Maximal Miscommunication *

Shintaro Miura† Takuro Yamashita‡

December 2, 2019

Abstract

We consider a cheap-talk game à la Crawford and Sobel (1982), where the sender could be an honest type but this probability is not common knowledge. We show that there exists a Harsanyi type space with a unique equilibrium where the receiver may play *any* action under *any* state of nature.

1 Introduction

People tend to communicate honestly even though they may gain by not telling the truth.[1] The literature investigate impacts of such behaviors in cheap-talk games. The typical formulation assumes that the sender could be an honest type who tells the truth, with this

*This paper has been circulated as "On the Possibility of Information Transmission" and as "Divergent Interpretation and Divergent Prediction in Communication". We are grateful to Attila Ambrus, Ying Chen, Sidartha Gordon, Tadashi Hashimoto, Daisuke Hirata, Kazumi Hori, Hideshi Itoh, Atsushi Kajii, Shinsuke Kambe, Navin Kartik, Kohei Kawamura, Frederic Koessler, Botond Köszegi, Takashi Kunimoto, Christoph Kuzmics, Michel Le Breton, Ming Li, Shih En Lu, George Mailath, Kota Murayama, Marco Ottaviani, Daisuke Oyama, Di Pei, Takashi Shimizu, Joel Sobel, Satoru Takahashi, Takashi Ui, Yuichi Yamamoto, and all seminar participants at CTWE, EEA-ESEM 2014, 20th DC Conference, ASSET, Game Theory Workshop, and Bocconi, for their helpful comments. We thank the editor and referees for insightful suggestions. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 714693), and from the JSPS Grant-in-Aid for Young Scientists (16K17093, 19K13655). All remaining errors are our own.

†Department of Economics, Kanagawa University and Waseda Business School, smiura@kanagawa-u.ac.jp.

‡Toulouse School of Economics, takuro.yamashita@tse-fr.eu.

[1]See, for example, Abeler, Nosenzo, and Raymond (2019).

probability as common knowledge (e.g., Chen, 2011). While it appears to be a reasonable "first step" to provide tractable frameworks, there remains concerns about its robustness. That is, the results in this formulation may not be purely due to the existence of honest types, but rather due to a particular belief structure: the receiver correctly expects types of the sender, and the sender also correctly expects it, and so on. It is natural to study how the results would change with less common knowledge about the sender's honesty. Given this concern, in this paper, we study a cheap-talk game (Crawford and Sobel, 1982), where different receiver types may assign different probabilities to the sender's honesty and where different sender types may assign different probabilities to those receiver types.

We show that there exists a type space (à la Harsanyi, 1967-68) with a unique perfect Bayesian equilibrium (PBE, hereafter) that exhibits *maximal miscommunication*, in the sense that a receiver plays *any* action given *any* state. That is, the receiver's action can be arbitrarily different from the full-information action, even if the degree of disagreement among the players is small. This is in stark contrast to the standard models. We interpret this result as a caution for modeling with certain behavioral types: without additional information about the players' belief structures regarding those behavioral types, it may be difficult to obtain meaningful predictions.

2 Model

There are two players, a sender ($i = 1$) and a receiver ($i = 2$). The sender knows the true payoff-state, $\theta \in \Theta = \mathbb{R}$, while the receiver does not. They play a message game where the sender sends a message $m \in M = \mathbb{R}$ in the first stage, and then the receiver takes an action $a \in A = \mathbb{R}$ in the second stage after observing m. The sender's payoff is $u(a, \theta, \varepsilon) = -(a - \theta - \varepsilon)^2$, and the receiver's payoff is $v(a, \theta) = -(a - \theta)^2$, where $\varepsilon \in \mathbb{R}$ represents the difference in their preferences, called the *bias*. We assume that the sender privately knows his behavioral type $d \in D = \{0, 1\}$: an *honest type* ($d = 0$), who always sends $m = \theta$, and a *strategic type* ($d = 1$), whose strategy is determined in an equilibrium.

The players' belief structure is described by a Harsanyi type space, denoted by $\mathcal{T} = (T_1, T_2, b_1, b_2)$. For each $i = 1, 2$, player i's *type* is an element t_i of a measurable space T_i. Player i's belief is a measurable mapping $b_i : T_i \to \Delta(T_{-i})$. The sender knows the true state

θ and his behavioral type. Therefore, let $\theta(t_1) \in \Theta$ and $d(t_1) \in D$ denote the true state and the sender's behavioral type when his type is t_1, respectively. These mappings $\theta(\cdot)$ and $d(\cdot)$ are assumed to be measurable. We denote by $T = T_1 \times T_2$ the set of type profiles, and $\mathbb{T}^\varepsilon$ represents the class of the type spaces.

Given $\mathcal{T}$, let $\sigma_1 : T_1 \to M$ denote the sender's pure strategy, and $\sigma_2 : T_2 \times M \to A$ denote the receiver's pure strategy. The solution concept is PBE, denoted by $\sigma^* = (\sigma_1^*, \sigma_2^*)$.[2] Let $A(\theta|\sigma^*)$ denote the set of equilibrium actions of the receiver when the true state is θ in equilibrium σ^*, that is:

$$A(\theta|\sigma^*) = \left\{\, a \in A \mid \exists (t_1, t_2) \in T \text{ s.t. } \theta(t_1) = \theta,\ a = \sigma_2^* \left(t_2, \sigma_1^*(t_1)\right) \right\}. \tag{1}$$

3 Main results

We show that there exists a type space in $\mathbb{T}^\varepsilon$, where its unique PBE exhibits maximal miscommunication.

Theorem 1. For any $\varepsilon > 0$, there exists $\mathcal{T} \in \mathbb{T}^\varepsilon$ such that: (i) a PBE σ^* exists and unique; and (ii) $A(\theta|\sigma^*) = A$ for every θ.

We focus on a Harsanyi type space analogous to the level-k theory, as in Figure 1: "level-0" players agree that the sender is certainly honest, and the strategic type of "level-$k(\geq 1)$" sender (resp. receiver) certainly believes that the opponent is level-$(k-1)$ (resp. k). While the sender exaggerates messages and the receiver discounts them given their beliefs, they are not canceled out, which is a contrast to the standard models. The sender could send different messages under the same payoff-state and the receiver could differently respond to the same message. Such miscommunication is exacerbated as the "level" increases, implying the maximal miscommunication. Notice that the result holds for any $\varepsilon > 0$, however small it is. Thus, even if the players' preferences are almost aligned, maximal miscommunication might appear.[3]

[2] Because, in the equilibrium we study, the posterior is fully specified by Bayes' rule, its explicit representation is omitted.

[3] While one might think of Theorem 1 as reminiscent of Weinstein and Yildiz (2007), our result is not directly related. In Weinstein and Yildiz (2007), there exists some finite k where each player's k-th order

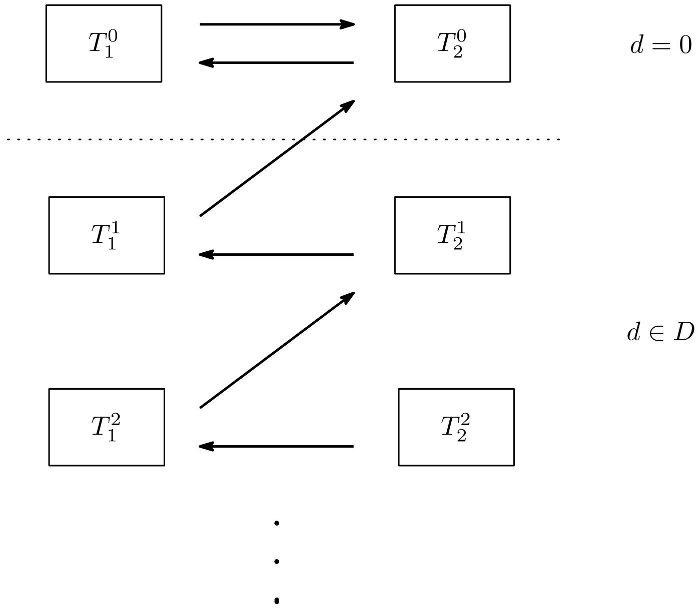

Figure 1: Construction of Harsanyi type space $\mathcal{T}$

Proof. We first construct a Harsanyi type space $\mathcal{T} \in \mathbb{T}^\varepsilon$ as follows. Let $T_1^0 = \{t_1^0(\theta) | \theta \in \Theta\}$ be a subset of the sender types, called "level-0" types. For each $t_1^0(\theta)$, we have:

$$d(t_1^0(\theta)) = 0, \;\; \theta(t_1^0(\theta)) = \theta, \;\text{ and } b_1\left(T_2^0 | t_1^0(\theta)\right) = 1. \tag{2}$$

Let $T_2^0 = \{t_2^0\}$, where t_2^0 is a "level-0" type of the receiver with $b_2\left(T_1^0 | t_2^0\right) = 1$.

Next, let $T_1^1 = \{t_1^1(\theta, d) | \theta \in \Theta, d \in D\}$ be a set of "level-1"types of the sender. For each $t_1^1(\theta, d)$, we have:

$$d(t_1^1(\theta, d)) = d, \;\; \theta(t_1^1(\theta, d)) = \theta, \;\text{ and } b_1\left(T_2^0 | t_1^1(\theta, d)\right) = 1. \tag{3}$$

Likewise, let $T_2^1 = \{t_2^1(p) | p \in [0, 1]\}$ be a set of "level-1" types of the receiver, where $t_2^1(p)$ believes that the sender is either an honest type in T_1^1 with probability p or a strategic type in T_1^1 with probability $1 - p$.

Inductively, given T_2^k for each $k = 1, 2, \ldots$, let T_1^{k+1} be a set of "level-$(k+1)$" types of the sender as follows. For each $t_2 \in T_2^k$, let $T_1^{k+1}(t_2) = \left\{t_1^{k+1}(\theta, d, t_2) \middle| \theta \in \Theta, d \in D\right\}$ be a

belief is *arbitrarily* different from the baseline common-knowledge case, and particularly, it can believe in "crazy" types which have dominant strategies. In our case, it is common knowledge that the sender has either $d = 0$ or 1, and thus, such crazy types cannot exist.

subset of types of the sender, where $t_1^{k+1}(\theta, d, t_2)$ satisfies:

$$d(t_1^{k+1}(\theta, d, t_2)) = d,\ \theta(t_1^{k+1}(\theta, d, t_2)) = \theta,\ \text{and } b_1\left(\{t_2\}\middle|t_1^{k+1}(\theta, d, t_2)\right) = 1. \tag{4}$$

Let $T_1^{k+1} = \bigcup_{t_2 \in T_2^k} T_1^{k+1}(t_2)$. Similarly, let $T_2^{k+1} = \left\{t_2^{k+1}(p, t_2)\middle|p \in [0,1], t_2 \in T_2^k\right\}$ be a set of "level-$(k+1)$" types of the receiver, where $t_2^{k+1}(p, t_2)$ believes that the sender's type is either an honest type in T_1^{k+1} with probability p or a strategic type in $T_1^{k+1}(t_2)$ with probability $1-p$. Finally, we define $T_i = \bigcup_{k=0}^{\infty} T_i^k$ for each i.

Now we construct a unique PBE σ^* given this type space. Define:

$$A^k(\theta|\sigma^*) = \left\{a \in A\middle|\exists t_1 \in T_1^k, t_2 \in T_2^k \text{ s.t. } \theta(t_1) = \theta, a = \sigma_2^*\left(t_2, \sigma_1^*(t_1)\right)\right\}. \tag{5}$$

For the level-0 types, we have $\sigma_1^*\left(t_1^0(\theta)\right) = \theta$ for each $\theta \in \Theta$, and $\sigma_2^*\left(t_2^0, m\right) = m$ for each $m \in M$. Thus, $A^0(\theta|\sigma^*) = \{\theta\}$.

For the sender with $t_1^1(\theta, d = 1) \in T_1^1$, because he believes that the receiver's type is t_2^0, his unique best response is $\sigma_1^*\left(t_1^1(\theta, d = 1)\right) = \theta + \varepsilon$. For the receiver with type $t_2^1(p) \in T_2^1$, after observing message m, she believes that $\theta = m$ and $m - \varepsilon$ with probabilities p and $1-p$, respectively. Therefore, $\sigma_2^*\left(t_2^1(p), m\right) = pm + (1-p)(m-\varepsilon)$, implying that $A^1(\theta|\sigma^*) = [\theta - \varepsilon, \theta + \varepsilon]$.[4]

By induction, suppose that for each $k = 1, 2, \ldots$ and $\delta_k \in [-k\varepsilon, k\varepsilon]$, there exists $t_2 \in T_2^k$ such that $\sigma_2^*(t_2, m) = m - \delta_k$ for each $m \in M$. Because the sender with type $t_1^{k+1}(\theta, d = 1, t_2) \in T_1^{k+1}(t_2)$ believes that the receiver's type is t_2, his unique best response is to send $\sigma_1^*\left(t_1^{k+1}(\theta, t_2)\right) = \theta + \varepsilon + \delta_k$.[5] For the receiver with type $t_2^{k+1}(p, t_2) \in T_2^{k+1}$, after observing message m, she believes that $\theta = m$ and $\theta - \varepsilon - \delta_k$ with probabilities p and $1-p$, respectively. Therefore, $\sigma_2^*\left(t_2^{k+1}, m\right) = pm + (1-p)(\theta - \varepsilon - \delta_k)$, implying that $A^{k+1}(\theta|\sigma^*) = [\theta - (k+1)\varepsilon, \theta + (k+1)\varepsilon]$. Thus, we conclude that $A(\theta|\sigma^*) = A$ for each θ because $A(\theta|\sigma^*) \supseteq \bigcup_k A^k(\theta|\sigma^*) = A$. $\square$

[4] The lower bound is attained when the sender is honest and the receiver has $p = 0$. Likewise, the upper bound is attained when the sender is strategic and the receiver has $p = 1$.

[5] For each $k \geq 1$, equilibrium messages are more exaggerated compared with those of level-0 types, interpreted as *language inflation* (Kartik, Ottaviani, and Squintani, 2007; Kartik, 2009).

4 Conclusion

The paper shows that the existence of a truth-telling type can imply maximal miscommunication. We conjecture that the "level-k" type space introduced here plays an important role in the study of more general games without common knowledge of the entire payoff functions.

References

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for Truth-Telling," *Econometrica*, 87(4), 1115–1153.

Chen, Y. (2011): "Perturbed Communication Games with Honest Senders and Naive Receivers," *Journal of Economic Theory*, 146(2), 401–424.

Crawford, V. P., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451.

Harsanyi, J. C. (1967-68): "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, 159–182,320–334,486–502.

Kartik, N. (2009): "Strategic Communication with Lying Costs," *Review of Economic Studies*, 76(4), 1359–1395.

Kartik, N., M. Ottaviani, and F. Squintani (2007): "Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, 134(1), 93–116.

Weinstein, J., and M. Yildiz (2007): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75(2), 365–400.