Notes on Kamenica (2019)

Bayesian Persuasion and Information Design

Quan Li

September 3, 2021

Abstract

This article review the literature that answers the questions of what is the optimal information that should be revealed.[1]

1 What is Bayesian Persuasion?

Economic behavior may be driven by three factors:

- Preference
- Technology
- Information

Consequently, there are three broad ways of influencing economic outcomes:[2]

[1]For instance, *"A school may improve its students' job outcomes if it issues only coarse grades. Google can reduce congestion on roads by giving drivers noisy information about the state of traffic. A social planner might raise everyone's welfare by providing only partial information about solvency of banks. All of this can happen even when everyone is **fully rational** and **understands the data-generating process**."*

[2]These methods are not necessarily isolated, Li (2017) add transfers to the persuasion model, Lewis and Sappington (1994) study a monopolist who chooses both what price to charge and what information to provide to consumers about their valuations for the firm's product.

- Change the (induced) preference over actions via incentives
  - contingent payments
  - threat of violence
  - supply of complementary goods
- Technology: make it easier for a decision maker to achieve it[3]
- **Persuasion: influencing behavior via provision of information (Kamenica and Gentzkow, 2011)**

Bayesian Persuasion: **Bayesian** means the decision maker is standard, i.e. she understands how information is generated and react to information in a rational (**Bayesian**) manner.[4]

The comparison between *Information Design* and *Mechanism Design* (?Taneva, 2019)

- **Mechanism Design**: the allocation of information (i.e., who knows what) is given, the designer select the game that the agents will play to influence the outcome.
- **Information Design**: the game is given, the designer influences the outcome by specifying the allocation of information.

Bayesian Persuasion is a communication protocol

Before, we have models of communication:

- Cheap talk (Crawford and Sobel, 1982)
- Verifiable message (Grossman, 1981; Milgrom, 1981)
- Signaling game (Spence, 1978)

[3]The methods used in nudging and choice structure (Thaler and Sunstein, 2008) can be seen as technological interventions, as argued by Kamenica (2012). For instance, teaching the taxi drivers to open the car door with their right hand in order to reduce the chance that they cause an accident can be regarded as a technological innovation.

[4]It is also referred to as *Information Design*, but the former is probably used more when the designer is one of the player and there is a single receiver, while the latter is used more when the designer is a social planner or there are multiple receivers.

The full commitment formulation of Bayesian Persuasion

Relative the above models, Bayesian Persuasion endows the sender with more commitment power. It allows the sender to commit to sending any distribution of messages as a function of the state of the world. [5]

The equivalence between the alternative models and Bayesian persuasion (??) (Gentzkow and Kamenica, 2017)

Other surveys of Information Design

This review exclusively focus on the desire of the sender to influence.

Bergemann and Bonatti (2019) studies the markets where a seller design information to sell it. [6]

The relationship with Bayes correlated equilibria (Bergemann and Morris, 2013)

Bayes correlated equilibria take as given a basic game:

- a set of players
- a set of feasible actions for each player
- players' payoffs as a function of the state of the world and the actions taken

Describe the set of all possible outcomes that could arise (as Bayes Nash equilibria) ***regardless of what each player knows(about the state and the other player know)***

It predicts the outcome of the basic game that is robust to the uncertainty about the knowledge, and by definition, it coincides the outcome can be attained through information design.

A Bayesian persuasion problem

[5] Min (2020) relax the assumption, study the results under *partial commitment*

[6] For example, Bergemann et al. (2018) and Kastl et al. (2018)

$$\Updownarrow$$

A problem of selecting an optimal *Bayes correlated equilibrium* given an objective function

More applied strand

Include the following interesting papers:

- Financial stress tests (Goldstein and Leitner, 2018)
- Grading in schools (Ostrovsky and Schwarz, 2010)
- Matching platforms (Romanyuk and Smolin, 2019)
- Price discrimination (Bergemann et al., 2015)

2 The Model and its Interpretations

2.1 The Basic Model

The specification:

- Agents: {Sender, Receiver}
- Action: $a \in A$
- State: $\omega \in \Omega$
- Utility:
  - Sender: $v(a, \omega)$[7]
  - Receiver: $u(a, \omega)$
- Belief: common prior $\mu_0 \in \Delta\Omega$

[7]We may be willing to view v as the *welfare function* if we regard the sender as a social designer.

The key object: *Signal* [8]

- Signal realizations: S[9]
- Signal: $\pi : \Omega \rightarrow \Delta(S)^{10} \Leftrightarrow {}^{11}\Delta(\Omega \times S) \in \Pi$

Timing

1. Sender chooses a signal π
2. Receiver observes which signal was chosen
3. ω is realized by μ_0
4. s is realized according to $\pi(\omega)$[12]
5. Receiver observes the realized s
6. Receiver takes action a

The receiver's mechanical behavior

- Bayesian Updating:

$$\mu_\pi(\omega|s) = \frac{\pi(s|\omega)\mu_0(\omega)}{\sum_{\omega'} \pi(s|\omega')\mu(\omega')}$$

- Optimization[13]:

$$a^*(\mu_\pi(\cdot|s)) \in argmax_{a\in A}\mathbb{E}_{\omega\sim\mu_\pi(\cdot|s)}u(a,\omega)$$

[8] It goes by many other names, including signal structure, information structure, experiment, Black-well experiment, data-generating process

[9] it suffices to assume $|S| \geq min\{|A|,|\Omega|\}$, (??) *Is there some economic consider relevant to care about the cardinality of S?*

[10] It takes form like $P(s|\omega)$

[11] For $P(s|\omega) = \frac{P(s,\omega)}{\mu_0(\omega)}$, and μ_0 is given at the very beginning

[12] Remember $\pi(\omega) \in \Delta(S)$

[13] The receiver's strategy is a mapping from Π to the mapping from the set of signal realizations S to the set of actions A

The sender's backward induction problem

$$\pi^* \in argmax_{\pi \in \Pi} \mathbb{E}_{\omega \sim \mu_0} \mathbb{E}_{s \sim \pi(\omega)} v\left(a^*\left(\mu_\pi(\cdot \mid s)\right), \omega\right) \tag{1}$$

What should we care about?

The model is quite general, it may be daunting or meaningless to find out a general optimal signal. So we are willing to recasting the problem to make it more precise and more approachable.(See Section 3)

For the following several reasons:

- Π is a pretty large set
- The choice of π influences the sender's payoff both by changing the second expectation ($\mathbb{E}_{s \sim \pi(\omega)}$) and by the optimal action $a^*\left(\mu_\pi(\cdot \mid s)\right)$ chosen by the receiver

2.2 Interpretations

The courtroom example

- Agents: **{prosecutor, judge}**
- Action: **convict or not**
- State: **the guilt of the defendant**
- Utility:
  - Prosecutor: **always prefers conviction**
  - Judge: **match the truth**
- Signal: **{forensic tests, questions asked to witnesses}**[14]

[14]If the assumption of finding out the whole truth is implausible, we can redefine the state as the realizations of the most informative signal.

Grades in schools(??[15])

- Agents: **{school, the labor market}**
- Action: **placement of the student**
- State: **the ability of the student**
- Utility:
  - school: ague by yourself
  - the labor market: think about it by yourself
- Signal: **the school's grading policy**

It maybe combined with the job market matching model (Chen and Hu, 2020) and delegation model (Frankel, 2014) to study a model with grading policy and job market.

Speeding of drivers (Lazear, 2006)

Original model:

- Z miles of road
- A driver choose to speed or not on each mile
- Speeding utility $V \leq K$ Fine for Speeding
- There $G \leq Z$ police, and each can patrol one mile of road
- The police wish to minimize the number of miles over which drivers speed.

It is equivalent to the following Bayesian persuasion model:

- Agents: **{the police, the driver}**
- Action: **speed or not**

[15]There are some tricky details I cannot understand, for example, the author assume schools know the ability (true state) before choosing the grading policy.

- State: **the presence of a policeman on a given mile**
- Utility:
  -
  -
- Common belief: $\mu_0 = G/Z$
- Signal: **the predictability of the patrolling strategy**

Typical signals:

- Randomly set up the speed traps $\Rightarrow$ uninformative signal $\underline{\pi} \Rightarrow \mu_{\underline{\pi}}(\cdot|s) = G/Z$
- Always patrol the exact same locations $\Rightarrow$ fully informative signal $\overline{\pi} \Rightarrow \mu_{\overline{\pi}}(\cdot|s) = 0$ or 1

Other applications involve yet other interpretations of how a signal is generated (Best and Quigley, 2020).

The justification of commitment assumption

It varies substantially across applications.

3 Sender's Optimization Problem

3.1 Concavification

From the sender's view, what matters is the *posterior beliefs*, not the signals. So we reformulate the problem as follows:

An important observation[16]

$$\mathbb{E}_s \mu(\omega \mid s) = \mu_0(\omega)$$

[16] $\mathbb{E}_s$ means take expectation with respect to s, that is, use the probability of appearance of s to calculate the expectation

If we use a rather formal notation $\tau =< \pi >$ to indicate that a distribution of posterior μ induced by π [17]

The equation transforms to

$$\mathbb{E}_{\mu \sim \tau} \mu = \mu_0$$

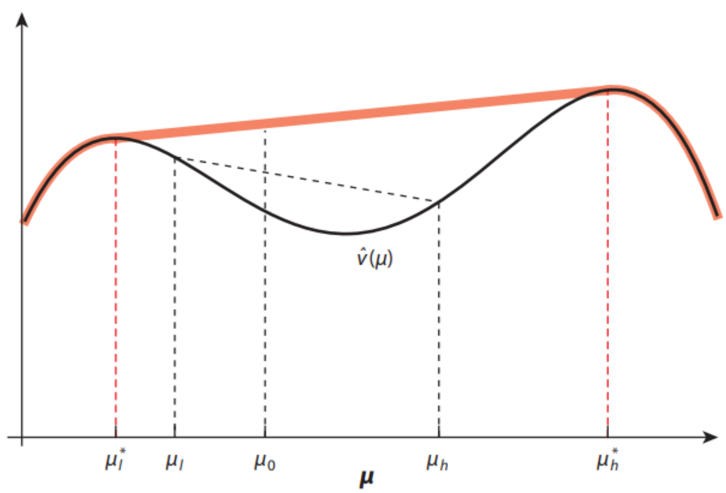
And for the expectation is commutative, then we can reformulate the sender's expected utility (1) as:

$$max_{\pi \in \Pi} \mathbb{E}_{\omega \sim \mu_0} \mathbb{E}_{s \sim \pi(\omega)} v\left(a^*\left(\mu_\pi(\cdot \mid s)\right), \omega\right)$$

$$= \max_{\tau} \mathbb{E}_{\mu \sim \tau} \hat{v}(\mu) \tag{2}$$

s.t. $\mathbb{E}_{\mu \sim \tau} \mu = \mu_0$, where as $\hat{v}(\mu) = \mathbb{E}_{\omega \sim \mu} v(a^*(\mu), \omega)$

A nice geometric interpretation



- The concavification of $\hat{v}$ evaluated at μ_0 equals $max\{z \mid (\mu_0, z) \in co(\hat{v})\}$[18]

[17] Algebraically, a signal induces a distribution of posterior τ if $\tau(\mu) = \sum_{s:\mu_{\pi(\cdot|s)}=\mu} \sum_{\omega' \in \Omega} \pi(s|\omega')\mu_0(\omega')$

[18] $co(\hat{v})$ denotes the convex hull of the graph of $\hat{v}$

- Send's utility under the optimal signal is precisely the concavification of $\hat{v}$ evaluated at μ_o

Remark. If there is more than three states, we cannot plot $\hat{v}$ anymore, but the concavification approach can still be used to derive some qualitative features of the optimal signal(see section 4).

The intellectual history of concavification approach

Characterizing the equilibria of the following repeated games of incomplete information is analogous to the Equation 1 (Aumann and Maschler, 1966)

- Two players: Informed player(I), Uninformed player(U)
- Two *zero-sum* games: G_A and G_B
- Identical action spaces
- With μ_0, the players will repeatedly play G_A ad infinitum, and G_B otherwise.
- Before each period, I knows which game they will be playing
- After each period, U observes the action of I, but she does not observe her payoff, nor which game they had played
- Each player seek to maximize their undiscounted average payoff (??[19])

This concern comes from the namely *"that the negotiating strategy used by the Americans in a series of arms control conferences might implicitly send signals to the Russians about the nature of the US arsenal"*

Brocas and Carrillo (2007) consider a sender select how many i.i.d draws of a fixed signal to generate rather than choose signal from Π

[19]Why is it undiscounted?

3.2 When the concavification approach fails

If the state space is *uncountable*[20], we cannot use the concavification approach even theoretically.

But in a special case, we can decrease the dimension of μ to one dimension.

Suppose $\Omega = [0,1]$, and $a^*(\mu) = f(\mathbb{E}_{\omega\sim\mu}\omega)$, then there must exist a function s.t. $\tilde{v}(\mathbb{E}_{\omega\sim\mu}\omega) = \hat{v}(\mu)$

3.3 Computational methods

Dughmi (2017) provides an excellent survey.

4 Extensions

Three main extensions:[21]

- Multiple receivers
- Multiple senders
- Dynamic environments

And there are some other extensions[22], for example:

Private information

- Allow the receiver to have their own *private information*
  - Preferences (Kolotilin, 2018; Rayo and Segal, 2010)
  - State of the world (Guo and Shmaya, 2019)

[20]What if it is countable infinite? How we address this problem theoretically? Though it is likely to be manageable intuitively.

[21]These extensions are sometimes combined, for instance, Koessler et al. (2021) examine a basic games with multiple senders and multiple receivers. Ely (2017) considers a dynamic model with multiple receivers and derives insights about information policies that reduce the likelihood of bank runs.

[22]There are also some paper cannot be included in the main branch, for example Au and Li (2018) add reciprocity to the receiver's preference(rather more *experimental*), and Tsakas et al. (2021) provides some commitment power to receiver(??)

- There are some interesting results under the circumstances that we allow the sender is able to elicit information
  - Kolotilin et al. (2017) consider the case where the sender observes the reported type of the receiver's preference, the results shows *the sender cannot benefit from the ability to elicit information*.
  - Li and Shi (2017) consider the case where sender conditions his signal on receiver's report but does not observe the report prior to setting the price, and they show that *discriminatory disclosure* dominates *full disclosure*.

A more complicated but rather interesting model is examined by Matyskova (2018) Settings:

- Receiver has no private information at outset
- *Can gather additional costly information after observing the realization of sender's signal*

Results:

- Receiver never gather information on the equilibrium path
- The threat of information gathering, weakly harms sender and can be beneficial or harmful for receiver (??[23])

Heterogeneous prior

Alonso and Câmara (2016) The sender and receiver have *heterogeneous* priors[24].

References

Alonso, R. and Câmara, O. (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory*, 165:672–706.

[23]Why can it be harmful?

[24]This paper has cited the paper Aumann and Maschler (1966) to discuss the heterogeneous prior, it may be very interesting

Au, P. H. and Li, K. K. (2018). Bayesian persuasion and reciprocity: theory and experiment. *Available at SSRN 3191203*.

Aumann, R. J. and Maschler, M. (1966). Game theoretic aspects of gradual disarmament. *Report of the US Arms Control and Disarmament Agency*, 80:1–55.

Bergemann, D. and Bonatti, A. (2019). Markets for information: An introduction. *Annual Review of Economics*, 11:85–107.

Bergemann, D., Bonatti, A., and Smolin, A. (2018). The design and price of information. *American economic review*, 108(1):1–48.

Bergemann, D., Brooks, B., and Morris, S. (2015). The limits of price discrimination. *American Economic Review*, 105(3):921–57.

Bergemann, D. and Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308.

Best, J. and Quigley, D. (2020). Persuasion for the long run. *Available at SSRN 2908115*.

Brocas, I. and Carrillo, J. D. (2007). Influence through ignorance. *The RAND Journal of Economics*, 38(4):931–947.

Chen, Y.-C. and Hu, G. (2020). Learning by matching. *Theoretical Economics*, 15(1):29–56.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451.

Dughmi, S. (2017). Algorithmic information structure design: a survey. *ACM SIGecom Exchanges*, 15(2):2–24.

Ely, J. C. (2017). Beeps. *American Economic Review*, 107(1):31–53.

Frankel, A. (2014). Aligned delegation. *American Economic Review*, 104(1):66–83.

Gentzkow, M. and Kamenica, E. (2017). Disclosure of endogenous information. *Economic Theory Bulletin*, 5(1):47–56.

Goldstein, I. and Leitner, Y. (2018). Stress tests and information disclosure. *Journal of Economic Theory*, 177:34–69.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics*, 24(3):461–483.

Guo, Y. and Shmaya, E. (2019). The interval structure of optimal disclosure. *Econometrica*, 87(2):653–675.

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annu. Rev. Econ.*, 4(1):427–452.

Kamenica, E. (2019). Bayesian persuasion and information design. *Annual Review of Economics*, 11:249–272.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kastl, J., Pagnozzi, M., and Piccolo, S. (2018). Selling information to competitive firms. *The RAND Journal of Economics*, 49(1):254–282.

Koessler, F., Laclau, M., and Tomala, T. (2021). Interactive information design. *Mathematics of Operations Research*.

Kolotilin, A. (2018). Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635.

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., and Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964.

Lazear, E. P. (2006). Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*, 121(3):1029–1061.

Lewis, T. R. and Sappington, D. E. (1994). Supplying information to facilitate price discrimination. *International Economic Review*, pages 309–327.

Li, C. (2017). A model of bayesian persuasion with transfers. *Economics Letters*, 161:93–95.

Li, H. and Shi, X. (2017). Discriminatory information disclosure. *American Economic Review*, 107(11):3363–85.

Matyskova, L. (2018). Bayesian persuasion with costly information acquisition. *CERGE-EI Working Paper Series*, (614).

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391.

Min, D. (2020). Bayesian persuasion under partial commitment. *Work. Pap., NYU*.

Ostrovsky, M. and Schwarz, M. (2010). Information disclosure and unraveling in matching markets. *American Economic Journal: Microeconomics*, 2(2):34–63.

Rayo, L. and Segal, I. (2010). Optimal information disclosure. *Journal of political Economy*, 118(5):949–987.

Romanyuk, G. and Smolin, A. (2019). Cream skimming and information design in matching markets. *American Economic Journal: Microeconomics*, 11(2):250–76.

Spence, M. (1978). Job market signaling. In *Uncertainty in economics*, pages 281–306. Elsevier.

Taneva, I. (2019). Information design. *American Economic Journal: Microeconomics*, 11(4):151–85.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Tsakas, E., Tsakas, N., and Xefteris, D. (2021). Resisting persuasion. *Economic Theory*, pages 1–20.