Natural Language and Multi-Agent Systems

Blair Tennessy
Department of Computer Science
University of British Columbia
tennessy@cs.ubc.ca

ABSTRACT

This paper explores the role of communication and natural language in social systems composed of multiple autonomous agents. We summarize a number of models for language and communication with special attention on optimization and game-theoretic concepts. We draw our survey from a gamut of literature, touching on books and papers from diverse fields such as linguistics, philosophy, computer science, and economics.

Keywords

Language, Communication, Decision Theory, Game Theory

1. INTRODUCTION

Language is an important cognitive tool and its emergence is closely linked to the largest recent evolutionary step, namely conceptual consciousness. Language is the medium of thought and for the interaction of intelligent individuals. Fluency in a popular language, especially today's *lingua franca*, English, endows one with a wider range of action and possibility:

> "Unlike many other kinds of competence, the knowledge of a language yields more benefits to an individual the larger the number of people who share it. Knowing a widely spoken language enables the individual to communicate with a larger number of persons and widens the set of possible interactions."
> - Silvana Dalmazzone

This survey ranges from low-level communication in terms of "meaningless" signals and simple messages, to high-level interest in the pragmatic effect of speech acts and social coordination through law. The reader will not require any sophisticated knowledge concerning linguistics; the underlying processes involved in natural language understanding are not discussed and are conveniently abstracted out of the formal treatment. Regarding theme, the central idea is that communication, and especially natural language communication, is a means to coordinating action.

True to form, this monologue may be considered a cooperative communication game in natural language between you, dear reader, and the author. May you derive utility from the text that follows.

1.1 Survey Outline

We begin by examining agent interaction within games which include the possibility of communication. When do optimal strategies demand that we remain mute or act as though we are deaf? What models are used to quantify success in communication? What different forms of equilibria arise in these games? Can evolutionary game theory drive us towards new terms, signals, and behaviors?

We then diverge into a discussion about discussions: conversations, debates, and arguments, all of which taking place in natural language, have been the recent focus of philosophers and economists. Here we include thoughts on the work in pragmatics by the English philosopher Paul Grice, and review his followers' attempts to formalize his inquiries into the ways of words within a game theoretic framework. We then investigate two models of debate and cite results about truth revelation strategies for the observer of a debate.

The final section contains a discussion of law and contracts—a device formalizing the notions of intent, with a formidable agency empowered and charged with the duties of interpretation and enforcement. Recently, work in economics has focused on the role of ambiguity of natural language and its impact on the contract, leading to incompleteness or multiple interpretations. We will examine, from a game-theoretic standpoint, concepts of contract readings, judge impartiality, and enforcement.

1.2 An Introductory Example

Let us conclude our introduction with an example of optimal linguistic behavior. A commonly observed phenomenon across all natural languages is the fact that frequently used words receive succinct orthographic forms. The corresponding terse, short vocalizations serve the purpose of verbal economy. Let us recall Zipf's Law:

> The probability of occurrence of words or other items starts high and tapers off. Thus, a few occur very often while many others occur rarely.

For example, the closed class of articles in English ("a/an" and "the") are high-frequency items; correspondingly, they enjoy short, monosyllabic verbalizations.

2. COMMUNICATION IN GAMES

This section is purposed on an investigation of the advantages of communication—natural or otherwise—on the outcome of a game. Certainly there are situations in which communication confers a benefit to those involved, for the very reason that it allows one to signal intent to play a certain action. We will also look at games of communication in which information is known by one party and not the other (for example, one player may know the particular game being played in a game of incomplete information).

We will review work in pre-play communication games, games of understanding, and *signalling games*. Our approach will borrow examples from studies in the social behavior of primates.

2.1 Pre-play Communication

Let us survey some results for various *costless* pre-play communication in games. We will look at situations in which agents have a game to play but engage in a series of messaging rounds prior to playing the actual game. We start with one-sided communication, then move to multi-sided communication. We will find that the value of communication in this setting is related to a measure of *risk* [1].

Consider an arbitrary two-player finite game G with multiple strict Nash equilibria. Assume that there is a unique Nash equilibrium strategy profile $s^* = (s_1^*, s_2^*)$ which maximizes the payoff to the first agent. This agent is to choose one message m from a set M, send it, and play some action in the (simultaneous) game. The second player receives the message and plays an action which may depend on m. Formally, let $F(G, M)$ be the set of pure strategies for the second player in the communication game induced by G and M. Define

$$\hat{F}(G, M) = \{f \in F(G, M) \mid \exists m \in M : f(m) = s_2^*\}$$

One can (conservatively) measure risk ρ relative to the first player of an equilibrium $s^* = (s_1^*, s_2^*)$ maximizing that player's payoff:

$$\rho((s_1^*, s_2^*), G) = \max_{s_1 \neq s_1^*} \frac{\max_{s_2} u_1(s_1, s_2) - \min_{s_2} u_1(s_1^*, s_2)}{u_1(s_1^*, s_2^*) - \max_{s_2} u_1(s_1, s_2)}$$

This measure is invariant to positive affine transformations of the payoff function. Using this measure, one can show that, when the following condition involving the size of the message space obtains,

$$\rho((s_1^*, s_2^*), G) < \frac{1}{|\hat{F}(G, M)| - 1}$$

then s_1 can communicate effectively (that is, agents play in an equilibria of the induced communication game in which the payoff vector is $u(s^*)$).

The general situation in which all n players may talk, termed multi-sided communication, is different in nature than the single-sided communication scenario. Blume [1] argues "if all players can talk, communication can reduce strategic uncertainty enough to ensure that a unique efficient equilibrium in the underlying game will be played." He introduces a modification which favors "efficient action"; a message is thought of as *affirming an intent* to play according to the preferred equilibrium.

Again, start by assuming a unique Nash equilibrium strategy profile s^* in the underlying game such that playing according to that strategy delivers the maximum payoff possible for each agent. The trick is to assume an (arbitrarily small) bonus of $\epsilon > 0$ for the use of a strategy where the message m_i affirms intent to play according to this unique equilibrium, regardless of the messages of the other agents m_{-i}. Formally, the payoffs are slightly modified so that

$$U_i(m, f) = \begin{cases} u_i(f(m)) + \epsilon & \text{if } \forall m_{-i} : f_i(m_{-i}) = s_i^* \\ u_i(f(m)) & \text{else} \end{cases}$$

Agents engage in a single round of simultaneous communication by announcing a message m_i from their message space M_i. Blume proves that these "intent affirming" preferences lead to coordination on the optimal strategy.

2.2 Meaning and Forms

Prior to saying something, one must have something to say. And when one has something to say, one wants to be understood. Let us consider a simplified model [6] in which we have a fixed set of meanings M (things to say), and fixed set of forms F (ways of saying them). A *speaker strategy* is a function $s : M \to F$, and a listener strategy is a function $h : F \to M$.

We assume that one can decide on what to say (some object $m \in M$). Consider the game in which nature presents the speaker with m. Communication is modelled as an encoding of m through the speaker strategy s, transmission, and a subsequent decoding of $s(m)$ on the part of the listening agent. Communication is successful in the case when the listener recovers the meaning from the observed signal, and measured with

$$\delta_m(s, h) = \begin{cases} 1 & h(s(m)) = m \\ 0 & \text{else} \end{cases}$$

However, communication is a process and thus requires work. There is a cost, and one may assume that the cost of producing a signal is related to the complexity of that signal. We assume cost function $c : F \to \mathbf{R}^+$. Now the speaker has competing interests: he wants to be understood while minimizing the complexity of the form used. In a game of meaning understanding, the speaker utility is

$$u_1(m, s, h) = \delta_m(s, h) - c(s(m))$$

Similarly, for the hearer, decoding is a process with some cost, and his utility is

$$u_2(m, s, h) = \delta_m(s, h) - d(s(m))$$

2.3 Signalling Games

Coordinated communication is a phenomenon observed within many animal species in nature, as well as humans. While not as rich as natural language, the emergence and maintenance of specific signals—vocalizations or through some other means—communicating a certain meaning serves as an interesting parallel worthy of consideration at this early point of our foray into information exchange. One of the most common forms of this communication in the animal world is the alarm signal, studied here.

Put simply: some animals are capable of forming distinct sounds with their vocal tract and of detecting sound vibrations with their ears; to harness this ability to individual and social benefit, to gain from the exchange mechanism made possible by his physique, the exchange must be coordinated. When a member of a population in a certain situation emits a signal to his group, the members of that group must respond accordingly to that situation.

Coordinated communicative behavior is witnessed in a host of animal populations, especially mammals and birds. One example is the vervet monkey (Ceropithecus aethiops), who make distinct alarm calls if they detect predation by leopards, eagles, or snakes. The monkeys respond to each signal in a manner unique to the type of danger. For example, upon hearing a leopard call, vocalized as a bark, the members of the troupe will run for the trees, seeking shelter [9].

First, let us review one formulation of a signalling game [12]. This is a two-player sequential game. There is a *sender* with knowledge of his type t but no payoff-relevant actions, and a *receiver* with payoff-relevant actions. Common knowledge to both is the prior beliefs on the state of the sender. The sender attempts to influence the action of the receiver by sending a signal with form $f \in F$, and the receiver responds by taking an action $a \in A$. Denoting the set of types by Θ, we have utility evaluated on the triple $\Theta \times F \times A$.

There are various instances of models for signalling. One can allow behavioral strategies in which the sender may pick different signals from within a single state. However, let us assume a simple model, which is phrased in terms of strategies and only associates one form with type. Formally, a *speaker strategy* is $s : \Theta \to F$, and a *receiver strategy* is a mapping $r : F \to A$.

One important assumption in the original work on conventional signalling in [7] is that messages are *costless*, formally

$$u_i(t, f, a) = u_i(t, a)$$

It has been shown that, for this situation, the sender has influence on the action of the receiver only when there is common interest between the two [3]. If we are playing a game of interpretation, then only the case in which communication is successful has bearing on the utility of the communication strategies; similar to the δ of our game of understanding:

$$u(t, s(t), r(s(t))) = \begin{cases} 1 & r(s(t)) = t \\ 0 & \text{else} \end{cases}$$

There are several varieties of equilibria in signalling games. These include *pooling* equilibria, where the sender emits the same signal in all states, and *babbling* equilibria, where the receiver ignores the signal (the sender is babbling). However, we are most interested in *separating* equilibria, roughly corresponding to the notion of proper interpretation of a signal. It can be shown that not all such games have a separating equilibrium.

Revisiting our primate example, researchers have developed a number of models for the emergence and maintenance of coordinated communication systems. We will review the evolutionary argument for the invention of a warning signal and an appropriate response [10]:

Say there were no word in our language for alert. A mutation would create a small mutant group using a new costless signal and responding appropriately. Now, many games are played in which agents are paired and nature selectively presents a information about a grave danger to one of these agents. Their fortunes fall together: this agent must signal his buddy about the danger, who decides on either avoiding the danger, ignoring the message, or panicking. In terms of equilibria, the second may respond by avoiding the danger, leading to a separating equilibrium, or he may enter a pooling equilibrium, in which he ignores or panics.

Standard definitions of *evolutionary stable strategies* and even *evolutionary stable sets* do not eliminate the pooling equilibrium. Researchers have had to extend their models to a *modified evolutionary stable strategy with complexity considerations* in order to give support to the separating equilibrium. The rationale behind the complexity considerations is that evolutionary forces should eliminate strategies which require mental resources to respond to messages that never occur. Thus evolution can be used to cut the threat of reacting in panic, paving the way for mutations that actually transfer and make use of information.

3. GAMES IN NATURAL LANGUAGE

The conventional meaning of a sentence underdetermines the interpreted content of that expression. For example, consider the following bit of roadside conversation:

> A: I am passing through and ran out of petrol.
> B: There is a station two blocks down.

In the normal course of affairs, one would assume that speaker A meant "I have no knowledge of this town, I intend on continuing on my journey, and I require the location of a store selling gasoline in order to get moving," while B meant "There is a gas station two blocks down with fuel and it is serving customers."

In an effort to account for this phenomenon, the English philosopher Paul Grice developed the idea of *conversation implicature*, a rough treatment explaining this form of non-conventional implicature in dialogue. Implicature is a term used to refer to a licensed set of inferences made possible under certain body of assumptions.

In his essay "Logic and Conversation", included in his book "Studies in the Ways of Words" [5], Grice posited the *Cooperative Principle*:

> Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

Following this, and "echoing Kant", he listed a number of categories containing a variety of maxims and submaxims. These categories, and their associated maxims are:

- **Quantity**: make your contribution as informative as required by the current exchange, and not more so.
- **Quality**: make you contribution one that is true.
- **Relation**: be relevant.
- **Manner**: be perspicuous: avoid obscurity and ambiguity; be brief and orderly.

At the very least, these serve as some rough guidelines for conducting oneself in conversation, seen as a cooperative communication game. They also guide the interpretation process to determine the pragmatic, context-dependent meaning of a speaker's utterance.

3.1 Grice's Maxim of Relevance

Grice approached the problem of underspecified conventional meaning in conversation philosophically. A number of researchers following Grice have attempted to ground his common-sense notions of conversational practice in a formal, game theoretic framework. Many of his maxims have been tangentially related to the communication considerations considered earlier in this document. For example, the maxim of Quality—that one should speak truth to one's fellows—is naturally defensible in a cooperative game. It is obvious that the maxims of Quantity and Manner are related to the cost of producing and understanding an utterance, and also bear some relation to whether one will be properly understood.

Let us examine the maxim of Relation from the standpoint of game-theory and optimization [11]. We will attempt to quantify relevance of various types of dialogue acts and generalize them from purely cooperative situations to strictly opposing games. During our discussion, it will become apparent that the best strategies in these games are linked to relevance maximization.

First, assume that we are in a cooperative situation and must decide on a course of action. Given a distribution p over the set of information sets W and a set of actions A, we can compute the expected utility of an action $a \in A$:

$$\mathbf{E}[u(a)] = \sum_{w \in W} p(w) \cdot u_w(a)$$

Now assume that, in the same situation, our agent is informed of a proposition c via an assertion speech act by a comrade. Say that, as we are about to shop for a favorite item, a friend mentions that one of the competing stores in the neighborhood is having a sale today of this good at a low price. Conditionalizing the probability function by this new proposition, we re-evaluate our expected utility:

$$\mathbf{E}[u(a) \mid c] = \sum_{w \in W} p(w \mid c) \cdot u_w(a)$$

We arrive at the *relevance* R of the assertion c as the difference between the utilities of the optimal action a^* in the informed state and the optimal action $\bar{a}$ from the uninformed state performed within the conditionalized universe:

$$R(c) = \mathbf{E}[u(a^*) - u(\bar{a}) \mid c]$$

This value is also known in statistical decision theory as the *value of sample information*. Continuing with our example, prior to the "good buy" assertion, we might have been indifferent to shopping at the stores, as they offer competitive prices; with the knowledge of the sale, we will probably go for the deal.

Viewing an assertion as an answer to a question, we can use the value of assertions to determine the utility of questions. Let a question Q consist of a partition of answers $Q = \{q_1, \ldots, q_n\}$. The *expected value of sample information*, $ER(Q)$, or value of the question Q, is the average utility value of the possible answers:

$$ER(Q) = \sum_{q \in Q} p(q) \cdot R(q)$$

A question is deemed *relevant* when asking it might provoke an answer that changes the action taken by the decision process, that is, when $ER(Q) > 0$.

Suppose now that we are engaged in an argument, an instance of a strictly opposing game. Say that an agent wants to argue for a particular hypothesis h and against $\neg h$. Model the common ground as an information state by probability function p. A proposition c may support h, in which case $p(h \mid c) > p(h)$, or it may detract from h when $p(h \mid c) < p(h)$. Define the *argumentative force* of c with respect to hypothesis h as

$$F_h(c) = p(h \mid c) - p(h)$$

It follows that an agent arguing for h should compose an argument with the proposition which has the greatest argumentative force. We can also evaluate arguments and order in terms of force. To make the most convincing argument, the most relevant, forceful proposition should be discussed.

This is a curious notion: when one finds that one is defending an untenable position, it is best to cite arguments that are irrelevant to the discussion at hand. Indeed, this seems to be the current strategy of our Prime Minister during Question Period.

3.2 Debate and Inference

Debate is interesting in that it is a natural language game, and it has certain sequencing properties: the persuasive force of an argument may be stronger when used as a counter-argument rather than an argument [10, 8].

Consider the following scenario: there is some finite set of possible states S of the world and some true state s^*. An uninformed listener does not know the state of the world and is to choose some state; the optimal state providing the greatest utility is s^*, the actual state of the world. There are two debaters—in full possession of information about the actual state of the world—intent on posing arguments (and counter-arguments) with the desire of persuading the listener to choose their preferred outcome. We assume that the preferences of each debater do not depend on the true state.

In terms of communication, each agent is endowed with some set of messages $M_i(s)$ in a given state $s \in S$. Further, evidence can be entered into the debate which may refute previous claims, or equivalently, a message $m_{\neg s}$ which proves that the state is not s. Formally, these are messages that, for every state $s \in S$, there is a message $m_{\neg s}$ with $m_{\neg s} \in M_i(s')$ if and only if $s \neq s'$.

Let us consider two people settling damages before a judge. The second agent has caused some amount of damage in $S = \{1, \ldots, 10\}$. The first player may claim anything as his damage but has no evidence, while the second has evidence which refutes any false claim by the first. The civil court system is backlogged, so the judge conducting the arbitration can hear only two messages. As an impartial judge, a finding of the truthful amount of damage is of paramount importance, but how can his honor elicit the truth and render an honest verdict?

To find a separating equilibrium, the judge asks for the value of the damages from the first player. He then asks the second to refute this claim. If the second cannot refute it, the judge has found his verdict. When the second player does refute it, the judge penalizes the first as much as possible. A policy of truthfulness is in the first player's interest. Crucially, asking for damage statements in this order is far better than spending many messages examining the negative evidence of the second.

The rule used above is an instance of a "believe unless refuted" (BUR) inference rule [8]. These rules have been shown to cause full information revelation in a general set of argument situations. Basically, the idea is to conditionally accept a statement until it is later refuted. If some debater provides refutation with evidence, his claim is now conditionally accepted. This process continues until the last round, at which point the observer infers the last accepted claim.

The general argument game has N debaters. There is some schedule to the debate, i.e. a speaking order. The BUR result requires the schedule to be one in which each debater gets at least one chance to speak. Also, the observer is required to know nothing more beyond the fact that debaters have conflicting preferences.

A key notion here is *burden of proof*. Consider two states s and s' with the property that $M(s) \subset M(s')$ (strict inclusion). If the true state is s, there is no evidence to refute a claim by an agent that it is s'. The agent then must provide an argument for s from $B(s)$, the set of messages in $M(s)$ which are not elements of any $M(s')$ where $M(s') \subset M(s)$. Lipman-Seppi assume *refutability*: for all s, s' with $s \neq s'$ and $M(s) \not\subset M(s')$, we have $B(s) \subseteq M(s')$. When refutability holds, and the observer is inferring using the BUR rule, every equilibrium among the debaters permits the observer to learn the true state of the world.

3.3 Optimal Debate

A different approach to modelling debate is taken in [10]. Formally, debate is a mechanism for extracting information from the debaters. The mechanism designer is intent on increasing the probability that the correct conclusion will be selected by a listener. As with any debate, there are limits to the time allowed for argument and on the participants cognitive abilities.

The authors model the situation as follows. A listener with no knowledge of the state of the world is to choose between two alternative outcomes o_1 and o_2 corresponding to the two informed debaters. To the listener, the correct outcome is determined by five *aspects*, a five-tuple which may take values in $\{1, 2\}$ (an aspect having value i supports outcome o_i). A state ω is also a five-tuple of realizations of the five aspects, and the correct outcome in state w is the one which is supported by a majority of arguments.

In terms of language, debaters raise arguments of the form "argument j is in my favor", and that they must prove their claims, thus disallowing a debater to claim an aspect in their favor unless it is so. Note here that debaters are not allowed to make a claim of the state.

With unlimited debate length, the listener could simply request three debates from one of the agents; this is enough for either debater to prove a majority case in their favor, when possible. The problem is made interesting by constraint: the length is restricted to two arguments [4]. This means that not all information can be revealed, and we are interested in inference rules which generate the maximum amount of revelation (equivalently, rules for which an observer minimizes mistakes in judgment).

The situation becomes an extensive form game. The set of feasible moves is a subset of $S = \{1, 2, 3, 4, 5\}$. Exactly one of the outcomes is attached to each terminal history of the game signifying the winner of the debate. There are at most two moves, and this allows a few different debate formats. First, we could have a *one-sided debate*, in which a debater makes at most two arguments, or we could have a *simultaneous debate*, in which each debater makes at most one argument, or we can have a *sequential debate*, in which on debater can make at most one argument, then the other has a turn at offering at most one argument. Note that, in this formulation, the observer is not actually playing; the game is two-player.

A *mistake* is made when the probability that we arrive at the incorrect outcome in a given state is 1. An optimal debate is one which minimizes the number of mistakes over all possible states. So, which debate procedure is optimal? It has been shown that the minimum number of mistakes possible in a one-sided debate is 4, the minimum number in the simultaneous debates is 5, and that there is a sequential debate format which admits only 3 mistakes. Thus the optimal debate procedure is sequential.

Let us end this section with a conclusion about debates made in [4] and summarized in [10] concerning the asymmetry of the optimal debate procedural rule. The persuasion rule is such that "there is a pair of aspects i and j such that when presented in sequence, i is a persuasive counter-argument against j and j is a persuasive counter-argument against i." Thus, under the optimal debate, players are not treated symmetrically.

4. LEGAL SYSTEMS: CONTRACT

Game theory is concerned principally with agents acting rationally, or with agents attempting to act towards their selfish maximal good. However, games in the real world do not occur in a vacuum, and unless we insulate players against injustices such as "utility theft", there would be little value in playing games. The means of developing a just environment is law, which is a natural language encoding spelling out rights accorded to each agent, and outlawing certain actions. The agent charged with *enforcement* is the state.

Consider, for example, the game of anarchy. A population is composed of two types: producers and criminals. There is no government, so there is no limit to the amount of force a particular criminal may visit upon a producer. It is the lot of the producer to toil in the game of production, in which utility is mined from the mechanistic environment (value at the price of time). It is the lot of the criminal to rob a producer who has accrued utility. For good measure, the criminal also breaks the producer's nose, leaving him with negative utility yet still fit for work. It is obvious that, in a rational sense, the producer ought not produce; when he does, his values are plundered.

Law spells out the permissible actions and the proper conduct of the agents; it is the first moral step of a primitive society. One finds constitutions in ancient Greece, the Code of Hammurabi of ancient Babylon, and the Decalogue of Moses. It is important to note some characteristics of these laws: they are written and therefore recorded and duplicated, and they have an air of permanence. For example, the Decalogue was commanded unto Moses and his encampment from God himself upon Mount Sinai. It is of no small import that the medium, two stone tablets, confer the distinction of permanence (His law is "set in stone").

Consider now how the producer and the criminal of our game of anarchy would fare were Hammurabi to install the following injunction and prohibition against theft:

> If any one steal cattle or sheep, or an ass, or a pig or a goat, if it belong to a god or to the court, the thief shall pay thirtyfold; if they belonged to a freed man of the king he shall pay tenfold; if the thief has nothing with which to pay he shall be put to death.

4.1 Contracts

On the whole, the criminal represents a tiny portion of the population. However, trade is a valid, potent source of wealth, and some coded, formal device is necessary to commit parties to agreement and obligations in future periods of time.

We will examine the notions of an agreement, contract, and enforcement as presented in the literature [2]. We then discuss notions of *polysemy* and *ambiguity* in contract games, phenomena which lead to multiple interpretations and thus possible misunderstandings regarding contractual obligation and fulfillment on behalf of the involved parties.

A *reference game* is a two-player perfect-information sequential game. Each party $i \in \{1, 2\}$ has a set of actions A_i. An *agreement* α is a set of terminal histories of the game. An agreement comprised of more than one terminal history permits a choice of action for a party at some stage. Such an agreement is termed *incomplete*, while an agreement with one history is termed *complete*.

In order to form a *written contract*, one must encode the meaning of an agreement into a coded document. Given a joint action space $A = A_1 \cup A_2$, an (exogenously given) encoding device is a pair (C, E) where C is a lexicon (a set of words) and E is an onto mapping $E : A \to C$. This mapping allows us to model situations in which different actions share the same code, in which case the encoding is *polysemic*.

We are now in a position to write a contract. Take the set of tokenized string sequences $S(C)$ generated from lexicon C. For a string $s \in S(C)$, let s^k denote token k. Let $\bar{E}(A_i)$ represent all codes associated with elements A_i by encoding E. A written contract is a pair (w_1, w_2) where $w_i \in S(E(A_i))$, a pair of written sentences describing behavior for both parties.

We now move to interpretation of a contract. A *reading* of contract w is a pair $(r(w_1), r(w_2))$ of tokenized action strings where the reading of each individual $r_i(w_i) \in S(A_i)$ is the same length as the written component, with the property that the reading for each token w_i^k from the written component belongs to the preimage of $\bar{E}$:

$$r^k(w_i) \in \bar{E}^{-1}(w_i^k)$$

Let the terminal histories induced by this reading $T_R(w, r)$ be made up of actions in the union of these action strings.

A reading of a contract is *admissible* if and only if it is nonempty. An admissible reading is *complete* when $T_R(w, r)$ contains exactly one history. A contract is *ambiguous* if it permits more than one admissible reading. A contract is termed *complete* if all its admissible readings are complete.

The *contracting game* is an extensive form game with three players with different information. We now describe the game in terms of two stages: the *proposal* and the *enforcement*.

First, nature assigns "bargaining-power" to one party, who will propose a contract to the other. This contractual proposal is a pair consisting of an agreement and a written contract document. This pair must be compatible. More precisely, given a contract w and a reading $r(w)$, define the continuation game-form $\mathbf{CG}(r(w))$ as the game-form constructed from $T_R(w, r)$. An agreement and contract are *compatible* if and only if the game made up of the terminal histories of a reference game consistent with the agreement is a continuation game under some admissible reading of the contract.

The other party may reject the proposal (in which case the game is over and the parties receive no payoff), or the game will continue to the second stage, enforcement. In the enforcement stage, a judge decides on a continuation game corresponding to the admissible readings. Equivalently, a judge chooses a particular admissible reading of the contract. When w is not ambiguous, the judge has no choice but to honor the lone admissible reading of the contract.

The judge thus has the power to pick the continuation game for the parties. But, strategically, what is he maximizing? One approach is to limit the judge to knowing only the written component of the game, and to state that he is to implement the will of the parties as accurately as possible. Thus the judge attempts to choose the continuation game which matches the agreement between the parties. In order to keep him impartial, it has been suggested that he receive an infinite payoff when he does this, and negative infinity when he doesn't. The compatibility requirement ensures that at least one continuation game matching the agreement is choosable.

5. REFERENCES

[1] A. Blume. Communication, risk and efficiency in games. Technical report, Economics Working Paper Archive at WUSTL, Dec. 1993.

[2] S. Chatterji and D. Filipovich. Ambiguous contracting: Natural language and judicial interpretation. Technical report, Econometric Society, Aug. 2004.

[3] V. Crawford and J. Sobel. Strategic information transmission. *Econometrica*, 50:1431–1451, 1982.

[4] J. Glazer and A. Rubenstein. Debates and decisions: On a rationale of argumentation rules. *Games and Economic Behavior*, 36:158–173, August 2001.

[5] P. Grice. *Studies in the Way of Words*. Harvard University Press, 1989.

[6] G. Jager. Evolutionary game theory and linguistic typology: A case study. In *Proceedings of the 14th Amsterdam Colloquium*, 2003.

[7] D. Lewis. *Convention: A philosophical study*. Harvard University Press, 1969.

[8] B. Lipman and D. Seppi. Robust inference in communication games with partial provability. *Journal of Economic Theory*, 66:370–405, August 1995.

[9] M. Oliphant and J. Batali. Learning and the emergence of coordinated communication. *The newsletter of the Center for Research in Language*, 11(1), 1997.

[10] A. Rubenstein. *Economics and Language*. Cambridge University Press, Cambridge, UK., 2000.

[11] R. van Rooy. Relevance of communicative acts. In *Proceedings of Tark 2001*, 2001.

[12] R. van Rooy. Signalling games select horn strategies. *Linguistics and Philosophy*, 27:493–527, 2004.