# Information Design: Single Agent

Renjie Zhong

Renmin University of China

December 2023

# Outline

# Illustrative Example

# Illustrative Example[1]


[1] From Kamenica and Gentskow (2011); Credit to Starkov′s slides

# Illustrative Example

## Prosecutor′s Persuasion: Call up a Witness

1. A suspect is on trial, accused of murder.
2. Judge must decide whether to convict or acquit him, wants to make the right decision.
3. Prosecutor is paid per cases won, so wants to convict the suspect regardless of guilt.
4. Prosecutor can call up a witness. What kind of witness should he summon?

Frame the above as an information design problem.

This example:

1. Designer = prosecutor
2. State $\omega \in \{G, N\}$ represents true guilt.
3. Let $\phi_0 = \mathbb{P}\{\omega = G\}$ denote the common prior belief (probability that the prosecutor and the judge assign to the suspect being guilty).
4. $N = 1$ (judge), $A = \{g, n\}$ (verdicts "guilty", "not guilty");
5. The judge's utility is $v_1(a, \omega) = \mathbb{I}(a = \omega)$ .
6. The prosecutor's objective function is $v_0(a, \omega) = \mathbb{I}(a = G)$ .
7. The witness was at a certain place on the night of murder - this determines $\mu$
    1. If W was around the place of murder, can confirm or deny the suspect was there.
    2. If W was in a random pub, can do the same, but this conveys different information.

## Timing

To be clear, the timing in this example (as well as in the general model) is as follows:

1. prosecutor chooses the witness $\pi$ and publicly commits to it
2. state $\omega$ is determined
3. witness reveals message m to the court according to $\pi(m \mid \omega)$
4. the judge observes m and chooses decision a
5. payoffs are realized

- Let $\phi$ denote the judge's posterior belief (after she observes $m$ ). What action does she choose?
- Denote $\hat{a}(\phi) \equiv \arg\max \mathbb{E}_{\phi(\omega)}\left[v_1(a,\omega)\right]$
- If there are many optimal actions, choose the best for the prosecutor.
- For the first time ever we want to fix the tie-breaking rule. The reason will be evident later.
- In our example: $\hat{a}(\phi) = \begin{cases} g & \text{if } \phi \geq 1/2 \\ n & \text{if } \phi < 1/2 \end{cases}$
- Knowing $\hat{a}(\phi)$ means we can write the prosecutor's utility as a function of $\phi$ : let

$$V_0(\phi) \equiv v_0(\hat{a}(\phi)) = \begin{cases} 1 & \text{if } \phi \geq 1/2; \\ 0 & \text{if } \phi < 1/2. \end{cases}$$

- In our example: $\hat{a}(\phi) = \begin{cases} g & \text{if } \phi \geq 1/2 \\ n & \text{if } \phi < 1/2 \end{cases}$
- Knowing $\hat{a}(\phi)$ means we can write the prosecutor's utility as a function of $\phi$ : let

$$V_0(\phi) \equiv v_0(\hat{a}(\phi)) = \begin{cases} 1 & \text{if } \phi \geq 1/2; \\ 0 & \text{if } \phi < 1/2. \end{cases}$$

- By choosing an experiment $(\mu, M)$ the prosecutor induces some distribution $\tau$ over posteriors $\phi$
- Trick: forget about $\mu$ and focus on this distribution $\tau$ as the choice object
  1. What if P could choose any distribution? Would want $\phi \geq 1/2$ always (after any message $m$ ).
  2. So if $\phi_0 \geq 1/2$ then optimal for P to do nothing (choose uninformative experiment).
  3. But the ideal is unattainable if $\phi_0 < 1/2$ because beliefs must be consistent!

- When $\phi_0 < 1/2$, any improvable space?
- Suppose that the prosecutor commits that:
  1. if innocent - obfuscate it (mix between sending guilty signals with innocent signals with fixed, committed prob)
  2. if guilty - claim guilty
- In other words, the optimal strategy is:
  1. if state favorable to prosecutor then disclose it truthfully;
  2. if state bad for prosecutor then try to obfuscate it.
  3. Need commitment to mix in $\omega = N$; message $m = g$. gives higher payoff, so without commitment the prosecutor would never send $m = n$.
- The judge is granted full confidence when taking action that is undesirable for designer; is made barely indifferent when taking action desired by the designer.
- We then formalize it.

# Introduction

# Driving Forces of Micro Behavioral

- The micro-behavior of an agent depends on his beliefs $\mu_i$, his feasible choices $A_i$, the resulting final payoffs $u_i$, neighbors $G_i$ and some idiosyncratic constraints.
- Design Problem: Designer design the game structure to implement/realize optimal/revenue-maximizing outcome

# Design Problem

- Mechanism/Market/Network Design as Institution/Organization Design
  1. mechanism design with “monetary”incentives (transferable or non-transferable utility): steer the agent(s) decisions by changing their payoff consequences
  2. delegation/redistribution policy deisgn: steer the agent(s) decisions by constraining the set of feasible actions
  3. matching/market design without “money”: steer the agent(s) decisions by designing the rules whereby reports about preferences map to final allocations of objects
  4. network intervention/design: steer the agent(s) decisions by constraining the set of players/ changing their payoff consequences

# Information Design: Motivation

- An agent's beliefs are an important driver of his behavior and can be influenced by information transmission from another agent, motivating the problem of **information design**
- In information design, payoff functions and feasible outcomes (i.e., the game) taken as given
- Object of design: information of the agent(s)—hence, the beliefs driving choices
  1. different characteristics of information (public/private, hard/soft, ambiguous/certain)
  2. commitment/no commitment
  3. decision rule

# Information Design: Focus

- Desipite of different situations, we always concern these problems in information design:
  1. Feasibility: what is the scope for changing the agent's behavior by designing his information environment?
  2. Optimality: what is the optimal information for the agent from the viewpoint of its designer?
  3. Welfare: when persuasion is beneficial/detrimental to the sender/receiver?
  4. Robustness

# Group Persuasion: Focus

- With multiple agents, we also care about the timing/sequence of the persuasion
- How setup affects the information releavation?
  1. the alignment/congruence of preference between senders and receivers/ within senders/receivers
  2. the number of receivers/senders
  3. the correlation structure of the information designed

# Bayesian Persuasion

- Bayesian Persuasion impose a critical assumption on the general information design: commitment
- We can intepretate it as a persuasion problem under the constraint on information structure – bayesian plausibility (martingale property/consistency/committment)
- Other interpretations:
  1. Correlation games: a correlated recommendation system
  2. Persuasion economics: duality theory
  3. Behavioral economics: a dynamically inconsistency model
  4. Operation Researches: optimal transportation problem

# This Lecture

- This mini lecture focuses on:
  1. single agent information design (setup and interpretations)
  2. a very simple but comprehensive survey of **methods and perspectives**
  3. several applications

# Related Surveys and Notes

- Kamenica (2019; 2022): concavification, its extensions (multiple players and dynamics) and leading economic examples
- Bergemann and Morris (2019)
  1. literal: optimal choice of information structure
  2. metaphorical: optimal (action recommendation) mechanism under different information structures
- Bergemann and Bonatti (2019): a framework of information selling
- Lecture note/slides:
  1. Introductory slides and focusing on BCE and Concavification: Morris-Bonn Lectures(2018),Sandomirskiy(2020),Starkov(2022)
  2. A systematical exploration and focusing on single/multiple agent(s): Galerpti(2022)

# Benchmark Model

# Benchmark Model: Setup[4]

- Players: **one** sender $S$/receiver $R$ with prior [2] $\mu_0 \in \Delta\Omega$,
- Notations: $\omega \in \Omega$: state space; $v(a,\omega), u(a,\omega)$: sender,receiver′s payoff
- Action[3] Space: $\pi : \Omega \to \Delta S$ ($S$: the set of signal realizations)
  1. zero marginal/common fixed cost of signals
  2. all information structures are feasible
  3. public signals
- Updating Rules: **Bayesian Updating** $\mu = \frac{\pi(s|\omega)\mu_0(\omega)}{\sum_{\omega' \in \Omega} \pi(s|\omega')\mu_0(\omega')}$

  Every signal induces a posterior $\Rightarrow$ the experiment induces a distribution of posteriors

[2] we take the fashion of (Alonso and Camara,2016a) without adding too much complexity to (Kamenica and Gentskov,2011)

[3] also called signal structure, information structure, experiment, Blackwell experiment, or data-generating process

[4] an axiomic representation in Jakobsen(2021)

# Benchmark Model: Setup

- Game Rules:
  1. commitment power[5][6]: denote $\tau(\mu) = \sum_{\omega\in\Omega}\sum_{s:\mu_s=\mu}\pi(s \mid \omega)\mu_0(\omega)$, then $\sum_{\mu\in \text{ supp } \tau}\mu(\omega)\tau(\mu) = \mu_0(\omega), \quad \omega\in\Omega$
  2. Designer-preferred equilibrium: choose $a \in A^*(\mu) = \arg\max_{a\in A}\mathbb{E}_\mu[u(a,\omega)]$ that maximizes $\mathbb{E}_\mu[v(a,\omega)]$ when $|A^*(\mu)| \geq 2$
- Timeline: designer commits $\pi \Rightarrow \omega$ realizes $\Rightarrow$ agent observes $s$, updates her belief and chooses her action $\Rightarrow$ payoffs realized
- Statics: only one period

[5] all called bayesian plausibility/consistency/martingale property (especially in dynamic setting (Ely et al., 2015))

[6] compared to other info, we can interpret it as no signaling through info structure, signals with objective meaning, info transmission with reputation foundation (Best and Quigley(2017), Mathevet et al.(2019))

# Setup: Returning to the Introductory Example

- When $\phi_0 = 0.3 < 0.5$, any improvable space?
- It only matters for $V_0(\phi)$ whether $\phi < 1/2$ or $\phi \geq 1/2$ .
- So suppose there are two possible posteriors induced by the experiment: $\phi_1 < 1/2$ and $\phi_2 \geq 1/2$ , occurring with respective probabilities $\tau_1$ and $\tau_2 = 1 - \tau_1$.
- P gets payoff 1 whenever $\phi_2$ is induced and 0 in case of $\phi_1$ :

$$\mathbb{E}_\phi V_0(\phi) = \tau_1 \cdot 0 + (1 - \tau_1) \cdot 1 = 1 - \tau_1$$

Consistency:

$$\tau_1 \phi_1 + (1 - \tau_1)\, \phi_2 = \phi_0$$
$$\Leftrightarrow \tau_1 = \tfrac{\phi_2 - \phi_0}{\phi_2 - \phi_1} = 1 - \tfrac{\phi_0 - \phi_1}{\phi_2 - \phi_1}$$

# Setup: Returning to the Introductory Example

- designer′ problem:

$$\mathbb{E}_{\phi} V_0(\phi) = 1 - \tau_1$$

$$\text{Consistency: } \tau_1 = \frac{\phi_2 - \phi_0}{\phi_2 - \phi_1} = 1 - \frac{\phi_0 - \phi_1}{\phi_2 - \phi_1}$$

- optimal signal: $\phi_1 = 0$ and $\phi_2 = 1/2$

$$\frac{\phi_0 \pi(n \mid G)}{\phi_0 \pi(n \mid G) + (1 - \phi_0)\, \pi(n \mid N)} = \phi_1 = 0$$

$$\frac{\phi_0 \pi(g \mid G)}{\phi_0 \pi(g \mid G) + (1 - \phi_0)\, \pi(g \mid N)} = \phi_2 = 1/2$$

$$\pi(n \mid N) + \pi(g \mid N) = 1$$

$$\pi(n \mid G) + \pi(g \mid G) = 1$$

- optimal signal: $\begin{matrix} \pi(n \mid N) = \frac{4}{7} & \pi(n \mid G) = 0 \\ \pi(g \mid N) = \frac{3}{7} & \pi(g \mid G) = 1. \end{matrix}$

# Belief Approach

# Concavification (KG,2011)

**Sender′ Problem**

Define $\hat{v}(\mu) = \mathbb{E}_\mu[v(\hat{a}(\mu), \omega)]$, then sender′ problem is:

$$v^* = \max_\tau \sum_{\mu \in \operatorname{supp} \tau} \hat{v}(\mu)\tau(\mu)$$
$$\text{s.t. } \sum_{\mu \in \operatorname{supp} \tau} \mu\tau(\mu) = \mu_0$$

- Concavification: $[\mathrm{CAV}(f)](\mu) = \sup\{z \mid (\mu, z) \in \mathrm{co}(f)\}$
- Considering $[\mathrm{CAV}(\hat{v})](\mu)$ and pick up the optimal signal![7]

[7] We need two additional assumptions to guarantee the convex combination, one is the willingness to share $(\exists \mu, \hat{v}(\mu) > \mathbb{E}_\mu\left[v\left(\hat{a}\left(\mu_0\right), \omega\right)\right])$, another is a technical assumption called "local continuity" at $\mu_0$ $(\exists \varepsilon > 0 s.t. \mathbb{E}_\mu[u(\hat{a}(\mu), \omega)] > \mathbb{E}_\mu[u(a, \omega)] + \varepsilon, \quad \forall a \neq \hat{a}(\mu))$

# Concavification: A Graphical Illustration

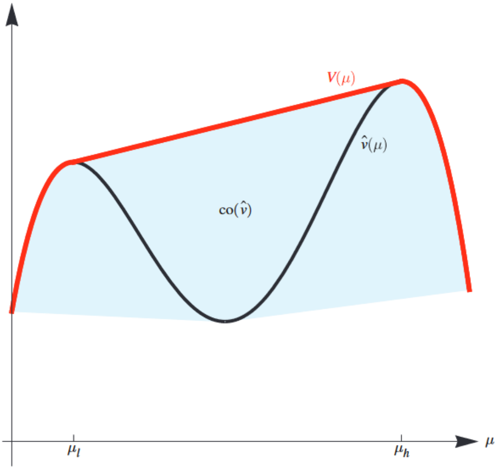


Figure 1. An Illustration of Concave Closure

# Concavification: Returning to the Introductory Example

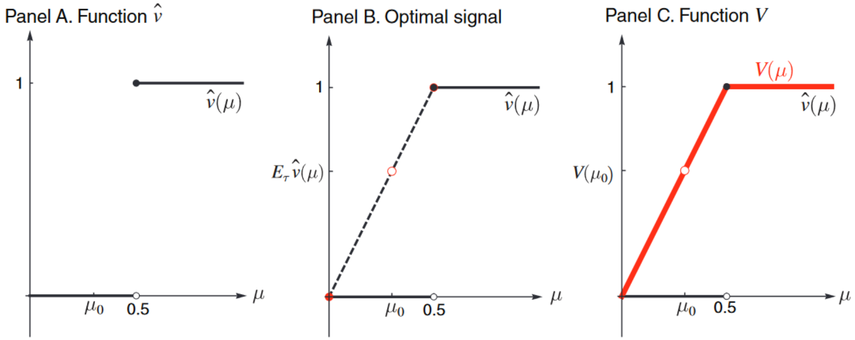


FIGURE 2. THE MOTIVATING EXAMPLE

# Pros and Cons of Concavification

- pros: always nice interpretation and allows for convex analysis; robustness to the analysis of various extensions
- cons: difficulty to characterize/derive the optimal signals; high dimensions/multiple players worsens it

# Birdview of Other Approaches

- Other approaches[8]:
  1. Myersonian and Duality Approach(Bergemann and Morris,2016; Morris,2019; Kolotilin,2018;Dworczak and Kolotilin,2023)
  2. For some speicificed utility function (posterior mean or quantile):
     1. Rothschild-Stiglitz Approach (Gentzkow and Kamenica,2016)
     2. Price Approach (Duality) (Dworczak and Martini,2019)
     3. Majorization and Extremal Distributions (Kleiner et al,2021; Yang and Zentefis,2023)
  3. Dynamic Inconsistency (Jakobsen,2021)
  4. Optimal Transportation (Arieli et al,2022)
- Considering the time limit, we only introduce the Myersionian and Duality Approach

[8] not mention some computational methods like Dughmi and Xu(2016) and Dughmi(2017)

# Myersonian Approach and Duality

- Bergemann and Morris (2016,2018) transform the seletion of optimal signals into the selection of optimal mechanism under different information structures
- Each $\pi(\cdot \mid \omega)$ induces a distribution $x(\cdot \mid \omega) \in \Delta(A)$ over actions:

$$x(a \mid \omega) = \sum_{s:\hat{a}(\mu_s)=a} \pi(s \mid \omega)$$

**Primal problem**

choose $x \in \mathbb{R}^{A \times \Omega}$ to solve:

$$\max \mathcal{V}(x) = \sum_{\omega \in \Omega, a \in A} v(a, \omega) x(a \mid \omega) \mu_0(\omega) \quad \text{s.t.}$$

1. (O) Obedience: $\sum_{\omega \in \Omega} \left[u(a, \omega) - u\left(a', \omega\right)\right] x(a \mid \omega) \mu_0(\omega) \geq 0$ for all $a, a' \in A$
2. (C) Consistency: $\sum_{a \in A} x(a \mid \omega) = 1$ for all $\omega \in \Omega$
3. (NN) Non-negativity: $x(a \mid \omega) \geq 0$ for all $(a, \omega) \in A \times \Omega$

## Dual problem

choose $p \in \mathbb{R}^{\Omega}$ and $\lambda \in \mathbb{R}^{A\times A}$ to solve

$$\min \mathcal{V}^*(p,\lambda) = \sum_{\omega\in\Omega} p(\omega)\mu_0(\omega) \quad \text{s.t.}$$

1. $(\lambda - NN)\lambda$ -Non-negativity: $\lambda(a' \mid a) \geq 0$ for all $(a, a') \in A \times A$
2. (DC) Dual constraint: for all $(a,\omega) \in A \times \Omega$
   $p(\omega) \geq v(a,\omega) + \sum_{a'\in A}\left[u(a,\omega) - u(a',\omega)\right]\lambda(a' \mid a)$

1. associate to each $\omega$ a monopolist seller of $\omega$-quality probability
2. designer buys probability $\pi(s|\omega)$ from seller $\omega$, whose stock is $\mu_0(\omega)$
3. designer pays unitary price $p(\omega)$ to seller
4. goal = minimize value of extra unit of probability evenly spread across sellers/states (otherwise, current stock not used in best way)

# Complementary Slackness

Suppose $x$ satisfies (O), (C) , and (NN) and $(p, \lambda)$ satisfy $(\lambda - \text{NN})$ and (DC).

Then, $x$ and $(p, \lambda)$ optimal iff

1. for all $a, a' \in A$ $\lambda\left(a' \mid a\right)\left[\sum_{\omega \in \Omega}\left[u(a, \omega)-u\left(a', \omega\right)\right] x(a \mid \omega) \mu_0(\omega)\right]=0$

   $\Rightarrow \lambda\left(a' \mid a\right)>0$ only if agent indifferent when recommended $a$

2. for all $(a, \omega) \in A \times \Omega$

   $x(a \mid \omega) \mu_0(\omega)\left[p(\omega)-v(a, \omega)-\sum_{a' \in A}\left[u(a, \omega)-u\left(a', \omega\right)\right] \lambda\left(a' \mid a\right)\right]=0$

   $\Rightarrow a \notin \operatorname{supp} x(a \mid \omega)$ if dual constraint cannot hold with equality.

Note: CS provides connection from dual variables to solution $x$.

# Other Perspectives

- Gentzkow and Kamenica(2016) proposes a way to tackle a special class with uncountable state spaces and Sender's payoff depends only on the mean of Receiver's posterior.
- Dworczak and Martini(2018) proposes a price-theoretic approach to Bayesian persuasion by establishing an analogy between the Sender's problem and finding Walrasian equilibria of a Persuasion Economy.
- Kleiner et al(2021), Yang and Zentefis(2023) characterize the set of extreme points of monotonic functions for majorization/ FOSD intervals, thus discovering the underlying solution structure of persuasion problem.
- Arieli et al(2022) reduce the persuasion problem to the Monge-Kantorovich problem of optimal transportation.

# Applications

# Applications

Examples of such problems (Kamenica,2019):

- financial sector stress tests (Goldstein and Leitner,2018; Inostroza-Pavan,2018; Orlov et al.,2018)
- grading in schools (Boleslavsky-Cotton,2015; Ostrovsky-Schwarz,2010)
- employee feedback (Habibi,2018; Smolin 2017)
- law enforcement deployment (Hernandez-Neeman 2017; Lazear,2006; Rabinovich et al.,2015)
- censorship (Gehlbach-Sonin,2014)
- entertainment (Ely et al.,2015)
- voter coalition formation (Alonso-Camara 2016)
- research procurement (Yoder 2018)
- medical research or testing (Kolotilin 2015,Schweizer-Szech 2019)
- matching platforms (Romanyuk-Smolin 2019)
- price discrimination (Bergemann et al.,2015)
- insurance (Garcia-Tsur,2018)
- transparency in organizations (Jehiel,2015)
- contest design (Zhang and Zhou,2016)