# ECON813 Game Theory
# Part A: Interactive Reasoning and Choice
## Topic 3 Correct Beliefs

Christian W. Bach

University of Liverpool & EPICENTER

# Nash Equilibrium – Classically

- The **refinement program** of classical game theory builds on the idea of NASH EQUILIBRIUM.
- Intuitively, a tuple of mixed choices (“one per player”) constitutes a NASH EQUILIBRIUM, whenever every mixed choice only assigns positive probability to pure best responses.
- The **refinement program** attempts to add conditions to NASH EQUILIBRIUM thereby further restricting the “surviving choices” with the ultimate objective of a unique solution for every game.

# Nash Equilibrium – Epistemically

- The **epistemic program** interprets NASH EQUILIBRIUM as a tuple of marginal conjectures.
- From the epistemic perspective NASH EQUILIBRIUM imposes rather strong conditions on interactive reasoning, notably a

  **correct beliefs assumption**

  - *Loosely speaking, a player believes his opponents' beliefs only deem possible his belief hierarchy,*
  - *and he also believes his opponents to believe their opponents beliefs' only deem possible their respective belief hierarchies.*
- In terms of belief hierarchies these **psychological conditions** can be represented by the rather vivid yet technical notion of

  **simple belief hierarchy**

  - *Intuitively, a player's entire belief hierarchy is spanned by a unique marginal conjecture per player.*

# Illustration

|  |  | *Bob* |  |
|---|---|---|---|
|  |  | $c$ | $d$ |
| *Alice* | $a$ | $2, 1$ | $0, 0$ |
|  | $b$ | $0, 0$ | $1, 2$ |

- First of all, note that $ISD = \{a, b\} \times \{c, d\}$, i.e. all pure choices can be rationally made under common belief in rationality.

- There exists three Nash Equilibria in this game:
  - $NE_1 = \big(1 \cdot a + 0 \cdot b, 1 \cdot c + 0 \cdot d\big)$ with optimal pure choices $c_{Alice} \in \{a\}$ and $c_{Bob} \in \{c\}$.
  - $NE_2 = \big(\frac{2}{3} \cdot a + \frac{1}{3} \cdot b, \frac{1}{3} \cdot c + \frac{2}{3} \cdot d\big)$ with optimal pure choices $c_{Alice} \in \{a, b\}$ and $c_{Bob} = \{c, d\}$.
  - $NE_3 = \big(0 \cdot a + 1 \cdot b, 0 \cdot c + 1 \cdot d\big)$ with optimal pure choices $c_{Alice} \in \{b\}$ and $c_{Bob} \in \{d\}$.

# Illustration

| | | | |
|---|---|---|---|
| | | *Bob* | |
| | | *c* | *d* |
| *Alice* | *a* | 2, 1 | 0, 0 |
| | *b* | 0, 0 | 1, 2 |

- In epistemic game theory Nash Equilibria are interpreted as tuples of marginal conjectures.
- Note that in two player games marginal conjectures and conjectures are identical, as every player only faces a single opponent.
- In the concrete game above:
  - $NE_1 = (\sigma^{Alice}, \sigma^{Bob})$ thus contains the marginal conjectures $\sigma^{Alice} \in \Delta(C_{Alice})$ about *Alice*´s choices where $\sigma^{Alice}(a) = 1$ and $\sigma^{Bob} \in \Delta(C_{Bob})$ about *Bob*´s choices where $\sigma^{Bob}(c) = 1$.
  - $NE_2 = (\hat{\sigma}^{Alice}, \hat{\sigma}^{Bob})$ thus contains the marginal conjectures $\hat{\sigma}^{Alice} \in \Delta(C_{Alice})$ about *Alice*´s choices where $\hat{\sigma}^{Alice}(a) = \frac{2}{3}$ as well as $\hat{\sigma}^{Alice}(b) = \frac{1}{3}$ and $\hat{\sigma}^{Bob} \in \Delta(C_{Bob})$ about *Bob*´s choices where $\hat{\sigma}^{Bob}(c) = \frac{1}{3}$ as well as $\hat{\sigma}^{Bob}(c) = \frac{2}{3}$.
  - $NE_3 = (\tilde{\sigma}^{Alice}, \tilde{\sigma}^{Bob})$ thus contains the marginal conjectures $\tilde{\sigma}^{Alice} \in \Delta(C_{Alice})$ about *Alice*´s choices where $\tilde{\sigma}^{Alice}(b) = 1$ and $\tilde{\sigma}^{Bob} \in \Delta(C_{Bob})$ about *Bob*´s choices where $\tilde{\sigma}^{Bob}(d) = 1$.

# Illustration

| | | *Bob* | |
|---|---|---|---|
| | | $c$ | $d$ |
| *Alice* | $a$ | $2, 1$ | $0, 0$ |
| | $b$ | $0, 0$ | $1, 2$ |

- So what does Nash Equilibrium mean in terms of reasoning?
- Consider $NE_1 = (\sigma^{Alice}, \sigma^{Bob}) = \left(1 \cdot a + 0 \cdot b, 1 \cdot c + 0 \cdot d\right)$ and suppose that

$$\beta_{Alice} = \sigma^{Bob} \text{ as well as } \beta_{Bob} = \sigma^{Alice}.$$

- Then, *Alice*´s reasoning can be described as follows:
  - *Alice* believes *Bob* to choose $c$.
  - *Alice* believes *Bob* to believe *her* to choose $a$.
  - Thus, *Alice* also believes that *Bob* acts rationally. *(as $c$ is optimal for him with conjecture $a$)*.
  - *Alice* believes *Bob* to believe *her* to believe *him* to choose $c$.
  - Thus, *Alice* also believes *Bob* to believe that *Alice* acts rationally *(as $a$ is optimal for her with conjecture $c$)*.
- Accordingly, the Nash equilibrium $NE_1$ is characterizable from a **one-person perspective** in terms of a sinlge player's – in this case Alice's – interactive thinking.

# The Case of More than Two Players

- With more than two players it no longer holds that the marginal conjectures and conjectures of a player coincide.

- In general, the reasoning side of Nash Equilibrium thus requires further properties beyond the **correct beliefs assumption**.

- **Projective beliefs:** *if a player holds some belief – about an opponent´s choices or beliefs – then he believes all other opponents to also hold this belief.*

- Nash Equilibrium needs two **projective beliefs** conditions:
  - Player $i$ entertains marginal conjecture $\sigma^j$ about every opponent $j \neq i$, and believes every $k \neq j$ does so too.
  - He also believes every opponent $j \neq i$ to entertain marginal conjecture $\sigma^i$ about himself.

# The Case of More than Two Players

- With more than two players Nash Equilibrium assumes that a given player´s mixed choice is optimal against the product measure of the opponents´mixed choices.

- Yet another epistemic property thus needs to be imposed.

- **Independent beliefs:** *a player's belief about some characteristic of all opponents equals the product of his marginal beliefs about each opponent's particular characteristic.*

- Nash Equilibrium needs two **independent beliefs** conditions:
    - Player $i$'s marginal conjectures are independent, i.e. $\beta_i = \bigotimes_{j \in I \setminus \{i\}} \sigma^j$.
    - He also believes the marginal conjectures of every oponent $j \neq i$ to be independent, i.e. $\beta_j = \bigotimes_{k \in I \setminus \{j\}} \sigma^k$.

# Outline

- Simple Belief Hierarchy
- Correct Beliefs
- Nash Equilibrium
- Charaterization

# SIMPLE BELIEF HIERARCHY

# The Idea of Simple Belief Hierarchy

- Let $\Gamma$ be a game with player set $I = \{1, \ldots, n\}$.
- A belief hierarchy of player $i$ is called **simple**, whenever it is entirely generated by some combination of marginal conjectures

$$(\sigma^1, \ldots, \sigma^n)$$

  for all players.
- Every layer in this belief hierarchy is pointing back to elements in $(\sigma^1, \ldots, \sigma^n)$ only:
  - FIRST-ORDER BELIEF: player $i$'s conjecture is given by $\bigotimes_{j \in I \setminus \{i\}} \sigma^j$,
  - SECOND-ORDER BELIEF: player $i$ believes that every opponent $j \in I \setminus \{i\}$ entertains conjecture $\bigotimes_{k \in I \setminus \{j\}} \sigma^j$,
  - THIRD-ORDER BELIEF: player $i$ believes that every opponent $j \in I \setminus \{i\}$ believes that every player $k \in I \setminus \{j\}$ entertains conjecture $\bigotimes_{l \in I \setminus \{k\}} \sigma^j$,
  - etc.

# Formal Definition of Generation

**Definition 1**

Let $\Gamma$ be a game, $(\sigma^1, \ldots, \sigma^n)$ some tuple of marginal conjectures, and $i \in I$ some player. A belief hierarchy of player $i$ is called *generated* by $(\sigma^1, \ldots, \sigma^n)$, if

- player $i$´s conjecture is given by $\bigotimes_{j \in I \setminus \{i\}} \sigma^j$ with marginal conjectures $\sigma^j$ for all $j \in I \setminus \{i\}$,
- player $i$ believes that every opponent $j \in I \setminus \{i\}$ has conjecture $\bigotimes_{k \in I \setminus \{j\}} \sigma^k$ with marginal conjectures $\sigma^k$ for all $k \in I \setminus \{j\}$,
- player $i$ believes that every opponent $j \in I \setminus \{i\}$ believes that every player $k \in I \setminus \{j\}$ has conjecture $\bigotimes_{l \in I \setminus \{k\}} \sigma^l$ with marginal conjectures $\sigma^l$ for all $l \in I \setminus \{k\}$,
- etc.

# Formal Definition of Simple Belief Hierarchy

**Definition 2**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ holds a *simple belief hierarchy*, if $t_i$'s induced belief hierarchy is generated by some tuple $(\sigma^1, \ldots, \sigma^n)$ of marginal conjectures.

# Illustration: A 2-Player Game

|  |  | *Bob* |  |
|---|---|---|---|
|  |  | $c$ | $d$ |
| *Alice* | $a$ | $2, 1$ | $0, 0$ |
|  | $b$ | $0, 0$ | $1, 2$ |

- Consider the following epistemic model
  - $T_{Alice} = \{t_{Alice}, t'_{Alice}, t''_{Alice}\}$ and $T_{Bob} = \{t_{Bob}, t'_{Bob}, t''_{Bob}\}$
  - $b_{Alice}(t_{Alice}) = (c, t_{Bob})$ and $b_{Alice}(t'_{Alice}) = (\frac{1}{3} \cdot c + \frac{2}{3} \cdot d, t'_{Bob})$ and $b_{Alice}(t''_{Alice}) = (d, t''_{Bob})$
  - $b_{Bob}(t_{Bob}) = (a, t_{Alice})$ and $b_{Bob}(t'_{Bob}) = (\frac{2}{3} \cdot a + \frac{1}{3} \cdot b, t'_{Alice})$ and $b_{Bob}(t''_{Bob}) = (b, t''_{Alice})$
- All types in this epistemic model hold a simple belief hierarchy.
- The types $t_{Alice}$ and $t_{Bob}$ form some epistemic counterpart to $NE_1 = \big(1 \cdot a + 0 \cdot b, 1 \cdot c + 0 \cdot d\big)$.
- The types $t'_{Alice}$ and $t'_{Bob}$ form some epistemic counterpart to $NE_2 = \big(\frac{2}{3} \cdot a + \frac{1}{3} \cdot b, \frac{1}{3} \cdot c + \frac{2}{3} \cdot d\big)$.
- The types $t''_{Alice}$ and $t''_{Bob}$ form some epistemic counterpart to $NE_3 = \big(0 \cdot a + 1 \cdot b, 0 \cdot c + 1 \cdot d\big)$.

# Illustration: A 3-Player Game

| | | Bob | |
|---|---|---|---|
| | | *c* | *d* |
| Alice | *a* | 2, 1, 1 | 0, 0, 0 |
| | *b* | 0, 0, 1 | 1, 2, 0 |

*Claire l*

| | | Bob | |
|---|---|---|---|
| | | *c* | *d* |
| Alice | *a* | 0, 0, 0 | 1, 2, 1 |
| | *b* | 2, 1, 0 | 0, 0, 1 |

*Claire r*

- Consider the following epistemic model
  - $T_{Alice} = \{t_{Alice}\}$, $T_{Bob} = \{t_{Bob}\}$, and $T_{Claire} = \{t_{Claire}\}$
  - $b_{Alice}(t_{Alice}) = (c, t_{Bob}) \bigotimes (l, t_{Claire})$
  - $b_{Bob}(t_{Bob}) = (a, t_{Alice}) \bigotimes (l, t_{Claire})$
  - $b_{Claire}(t_{Claire}) = (a, t_{Alice}) \bigotimes (c, t_{Bob})$
- All types in this epistemic model hold a simple belief hierarchy.
- The three types form some epistemic counterpart to the Nash equilibrium

$$(1 \cdot a + 0 \cdot b, 1 \cdot c + 0 \cdot d, 1 \cdot l + 0 \cdot r).$$

# Example: Teaching a Lesson

**Story**

- It is Friday and your teacher announces a surprise exam for next week.
- You must decide on what day you start preparing for the exam.
- In order to pass the exam you must study for at least two days.
- For a perfect exam and a subsequent compliment by your father you need to study for at least six days.
- Passing the exam increases your utility by 5.
- Failing the exam increases the teacher's utility by 5.
- Every day you study decreases your utility by 1, but increases the teacher's utility by 1.
- A compliment by your father increases your utility by 4.

# Example: Teaching a Lesson

| | | Teacher | | | | |
|---|---|---|---|---|---|---|
| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
| | *Sat* | $3,2$ | $2,3$ | $1,4$ | $0,5$ | $3,6$ |
| | *Sun* | $-1,6$ | $3,2$ | $2,3$ | $1,4$ | $0,5$ |
| You | *Mon* | $0,5$ | $-1,6$ | $3,2$ | $2,3$ | $1,4$ |
| | *Tue* | $0,5$ | $0,5$ | $-1,6$ | $3,2$ | $2,3$ |
| | *Wed* | $0,5$ | $0,5$ | $0,5$ | $-1,6$ | $3,2$ |

# Example: Teaching a Lesson

- Consider the following epistemic model

$T_{You} = \{t_{You}^{Sat}, t_{You}^{Sun}, t_{You}^{Mon}, t_{You}^{Tue}, t_{You}^{Wed}\}$ and $T_{Teacher} = \{t_{Teacher}^{Mon}, t_{Teacher}^{Tue}, t_{Teacher}^{Wed}, t_{Teacher}^{Thu}, t_{Teacher}^{Fri}\}$

$$b_{You}(t_{You}^{Sat}) = (Fri, t_{Teacher}^{Fri})$$
$$b_{You}(t_{You}^{Sun}) = (Tue, t_{Teacher}^{Tue})$$
$$b_{You}(t_{You}^{Mon}) = (Wed, t_{Teacher}^{Wed})$$
$$b_{You}(t_{You}^{Tue}) = (Thu, t_{Teacher}^{Thu})$$
$$b_{You}(t_{You}^{Wed}) = (Fri, t_{Teacher}^{Fri})$$

$$b_{Teacher}(t_{Teacher}^{Mon}) = (Sun, t_{You}^{Sun})$$
$$b_{Teacher}(t_{Teacher}^{Tue}) = (Mon, t_{You}^{Mon})$$
$$b_{Teacher}(t_{Teacher}^{Wed}) = (Tue, t_{You}^{Tue})$$
$$b_{Teacher}(t_{Teacher}^{Thu}) = (Wed, t_{You}^{Wed})$$
$$b_{Teacher}(t_{Teacher}^{Fri}) = (Sat, t_{You}^{Sat})$$

- Every type in the epistemic model believes in the opponent's rationality.
- Hence, all types express common belief in rationality.
- As for every choice there is a type for which it is optimal, all choices can be rationally made under CBR.
- However, only the types $t_{you}^{Sat}$ and $t_{you}^{Wed}$ and $t_{Teacher}^{Fri}$ hold a simple belief hierarchy.

# CORRECT BELIEFS

# What does Simple Belief Hierarchy mean Psychologically?

- The notion of simple belief hiearchy is transparent and convenient from an operational perspective.
- However, how can a simple beilef hierarchy be conceived of psychologically, i.e. in terms of interactive thinking?
- In this section a psychological characterization of simple belief hierarchy is given unveiling a

  **correct beliefs assumption**

  as its essence.
- In the case of more than 2 players two further psychological conditions need to be imposed:

  **projective beliefs** as well as **independent beliefs**

# Believing Others to be Correct about One's Beliefs

**Definition 3**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ believes his opponents to be correct about his beliefs, if $t_i$ believes that his opponents believe that his type is $t_i$.

Thus, if a player believes his opponents to hold correct beliefs, then he believes them to be correct about his entire belief hierarchy.

# Correct Beliefs Assumption

**Definition 4**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ satisfies the correct beliefs assumption, if $t_i$ believes his opponents to be correct about his beliefs and believes every opponent to believe his respective opponents to be correct about his beliefs.

Intuitively, the **correct beliefs assumption** imposes two layers of correctness conditions on a type.

# Psychological Characterization of Simple Belief Hierarchy for the 2-Player Case

**Theorem 5**

*Let* $\Gamma$ *be a two player game,* $\mathcal{M}^{\Gamma}$ *some epistemic model of it,* $i \in I$ *some player, and* $t_i \in T_i$ *some type of player* $i$*. The type* $t_i$ *holds a simple belief hierarchy, if and only if,* $t_i$ *satisfies the correct beliefs assumption.*

# Proof for: *If* Direction

- Suppose that type $t_i$ believes his opponent $j$ to be correct about his beliefs, and believes his opponent $j$ to believe that $i$ is correct about $j$'s beliefs too.

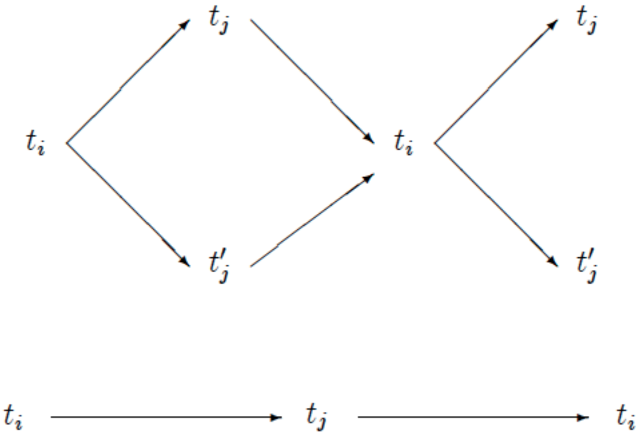
- By Definition 3 it directly follows that $\text{marg}_{T_i} b_j(t_j)(t_i) = 1$ for all $t_j \in \text{supp}\big(\text{marg}_{T_j} b_i(t_i)\big)$.

- It is now shown that $|\ \text{supp}\big(\text{marg}_{T_j} b_i(t_i)\big)\ | = 1$ .
  - Towards a contradiction, suppose that $t_i$ assigns positive probability to at least two distinct types $t_j, t_j' \in T_j$ such that $t_j \neq t_j'$.
  - As $t_i$ believes $j$ to be correct about his beliefs, both types must believe that $i$'s type is $t_i$.
  - In particular, type $t_j$ then believes that $i$ considers it possible that $j$'s type may be $t_j'$.
  - Hence, $t_j$ does not believe that $i$ is correct about his beliefs.
  - But then, as $t_i$ considers possible type $t_j$, it follows that $t_i$ does not believe $j$ to believe that $i$ is correct about $j$'s beliefs, a contradiction.

↯

- Consequently, $|\ \text{supp}\big(\text{marg}_{T_j} b_i(t_i)\big)\ | = 1$ and denote this single type of player $j$ deemed possible by $t_i$ as $t_j$.

# Proof for *If* Direction



# Proof for *If* Direction

- Let $\sigma^j$ be the indcuded conjecture of $t_i$ and $\sigma^i$ the induced conjecture of $t_j$.
- The belief hierarchy of $t_i$ then reads as follows:
  - Type $t_i$ has conjecture $\sigma^j$ *($t_i$'s first-order belief)*
  - As $t_i$ believes that $j$ is of type $t_j$, it follows that $t_i$ believes that $j$ has conjecture $\sigma^i$ *($t_i$'s second-order belief)*
  - As $t_i$ believes that $j$ believes that $i$ is of type $t_i$, it follows that $t_i$ believes that $j$ believes that $i$ has conjecture $\sigma^j$ *($t_i$'s third-order belief)*
  - etc.
- Therefore, type $t_i$'s induced belief hierarchy is generated by $(\sigma^i, \sigma^j)$ and hence simple.

# Proof for *Only If* Direction

- Suppose that $t_i$ holds a simple belief hierarchy generated by $(\sigma^i, \sigma^j)$.
- It follows that:
    - Type $t_i$ does not only have conjecture $\sigma^j$ but also believes that $j$ believes that, **indeed**, $i$'s conjecture is $\sigma^j$.
    - Type $t_i$ does not only believe that $j$ has conjecture $\sigma^i$ but also believes that $j$ believes that, **indeed**, $i$ believes that $j$ has conjecture $\sigma^i$.
    - Type $t_i$ does not only believe that $j$ believes that $i$ has conjecture $\sigma^j$ but also believes that $j$ believes that, **indeed**, $i$ believes that $j$ believes that $i$ has conjecture $\sigma^j$.
    - etc.
- Consequently, type $t_i$ believes $j$ to be correct about his entire belief hierarchy, i.e. type, and hence $t_i$ believes $j$ to be correct about his beliefs.

# Proof for *Only If* Direction

- Next, let $t_j \in T_j$ be some type of player $j$ that $t_i$ considers possible.
- It follows that:
  - Type $t_j$ does not only have conjecture $\sigma^i$ but also believes that $i$ believes that, **indeed**, $j$ has conjecture $\sigma^i$.
  - Type $t_j$ does not only believe that $i$ has conjecture $\sigma^j$ but also believes that $i$ believes that, **indeed**, $j$ believes that $i$ has conjecture $\sigma^j$.
  - Type $t_j$ does not only believe that $i$ believes that $j$ has conjecture $\sigma^i$ but also believes that $i$ believes that, **indeed**, $j$ believes that $i$ believes that $j$ has conjecture $\sigma_i$.
  - etc.
- Consequently, type $t_j$ believes $i$ to be correct about his entire belief hierarchy, i.e. type, and hence $t_j$ believes $i$ to be correct about his beliefs.
- Since this holds for every type $t_j$ considered possible by $t_i$, it follows that type $t_i$ believes $j$ to believe that $i$ is correct about $j$'s beliefs.

# Example: Teaching a Lesson

- Consider the following epistemic model

$$T_{You} = \{t_{You}^{Sat}, t_{You}^{Sun}, t_{You}^{Mon}, t_{You}^{Tue}, t_{You}^{Wed}\} \text{ and } T_{Teacher} = \{t_{Teacher}^{Mon}, t_{Teacher}^{Tue}, t_{Teacher}^{Wed}, t_{Teacher}^{Thu}, t_{Teacher}^{Fri}\}$$

$$b_{You}(t_{You}^{Sat}) = (Fri, t_{Teacher}^{Fri})$$
$$b_{You}(t_{You}^{Sun}) = (Tue, t_{Teacher}^{Tue})$$
$$b_{You}(t_{You}^{Mon}) = (Wed, t_{Teacher}^{Wed})$$
$$b_{You}(t_{You}^{Tue}) = (Thu, t_{Teacher}^{Thu})$$
$$b_{You}(t_{You}^{Wed}) = (Fri, t_{Teacher}^{Fri})$$

$$b_{Teacher}(t_{Teacher}^{Mon}) = (Sun, t_{You}^{Sun})$$
$$b_{Teacher}(t_{Teacher}^{Tue}) = (Mon, t_{You}^{Mon})$$
$$b_{Teacher}(t_{Teacher}^{Wed}) = (Tue, t_{You}^{Tue})$$
$$b_{Teacher}(t_{Teacher}^{Thu}) = (Wed, t_{You}^{Wed})$$
$$b_{Teacher}(t_{Teacher}^{Fri}) = (Sat, t_{You}^{Sat})$$

- Observe that the types $t_{you}^{Sat}$, $t_{you}^{Wed}$, and $t_{Teacher}^{Fri}$ believe the opponent to be correct about his beliefs.

- Moreover, these types all believe that the opponent believes him to be correct about the opponent's beliefs, and consequently satisfy the **correct beliefs assumption**.

- Indeed, recall that $t_{you}^{Sat}$, $t_{you}^{Wed}$, and $t_{Teacher}^{Fri}$ are the only types in this epistemic model to hold a simple belief hierarchy.

# The General Case: More Conditions are Needed

- **Problem**
  - In games with more than two players the **correct beliefs assumption**, no longer implies that a player holds a simple belief hierarchy.
  - In fact, player $i$ might believe that opponent $j$ holds a marginal conjecture about a third player $k$ **distinct** from $i$'s marginal conjecture about $k$.
- **Remedy:** **projective beliefs assumption**

**Definition 6**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ holds projective beliefs, if for every opponent $j \in I \setminus \{i\}$ type $t_i$ believes every player $k \in I \setminus \{i, j\}$ to hold the same marginal belief hierarchy as himself about player $j$.

**Definition 7**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ satisfies the projective beliefs assumption, if $t_i$ holds projective beliefs and believes every opponent to hold projective beliefs too.

# The General Case: More Conditions are Needed

- **Problem**
  - In games with more than two players the **correct beliefs assumption**, no longer implies that a player holds a simple belief hierarchy.
  - In fact, player $i$ marginal conjectures about some opponents $j$ and $k$ might be **correlated**.

- **Remedy: independent beliefs assumption**

**Definition 8**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ holds independent beliefs, if his marginal conjectures are stochastically independent.

**Definition 9**

Let $\Gamma$ be a game, $\mathcal{M}^{\Gamma}$ some epistemic model of it, $i \in I$ some player, and $t_i \in T_i$ some type of player $i$. The type $t_i$ satisfies the independent beliefs assumption, if $t_i$ holds independet beliefs and believes every opponent to hold independent beliefs too.

# The General Case: Psychological Characterization of Simple Belief Hierarchy

**Theorem 10**

*Let* $\Gamma$ *be a game,* $\mathcal{M}^{\Gamma}$ *some epistemic model of it,* $i \in I$ *some player, and* $t_i \in T_i$ *some type of player* $i$*. The type* $t_i$ *holds a simple belief hierarchy, if and only if,* $t_i$ *satisfies the correct beliefs assumption, the projective beliefs assumption, and the independent beliefs assumption.*

# Illustration

| | | *Bob* | |
|---|---|---|---|
| | | *c* | *d* |
| *Alice* | *a* | 2, 1, 1 | 0, 0, 0 |
| | *b* | 0, 0, 1 | 1, 2, 0 |

*Claire l*

| | | *Bob* | |
|---|---|---|---|
| | | *c* | *d* |
| *Alice* | *a* | 0, 0, 0 | 1, 2, 1 |
| | *b* | 2, 1, 0 | 0, 0, 1 |

*Claire r*

- Consider the following epistemic model
    - $T_{Alice} = \{t_{Alice}\}$, $T_{Bob} = \{t_{Bob}\}$, and $T_{Claire} = \{t_{Claire}\}$
    - $b_{Alice}(t_{Alice}) = (c, t_{Bob}) \bigotimes (l, t_{Claire})$
    - $b_{Bob}(t_{Bob}) = (a, t_{Alice}) \bigotimes (l, t_{Claire})$
    - $b_{Claire}(t_{Claire}) = (a, t_{Alice}) \bigotimes (c, t_{Bob})$
- Observe that $t_{Alice}$, $t_{Bob}$, and $t_{Claire}$ all express the **correct beliefs assumption**, the **projective beliefs assumption**, as well as the **independent beliefs assumption**.
- Indeed, recall that all types in this epistemic model hold a simple belief hierarchy.

# Semantics of Simple Belief Hierarchy

- The **essence** of a simple belief hierarchy lies in the

  **correct beliefs assumption**.

- This represents a rather strong restriction on a player's interactive thinking as it excludes that he might **err** about properties external to his mind.

- With more than two players a simple belief hierarchy also *requires* a player to believe his opponents to share his beliefs *as well as* the stochastic independence of his marginal conjectures.

- These conditions are – to say the least – non-trivial too and contexts can be easily envisioned where they are not met.

# NASH EQUILIBRIUM

# Nash Equilibrium

**Definition 11**

Let $\Gamma$ be a game, and $(\sigma^i)_{i \in I} \in \times_{i \in I} \Delta(C_i)$ some tuple of marginal conjectures. The tuple $(\sigma^i)_{i \in I}$ is called Nash equilibrium, if for all $i \in I$ the marginal conjecture $\sigma^i$ only assigns positive probability to choices $c_i \in C_i$ such that $c_i$ is optimal given the product conjecture $\bigotimes_{j \in I \setminus \{i\}} \sigma^j$.

# Intelligibility

**Theorem 12 (Nash, 1950)**

*Let $\Gamma$ be a game. There exists a Nash equilibrium.*

# Rational Choice under Nash Equilibrium

**Definition 13**

Let $\Gamma$ be a game, $i \in I$ some player, and $c_i \in C_i$ some choice of player $i$. The choice $c_i$ is rational under Nash equilibrium, if there exists a Nash equilibrium $(\sigma^j)_{j \in I}$ such that $c_i$ is optimal given the product conjecture $\bigotimes_{j \in I \setminus \{i\}} \sigma^j$.

# Rational Choice under Nash Equilibrium

- If a choice $c_i$ receives positive probability in some Nash equilibrium $(\sigma^j)_{j \in I}$, then it is also rational under Nash equilibrium.
- Indeed, by Nash equilibrium itselt (cf. Definition 11) it is already ensured that $c_i$ is optimal given $\bigotimes_{j \in I \setminus \{i\}} \sigma^j$.
- However, if a choice $c_i$ is ratinonal under Nash equilibrium, then it does not always receive positive probability in some Nash equilibrium.
- Indeed, consider the following game:

| | | *Bob* | |
|---|---|---|---|
| | | *c* | *d* |
| *Alice* | *a* | 2, 0 | 0, 1 |
| | *b* | 1, 0 | 1, 0 |

- The tuple $(b, \frac{1}{2} \cdot c + \frac{1}{2} \cdot d)$ constitutes a Nash equilibrium and $a$ is optimal given the conjecture $\frac{1}{2} \cdot c + \frac{1}{2} \cdot d$ thus qualifying as rational under Nash equilibrium.
- However, there exists no other Nash equilibrium $(\sigma^{Alice}, \sigma^{Bob})$ such that $\sigma^{Alice}(a) > 0$.
- Towards a contradiction suppose that $\sigma^{Alice}(a) > 0$.
- Then, only $d$ is optimal for *Bob* and thus $\sigma^{Bob}(d) = 1$ which in turn implies only $b$ to be optimal for *Alice* yielding $\sigma^{Alice}(b) = 1$, a contradiction.

$\lightning$

# Nash Equilibrium Method

## Nash Equilibrium Method

Let $\Gamma$ be a game, and $i \in I$ some player.

- **Step 1:** Compute all Nash equilibria of $\Gamma$.
- **Step 2:** For every Nash equilibrium $(\sigma^j)_{j \in I}$ found in **Step 1**, determine all choices of $i$ that are optimal given the product conjecture $\bigotimes_{j \in I \backslash \{i\}} \sigma^j$.

The choices selected by **Step 2** are the choices of player $i$ that are rational under Nash equilibrium.

However, there exists no simple algorithm to identify all Nash equilibria of a given game.

# Illustration

| | | *Bob* | |
|---|---|---|---|
| | | *c* | *d* |
| *Alice* | *a* | 2, 1 | 0, 0 |
| | *b* | 0, 0 | 1, 2 |

- Recall that the Nash equilibria of this game are as follows:
  - $NE_1 = \left(1 \cdot a + 0 \cdot b, 1 \cdot c + 0 \cdot d\right)$ with optimal pure choices $c_{Alice} \in \{a\}$ and $c_{Bob} \in \{c\}$.
  - $NE_2 = \left(\frac{2}{3} \cdot a + \frac{1}{3} \cdot b, \frac{1}{3} \cdot c + \frac{2}{3} \cdot d\right)$ with optimal pure choices $c_{Alice} \in \{a, b\}$ and $c_{Bob} = \{c, d\}$.
  - $NE_3 = \left(0 \cdot a + 1 \cdot b, 0 \cdot c + 1 \cdot d\right)$ with optimal pure choices $c_{Alice} \in \{b\}$ and $c_{Bob} \in \{d\}$.
- Consequently, for both players all choices are rational under Nash equilibrium.

# Illustration

| | | *Bob* | |
|---|---|---|---|
| | | *c* | *d* |
| *Alice* | *a* | $2, 1$ | $0, 0$ |
| | *b* | $0, 0$ | $1, 2$ |

- It is now shown that there are no other Nash equilibria.
- Suppose that $(\sigma^{Alice}, \sigma^{Bob})$ is a Nash equilibrium.
- **Case 1:** $\sigma^{Alice}(a) = 1$.
  - Then, $a$ must be optimal given $\sigma^{Bob}$.
  - This is only possible if $\sigma^{Bob}(c) \geq \frac{1}{3}$.
  - The choice $c$ is indeed optimal against $\sigma^{Alice}$, yet $d$ is not.
  - Consequently, $\sigma^{Bob}(c) = 1$ and $NE_1 = \big((1,0),(1,0)\big)$ ensues.

# Illustration

| | | *Bob* | |
|---|---|---|---|
| | | $c$ | $d$ |
| *Alice* | $a$ | $2, 1$ | $0, 0$ |
| | $b$ | $0, 0$ | $1, 2$ |

- **Case 2:** $0 < \sigma^{Alice}(a) < 1$.
    - Then, $\sigma^{Alice}(b) > 0$ too.
    - Both $a$ and $b$ must thus be optimal given $\sigma^{Bob}$, i.e.

$$2 \cdot \sigma^{Bob}(c) + 0 \cdot (1 - \sigma^{Bob}(c)) = 0 \cdot \sigma^{Bob}(c) + 1 \cdot (1 - \sigma^{Bob}(c))$$

    - This equation is only satisfied if $\sigma^{Bob}(c) = \frac{1}{3}$.
    - Hence, $\sigma^{Bob}(d) = \frac{2}{3}$.
    - Both $c$ and $d$ must thus be optimal given $\sigma^{Alice}$, i.e.

$$1 \cdot \sigma^{Alice}(a) + 0 \cdot (1 - \sigma^{Alice}(a)) = 0 \cdot \sigma^{Alice}(a) + 2 \cdot (1 - \sigma^{Alice}(a))$$

    - This equation is only satisfied if $\sigma^{Alice}(a) = \frac{2}{3}$.
    - Consequently, $\sigma^{Alice}(a) = \frac{2}{3}$ and $NE_2 = ((\frac{2}{3}, \frac{1}{3}), (\frac{2}{3}, \frac{1}{3}))$ ensues.

# Illustration

| | | | |
|---|---|---|---|
| | | *Bob* | |
| | | $c$ | $d$ |
| *Alice* | $a$ | $2,1$ | $0,0$ |
| | $b$ | $0,0$ | $1,2$ |

- **Case 3:** $\sigma^{Alice}(a) = 0$.
  - Then, $b$ must be optimal given $\sigma^{Bob}$.
  - This is only possible if $\sigma^{Bob}(d) \geq \frac{2}{3}$.
  - The choice $d$ is indeed optimal against $\sigma^{Alice}$, yet $c$ is not.
  - Consequently, $\sigma^{Bob}(d) = 1$ and $NE_3 = \big((0,1),(0,1)\big)$ ensues.
- Therefore, there do not exist any Nash equilibria other than $NE_1$, $NE_2$, and $NE_3$.

# Example: Teaching a Lesson

|  |  | Teacher |  |  |  |  |
|---|---|---|---|---|---|---|
|  |  | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
| You | *Sat* | 3, 2 | 2, 3 | 1, 4 | 0, 5 | 3, 6 |
|  | *Sun* | −1, 6 | 3, 2 | 2, 3 | 1, 4 | 0, 5 |
|  | *Mon* | 0, 5 | −1, 6 | 3, 2 | 2, 3 | 1, 4 |
|  | *Tue* | 0, 5 | 0, 5 | −1, 6 | 3, 2 | 2, 3 |
|  | *Wed* | 0, 5 | 0, 5 | 0, 5 | −1, 6 | 3, 2 |

# Example: Teaching a Lesson

Teacher

| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
|---|---|---|---|---|---|---|
| | *Sat* | 3, 2 | 2, 3 | 1, 4 | 0, 5 | 3, 6 |
| | *Sun* | −1, 6 | 3, 2 | 2, 3 | 1, 4 | 0, 5 |
| You | *Mon* | 0, 5 | −1, 6 | 3, 2 | 2, 3 | 1, 4 |
| | *Tue* | 0, 5 | 0, 5 | −1, 6 | 3, 2 | 2, 3 |
| | *Wed* | 0, 5 | 0, 5 | 0, 5 | −1, 6 | 3, 2 |

- According to Step 1 of the Nash equilibrium method, all Nash equilibria $(\sigma^y, \sigma^T)$ of the game are computed first, where $\sigma^y \in \Delta(C_{you})$ and $\sigma^T \in \Delta(C_{Teacher})$.

- Suppose that $(\sigma^y, \sigma^T)$ is a Nash equilibrium.

- **Step 1:** it is shown that $\sigma^T(Thu) = 0$.
  - Suppose that $\sigma^T(Thu) > 0$.
  - Then, *Thu* must be optimal for the teacher under the conjecture $\sigma^y$.
  - This is only possible, if $\sigma^y(Wed) > 0$; otherwise *Fri* would be strictly better than *Thu* for the teacher.
  - Then, *Wed* must be optimal for you with conjecture $\sigma^T$.
  - Yet, *Wed* is only optimal, if $\sigma^T(Fri) = 1$, otherwise *Sat* would be strictly better than *Wed* for you.
  - **Contradiction!** Hence, $\sigma^T(Thu) = 0$.

# Example: Teaching a Lesson

| | | Teacher | | | | |
|---|---|---|---|---|---|---|
| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
| You | *Sat* | $3, 2$ | $2, 3$ | $1, 4$ | $0, 5$ | $3, 6$ |
| | *Sun* | $-1, 6$ | $3, 2$ | $2, 3$ | $1, 4$ | $0, 5$ |
| | *Mon* | $0, 5$ | $-1, 6$ | $3, 2$ | $2, 3$ | $1, 4$ |
| | *Tue* | $0, 5$ | $0, 5$ | $-1, 6$ | $3, 2$ | $2, 3$ |
| | *Wed* | $0, 5$ | $0, 5$ | $0, 5$ | $-1, 6$ | $3, 2$ |

- **Step 2:** it is shown that $\sigma^T(\mathit{Wed}) = 0$.
  - Suppose that $\sigma^T(\mathit{Wed}) > 0$.
  - Then, $\mathit{Wed}$ must be optimal for the teacher with conjecture $\sigma^y$.
  - This is only possible, if $\sigma^y(\mathit{Tue}) > 0$; otherwise $\mathit{Thu}$ would be strictly better than $\mathit{Wed}$ for the teacher.
  - Then, $\mathit{Tue}$ must be optimal for you with conjecture $\sigma^T$.
  - However, $\mathit{Tue}$ is only optimal, if $\sigma^T(\mathit{Thu}) > 0$; otherwise $\mathit{Sat}$ is strictly better than $\mathit{Tue}$ for you.
  - **Contradiction with Step 1!** Hence, $\sigma^T(\mathit{Wed}) = 0$.

# Example: Teaching a Lesson

| | | | | Teacher | | |
|---|---|---|---|---|---|---|
| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
| | *Sat* | 3, 2 | 2, 3 | 1, 4 | 0, 5 | 3, 6 |
| | *Sun* | −1, 6 | 3, 2 | 2, 3 | 1, 4 | 0, 5 |
| You | *Mon* | 0, 5 | −1, 6 | 3, 2 | 2, 3 | 1, 4 |
| | *Tue* | 0, 5 | 0, 5 | −1, 6 | 3, 2 | 2, 3 |
| | *Wed* | 0, 5 | 0, 5 | 0, 5 | −1, 6 | 3, 2 |

- **Step 3:** it is shown that $\sigma^T(\mathit{Tue}) = 0$.
  - Suppose that $\sigma^T(\mathit{Tue}) > 0$.
  - Then, $\mathit{Tue}$ must be optimal for the teacher with conjecture $\sigma^y$.
  - This is only possible, if $\sigma^y(\mathit{Mon}) > 0$; otherwise $0.9 \cdot \mathit{Wed} + 0.1 \cdot \mathit{Thu}$ would be strictly better than $\mathit{Tue}$ for the teacher.
  - Then, $\mathit{Mon}$ must be optimal for you with conjecture $\sigma^T$.
  - However, $\mathit{Mon}$ is only optimal, if $\sigma^T(\mathit{Wed}) > 0$ or $\sigma^T(\mathit{Thu}) > 0$; otherwise $\mathit{Sat}$ is strictly better than $\mathit{Mon}$ for you.
  - **Contradiction with Step 1 or 2!** Hence, $\sigma^T(\mathit{Tue}) = 0$.

# Example: Teaching a Lesson

| | | Teacher | | | | |
|---|---|---|---|---|---|---|
| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
| | *Sat* | $3, 2$ | $2, 3$ | $1, 4$ | $0, 5$ | $3, 6$ |
| | *Sun* | $-1, 6$ | $3, 2$ | $2, 3$ | $1, 4$ | $0, 5$ |
| You | *Mon* | $0, 5$ | $-1, 6$ | $3, 2$ | $2, 3$ | $1, 4$ |
| | *Tue* | $0, 5$ | $0, 5$ | $-1, 6$ | $3, 2$ | $2, 3$ |
| | *Wed* | $0, 5$ | $0, 5$ | $0, 5$ | $-1, 6$ | $3, 2$ |

- **Step 4:** it is shown that $\sigma^T(Mon) = 0$.
    - Suppose that $\sigma^T(Mon) > 0$.
    - Then, $Mon$ must be optimal for the teacher with conjecture $\sigma^y$.
    - This is only possible, if $\sigma^y(Sun) > 0$; otherwise $0.9 \cdot Tue + 0.09 \cdot Wed + 0.01 \cdot Thu$ would be strictly better than $Mon$ for the teacher.
    - Then, $Sun$ must be optimal for you with conjecture $\sigma^T$.
    - However, $Sun$ is only optimal, if $\sigma^T(Tue) > 0$; otherwise $Mon$ is strictly better than $Sun$ for you.
    - **Contradiction with Step 3!** Hence, $\sigma^T(Mon) = 0$.

# Example: Teaching a Lesson

Teacher

| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
|---|---|---|---|---|---|---|
| | *Sat* | 3, 2 | 2, 3 | 1, 4 | 0, 5 | 3, 6 |
| | *Sun* | −1, 6 | 3, 2 | 2, 3 | 1, 4 | 0, 5 |
| You | *Mon* | 0, 5 | −1, 6 | 3, 2 | 2, 3 | 1, 4 |
| | *Tue* | 0, 5 | 0, 5 | −1, 6 | 3, 2 | 2, 3 |
| | *Wed* | 0, 5 | 0, 5 | 0, 5 | −1, 6 | 3, 2 |

- Therefore, if $(\sigma^y, \sigma^T)$ is a Nash equilibrium, then $\sigma^T$ must assign probability 0 to $Mon$, $Tue$, $Wed$, and $Thu$: hence, $\sigma^T(Fri) = 1$.
- Since your optimal choices with conjecture $\sigma^T = Fri$ are $Sat$ and $Wed$, the conjecture $\sigma^y$ can only assign positive probability to these choices, i.e. $\sigma^y(Sat) + \sigma^y(Wed) = 1$.
- As $\sigma^T(Fri) = 1$, it must be the case that $Fri$ is optimal for the teacher with conjecture $\sigma^y$.
- Note that with conjecture $\sigma^y$, the choice $Thu$ is strictly better than $Mon$, $Tue$, and $Wed$ for the teacher.
- For $Fri$ to be optimal with conjecture $\sigma^y$ it thus needs to hold that , i.e.

  $$u_{Teacher}(Fri, \sigma^y) = \sigma^y(Sat) \cdot 6 + \big(1 - \sigma^y(Sat)\big) \cdot 2 \geq \sigma^y(Sat) \cdot 5 + \big(1 - \sigma^y(Sat)\big) \cdot 6 = u_{Teacher}(Thu, \sigma^y)$$

  which is equivalent to $4 \cdot \sigma^y(Sat) + 2 \geq 6 - \sigma^y Sat$ and thus amounts to $\sigma^y(Sat) \geq 0.8$.

# Example: Teaching a Lesson

| | | Teacher | | | | |
|---|---|---|---|---|---|---|
| | | *Mon* | *Tue* | *Wed* | *Thu* | *Fri* |
| | *Sat* | 3, 2 | 2, 3 | 1, 4 | 0, 5 | 3, 6 |
| | *Sun* | −1, 6 | 3, 2 | 2, 3 | 1, 4 | 0, 5 |
| You | *Mon* | 0, 5 | −1, 6 | 3, 2 | 2, 3 | 1, 4 |
| | *Tue* | 0, 5 | 0, 5 | −1, 6 | 3, 2 | 2, 3 |
| | *Wed* | 0, 5 | 0, 5 | 0, 5 | −1, 6 | 3, 2 |

- Consequently the set of all Nash equilibria of the *Teaching a Lesson* game reads as follows:

$$NE = \{(\alpha \cdot Sat + (1-\alpha) \cdot Wed, 1 \cdot Fri) : 0.8 \leq \alpha \leq 1\}$$

- Then, the choices rational under Nash equilibrium for you are *Sat* and *Wed*, while they are *Thu* and *Fri* for the teacher.

# CHARACTERIZATION

# The “Two-Edged Sword” Again

- The CLASSICAL and EPISTEMIC perspectives are now conjoined.
- The solution concept of Nash equilibrium turns out to be epistemically characterizable by the three earlier introduced
  - **correct beliefs assumption**
  - **projective beliefs assumption**
  - **independent beliefs assumption**

**PLUS** up to 2-fold belief in rationality (“rationality assumption”).

# Rational Choice under Simple Belief Hierarchy

**Definition 14**

Let $\Gamma$ be a game, $i \in I$ some player, and $c_i \in C_i$ some choice of player $i$. The choice $c_i$ is rational under simple belief hierarchy and up to 2-fold belief in raitonality, if there exists an epistemic model $\mathcal{M}^\Gamma$ of $\Gamma$ with some type $t_i \in T_i$ of player $i$ such that

- $t_i$ holds a simple belief hierarchy,
- $t_i$ expresses up to 2-fold belief in rationality,
- $c_i$ is optimal for $t_i$.

# Nash Equilibrium and Simple Belief Hierarchy

**Lemma 15**

*Let* $\Gamma$ *be a game,* $i \in I$ *some player, and* $c_i \in C_i$ *some choice of player* $i$*. The choice* $c_i$ *is rational under simple belief hierarchy and up to 2-fold belief in raitonality, if and only if,* $c_i$ *is rational under Nash equilibrium.*

# Intuition for the *Epistemic Foundation* Direction

- Consider a simple belief hierarchy for player $i$ generated by $(\sigma^j)_{j\in I}$:
  - player $i$'s conjecture is $\bigotimes_{j\in I\setminus\{i\}} \sigma^j$,
  - player $i$ believes that every opponent $j \in I \setminus \{i\}$ has conjecture $\bigotimes_{k\in I\setminus\{j\}} \sigma^k$,
  - player $i$ believes that every opponent $j \in I \setminus \{i\}$ believes that every player $k \in I \setminus \{j\}$ has conjecture $\bigotimes_{l\in I\setminus\{k\}} \sigma^l$,
  - etc.
- Suppose that the belief hierarchy also expresses up to 2-fold belief in rationality.
- Let $j \in I \setminus \{i\}$ and consider $c_j \in C_j$ such that $\sigma_j(c_j) > 0$, i.e. player $i$ assigns positive probability to $c_j$.
- By 1-fold belief in rationality, player $i$ believes that $c_j$ is optimal for $j$ with conjecture $\bigotimes_{k\in I\setminus\{j\}} \sigma^k$ (which player $i$ believes player $j$ to have as conjecture).
- Now, consider $c_i \in C_i$ such that $\sigma^i(c_i) > 0$ and let $j \in I \setminus \{i\}$, i.e. player $i$ believes player $j$ assigns positive probability to $c_i$.
- By 2-fold belief in rationality, player $i$ believes $j$ to believe that $c_i$ is optimal for $i$ with conjceture $\bigotimes_{k\in I\setminus\{i\}} \sigma^k$ (which player $i$ believes player $j$ to believe him to have as conjecture).
- **Conclusion:** if a belief hierarchy of player $i$ is simple – generated by $(\sigma^j)_{j\in I}$, and expresses up to 2-fold belief in rationality, then for all $j \in I$ the marginal conjecture $\sigma_j$ only assigns positive probability to choices $c_j \in C_j$ such that $c_j$ is optimal given $\bigotimes_{k\in I\setminus\{j\}} \sigma^k$, i.e. $(\sigma^j)_{j\in I}$ constitutes a Nash equilibrium.

# Intuition for the *Existence* Direction

- Let $(\sigma_j)_{j \in I}$ be a Nash equilibrium, i.e. for all $j \in I$, the marginal conjecture $\sigma_j$ only assigns positive probability to choices $c_j \in C_j$ such that $c_j$ is optimal given $\bigotimes_{k \in I \setminus \{j\}} \sigma^k$.
- Consider some player $i \in I$ and the simple belief hierarchy for $i$ generated by $(\sigma^j)_{j \in I}$.
- Let $j \in I \setminus \{i\}$ be some opponent of player $i$, and $c_j \in C_j$ such that $\sigma^j(c_j) > 0$.
- Since by Nash equilibrium $c_j$ is optimal for $j$ with conjecture $\bigotimes_{k \in I \setminus \{j\}} \sigma^k$ and by simple belief hierarchy $i$ believes $j$ to have conjecture $\bigotimes_{k \in I \setminus \{j\}} \sigma^k$, it follows that $i$ believes in $j$'s rationally.
- Hence, $i$ expresses 1-fold belief in rationality.
- Now, let $k \in I \setminus \{j\}$ be some player other than $j$, and $c_k \in C_k$ such that $\sigma^k(c_k) > 0$
- Since by Nash equilibrium $c_k$ is optimal for $k$ with conjecture $\bigotimes_{l \in I \setminus \{k\}} \sigma^l$ and by simple belief hierarchy $i$ believes $j$ to have conjecture $\bigotimes_{k \in I \setminus \{j\}} \sigma^k$ as well as $j$ to believe $k$ to have conjecture $\bigotimes_{l \in I \setminus \{k\}} \sigma^l$, it follows that $i$ believes that $j$ believes in $k$'s rationally.
- Hence, $i$ expresses 2-fold belief in rationality.
- **Conclusion:** if a tuple $(\sigma^j)_{j \in I}$ of marginal conjectures constitutes a Nash equilibrium, then a simple belief hierarchy generated by $(\sigma^j)_{j \in I}$ expresses up to 2-fold belief in rationality.

# Rational Choice under 4 Psychological Conditions

**Definition 16**

Let $\Gamma$ be a game, $i \in I$ some player, and $c_i \in C_i$ some choice of player $i$. The choice $c_i$ is rational under correctness, projection, independence and up to 2-fold belief in rationality, if there exists an epistemic model $\mathcal{M}^\Gamma$ of $\Gamma$ with some type $t_i \in T_i$ of player $i$ such that

- $t_i$ satisfies the correct beliefs assumption,
- $t_i$ satisfies the projective beliefs assumption,
- $t_i$ satisfies the independent beliefs assumption,
- $t_i$ expresses up to $2$-fold belief in rationality,
- $c_i$ is optimal for $t_i$.

# Nash Equilibrium Psychologically

**Theorem 17**

*Let* $\Gamma$ *be a game,* $i \in I$ *some player, and* $c_i \in C_i$ *some choice of player* $i$. *The choice* $c_i$ *is rational under correctness, projection, independence and up to 2-fold belief in rationality, if and only if,* $c_i$ *is rational under Nash equilibrium.*

- Reasoning in line with Nash equilibrium thus requires rather substantial conditions to be met by the players.
- In particular, the **correct beliefs assumptions** seems strong.

# Intelligibility

**Corollary 18**

*Let* $\Gamma$ *be a game. There exists an epistemic model* $\mathcal{M}^{\Gamma}$ *of* $\Gamma$ *in which all types satisfy the correct beliefs assumption, the projective beliefs assumption, the independent beliefs assumption, and expresses up to* $2$*-fold belief in rationality.*

# Background Reading

- PEREA, A. (2012): *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press. **Chapter 4.**