Game Theory

Bayesian Games Revisited

Tetsuya Hoshino

April 2, 2022

In this note, we discuss the conceptual differences between incomplete-information games and Bayesian games.[1]

1 Incomplete-Information Games*

Roughly speaking, an incomplete-information game is a game whose game structure is not common knowledge among all players. A straightforward approach to represent such a situation—if we do not know Bayesian games—is to explicitly specify all relevant uncertainty that matters for player i's decision-making.

Example 1. There are two players, denoted $i = 1, 2$. Player 1 (she) chooses either action T or B, and player 2 (he) chooses either action L or R. Their payoff structure depends on a state $\theta \in \Theta = \{1, 2\}$. Assume that player 1 has her belief $\mu_1^1 \in \Delta(\Theta)$ and that player 2 has his belief $\mu_2^1 \in \Delta(\Theta)$. Their belief may not coincide.[2]

	L	R
T	$1, 0$	$0, 1$
B	$0, 1$	$1, 0$

(a) state $\theta = 1$

	L	R
T	$0, 1$	$1, 0$
B	$1, 0$	$0, 1$

(b) state $\theta = 2$

Player 1's choice of action depends not only on her belief μ_1^1 but also on player 2's choice of action, which depends on his belief μ_2^1. Then, she must have a belief about his belief μ_2^1. This belief—player 1's belief about player 2's first-order belief—is called her **second-order belief** μ_1^2. Similarly, player 2 must have his second-order belief μ_2^2. This urges player 1 to form her belief about player 2's second-order belief, which is called her **third-order belief** μ_1^3. Similarly, player 2 also forms his third-order belief μ_2^3. This argument continues ad infinitum. Each player i has her k**th-order belief** for every $k \in \mathbb{N}$. Then, player i's belief of every order is specified by her **belief hierarchy** $\mu_i = (\mu_i^1, \mu_i^2, \ldots)$. □

[1]In some references, these two games are treated as if they were synonymous, but actually not.

[2]Here, we do not ask how they reach those beliefs. For example, they might have simply started from different priors, or they might have started from a common prior and then received different information.

2 Bayesian Games Revisited*

As shown by Example 1, the approach that describes all relevant uncertainty requires spelling out the infinite hierarchies of beliefs for all players. However, the approach is intractable. The approach of Bayesian games, however, avoids this intractability by encapsulating the uncertainty, or players' belief hierarchies, in their types.

Next, we see how the Bayesian-game approach deals with belief hierarchies.

Example 2. There are two players, denoted $i \in I = \{1, 2\}$. Here is the specification of states of nature and types:

- $\Omega = \Theta \times T$ is the set of states of the world.
- $\Theta = \{(1,1), (1,2), (2,1), (2,2)\}$ is the set of states of nature.
- $T = \{(1,1), (1,2), (2,1), (2,2)\}$ is the set of type profiles.
  - Player i learns her type $t_i = \theta_i$ at state $\theta = (\theta_1, \theta_2)$.
- $\mathbb{P} \in \Delta(\Omega)$ is the (common) prior defined by the table below:

$\theta_1 \backslash \theta_2$	1	2
1	$\frac{1}{4}$	$\frac{1}{4}$
2	$\frac{1}{3}$	$\frac{1}{6}$

Figure 2: the state distribution

Since a state θ is bijective to a type profile t, this table pins down the type distribution.[3]

In this example, we write player i's belief $\mu_i^1 = (\mu_i^1(1,1), \mu_i^1(1,2), \mu_i^1(2,1), \mu_i^1(2,2))$, with her belief $\mu_i^1(\theta)$ of state θ.

First-order beliefs

- Player 1 of type $t_1 = 1$ has $\mu_1^1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$.
- Player 1 of type $t_1 = 2$ has $\mu_1^1 = (0, 0, \frac{2}{3}, \frac{1}{3})$.
- Player 2 of type $t_2 = 1$ has $\mu_2^1 = (\frac{3}{7}, 0, \frac{4}{7}, 0)$.
- Player 2 of type $t_2 = 2$ has $\mu_2^1 = (0, \frac{3}{5}, 0, \frac{2}{5})$.

Second-order beliefs

- Player 1 of type $t_1 = 1$ has μ_1^2 that assigns:
  - probability $\frac{1}{2}$ to $\mu_2^1 = (\frac{3}{7}, 0, \frac{4}{7}, 0)$.
  - probability $\frac{1}{2}$ to $\mu_2^1 = (0, \frac{3}{5}, 0, \frac{2}{5})$.

[3]Because of this reason, this formulation is redundant, but we distinguish them for pedagogical reasons.

- Player 1 of type $t_1 = 2$ has μ_1^2 that assigns:
  - probability $\frac{2}{3}$ to $\mu_2^1 = (\frac{3}{7}, 0, \frac{4}{7}, 0)$.
  - probability $\frac{1}{3}$ to $\mu_2^1 = (0, \frac{3}{5}, 0, \frac{2}{5})$.
- Player 2 of type $t_2 = 1$ has μ_2^2 that assigns:
  - probability $\frac{3}{7}$ to $\mu_1^1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$.
  - probability $\frac{4}{7}$ to $\mu_1^1 = (0, 0, \frac{2}{3}, \frac{1}{3})$.
- Player 2 of type $t_2 = 2$ has μ_2^2 that assigns:
  - probability $\frac{3}{5}$ to $\mu_1^1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$.
  - probability $\frac{2}{5}$ to $\mu_1^1 = (0, 0, \frac{2}{3}, \frac{1}{3})$.

Analogously, we have player i's kth-order beliefs for each $k \geq 1$.[4] $\square$

3 Bayesian Games versus Incomplete-Information Games*

As shown by Example 2, Bayesian games induce the (entire) hierarchies of beliefs, but do Bayesian games induce *all* possible belief hierarchies? That is, given any belief hierarchies about a state θ of nature, does there exist a Bayesian game that induces these belief hierarchies?

3.1 Belief Hierarchies

There are two players, denoted $i = 1, 2$. There is an unknown parameter $\theta \in \Theta$, which determines a game structure. Player 1 forms her belief about θ. This belief, denoted $\mu_1^1 \in \Delta(\Theta)$, is called her **first-order belief**. Similarly, player 2 forms his first-order belief μ_2^1.

Player 1 notices that since player 2's choice of action depends on his first-order belief, player 1 has to form her belief about player 2's first-order belief μ_2^1. Player 1 then forms her **second-order belief** $\mu_1^2 \in \Delta(\Theta \times \Delta(\Theta))$, which assigns probability $\mu_1^2(\theta, \mu_2^1)$ to a pair (θ, μ_2^1). Similarly, player 2 forms his second-order belief μ_2^2.

Player 1 notices that since player 2's choice of action depends on his second-order beliefs, player 1 has to form her belief about player 2's second-order beliefs μ_2^2. Player 1 then forms her **third-order belief** $\mu_1^3 \in \Delta(\Theta \times \Delta(\Theta) \times \Delta(\Theta \times \Delta(\Theta)))$, which assigns probability $\mu_1^3(\theta, \mu_2^1, \mu_2^2)$ to a pair $(\theta, \mu_2^1, \mu_2^2)$. Similarly, player 2 forms his third-order belief μ_2^3. This argument continues ad infinitum. Each player i has to form her k**th-order belief** μ_i^k for every $k \in \mathbb{N}$.

Player i's **belief hierarchy** $\mu_i = (\mu_i^1, \mu_i^2, \ldots)$ specifies her belief of every order.

[4] For example, player 1 of type $t_1 = 1$ has μ_1^3 that assigns:
- probability $\frac{1}{2}$ to player 2 having a second-order belief that assigns:
  - probability $\frac{3}{7}$ to $\mu_1^1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$.
  - probability $\frac{4}{7}$ to $\mu_1^1 = (0, 0, \frac{2}{3}, \frac{1}{3})$.
- probability $\frac{1}{2}$ to player 2 having a second-order belief that assigns:
  - probability $\frac{3}{5}$ to $\mu_1^1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$.
  - probability $\frac{2}{5}$ to $\mu_1^1 = (0, 0, \frac{2}{3}, \frac{1}{3})$.

Remark 1. The above definitions of player i's kth-order belief for each $k \geq 2$ is different from those in Example 1. For instance, the above definition defines player 1's second-order belief $\mu_1^2 \in \Delta(\Theta \times \Delta(\Theta))$, while Example 1 defines her second-order belief $\mu_1^2 \in \Delta(\Delta(\Theta))$. □

3.2 Bayesian Games versus Incomplete-Information Games

We may choose the largest possible type space for each player—that is, we may take the set of all her hierarchies μ_i as her type space T_i^0. Then, we can represent any belief by definition. That is, if we have player i's belief μ_i^k for all $k \in \mathbb{N}$ then we can determine her "type" μ_i. However, the formulation of a Bayesian game requires more. Player i's hierarchy μ_i must pin down her belief about a state θ and player $-i$'s hierarchies μ_{-i}. That is, we must find a function $f^0 : T_i^0 \to \Delta(\Theta \times T_{-i}^0)$ that maps each $\mu_i \in T_i^0$ into a unique belief $f^0(\mu_i) \in \Delta(\Theta \times T_{-i}^0)$, but, as we shall see, this approach does not work.

Universal Type Spaces It is instructive to summarize our question.

> Does there exist a collection $\{T_i^*\}_i$ such that for each $i \in I$, T_i^* is the set of all possible beliefs and is isomorphic to the set $\Delta(\Theta \times \prod_{j \neq i} T_j^*)$?

If we have such a collection $\{T_i^*\}_i$, then with the state space $\Omega = \Theta \times \prod_i T_i^*$, we can use a Bayesian game that captures any situation in which players are uncertain not only about each other's payoffs but also about each other's beliefs.

3.3 Existence of Universal Type Spaces

Following Brandenburger & Dekel (1993), we construct a "universal" type space T_i^* and an isomorphism $f^* : T_i^* \to \Delta(\Theta \times T_{-i}^*)$.[5]

Assume that there are two players, denoted $i = 1, 2$.[6] Assume that a state space Θ is Polish. Then, we recursively define the following sets:

$$\begin{aligned} X^0 &= \Theta, \\ X^k &= X^{k-1} \times \Delta(X^{k-1}) \qquad \forall\, k \geq 1. \end{aligned}$$

Then, $\mu_i^k \in \Delta(X^{k-1})$ is player i's kth-order belief. For example, $\mu_i^1 \in \Delta(\Theta)$ is player i's first-order belief, and $\mu_i^2 \in \Delta(\Theta \times \Delta(\Theta))$ is player i's second-order belief.

We define the set of all player i's belief hierarchies $\mu_i = (\mu_i^1, \mu_i^2, \ldots)$ as follows:[7]

$$T^0 = \prod_{k=0}^{\infty} \Delta(X^k).$$

[5] Mertens & Zamir (1985) provide a different construction. They assume that a state space is compact, while Brandenburger & Dekel (1993) assume that a state space is Polish, as detailed below. Neither is stronger.

[6] It is possible to extend the result to any finite number of players.

[7] Since each player has the same set T^0, we do not add a player subscript.

Note (Polish Space):

Definition 1. A topological space Z is **separable** if it contains a countable dense subset.

Definition 2. A topological space Z is **completely metrizable** if it admits a metric such that the induced topology coincides with the topology on Z and that every Cauchy sequence has a limit in Z.

Definition 3. A topological space Z is **Polish** if it is separable and completely metrizable.

Example 3. The Euclidean space $\mathbb{R}$ (with the usual topology) is Polish. First, it is separable because the set $\mathbb{Q}$ of all rational numbers is countable and dense. Second, it is completely metrizable because the usual topology is induced by the usual metric and every Cauchy sequence under the usual metric has a limit in $\mathbb{R}$. $\square$

Remark 2. For a Polish space Z, let $\Delta(Z)$ denote the set of probability measures on the Borel sigma-algebra, endowed with the weak* topology (the topology of weak convergence). Then, $\Delta(Z)$ is also completely metrizable (by the Prokhorov metric) and thus is Polish. $\square$

Remark 3. The product of countable Polish spaces is Polish (in the product topology). $\square$

Then, T^0 is Polish.

Coherency Consider any hierarchy $\mu_i = (\mu_i^1, \mu_i^2, \dots) \in T^0$. Take, for example, the first-order belief $\mu_i^1 \in \Delta(\Theta)$ and the second-order belief $\mu_i^2 \in \Delta(\Theta \times \Delta(\Theta))$. These are redundant in the sense that both must specify the same marginal beliefs on $X_0 = \Theta$. If, on the contrary, μ_i^1 and μ_i^2 assigned different probabilities to the same state, then something would be wrong. Since the set T^0 constructed above include such "irrational" belief hierarchies, we exclude them by requiring that marginal beliefs must coincide.

Definition 4. A belief hierarchy $\mu_i = (\mu_i^1, \mu_i^2, \dots) \in T^0$ is **coherent** if for each $k \geq 2$,

$$\operatorname*{marg}_{X^{k-2}} \mu_i^k = \mu_i^{k-1}.$$

Let T^1 be the set of all **coherent** belief hierarchies:

$$T^1 = \Big\{ \mu \in T^0 : \mu \text{ is coherent} \Big\}.$$

Belief-Preservation Suppose that player i has a coherent hierarchy $\mu_i = (\mu_i^1, \mu_i^2, \dots) \in T^1$. Then, she must have a belief, denoted $f(\mu_i) \in \Delta(\Theta \times T^0)$, about a state θ and player $-i$'s (possibly incoherent) belief hierarchy $\mu_{-i} \in T^0$. How does this belief look like?

We want to have a mapping $f : T^1 \to \Delta(\Theta \times T^0)$ that assigns to each $\mu_i \in T^1$ a belief $f(\mu_i) \in \Delta(\Theta \times T^0)$. Under this belief $f(\mu_i)$, what probability does player i assign to "each $\theta \in \Theta$"?[8] By taking the marginal, player i (with hierarchy μ_i) has the first-order belief,

[8] This "each $\theta \in \Theta$" may not make rigorous sense, because if this measure $\operatorname{marg}_\Theta f(\mu_i) \in \Delta(\Theta)$ is atomless, it assigns probability zero to each $\theta \in \Theta$. We should define the probability assigned to any Borel set.

$\text{marg}_\Theta f(\mu_i) \in \Delta(\Theta)$. Recall that player i with hierarchy $\mu_i = (\mu_i^1, \mu_i^2, \ldots)$, by definition, has the first-order belief, $\mu_i^1 \in \Delta(\Theta)$. These beliefs must agree: For $X^0 = \Theta$,

$$\operatorname*{marg}_{X^0} f(\mu_i) = \mu_i^1.$$

Under the belief $f(\mu_i)$, what probability does player i assign to "each $(\theta, \mu_{-i}^1) \in X^1$"?[9] By taking the marginal, player i (with hierarchy μ_i) has the second-order belief, $\text{marg}_{X^1} f(\mu_i) \in \Delta(X^1)$. Recall that player i with hierarchy $\mu_i = (\mu_i^1, \mu_i^2, \ldots)$, by definition, has the second-order belief, $\mu_i^2 \in \Delta(X^1)$. These beliefs must agree:

$$\operatorname*{marg}_{X^1} f(\mu_i) = \mu_i^2.$$

We must have analogous agreement for each order, reaching the following definition:

Definition 5. A mapping $f : T^1 \to \Delta(\Theta \times T^0)$ is **belief-preserving** if for each $k \geq 1$,

$$\operatorname*{marg}_{X^{k-1}} f(\mu_i^1, \mu_i^2, \ldots) = \mu_i^k.$$

Now that we have defined the desirable belief-preserving property, we are concerned with the existence of such a nice mapping? From **Kolmogorov's Extension Theorem**, we obtain the following lemma:

Lemma 1. *There exists a unique belief-preserving mapping $f : T^1 \to \Delta(\Theta \times T^0)$, and it is a homeomorphism.*

Common Knowledge of Coherency The function f of Lemma 1 assigns to each coherent hierarchy μ_i her belief over states and (possibly incoherent) hierarchies of the opponent. Since the opponent also has a coherent belief, we restrict the domain (from the original T^1) to $T^2 = \{\mu_i \in T^1 : f(\mu_i)(\Theta \times T^1) = 1\}$ consisting of all (coherent) hierarchies that are mapped to coherent hierarchies of the opponent. That is, the restricted $f|_{T^2} : T^2 \to \Delta(\Theta \times T^1)$ assigns to each hierarchy $\mu_i \in T^2$ her belief over states and *coherent* hierarchies of the opponent.

However, we are not yet done. The restricted $f|_{T^2}$ assigns to player $-i$'s hierarchy $\mu_{-i} \in T^2$ his belief $f|_{T^2}(\mu_{-i}) \in \Delta(\Theta \times T^1)$, but this allows player $-i$ to believe that player i has a hierarchy $\mu_i \in T^1$ in which player i may believe player $-i$'s hierarchy is incoherent. To prevent that—in other words, to ensure that player $-i$ assigns probability 1 to player i's hierarchies such that player i assigns probability 1 to player $-i$'s coherent hierarchies—we restrict the domain to $T^3 = \{\mu_i \in T^1 : f(\mu_i)(\Theta \times T^2) = 1\}$.

[9] This "each $(\theta, \mu_{-i}^1) \in X^1$" may not make rigorous sense, because if this measure $\text{marg}_{X^1} f(\mu_i) \in \Delta(X^1)$ is atomless, it assigns probability zero to each each $(\theta, \mu_{-i}^1) \in X^1$. We should define the probability assigned to any Borel set.

Note (Homeomorphism):

Definition 6. Two topological spaces X, Y are **homeomorphic** if there exists a function $h : X \to Y$ such that:

1. h is bijective (so that the inverse h^{-1} is well-defined).
2. h and h^{-1} are continuous.

This function h is called a **homeomorphism** (between X and Y).

However again and again.... We recursively define the following sets:

$$T^k = \left\{ \mu_i \in T^1 : f(\mu_i)(\Theta \times T^{k-1}) = 1 \right\} \qquad \forall\, k \geq 2.$$

Then, we define the set

$$T^* = \bigcap_{k \geq 1} T^k.$$

Intuitively, T^* is the set of hierarchies consistent with "common knowledge" of coherency. This is the **universal type space** (over the state space Θ) that we have been looking for.

Finally, we define the restricted $f^* \equiv f|_{T^*} : T^* \to \Delta(\Theta \times T^*)$. Then, it is immediate that the restricted f^* is belief-preserving and homeomorphism (between T^* and $\Delta(\Theta \times T^*)$).

References

Brandenburger, A., & Dekel, E. (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, *59*(1), 189–198.

Mertens, J.-F., & Zamir, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory*, *14*(1), 1–29.