Mini-course on Epistemic Game Theory
Lecture 4: Forward Induction Reasoning

Andrés Perea
EpiCenter & Dept. of Quantitative Economics

Maastricht University

Toulouse, June/July 2015

Introduction

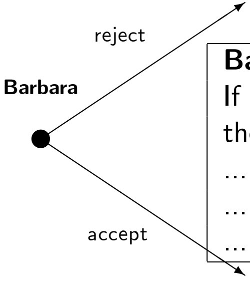
- In the previous chapter, we have discussed the concept of **common belief in future rationality.**
- **Main idea:** Whatever you observe in the game, you always believe that your opponents will choose rationally **from now on.**
- It represents a **backward induction-type** of reasoning.
- It may **not** be the **only plausible way** of reasoning in a dynamic game!

Example: Painting Chris' house

Story

- Chris is planning to paint his house tomorrow, and needs someone to help him.
- You and Barbara are both interested. This evening, both of you must come to Chris' house, and whisper a price in his ear. Price must be either 200, 300, 400 or 500 euros.
- Person with lowest price will get the job. In case of a tie, Chris will toss a coin.
- Before you leave for Chris' house, Barbara gets a phone call from a colleague, who asks her to repair his car tomorrow at a price of 350 euros.
- Barbara must decide whether or not to accept the colleague's offer.

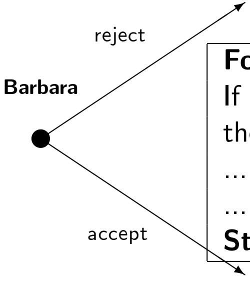
	200	300	400	500
200	100, 100	200, 0	200, 0	200, 0
300	0, 200	150, 150	300, 0	300, 0
400	0, 200	0, 300	200, 200	400, 0
500	0, 200	0, 300	0, 400	250, 250



Backward induction:
If you observe that Barbara has rejected offer, then you believe that
... rejecting offer was a mistake,
... Barbara chooses rationally in subgame,
... Barbara believes that you choose rationally.

You will choose price 200.

	200	300	400	500
200	100, 100	200, 0	200, 0	200, 0
300	0, 200	150, 150	300, 0	300, 0
400	0, 200	0, 300	200, 200	400, 0
500	0, 200	0, 300	0, 400	250, 250



Forward induction:
If you observe that Barbara has rejected offer,
then you believe that
... rejecting offer is part of a **rational strategy,**
... Barbara will choose price 400.
Strong belief in Barbara's rationality.

You will choose price 300.

Strong belief in the opponents' rationality

- **If** at information set $h \in H_i$, it is possible for player i to believe that each of his opponents is implementing a **rational** strategy,

  then player i **must** believe at h that each of his opponents is implementing a **rational** strategy.

- How can we **formalize** this idea within an epistemic model?

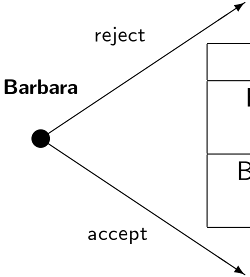
- **Attempt:** Consider an epistemic model M, a type t_i and an information set $h \in H_i$.

  If there is an opponents' strategy-type combination in M where (a) the opponents' strategy combination leads to h, and (b) the strategies are optimal for the types,

  then type t_i must at h only assign positive probability to opponents' strategy-type combinations that satisfy (a) and (b).

- This will **not work!**

200	100, 100	200, 0	200, 0	200, 0
300	0, 200	150, 150	300, 0	300, 0
400	0, 200	0, 300	200, 200	400, 0
500	0, 200	0, 300	0, 400	250, 250



350, 500

Types	$T_1 = \{t_1\},\ T_2 = \{t_2\}$
Beliefs for Barbara	$b_1(t_1, \varnothing) = (200, t_2)$ $b_1(t_1, h_1) = (200, t_2)$
Beliefs for you	$b_2(t_2, h_1) = ((reject, 200), t_1)$

Your type t_2 satisfies conditions, but does not strongly believe in Barbara's rationality.

Problem: Not sufficiently many types in epistemic model M !

- To make the definition of **strong belief in the opponents' rationality** work, we must require that the epistemic model M contains **sufficiently many types.**

- Consider an epistemic model M, and an information set $h \in H_i$:

  If we can find a combination of opponents' types – **possibly outside** M – for which there is a combination of **optimal** strategies leading to h,

  then the epistemic model M must contain at least one such combination of opponents' types.

Definition (Strong belief in the opponents' rationality)

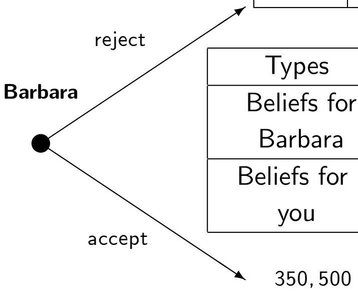
Type t_i **strongly believes in the opponents' rationality** at h if, whenever we can find a combination of opponents' types, possibly outside M, for which there is a combination of optimal strategies leading to h, then

(1) the epistemic model M must contain at least one such combination of opponents' types, and

(2) type t_i must at h only assign positive probability to opponents' strategy-type combinations where the strategy combination leads to h, and the strategies are optimal for the types.

- Definition is based on Battigalli and Siniscalchi (2002). However, they require a **complete** type space. We do not.
- Idea is implicitly present in Pearce's (1984) **extensive form rationalizability** concept.

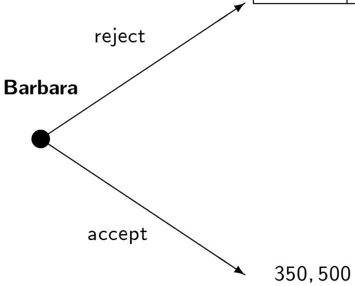
	200	300	400	500
200	100, 100	200, 0	200, 0	200, 0
300	0, 200	150, 150	300, 0	300, 0
400	0, 200	0, 300	200, 200	400, 0
500	0, 200	0, 300	0, 400	250, 250



Types	$T_1 = \{t_1\},\ T_2 = \{t_2\}$
Beliefs for Barbara	$b_1(t_1, \varnothing) = (200, t_2)$ $b_1(t_1, h_1) = (200, t_2)$
Beliefs for you	$b_2(t_2, h_1) = ((\mathit{reject}, 200), t_1)$

Your type t_2 does **not** strongly believe in Barbara's rationality.

200	100, 100	200, 0	200, 0	200, 0
300	0, 200	150, 150	300, 0	300, 0
400	0, 200	0, 300	200, 200	400, 0
500	0, 200	0, 300	0, 400	250, 250



$$T_1 = \{t_1^a, t_1^r\},\ T_2 = \{t_2\}$$

$$\begin{aligned} b_1(t_1^a, \varnothing) &= (300, t_2) \\ b_1(t_1^a, h_1) &= (300, t_2) \\ \\ b_1(t_1^r, \varnothing) &= (500, t_2) \\ b_1(t_1^r, h_1) &= (500, t_2) \end{aligned}$$

$$b_2(t_2, h_1) = ((\textit{reject}, 400), t_1^r)$$

Your type t_2 strongly believes in Barbara's rationality.

Common strong belief in rationality

- **Two-fold strong belief in rationality:**

  Consider an information set h for player i.

  If there is an opponents' strategy-type combination where (a) the opponents' strategy combination leads to h, (b) the strategies are optimal for the types, and (c) the types **strongly believe in the opponents' rationality,**

  then type t_i must at h only assign positive probability to opponents' strategy-type combinations that satisfy (a), (b) and (c).

- To make this definition work, we must require that the epistemic model M contains **sufficiently many types.**

Definition (Common strong belief in rationality)

(Induction start) Type t_i is said to express 1-**fold strong belief in rationality** if t_i strongly believes in the opponents' rationality.

(Inductive step) For $k \geq 2$, say that type t_i expresses k-**fold strong belief in rationality** at h if, whenever we can find a combination of opponents' types, possibly outside M, that express up to $(k-1)$-fold strong belief in rationality, and for which there is a combination of optimal strategies leading to h, then

(1) the epistemic model M must contain at least one such combination of opponents' types, and

(2) type t_i must at h only assign positive probability to opponents' strategy-type combinations where the strategy combination leads to h, the types express up to $(k-1)$-fold strong belief in rationality, and the strategies are optimal for the types.

Type t_i expresses **common strong belief in rationality** if it expresses k-fold strong belief in rationality for every k.

Algorithm

- We wish to find those strategies you can rationally choose under **common strong belief in rationality.**
- Is there an **algorithm** that helps us find these strategies?
- Yes. Algorithm is similar in flavor to the **backward dominance** procedure.

Important ingredients:

- The **full decision problem** for player i at h is $\Gamma^0(h) = (S_i(h), S_{-i}(h))$, where $S_i(h)$ is the set of strategies for player i that lead to h, and $S_{-i}(h)$ is the set of opponents' strategy combinations that lead to h.

- A **reduced decision problem** for player i at h is $\Gamma(h) = (D_i(h), D_{-i}(h))$, where $D_i(h) \subseteq S_i(h)$ and $D_{-i}(h) \subseteq S_{-i}(h)$.

Algorithm (Iterated conditional dominance procedure)

(Induction start) At every information set h, let $\Gamma^0(h)$ be the full decision problem at h.

(Inductive step) Let $k \geq 1$. At every reduced decision problem $\Gamma^{k-1}(h)$, eliminate for every player i those strategies that are strictly dominated at some reduced decision problem $\Gamma^{k-1}(h')$ at which player i is active, **unless** *this would remove* **all** *strategy combinations that lead to h. In the latter case, we remove nothing from $\Gamma^{k-1}(h)$. This leads to new reduced decision problems $\Gamma^k(h)$ at every information set.*

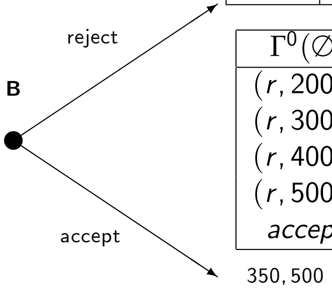
- Algorithm is due to Shimoji and Watson (1998), and is based on earlier procedures by Pearce (1984) and Battigalli (1997).
- The **order of elimination** is **crucial** for the strategies that survive this algorithm!

Theorem (Battigalli and Siniscalchi (2002))

*(1) For every $k \geq 1$, the strategies that can rationally be chosen by a type that expresses up to k-**fold** strong belief in rationality are precisely the strategies in $\Gamma^{k+1}(\emptyset)$.*

*(2) The strategies that can rationally be chosen by a type that expresses **common** strong belief in rationality are exactly the strategies that are in $\Gamma^k(\emptyset)$ for every k.*

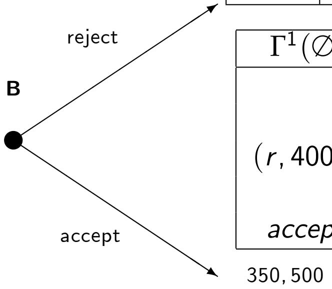
$\Gamma^0(h_1)$	200	300	400	500
$(r, 200)$	$100, 100$	$200, 0$	$200, 0$	$200, 0$
$(r, 300)$	$0, 200$	$150, 150$	$300, 0$	$300, 0$
$(r, 400)$	$0, 200$	$0, 300$	$200, 200$	$400, 0$
$(r, 500)$	$0, 200$	$0, 300$	$0, 400$	$250, 250$



$\Gamma^0(\emptyset)$	200	300	400	500
$(r, 200)$	$100, 100$	$200, 0$	$200, 0$	$200, 0$
$(r, 300)$	$0, 200$	$150, 150$	$300, 0$	$300, 0$
$(r, 400)$	$0, 200$	$0, 300$	$200, 200$	$400, 0$
$(r, 500)$	$0, 200$	$0, 300$	$0, 400$	$250, 250$
accept	$350, 500$	$350, 500$	$350, 500$	$350, 500$

Step 1

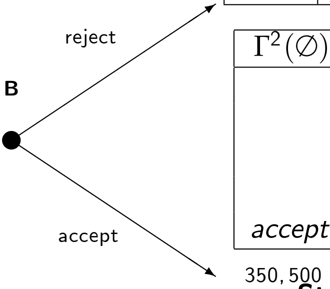
$\Gamma^1(h_1)$	200	300	400	
$(r, 400)$	$0, 200$	$0, 300$	$200, 200$	



$\Gamma^1(\emptyset)$	200	300	400
$(r, 400)$	$0, 200$	$0, 300$	$200, 200$
accept	$350, 500$	$350, 500$	$350, 500$

Step 1

$\Gamma^2(h_1)$		300		
$(r, 400)$		$0, 300$		



$\Gamma^2(\varnothing)$	300
accept	$350, 500$

Step 2: Algorithm stops.

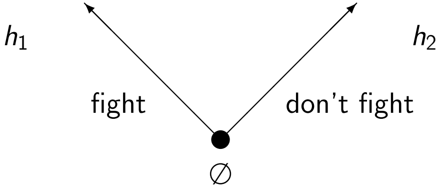
Example: Watching TV with Barbara

Story

- Barbara and you must decide with TV program to watch: *Blackadder* or *Dallas*.
- You prefer *Blackadder* (utility 6) to *Dallas* (utility 3).
- Barbara prefers *Dallas* (utility 6) to *Blackadder* (utility 3).
- You both must write down a program on a piece of paper. If you both write the same program, you will watch it together. Otherwise, you will play a game of cards (utility 2 for both).
- Before writing down a program, you have the option to **start a fight** with Barbara to convince her to watch your favorite program. This would reduce your utility and Barbara's utility by 2.

$\Gamma^0(h_1)$	B	D
B	4, 1	0, 0
D	0, 0	1, 4

$\Gamma^0(h_2)$	B	D
B	6, 3	2, 2
D	2, 2	3, 6

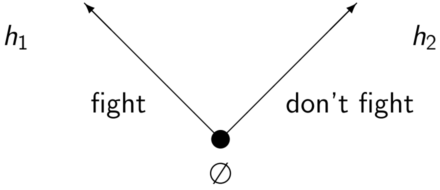


$\Gamma^0(\varnothing)$	(B, B)	(B, D)	(D, B)	(D, D)
$(fight, B)$	4, 1	4, 1	0, 0	0, 0
$(fight, D)$	0, 0	0, 0	1, 4	1, 4
$(don't, B)$	6, 3	2, 2	6, 3	2, 2
$(don't, D)$	2, 2	3, 6	2, 2	3, 6

Step 1

$\Gamma^1(h_1)$	B	D
B	4, 1	0, 0

$\Gamma^1(h_2)$	B	D
B	6, 3	2, 2
D	2, 2	3, 6

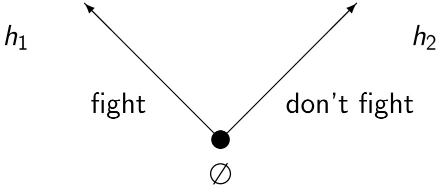


$\Gamma^1(\emptyset)$	(B, B)	(B, D)	(D, B)	(D, D)
$(fight, B)$	4, 1	4, 1	0, 0	0, 0
$(don't, B)$	6, 3	2, 2	6, 3	2, 2
$(don't, D)$	2, 2	3, 6	2, 2	3, 6

Step 1

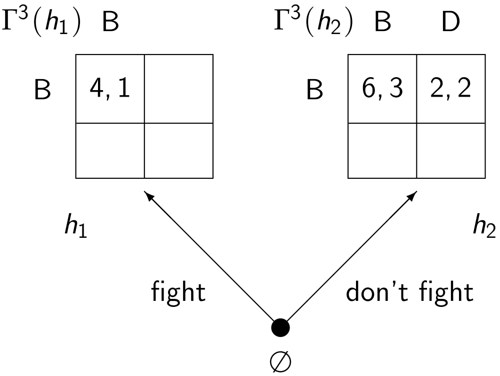
$\Gamma^2(h_1)$	B	
B	4, 1	

$\Gamma^2(h_2)$	B	D
B	6, 3	2, 2
D	2, 2	3, 6



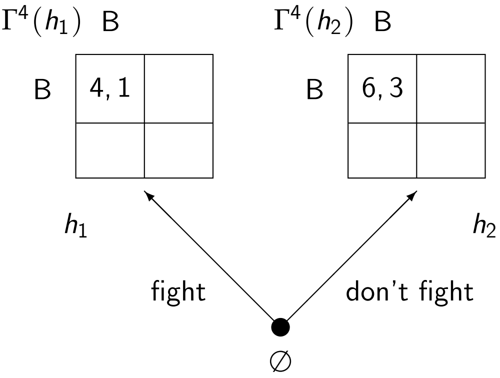
$\Gamma^2(\varnothing)$	(B, B)	(B, D)
$(fight, B)$	4, 1	4, 1
$(don't, B)$	6, 3	2, 2
$(don't, D)$	2, 2	3, 6

Step 2



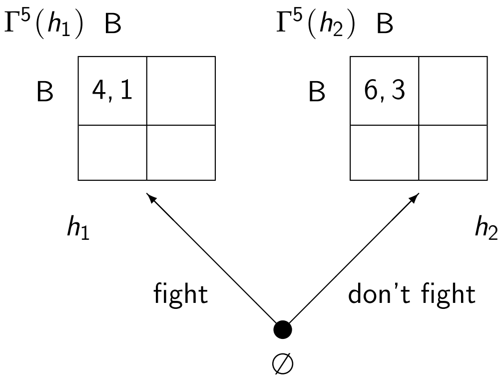
$\Gamma^3(\varnothing)$	(B, B)	(B, D)
$(fight, B)$	4, 1	4, 1
$(don't, B)$	6, 3	2, 2

Step 3



$\Gamma^4(\emptyset)$	(B, B)
$(fight, B)$	$4, 1$
$(don't, B)$	$6, 3$

Step 4



$\Gamma^5(\varnothing)$	(B, B)
$(don't, B)$	6, 3

Step 5: Algorithm stops.

Relation with backward induction reasoning

- You **initially deem possible an outcome** z under common strong belief in rationality, if there is a strategy combination leading to z, where every strategy can rationally be chosen under common strong belief in rationality.

Theorem (Outcomes under common strong belief in rationality and common belief in future rationality)

Every outcome you initially deem possible under common strong belief in rationality, is also initially deemed possible under common belief in future rationality.

- Proof can be found in my book Perea (2012) and in Chen and Micali (2011, 2012).

- The opposite direction is not true! See the example "Watching TV with Barbara".

- Remember that in dynamic games with **perfect information**, **common belief in future rationality** selects exactly the **backward induction strategies.**

Corollary (Common strong belief in rationality leads to backward induction outcomes)

In a game with perfect information, every outcome you initially deem possible under common strong belief in rationality must be a backward induction outcome.

- The first proof for this result is in Battigalli (1997).
- A more direct proof can be found in Heifetz and Perea (2015).
- However, common strong belief in rationality may **not** lead to the **backward induction strategy** for every player!

The End

Thank you for your attention!

- P. Battigalli, 'On rationalizability in extensive games', *Journal of Economic Theory*, 74 (1997), 40–61
- P. Battigalli and M. Siniscalchi, 'Strong belief and forward induction reasoning', *Journal of Economic Theory*, 106 (2002), 356–391
- J. Chen and S. Micali, 'The robustness of extensive-form rationalizability', Working paper (2011)
- J. Chen and S. Micali, 'The order independence of iterated dominance in extensive games', *Theoretical Economics*, 8 (2012), 125–163

- A. Heifetz and A. Perea, 'On the outcome equivalence of backward induction and extensive-form rationalizability', *International Journal of Game Theory*, 44 (2015), 37–59
- D. Pearce, 'Rationalizable strategic behavior and the problem of perfection', *Econometrica*, 52 (1984), 1029–1050
- A. Perea, '*Epistemic Game Theory: Reasoning and Choice*', *Cambridge University Press* (2012)
- M. Shimoji and J. Watson, 'Conditional dominance, rationalizability, and game forms', *Journal of Economic Theory*, 83 (1998), 161–195