A Study in the Pragmatics of Persuasion: a Game Theoretical Approach

Jacob Glazer         Ariel Rubinstein

Presented by: Alireza Bakhtiari

Laboratoire DAMAS, Département d’Informatique et de Génie Logiciel
Université Laval

September 18th, 2009
Problem

- In a persuasion situation, what are the rules that determine which of the facts, presented by the speaker, the listener will find persuasive

Approach

- Defining the persuasion rule based on the speaker’s beliefs about the persuasiveness of the sentences
- Defining the listeners error probability
- Finding an optimal persuasion rule that minimizes the listener’s error probability
- Demonstrating that optimal persuasion rules are ex-post optimal
A persuasion situation involves an agent (the speaker) who attempts to persuade another agent (the listener) to take a certain action.

The speaker is restricted as to how many pieces of evidence he can present:

- time constraints, limitations on the listener’s capability to process information

The listener can either accept or reject the speaker’s suggestion (no partial acceptance).
The Model

A persuasion problem is modeled as a four-tuple \( \langle X, A, p, \sigma \rangle \)

- \( X \) : The set of all states
- \( A \subset X \) : All the states in which the listener would wish to accept the speaker’s suggestion
  - \( R = X \setminus A \) : All the states in which the listener would wish to reject the speaker’s request
- \( p \) : The listener’s initial beliefs about the state (\( p_x \) is the probability of state \( x \))
- \( \sigma(x) \) : The set of statements that the speaker can make
  - \( S = \bigcup_{x \in X} \sigma(x) \)
Persuasion Rules

- A persuasion rule determines how the listener responds to each of the speaker’s possible statements.
- A function $f : S \rightarrow [0, 1]$ that specifies the speaker’s beliefs about how the listener will interpret each of his possible statements.
- $f(s) = q$ gives the probability that the listener is persuaded following a statement $s$.
- $f$ is deterministic if $f(s) \in \{0, 1\}$ for all $s \in S$. 

Listeners error probability

- In each state, the speaker solves the following problem:

\[ \alpha(f, x) = \max_{s \in \sigma(x)} f(s) \]

- The listener’s error probability \( \mu_x(f) \) in state \( x \) is defined as follows:

\[
\mu_x(f) = \begin{cases} 
1 - \alpha(f, x) & x \in A \\
\alpha(f, x) & x \in R
\end{cases}
\]

- Therefore, the *error probability* induced by \( f \) is:

\[ m(f) = \sum_{x \in X} p_x \mu_x(f) \]

- Given a problem \( \langle X, A, p, \sigma \rangle \), an optimal persuasion rule is one that minimizes \( m(f) \).
**Example 1** ”The majority of the facts supports my position”

- 5 independent random variables, each taking the values 1 or 0 with equal probability
- The speaker can present $m = 2$ variables to the listener
- The listener is persuaded iff there are at least 3 variables that take the value 1
- Formally:

\[
X = \{(x_1, \cdots, x_5) | x_k \in \{0, 1\} \text{ for all } k\}
\]

\[
A = \{x | n(x) \geq 3\} \text{ where } n(x) = \sum_k x_k
\]

\[
p_x = \frac{1}{32}
\]

\[
\sigma(x) = \{\kappa | \kappa \subseteq \{k | x_k = 1\} \text{ and } |\kappa| \leq m\}
\]
Example

- If the listener is persuaded by the presentation of any two variables, then the error probability is $\frac{10}{32}$.
- If he is persuaded by the presentation of two neighboring variables ($\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}$) with the value 1, the error probability reduces to $\frac{5}{32}$.
- An optimal persuasion rule can achieve an error probability of $\frac{4}{32}$. 
A pair \((x, T)\) with \(x \in A\) and \(T \in R\) is called an "L" if for any \(s \in \sigma(x)\), there is \(t \in T\) such that \(s \in \sigma(t)\).

**The L-Principle:** Let \((x, T)\) be an L in the persuasion problem \(\langle X, A, p, \sigma \rangle\) and let \(f\) be a persuasion rule. Then \(\Sigma_{t \in \{x\} \cup T} \mu_t(f) \geq 1\).

If a particular persuasion rule induces an error probability equal to the lower bound derived from the L-principle, then this persuasion rule is optimal.
According to the following proposition, uncertainty is not useful to the listener in the present context.

**Proposition**: For every finite persuasion problem \( \langle X, A, p, \sigma \rangle \), there exists an optimal persuasion rule \( f \) that is deterministic.
Example 1 revisited: Consider a deterministic optimal persuasion rule; i.e. $\mu_x$ is either 0 or 1 for any state. By the L-principle we have:

$$\mu(1,1,0,0) + \mu(1,1,0,0) + \mu(1,0,1,0) + \mu(0,1,1,0) \geq 1$$

similar inequalities hold for the other 9 states that exactly 3 variables support the speaker

Therefore:

$$\sum_{n(x)=3} \mu_x + 3 \sum_{n(x)=2} \mu_x \geq 10$$

Since $\mu_x$ is either 0 or 1, we have $\sum_{n(x)=3} \mu_x + \sum_{n(x)=2} \mu_x \geq 4$ and thus $\sum_x p_x \mu_x \geq \frac{4}{32}$
Example

- The optimal persuasion rule would be to partition the set of variables into two sets \{1, 2, 3\} and \{4, 5\}

- The listener is persuaded only if the speaker can show him that two random variables from the same cell of the partition support him
Example

Example 2 "Persuading someone that the median is above the expected value"

- Three independent random variables uniformly distributed over $[0, 1]$
- The expected value is 0.5
- The speaker can reveal the value of only one variable
- Formally:

$$A = \{(x_1, x_2, x_3) \mid \text{two of the values are above } 0.5\}$$

- The persuasion rule according to which the listener is persuaded iff the revealed value is above 0.5 yields a probability of error of $\frac{1}{4}$
- The L-principle shows that this persuasion rule is optimal
The space $X = [0, 1] \times [0, 1] \times [0, 1]$ is isomorphic to the probabilistic space $Y \times Z$ where $Y = [0, \frac{1}{2}] \times [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ and $Z = \{-1, 1\} \times \{-1, 1\} \times \{-1, 1\}$ by the following identification:

$$x = \left( \frac{1}{2} + y_i z_i \right)_{i=1,2,3}$$

Therefore $(y, (1, 1, -1)) \in A$ is part of an $L$ with $(y, (-1, 1, -1)) \in R$ and $(y, (1, -1, -1)) \in R$. So we obtain:

$$\mu(y,(1,1,-1)) + \mu(y,(-1,1,-1)) + \mu(y,(1,-1,-1)) \geq 1$$

$$\mu(y,(1,-1,1)) + \mu(y,(-1,-1,1)) + \mu(y,(1,-1,-1)) \geq 1$$

$$\mu(y,(-1,1,1)) + \mu(y,(-1,1,-1)) + \mu(y,(-1,-1,1)) \geq 1$$
Example

- By adding the inequalities we obtain

\[ \mu(y,(1,1,-1)) + \mu(y,(1,-1,1)) + \mu(y,(-1,1,1)) \]
\[ + 2\mu(y,(-1,1,-1)) + 2\mu(y,(1,-1,-1)) + 2\mu(y,(-1,-1,1)) \geq 3 \]

- The persuasion rule is deterministic, so at least two of the variables \( \mu(y,z) \) must take the value 1

\[ \sum_z p(y,z) \mu(y,z) \geq \frac{2}{8} = \frac{1}{4} \]
Finding the optimal persuasion rule

- **Proposition:** Let $\langle X, A, p, \sigma \rangle$ be a finite persuasion problem. Let $(\mu_x^*)_{x \in X}$ be a solution to the optimization problem

\[
\min_{\{\mu_x\}_{x \in X}} \sum_{x \in X} p_x \mu_x \text{ s.t. } \mu_x \in \{0, 1\} \text{ for all } x \in X \text{ and } \sum_{t \in \{x\} \cup T} \mu_t \geq 1 \text{ for any minimal } L, (x, T)
\]

Then there is an optimal persuasion rule that induces the probabilities of errors $(\mu_x^*)_{x \in X}$
Ex-post optimality

- So far the listener was committed to a persuasion rule
- Would the listener’s optimal persuasion rule differ if he were able to reconsider his commitment after the speaker has made his statement?

**Example**

- A listener wishes to choose a guest with strong views about the issues of the day for a TV news program
- Four types: ”hawk” (H), ”dove” (D), ”pretender” (M), ”ignorant” (I)
- \( p(H) = p(D) = 0.2 \) and \( p(M) = p(I) = 0.3 \)
- \( \sigma(H) = \{h, 0\}, \sigma(D) = \{d, 0\}, \sigma(M) = \{h, d, 0\}, \sigma(I) = \{0\} \)
Example

- *naive approach*: Given the statement $s$, the listener excludes the types that cannot make the statement $s$ and makes the optimal decision given the probabilities.
  - probability of error: 0.4

- If the listener can commit to how he will respond to the speaker, then the best persuasion rule is to invite him iff he makes the statement $d$ or $h$
  - probability of error: 0.3
Example

- Suppose the listener has no commitment

- If he believes that M’s strategy is to utter $d$, then he should assign a higher probability to the possibility that he is facing an M rather than a D. Therefore, he should reject the speaker

- If he believes that M randomizes between $d$ and $h$, then upon hearing $d(h)$, the listener attributes the probability $\frac{4}{7}$ to the possibility that he is facing D (H). Therefore he accepts the speaker
Consider the corresponding extensive persuasion game $\Gamma(\langle X, A, p, \sigma \rangle)$

- First nature chooses the state according to $p$
- Then the speaker is informed of the state $x$ and utters a sentence from $\sigma(x)$
- Finally, the listener chooses between $a$ and $r$

Payoffs:

$$\text{Speaker’s payoff} = \begin{cases} 1 & \text{if listener takes action } a \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Listener’s payoff} = \begin{cases} 1 & \text{if } x \in A \text{ and action } a \text{ is taken} \\ 0 & \text{otherwise} \end{cases} \text{ or } x \in R \text{ and action } r \text{ is taken}$$
A persuasion rule $f$ is *credible* if there exists a sequential equilibrium of \( \Gamma(\langle X, A, p, \sigma \rangle) \) such that the listener’s strategy is $f$

**Proposition:** If the persuasion problem is finite, then any optimal persuasion rule is credible
Concluding remarks

- The authors present a model for persuasion situations
- The model is basically a one stage game, so after the speaker has uttered a sentence, the listener must either accept or reject his position
- One of the main results of the article is that any optimal persuasion rule is also ex-post optimal
- This finding is demonstrated with an example, but ”the generalizability of this result is still an open question”
Questions?