人　工　知　能 99−18
（1995. 3. 7）

コミュニケーションのゲーム理論的分析

橋田 浩一
〒305 つくば市梅園 1-1-4
電子技術総合研究所
自然言語研究室

長尾 確
〒141 品川区東五反田 3-14-13
高輪ミューズビル
ソニーコンピュータサイエンス研究所

宮田 高志
〒113 文京区本郷 7-3-1
東京大学
理学部情報科学科

コミュニケーションにおいては、メッセージの送り手も受け手も、送り手の意図した意味内容を受け手が正しく理解することを望むから、コミュニケーションは必然的に協調的な作業となる。本稿では、利己的に効用最大化を図る行為者がメッセージとその内容との間の最適な対応関係の共有を行ない、コミュニケーションの頑健性を高める様子をゲーム理論的に解析する。2 個の行為者の間での 1 個のメッセージの送受信は、一般に、n 人非協力ゲームとして定式化でき、そのゲームにおける Nash 均衡解のうち両方の行為者の何らかの共同効用の期待値を最大にするものが最適な対応関係を与える。自然言語における照応現象の一部はこの理論によって一般的に説明できる。

A Game-Theoretic Account of Communication

HASIDA Kôiti
Electrotechnical Lab.
1-1-4 Umezono, Tukuba
Ibaraki 305, Japan
hasida@etl.go.jp

NAGAO Katashi
Sony Computer Science Lab. Inc.
3-14-13 Higashi-gotanda,
Shinagawa-ku, Tokyo 141, Japan
nagao@csl.sony.co.jp

MIYATA Takashi
University of Tokyo
7-3-1 Hongo, Bunkyo-ku,
Tokyo 113, Japan
mya-u@is.s.u-tokyo.ac.jp

Communication inherently tends to be cooperative. Not only the sender of a message intends to communicate, but also the receiver is normally motivated to know the semantic content of the message intended by the sender. The present paper accounts for how autonomous agents as selfish utility maximizers naturally cooperate in reaching a common optimal mapping between messages and their contents, raising the robustness of communication. An occasion of communication between two agents can be generally formalized as a non-cooperative n-person game, and the optimal mapping is shown to be obtained as a Nash equilibrium which maximizes the agents' expected utility over all the possible occasions of communication. Some regularities in natural language anaphora are demonstrated to follow from this account.

1 Introduction

In communication, both the sender and the receiver of a message are motivated to communicate. That is, not only the sender wants to communicate a semantic content by sending a message, but also the receiver normally wants to interpret that message correctly. Communication is cooperative in this sense. More technically speaking, illocutionary act in the sense of Austin (1962) is performed by cooperation of the communicating agents.

The present paper discusses how such cooperation may contribute the reliability of communication. As is the case with natural language, a message may be ambiguous, potentially denoting several different semantic contents. This ambiguity accounts for efficiency of language (Barwise & Perry, 1983), because it allows the same message may be used differently in different occasions. But of course the problem here is how to disambiguate. It is this disambiguation that the rest of the paper concerns. Such a study should be useful not just for explaining communication among humans and other living beings, but also for designing artificial agents communicating with each other or with humans.

Below we are going to formalize a case of communication as a non-cooperative game, and investigate what kind of protocols for disambiguation should be stably agreed upon among the communicating agents. A key assumption here is that the agents recognize each other as selfish utility maximizers, and that this mutual recognition is exploited in mutually recognizing plans. We will mainly investigate Nash equilibria of the game to characterize the optimal protocol, unlike in the previous work (Etzioni, 1991; Durfee, Gmytrasiewicz, & Rosenschein, 1994; Nagao, Hasida, & Miyata, 1993) on communication by utility maximizers.

2 Communication as Game

The range of ambiguity of each message may vary from one context to another. That is, it may be the case that a message is ambiguous across a set of contents in a context, whereas it is ambiguous across a different set of contents in a different context. In English, for instance, 'you' may refer to Tom or Bill in one occasion but to Mary or Kim in another. The problem here is how to resolve this ambiguity.

Let us consider a very simple example of how agents are able to 'cooperate' in disambiguation. Suppose that, as shown in Figure 1, message m_1 can mean either content c_1 or c_2,

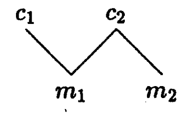

Figure 1: A simple pattern of possible encoding of contents by messages.

whereas m_2 means c_2 only. In this case, the sender as a rational agent will use m_2 to communicate c_2, because using m_1 instead would make the receiver face the ambiguity between the two interpretations corresponding to c_1 and c_2, thus lowering the probability of correct interpretation. On the other hand, the receiver as a similar rational agent will interpret m_1 as meaning c_1, by recognizing that, for the above reason, the sender would have used m_2 in place of m_1 if she had wanted to communicate c_2.

Here let us formalize a notion of communication. An entire system of communication is defined as a set of *contexts* of communication. A context Φ of communication is a 7-tuple $[S, R, C, P, M, E, U]$. S is the *sender*, R the *receiver*, C the set of semantic contents, P the probability assignment over C, M the set of messages, and E the *encoding*. We assume the following regarding them.

(1) a. C and M are finite sets.
b. P is a function from C to non-negative real numbers not exceeding 1.[1]

[1] We need not assume $\sum_{c \in C} P(c) = 1$, because S might want to communicate several semantic contents at one time, as discussed in Section 2.2.

c. E is a binary relation between C and M. That is, $E \subseteq C \times M$.

We say m *encodes* c when $\langle c, m\rangle \in E$. E specifies the patterns of ambiguity of the messages, as in Figure 1. The contextual differences are partially reflected in E.

An *occasion* of communication under Φ is a triple $[c_S, m, c_R]$, where $\langle c_S, m\rangle \in E$ and $\langle c_R, m\rangle \in E$. This is to be interpreted as a situation where S, intending to communicate a semantic content c_S, sends a message m to R, who interprets m as meaning c_R. We say this occasion of communication is *successful* when $c_S = c_R$. We call U the *utility assignment*, and assume the following.

(2) U defines both S's and R's utility (a real number) of every occasion of communication under Φ.

The success of the occasions of communication is normally considered as having the greatest contribution to utility.

As for the agents, let us assume:

(3) S and R share all the information in P, E and U.

That is, they mutually believe that their inferences presuppose the same probability distribution over the contents, the same encoding, and the same utility assignment. So the agents are allowed to infer each other's plan to arbitrary depths of belief embedding. Next, we assume:

(4) S and R are selfish utility maximizers, and they recognize each other as such.

So each of them tries to maximize her own expected utility, knowing that the other is doing so. The expected utility is accounted for by the utilities and the probabilities of the occasions of communication under Φ. The probability of an occasion $[c_S, m, c_R]$ of communication is the product of $P(c_S)$, the probability of S's sending m under the condition that she wants to communicate c_S, and the probability of R's interpreting m as meaning c_R, the latter two of which are up to S and R's decision. Finally, for the time being let us suppose:

(5) S and R are logically omniscient, and can carry out any computation with no cost.

This allows us to disregard computational cost when considering the agents' utility, greatly simplifying the analysis. However, postulating logical omniscience here will not diminish the significance of the discussion that follows, because, as shown later, logically omniscient agents should be willing to agree upon some protocol for the sake of reliable communication, and such a protocol is often executable for agents with limited computational capacity, allowing them to simulate logically omniscient ones and hence to obtain the best possible utility.

Below we are going to consider two different types of situations in communication, and formalize them as two different types of games, though it will eventually turn out that those two are equivalent. In either case, S and R are to decide their optimal strategies in the setting of some game. Here let us postulate:

(6) S and R have agreed upon following a certain protocol to determine their strategies.

We suppose that this protocol is a best possible one, in the following sense.

(7) The protocol ensures that each of S and R can uniquely decide her strategy.

Being logically omniscient, the agents should be willing to agree upon respecting such a protocol, because the success of communication contributes best to their utility. The probability of successful communication is maximized when the agents know each other's strategy, which is possible if (7) holds.

We consider that a context is too transient to have any room for negotiation. So it is not that different protocols are agreed upon for different contexts of communication. We assume the following instead.

(8) One common protocol is agreed upon for all the contexts.

S and R as selfish utility maximizers, however, may break the agreement and use different protocols. For the agreement to be effective, the protocol hence should be *stable* in the following sense.

(9) Neither S nor R can expect to profit more by adopting other protocols.

In the following section, we want to investigate what such protocols are.

2.1 $|C| + |M|$-Person Game

When S wants to communicate a semantic content c, she will try to figure out the message m^* which maximizes her expected utility. Her expected utility is influenced by the probability of the success of communication. But the success of communication means R's interpreting m^* as meaning c. So S must calculate, for each candidate m for m^*, the probability of R's interpreting m as meaning c. In turn, when R wants to understand message m, she will try to figure out the semantic content c^* which maximizes her expected utility, which is influenced by the probability of the success of communication, which means here S's having sent m intending c^*. So R must calculate, for each candidate c for c^*, the probability of S's sending m as meaning c.

Consequently, the search by each agent will constitute an infinite tree. Figure 2 depicts such

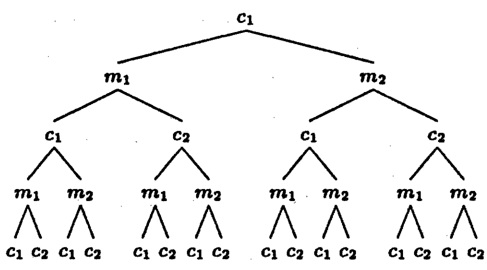

Figure 2: Inference by S to communicate semantic content c_1

a tree for S when she wants to communicate c_1, where $C = \{c_1, c_2\}$, $M = \{m_1, m_2\}$, and $E = C \times M$. The nodes labeled by c_i represent S when she wants to communicate c_i, and those labeled by m_i represent R when she wants to interpret m_i, for $i = 1, 2$. An agent at each such state searches over the subtree dominated by the corresponding node.

Such an infinite tree is not a so-called game tree or decision tree, because it is not the case that S and R are taking turn through the path running down from the root node. The behaviors of S and R in such a setting cannot be regarded as sequences of turns, and hence the situation cannot be regarded as a game of an extensive form.

Incidentally, that logically omniscient agents should search over infinite trees entails that protocols to enforce a finite search depth are unstable. For example, a protocol is not stable if it requires S to disregard the contents which R may wrongly assign to the message. This is because the same message may have different sets of potential interpretations in different occasions. In the situation shown in Figure 1, for instance, S should dynamically take c_1 into consideration when she wants to communicate c_2, because in some other context m_2 instead of m_1 might be ambiguous between c_1 and c_2.

The kind of communication between two logically omniscient agents, which involves inferences over such an infinite tree as discussed above, can instead be regarded as a non-cooperative game with infinitely many players. Here the players are the states of the agents represented by the nodes of the infinite tree. The simple strategies of each player labeled with content c' are the ways of encoding c' into some message m' such that $\langle c', m' \rangle \in E$, and the simple strategies of each player labeled with a message m' are the ways of decoding m' into some content c' such that $\langle c', m' \rangle \in E$. Considering mixed strategies, the utility function of each player is naturally obtained as that of the corresponding state of the agent. The protocol for the agents is shared among all the players. Also, all the players are logically omniscient utility maximizers, since the agents are. This game is non-cooperative, because the players cannot communicate but by knowing the strategies of the players located lower in the tree.

Then the protocol must be *deterministic*, in the sense that it makes all the players with the same label choose the same strategy. Otherwise, a player X, being logically omniscient, would notice that there are several different strategies she could choose.[2] This means that X cannot uniquely decide her strategy under the protocol. But then it follows that S or R cannot uniquely decide her strategy given some content or message, because X represents either S when wanting to communicate a certain content or R when wanting to interpret a certain message, and logically omniscient utility maximizers should come to the same conclusion when situated in the same circumstance. This contradicts (7). S may be able to uniquely determine her strategy for communicating c (the semantic content she really wants to communicate now), but the strategy must be agreed upon without knowing what c is, so that (7) entails that S must be able to determine her strategy for communicating any content in C. Similarly, (7) entails that R must be able to determine her strategy for interpreting any message in M.

Given that a useful protocol must be deterministic, below we shall consider what a stable protocol may be. Note that, under a deterministic protocol, the infinite tree as in Figure 2 is rendered a finite bipartite graph as in Figure 3, because the players with the

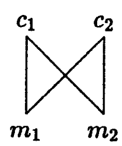

Figure 3: The finite bipartite graph as the reduction of the tree in Figure 2.

same label decide on the same strategy, the nodes with the same label reducing to one and the same node. The game hence reduces to a non-cooperative game with finite ($|C| + |M|$) players, who are S and R when wanting to communicate a certain content and interpret a certain message, respectively. Their strategies are defined straightforwardly as before.

The search space now being finite, the agents often need not be logically omniscient to simulate the behavior of logically omniscient agents. Note that the reduction of an infinite tree to a finite bipartite graph follows if the agents were logically omniscient utility maximizers and thus are willing to agree upon using a protocol meeting (7) and (8). This means that agents with limited computational power can often do as good as logically omniscient agents in raising the reliability of communication, by adopting the same protocol for disambiguation.

Then what should the protocol be like? First, this reduced game must have some Nash equilibrium, because there are only finitely many players and simple strategies, and so the protocol, being stable, must yield one Nash equilibrium. Otherwise, some player would be motivated to change her strategy, hence deviating from the protocol. If some player were motivated to break the agreement and deviate from the protocol, then either S or R should have that chance, contradicting (9), which requires the protocol be stable.

Second, the protocol should maximize some joint expected utility of S and R over the whole context, rather than for a particular case where S wants to communicate some particular content or where R wants to interpret some particular message. This is simply because the same protocol must be shared among all the players in the $|C| + |M|$-person game under consideration, due to (8). Also due to (8), the protocol should maximize the agents' joint utility over all the contexts.

2.2 Multiple Communication

The above discussion extends to the cases where S wants to communicate k different contents at one time. In such a case, an occasion of

[2]Those strategies must be equally profitable to X. Otherwise X could disregard the less profitable strategies for the sake of utility maximization.

communication is regarded as a two-person game, where the players are S and R, each of S's strategies is to choose several messages to encode the contents she wants to communicate, and each of R's strategies is to assign a content to each of the messages she receives.

Through essentially the same discussion as before, the search space for each player in this game reduces to the bipartite graph as in Figure 3. We can also show that two corresponding games of the two types have the same set of equilibria. Note that in the $|C| + |M|$-person game, due to the bipartite configuration of the players, a change of the strategy of S wanting to communicate a content does not affect the utility of S wanting to communicate a different content, and similarly a change of the strategy of R wanting to interpret a message does not affect the utility of R wanting to interpret a different message. So if a player can expect more profit by changing her strategy in the two-person game, then that change corresponds to changes of several players' strategies and some of these players should expect more profit in the $|C| + |M|$-person game. That is, if a state (an assignment of strategies to the players) is a Nash equilibrium in the two-person game, then the corresponding state in the $|C|+|M|$-person game must be a Nash equilibrium as well. The opposite is similarly true.

Consequently, the two protocols, one for the two-person game and the other for the $|C|+|M|$-person game, are essentially the same in the sense that they yield the same assignment of strategies to the communicating agents. In summary, the two settings of communication discussed so far are treated uniformly as respecting the same protocol, which is to yield a Nash equilibrium (of either type of game) which maximizes some joint utility of S and R.

3 Case of Natural Language

Anaphora in natural language provides an example of a context of communication regarded as a two-person game. Several proposals have been done to account for anaphora. Let us take *centering theory* (Joshi & Weinstein, 1981; Brennan, Friedman, & Pollard, 1987; Walker, Iida, & Cote, 1994) as an example, and show that the core of the theory follows from our account developed above.

Let U_i denote the i-th utterance. Centering theory considers list $Cf(U_i)$ of ***forward-looking centers***, which are the semantic entities referred to in U_i. The forward-looking centers of utterance U are ranked in $Cf(U)$ according to the salience. In English, this ranking is determined by grammatical functions of the referring expressions in the utterance, as below.

> subject > direct object > indirect object > other complements > adjuncts

The highest-ranked element of $Cf(U)$ is called the ***preferred center*** of U and written $Cp(U)$. ***Backward-looking center*** $Cb(U_i)$ of utterance U_i is the highest-ranked element of $Cf(U_{i-1})$ that is referred to in U_i.[3] $Cb(U)$ is the entity which the discourse is most centrally concerned with at U.

Further, centering theory stipulates the following defeasible 'rules.'

(10) If an element of $Cf(U_{i-1})$ is referred to by a pronoun in U_i, then so is $Cb(U_i)$.

(11) Types of transition between utterances are preferred in the ordering:

CONTINUE > RETAIN > SMOOTH-SHIFT > ROUGH-SHIFT

The types of transition are defined in Figure 4.

	$Cb(U_i) = Cb(U_{i-1})$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	CONTINUE	SMOOTH-SHIFT
$Cb(U_i) \neq Cp(U_i)$	RETAIN	ROUGH-SHIFT

Figure 4: Types of transition.

Consider the following discourse for example.

[3]We are simplifying the account here by replacing 'refer to' for 'realize' in the original account, but this does not influence the significance of our discussion.

U_1: Tom was late.
Cf =[Tom]

U_2: Bob scolded him.
Cb =Tom, Cf =[Bob,Tom]

U_3: The man was angry with him.
Cb =Bob, Cf =[Tom,Bob]

Suppose that both Bob and Tom are referred to in U_3. Then either 'the boy' refers to Bob and 'him' refers to Tom, or vice versa, but the former violates (10), because $Cb(U_3)$ is Bob. Hence centering theory prefers the latter reading: Tom was angry with Bob.

The case of anaphora considered in centering theory is formulated as a two-person game of the sort discussed above, with the following correspondence.

- C contains the forward-looking centers of U_{i-1}.
- M consists of the parts of U_i.
- P represents the prior probability of each element of C being referred to by an utterance.
- E is the semantically possible referring relation between C and M.

Concerning U, let us assume that referring to salient semantic entity by 'light' linguistic expression yields a very large joint utility for the conversation participants, in such a way that the combination of referring to c_1 by m_1 and referring to c_2 by m_2 has a larger joint utility than the combination of referring to c_1 by m_2 and referring to c_2 by m_1, when c_1 is more salient than c_2 and m_1 is lighter than m_2. Note that this is the case if the utility is the product of salience and lightness, which is captured by considering that salience and lightness correspond to probability and basic profit, respectively. Also, it would be natural to consider that pronouns are lighter than full noun phrases. Then (10) directly follows from maximization of joint utility.

(11) follows as well. The preference for $Cb(U_i) = Cb(U_{i-1})$ results from the natural assumption that $Cb(U_{i-1})$ is highly salient in U_{i-1}, because that raises the utility of reference to $Cb(U_{i-1})$ in U_i. Also, the preference for $Cb(U_i) = Cp(U_i)$ results, because if the same entity were both $Cb(U_i)$ and $Cp(U_i)$ then that would raise the utility of referring to it in U_{i+1}, raising the expected utility of U_{i+1} as a whole. But this preference is considered weaker than the former, because it is based on a prediction of a future utterance.

Several different approaches have been proposed on anaphora resolution (Hobbs, 1978; Sidner, 1983; Kameyama, 1986; Suri & McCoy, 1994). Although we have considered as an example a version of centering approach in the above, it is not at all our intent to claim that it is the right one and the others are wrong. Probably it is not the case that just one of these different accounts is the right one, but all of them capture some truth of natural language discourse. Anaphora, like other linguistic phenomena, is sensitive to various sorts of contextual information, which each of those accounts seems to partially take into consideration. A major advantage of our game-theoretic account is that it reduces the specifics of anaphora to general properties, such as utility and probability, of various linguistic elements, such as pronouns and thematic roles, thus reducing the theory of anaphora to a general theory of communication.

4 Final Remarks

The type of communication where each context involves only finitely many possible semantic contents and possible messages has been formulated as a sort of non-cooperative games. Every effective protocol for disambiguation agreed upon among the communicating agents should yield a Nash equilibrium which maximizes some joint utility of the agents. This account has been applied to a pragmatic aspect of natural language, and some core stipulations of a previous theory have been derived from our theory.

Possible targets of similar applications in linguistics include binding theory, conversational implicature, and so on. Further, such a game theoretic account of natural language may reveal non-separability of syntax, semantics and pragmatics. The account could be also applicable to the design of communicating artificial agents, for the sake of reliable communication among them and with humans.

As for either natural or artificial languages, it is unclear to what extent our theory is applicable when the contexts of communication are complex, either with many contents and messages or with complex contents and messages. Taking computational cost into consideration with respect to utility assignment seems to be a difficult issue. Constraint-based approaches (Hasida, Nagao, & Miyata, 1993; Hasida, 1994) could be extended to embed the equilibration mechanism for disambiguation into the language faculty without complicating too much the design of the computational model.

Reference

Austin, J. L. (1962). *How to Do Things with Words.* Oxford.

Barwise, J., & Perry, J. (1983). *Situations and Attitudes.* MIT Press.

Brennan, S. E., Friedman, M. W., & Pollard, C. J. (1987). A Centering Approach to Pronouns. *Proceedings of the 25th Annual Meeting of the ACL*, pp. 155–162.

Durfee, E. H., Gmytrasiewicz, P., & Rosenschein, J. S. (1994). The Utility of Embedded Communications and the Emergence of Protocols. *Proceedings of AAAI '94 Workshop on Planning for Interagent Communication.*

Etzioni, O. (1991). Embedding Decision-Analytic Control in a Learning Architecture. *Artificial Intelligence*, **49**, 129–159.

Hasida, K., Nagao, K., & Miyata, T. (1993). Joint Utterance: Intrasentential Speaker/Hearer Switch as an Emergent Phenomenon. In R. Bajcsy (Ed.), *Proceedings of the 13th International Joint Conference on Artificial Intelligence.* Chambéry:

Hasida, K. (1994). Emergent Parsing and Generation with Generalized Chart. *Proceedings of the Fifteenth International Conference on Computational Linguistics*, pp. 468–474.

Hobbs, J. R. (1978). Resolving Pronoun Reference. *Lingua*, **44**, 311–338.

Joshi, A. K., & Weinstein, S. (1981). Control of Inference: Role of Some Aspects of Discourse Structure — Centering. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pp. 385–387.

Kameyama, M. (1986). A Property-Sharing Constraint in Centering. *Proceedings of the 24th Annual Meeting of ACL*, pp. 200–206.

Nagao, K., Hasida, K., & Miyata, T. (1993). Emergent Planning: A Computational Architecture for Situated Behavior. In J.-P. Muller (Ed.), *Proceedings of Fifth European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pp. 1–13. Neuchatel, Switzerland:

Sidner, C. L. (1983). Focusing in the Comprehension of Definite Anaphora. In M. Brady, & R. C. Berwick (Eds.), *Computational Models of Discourse*, pp. 267–330. MIT Press.

Suri, L. Z., & McCoy, K. (1994). RAFT/RAPR and Centering: A Comparison and Discussion of Pronouns Related to Processing Complex Sentences. *Computational Linguistics*, **20** (3), 301–317.

Walker, M., Iida, M., & Cote, S. (1994). Japanese Discourse and the Process of Centering. *Computational Linguistics*, **20** (2), 193–232.