Signaling Games and Equilibrium

Pierpaolo Battigalli
Bocconi University
Game Theory: Analysis of Strategic Thinking

December 5, 2023

Abstract
These slides focus on **signaling games**, that is, leader-follower games with incomplete information. To simplify the analysis, it is assumed that the first mover knows the state of nature, that is, only the second mover is uninformed. The perfect Bayesian equilibrium (PBE) concept is analyzed. A kind of "case-by-case" backward-induction algorithm allows to compute all (pure) PBEs. On top of standard (PBE) equilibrium analysis, we give hints on so called "forward-induction reasoning."
[These slides summarize and in part complement Section 15.8 of Chapter 15 of GT-AST.]

- Like Nash equilibrium, Bayesian equilibrium allows for non-maximizing choices at histories that are not supposed to occur in equilibrium.

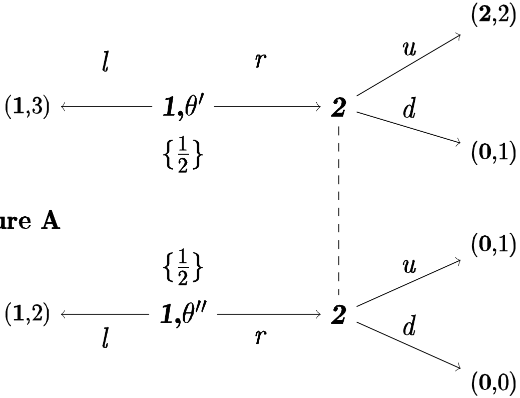
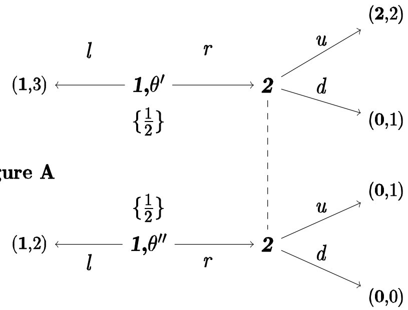


Figure A

- $(\ell'.\ell'', d)$ is an "imperfect" Bayesian equilibrium: d is a "non-credible" threat (u dominates d *given* r).

(Im-)perfect Bayesian equilibria: an example, 2



- Only $(r'.\ell'', u)$ is a "perfect" Bayesian equilibrium (PBE) strategy profile.
- Figure A shows an example of "**signaling game**": player 1's action may "signal" 1's private information.
- In particular, it does in equilibrium $(r'.\ell'', u)$: pl. 1 anticipates that 2 would reply to r with u; thus he chooses r if $\theta = \theta'$ and ℓ if $\theta = \theta''$. If player 2 understands this, r "signals" to 2 that $\theta = \theta'$.

Signaling games

- We consider a simple class of games, *for which everybody agrees on the meaning of PBE*:

Definition

A **leader-follower game** is a two-person, two-stage game with observable actions where only one player is active at each stage and the second mover is different from the first mover. A (Bayesian) **signaling game** (or **sender-receiver game**) is a leader-follower Bayesian game.

- Simplifying *assumption*: $\Theta \cong \Theta_1$. With this:
  - First mover: $i = 1$, called **Sender,** *knows* θ and chooses actions, called **messages** or **signals**, $a_1 \in A_1$.
  - Second mover $i = 2$, called **Receiver**: given $a_1 \in A_1$,chooses $a_2 \in \mathcal{A}_2(a_1) \subseteq A_2$ (if a_1 ends the game, let $\mathcal{A}_2(a_1) = \{\text{wait}\}$ by convention).
  - Thus,

$$Z = \{(a_1, a_2) \in A_1 \times A_2 : a_2 \in \mathcal{A}_2(a_1)\}, \; u_i : \Theta \times Z \to \mathbb{R} \; (i = 1, 2).$$

Signaling and Perfect Bayesian Equilibrium, 1

- **Prior**=*initial* exogenous (strictly positive) *belief* of receiver $\mu(\cdot|\varnothing) = p \in \Delta^\circ(\Theta)$. **(Extended) Behavior strategies**: $\beta_1 \in \Delta(A_1)^\Theta$, $\beta_2 \in \times_{a_1 \in A_1} \Delta\left(\mathcal{A}_2(a_1)\right)$.
- **Expected payoffs of actions** given (β, μ), $\forall(\theta, a_1, a_2) \in \Theta \times Z$

$$\begin{aligned}\mathbb{E}_{\beta_2}\left(u_1|\theta, a_1\right) &= \sum_{a_2 \in \mathcal{A}_2(a_1)} u_1(\theta, a_1, a_2)\beta_2(a_2|a_1), \\ \mathbb{E}_{\mu}\left(u_2|a_1, a_2\right) &= \sum_{\theta' \in \Theta} u_2(\theta', a_1, a_2)\mu(\theta'|a_1).\end{aligned}$$

- **Bayes consistency:** $\mathbb{P}_{\beta_1}(\theta, a_1) = \beta_1(a_1|\theta)p(\theta)$, $\mathbb{P}_{\beta_1}(a_1) = \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')$ (predictive probability of a_1). Thus,

$$\mathbb{P}_{\beta_1}(a_1) > 0 \Rightarrow \mu(\theta|a_1) = \frac{\beta_1(a_1|\theta)p(\theta)}{\sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')}. \qquad \text{(BR-SG)}$$

Signaling and Perfect Bayesian Equilibrium, 2

Definition

A PBE (or *sequential equil.*) of (Bayesian) signaling game $\langle A_1, A_2, \Theta, p, \mathcal{A}_2(\cdot), u_1, u_2 \rangle$ is an assessment (β_1, β_2, μ) such that
(1) (rat. of 1) $\forall \theta \in \Theta$, $\mathrm{supp}\beta_1(\cdot|\theta) \subseteq \arg\max_{a_1 \in A_1} \mathbb{E}_{\beta_2}(u_1|\theta, a_1)$,
(2) (rat. of 2) $\forall a_1 \in A_1$, $\mathrm{supp}\beta_2(\cdot|a_1) \subseteq \arg\max_{a_2 \in \mathcal{A}_2(a_1)} \mathbb{E}_{\mu}(u_2|a_1, a_2)$,
(3) (Bayes consistency) $\forall (\theta, a_1) \in \Theta \times A_1$, (BR-SG) holds.

- **Comment:** 3 (vectors of) unknowns, β_1, β_2, μ. Each equilibrium condition involves 2 out of 3 unknowns: (1) β_1, β_2, (2) β_2, μ (3) β_1, μ.
- **Example:** The unique PBE of the game in Fig. A is $(r'.\ell'', u, \mu(\theta'|r) = 1)$.

Comment on PBE and Bayes Consistency

- (BR-SG) is the version of *Bayes rule* for *signaling games*.
- *Analogy*:
  - Given β_1, each θ is a kind of statistical model (like the composition of an urn), it determines the probabilities of the pieces of "evidence" observable by pl. 2, that is, the probabilities $(\beta_1(a_1|\theta))_{a_1 \in A_1}$. The likelihood function is $(\theta, a_1) \mapsto \beta_1(a_1|\theta)$.
  - The **prior** belief of 2 about statistical models is $\mu(\cdot|\varnothing) = p \in \Delta^\circ(\Theta)$. The **posterior** belief conditional on "evidence" a_1 is $\mu(\cdot|a_1)$.
- **Bayes rule:**
  - It requires to *use* **Bayes formula**

$$\mu(\theta|a_1) = \frac{\beta_1(a_1|\theta)p(\theta)}{\sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')} \qquad \text{(BF)}$$

    whenever possible, i.e., *whenever the denominator is positive.*
  - The possibility of denominator=0 (in discrete models) is typically ruled out in statistics, but we must allow for it because β_1 is *endogenous* and may assign probability 0 to some actions/messages.

Pooling and Separating Equilibria

- A PBE is **separating** if β_1 is pure and *different types choose different actions.*
- A PBE is **pooling** if β_1 is pure and *all types choose the same action.*
- In a separating equilibrium, $\mu(\cdot|a_1)$ is degenerate for each action a_1 "on the equilibrium path," i.e., chosen by some type $\theta = \vartheta(a_1)$: $\mu(\vartheta(a_1)|a_1) = 1$.
- In a pooling equilibrium where each type θ chooses a_1^*, $\mu(\cdot|a_1^*) = p(\cdot)$, whereas $\mu(\cdot|a_1)$ is not pinned down by Bayes rule if $a_1 \neq a_1^*$ (if a_1 is "off the equilibrium path").

Example with Multiple PBEs

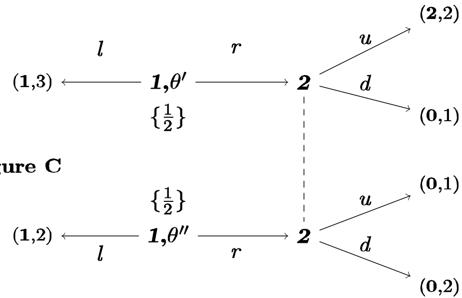


Figure C

- Here, each action of the receiver is justifiable given r. Given ℓ, the receiver can only "wait". Separating and pooling PBEs:
  - separating: $\left\{(\beta,\mu): \beta_1\left(r|\theta'\right)=\beta_1\left(\ell|\theta''\right)=1=\beta_2\left(u|r\right),\mu\left(\theta'|r\right)=1\right\}$;
  - pooling-1: $\left\{(\beta,\mu): \beta_1\left(\ell|\theta'\right)=\beta_1\left(\ell|\theta''\right)=1,\beta_2\left(d|r\right)\geq 0.5,\mu\left(\theta'|r\right)=0.5\right\}$,
  - pooling-2: $\left\{(\beta,\mu): \beta_1\left(\ell|\theta'\right)=\beta_1\left(\ell|\theta''\right)=1,\beta_2\left(d|r\right)=1,\mu\left(\theta'|r\right)\leq 0.5\right\}$.

Backward Computation of PBEs: Dominance

- One can use a kind of *"case-by-case backward induction"* algorithm to compute PBEs, starting from Stage 2 and then going back to Stage 1 (focus on *pure* PBEs for simplicity). Preliminary:

Definitions

Action $a_2 \in \mathcal{A}_2(a_1)$ is **conditionally dominated** given a_1 if

$$\exists \alpha_2 \in \Delta(\mathcal{A}_2(a_1)), \forall \theta, \sum_{a_2' \in \mathcal{A}_2(a_1)} u_2(\theta, a_1, a_2') \alpha_2(a_2') > u_2(\theta, a_1, a_2);$$

Denote by $\mathcal{ND}_2(a_1) \subseteq \mathcal{A}_2(a_1)$ the set of actions that are *not* conditionally dominated given a_1. For any $a_1 \in A_1$, $a_2 \in \mathcal{A}_2(a_1)$,

$$J(a_2|a_1) := \left\{ \nu \in \Delta(\Theta) : a_2 \in \arg\max_{a_2' \in \mathcal{A}_2(a_1)} \mathbb{E}_\nu(u_2(\cdot, a_1, a_2')) \right\}$$

denotes the set of **beliefs justifying** a_2 **given** a_1.

Backward computation of PBEs: Stage 2 of the game

Lemma

(Cf. Wald-Pearce Lemma) An action of the receiver is justifiable IFF it is undominated:

$$\forall a_1 \in A_1, \forall a_2 \in \mathcal{A}_2(a_1),\ J(a_2|a_1) \neq \emptyset \iff a_2 \in \mathcal{ND}_2(a_1).$$

Analysis of stage 2 of the game:

- *Each conditionally undominated* (deterministic behavior) *strategy*

$$s_2 \in \mathbf{ND}_2 := \times_{a_1 \in A_1} \mathcal{ND}_2(a_1)$$

  is a "case" to start from in the BI computation.

- For every case s_2, the set of **systems of beliefs justifying** s_2 is

$$\mathbf{J}(s_2) := \times_{a_1 \in A_1} J(s_2(a_1)|a_1) \subseteq \Delta(\Theta)^{A_1}.$$

Analysis of stage 1 of the game:

- **Sender's best reply.** Given "case" $s_2 \in \mathbf{ND}_2$, *consider the set of* **decision functions** *of the Sender* $\mathbf{a}_1 \in A_1^\Theta$ (deterministic extended behavior strategies) *that maximize* EU_1 *given* s_2:

$$\mathbf{BR}_1(s_2) := \times_{\theta\in\Theta} \arg\max_{a_1\in A_1} u_1(\theta, a_1, s_2(a_1))$$

(typically, there is only one $\mathbf{a}_1 \in \mathbf{BR}_1(s_2)$).

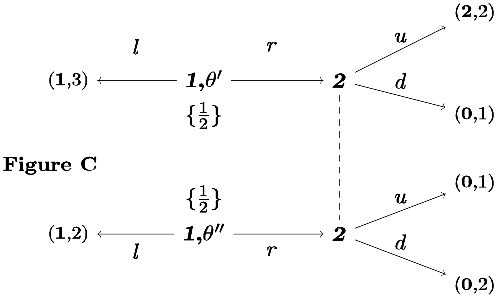
- **Bayes consistency.** For every $\mathbf{a}_1 \in A_1^\Theta$, let $\mathbf{BC}(\mathbf{a}_1)$ denote the set of $\mu \in \Delta(\Theta)^{A_1}$ *consistent with Bayes rule given* $\mathbf{a}_1$: $\mathbf{BC}(\mathbf{a}_1) :=$

$$\left\{\mu : \forall (a_1, \theta), \mu(\theta|a_1)\left(\sum_{\theta'} \beta_1^{\mathbf{a}_1}(a_1|\theta') p(\theta')\right) = \beta_1^{\mathbf{a}_1}(a_1|\theta) p(\theta)\right\}$$

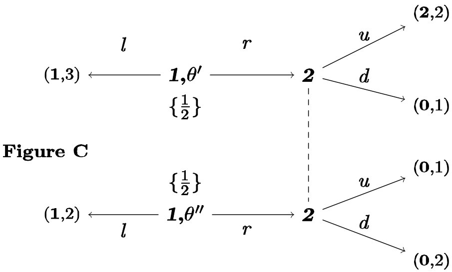
where $\beta_1^{\mathbf{a}_1}(a_1|\theta) := \mathbf{1}_{[a_1=\mathbf{a}_1(\theta)]} = 1$ if $\mathbf{a}_1(\theta) = a_1$ and $= 0$ otherwise.

- $(\mathbf{a}_1, s_2)$ *is part of a* (pure) *PBE assessment* if and only if $s_2 \in \mathbf{ND}_2$, $\mathbf{a}_1 \in \mathbf{BR}_1(s_2)$ *and* $\mathbf{BC}(\mathbf{a}_1) \cap \mathbf{J}(s_2) \neq \emptyset$.

Backward-Computation Example: Receiver



- The receiver has no dominated action. Note, "wait" after ℓ: $J\left(\text{wait}|\ell\right) = \Delta\left(\Theta\right)$ (pedantic detail added to better connect with the general algorithm). Two cases:
  1. $s_2 = u$ (u if r, wait if ℓ), $J\left(u|r\right) = \left\{\nu \in \Delta\left(\Theta\right) : \nu\left(\theta'\right) \geq 0.5\right\}$, $\mathbf{J}\left(s_2\right) = \Delta\left(\Theta\right) \times J\left(u|r\right) \subseteq \Delta\left(\Theta\right)^{\{\ell,r\}}$.
  2. $s_2 = d$ (d if r, wait if ℓ), $J\left(d|r\right) = \left\{\nu \in \Delta\left(\Theta\right) : \nu\left(\theta'\right) \leq 0.5\right\}$, $\mathbf{J}\left(s_2\right) = \Delta\left(\Theta\right) \times J\left(d|r\right) \subseteq \Delta\left(\Theta\right)^{\{\ell,r\}}$.



- $\mathbf{a}_1(\theta'') = \ell$ (dominant for θ''), $\mathbf{a}_1(\theta')$, is case-dependent.
  1. $s_2 = u$, $\mathbf{a}_1(\theta') = r$, $\mathbf{BC}(r'.\ell'') = \{\mu : \mu(\theta''|\ell) = \mu(\theta'|r) = 1\}$, $\mathbf{BC}(r'.\ell'') \cap \mathbf{J}(u) = \mathbf{BC}(r'.\ell'')$, $(r'.\ell'', u, \mu(\theta'|r) = 1)$ is a PBE.
  2. $s_2 = d$, $\mathbf{a}_1(\theta') = \ell$, $\mathbf{BC}(\ell'.\ell'') = \{\mu : \mu(\theta'|\ell) = 0.5\}$, $\mathbf{BC}(\ell'.\ell'') \cap \mathbf{J}(d) = \{\mu : \mu(\theta'|\ell) = 0.5, \mu(\theta'|r) \leq 0.5\}$, each $(\ell'.\ell'', d, \mu)$ with $\mu(\theta'|\ell) = 0.5$ and $\mu(\theta'|r) \leq 0.5$ is a PBE.

Forward Induction

- Intuitively, *forward-induction reasoning* means that—whenever possible—players interpret the observed actions of their opponents as the result of *intentional* and "*strategically sophisticated* choices" (cf., best rationalization principle).
- The cleanest and most rigorous way to capture forward-induction reasoning is to use the epistemic assumption of *common strong belief in rationality*, possibly with some transparent belief restrictions [cf. strong (directed) rationalizability].
- Such ideas can be used to characterize "forward-induction refinements" of the PBE concept (e.g., the "(Iterated) Intuitive Criterion" of Cho & Kreps 1987, cf Battigalli & Siniscalchi 2002/3).
- In games with complete information, there is a tension between forward and backward-induction reasoning; but the *predictions (possible paths) of FI* (weakly) *refine the predictions of BI* (see Ch. 12). Such considerations and results extend to games with incomplete information.

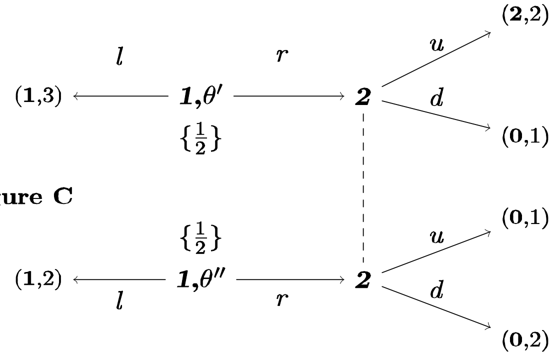


Figure C

- According to the PBE concept, d is a "credible threat" of *2*: $((\ell'.\ell'', d), \mu(\theta'|r) < \frac{1}{2})$ is a PBE. Yet:
- $\mathrm{SB}_2(R_1) \Rightarrow \mu(\theta'|r) = 1$ (r is dominated for θ'', not for θ');
- $R_2 \cap \mathrm{SB}_2(R_1) \Rightarrow u$, $R_1 \cap \mathrm{B}_1(R_2 \cap \mathrm{SB}_2(R_1)) \Rightarrow r'.\ell''$.

Given the *ex ante strategic form* (assuming, w.l.o.g. in this case, a common prior) of the signaling game depicted in Fig. C, (maximal) *iterated weak dominance* yields:

$\mathbf{1}\backslash 2$	u	d
$\ell'.\ell''$	**1**, *2.5*	**1**, *2.5*
$\ell'.r''$	**0.5**, *2*	**0.5**, *2.5*
$r'.\ell''$	**1.5**, *2*	**0.5**, *1.5*
$r'.r''$	**1**, *1.5*	**0**, *1.5*

$\Rightarrow$

$\mathbf{1}\backslash 2$	u	d
$\ell'.\ell''$	**1**, *2.5*	**1**, *2.5*
$r'.\ell''$	**1.5**, *2*	**0.5**, *1.5*

$\Rightarrow$

$\mathbf{1}\backslash 2$	u
$\ell'.\ell''$	**1**, *2.5*
$r'.\ell''$	**1.5**, *2*

$\Rightarrow$

$\mathbf{1}\backslash 2$	u
$r'.\ell''$	**1.5**, *2*

Note: Same solution $(r'.\ell'', \mathrm{u})$ as with FI reasoning!

THAT'S ALL FOLKS!

Battigalli, P., E. Catonini, and N. De Vito (2023): *Game Theory: Analysis of Strategic Thinking*. Typescript, Bocconi University.

Battigalli, P. (2023): *Mathematical Language and Game Theory*. Typescript, Bocconi University.

Battigalli, P. and M. Siniscalchi (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, **106**, 356-391.

Battigalli, P. and M. Siniscalchi (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, **3** (1), Article 3.

- Cho, I.K. and D. Kreps (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, **102**, 179-222.
- Dekel, E., and M. Siniscalchi (2015): "Epistemic Game Theory," in *Handbook of Game Theory with Economic Applications, Volume 4*, ed. by P. Young and S. Zamir. Amsterdam: North-Holland, 619-702.