December 1997

Truth-telling Problems in Virtual Society: Higher Order Hypergame

Kyoichi Kijima
Tokyo Institute of Technology

Truth-telling Problems in Virtual Society: Higher Order Hypergame

Kyoichi Kijima
Dept. of Value and Decision Science
Graduate School of Decision Science and Technology
Tokyo Institute of Technology

This research was in part supported by The Japan Securities Scholarship Foundation

Key words: virtual (or network) society, conflict, diversified preference, truth-telling, hypergame, game theory

Executive Summary

In this paper we will argue truth-telling problems in virtual (or network) society by employing what we call hypergame framework.

In network societies, the people involved may not have incentive to behave seriously particularly because there is no substantial mechanism to impose sanctions to those who report the false. For instance, according to recent newspaper, a thirty-eight-years-old Polish man was arrested because he pretended a twenty-five-years-old female doctor in the internet. Through the internet she (he?) announced that she was looking for a nice boy friend. She copied a photo of a beautiful girl from a fashion magazine as her own portrait in the self introduction. By claiming that the hospital she worked for was trying to raise money for sick children, she cheated those who got interested in her out of their money. This is just an instance and to our surprise we can see similar phenomena quite often, though most of them are not laid before the public.

The episode above typically suggests that it is critically important to motivate the people involved to reveal their private information honestly for successful realization of virtual society supported by information technology. This motivation and realization problem is called a truth-telling problem. Truth-telling constitutes one of the fundamentals for success of a network society supported by electric commerce and internet communication.

To investigate the truth-telling problems we first develop a new idea called higher order hypergame and apply it to them. We began with formulating relationship between a sender and a receiver of message in the network society in the framework of simple hypergame. We, then, introduced a new idea called higher order hypergame to describe the interpretation by the players. By this analysis we pointed out that in the network society it is very reasonable for the senders to prepare that the messages are ignored. We also clarified that the sender has some initiative over the receiver and the receiver is inevitably on the defensive.

Truth-telling Problems in Virtual Society: Higher Order Hypergame Approach*

Kyoichi Kijima
Tokyo Institute of Technology

Abstract

In this paper we will argue truth-telling problems in virtual (or network) society by employing hypergame framework. In network societies, the people involved may not have incentive to behave seriously particularly because there is no substantial mechanism to impose sanctions to those who report the false. However, for successful realization of virtual society supported by information technology including electric commerce and internet communication, it is critically important to motivate the people to reveal their private information honestly. This motivation and realization problem is called truth-telling problem. To investigate the truth-telling problems we first introduce a new idea called higher order hypergame and apply it to them.

key words: virtual (or network) society, conflict, diversified preference, truth-telling, hypergame, game theory

1 Introduction

The purpose of this paper is to formally discuss truth-telling problems in network (or virtual) societies by employing a new concept called higher order hypergame framework.

In network societies, the people involved may not have incentive to behave seriously particularly because there is no substantial mechanism to impose sanctions to those who report the false. For instance, according to recent newspaper [Asahi Newspaper, 1996], a thirty-eight-years-old Polish man was arrested because he pretended a twenty-five-years-old female doctor in the internet. Through the internet she (he?) announced that she was looking for a nice boy friend. She copied a photo of a beautiful girl from a fashion magazine as her own portrait in the self introduction. By claiming that the hospital she worked for was trying to raise money for sick children, she cheated those who got interested in her out of their money up to twenty six thousand US dollars in total. This is just an instance and to our surprise we can see similar phenomena quite often, though most of them are not laid before the public.

The episode above typically suggests that it is critically important to motivate the people involved to reveal their private information honestly for successful realization of virtual society

*This research was in part supported by The Japan Securities Scholarship Foundation.

supported by information technology. This motivation and realization problem is called a truth-telling problem. Truth-telling constitutes one of the fundamentals for success of a network society supported by electric commerce and internet communication.

If a situation involves a variety of people with essentially diversified and conflicting preference on the outcomes, it is often referred to as a *soft* problematic situation in the fields of systems science and Operational Research [Checkland, 1981; Rosenhead, 1989]. A network society is a typical soft problematic situation. In order to explore the essential features of a network society mentioned above we need to focus on subjective perception and interpretation by the people participating in them.

Game theory is probably the only theoretical framework that can accommodate such conflicting situations. The theory usually analyzes rational behavior of players under the assumption that they can observe the problematic situation objectively even if information about it is not sufficient. For example, in a game with incomplete information [Harsanyi, 1967] it is assumed that at least one player uncertain about another players payoff function. However, though such a game situation inherently involves dynamic issues in the sense that the existence of private information leads naturally to attempts by informed parties to communicate (or mislead) and to attempts by informed parties to learn and respond, in a game with incomplete information even in dynamic cases rationality concept is usually fixed throughout consideration, *e.g.*, perfect Bayesian equilibrium. It does not necessarily reflect reality.

In order to overcome the limitations of the analytical approach and to describe soft problematic situations more realistically by focusing on irrational behavior and subjective perception by players, several lines of research have been conducted. These include hypergame analysis [Bennett, 1980; Bennett et al. 1979; Bennett et al. 1980; Bennett, et al. 1989], metagame theory [Howard, 1987; Howard, 1989] and conflict analysis [Fraser et al. 1984]. The interests of these approaches are quite near to ours in the sense that they all try to deal with what we call soft aspects of decision makers [Wang et al. 1988].

The fundamental idea of hypergame analysis is that it is better to model interactive decision situations not as a single game, but as a collection of subjective games. The basic model of hypergame analysis assumes, among others, that decision makers may conceptualize problems in a similar manner to that of game theory, but they see different games. The basic hypergame framework has been extended mainly in two ways; one is to allow for more radical differences in players' perceptions, while the other is to consider systems of linked interactions, rather than just isolated hypergames [Bennett et al. 1989]. We call the latter symbiotic hypergame [Kijima, 1996].

Metagame approach tries to generalize Von Neumann and Morgenstern's recommendation that to study a given two person zero-sum game we should study the minorant and majorant games based on it, and to apply them recursively. These are games in which player 1 (and also, player 2) is able to predict the other player's strategy. Conflict analysis presents metagame approach in a more algorithmic way with further development [Fraser et al. 1984; Fraser et al. 1991]. The metagame approach has been also extended to soft game theory [Howard, 1990] by introducing credibility and emotions. More recently Howard advocated drama theory [Howard et al. 1993; Howard, 1994].

However, the basic attitudes of the drama theory and the symbiotic hypergame analysis are quite different: Drama theory abandons the rationality assumption and, instead, claims that emotions are very often strategically generated by the tension between initial preferences and the preferences it would 'pay' one to have [Howard et al. 1993]. By introducing emotion, preference change and irrationality, drama theory tries to overcome the various kinds of paradoxes of game theory.

On the other hand, the symbiotic hypergame analysis assumes that though optimization is not meaningfully available when dealing with soft problematic situations, rationality rather than emotion and feeling plays a crucial role for describing players' behavior.

In this paper we follow the line basically similar to the symbiotic hypergame analysis. That is, in the following sections we will explain the concept of simple and symbiotic hypergame and then expand it further to higher order hypergame. Then, we will discuss the truth-telling problems by using it.

2 Simple Hypergame in Network Society

Let $N = \{1, 2, \ldots, n\}$ be a set of citizens (or netizens = *net*work cit*izens*) in a network society. Assume p and q be in N and suppose p sends a message to q. p is called a sender of the message while q is referred to as a receiver of it. The message may be true or false while the receiver may believe it or not, especially when they believe they come across at the first time. Which of those behavior, or strategies, they chooses entirely depends on how they perceive the network society and how they commit themselves to it. We cannot fully explain the behavior in this situation simply by rationality but need to take into account their emotional stance towards the network society.

We will formulate this situation by simple hypergame. A simple hypergame explicitly presumes that each player is supposed to perceive the problematic situation independently. Since simple hypergame is essentially based on a two person non-cooperative game, we begin with the definition of it.

A two person non-cooperative game by p and q is defined by

Definition 1 *A two person non-cooperative game with players p and q is a quadruple $G = (S_p, S_q, \geq_p, \geq_q)$.*

For each $i \in \{p, q\}$ S_i denotes a set of strategies of player i, while $\geq_i$ is i's preference ordering on $S_p \times S_q$. For any (s_p, s_q) and $(s'_p, s'_q) \in S_p \times S_q$, $(s_p, s_q) \geq_i (s'_p, s'_q)$ means that i prefers (s_p, s_q) to (s'_p, s'_q) or is indifferent between (s_p, s_q) and (s'_p, s'_q). We assume that S_p and S_q are finite while $\geq_p$ and $\geq_q$ are linear orderings and can be represented by some ordinal utility functions.

The most well-known rationality concept for a two person non-cooperative game is Nash equilibrium. It is given by:

Definition 2 *For a two person non-cooperative game $G = (S_p, S_q, \geq_p, \geq_q)$ $s^* = (s^*_p, s^*_q) \in S_p \times S_q$ is called Nash equilibrium of G iff we have*

$$(\forall s_p \in S_p)((s^*_p, s^*_q) \geq_p (s_p, s^*_q))$$

and

$$(\forall s_q \in S_q)((s_p^*, s_q^*) \geq_q (s_p^*, s_q))$$

hold.

The definition implies that if $s^* = (s_p^*, s_q^*)$ is Nash equilibrium then there is no incentive for either of the players to change their strategy as long as the other does not change its strategy.

Now a simple hypergame played by p and q is formally defined by

Definition 3 *A simple hypergame of players p and q is a pair of* (G_p, G_q)*, where* $G_p = (S_p, S_{qp}, \geq_p, \geq_{qp})$ *is a game that p believes both sides perceive while* $G_q = (S_{pq}, S_q, \geq_{pq}, \geq_q)$ *is a game that q believes both sides perceive.*

In G_p, S_p denotes a set of strategies for p while S_{qp} denotes a set of strategies which the player p assumes that q can prepare. That is, p perceives that q's strategy set is S_{qp}. $\geq_p$ denotes p's preference ordering on $S_p \times S_{qp}$ while $\geq_{qp}$ is a preference ordering on $S_p \times S_{qp}$ which p assumes that q holds. That is, p perceives that q's preference ordering is $\geq_{qp}$. We similarly define G_q. It is quite natural to assume that $S_{ii} = S_i$ and $\geq_{ii}=\geq_i$ for $i = p$ and q. We assume all the preference orderings are linear orderings and can be represented by some ordinal utility functions.

One of the most natural ways to describe rational behavior of p and q in a simple hypergame is obtained by modifying the concept of Nash equilibrium for G_p and G_q in the following manner.

Definition 4 *For* $G_p = (S_p, S_{qp}, \geq_p, \geq_{qp})$ (s_p^*, s_{qp}^*) *is called Nash equilibrium of* G_p *iff we have*

$$(\forall s_p \in S_p)((s_p^*, s_{qp}^*) \geq_p (s_p, s_{qp}^*))$$

and

$$(\forall s_{qp} \in S_{qp})((s_p^*, s_{qp}^*) \geq_{qp} (s_p^*, s_{qp}))$$

hold.

The definition claims the following: If (s_p^*, s_{qp}^*) is Nash equilibrium of G_p, then p believes that there is no incentive for either of the players to change their strategy as long as the other does not change its strategy. By applying symmetric arguments to G_q we can define Nash equilibrium of G_q.

Now let us apply the hypergame framework above to our situation. We focus on the sender p and the receiver q and represent the interactive activity by a simple hypergame (G_q, G_q) (Refer to Tables 1 and 2). In G_p, we assume p possesses three strategies, namely, T, OM and M. T means telling the truth while OM indicates the message contains some false but p has no malice and just wants to enjoy the false. M implies the message is filled with malice. Furthermore, we presume that p believes q has two alternatives B and NB, where B denotes a strategy to believe the message while NB does a strategy not to believe the message. The numbers in the matrix G_p show the utilities p believes. For example, p believes that the utility from (T, B) for himself is 3 while that for q is 6. The matrix G_p indicates that the sender p is a rather mischievous person since p has the highest preference towards OM and M. That is, for p (OM, B) is the most preferable situation.

Table 1 G_p of Symbiotic game in Network Society

Table 2 G_q of Symbiotic game in Network Society

On the other hand, G_q shows that q believes that ps message is either acceptable (A) or unacceptable(UA). G_q also implies q has two choices, *i.e.*, to respond to the message (R) or to ignore it (I). The numbers in the matrix show the utilities q believes.

The matrix G_q represents that the receiver q is a somewhat serious person since q likes to receive acceptable message and to believe it the most. As a matter of fact, for the receiver q (A, R) is the most preferable situation.

Then we can easily find out that Nash equilibria of G_p are (M, NB) and (OM, B) while that of G_q is (UA, I).

3 Symbiotic Hypergame in Network Society

As the time goes on, p and q may begin to interpret other's identification of the situation after several exchange of messages.

We describe this as a *symbiotic* hypergame. In the symbiotic hypergame, the players understand that they are concerned with a common situation but they allow for different ways of identifying the situation. The symbiotic hypergame consists of a simple hypergame and interpretation functions between them. Interpretation functions represent how each player interprets the other's game.

A formal definition of symbiotic hypergame is given by:

Definition 5 *A symbiotic hypergame with players p and q is a pair* $((G_p, f), (G_q, g))$, *where we have* $G_p = (S_p, S_{qp}, \geq_p, \geq_{qp})$ *and* $f: S_q \rightarrow S_{qp}$, *while* $G_q = (S_{pq}, S_q, \geq_{pq}, \geq_q)$ *and* $g: S_p \rightarrow S_{pq}$ *hold.*

In $((G_p, f), (G_q, g))$, (G_p, G_q) is a simple hypergame defined by Definition 3. The function f represents how p interprets the set S_q of strategies of q. Similarly, g formulates how q interprets the set S_p of strategies of p. has a symmetric interpretation. We refer to f and g as interpretation functions of p and q, respectively.

We may consider several ways of defining overall rationality for dealing with (G_p, f) and (G_q, g), all of which should depend on f and g as well as on G_p and G_q.

The following is a natural and straightforward way to define such rationality.

Definition 6 *Let* $((G_p, f), (G_q, g))$ *be a symbiotic hypergame where* $G_p = (S_p, S_{qp}, \geq_p, \geq_{qp})$ *and* $G_q = (S_{pq}, S_q, \geq_{pq}, \geq_q)$, *while* $f: S_q \rightarrow S_{qp}$ *and* $g: S_p \rightarrow S_{pq}$. (s_p^*, s_q^*) *is called symbiotic Nash equilibrium of* $((G_p, f), (G_q, g))$ *iff*

$$(s_p^*, f(s_q^*)) \text{ is Nash equlibrium of } G_p$$

and

$$(g(s_p^*), s_q^*) \text{ is Nash equlibrium of } G_q.$$

Let $(s_p^*, s_q^*) \in S_p \times S_q$ be symbiotic Nash equilibrium of $((G_p, f), (G_q, g))$. Then, player p, who perceives it as $(s_p^*, f(s_q^*)) \in S_p \times S_{qp}$, has no incentive to change its strategy from s_p^* as long as p believes that q will not change its strategy from $f(s_q^*)$. Since a similar argument holds for q, a symbiotic Nash equilibrium can be seen as a natural extension of the Nash equilibrium by taking into account interpretation by the players.

In our situation the following f_1 and f_2 are meaningful candidates of the interpretation functions for p while g_1 and g_2 are those for q:

$$f_1(R) = B, f_1(I) = NB; \quad f_2(R) = NB, f_2(I) = B$$

$$g_1(T) = g_1(OM) = A, g_1(M) = UA; \quad g_2(T) = A, g_2(OM) = g_2(M) = UA$$

f_1 shows p believes the response from q is a sign of q's belief of the message while f_2 implies p assumes that the response is understood as a sign of no belief by q. On the other hand, g_1 indicates q assumes that both T and OM are acceptable while g_2 represents q believes that only T is acceptable.

Then we have four types of symbiotic hypergames. One of them is $((G_p, f_1), (G_q, g_1))$. In this game (M, I) is symbiotic Nash equilibrium since we have

$$(M, f_1(I)) = (M, NB) \text{ is Nash equlibrium of } G_p$$

and

$$(g_1(M), I) = (UA, I) \text{ is Nash equlibrium of } G_q.$$

We can apply similar arguments to other three cases, *i.e.*, $((G_p, f_1), (G_q, g_2))$, $((G_p, f_2), (G_q, g_1))$ and $((G_p, f_2), (G_q, g_2))$.

4 Higher order Hypergame in Network Society

Our main concern of this paper is with which of these four symbiotic hypergames are expected to happen with the highest likelihood. In order to investigate the problem, we define a higher order hypergame game H (Refer to Table 3). It is essentially a two person non-cooperative game played by p and q. However, we should notice that strategies of H, we denote them by I_p and I_q, consist of interpretation functions between the symbiotic hypergames, *i.e.*, $I_p = \{f_1, f_2\}$ and $I_q = \{g_1, g_2\}$, respectively. In H the cell corresponding to (f_1, g_1), for example, indicates the set of symbiotic Nash equilibria of $((G_p, f_1), (G_q, , g_1))$, which is $\{(M, I)\}$ from the discussion in Section 3. We can similarly calculate solution sets for each of the remaining three cases.

By analyzing the higher order hypergame H we can see that the interpretation from S_q into S_{qp} by the sender p plays a critical role for determining the stability of the game. That is, if the sender believes that ignorance by the receiver is simply a sign of the receiver's belief (*i.e.*, f_2 is adopted), then the situation becomes unstable. On the other hand, as far as p employes f_1, the situation always has equilibrium, no matter how the receiver q behaves. That is, the belief of the receiver does not give essential influence on the stability.

Table 3 Higher-order Hypergame Game H

The results of our analysis are rather pessimistic to the truth-telling problems in dual sense. First, we have only (M, I) and (OM, I) as equilibria. It follows that in the network society it is quite reasonable for the sender to prepare that the messages are ignored by the reciever.

Second, the higher order hypergame H suggests that the relationship between the sender and the receiver is not symmetric. f_1 generates a stable situation while f_2 leads to an unstable one whichever interpretation functions the reciever adopts. On the other hand, the results led by g_1 and g_2 are quite sensitive to f_1 and f_2, *i.e.*, under f_1 the both are indifferent while under f_2 g_1 and g_2 lead to significant difference. These observations imply that the sender has some initiative over the receiver while the receiver is inevitably on the defensive.

We can use the framework for what-if analysis as well. As far as the payoff matrices are fixed as Tables 1 and 2, the situation is rather pessimistic, as discussed above. It implies that to emerge other equilibria other than (M, I) and (OM, I), the belief of the players should be changed. For example, if p changes his/her mind to become a little bit serious by rewriting the associated payoff matrix, then another situation may become an equilibrium state.

5 Conclusions and Further Research

In this paper we formally discussed what we call truth-telling problems in virtual (or network) society by extending hypergame framework.

Since in the network society the people involved may not have incentive to behave seriously, it is critically important to motivate them to reveal their private information honestly. We tackled this truth-telling problem since we believe the resolution is crucial for successful realization of virtual society supported by information technology including electric commerce and internet communication.

We began with formulating relationship between a sender and a receiver of message in the network society in the framework of simple hypergame. We, then, introduced a new idea called higher order hypergame to describe the interpretation by the players. By this analysis we pointed out that in the network society it is very reasonable for the senders to prepare that the messages are ignored. We also clarified that the sender has some initiative over the receiver and the receiver is inevitably on the defensive.

The research presented here is just a first step toward the truth-telling problems in the network society. There still remain quite a lot of topics as targets of future research. One of them, which we are now undertaking, is as follows: Suppose that all the receivers are serious while the senders can be classified into to categories; the serious and the mischievous, though in this paper we focused only on two players and the interactions between them. Furthermore, assume that each sender can change their type according to the payoff obtained from the interactions with the receivers, then which type of the two survive after a certain time period? By investigating this question through simulation, we believe to dig out an interesting and new findings suggesting the future of network society.

Table 1: G_p of Symbiotic game in Network Society

S_p \ S_{qp}	Do believe (B)	Do not believe (NB)
Tell truth (T)	3 6	1 5
Out of malice (OM)	6 4	2 3
Malice (M)	5 1	4 2

Table 2: G_q of Symbiotic game in Network Society

S_{pq} \ S_q	Do response (R)	Ignore (I)
Acceptable (A)	3 4	1 3
Unacceptable (UA)	4 1	2 2

Table 3: Higher- order Hypergame Game H

I_p \ I_q	g_1	g_2
f_1	$\{(M, I)\}$	$\{(M, I)\}$
f_2	ϕ	$\{(OM, I)\}$

References

[1] Asahi Newspaper (in Japanese), from AP Network, Oct. 20, 1996

[2] Bennett, P.G., Hypergames: Developing a Model of Conflict, *Futures*, 12(6)(1980) 489-507

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

[4] Bennett, P.G., Dando, M.R. and Sharp, R.G., Using Hypergames to Model Difficult Social Issues: An Approach to the Case of Soccer Hooliganism, *J. Opl. Res. Soc.*, 31(7) (1980) 621-635

[5] Bennett, P.G., Cropper, S. and Huxham, C., Modelling Interactive Decisions: The Hypergame Focus, in *Rational Analysis for a Problematic World*, John Wiley and Sons, Chichester, (1989)

[6] Checkland, P. B., *Systems Thinking, Systems Practice*, John Wiley, Chichester, (1981)

[7] Checkland, P. B., *Soft Systems Methodology in Action*, John Wiley, Chichester, (1990)

[8] Fraser, N.M. and Hipel, K.W., *Conflict Analysis: Models and Resolutions*, North-Holland, Amsterdam, (1984)

[9] Fraser, N.M., Wang, M. and Hipel, K.W., Hypergame theory in 2-person conflicts, *Information and Decision Technology*, 16 (1991) 301-319

[10] Gibbons, R., *A Primer in Game Theory*, Harvester Wheatsheaf, London, (1992)

[11] Harsanyi, J., Games with Incomplete Information played by Bayesian Players, Parts I, II and III, *Management Science*, 14 (1967) 159-182, 320-334, 486-502

[12] Howard, N., The Present and Future of Metagame Analysis, *European Journal of Operational Research*, 32 (1987) 1-25

[13] Howard, N., The Manager as Politician and General: the Metagame Approach to Analysing Cooperation and Conflict, in *Rational Analysis for a Problematic World*, John Wiley and Sons, Chichester, (1989)

[14] Howard, N, Soft Game Theory, *Information and Decision Technologies*, 16 (3), (1990) 215-227

[15] Howard, N., Drama Theory and its Relation to Game Theory: Part One, *Group Decision and Negotiation*, 3, (1994) 187-206

[16] Howard, N., Bennett, P., Bryant, J. and Bradley, M., Manifesto for a Theory of Drama and Irrational Choice, *Systems Practice*, 6(4), (1993) pp429-434

[17] Kijima, K, Decision Making based on Subjective Evaluations of Problem Situation (in Japanese), *T. IEE, Japan* 111-C(3)(1991) 98-106

[18] Kijima, K, Decision Making based on Subjective Evaluations of Problem Situation (in Japanese), *T. IEE, Japan* 111-C(3)(1991) 98-106

[19] Kijima, K., Intelligent Poly-agent Learning Model, *Proceedings of Workshop for Cooperation between Japan and the United Kingdom on Soft Science and Technology*, Osaka (1994)

[20] Kijima, K., Intelligent Poly-agent Learning Model and its Application, *Information and Systems Engineering*, 2(1996) 47-61

[21] Neches, R. et al., Enabling Technology for Knowledge Sharing, *AI Magazine*, 12(3) (1991) 36-56

[22] Neches, R. , Knowledge Sharing Effort, *Research Report*, in /pub/knowledge-sharing/papers, hpp.stanford.edu, (1993)

[23] Rosenhead, J.,(ed.) *Rational Analysis for Problematic World*, (John Wiley, Chichester, 1989)

[24] Van Gigch, J. P., *System Design Modeling and Metamodeling*, Plenum Press, New York, (1991)

[25] Wang, M., Hipe, K.W. and Frase, N. M., Modeling Misperceptions in Games, *Journal of Behavior Science*, 33, (1988) 207-223

[26] Wang, M. and Hipe, K.W., Misperception and Bargaining in the Persian Gulf war, *Control and Cybernetics*, 10(2), (1992) 1-26