DIFFERENT TYPES OF MISPERCEPTION REGARDING BENEVOLENCE

Shutaro Kawabe, Norimasa Kobayashi
Graduate School of Decision Science and Technology, Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan
shuta_raskolnikov@wit.ocn.ne.jp, nkoba@valdes.titech.ac.jp

ABSTRACT

This paper introduces misperception to bilateral decision situations with benevolent players (Bergstrom, 1999) based on hypergame framework (Bennett, 1979). In Bergstrom (1999), a benevolent player not only considers the opponent's private subutility but also his benevolence towards her. Accordingly, we analyze two cases in which the player misperceives only the opponent's private subutility and only the opponent's benevolence respectively. An interesting result we obtained is that misperception of the opponent's benevolence only has the unit change effect, such that in many applications the effect of opponent's benevolence may be ignored.

Keywords: Hypergames, benevolence, misperception.

INTRODUCTION

In this paper, we generalize the bilateral model of benevolence by Bergstrom (1999), introducing misperception based on hypergame framework (Bennett, 1979). We then specifically ask what effect different types of misperception have.

In Bergstrom (1999), the overall utility of a benevolent player is determined by two factors: her *private subutility* caused by the physical environment (allocation) and *apparent happiness* of other players that she perceives. While Bergstrom (1999) well understands the difficulty of correctly perceiving the individual utility of the opponent, he deals with the simplest case in which apparent happiness and the real individual utility of the opponent approximately coincide. In this paper, we relax this assumption and introduce possible misperception regarding the opponent's utility.

The above complex structure of Bergstrom (1999) motivated us to analyze different types of misperception. Recall that a benevolent player must consider the *overall* utility of the other players *who are also benevolent* (many other models of benevolence and altruism ignore this point). In general, therefore, a player may misperceive both the private subutility and the benevolence level of the other players. We ask specifically what effect each of these two types of misperceptions has on decision making respectively. An interesting result we obtained is that the misperception of the opponent's benevolence level has no effect except affine transformation of utility (von Neumann and Morgenstern, 1953).

The paper proceeds as follows. We first introduce the necessary notations and the preceding concepts in the literature for the introduction of our own model. Then we introduce our model, after which we analyze different types of misperception.

PRELIMINARY

Notations

Let

- R be the set of real numbers
- $N = \{1,2\}$ be the set of players
- S be the set of possible consequences
- $v = (v_i)_{i \in N}$ denote the abbreviated representation of vectors

Benevolence Model by Bergstrom (1999)

Here, we formally introduce the bilateral model of benevolent decision makers by Bergstrom (1999). The italicized terms are from Bergstrom (1999).

Denote

- $p_i : S \to R$ player i's *private subutility,*
- $H_i \in R$ player i's *apparent happiness*, and
- $u_i : S \to R$ player i's (overall) benevolent utility
- k_i player i's rate of benevolence satisfying $0 < k_i, 0 < k_i k_j < 1 (j \neq i)$

We deal with the additively separable model such that the players' utilities are represented by

$$u_i(s) = p_i(s) + k_i H_j(s)$$

$$u_j(s) = p_j(s) + k_j H_i(s)$$

Bergstrom (1999) deals with the case in which $H_l = u_l (l = 1,2)$. In this case, the following simultaneous equation is obtained.

$$u_i(s) = p_i(s) + k_i u_j(s)$$

$$u_j(s) = p_j(s) + k_j u_i(s)$$

It is straightforward, as in Bergstrom (1999), to check that the solution of the above equation is

$$u_i(s) = \frac{1}{1 - k_i k_j}(p_i(s) + k_i p_j(s))$$

$$u_j(s) = \frac{1}{1 - k_i k_j}(p_j(s) + k_j p_i(s))$$

Equivalence of Utility Functions

Equivalence of cardinal utility functions is characterized with affine transformation in von Neumann and Morgenstern (1953) as follows.

Definition. Equivalence

Two utility functions u and v *are equivalent except for affine transformation if there exists a positive real number* α *and there exists real number* β *for all* s *in* S *such that* $u(s) = \alpha v(s) + \beta$.

MODEL

We introduce misperception to opponent's utility such that opponent's apparent happiness and opponent's utility are not identical. Conforming to the tradition of the literature of simple hypergames, we assume that the players themselves are not aware of the misperception. Hence, players continue to think that apparent happiness of the opponent is approximated by his real overall benevolent utility. Consequently, each of apparent happiness interpreted by player $i(i \neq j)$ is represented as follows

$$H_i(s) = u_i(s) = p_i(s) + k_i u_j^i(s)$$

$$H_j^i(s) = u_j^i(s) = p_j^i(s) + k_j^i u_i(s)$$

Whereas the subscripts represent the players' physical components, superscripts represent the interpretation.

Substituting the above apparent happiness function into the system of interdependent utility functions of Bergstrom (1999) introduced in the Preliminary, the following subjective utility function is obtained.

Definition 1. The benevolent utility functions interpreted by player i

$$u_i(s) = \frac{1}{1 - k_i k_j^i}(p_i(s) + k_i p_j^i(s))$$

Different Types of Misperception regarding Benevolence

$$u_j^i(s) = \frac{1}{1 - k_i k_j^i}(p_j^i(s) + k_j^i p_i(s))$$

In the definition, we implicitly assume as the standard hypergame literature assumes that the players perceive their own components such as private subutilities correctly. Hence, $p_i^i(s) = p_i(s)$ and $k_i^i = k_i$. Whether a player perceives her overall benevolent utility correctly is another issue, and that is exactly one of the questions we will answer in the analysis section.

The definition implies that each player subjectively percept opponent's benevolence and private subutility.

The explicit definitions about two type misperception and concerned concepts

For more explicit discussion we introduce definitions about two type misperception and concerned concepts.

Definition 3. A misperception about opponent's benevolence

Let k_j *be a level of benevolence of player* j. *If a player* $i(i \neq j)$ *has misperception about opponent's benevolence, then there exist a real number* $\alpha(\neq 0)$ *such that* $k_j^i = k_j + \alpha$ *and* $0 < k_j^i < 1$.

Definition 4. A misperception about opponent's private subutility

Let p_j *be a private subutility function of player* j. *If a player* $i(i \neq j)$ *has misperception about opponent's private subutility, then for all positive real number* α *and for all real number* β *there exists an element* s *in* S *such that* $p_j^i(s) \neq \alpha p_j(s) + \beta$.

Next, we define the effect of two type misperceptions for an opponent's benevolent utility function interpreted by a player and the equivalence between these misperceptions.

Definition 5. The effect of two type misperceptions for an opponent's benevolent utility function interpreted by a player

Let u_j *be a benevolent utility function of player* j. *If there exists the effect of misperceptions about* j*'s benevolence or private subutility for an* j*'s benevolent utility function interpreted by player* $i(i \neq j)$, *then for all positive real number* α *and for all real number* β *there exists an element* s *in* S *such that* $u_j^i(s) = \alpha u_j(s) + \beta$.

Definition 6. The equivalence between misperception about opponent's benevolence and private subutility

Different Types of Misperception regarding Benevolence

Let $\overline{u}_l^i$ *be a benevolent utility function of player* l *interpreted player* i *such that there exists a misperception about opponent's benevolence and* $\hat{u}_l^i$ *be a benevolent utility function of player* l *interpreted player* i *such that there exists a misperception about opponent's private subutility. If the misperception about opponent's benevolence is equivalent to the misperception about opponent's private subutility for player* i*, then for all player* l *there exists positive real number* α_l^i *and there exists real number* β_l^i *for all* s *in* S *such that* $\overline{u}_l^i(s) = \alpha_l^i \hat{u}_l^i(s) + \beta_l^i$.

In general utility theory, positive affine transformation simply means multiplying by a positive number and adding a constant [5]. It turns out that if you subject a utility function to a positive affine transformation, it not only represents the same preferences (this is obvious since an affine transformation is just a special kind of monotonic transformation). Therefore definition 6 suggest that if for a player i a misperception about player $j's(j \neq i)$ benevolence equivalent to misperception about player $j's$ private subutility, then for all player l the benevolent utility function $\overline{u}_l^i$ that the player $i's$ has the misperception about player $j's$ benevolence can a positive affine transformation with $\hat{u}_l^i$ that the player $i's$ has the misperception about player $j's$ private subutility.

In next chapter, we discuss some characteristics about a misperception about opponent's benevolence and a misperception about opponent's private subutility. And we suggest difference between theses two type misperceptions by using these discussions.

ANALYSES OF DIFFERENT TYPES OF MISPERCEPTION

Discussion about characteristics of misperception about opponent's benevolence

Proposition 1. The effect of misperception about opponent's benevolence for own benevolent utility function

Any misperception about opponent's benevolence does not effect own benevolent utility function.

Proof:

Let $\alpha(\neq 0)$ be an arbitrary misperception about $j's$ benevolence of player $i(i \neq j)$ and $\overline{u}_i$ is the benevolent utility function of player i such that player i has a misperception about $j's$ benevolence α and u_i is a benevolent utility function of player i such that there are no misperceptions.

Since $k_j^i = k_j + \alpha$, for all $s \in S$

$$\overline{u}_i(s) = \frac{1}{1-k_i(k_j+\alpha)}(p_i(s)+k_i p_j(s)) .$$

By definition of benevolent utility function,

$$\overline{u}_i(s) = \frac{1-k_i k_j}{1-k_i(k_j+\alpha)} u_i(s) .$$

Therefore for all misperception about opponent's benevolence there exist a positive affine transformation from u_i to $\overline{u}_i$ and the proof of this proposition is complete.

Proposition 2. The effect of misperception of a player about opponent's benevolence for opponent's benevolent utility function interpreted by the player.

Every misperception about opponent's benevolence affects opponent's benevolent utility function interpreted by a player.

Proof:

Let $\alpha(\neq 0)$ be an arbitrary misperception about $j's$ benevolence of player $i (i \neq j)$ and $\overline{u}_j^i$ is the benevolent utility function of player j interpreted player i such that player i has a misperception about $j's$ benevolence α and u_j is a benevolent utility function of player j such that there are no misperceptions.

If there is no effect of misperception about $j's$ benevolence for a $j's$ benevolent utility function interpreted player i, then $\overline{u}_j^i$ is equivalent to u_j. Therefore there exists a positive real number a and a real number b such that for all $s \in S$

$$\overline{u}_j^i(s) = a u_j(s) + b$$

$$= \frac{a}{1-k_i k_j}(p_j(s)+k_j p_i(s)) + b .$$

Since

$$\overline{u}_j^i(s) = \frac{1}{1-k_i(k_j+\alpha)}(p_j(s)+(k_j+\alpha)p_i(s)) ,$$

It is true that

$$\frac{1}{1-k_i(k_j+\alpha)}=\frac{a}{1-k_ik_j}$$

and,

$$\frac{(k_j+\alpha)}{1-k_i(k_j+\alpha)}=\frac{ak_j}{1-k_ik_j},$$

$$b=0.$$

Hence $\alpha=0$ and $a=1, b=0$. It is contradicts that $\alpha\neq 0$. The proof of this proposition is complete.

Above two propositions suggest that any misperception of a player about opponent's benevolence affect opponent's benevolent utility function interpreted the player while it does not affect own benevolent utility function of the player.

Discussion about characteristics misperception about opponent's private subutility

Proposition 3. The effect of misperception about opponent's private subutility for own benevolent utility function

Every misperception about opponent's private subutility affects own benevolent utility function.

Proof:

Let $\hat{u}_i$ is the benevolent utility function of player i such that player i has a misperception about $j's(j\neq i)$ private subutility and u_i is a benevolent utility function of player i such that there are no misperceptions and $\hat{p}_j^i$ is an $j's$ private subutility function that exists a misperception about $j's$ private subutility.

If there is a misperception about opponent's private subutility such that it does not affect own benevolent utility function; that is, $\hat{u}_i$ is equivalent to u_i, then there exists a positive real number a and a real number b such that for all $s\in S$

$$\hat{u}_i(s)=au_i(s)+b$$

$$=\frac{a}{1-k_ik_j}(p_i(s)+k_ip_j(s))+b.$$

By definition of benevolent utility function,

$$\hat{u}_i(s) = \frac{1}{1 - k_i k_j}(p_i(s) + k_i \hat{p}^i_j(s)).$$

By solving above two identical equations, we obtain that $a = 1$ and $\hat{p}^i_j(s) = p_j(s) + \frac{(1 - k_i k_j)}{k_i} b$.

Since $\frac{(1 - k_i k_j)}{k_i} b$ is a fixed number for s, there exist a positive affine transformation from p_i to $\hat{p}^i_j$ and it contradicts that $\hat{p}^i_j$ is a $j's$ private subutility function that exists a misperception about $j's$ private subutility. Therefore for all misperception about opponents private subutility effect owns benevolent utility function and the proof of this proposition is complete.

Proposition 4. The effect of misperception about opponent's private subutility for opponent's benevolent utility function interpreted by a player.

Every misperception about opponent's private subutility affects opponent's benevolent utility function interpreted by a player.

Proof:

Let $\hat{u}^i_j$ be the benevolent utility function of player j interpreted by a player i such that player i has a misperception about $j's(j \neq i)$ private subutility and u_j be a benevolent utility function of player i such that player i has no misperceptions about opponent's benevolence and private subutility and $\hat{p}^i_j$ be an $j's$ private subutility function such that the player i has a misperception about $j's$ private subutility.

If there exists a misperception of the player i about $j's$ private subutility such that it does not affect $j's$ benevolent utility function; that is, there exists $j's$ benevolent utility function $\hat{u}^i_j$ interpreted by player i that is equivalent to u_j, then there exists a positive real number a and a real number b such that for all $s \in S$

$$\hat{u}^i_j(s) = a u_j(s) + b$$

$$= \frac{a}{1 - k_i k_j}(p_j(s) + k_j p_i(s)) + b.$$

By definition of benevolent utility function,

$$\hat{u}_j^i(s) = \frac{1}{1-k_ik_j}(\hat{p}_j^i(s) + k_j p_i(s)) .$$

By solving above two identical equations, we obtain that $a = 1$ and $\hat{p}_j^i(s) = p_j(s) + \frac{(1-k_ik_j)}{k_j}b$.

Since $\frac{(1-k_ik_j)}{k_j}b$ is a fixed number for every $s \in S$, there exist a positive affine transformation from p_j to $\hat{p}_j^i$ and it contradicts that $\hat{p}_j^i$ is a $j's$ private subutility function interpreted by player i such that player i has a misperception about $j's$ private subutility. Therefore every misperception of a player about opponent's private subutility affect opponent's benevolent utility function interpreted by the player and the proof of this proposition is complete.

Above two propositions suggest that any misperception of a player about opponent's private subutility affect not only opponent's benevolent utility function interpreted the player, but also own benevolent utility function of the player. Since any misperception of a player about opponent's private subutility affects own benevolent utility function of the player while any misperception of a player about opponent's benevolent does not affect own benevolent utility function of the player, we can derive the following proposition.

The difference between misperception about opponent's benevolence and about opponent's private subutility

Proposition 5. The difference between misperception about opponent's benevolence and about opponent's private subutility.

Every misperception about opponent's benevolence is different to every misperception about opponent's private subutility.

Proof:

Let $\overline{u}_i$ be a benevolent utility function of player i such that player i has a misperception about $j's(j \neq i)$ benevolence and $\hat{u}_i$ be a benevolent utility function of player i such that player i has a misperception about $j's$ private subutility and u_i be a benevolent utility function of player i such that player i has no misperception. We prove that if there exists a misperception about $j's$ benevolence and about $j's$ private subutility such that $\overline{u}_i$ is equivalent to $\hat{u}_i$, then it is contradiction. Suppose that there exist a misperception about $j's$ benevolence and $j's$ misperception about private subutility such that $\overline{u}_i$ is equivalent to $\hat{u}_i$. Since for every misperception about $j's$ benevolence $\overline{u}_i$ is equivalent to u_i and equivalent relation on the set of all benevolent utility function satisfy transitive, u_i is equivalent to $\hat{u}_i$ and it contradicts that every misperception about opponent's private subutility affects own benevolent utility function (Proposition3).Therefore every misperception about opponent's benevolence is different to every misperception about opponent's private subutility and the proof of this proposition is complete.

EXAMPLE

We see an example that suggests difference between misperception about opponent's benevolence and about opponent's private subutility. We consider a situation such that a man, say player 1, deals with a swindler, say player 2 ; the set of players N equal to $\{1,2\}$.An action A of player 1 means that player 1 trusts player 2 and action B of player 1 means that player 1 does not trust player 2 . An action X of player 2 means that player 2 tricks player 1 and action Y of player 2 means that player 2 does not trick player 1; that is, The set of all actions of Bob S_1 equal to $\{A, B\}$ and the set of all actions of Mary S_2 equal to $\{X, Y\}$.Figure1 suggests the payoff matrix in private subutility level of the above situation. Real payoff matrix suggests that each player has no misperception about opponent's information.

1/2	X	Y
A	1,4	4,3
B	3,2	2,1

Fig. 1 Real payoff matrix in private subutility level

Let player 1 be good-natured to a fault; that is, his benevolence $k_1 = 0.5$ and player 1 be selfish; that is, his benevolence $k_2 = 0.01$ .Then we can suggest the payoff matrix in benevolence utility level as Figure2.

1/2	X	Y
A	3.02,4.03	5.53,3.06
B	4.02,2.04	2.51,1.03

Fig. 2 Real payoff matrix in benevolent utility level

- Player 1 has misperception about benevolence of player2

If player 1 has misperception about benevolence of player2, $k_2^1 = 0.5$ then payoff matrix in benevolent utility level interpreted by player 1 is represented by Figure3.

1/2	X	Y
A	4.00,6.00	7.33,6.67
B	5.33,4,67	3.33,2,67

Fig. 3 Payoff matrix in benevolent utility level interpreted by player 1

Different Types of Misperception regarding Benevolence

Here benevolent utility of player1 in Fig.3 corresponds to one in Fig.2 by affine transformation while benevolent utility of player2 interpreted by player1 in Fig.3 does not corresponds to one in Fig.2.This example suggest that misperception of a player about opponent's benevolence does not affect his benevolent utility.

- Player1 has misperception about private subutility of player2

If player1 has misperception about private subutility of player2, the case is suggested by Fig.4, then game in benevolent utility level interpreted by player1is represented by Figure 5.

1/2	X	Y
A	1,1	4,3
B	3,2	2,4

Fig. 4 Payoff matrix in private subutility level interpreted by player 1

1/2	X	Y
A	1.51,1.02	5.53,3.06
B	4.02,2.04	4.02,4.04

Fig. 5 Payoff matrix in benevolent utility level interpreted by player 1

Here benevolent utility of player1 in Fig.5 does not correspond to one in Fig.2 by affine transformation. This example suggest that misperception of a player about opponent's private subutility affects his benevolent utility and therefore his best response for opponent's action changes from the case that he has no misperception about opponent's benevolence.

REFERENCES

P. G.Bennett and M.R.Dando (1979).Complex Strategic Analysis: A Hypergame Study of the Fall of France, J.Opl Res.Soc,30(1),23-32

P.,G. Bennett (1980),.Bidders and dispenser: manipulative hypergames in a multinational context. European Journal of Operational Research, 4(5),293-306.

Bergstrom, T. C. (1989). Puzzles: Love and Spaghetti, The Opportunity Cost of Virtue, *Journal of Economic Perspectives*, 3(2), 165-173

Bergstrom, T. C. (1999). Systems of Benevolent Utility Functions, *Journal of Public Economic Theory*,1(1), 71-100

von Neumann, J. and Morgenstern, O. (1953). *Games and Economic Behavior*, Princeton

Shutaro. K(2008). Two player hypergames with benevolent players: Misperception about benevolence vs. misperception about monetary payoff, *In the Proceedings of Annual Conference of Japan Society for Management Information*