Information Design in Binary-Action Supermodular Games

Stephen Morris, Daisuke Oyama, and Satoru Takahashi

September 2019
Fix a decision problem. Which outcome can be induced if we design an information structure?

Kamenica and Gentzkow (2011) introduce the basic model of information design (Bayesian persuasion).

- The information designer commits himself to a signal-generating mechanism at no cost.
- Nature draws a state. A signal is generated.
- The agent observes the signal, and makes a decision.

This paper extends Kamenica-Gentzkow to multiple agents.
Two New Issues

A large class of signal-generating mechanisms: public vs private disclosure.

- With private signals, agents have beliefs and higher order beliefs about states.

Multiple equilibria in the induced game, leading to two different objective functions:

\[
\sup_{\text{signal-generating mechanism}} \left( \text{max or min}\right) \text{“info designer’s payoff”}.
\]

- **Sup-max**: the information designer maximizes his objective at the best equilibrium. \(\rightarrow\) Bayesian correlated equilibrium.
- **Sup-min**: the information designer maximizes his objective at the worst equilibrium. \(\rightarrow\) our paper.
Overview of Our Results

We focus on general games with binary actions and strategic complementarity.

We define $SI$ as the set of smallest-BNE implementable outcomes.

We characterize $SI$ by sequential obedience, a strengthening of standard obedience condition.

**Sequential obedience** can be simplified to coalitional obedience in state-wise potential games.

We use $SI$ to solve the sup-min information design problem.
Literature


Mathevet, Perego, and Taneva (2019). Analyzed a sup-min information design problem in an example with two players, two actions, two states, and symmetric payoffs.

Bergemann and Morris (2016). Characterized partially implementable outcomes (the sup-max information design problem) by obedience, an incomplete-information analogue of Aumann (1987).
The information design problem can be interpreted as robustness to information structures a la Kajii and Morris (1997).

- Kajii-Morris focus on near-complete information, whereas we consider general incomplete information.
- Some techniques in Kajii-Morris’ so called “critical path theorem” turn out to be useful for us.
Consider a single-agent information design problem.

\[
\begin{array}{c|c|c|c|c}
\hline
I & 2 & I & -2 \\
N & 0 & N & 0 \\
\hline
\end{array}
\]

\[\theta = G \text{ with prob } 10\% \quad \theta = B \text{ with prob } 90\%\]

The information designer maximizes the probability of \( I \).
Full vs No Disclosure

\[
\begin{array}{c|c|c|c|c}
  & I & 2 & I & -2 \\
\hline
N & 0 & 0 & N & 0 \\
\end{array}
\]

\(\theta = G\) with prob 10\% \hspace{1cm} \(\theta = B\) with prob 90\%

**No information disclosure.** The expected payoffs are given by

\[
\begin{array}{c|c|c|c|c}
  & I & -1.6 & I & -1.6 \\
\hline
N & 0 & 0 & N & 0 \\
\end{array}
\]

→ The agent plays I with probability 0\%.

**Full information disclosure.**

→ The agent plays I with probability 10\%. 
Partial Disclosure

\[
\begin{array}{c|cc}
I & 2 & -2 \\
N & 0 & 0 \\
\end{array}
\]

\[
\theta = G \text{ with prob } 10\
\theta = B \text{ with prob } 90\%
\]

**Partial information disclosure.** The information designer commits himself to the following signal-generating mechanism.

- If \( \theta = G \) realizes, then signal \( g \) is sent to the agent.
- If \( \theta = B \) realizes, then signal \( g \) is sent to the agent with probability \( 10\%/90\% = \varepsilon \); signal \( b \) is sent with the remaining probability.

Conditional on receiving signal \( g \), the agent strictly prefers \( I \).

\[
\begin{array}{c|cc}
I & \varepsilon' \\
N & 0 \\
\end{array}
\]

\( \rightarrow \) The agent plays \( I \) with probability arbitrarily close to 20%. 
A Two-Agent Example

Now we extend to two agents, an asymmetric variant of Mathevet, Perego, and Taneva (2019).

\[\theta = G \text{ with prob 10\%} \quad \theta = B \text{ with prob 90\%}\]

The information designer maximizes the probability of \((l, l)\) in the worst equilibrium.
Public Disclosure: Dominance

\[
\begin{array}{c|c|c|c|c}
& I & N \\
\hline
I & 3, 4 & 1, 0 \\
N & 0, 2 & 0, 0 \\
\end{array}
\quad
\begin{array}{c|c|c|c|c}
& I & N \\
\hline
I & -2, -1 & -4, 0 \\
N & 0, -3 & 0, 0 \\
\end{array}
\]

\[\theta = G \text{ with prob } 10\%\]
\[\theta = B \text{ with prob } 90\%\]

**Public disclosure.**
- If \(\theta = G\) realizes, then signal \(g\) is sent to the both agents.
- If \(\theta = B\) realizes, then signal \(g\) is sent to the both agents with probability \(2.5\%/90\% - \varepsilon\); signal \(b\) is sent with the remaining probability.

Conditional on \(g\), \(I\) is the **dominant** action for both agents.

\[
\begin{array}{c|c|c|c|c}
& I & N \\
\hline
I & 2 + \varepsilon', 3 + \varepsilon' & \varepsilon', 0 \\
N & 0, 1 + \varepsilon' & 0, 0 \\
\end{array}
\]

→ The agents play \((I, I)\) with probability close to **12.5\%**.
Public Disclosure: Iterative Dominance

Public disclosure.

- If $\theta = G$ realizes, then signal $g$ is sent to the both agents.
- If $\theta = B$ realizes, then signal $g$ is sent to the both agents with probability 6.6%/90%; signal $b$ is sent with the remaining probability.

Conditional on $g$, $(I, I)$ is iteratively dominant.

\[
\begin{array}{c|cc}
 & I & N \\
\hline
I & 1 + \varepsilon', 2 + \varepsilon' & -1 + \varepsilon', 0 \\
N & 0, \varepsilon' & 0, 0 \\
\end{array}
\]

→ The agents play $(I, I)$ with probability 16.6%. 

---

Public Disclosure: Iterative Dominance

Public disclosure.

- If $\theta = G$ realizes, then signal $g$ is sent to the both agents.
- If $\theta = B$ realizes, then signal $g$ is sent to the both agents with probability 6.6%/90%; signal $b$ is sent with the remaining probability.

Conditional on $g$, $(I, I)$ is iteratively dominant.

\[
\begin{array}{c|cc}
 & I & N \\
\hline
I & 1 + \varepsilon', 2 + \varepsilon' & -1 + \varepsilon', 0 \\
N & 0, \varepsilon' & 0, 0 \\
\end{array}
\]

→ The agents play $(I, I)$ with probability 16.6%. 

---
Private Noise: Risk Dominance

Public disclosure.

- If $\theta = G$ realizes, then signal $g$ is sent to the both agents.
- If $\theta = B$ realizes, then signal $g$ is sent to the both agents with probability 10%/90%; signal $b$ is sent with the remaining probability.

Conditional on $g$, $(I, I)$ is (weakly) risk-dominant.

$$
\begin{array}{ccc}
I & N \\
I & | & 0.5, 1.5 & -1.5, 0 \\
N & | & 0, -0.5 & 0, 0 \\
\end{array}
$$

By adding private noise à la email/global games, we can induce $(I, I)$ as a unique equilibrium outcome.

$\rightarrow$ The agents play $(I, I)$ with probability close to 20%.
Two issues remains.

- Can we achieve more than 20%? If not, how can we show that?
- Can we generalize (weak) risk dominance in games with more than two players?
General Framework

$I = \{1, \ldots, |I|\}$: the set of players.

$\Theta$: a finite set of states.

$\mu \in \Delta(\Theta)$: a **common** prior.

- Without loss of generality, we assume $\mu(\theta) > 0$ for any $\theta$.

$A_i = \{0, 1\}$: the **binary-action** set for player $i$.

- $A = \{0, 1\}^I$.

$u_i: A \times \Theta \rightarrow \mathbb{R}$: player $i$’s payoff, **supermodular**.
Information Structures

$T_i$: a countable set of signals for player $i$.
- $T = \prod_{i \in I} T_i$.

$P \in \Delta(T \times \Theta)$: a common prior.
- Without loss of generality, we assume $P(\{t_i\} \times T_{-i}) > 0$ for any $t_i$.
- **Consistency:** $P(T \times \{\theta\}) = \mu(\theta)$ for any $\theta \in \Theta$.

Given $(T, P)$, the notion of Bayesian Nash equilibrium $\sigma = (\sigma_i)_{i \in I}$, $\sigma_i: T_i \rightarrow \Delta(A_i)$, is defined as usual.

Let $\sigma_P \in \Delta(A \times \Theta)$ denote the induced outcome distribution:

$$\sigma_P(a, \theta) = \sum_t P(t, \theta) \prod_{i \in I} \sigma_i(t_i)(a_i).$$
Partial Implementability

Let $PI$ be the set of partially implementable outcomes:

$$PI = \{ \nu \in \Delta(A \times \Theta) \mid \nu = \sigma_{P} \text{ with some BNE } \sigma \text{ of some } (T, P) \text{ consistent with } \mu \}.$$  

Bergemann and Morris (2016) characterize $PI$ by Bayes correlated equilibrium, i.e.,

- **Consistency**: $\nu(A \times \{ \theta \}) = \mu(\theta)$ for any $\theta \in \Theta$.
- **Obedience**:

$$\sum_{a_{-i} \in A_{-i}, \theta \in \Theta} \nu(a_{i}, a_{-i}, \theta)(u_{i}(a_{i}, a_{-i}, \theta) - u_{i}(a'_{i}, a_{-i}, \theta)) \geq 0$$

for any $i \in I$ and $a_{i}, a'_{i} \in A_{i}$. 

Smallest-BNE Implementability

Let $SI$ be the set of smallest-BNE implementable outcomes:

$$SI = \{ \nu \in \Delta(A \times \Theta) \mid \nu = \sigma P \text{ with the smallest BNE } \sigma \text{ of some } (T, P) \text{ consistent with } \mu \}.$$ 

Note that for each $(T, P)$, by the supermodularity of $u$,

- the smallest BNE exists in pure strategies;\(^1\)
- the smallest BNE is the limit of iterative applications of best responses to constant 0 strategies;
- the limit is order independent, as long as best responses are applied to each player infinitely many times.

\(^1\)We define partial order $\sigma \succeq \sigma'$ if $\sigma_i(t_i)(1) \geq \sigma'_i(t_i)(1)$ for any $i \in I$ and $t_i \in T_i$. 
Ordering of Players

Let $\Gamma$ be the set of all finite sequences of distinct players.

- For example, if $I = \{1, 2, 3\}$, then

$$\Gamma = \{\emptyset, 1, 2, 3, 12, 13, 21, 23, 31, 32, 123, 132, 213, 231, 312, 321\}.$$

- For $\gamma \in \Gamma$, $\bar{a}(\gamma)$ denotes the action profile where player $i$ plays action 1 iff player $i$ appears in $\gamma$;

- Each $\nu_\Gamma \in \Delta(\Gamma \times \Theta)$ induces $\nu \in \Delta(A \times \Theta)$ by forgetting the ordering, i.e., $\nu(a, \theta) = \nu_\Gamma(\bar{a}^{-1}(a) \times \{\theta\})$.

Let $\Gamma_i = \{\gamma \in \Gamma \mid \text{player } i \text{ appears in } \gamma\}$.

- For $\gamma \in \Gamma_i$, $a_{-i}(\gamma)$ denotes the action profile of player $i$’s opponents where player $j$ plays action 1 iff player $j$ appears in $\gamma$ before player $i$. 
A Characterization of Smallest-BNE Implementability

We characterize smallest-BNE implementability by the following properties:

- **Consistency**: \( \nu(A \times \{\theta\}) = \mu(\theta) \) for any \( \theta \in \Theta \).

- **0-obedience**:

  \[
  \sum_{a_{-i} \in A_{-i}, \theta \in \Theta} \nu(0, a_{-i}, \theta)(u_i(0, a_{-i}, \theta) - u_i(1, a_{-i}, \theta)) \geq 0
  \]

  for any \( i \in I \).

- **Sequential obedience**: there exists \( \nu_\Gamma \in \Delta(\Gamma \times \Theta) \) that induces \( \nu \) and satisfies

  \[
  \sum_{\gamma \in \Gamma_i, \theta \in \Theta} \nu_\Gamma(\gamma, \theta)(u_i(1, a_{-i}(\gamma), \theta) - u_i(0, a_{-i}(\gamma), \theta)) > 0
  \]

  for any \( i \in I \) such that \( \nu_\Gamma(\Gamma_i \times \Theta) > 0 \).

  - Recall that \( a_{-i}(\gamma) \) denotes the action profile where player \( j \) plays action 1 iff player \( j \) appears in \( \gamma \) before player \( i \).
Sequential Obedience

Sequential obedience captures the iterative procedure at the outcome level.

Sequential obedience is a strengthening of 1-obedience, as

\[
\sum_{a_{-i}, \theta} \nu(1, a_{-i}, \theta) (u_i(1, a_{-i}, \theta) - u_i(0, a_{-i}, \theta)) \\
= \sum_{\gamma, \theta} \nu_\Gamma(\gamma, \theta) (u_i(1, \bar{a}_{-i}(\gamma), \theta) - u_i(0, \bar{a}_{-i}(\gamma), \theta)) \\
\geq \sum_{\gamma, \theta} \nu_\Gamma(\gamma, \theta) (u_i(1, a_{-i}(\gamma), \theta) - u_i(0, a_{-i}(\gamma), \theta)) \\
> 0,
\]

where \(\bar{a}_{-i}(\gamma)\) is the action profile of player \(i\)'s opponents where player \(j\) plays action 1 iff player \(j\) appears in \(\gamma\) (regardless of his relative position to player \(i\)).
The Main Result

**Theorem 1a.** Every $\nu \in SI$ satisfies consistency, 0-obedience, and sequential obedience.

We say that $\Theta$ is **rich** if there exists $\bar{\theta} \in \Theta$ such that $u_i(1, 0_{-i}, \bar{\theta}) > u_i(0, \bar{\theta})$ for any $i \in I$.

**Theorem 1b.** If $\Theta$ is rich, then every $\nu$ that satisfies consistency, 0-obedience, and sequential obedience is in the closure of $SI$.

Thus smallest-BNE implementability is essentially characterized by consistency, 0-obedience, and sequential obedience.

- By definition, we have $SI \subset PI$.
- Accordingly, we strengthen 1-obedience to sequential obedience.
- Similarly, we can characterize **full implementability** (outcomes that can be induced by the unique BNE) by consistency and “two way” sequential obedience.
The Proof of Theorem 1a

Fix any type space \((T, P)\) consistent with \(\mu\).

Apply best responses iteratively to constant 0 strategies. For each type \(t_i \in T_i\), if type \(t_i\) changes from action 0 to action 1 in the \(n\)-th step, we denote by \(n_i(t_i) = n\); if he never changes, then we denote by \(n_i(t_i) = \infty\).

Define
\[
\nu_\Gamma(\gamma, \theta) = \sum_{t: (n_i(t_i)) \text{ is ordered according to } \gamma} P(t, \theta),
\]
\[
\nu(a, \theta) = \nu_\Gamma(\bar{a}^{-1}(a) \times \{\theta\}).
\]

It is easy to show that \(\nu\) satisfies consistency and 0-obedience.
To show sequential obedience, note that for each $t_i \in T_i$ with $n_i(t_i) < \infty$, we have

$$\sum_{t_{-i}, \theta} P(t, \theta)(u_i(1, a_{-i}(t), \theta) - u_i(0, a_{-i}(t), \theta)) > 0,$$

where $a_{-i}(t)$ is the action profile of player $i$’s opponents where player $j$ plays action 1 iff $n_j(t_j) < n_i(t_i)$.

By adding up the inequality over all such $t_i$, we have

$$\sum_{\gamma \in \Gamma_i, \theta} \nu_{\Gamma}(\gamma, \theta)(u_i(1, a_{-i}(\gamma), \theta) - u_i(0, a_{-i}(\gamma), \theta))$$

$$= \sum_{t_i : n_i(t_i) < \infty} \sum_{t_{-i}, \theta} P(t, \theta)(u_i(1, a_{-i}(t), \theta) - u_i(0, a_{-i}(t), \theta))$$

$$> 0$$

for any $i \in I$ such that $\nu_{\Gamma}(\Gamma_i \times \Theta) > 0$. 
A Sketch of the Proof of Theorem 1b

We construct an information structure as follows.

- With probability $1 - \varepsilon$, we draw $\gamma$ according to $\nu_{\Gamma}$, and inform each player $i$ of

  $$t_i = \begin{cases}    m + \text{ranking of } i \text{ in } \gamma & \text{if } \gamma \in \Gamma_i, \\    \infty & \text{otherwise,} \end{cases}$$

  where $m$ is drawn from the geometric distribution on $\mathbb{N} = \{0, 1, 2, \ldots\}$ with pmf $\eta(1 - \eta)^m$.

- With the remaining probability $\varepsilon$, we inform each player of
  - $t_i = \infty$ if $\theta \neq \bar{\theta}$;
  - $t_i = \tau$ if $\theta = \bar{\theta}$, where $\tau$ is drawn uniformly from $\{1, \ldots, |l| - 1\}$.

Mimicking the arguments in the email/global game literature, we can show that each player of type $t_i < \infty$ plays action 1 in any equilibrium.
State-wise Potential Games

Sequential obedience is a system of linear inequalities involving super-exponentially many variables \((\approx 2.7 \times |l|! \times |\Theta|)\). Can we reduce the size of linear inequalities?

Suppose that for each \(\theta\), \(u(\cdot, \theta)\) admits a potential \(\Phi(\cdot, \theta): A \to \mathbb{R}\):

\[
u_i(1, a_{-i}, \theta) - u_i(0, a_{-i}, \theta) = \Phi(1, a_{-i}, \theta) - \Phi(0, a_{-i}, \theta)
\]

for any \(i \in I\) and \(a_{-i} \in A_{-i}\).
Coalitional Obedience

For $\nu \in \Delta(A \times \Theta)$ and $a \in A$, define

$$\Phi_\nu(a) = \sum_{a', \theta} \nu(a', \theta) \Phi(a \land a', \theta).$$

**Coalitional Obedience:** $\Phi_\nu(1) > \Phi_\nu(a)$ for any $a \neq 1$ such that $\nu(\{a\} \times \Theta) > 0$.

**Theorem 2.** In a state-wise potential game, sequential obedience is equivalent to coalitional obedience.

The number of variables is reduced to **exponential**, $2^{|I|} \times |\Theta|$. 
Revisit to the Two-Agent Example

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>3,4</td>
<td>1,0</td>
</tr>
<tr>
<td>N</td>
<td>0,2</td>
<td>0,0</td>
</tr>
</tbody>
</table>

θ = G with prob 10%

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-2,-1</td>
<td>-4,0</td>
</tr>
<tr>
<td>N</td>
<td>0,-3</td>
<td>0,0</td>
</tr>
</tbody>
</table>

θ = B with prob 90%

Note that

$$\sup_{(T,P)} \max_{\sigma : \text{BNE}} \text{“probability of } (I,I)\text{”} = \sup_{\nu \in PI} \nu(\{(I,I)\} \times \Theta),$$

$$\sup_{(T,P)} \min_{\sigma : \text{BNE}} \text{“probability of } (I,I)\text{”} = \sup_{\nu \in SI} \nu(\{(I,I)\} \times \Theta).$$
Without loss of generality, we focus on \( \nu \in \Delta(A \times \Theta) \) in the form of:

\[
\begin{array}{c|c|c}
I & N \\
\hline
I & x & 0 \\
N & 0 & 0.1 - x \\
\end{array}
\quad
\begin{array}{c|c|c}
I & N \\
\hline
I & y & 0 \\
N & 0 & 0.9 - y \\
\end{array}
\]

\[\theta = G\quad \theta = B\]

Let \( \nu_\Gamma \in \Delta(\Gamma \times \Theta) \) be given by

\[
\begin{align*}
\nu_\Gamma(12, G) &= x_{12}, & \nu_\Gamma(21, G) &= x_{21} = x - x_{12}, \\
\nu_\Gamma(12, B) &= y_{12}, & \nu_\Gamma(21, B) &= y_{21} = y - y_{12}.
\end{align*}
\]
0-obedience:

\[(0.1 - x) \times 1 + (0.9 - y) \times (-4) \leq 0,\]
\[(0.1 - x) \times 2 + (0.9 - y) \times (-3) \leq 0.\]

1-obedience:

\[x \times 3 + y \times (-2) \geq 0,\]
\[x \times 4 + y \times (-1) \geq 0.\]
Together with non-negativity constraints ($0 \leq x \leq 0.1$ and $0 \leq y \leq 0.9$), we have

$$\nu \in PI \iff 0 \leq x \leq 0.1 \text{ and } 0 \leq y \leq 1.5x.$$ 

Therefore,

$$\sup_{\nu \in PI} (x + y) = 0.25.$$
Sequential obedience: if $x > 0$ or $y > 0$, then

\[
x_{12} \times 1 + x_{21} \times 3 + y_{12} \times (-4) + y_{21} \times (-2) > 0,
x_{12} \times 4 + x_{21} \times 2 + y_{12} \times (-1) + y_{21} \times (-3) > 0.
\]

Adding them up (together with $x_{12} + x_{21} = x$ and $y_{12} + y_{21} = y$), we have

\[5x - 5y > 0.\]
In fact, it is easy to show that

\[ \nu \in \text{cl } SI \iff 0 \leq x \leq 0.1 \text{ and } 0 \leq y \leq x. \]

Therefore,

\[ \sup_{\nu \in SI} (x + y) = 0.2. \]
Conclusion

In binary-action supermodular games, we characterized smallest-BNE implementability by consistency, 0-obedience, and sequential obedience.

- **Sequential obedience** captures the iteration procedure at the outcome level.
- **Sequential obedience** may be difficult to compute, but can be simplified to coalitional obedience in state-wise potential games.

We used $SI$ to solve the sup-min information design problem.
“All or Nothing”

Assume that \( u \) is additively separable:

\[
\begin{aligned}
    u_i(a, \theta) &= \begin{cases} 
        \theta + h(m) - c_i & \text{if } a_i = 1, \\
        0 & \text{if } a_i = 0
    \end{cases}
\end{aligned}
\]

with \( m = \sum_i a_i \) and hence

\[
\begin{aligned}
    \Phi(a, \theta) &= m\theta + \sum_{k=1}^{m} h(k) - \sum_i a_i c_i.
\end{aligned}
\]

Assume also that the information designer’s objective is independent of \( \theta \) and convex in \( m \).

**Theorem 3.** In the above game, we can assume wlog \( \nu(a, \theta) = 0 \) for any \( a \neq 0, 1 \) if and only if “\( h(\cdot) \) increases fast” and “\( c_1, \ldots, c_{|I|} \) are not so asymmetric”.

- The number of variables is further reduced to linear, \( |\Theta| \).