A Game-Theoretic Analysis on the Use of Indirect Speech Acts

Mengyuan Zhao
School of Social Sciences
Shanghai, China 200093
mengyuan.zhao@usst.edu.cn

University of Shanghai for Science and Technology

Abstract

In this paper we will discuss why in some circumstances people express their intentions indirectly: the use of Indirect Speech Acts (ISA). Based in Parikh's game of part information and Franke's IBR model, we develop a game-theoretic model of ISA, which is divided into two categories, namely non-conventional ISA and conventional ISA. We assume that non-conventional ISA involves two types of communication situations: communication under certain cooperation and that under uncertain cooperation. We will analyze the cases of ironical request and implicit bribery as typical instances of non-conventional ISA of each situation type, respectively. We then apply our model to analyze the use of conventional ISA from an evolutionary perspective, which is inspired by Lewisian convention theory. Our model yields following predictions: the use of non-conventional ISA under certain cooperation relies on the sympathy between interlocutors, which blocks its evolution towards conventional ISA; in uncertain cooperation situations, people are more likely to use ISA, which helps its conventionalization.

1 Introduction

Yesterday, my husband and I went out for lunch. I could not reach the chopstick box, so I talked to my husband: " I do like eating noodle with a spoon! " My husband stared at me, laughed, and passed me the chopsticks.

I did not ask my husband to pass me the chopsticks directly, but intended to make a request in an ironical way. And he understood my intention correctly.

Like the example above, we often express our intention indirectly rather than mean what the utterance literally says. According to the speech act theory, which is introduced by Austin (1962) and developed by Searle (1969, 1975), this kind of pragmatic phenomenon is called indirect speech act (ISA). Seale (1975) proposes an explanation to the use of ISA, that is, an apparatus based in Gricean principles of cooperative conversation (see Grice 1975). Then here comes the puzzle of indirect speech (Terkourafi 2011): as Gricean principle suggests, cooperative interlocutors should communicate with informative, truthful, relevant and succinct message, but why is indirectness commonly used in our daily communication?

In: T. Ågotnes, B. Liao, Y.N. Wang (eds.): Proceedings of the first Chinese Conference on Logic and Argumentation (CLAR 2016), Hangzhou, China, 2-3 April 2016, published at http://ceur-ws.org

According to Brown and Levinson (1987), ISA is a strategy of politeness. In their *politeness theory*, people would like to adopt some strategies to save each other's face when their communication involves face-threatening acts, such as criticism, insults, disagreement, suggestions, refusal, requests etc. Of all the four strategies, they distinguish between on-record and off-record ISA, which roughly correspond to conventional ISA and non-conventional ISA, respectively. Clark (1996) also argues that the main reason for the use of ISA is to mitigate the threat to face and then to maintain social equity between interlocutors.

However, Pinker and his colleagues (Pinker *et al.*, 2008; Lee and Pinker, 2010) point out that neither Searle's theory nor politeness theory is comprehensive enough to account for the motivation for use of ISA: they both presuppose pure cooperation in human communication, which is not always the case during instances of ISA (e.g. sexual comes-ons, veiled threats and implicit briberies). They propose the *theory of strategic speaker*: in communication games under uncertain cooperation, the speaker chooses the strategy of ISA because it allows for plausible deniability facing an uncooperative hearer. Rather than appealing to a social ritual, the theory offers a strategic rationale to the use of non-conventional ISA by introducing static game model and by building decision functions to represent plausible deniability of ISA.

Actually, Pinker's work is originated in a tradition of game-theoretic pragmatics (see Jäger 2008 for a selective review). The idea of using game as a model of language communication goes back to Wittgenstein (1953). Inspired by this, many attempt to construct game-theoretic models of communication, among which Lewis (1969) started the tradition that communication is taken as at least partially cooperative effort in the model. He not only builds signaling games to solve coordination problem in communication, but gives a game-theoretic interpretation on convention as well: Convention is Nash equilibrium in a special sense. Lewisian convention theory explains the rationale of how meaning is assigned to natural language through their conventional use. Following Lewisian tradition, Parikh (2001, 2007) constructs *game of partial information*, in which he introduces literal meaning of message and makes two types of distinctions, namely, distinction between literal meaning and speaker's intention and that between literal meaning and hearer's interpretation. Van Rooij (2003, 2004, 2008) analyzes on-record indirect request and Horn's strategy (see Horn 1984) in terms of signaling game. By introducing the concept of *risk dominance* as equilibrium selection standard and through the introduction of *super conventional signaling game* model, Sally (2003) and van Rooij (2006) study how sympathy between interlocutors may affect the use of indirectness such as irony. Franke (2009) criticizes that equilibrium as the traditional solution concept does not correspond to actual reasoning process during communication. He introduces *iterated best response* (IBR) reasoning, which formally illustrate how interlocutors departing from literal meaning of messages pick out their strategies based in their belief in each other's rationality and through a process of iterated reasoning. Blume & Board (2014) analyze off-record indirectness through evolutionary game-theoretic model. By adopting vagueness dynamics, they solve the game and explain why interest conflict may encourage speaker to use indirectness. Mialon & Mialon (2013) study analytical conditions of ISA such as terseness, irony and implicit bribery through construction of signaling game and solution of *perfect Bayesian equilibrium* (PBE).

This paper develops a game-theoretic model for ISA. Our model is basically composed of two parts, namely, describing communication situations in terms of signaling games and solving the situations through a reasoning framework. We introduce higher-order belief as the quantification of sympathy between interlocutors, and thus study how sympathy affect player's choice on ISA strategy. The next section describes our model for two situation types: communication under certain cooperation (in the basic model) and communication under uncertain cooperation (in the extended model). For each type, a signaling game is first built, then it is solved. At the end of Section 2, we compare our model to other related models proposed in game theoretic pragmatics. In Section 3 we apply our model to analyze the cases of ironical request and implicit bribery. Based in an evolutionary consideration of our model, Section 4 predicts conventionalization of ISA under the two situation types, respectively. Section 5 provides a summary and suggestions for future work.

2 The Model

2.1 Game of Basic Model

In the game of basic model, we assume that the interlocutors are under certain cooperation, that is, both of them hope that the hearer correctly understands the speaker's intention. Given this, we further assume that a successful communication using ISA will bring extra gain, say $\varepsilon(> 0)$, to both interlocutors. A speaker, S, may have two possible intentions, say $\{t_1, t_2\} \in T$, which he would like to express to a hearer, H. When S has t_1, he may utter a direct message m_1 or an indirect message $\overline{m}$; when S has t_2, he may send a direct message m_2 or $\overline{m}$. We assume that $\overline{m}$ has a literal meaning of m_2. We also assume that H is a sophisticated hearer whose strategy conforms to the following rule: He performs action $a(t_1)$ while hearing m_1, $a(t_2)$ while m_2 and either $a(t_1)$ or $a(t_2)$ while $\overline{m}$.

S and H are under certain cooperation where both prefer the action $a(t_i)$ to be taken with the correspondent intention t_i, where $i = 1, 2$. Taking this along with our assumption of ε above, we define interlocutors' payoff functions as follows.

Definition 1 In basic model, let $U_N(t_i, m, a(t_j))$ be payoff of $N \in \{S, H\}$ given t_i, m and $a(t_j)$, where $i, j = 1, 2$.

$$U_N(t_i, m, a(t_j)) = \begin{cases} 1 & i = j, m \in \{m_1, m_2\} \\ 1+\varepsilon & i = j, m = \overline{m} \\ 0 & i \neq j, m = \overline{m} \end{cases}$$

Definition 1 suggests: Both interlocutors will gain 1 using direct speech; both will earn $1+\varepsilon$, if indirectness is involved and communication goes through successfully; both will get 0, if the use of ISA leads to misunderstanding. We denote by $p_1 \in (0, 1)$ H's prior belief that S has an intention of t_1. Figure 1 illustrates the extensive form of this signaling game.

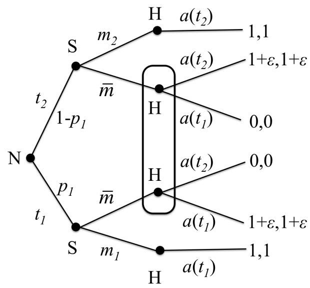

Figure 1: Game of Basic Model under Certain Cooperation

2.2 Solution to the Game of Basic Model

P-added IBR Reasoning Framework

To solve the game of basic model, we introduce the *P-added IBR reasoning framework*. The framework contains two parallel reasoning sequences, namely, the S_0-sequence and the H_0-sequence. We will define the scaffolding of P-added IBR reasoning framework by induction.

Base: Level-Zero Players

S_0-sequence starts from a naïve speaker S_0. We assume that S_0 arbitrarily plays an *intentionally consistent strategy*, which is defined as follows.

Definition 2 A speaker strategy $s(s(t) = m)$ is *intentionally consistent* iff

(I) When $s(t) \in \{m_1, m_2, \cdots m_i, \ldots, m_n\}$, $t = [\![s(t)]\!]$. $[\![\cdot]\!] : M \to T$ is called denotation function that maps the literal meaning of a message to speaker's intention. m_i denotes direct message.

(II) When $s(t) = \overline{m}$, $t \in \{t_1, t_2, \cdots t_i, \cdots, t_n\}$. $\overline{m}$ denotes indirect message.

Definition 2 suggests that indirect message may be used to express all possible intentions in the context. It is also noted that S_0 is not rational in the sense that she chooses the strategy not for it guarantees her a better payoff, but for it corresponds to the general rule of language use given the literal meaning of a message and the context.

H_0-sequence starts from a naïve hearer H_0. We assume that H_0 chooses an arbitrary strategy that will offer her the highest expected payoff given an *intentionally consistent interpretation* of message.

Definition 3 An *intentionally consistent interpretation* is a posterior belief $\mu_0(t|m) = Pr(t|m)$, which results from updating prior beliefs with the intentionally consistent meaning of the observed message. t is the *intentionally consistent meaning* of m iff

(I) When $m \in \{m_1, m_2, \cdot m_i, \cdot, m_n\}$, $t = [\![m]\!]$. m_i denotes direct message.

(II) When $m = \overline{m}$, $t \in \{t_1, t_2, \cdot t_i, \cdot, t_n\}$. $\overline{m}$ denotes indirect message.

Step: Higher Level Types

We assume that player type of level $k+1$ gives best responses to their belief in their opponent type of level k.

Let us take S_0-sequence first. After S_0 sends a message m, H_1, who is strategically one-level higher than S_0, will act according to her posterior belief $\mu_1(t|m)$. We assume that hearers will adapt their posterior belief in a sophisticated way, that is, $\mu_1(t|m)$ is consistent with her belief in speaker's behavior, say $s(t) = m$, as well as her prior belief in t, say $Pr(t)$:

$$\mu_{k+1}(t_j|m_i) = \frac{Pr(t_j) \times S_k(m_i|t_j)}{\sum_{t' \in T} Pr(t') \times S_k(m_i|t')} \tag{1}$$

H_1 will choose the strategy $h(m)$ that offers her the highest expected payoff $EU_{H_1}(a(t))$, which depends on her posterior belief, $\mu(t|m)$, and the corresponding payoff, $U_H(t, m, a(t))$:

$$EU_H(a(t)|m) = \sum_{t_i \in T} \mu(t_i|m) \times U_H(t_i, m, a(t)) \tag{2}$$

$$h(m) = BR(\mu) \in \arg\max_{a(t) \in A} EU_H(a(t)|m) \tag{3}$$

From (1), (2) and (3), H_1's strategy is dependent on $Pr(t)$, given $s(t) = m$ and $U_H(t, m, a(t))$. Let $p_1 = Pr(t)$ and A_1 ($A_1 \subseteq U = \{0, 1\}$) be some interval. H_1's strategy is a mixed strategy of H_1^1 and H_1^2:

$$H_1 = \begin{cases} H_1^1 & \text{if } p_1 \in A_1 \\ H_1^2 & \text{if } p_1 \in \complement_U A_1 \end{cases}$$

We denote by $p_2 \in (0, 1)$ the probability that p_1 falls in A_1. S_2, who is strategically one-level higher than H_1, will play $s(t)$ that guarantees her the highest expected payoff $EU_{S_2}(s(t) = m)$, given her belief in H_1, $\rho_2 = \langle H_1, p_2 \rangle$, and the corresponding payoff, $U_S(t, m, a(t))$:

$$\begin{aligned} EU_S(s(t) = m) &= (p_2 \times \sum_{t_i \in T} \rho_2^1(a(t_i)|m)) \times U_S(t, m, a(t_i)) \\ &\quad + ((1 - p_2) \times \sum_{t_i \in T} \rho_2^2(a(t_i)|m)) \times U_S(t, m, a(t_i)) \qquad (4) \\ s(t) &= BR(\rho) \in \arg\max_{m \in M} EU_S(s(t) = m) \qquad (5) \end{aligned}$$

From (4) and (5), S_2's strategy is dependent on p_2, given H_1 and $U_S(t, m, a(t))$:

$$S_2 = \begin{cases} S_2^1, & \text{if } p_2 \in A_2 \\ S_2^2, & \text{if } p_2 \in \complement_U A_2 \end{cases}, \text{ where } A_2 \subseteq U = \{0, 1\}$$

.

Similarly, H_3's strategy is dependent on p_3, which denotes the probability distribution of p_2. Inductively, in S_0-sequence, $S_{2k+2} = \{s \in S | \exists \rho : \rho = \langle H_{2k+1}, p_{2k+2} \rangle, s \in BR(\rho)\}$ and $H_{2k+1} = \{h \in H | \exists \mu : \mu$ is consistent with Pr and $\sigma, \sigma = \langle S_{2k}, p_{2k+1} \rangle, h \in BR(\mu)\}$, where $k > 0$.

H_0-sequence follows exactly the same rule as that for S_0-sequence. The only difference is that we denote by $p'_{2k+2} \in (0, 1)$ the probability on which S_{2k+1} depends and $p'_{2k+3} \in (0, 1)$ the probability on which H_{2k+2} depends. Then $S_{2k+1} = \{s \in S | \rho : \rho = \langle H_{2k}, p'_{2k+2} \rangle, s \in BR(\rho)\}$ and $H_{2k+2} = \{h \in H | \exists \mu : \mu$ is consistent with Pr and $\sigma, \sigma = \langle S_{2k+1}, p'_{2k+3} \rangle, h \in BR(\mu)\}$, where $k > 0$.

In the inductive steps above, p_k (or p'_k)represents k-order belief in Pr of $N(\in \{S, H\})$.We define *higher-order belief* in Pr as follows.

Definition 4 P is higher-order belief in Pr iff

(I) In S_0-sequence, H's prior belief in $T = \{t_1, t_2, \cdots, t_n\}$ is $Pr = p_1$. p_{2k+2} (or p_{2k+1}) that determines S_{2k+2}'s (or H_{2k+1}'s) belief in H_{2k+1}'s (S_{2k}'s) behavior represents the probability distribution of p_{2k+1} (or p_{2k}), $P = \{p_1, p_2, \cdots, p_n\}$.

(II) In H_0-sequence, H's prior belief in $T = \{t_1, t_2, \cdots, t_n\}$ is $Pr = p_1$. p'_{2k+3} (or p'_{2k+2}) that determines S_{2k+1}'s (or H_{2k+2}'s) belief in H_{2k}'s (S_{2k+1}'s) behavior represents the probability distribution of p'_{2k+2} (or p'_{2k+1}), $P = \{p_1, p'_2, \cdots, p'_n\}$.

We further define players' *sympathy* towards each other as follows.

Definition 5 In P-added IBR reasoning framework, S and H share sympathy $\lambda \in (0,1)$ towards each other. When S has intention t_i, $\lambda = p_i(Pr(t_i))$, where $p_i \in P$.

Definition (5) suggests that as interlocutor' higher-order belief in speaker's real intention gets close to 1, their sympathy towards each other increases.

Limit

Since we assume finitely many pure player strategies for finite sets of T, M and A in the game of basic model, the P-added IBR sequence is bounded to repeat itself. We define the *idealized solution* of the reasoning framework as follows.

Definition 6 The *idealized solution* of P-added IBR reasoning framework are all infinitely repeated strategies S^* and H^*:

$$\begin{aligned} S^* &= \{s \in S | \forall i \exists j > i : s \in S_j\} \\ H^* &= \{h \in H | \forall i \exists j > i : h \in H_j\} \end{aligned}$$

The idealized solution can be explained in two senses: first, it represents the reasoning result of individual interlocutors with idealized rationality; second, it marks the final reasoning result of a group of infinite interlocutors after pair-wise plays. The latter has something to do with evolution in that it assumes players improve their strategy types level by level with increasing plays.

Solution Analysis

The following proposition shows a complete characterization of idealized solution to the game of basic model in terms of P-added IBR reasoning framework. Proofs are in the Appendix.

Proposition 1 *Suppose* $\varepsilon \in (0,1)$ *and* p_i *(or* p'_i*)* $\in (0,1)$

$$S^* = \begin{cases} \begin{cases} s(t_1) = \bar{m}, \\ s(t_2) = m_2, \end{cases} & \textit{if } p_i(t_1) > \frac{1}{1+\varepsilon} \textit{ and } p'_i(t_1) > \frac{1}{1+\varepsilon}, \\ \begin{cases} s(t_1) = m_1, \\ s(t_2) = \bar{m}, \end{cases} & \textit{if } p_i(t_1) < \frac{\varepsilon}{1+\varepsilon} \textit{ and } p'_i(t_1) < \frac{\varepsilon}{1+\varepsilon}, \\ \begin{cases} s(t_1) = m_1, \\ s(t_2) = m_2, \end{cases} & \textit{if } \frac{\varepsilon}{1+\varepsilon} < p_i(t_1) < \frac{1}{1+\varepsilon} \textit{ and } \frac{\varepsilon}{1+\varepsilon} < p'_i(t_1) < \frac{1}{1+\varepsilon}. \end{cases}$$

$$H^* = \begin{cases} \begin{cases} h(m_1) = a(t_1), \\ h(m_2) = a(t_2), \\ h(\bar{m}) = a(t_1), \end{cases} & \textit{if } p_i(t_1) > \frac{1}{2} \textit{ and } p'_i(t_1) > \frac{1}{2}, \\ \begin{cases} h(m_1) = a(t_1), \\ h(m_2) = a(t_2), \\ h(\bar{m}) = a(t_2), \end{cases} & \textit{if } p_i(t_1) < \frac{1}{2} \textit{ and } p'_i(t_1) < \frac{1}{2}. \end{cases}$$

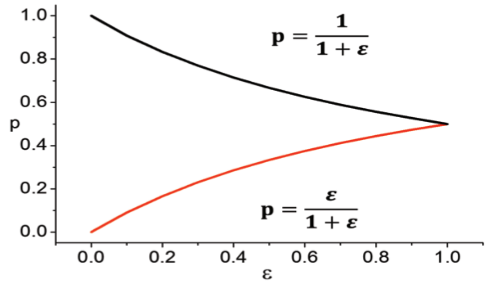

Figure 2: Functions of $p = \frac{1}{1+\varepsilon}$ (black curve) and that of $p = \frac{\varepsilon}{1+\varepsilon}$ (red curve)

Figure 2 illustrates functions of $p = 1/(1+\varepsilon)$ and $p = \varepsilon/(1+\varepsilon)$. Proposition 1 suggests that when the coordinates of $\langle \varepsilon, p \rangle$ falls in the area above the black curve or that below the red curve, interlocutors will use ISA, or else they will communicate explicitly. It is shown that with increase of ε, the area between black curve and red curve gets smaller. Furthermore, with p getting close to 1, we need smaller ε to satisfy $p > \frac{1}{1+\varepsilon}$.

As illustrated in Figure 2, the following corollary follows immediately from Proposition 1:

Corollary 2 *In game of basic model, where S and H are under certain cooperation, with N's higher-order belief in S's real intention $p(t_i)$ getting close to 1, there needs smaller stimulation from ε for N to play ISA strategy, namely $s(t_i) = \bar{m}$ and $h(\bar{m}) = a(t_i)$.*

Ceteris paribus, interlocutors are more likely to use ISA when their higher-order belief in speaker's real intention is more certain, which means that they know each other better, and they share more sympathy towards each other.

2.3 Game of Extended Model

the game of basic model, we assume that S and H are totally unknown to each other. With Definition 4 and Definition 5, we denote by $p_i = \frac{1}{2}$ their higher-order belief in t_1. S may have two asymmetric intentions t_1 and t_2 in the sense that she gets extra gain, ε, if H acts in favor of t_1 rather than in the case of t_2. We also assume that the interlocutors are under uncertain cooperation while facing t_1, that is, S is not sure whether H acts in favor of or adversely to her when H understands t_1. Given this, we introduce that H is one of two types, $\alpha \in \{\alpha_1, \alpha_2\}$: A non-cooperative type, α_1, who acts adversely to her belief in t_1, i.e. chooses $\bar{a}(t_1)$; and a cooperative type, say α_2, who acts in favor of her belief in t_1, i.e. chooses $a(t_1)$. We also assume that H is cooperative with her belief in t_2. We denote by $q \in (0,1)$ S's belief in H's type of α_1. The assumption on $\bar{m}$ is almost the same as in the basic model: In both t_1 and t_2, S may utter $\bar{m}$. The only difference is that in extended model, we assume that S may deny her original intention of t_1 with a cost, say ε', when she utters $\bar{m}$ and later finds out H's type of α_1. However, t_1 expressed by the direct message m_1 is undeniable, so S's explicit strategy $s(t_1) = m_1$ will lead to S's poor payoff, $-\varepsilon'$ with $\bar{a}(t_1)$. We assume that S's loss is greater in the case where α_1 performs $\bar{a}(t_1)$ towards m_1 than in that S denies t_1 after uttering $\bar{m}$, that is, $-\varepsilon'' < \varepsilon'$. α_1 earns 0 when S successfully denies t_1, otherwise, α_1 earns 1. α_2 earns $1+\varepsilon$ when she performs $a(t_1)$ in t_1. α_2 earns 0 when she misunderstands $\bar{m}$ in either t, otherwise, α_2 earns 1. Figure 3 illustrates the extensive form of this signaling game.

2.4 Solution to the Game of Extended Model

The following proposition shows a complete characterization of idealized solution to the game of extended model in terms of P-added IBR reasoning framework. Proofs are in the Appendix.

Proposition 3 *Suppose $\varepsilon \in (0,1)$, $-\varepsilon'' < \varepsilon' < 0$ and $q \in (0,1)$.*

$$S^* = \begin{cases} s(t_1) = \bar{m} \\ s(t_2) = m_2 \end{cases}, \ \alpha_1(H^*) = \begin{cases} h(m_1) = a(t_1) \\ h(m_2) = a(t_2) \\ h(\bar{m}) = a(t_2) \end{cases}, \ S^* = \begin{cases} s(t_1) = \bar{m} \\ s(t_2) = m_2 \end{cases}$$

.

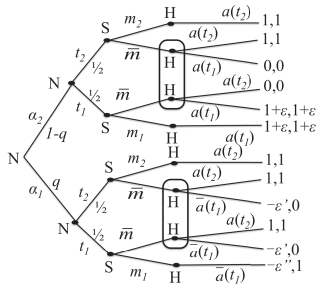

Figure 3: Game of Extended Model under Uncertain Cooperation.

The following corollary follows immediately from Proposition 3:

Corollary 4 *In game of extended model, where S and H are under uncertain cooperation, S will play ISA strategy with* t_1 *and she will play explicit strategy with* t_2*.*

Ceteris paribus, the speaker is more likely to use ISA when she has an intention that may induce adverse action from an uncooperative hearer. The non-cooperative hearer will not act adversely towards ISA, which is plausibly deniable.

2.5 Model Comparison

Parikhian Game of Partial Information

The main differences between our game model and Parikhian model (2001, 2007) are as follows. First, we assume that $t(\in T)$ represents speaker's intention, while Parikh assumes that t represents the game situation. Second, we add the collection variable P as a quantification of sympathy between players. P is the higher-order belief in speaker's intention, t. More specifically, p_1 $(\in P)$ is hearer's first-order belief or her prior belief in t, p_2 is the speaker's belief in hearer's prior belief, namely, p_2 is the speaker's second-order belief in t, etc. In Parikhian model, p denotes probability distribution on situation set T. Third, we introduce idealized solution of P-added IBR reasoning framework as the solution to the game, while Parikh adopts equilibrium as the solution concept and solve the game through equilibrium selection by introducing Pareto dominance as the selection standard.

Comparing to Parikhian model, our model consider how sympathy may affect use of ISA strategy, and we also take pragmatic reasoning as an iterated reasoning process.

Sally's Sympathy Theory

Sally (2003) studies pragmatic phenomena such as irony and indirectness in the term of his sympathy theory. The core idea of Sally's sympathy theory is that social interaction and intimacy between game players may affect the solution by influencing their payoff (Sally, 2000, 2001, 2002). More specifically, for player i and j, if $u_i\langle s_i, s_j\rangle$ indicates i's payoff independent of j, and λ_{ij} designates sympathy between i and j, then the final payoff of i is $u_i\langle s_i, s_j\rangle + \lambda_{ij} u_j\langle s_i, s_j\rangle$. Sally (2001, 2003) suggests that λ_{ij} depends on physical and psychological distance between players: λ_{ij} is 0 or negative for enemies or strangers, and is close to 1 for family or close friends.However, Sally's approach that models sympathy as the degree of common interest of players does not fit signaling games. The main reason is that signaling games involve multiple situations, which leads to multiple payoff matrixes. Different situations make equilibrium less practical as a solution, and thus block Sally's sympathy model.

In our model, collection variable P as higher-order belief in speaker's intention is used to reflect sympathy between interlocutors. We assume that people who share more sympathy know better about each other and thus have a greater chance to make a correct prediction of each other's belief.

Franke's IBR Model

The main differences between our P-added IBR reasoning framework and Franke's IBR model(2009) are as follows. First, Franke assumes that the reasoning starts from a focal point of message semantic meaning. He assumes that the naïve speaker S_0 may choose arbitrarily true message to express her intention and the naïve hearer H_0 may react towards her literal interpretation of an observed message. In our model, we assume that

S_0 plays intentionally consistent strategy (Definition 2) and H_0 makes best response to her belief updated by intentionally consistent interpretation of observed message (Definition 3). Second, Franke assumes that the player type of level $k+1$ gives best response to their unbiased belief in their opponent type of level k. Specifically, $N_{k+1}(N \in \{S, H\})$ will average over all possible actions she believes that N_k may take at every iterated reasoning step. In our model, we introduce higher-order belief, p, that represents the probability distribution on type of level $k+1$ player's belief in type of level k player's behavior. In other words, Franke simply assumes that $p = \frac{1}{2}$ for the corresponding p in our model.

Comparing to Franke's model, our model has a stronger pragmatic explanation power in two senses: His assumption of semantic meaning as a focal point blocks the way to analyze pragmatic phenomenon that involves use of messages going against their literal meaning (e.g. irony and metaphor),while our model gives up this assumption and allows ISA to express all possible intentions consistent with the context; unlike Franke's model, our introduction of higher-order belief enable our model to analyze how sympathy between interlocutors affect use of ISA strategy.

Mialon & Mialon's Model

Mialon & Mialon (2013) builds a signaling game model which yields analytical conditions for ISA, and they applyit to an analysis of terseness, irony and implicit bribery. They discuss the use of ISA strategy in the cases where successful communication provides greater benefit to the hearer than to the speaker. In comparison, weassumes symmetric payoff under certain cooperation situation (as in Definition 1) and in uncertain cooperation case, we consider how payoff may be affected by plausible deniability of ISA. In addition, Mialon & Mialon distinguish two hearer types, namely naïve type and sophisticated type, while we simply consider the sophisticated hearer type. Finally and most importantly, Mialon & Mialon adopts the traditional solution concept of PBE, while we use idealized solution in terms of our P-added IBR reasoning framework, which offers a more intuitive solution as discussed above.

3 Applications

3.1 Ironical Request

We now employ the basic model to provide a systematic analysis on a typical instance of ISA, ironical request.

Example Yesterday, my husband and I went out for lunch. I could not reach the chopstick box, so I talked to my husband: "I do like eating noodle with a spoon!"

Correspondence with the Basic Model

My husband believes that I may have two possible intentions: I request him to pass me the chopsticks, say t_1, or I sincerely express my preference to eating noodle with a spoon, say t_2. When I have t_1, I may explicitly utter, "Pass me the chopstick", say m_1, or ironically,"I do like eating noodle with a spoon", say $\bar{m}$. When I have t_2, I may explicitly utter "I plainly like eating noodle with a spoon", say m_2, or $\bar{m}$. It is obvious that $\bar{m}$ has the literal meaning of m_2. My husband will pass me the chopsticks when he hears m_1, say he performs $a(t_1)$, and he will not pass me the chopsticks when he hears m_2, say he performs $a(t_2)$. If he hears $\bar{m}$, he may perform either $a(t_1)$ or $a(t_2)$. My husband and I love each other and we both prefer that he understands my real intention and act accordingly. If I express explicitly, my husband will act according to my intention for sure and we both gain a plain payoff, say 1. If I express implicitly and my husband interprets it correctly, both of us will gain a better payoff, say $1 + \varepsilon$, where $\varepsilon > 0$. The values assigned are based in the following considerations: When I want my husband to pass me the chopsticks and express ironically, my husband's correct interpretation make us feel close to each other for he knows me well; when I just want to express my special preference and he does not interpret it ironically, we also feel happy for he is one of few people that knows about my preference. In contrast, if my husband misunderstands my implicit word, neither of us is happy, and thus we gain tiny, say 0.

Analysis

My husband and I are so close that we share high degree of sympathy with each other, say $\lambda = 1$. That means we have known each other for a long time, so we are more likely to correctly guess each other's intention in a certain context. He has a large chance to correctly get my intention, I have a large chance to correctly guess that he can correctly understand my intention, and so on. Namely, our higher-order belief in my real intention is certain, say $p = 1$. Thus according to Corollary 1, we are more like to use ISA strategy in the case of ironical request.■

3.2 Implicit Bribery

We now employ the extended model to provide a systematic analysis on another instance of ISA, implicit bribery. The example originally comes from Pinker *et al.* (2008) and Lee and Pinker (2010).

Example Bob is stopped by a police officer for running a red light. When the police officer asks him to show his driving license, Bob takes out his wallet and says, "Gee, officer, is there some way we could take care of the ticket here?" (Pinker *et al*, 2008:833)

Correspondence with the Extended Model

Bob never saw this police officer before, so they are totally unknown to each other. The officer guesses Bob may have two possible intentions: Bob intends to bribe him, say t_1, and Bob has no intention to bribe, say t_2. Both know that if Bob bribe successfully, Bob will gain more than he pays the ticket. Bob has no idea whether he is caught by an honest officer who does not accept bribery, say type α_1 officer, or by a corrupt officer, say type α_2 officer. When Bob intends to bribe, he may offer explicit bribery by saying, "I'll give you $50 if you let me go", say m_1, or he may bribe implicitly by uttering, "Gee, officer, is there some way we could take care of the ticket here?", say $\bar{m}$. When Bob does not intend to bribe, he may just say, "I'm sorry and I'll be more careful next time", say m_2, or he may use $\bar{m}$. While hearing m_1, an honest officer will arrest Bob for bribery, say performing $\bar{a}(t_1)$, which leads to a very low payoff for Bob, say $-\varepsilon''$, and a plain payoff for himself, say 1. But a corrupt officer will accept the bribery and let Bob go, say performing $a(t_1)$, which offers a good payoff for both Bob and himself, say $1+\varepsilon$. While hearing m_2, both honest and corrupt officers will ask Bob to pay the ticket, say performing $a(t_2)$, which gives both a plain payoff. When the honest officer hears $\bar{m}$, if he interprets it as a bribery, Bob will deny it, whether or not he actually intended to bribe, which not only gives Bob a relatively lower payoff with the effort of denial, say $-\varepsilon'$, but also gives himself a relatively lower payoff with a cost of accepting the denial. And if the honest officer interprets $\bar{m}$ as non-bribery, he will simply ask Bob to pay the ticket, which results in plain payoff for both Bob and himself. If corrupt officer correctly interprets $\bar{m}$ as bribery, he will accept it, which results in a good payoff for both. If he misunderstands $\bar{m}$ as bribery, though he is ready to accept it, he gets nothing, and both just get the plain payoff. If the corrupt officer correctly interprets $\bar{m}$ as non-bribery, Bob will pay the ticket and both get plain payoff. If corrupt officer misunderstands $\bar{m}$ as non-bribery, both gain less for they lose the chance of getting more, say they get 0.

Analysis

Bob and the police officer do not know each other, so for Bob's intention of bribery, he is not certain whether the officer will cooperate or not. According to Corollary 2, Bob will play ISA strategy with intention of bribery, and he will not play ISA strategy if he does not intend to bribe.■

4 An Evolutionary View

We now develop our model in an evolutionary view: We combine our model with Lewisian convention theory to analyze how conventional ISA evolves.

4.1 Analysis on Convention

Lewis (1969) gives a game-theoretic explanation of convention:

> A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that, and it is common knowledge in P that, in any instance of S among members of P,
>
> (1) everyone conforms to R;
>
> (2) everyone expects everyone else to conform to R;
>
> (3) everyone has approximately the same preferences regarding all possible combinations of actions;
>
> (4) everyone prefers that everyone conform to R, on condition that at least all but one conforms to R;

(5) everyone would prefer that everyone conform to R', on condition that at least all but one conforms to R',

where R' is some possible regularity in the behavior of members of P in S, such that no one in any instance of S among members of P could conform both to R' and to R. (Lewis, 1969:76)

Lewisian definition of convention suggests that the formation of convention originates in people's expectation towards each other and in their reasoning dependent on their own preference. He proposes that this expectation comes from precedence: In previous cases, if people respect a regularity that they express some intention by a specific message, and they expect that others prefer to conform to this regularity with the same expectation as themselves do, then they are prone to continuously conform to this regularity in order to maximize their common interest.

However, Lewis does not explain where precedence comes from. We propose that even this precedence comes from people's rationality and their belief in rationality. In process of iterated reasoning, people as a group will evolve towards reaching idealized rationality. Our P-added IBR reasoning framework, which starts from intentionally consistent use and interpretation of message offers an approach to model the formation of precedence. Figure 4 shows the schema of how convention forms in an evolutionary perspective of our model combining Lewisian convention theory.

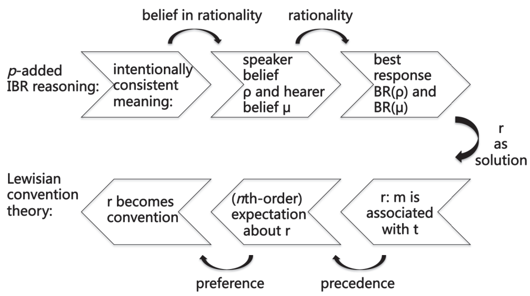

Figure 4: Schema of the formation of language convention.

4.2 Predictions

In the case of conventional ISA, people generally use it without considering about its literal meaning. For instance, when I say, "Can you pass the salt", you will take it as a request without reasoning on whether or not I ask about your ability to pass the salt. We predict that the formation of conventional use of ISA has the following rationale: In communication games, people follow a reasoning pattern that can be modeled as our P-added IBR reasoning framework; after repeated play, their strategies gradually evolve towards the model's idealized solution, which illustrates systematic conditions of use of ISA with ideal rationality; once the solution becomes a precedence, it satisfies self-perpetuating process in the continuous games, and the corresponding use of ISA strategy becomes a convention. Corollary 1 and Corollary 2 of our model come from idealized solution to the game of basic model and that of extended model. The following predictions follow immediately from those corollaries:

(I) The use of non-conventional ISA under certain cooperation relies on the sympathy between interlocutors, which blocks its evolution towards conventional ISA.

(II) In uncertain cooperation situations, people are more likely to use ISA, which helps its conventionalization.

5 Summary And Future Work

In this paper, we develops a game-theoretic model to analyze the rationale of ISA. The model provides analytical conditions for the use of ISA and predicts conventionalization of ISA in an evolutionary perspective. We propose that in situations under certain cooperation interlocutors who share more sympathy are more likely to use ISA, while in uncertain cooperation situations people are more likely to use ISA for its plausible deniability. We apply our model to the analysis of typical instances of non-conventional ISA, namely ironical request and implicit bribery. The solution of our model predicts that ISA used under uncertain cooperation (e.g. implicit bribery) is more likely to be conventionalized than that used under certain cooperation, because the latter depends on request on interlocutors' sympathy.

Our model can be further developed in at least the following three ways. First, it might be interesting to compare our predictions with research results from corpus study in ISA. Second, it might be fruitful to test the justification of our assumption that the use of ISA has something to do with our rationality in the area of neuroscience. For instance, fMRI experiments can be designed and performed to test whether there exists activation of the neuroanatomic regions related to decision making during our processing of ISA. Third, it might be meaningful to explore computer simulation of our model within the research area of artificial intelligence.

Appendix

Proof of Proposition 1

First look at S_0-sequence. Given Definition 2,

$$S_0 = \begin{cases} s(t_1) = m_1, \bar{m} \\ s(t_2) = m_2, \bar{m} \end{cases}$$

.

Given (1), $\mu_1 = (t_1|\bar{m}) = p_1$ and $\mu_1 = (t_2|\bar{m}) = 1 - p_1$. Given (2), $EU_{H_1}(a(t_1)|\bar{m}) = p_1 \times (1 + \varepsilon)$ and $EU_{H_1}(a(t_2)|\bar{m}) = (1 - p_1) \times (1 + \varepsilon)$. Given (3),

$$H_1 = \begin{cases} \begin{cases} h(m_1) = a(t_1), \\ h(m_2) = a(t_2), \\ h(\bar{m}) = a(t_1), \end{cases} & \text{if } p_1 > \frac{1}{2} \\ \begin{cases} h(m_1) = a(t_1), \\ h(m_2) = a(t_2), \\ h(\bar{m}) = a(t_2), \end{cases} & \text{if } p_1 < \frac{1}{2} \end{cases}$$

Let $p_2 = p(p_1 > \frac{1}{2})$ and given (4), $EU_{S_2}(\bar{m}|t_1) = p_2 \times (1+\varepsilon)$ and $EU_{S_2}(\bar{m}|t_2) = (1-p_2) \times (1+\varepsilon)$. Given (5),

$$S_2 = \begin{cases} \begin{cases} s(t_1) = \bar{m}, \\ s(t_2) = m_2, \end{cases} & \text{if } p_2(t_1) > \frac{1}{1+\varepsilon} \\ \begin{cases} s(t_1) = m_1, \\ s(t_2) = \bar{m}, \end{cases} & \text{if } p_2(t_1) < \frac{\varepsilon}{1+\varepsilon} \\ \begin{cases} s(t_1) = m_1, \\ s(t_2) = m_2, \end{cases} & \text{if } \frac{\varepsilon}{1+\varepsilon} < p_2(t_1) < \frac{1}{1+\varepsilon} \end{cases}$$

Let $p_3 = p(\frac{p(p_2(t_1) > \frac{1}{1+\varepsilon})}{p(p_2(t_1) < \frac{\varepsilon}{1+\varepsilon})} > 1)$ and given (2) and (3),

$$H_3 = \begin{cases} \begin{cases} h(m_1) = a(t_1), \\ h(m_2) = a(t_2), \\ h(\bar{m}) = a(t_1), \end{cases} & \text{if } p_3 > \frac{1}{2} \\ \begin{cases} h(m_1) = a(t_1), \\ h(m_2) = a(t_2), \\ h(\bar{m}) = a(t_2), \end{cases} & \text{if } p_3 < \frac{1}{2} \end{cases}$$

Notably, S_0-sequence starts repetition from H_3. Then $H_3 = H^*$ and $S_2 = S^*$. Similarly, H_0-sequence leads to the same solution.■

Proof of Proposition 2

First look at S_0-sequence. Given Definition 2,

$$S_0 = \begin{cases} s(t_1) = m_1, \bar{m} \\ s(t_2) = m_2, \bar{m} \end{cases}$$

.

Given(1), $\alpha_1(\mu_1(t_1|\bar{m})) = \alpha_1(\mu_1(t_2|\bar{m})) = \frac{1}{2}$ and $\alpha_2(\mu_1(t_1|\bar{m})) = \alpha_2(\mu_1(t_2|\bar{m})) = \frac{1}{2}$. Given (2), $EU_{\alpha_1(H_1)}(\bar{a}(t_1)|\bar{m}) = 0$, $EU_{\alpha_1(H_1)}(a(t_2)|\bar{m}) = 1$, $EU_{\alpha_2(H_1)}(a(t_1)|\bar{m}) = \frac{1}{2}(1+\varepsilon)$ and $EU_{\alpha_2(H_1)}(a(t_2)|\bar{m}) = \frac{1}{2}$. Given (3),

$$\alpha_1(H^*) = \begin{cases} h(m_1) = \bar{a}(t_1) \\ h(m_2) = a(t_2) \\ h(\bar{m}) = a(t_2) \end{cases}, \alpha_2(H^*) = \begin{cases} h(m_1) = a(t_1) \\ h(m_2) = a(t_2) \\ h(\bar{m}) = a(t_1) \end{cases}$$

.

Given (4), $EU_{S_2}(s(t_1) = \bar{m}) = q + (1-q) \times (1+\varepsilon)$, $EU_{S_2}(s(t_1) = m_1) = q \times (-\varepsilon'') + (1-q) \times (1+\varepsilon)$, $EU_{S_2}(s(t_2) = \bar{m} = q$, $EU_{S_2}(s(t_2) = m_2) = 1$. Given (5),

$$S_2 = \begin{cases} s(t_1) = \bar{m} \\ s(t_2) = m_2 \end{cases}$$

.

Obviously, S_0-sequence starts repetition from S_2. Then $H_1 = H^*$ and $S_2 = S^*$. Similarly, H_0-sequence leads to the same solution.■

References

[Aus62] J. L. Austin. *How to Do Things with Words– 2nd Edition.* Harvard University Press, 1962.

[BB14] A. Blume and O. Board. Intentional vagueness. *Erkenn*, 79(4 Supplement):855–899, 2014.

[BL87] P. Brown and S. C. Levinson. *Politeness: Some Universals in Language Usage.* Cambridge University Press, 1987.

[Cla96] H. H. Clark. *Using Language.* Cambridge University Press, 1996.

[Fra09] M. Frank. *Signal to Act: Game Theory in Pragmatics.* University of Amsterdam Institute for Logic, Language and Computation, 2009.

[Gri75] H. P. Grice. Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics – Volume 3 / Speech Acts.* Academic Press, 1975.

[Jäg08] G. Jäger. Applications of game theory in linguistics. *Language and Linguistics Compass*, 2(3):406–421, 2008.

[LP10] J. J. Lee and S. Pinker. Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review*, 117(3):785–807, 2010.

[Lew69] D. Lewis. *Convention: A Philosophical Study.* Harvard University Press, 1969.

[MM13] H. M. Mialon and S. H. Mialon. Go figure: The strategy of nonliteral speech. *American Economic Journal Microeconomics*, 5(2):186–212, 2013.

[Par01] P. Parikh. *The Use of Language.* CSLI Publications, 2001.

[Par07] P. Parikh. Situations, rules, and conventional meaning: Some uses of games of partial information. *Journal of Pragmatics*, 39(5):917–933, 2007.

[PNL08] S. Pinker, M. A. Nowak and J. J. Lee. The logic of indirect speech. *Proceedings of the National Academy of Sciences*, 105(105):833–838, 2008.

[Sal01] D. Sally. On sympathy and games. *Journal of Economic Behavior and Organization*, 44(1):1–30, 2001.

[Sal03] D. Sally. Risky speech: Behavioral game theory and pragmatics. *Journal of Pragmatics*, 35(8):1223–1245, 2003.

[Sea69] J. R. Searle. *Speech Acts: An Essay in the Philosophy of Language.* Cambridge University Press, 1969.

[Sea75] J. R. Searle. Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics – Volume 3 / Speech Acts*. Academic Press, 1975.

[Ter11] M. Terkourafi. Why direct speech is not a natural default: Rejoinder to Steven Pinker's "Indirect speech, politeness, deniability, and relationship negotiation". *Journal of Pragmatics*, 43(11):2869–2871, 2011.

[Roo03] R. van Rooij. Being polite is a handicap: Towards a game theoretical analysis of polite linguistic behavior. *Proceedingsof the 9th Conference on Theoretical Aspects of Rationality and Knowledge*. Los Angeles, 2003.

[Roo04] R. van Rooij. Signalling games select Horn strategies. *Linguistics and Philosophy*, 27(4):493–527, 2004.

[Roo08] R. van Rooij. Games and quantity implicatures. *Journal of Economic Methodology*, 15(3):261–274, 2004.

[Wit53] L. Wittgensten. *Philosophical Investigations*. Blackwell, 1953.