(s_i, G_i, G_{-i}) \text{ is } \neg \exists s'_i \in G_i, \forall s_{-i} \in G_{-i} u_i(s'_i, s_{-i}) > u_i(s_i, s_{-i}) \]

\[ br_i(s_i, G_i, G_{-i}) \text{ is } \exists \mu_i \in B_i(G_{-i}) \forall s'_i \in G_i, U_i(s_i, \mu_i) \geq U_i(s'_i, \mu_i). \]

\[ U_\varphi(G) = G' \text{ where } G'_i = \{ s_i \in G_i \mid \varphi_i(s_i, G_i, G_{-i}) \}. \]

Note: \( U_\varphi \) is not monotonic.
**Corollary.** For all belief models, \( G_{\text{Rat}(br) \cap \Box^* \text{Rat}(br)} \subseteq U_{sd}^\infty \). For all \( G \), we have

\[
T_{br}(G) \subseteq T_{sd}(G)
\]

\[
T_{sd}(G) \subseteq U_{sd}(G)
\]

Then, \( T_{sd}^\infty \subseteq U_{sd}^\infty \).
**Corollary.** For all belief models, $G_{\text{Rat}(br) \cap \Box^* \text{Rat}(br)} \subseteq U_{sd}$. For all $G$, we have

$$T_{br}(G) \subseteq T_{sd}(G)$$

$$T_{sd}(G) \subseteq U_{sd}(G)$$

Then, $T_{sd}^\infty \subseteq U_{sd}^\infty$.

**Fact.** Consider two operators $T_1$, $T_2$ on $(D, \subseteq)$ such that,

- for all $G$, $T_1(G) \subseteq T_2(G)$
- $T_1$ is monotonic
- $T_2$ is contracting

Then, $T_1^\infty \subseteq T_2^\infty$. 
This analysis does not work for weak dominance...
Common knowledge of rationality (CKR) in the tree.
Invented by Zermelo, Backwards Induction is an iterative algorithm for “solving” and extensive game.
RCK in the tree

Eric Pacuit and Olivier Roy
RCK in the tree

A

B

(1, 0) (2, 3) (1, 5) (3, 1) (4, 4)

Eric Pacuit and Olivier Roy
RCK in the tree
RCK in the tree

[(1, 0), (2, 3), (1, 5), (4, 4)]
RCK in the tree

The image contains a tree diagram with nodes labeled as follows:

- Node $A$: $(2, 3)$
  - Left subtree: $(1, 0)$
  - Right subtree: $(2, 3)$

- Node $B$: $(1, 5)$
  - Left subtree: $(4, 4)$
    - Sub-subtree 1: $(3, 1)$
    - Sub-subtree 2: $(4, 4)$

The tree structure is as follows: $A$ is connected to $B$, with $B$ having two subtrees: one to the left labeled $(1, 5)$ and one to the right labeled $(4, 4)$. The left subtree of $B$ has two nodes: $(3, 1)$ and $(4, 4)$. The left subtree of $A$ has two nodes: $(1, 0)$ and $(2, 3)$.
RCK in the tree

- **Node A**
  - (2, 3)
  - Subtrees:
    - (1, 0)
    - (2, 3)
    - (1, 5)
- **Node B**
  - (4, 4)
  - Subtrees:
    - (3, 1)
    - (4, 4)
RCK in the tree
RCK in the tree

A

(2, 3)

(1, 0) (2, 3)

(1, 5)

(4, 4)

(1, 5)

(3, 1) (4, 4)
RCK in the tree

(2, 3)

(1, 0)  (2, 3)  (1, 5)

(2, 3)

(1, 5)

(3, 1)  (4, 4)

(4, 4)
Bl Puzzle

A → B → A

R1 → r → R2

D1 → d → D2

(2,1) → (1,6) → (7,5) → (6,6)
Bl Puzzle

![Diagram of a puzzle with nodes A, B, and A, connected by edges R1, r, and R2, with tuples (2,1), (1,6), and (7,5)]

RCK in the tree
RCK in the tree

Bl Puzzle

\[ A \quad R1 \quad B \quad r \quad (7,5) \]

\[ D1 \quad d \quad (2,1) \quad (1,6) \]
RCK in the tree

Bl Puzzle

A \quad R1 \quad B \quad r \quad (7,5)

D1 \quad d

(2,1) \quad (1,6)
RCK in the tree

Bl Puzzle

![Diagram](attachment:Bl_Puzzle_diagram.png)
RCK in the tree

Bl Puzzle

![Diagram of a puzzle with nodes A and D1, and edges labeled R1 and D1, with coordinates (1,6) and (2,1).]
RCK in the tree

Bl Puzzle

$A$

$D1$

(2,1)
BI Puzzle
But what if Bob has to move?

Either she doesn’t believe that he is rational and that he believes that she would choose $R_2$.  

Either way, rationality is not “common knowledge”.

Eric Pacuit and Olivier Roy
What should Bob think of Ann?

- Either she doesn’t believe that he is rational and that he believes that she would choose $R_2$.
- Or Ann made a “mistake” (= irrational move) at the first turn.

Either way, rationality is not “common knowledge”.


Models of Extensive Games

Let $\Gamma$ be a *non-degenerate* extensive game with perfect information. Let $\Gamma_i$ be the set of nodes controlled by player $i$. 

**M**($\Gamma$) = $\langle$ $W$, $\sim_i$, $\sigma$ $\rangle$ where $\sigma: W \to \text{Strat}(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation. 

If $\sigma(w) = \sigma$, then $\sigma_i(w) = \sigma_i$ and $\sigma_{-i}(w) = \sigma_{-i}$.

References: Eric Pacuit and Olivier Roy
Models of Extensive Games

Let $\Gamma$ be a non-degenerate extensive game with perfect information. Let $\Gamma_i$ be the set of nodes controlled by player $i$.

A strategy profile $\sigma$ describes the choice for each player $i$ at all vertices where $i$ can choose.
Models of Extensive Games

Let $\Gamma$ be a *non-degenerate* extensive game with perfect information. Let $\Gamma_i$ be the set of nodes controlled by player $i$.

A strategy profile $\sigma$ describes the choice for each player $i$ at all vertices where $i$ can choose.

Given a vertex $v$ in $\Gamma$ and strategy profile $\sigma$, $\sigma$ specifies a unique path from $v$ to an end-node.
Models of Extensive Games

Let $\Gamma$ be a non-degenerate extensive game with perfect information. Let $\Gamma_i$ be the set of nodes controlled by player $i$.

A strategy profile $\sigma$ describes the choice for each player $i$ at all vertices where $i$ can choose.

Given a vertex $v$ in $\Gamma$ and strategy profile $\sigma$, $\sigma$ specifies a unique path from $v$ to an end-node.

$\mathcal{M}(\Gamma) = \langle W, \sim_i, \sigma \rangle$ where $\sigma: W \to Strat(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation.
Let $\Gamma$ be a *non-degenerate* extensive game with perfect information. Let $\Gamma_i$ be the set of nodes controlled by player $i$.

A strategy profile $\sigma$ describes the choice for each player $i$ at all vertices where $i$ can choose.

Given a vertex $v$ in $\Gamma$ and strategy profile $\sigma$, $\sigma$ specifies a unique path from $v$ to an end-node.

$\mathcal{M}(\Gamma) = \langle W, \sim_i, \sigma \rangle$ where $\sigma : W \rightarrow \text{Strat}(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation.

If $\sigma(w) = \sigma$, then $\sigma_i(w) = \sigma_i$ and $\sigma_{-i}(w) = \sigma_{-i}$
Models of Extensive Games

Let $\Gamma$ be a non-degenerate extensive game with perfect information. Let $\Gamma_i$ be the set of nodes controlled by player $i$.

A strategy profile $\sigma$ describes the choice for each player $i$ at all vertices where $i$ can choose.

Given a vertex $v$ in $\Gamma$ and strategy profile $\sigma$, $\sigma$ specifies a unique path from $v$ to an end-node.

$\mathcal{M}(\Gamma) = \langle W, \sim_i, \sigma \rangle$ where $\sigma : W \rightarrow Strat(\Gamma)$ and $\sim_i \subseteq W \times W$ is an equivalence relation.

If $\sigma(w) = \sigma$, then $\sigma_i(w) = \sigma_i$ and $\sigma_{-i}(w) = \sigma_{-i}$

(A1) If $w \sim_i w'$ then $\sigma_i(w) = \sigma_i(w')$. 
Rationality

\( h^v_i(\sigma) \) denote “\( i \)’s payoff if \( \sigma \) is followed from node \( v \)”
Rationality

$h_i^v(\sigma)$ denote “$i$’s payoff if $\sigma$ is followed from node $v$”

\[ i \text{ is rational at } v \text{ in } w \text{ provided for all strategies } s_i \neq \sigma_i(w), \\
 h_i^v(\sigma(w')) \geq h_i^v((\sigma_{-i}(w'), s_i)) \text{ for some } w' \in [w]_i. \]
Substantive Rationality

\[
i \text{is substantively rational} \text{ in state } w \text{ if } i \text{ is rational at a vertex } v
\text{ in } w \text{ of every vertex in } v \in \Gamma_i;
\]
Stalnaker Rationality

For every vertex \( v \in \Gamma_i \), if \( i \) were to actually reach \( v \), then what he would do in that case would be rational.
For every vertex $v \in \Gamma_i$, if $i$ were to actually reach $v$, then what he would do in that case would be rational.

$f : W \times \Gamma_i \to W, \ f(w, v) = w'$, then $w'$ is the “closest state to $w$ where the vertex $v$ is reached.
For every vertex $v \in \Gamma_i$, if $i$ were to actually reach $v$, then what he would do in that case would be rational.

$f : W \times \Gamma_i \rightarrow W$, $f(w, v) = w'$, then $w'$ is the “closest state to $w$ where the vertex $v$ is reached.

(F1) $v$ is reached in $f(w, v)$ (i.e., $v$ is on the path determined by $\sigma(f(w, v)))$

(F2) If $v$ is reached in $w$, then $f(w, v) = w$

(F3) $\sigma(f(w, v))$ and $\sigma(w)$ agree on the subtree of $\Gamma$ below $v$
\[ s_1 = (d, a), \quad s_2 = (a, a), \quad s_3 = (a, d), \quad s_4 = (a, a), \quad s_5 = (d, a) \]

\[ W = \{ w_1, w_2, w_3, w_4, w_5 \} \]

A = \{ w_i \} for \( i = 1, 2, 3, 4, 5 \)

B = \{ w_2 \} for \( i = 1, 4, 5 \)

\[ \sigma(w_i) = s_i \]

(2, 2) (1, 1) (0, 0) (3, 3)
RCK in the tree

\[ s_1 = (da, d), \ s_2 = (aa, d), \ s_3 = (ad, d), \ s_4 = (aa, a), \ s_5 = (ad, a) \]
RCK in the tree

\[ W = \{ w_1, w_2, w_3, w_4, w_5 \} \] with \( \sigma(w_i) = s^i \)

\[ [w_i]_A = \{ w_i \} \] for \( i = 1, 2, 3, 4, 5 \)

\[ [w_i]_B = \{ w_i \} \] for \( i = 1, 4, 5 \) and \( [w_2]_B = [w_3]_B = \{ w_2, w_3 \} \)

\[ s^1 = (da, d), \; s^2 = (aa, d), \; s^3 = (ad, d), \; s^4 = (aa, a), \; s^5 = (ad, a) \]
It is common knowledge at $w_1$ that if vertex $v_2$ were reached, Bob would play down.
It is common knowledge at $w_1$ that if vertex $v_2$ were reached, Bob would play down.

$s^1 = (da, d), s^2 = (aa, d), s^3 = (ad, d), s^4 = (aa, a), s^5 = (ad, a)$
RCK in the tree

Bob is not rational at $v_2$ in $w_1$

$s^1 = (da, d), \ s^2 = (aa, d), \ s^3 = (ad, d), \ s^4 = (aa, a), \ s^5 = (ad, a)$
RCK in the tree

\[ s^1 = (da, d), \quad s^2 = (aa, d), \quad s^3 = (ad, d), \quad s^4 = (aa, a), \quad s^5 = (ad, a) \]

Bob is rational at \( v_2 \) in \( w_2 \)
Note that $f(w_1, v_2) = w_2$ and $f(w_1, v_3) = w_4$, so there is common knowledge of S-rationality at $w_1$. 

$s^1 = (da, d), s^2 = (aa, d), s^3 = (ad, d), s^4 = (aa, a), s^5 = (ad, a)$
**Aumann’s Theorem:** If $\Gamma$ is a non-degenerate game of perfect information, then in all models of $\Gamma$, we have $C(A - \text{Rat}) \subseteq BI$.

**Stalnaker’s Theorem:** There exists a non-degenerate game $\Gamma$ of perfect information and an extended model of $\Gamma$ in which the selection function satisfies F1-F3 such that $C(S - \text{Rat}) \nsubseteq BI$. 

Aumann’s Theorem: If $\Gamma$ is a non-degenerate game of perfect information, then in all models of $\Gamma$, we have $C(A - \text{Rat}) \subseteq BI$

Stalnaker’s Theorem: There exists a non-degenerate game $\Gamma$ of perfect information and an extended model of $\Gamma$ in which the selection function satisfies F1-F3 such that $C(S - \text{Rat}) \not\subseteq BI$.

Revising beliefs during play:

“Although it is common knowledge that Ann would play across if $v_3$ were reached, if Ann were to play across at $v_1$, Bob would consider it possible that Ann would play down at $v_3$”
F4. For all players $i$ and vertices $v$, if $w' \in [f(w, v)]_i$; then there exists a state $w'' \in [w]_i$; such that $\sigma(w')$ and $\sigma(w'')$ agree on the subtree of $\Gamma$ below $v$.

**Theorem** (Halpern). If $\Gamma$ is a non-degenerate game of perfect information, then for every extended model of $\Gamma$ in which the selection function satisfies F1-F4, we have $C(S – Rat) \subseteq BI$. Moreover, there is an extend model of $\Gamma$ in which the selection function satisfies F1-F4.

Suppose $w \in C(S - Rat)$. We show by induction on $k$ that for all $w'$ reachable from $w$ by a finite path along the union of the relations $\sim_i$, if $v$ is at most $k$ moves away from a leaf, then $\sigma_i(w)$ is $i$'s backward induction move at $w'$. 

Eric Pacuit and Olivier Roy
Proof of Halpern’s Theorem

- Base case: we are at most 1 move away from a leaf. Suppose $w \in C(S - \text{Rat})$. Take any $w'$ reachable from $w$. 
Proof of Halpern’s Theorem

Base case: we are at most 1 move away from a leaf. Suppose $w \in C(S - Rat)$. Take any $w'$ reachable from $w$. Since $w \in C(S - Rat)$, we know that $w' \in C(S - Rat)$. 
Base case: we are at most 1 move away from a leaf. Suppose \( w \in C(S - Rat) \). Take any \( w' \) reachable from \( w \). Since \( w \in C(S - Rat) \), we know that \( w' \in C(S - Rat) \). So \( i \) must play her BI move at \( f(w', v) \).
Base case: we are at most 1 move away from a leaf. Suppose \( w \in C(S - Rat) \). Take any \( w' \) reachable from \( w \). Since \( w \in C(S - Rat) \), we know that \( w' \in C(S - Rat) \). So \( i \) must play her BI move at \( f(w', v) \). But then by F3 this must also be the case at \( (w', v) \).
Proof of Halpern’s Theorem

Base case: we are at most 1 move away from a leaf. Suppose $w \in C(S - Rat)$. Take any $w'$ reachable from $w$. Since $w \in C(S - Rat)$, we know that $w' \in C(S - Rat)$. So $i$ must play her BI move at $f(w', v)$. But then by F3 this must also be the case at $(w', v)$. 
Proof of Halpern’s Theorem

Suppose $w \in C(S - Rat)$. Take any $w'$ reachable from $w$. Assume, towards contradiction, that $\sigma(w)_i(v) = a$ is not the BI move for player $i$.
Induction step. Suppose $w \in C(S - Rat)$. Take any $w'$ reachable from $w$. Assume, towards contradiction, that $\sigma(w)_i(v) = a$ is not the BI move for player $i$. Since $w$ is also in $C(S - Rat)$, we know by definition $i$ must be rational at $w'' = f(w', v)$. But then, by F3 and our IH, all players play according to the BI solution after $v$ at $w''$. 
Proof of Halpern’s Theorem

- $i$’s rationality at $w''$ means, in particular, that there is a $w_3 \in [w'']_i$ such that

$$h_i^V(\sigma_i(w''), \sigma_{-i}(w_3)) \geq h_i^V((bi_i, \sigma_{-i}(w_3)))$$

for $bi_i$ $i$’s backward induction strategy.
But then by F4 there must exists $w_4 \in [w]_i$ such that $\sigma(w_4)$ is strictly worse than $\sigma(w_3)$ at the same position in the sub-tree starting at $v$. 
But then by F4 there must exists \( w_4 \in [w]_i \) such that \( \sigma(w_4) \) \( \sigma(w_3) \) at the same in the sub-tree starting at \( v \). Since \( w_4 \) is reachable from \( w \), in that state all players play according to the backward induction after \( v \), and so this is also true of \( w_3 \).
Proof of Halpern’s Theorem

- But then by F4 there must exist $w_4 \in [w]_i$ such that $\sigma(w_4)$ at the same in the sub-tree starting at $v$. Since $w_4$ is reachable from $w$, in that state all players play according to the backward induction after $v$, and so this is also true of $w_3$. But then since the game is non-degenerate, playing something else than $b_i$ must make $i$ strictly worst off at that state, a contradiction.