Persuasion
Bayesian Persuasion
Robust Bayesian Persuasion

Jenny Zhang

CPSC 532L Presentation

Persuasion - Example

Prosecutor

2/3 innocent
1/3 guilty

Judge

- Judge
  - Get a payoff of 1 if they take the correct action, 0 otherwise

- Prosecutor
  - Wants to persuade the judge to convict
  - Gets a payoff of 1 if the judge convicts, 0 otherwise
  - Conducts an investigation, and report its outcome to the judge

Persuasion - Example

What is the prosecutor's optimal "recommendation strategy"?

- Attempt 1: always say "guilty" (equivalently, no information)
    - Judge never convicts
    - Prosecutor's expected utility 0

- Attempt 2: full information (i.e., being honest)
    - Judge convicts 1/3 of the time
    - Prosecutor's expected utility 1/3

- Attempt 3: noisy information ✨🪄
    - Prosecutor's expected utility 2/3 - ε
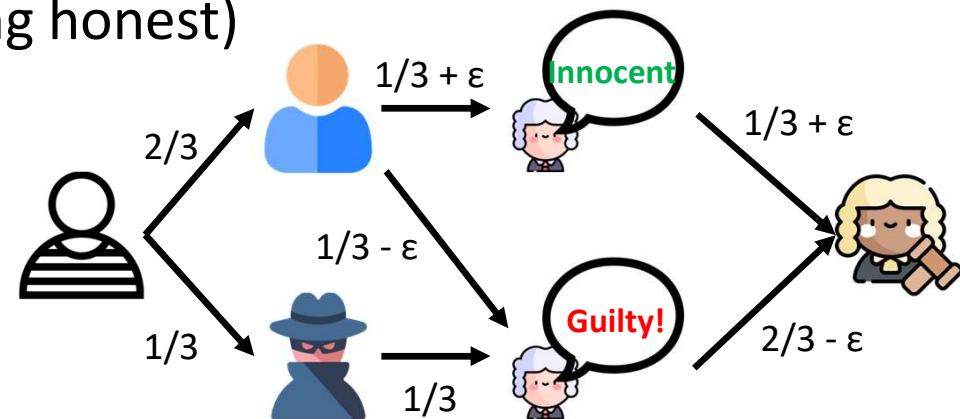
Persuasion
The act of exploiting an informational advantage in order to influence the decisions of others.

- Intrinsic in most human activities: advertising, politics, marketing, …
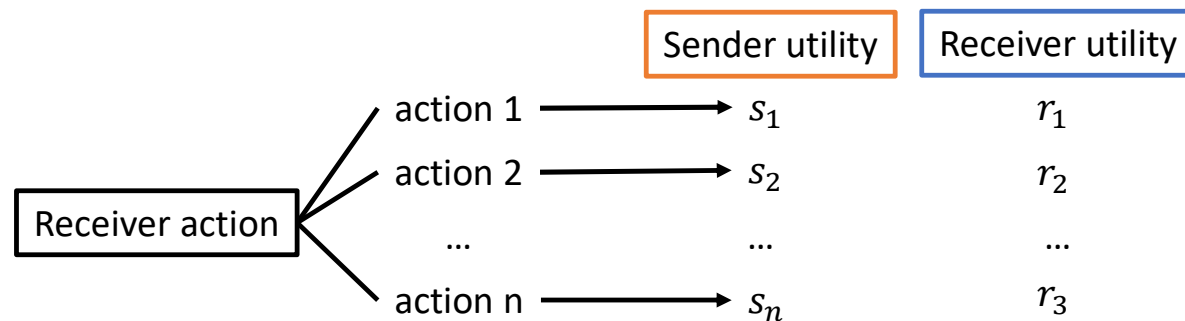- A large body of research

One Quarter of GDP Is Persuasion

By DONALD MCCLOSKEY AND ARJO KLAMER*

- The American Economic Review, 1995

Bayesian Persuasion

- *Kamenica and Gentzkow, American Economic Review, 2011*
- Two players: a persuader (sender), a decision maker (receiver)
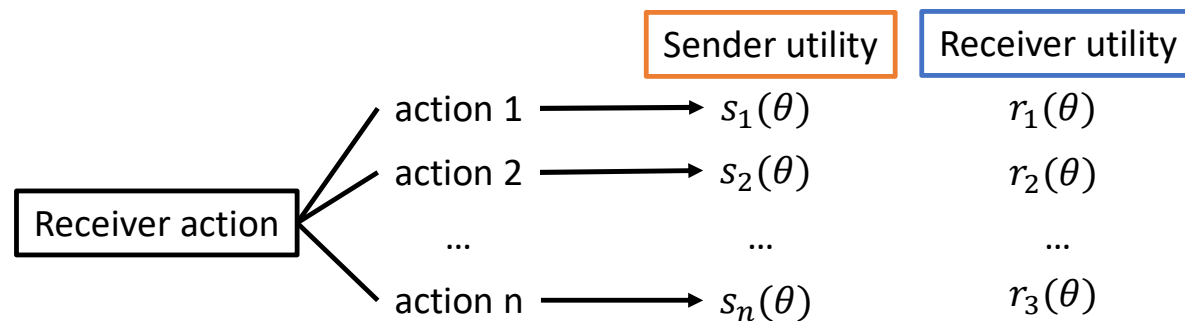 - In previous example: prosecutor = sender, judge = receiver

Bayesian Persuasion

- *Kamenica and Gentzkow, American Economic Review, 2011*
- Two players: a persuader (sender), a decision maker (receiver)
  - In previous example: prosecutor = sender, judge = receiver

| Sender utility | Receiver utility |

Receiver action → action 1 → $s_1(\theta)$ ... $r_1(\theta)$
action 2 → $s_2(\theta)$ ... $r_2(\theta)$
...
action n → $s_n(\theta)$ ... $r_3(\theta)$

- Receiver only knows the prior distribution of θ
- Sender can observe the realization of θ
- Sender can *commit* to a signaling scheme: randomized map $\theta \rightarrow \sum$
- Receiver: receive a signal → Bayes update → best respond
- $|\sum| = n$: signal i recommends action i as the receiver's best response

Variant of the Judge Example

- There are 2 suspects, but only one crime
- Judge wants to convict one of them
- State θ is a uniformly-random type from {N, G, F} for each suspect

	N	G	F
Judge utility	0	1 + ε	2
Prosecutor utility	0	1	0

Signaling Scheme 1

- No information
- Suspects are identical to the recruiter
- Judge randomly chooses a suspect to convict
- Expected prosecutor utility 1/3

Variant of the Judge Example

- There are 2 suspects, but only one crime
- Judge wants to convict one of them
- State θ is a uniformly-random type from {N, G, F} for each suspect

	N	G	F
Judge utility	0	1 + ε	2
Prosecutor utility	0	1	0

Signaling Scheme 2

- Full information
- Good cases for prosecutor: (G, G), (G, N), (N, G)
- Expected prosecutor utility 1/9 * 3 = 1/3

Variant of the Judge Example

- There are 2 suspects, but only one crime
- Judge wants to convict one of them
- State θ is a uniformly-random type from {N, G, F} for each suspect

	N	G	F
Judge utility	0	$1 + \varepsilon$	2
Prosecutor utility	0	1	0

Signaling Scheme 3 (Optimal)

- Properly correlate suspects' types:
  - Whenever there is exactly one type-G suspect, recommend him
  - Otherwise, recommend a suspect uniformly at random
- Prosecutor gets utility 1 whenever type-G shows up (with probability 5/9)

Bayesian Persuasion – In General

- 2 agents - Sender, Receiver
- Receiver has a continuous utility function $u(a, w)$
  - Depends on its action $a \in A$ and the state of the world $w \in \Omega$
- Sender has a continuous utility function $v(a, w)$
- Sender and Receiver share a prior $\mu_0 \in int(\Delta(\Omega))$
  - $int(X)$ denotes the interior of set X
  - $\Delta(X)$ denotes the set of all probability distributions on X
- Sender chooses a signal π
  - Consists of a finite realization space S and a family of distributions $\{\pi(\cdot | \omega)\}_{\omega \in \Omega}$ over S
- Receiver observes sender's signal and a signal realization $s \in S$
- Receiver forms the posterior μ_s using Bayes's rule
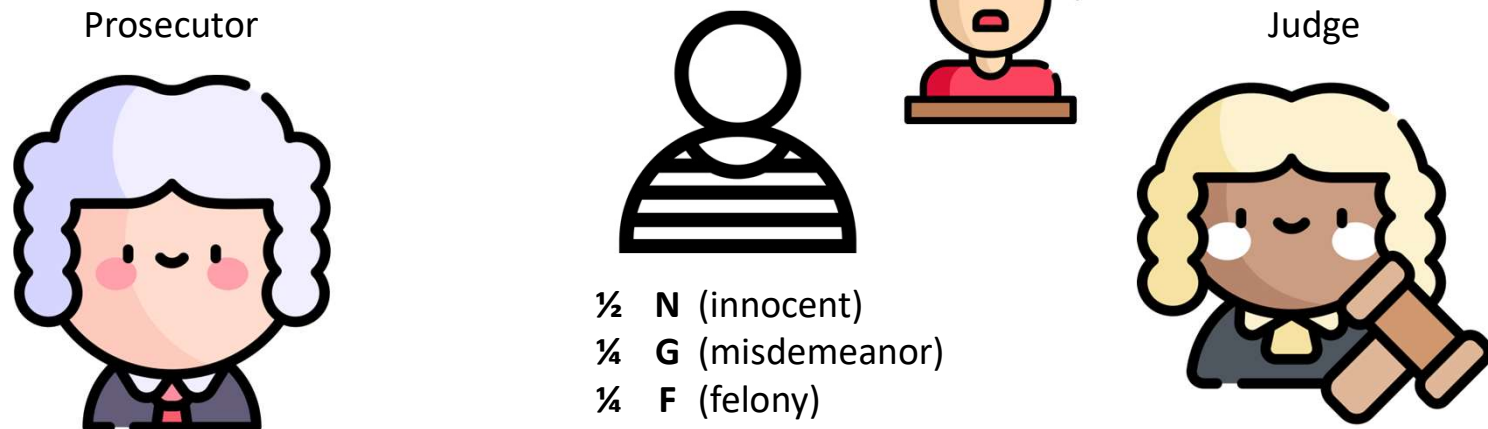- Receiver takes an action from the set $a^*(\mu_s) = argmax_{a \in A} E_{\mu_s}[u(a, \omega)]$

Robust Bayesian Persuasion

Robust Bayesian Persuasion

- In the previous example, the prosecutor believes ("conjectures") that they are the sole provider of information
- But what if the judge calls a witness?
- Let's consider this game:

Prosecutor

Judge

½ N (innocent)
¼ G (misdemeanor)
¼ F (felony)

Get payoff 1 if suspect is convicted as misdemeanor
Get payoff 2 if suspect is convicted as felony

Judge convicts as guilty if the belief of the suspect being guilty $\geq \frac{2}{3}$

Robust Bayesian Persuasion

Robustness

- The prosecutor might not know the likelihood of the witness appearing in court, the amount of information that they have, or their motives

- Assumption: The sender is concerned about the worst-case scenario

- In the worst case, the sender cannot do better than the full-disclosure payoff
  - They clearly cannot do strictly better
  - They can achieve that payoff by disclosing the state herself

- This was already observed by Hu and Weng (2019)

- Key idea:
  - The sender should **not** fully disclose the state in this case!

Robust Bayesian Persuasion

Superior policy Robust Solution

- The prosecutor reveals the state N, but pools together the states G and F

- In the worst case:
  - When the state is N, the witness has no additional information
  - When the state is G or F, the witness reveals the state

- In the worst case, the payoff is **exactly the same** as under full disclosure by the sender

- If the conjecture happens to be right:
  - When the state is G or F, the judge's belief is (0, ½, ½)
  - The expected payoff is $1 > \frac{3}{4}$

- This policy is just as good as full disclosure in the worst-case but strictly better when the sender's conjecture turns out to be right

Robust Bayesian Persuasion

Robust Solution
1. The sender secures the best possible payoff guarantee. Dismiss any policy that is not optimal in the "worst-case scenario."
2. When there are multiple policies that are worst-case optimal, the sender acts as in the standard Bayesian persuasion model. That is, it selects the policy that, among those that are worst-case optimal, maximizes her expected payoff under the conjecture.

Conclusion

- Persuasion is a powerful tool
- Senders can modify their signals to convince the receiver to act in a way that is more favorable to them
  - How to design these signals will be covered next week in mechanism design!
- There are robust solutions that can do just as well in the worst-case scenario and favor the sender when the conjecture is true (i.e., they are the sole information provider)

References

- [Bayesian Persuasion](), Kamenica and Gentzkow, 2011
- [Robust persuasion of a privately informed receiver](), Hu & Weng, 2019
- [Preparing for the Worst but Hoping for the Best: Robust (Bayesian) Persuasion](), Dworczak et al., 2022

Additional Resources

- [Bayesian persuasion with heterogeneous priors](), Alonso and Camara, 2016
- [Bayesian persuasion with multiple senders and rich signal spaces](), Kamenica and Gentzkow, 2017
- [Bayesian Persuasion with Costly Information Acquisition](), Matyskova, 2018
- [Online Bayesian Persuasion](), Castiglioni, Celli, Marchesi, Gatti, 2020
- [Bayesian Persuasion with Mediators](), Arieli, Babichenko, Sandomirskiy, 2022

Thank you! Questions?

All images are AI generated, Icons are from FlatIcon