Game Theory

Cheap Talk

Tetsuya Hoshino

May 15, 2022

In a signaling game, a player communicates her private information by taking costly actions. However, such actions are not always available. The **cheap-talk** model of Crawford & Sobel (1982) studies how much information can be transmitted when the sender's all actions are costless.

1 Cheap Talk

Definition 1. A **cheap-talk game** is a two-player extensive-form game that proceeds as follows:

1. Nature draws player 1's type t from a finite set T according to a uniform prior π.
   - Assume that $T = \{1, 2, \ldots, K\}$ for some $K \geq 2$.
   - Player 1's learns her type t but player 2's does not.
2. Player 1 chooses a message m from a finite set $M = T$.[1]
   - Player 1's message m is often called a (cheap-talk) message.
   - Player 1's message m affects neither player's payoff.
3. Player 2 observes message m and then takes an action a from the set $A = \mathbb{R}$.
4. Payoffs are realized, and then the game ends.
   - $u_1(a, t) = -(t + b - a)^2$ for a given "bias" $b \geq 0$.
   - $u_2(a, t) = -(t - a)^2$.

Remark 1. We refer to player 1's "action space" as her message space M, just because we emphasize that player 1's behavior affects neither player's payoff. It makes no technical difference whether we call it an action space or a message space. □

Remark 2. The payoff functions of Definition 1 are called **quadratic payoff functions** and widely used for tractability. Note that both players' interests are well-aligned in the sense that they prefer a higher action a as a type t gets higher, but player 1 is always biased (for an even higher action a) by a fixed amount b. □

Remark 3. The cheap-talk game of Definition 1 is the same as Crawford & Sobel's (1982) game except that they take player 1's type space $T = \mathbb{R}$. Our finiteness assumption simplifies the analysis. □

[1] All results remain true if $|M| \geq |T| = K$.

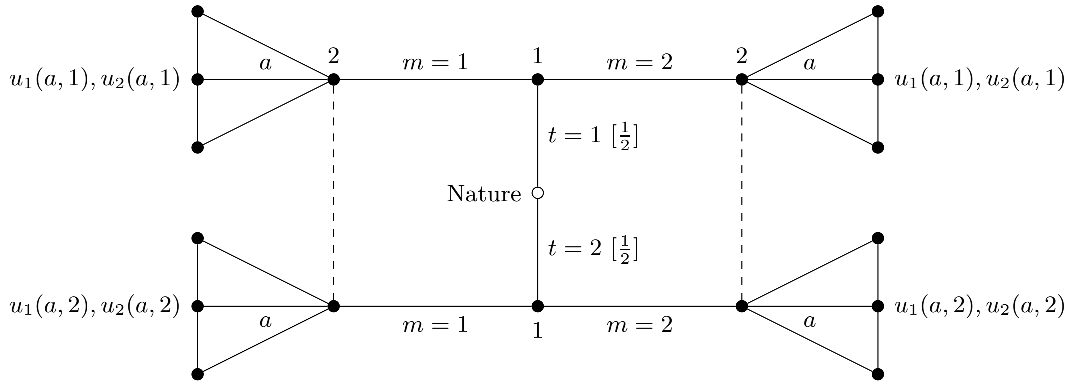

Figure 1: Cheap-talk model ($K = 2$)

In this note, we focus on (behavioral) pure strategies. Let $s_1 : T \to M$ be player 1's strategy, where $s_1(t)$ denotes her message when she has type t. Let $s_2 : M \to A$ be player 2's strategy, where $s_2(m)$ denotes his action when he observes message m.

Our equilibrium concept is a (pure-strategy) perfect Bayesian equilibrium.

1.1 Babbling Equilibrium

If player 1 sends the same message m regardless of her type t then her message conveys no information about her type t and thus player 2 will ignore it. This kind of (pooling) strategy is usually called a **babbling strategy** (in the context of cheap talk), and the equilibrium involving a babbling strategy is often called a **babbling equilibrium**.

Proposition 1. *For each $m \in M$, there exists a perfect Bayesian equilibrium such that:*

- *Player 1's strategy $s_1(t) = m$ for each $t \in T$.*
- *Player 2's strategy $s_2(m') = \frac{K+1}{2}$ for each $m' \in M$.*

Proof. Fix a message $m \in M$ and suppose that player 1's strategy $s_1(t) = m$ for each $t \in T$. It follows that player 2's belief is $\mu(t \mid m) = \frac{1}{K}$ for each $t \in T$. Player 2, with a belief $\mu(t \mid m)$, chooses action a^* such that:

$$a^* \in \operatorname*{argmin}_{a \in A} \quad \sum_{t \in T} \mu(t \mid m)(t - a)^2.$$

His sequentially rational action $a^* = \frac{K+1}{2}$ is independent of message m. Player 1's strategy $s_1(t) = m$ is sequentially rational (since her message does not affect player 2's behavior). Hence, this assessment (s, μ) is a perfect Bayesian equilibrium. ∎

Roughly speaking, this babbling perfect Bayesian equilibrium illustrates a situation in which since player 1's cheap-talk message conveys no information and thus player 2 just ignores it and takes his best action based on his prior belief.

1.2 Informative Equilibrium

Fully Informative Equilibrium If players' interests are well-aligned in the sense that a bias b is small, then player 1 may want to tell her true type t. Such a message strategy is called a **truth-telling strategy**. The equilibrium involving a truth-telling strategy is often called a **fully informative equilibrium**. This intuition is formalized as follows:

Proposition 2. *If $b \leq \frac{1}{2}$ then there exists a perfect Bayesian equilibrium such that:*

- *Player 1's strategy $s_1(t) = t$ for each $t \in T$.*
- *Player 2's strategy $s_2(m) = m$ for each $m \in M$.*

Proof. Given the said strategy profile s, a weakly consistent belief must be such that:

$$\mu(t \mid m) = \begin{cases} 1 & \text{if } t = m \\ 0 & \text{if } t \neq m. \end{cases}$$

It then suffices to show that this assessment (s, μ) is sequentially rational. First, we examine player 2's strategy s_2. When observing message m, player 2 realizes the type $t = m$, learning his bliss point $m = t$. Therefore, he chooses action $a^* = m$. That is, player 2's strategy $s_2(m) = m$ is sequentially rational.

Next, we turn to player 1's strategy s_1. If she sends message m then player 2 will choose action $a = m$. Thus, from a truth-telling strategy $m = t$, player 1 gains payoff $-b^2$. In contrast, if she deviates to sending message $m' \neq m$ then since player 2 will choose action $a = m'$, player 1 gains payoff $-(t - m' + b)^2$. To see that player 1 has no profitable deviation, it suffices to show that $-b^2 \geq -(t - m' + b)^2$. This inequality is equivalent to

$$(t - m')\left(b + \frac{t - m'}{2}\right) \geq 0.$$

To see this, we note that $|t - m'| \geq 1$ since player 2 will choose action $s_2(m') = m'$ when observing message $m' \neq m = t$. Therefore, she has no profitable deviation. ■

Partially Informative Equilibrium If players' interests are not well-aligned in the sense that a bias b is not small, then it seems impossible that player 1 can convey her information fully. This intuition is formalized as follows:

Proposition 3. *If $b > \frac{1}{2}$ then there exists no perfect Bayesian equilibrium such that:*

- *For each $t, t' \in T$, if $t \neq t'$ then $s_1(t) \neq s_1(t')$.*

Proof. Suppose, for a contradiction, that there is a perfect Bayesian equilibrium such that at any types $t \neq t'$, player 1 sends distinct messages $s_1(t) \neq s_1(t')$. Then, since the belief is weakly consistent, it must be that $\mu(t \mid s_1(t)) = 1$ for each $t \in T$.

We take two types $1, 2$, for example. Then, when observing message $m_1 = s_1(1)$ (resp. $m_2 = s_1(2)$), player 2 is sure about player 1's type $t = 1$ (resp. $t = 2$) and he takes a subsequent action $a = 1$ (resp. $a = 2$). However, player 1 of type $t = 1$ wants to deviate because if she sends message m_1 (as in the equilibrium), she will gain payoff $u_1(1, 1) = -b^2$, while if she deviates to sending message m_2, she will gain payoff $u_1(1, 2) = -(1-b)^2$, which is greater than $-b^2$. This means that the said assessment fails to be sequentially rational. ■

1.3 Interval Equilibria

Monotonicity In every perfect Bayesian equilibrium, as player 1's type gets higher, player 2 chooses a higher action. This is formally done below.

Lemma 1. *Every perfect Bayesian equilibrium (s, μ) is monotone in the sense that:*

- *For each $t, t' \in T$, if $t < t'$ then $s_2(s_1(t)) \leq s_2(s_1(t'))$.*

Proof. Suppose, for a contradiction, that $s_2(s_1(t)) > s_2(s_1(t'))$ at some types $t < t'$. It suffices to show either that player 1 of type t strictly prefers message $s_1(t')$ to message $s_1(t)$ or that player 1 of type t' strictly prefers message $s_1(t)$ to message $s_1(t')$. Let $a = s_2(s_1(t))$ denote player 2's action when player 1 chooses message $s_1(t)$, and let $a' = s_2(s_1(t'))$ denote player 2's action when player 1 chooses message $s_1(t')$.

There are two cases to consider:

1. Suppose $t' \geq a$. Then, $t' \geq a > a'$. Hence, player 1 of type t', with the bliss point $t' + b$, prefers action a to action a', because a is closer to $t' + b$ than a' is. Hence, player 1 of type t' prefers message $s_1(t)$.

2. Suppose $t' < a$. By sequential rationality, player 1 of type t prefers action a to action a', while player 1 of type t' prefers action a' to action a. That is,

$$(a - (t + b))^2 \leq (a' - (t + b))^2,$$
$$(a' - (t' + b))^2 \leq (a - (t' + b))^2.$$

By simple algebra, the first and second inequality respectively imply that

$$(a' - a)(a + a' - 2(t + b)) \geq 0,$$
$$(a - a')(a + a' - 2(t' + b)) \geq 0.$$

Since $a > a'$, it follows that:

$$a + a' - 2(t + b) \leq 0,$$
$$a + a' - 2(t' + b) \geq 0.$$

Hence, $2(t' + b) \leq a + a' \leq 2(t + b)$, but this inequality never holds since $t < t'$. This is a contradiction. ■

Partition with Intervals Lemma 1 implies that player 1's perfect Bayesian equilibrium strategy is characterized by partitioning her type space T with "intervals."

Definition 2. A subset I of player 1's (finite) type space T is an **interval** if for any $t, t' \in I$ such that $t < t'$ and for any $t'' \in T$ such that $t < t'' < t'$, it must be that $t'' \in I$.

Given this terminology, the following lemma is immediate from Lemma 1.

Proposition 4. *Every perfect Bayesian equilibrium* (s, μ) *is characterized by a partition of player* 1*'s type space* T *with* L *intervals, denoted* $I_1, I_2, \ldots, I_L$*. Moreover, it holds that:*

1. *For any types* $t, t' \in I_\ell$*, player* 2 *takes the same action* $s_2(s_1(t)) = s_2(s_1(t'))$*.*
2. *For all types* $t_\ell \in I_\ell$ *across all* $\ell = 1, 2, \ldots, L$*, player* 2 *will take different actions* $s_2(s_1(t_1)) < s_2(s_1(t_2)) < \cdots < s_2(s_1(t_L))$*.*

It is helpful to review the previous propositions in the language of intervals. Proposition 1 says that there exists a (babbling) equilibrium with a single interval $\{1, 2, \ldots, K\}$, and Proposition 2 says that if $b \leq \frac{1}{2}$ then there exists a (truth-telling) equilibrium with K intervals $\{1\}, \{2\}, \ldots, \{K\}$. Proposition 3 says that if $b > \frac{1}{2}$ then every perfect Bayesian equilibrium has at most $K - 1$ intervals.

References

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, *50*(6), 1431–1451.