On Information Design in Games

Mathevet, Perego and Taneva (2018)

Akira Matsushita

UTGSE M1

October 9, 2018

Overview

1. Summary

Settings

- This paper extends Kamenica and Gentzkow (2011)[1]'s single agent model into a multi-agent Bayesian game setting
- Each agent's utility depends on the unobservable state of the world θ and agents' action profile $\mathbf{a} = (a_i)$
- An information designer commits to a mechanism that observes the true state θ and sends signals to the agents
- The designer would like to find the best mechanism that maximizes the value function depending on θ and $\mathbf{a}$
- e.g. investment game under imcomplete information

[1] Emir Kamenica and Matthew Gentzkow. "Bayesian persuasion". In: *American Economic Review* 101.6 (2011), pp. 2590–2615.

Contributions

- Characterizing the feasible distributions of agents' beliefs that a designer can induce through the choice of information structure
  - A designer cannot persuade agents of arbitrary beliefs - there are constraints
- Eliciting the structure of agents' belief distributions and characterize the designer's problem in terms of it
  - Theorem 1 shows that any belief distribution that the designer can induce is a combination of basic communication schemes
  - And Corollary 1 suggests a two-step solution to the designer's problem: first selects an optimal components and second finds a best mixture of them

Contributions

- The results can apply to a variety of solution concepts and equilibrium selection rules
  - e.g. the case that agents only have bounded depths of reasoning, that they can deviate in coalitions, or that they can communicate
  - e.g. the robust mechanism that considers pessimistic equilibrium selection
- Finally we see an application to an investment game under an adversarial equilibrium selection (section 5)

2. Model

Model - 1

- Finite set of players: $N = \{1, 2, \dots, n\}$
- Uncertain state of the world $\theta \in \Theta$, where Θ is a finite set
- Finite action set of each player: A_i, and let $A = \prod_{i \in N} A_i$
- Utility function of each player: $u_i : A \times \Theta \to \mathbb{R}$, and let $u = \prod_{i \in N} u_i$
- θ is distributed according to $\mu_0 \in \Delta\Theta$, which is common knowledge
- Each player would like to maximize her expected payoff
- Refer to $G = (\Theta, \mu_0, N, A, u)$ as the **base game**

Model - 2

- A designer commits to disclosing information to the players about θ
- This is modeled by an **information structure** (S, π)
  - A finite set of signals that player i can receive: S_i, and $S = \prod_{i \in N} S_i$
  - An information map $\pi : \Theta \to \Delta S$
- For any state $\theta \in \Theta$, the message profile drawns from $\pi(s \mid \theta)$ and player i privately observes s_i
- The designer commits to a mechanism **before** she knows the true state θ

Model - 3

- The designer's payoff function: $v : A \times \Theta \to \mathbb{R}$
- The designer would like to maximize her expected payoff
- The combination of a base game and an information structure constitutes a Bayesian game $\mathcal{G} = \langle G, (S, \pi) \rangle$
- **Solution concept**: $\Sigma(\mathcal{G}) \subseteq \{\sigma = (\sigma_i) \mid \sigma_i : S_i \to \Delta A_i \text{ for all } i \in N\}$
  - Σ is a function that maps a Bayesian game to a set of equilibrium strategies
  - An equlibrium concept is arbitrary. e.g. BNE

Model - 4

- The resulting outcomes are represented by

$$O_\Sigma(\mathcal{G}) = \{\gamma \in \Delta(A \times \Theta) \mid \exists \sigma \in \Sigma(\mathcal{G})$$
$$\text{s.t. } \gamma(a, \theta) = \sum_s \sigma(a \mid s)\pi(s \mid \theta)\mu_0(\theta)\}$$

- Assumption: O_Σ is non-empty and compact-valued
  - Given G is finite, this holds when Σ is BNE (?)
  - Non-emptiness comes from the fact that $\Sigma(\mathcal{G}) \neq \emptyset$
  - $\Delta(A \times \Theta) = \{p \in \mathbb{R}_+^{|A| \times |\Theta|} \mid \sum_j p_j = 1\}$, so $O_\Sigma(\mathcal{G})$ is compact $\iff$ $O_\Sigma(\mathcal{G})$ is closed and bounded
  - If $\Sigma(\mathcal{G})$ is compact, then O_Σ is compact since γ is a linear function of σ
  - Boundedness comes from $\Sigma(\mathcal{G}) \subseteq \{\sigma = (\sigma_i) \in \mathbb{R}_+^{\prod_i (|A_i| \times |S_i|)} \mid \sum \sigma_i = 1\}$
  - (Maybe) $\Sigma(\mathcal{G})$ is closed...

Model - 5

- For a fixed game G, we just write $O_\Sigma(S, \pi)$ instead of $O_\Sigma(\mathcal{G})$
- When $O_\Sigma(\mathcal{G})$ is not singleton, the designer expects that one of them will happen, which is described by a **selection rule**:

$$g : D \ (\subseteq \Delta(A \times \Theta)) \mapsto g(D) \ (\in D)$$

  - e.g. A pessimistic designer, or one interested in robust information design, expects the worst outcome:

  $$g(D) = \arg\max_{\gamma \in D} \sum_{a,\theta} \gamma(a, \theta) v(a, \theta)$$

  for all compact (so a minimizer exists) $D \subseteq \Delta(A \times \Theta)$
  - Other rules, such as optimistic selection rule and random choise rule, can be considered

Model - 6

- Let $g^{(S,\pi)} = g(O_\Sigma(S,\pi))$
- Now the designer's expected payoff is given by

$$V(S,\pi) \equiv \sum_{a,\theta} g^{(S,\pi)}(a,\theta)v(a,\theta)$$

- And the information design problem is $\sup_{(S,\pi)} V(S,\pi)$

3. Belief Distributions

Belief Hierarchy

- A **belief hierarchy** t_i for player i is an infinite sequence $(t_i^1, t_i^2, \ldots)$, whose components are **coherent** beliefs of all orders
  - $t_i^1 \in \Delta(\Theta)$ is i's first order belief, $t_i^2 \in \Delta(\Theta \times (\Delta(\Theta))^{n-1})$ is i's second order belief, and so on

- A hierarchy t is **coherent** if any belief t_i^k coincides with all beliefs of lower order, $\{t_i^n\}_{n=1}^{k-1}$, on lower order events:

$$\underset{X_{k-1}}{\text{marg}}\ t_i^k \equiv \sum_{\Theta^{n-1}} t_i^k = t_i^{k-1}, \text{ where } X_{k-1} \equiv \text{supp } t_i^{k-1}$$

for all $k \geq 2$

Belief Hierarchy

- Whereas a player's belief hierarchies are coherent, they may assign positive probability to other players' belief hierarchies that are not coherent

- However Brandenburger and Dekel (1993)[2] showed that we can construct a set of coherent belief hierarchies T_i for every $i \in N$ such that there exists a homeomorphism (bijective and continuous mapping)

$$\beta_i^* : T_i \to \Delta(\Theta \times T_{-i})$$

  for all $i \in N$

  - β_i^* describes i's belief about (θ, t_{-i}) given t_i, and makes coherency common knowledge

- Define $T \equiv \prod_{i \in N} T_i$

[2] Adam Brandenburger and Eddie Dekel. "Hierarchies of beliefs and common knowledge". In: *Journal of Economic Theory* 59.1 (1993), pp. 189–198.

Belief Hierarchy

- Given (S, π) and μ_0, player i recieves s_i and use Bayes' rule to formulate beliefs $\mu_i(s_i) \in \Delta(\Theta \times S_{-i})$
- $\mu_i^1(s_i) \equiv \underset{\Theta}{\text{marg}}\ \mu_i(s_i)$: i's first-order belief about the state
- $\mu_i^2(s_i)$: i's second-order belief about the state, derived from i's first-order belief about s_{-i} and $j's$ $(\neq i)$ first-order beliefs about the state
- and so on...

Belief Hierarchy Distribution

- Every $s_i \in S_i$ induces a belief hierarchy $h_i(s_i) \in T_i$
- And so every $s \in S$ induces a profile of belief hierarchies $h(s) \equiv (h_i(s_i))_{i \in N}$

Definition (Definition 1)

An information structure (S, π) induces a distribution $\tau \in \Delta T$ over profiles of belief hierarchies, called a belief(-hierarchy) distribution, if

$$\tau(t) = \sum_{\theta} \pi(\{s \mid h(s) = t\} \mid \theta)\mu_0(\theta) \tag{3}$$

for all $t \in T$

Belief Hierarchy: Example

- The information structure given by the following table induces

$$\tau = \frac{3}{4} t_{1/3} + \frac{1}{4} t_1$$

when $\mu_0 \equiv \mu_0(\theta = 1) = \frac{1}{2}$, where t_μ is a hierarchy profile in which $\mu \equiv \mu(\theta = 1)$ is commonly believed

$\pi(\cdot\|0)$	s_1	s_2
s_1	1	0
s_2	0	0

$\pi(\cdot\|1)$	s_1	s_2
s_1	$\frac{1}{2}$	0
s_2	0	$\frac{1}{2}$

TABLE 1: A (Public) Information Structure

Public/Private Belief Distribution

- We categorize belief distributions into public and private

Definition (Definition 2)

A belief distribution τ is **public** if

- $t_i^1 = t_j^1$ for all $i, j \in N$
- $\operatorname*{marg}_{T_{-i}} \beta_i^*(\theta, t'_{-i} \mid t_i) = \mathbb{1}\{t'_{-i} = t_{-i}\}$ for all $i \in N$

for all $t \in \text{supp } \tau$ and $|\text{supp } \tau| \geq 2$. Otherwise τ is **private**

- we categorize the $|\text{supp } \tau| = 1$ degenerate case as private

Belief Manipulation: Example

- Consider the following 'guessing the state' game

Example

- Players: $N = \{1, 2\}$
- States: $\Theta = \{0, 1\}$
- Actions: $A_i = \{0, 1\}$ for $i = 1, 2$
- Utilities: $u_i(a_i, \theta) = -(a_i - \theta)^2$ for $i = 1, 2$
- Designer's payoff: $v = u_1 - u_2$
- $\mu_0(\theta = 0), \mu_0(\theta = 1) > 0$

Here the designer could obtain her maximal payoff of 1, if she could

- somehow reveal the state perfectly to player 1
- persuade player 2 that the opposite state has realized

Is it possible? - If they have a common prior, no.

Belief Manipulation: Common Priors

- Aumann (1976)[3] showed that Bayesian agents cannot agree to disagree if they have a common prior
- Say $p \in \Delta(\Theta \times T)$ is a **common prior** if

$$p(\theta, t) = \beta_i^*(\theta, t_{-i} \mid t_i) \times \underset{T_i}{\operatorname{marg}}\, p(t_i) \tag{4}$$

  for all θ, t and i
- That is, all players i obtain their belief map β_i^* by Bayesian updating of the same distribution p

[3] Robert J Aumann. "Agreeing to disagree". In: *The annals of statistics* (1976), pp. 1236–1239.

Belief Manipulation: Common Priors

- Denote by Δ^f the probability measures with finite support
- Define

$$\mathcal{C} \equiv \{\tau \in \Delta^f(T) \mid \exists \text{a common prior } \text{ s.t. } \tau = p\}$$

to be the space of **consistent** (belief-hierarchy) distributions

- In a consistent distribution, all players' beliefs arise from a common prior that draws every t with the same probability as τ
- Let p_τ be the unique distribution p in the above equation (uniquness comes from Mertens and Zamir (1985, Proposition 4.5)[4])

[4] Jean-François Mertens and Shmuel Zamir. "Formulation of Bayesian analysis for games with incomplete information". In: *International Journal of Game Theory* 14.1 (1985), pp. 1–29.

Belief Manipulation: Bayes plausible

- A distribution $\tau \in \Delta^f(T)$ is **Bayes plausible** if the expected first-order belief of at least one player equals the prior:

$$\forall \theta \in \Theta, \sum_{t_i} \operatorname*{marg}_{\Theta} \beta_i^*(\theta, t_{-i} \mid t_i)\tau_i(t_i) = \mu_0(\theta)$$

for some $i \in N$

Characterization of Beliefs

Proposition (Proposition 1)

There exists (S, π) that induces $\tau \in \Delta^f(T)$, if and only if, τ is consistent and Bayes plausible

- This proposition shows the limitation of the designer's belief manipulation
- it does not matter which player i satisfies Bayes plausibility, because by consistency, if it is true for one player, then it will hold for all

Characterization of Beliefs

- Before seeing the proof, let's go back to the example
- To reach the upper bound of 1 of the designer's payoff, it would have to be that

$$\beta_1^*(\theta = 1, t_2 \mid t_1) = 1 \wedge \beta_2^*(\theta = 0, t_1 \mid t_2) = 1$$

for some (t_1, t_2), which violates (4)

Consistency Problem

- Consistency is the main difference and technical challenge compared to the single-agent case
- One approach to implementing it is to design the individual distributions of players' beliefs and then couple them in a consistent way (Ely (2017) approach)
- Another approach: interpreting the designer's information design problem as a choice of a distribution over her own beliefs about the state, and then viewing players' first-order belief distributions as a special garbling of the designer's information

Consistency Problem

Proposition (Proposition 2)

If τ is consistent and Bayes plausible, then there exists $v : \operatorname{supp} \tau \to \Delta\Theta$ such that:

$$\sum_t \tau(t) v(t) = \mu_0$$

and

$$\sum_{t_{-i}} \tau(t_{-i} \mid t_i) v(t_i, t_{-i}) = \operatorname*{marg}_{\Theta} \beta_i^*(\cdot \mid t_i), \; \forall i, t_i.$$

Consistency Problem

Proposition (Proposition 2 (Cont.))

Conversely, if $\xi \in \Delta^f(\Delta\Theta)^n$ and $v : \text{supp } \xi \to \Delta\Theta$ satisfies

$$\sum_{\mu=(\mu_i)} \xi(\mu)v(\mu) = \mu_0$$

and

$$\sum_{\mu_{-i}} \xi(\mu_{-i} \mid \mu_i)v(\mu_{-i}, \mu_i) = \mu_i \ \forall i, \mu_i,$$

then there exist a consistent and Bayes plausible τ such that supp $\tau_i \simeq$ supp ξ_i (a bijection ϕ exists) and $\mu_i = \underset{\Theta}{\text{marg}}\ \beta_i^*(\cdot \mid \phi_i(\mu_i))$ for all i, and $\tau(t) = \xi(\phi^{-1}(t))$ for all t where $\phi = (\phi_i)$

4. Optimal Solutions

Some Assumptions

- **Assumption 1 (Linear Selection)**: g is linear
  - An assumption that requires the selection criterion to be independent of the subsets of outcomes to which it is applied
  - The best, worst and random selection satisfies linearlity
- **Assumption 2 (Invariant Solution)**: For all consistent τ, τ', if $\sigma \in \Sigma^B(\tau)$, then there exists $\sigma' \in \Sigma^B(\tau')$ such that $\sigma(t) = \sigma(t')$ for all $t \in \text{supp } \tau \cap \text{supp } \tau'$
  - This assumption says that play at a profile of belief hierarchies t under Σ^B is independent of the ambient distribution from which t is drawn

Some Assumptions

Proposition (Proposition 3)

If Σ^B is invariant, then O_{Σ^B} is linear

Representation Theorem

- Given any consistent distribution τ and the selected outcome $g^\tau \equiv g(O_{\Sigma^B}(\tau))$, the designer's ex ante expected payoff is given by

$$w(\tau) \equiv \sum_{\theta,a} g^\tau(a,\theta)v(a,\theta)$$

Theorem (Theorem 1)

The designer's maximization problem can be represented as

$$\sup_{(S,\pi)} V(S,\pi) = \sup_{\lambda\in\Delta^f(C^M)} \sum_{e\in C^M} w(e)\lambda(e)$$

$$\text{s.t.} \sum_{e\in C^M} \operatorname*{marg}_{\Theta} p_e\lambda(e) = \mu_0$$

Representation Theorem

- Theorem 1 states that the designer maximizes her expected payoff as if she were optimally randomizing over minimal consistent distributions, subject to posterior beliefs averaging to μ_0 across those distributions
- Every minimal distribution e induces a Bayesian game and leads to an outcome for which the designer receives expected payoff $w(e)$
- Every minimal distribution has a distribution over states, $\operatorname{marg}_{\Theta} p_e \lambda(e) = \mu_0$, and the further that is from μ_0, the more costly it is for designers
    - A kind of budget constraints

Within-Between Maximizations

Corollary (Corollary 1)

For any $\mu \in \Delta\Theta$, let

$$w^*(\mu) \equiv \sup_{e \in C^M | \underset{\Theta}{\mathrm{marg}}\ p_e = \mu} w(e)$$

Then the designer's maximization problem can be represented as

$$\sup_{(S,\pi)} V(S,\pi) = \sup_{\lambda \in \Delta^f \Delta\Theta} \sum_{\mathrm{supp}\ \lambda} w^*(e)\lambda(e)$$

$$\text{s.t.} \sum_{\mathrm{supp}\ \lambda} \mu\lambda(\mu) = \mu_0$$

- w^* is 'maximization within' and $\sup_{\lambda \in \Delta^f \Delta\Theta} \sum_{\mathrm{supp}\ \lambda} w^*(e)\lambda(e)$ is 'maximization between'

5. Application

Application: Adversarial Selection

- TBA