PRASHANT PARIKH

COMMUNICATION, MEANING, AND INTERPRETATION

1. INTRODUCTION

In this paper, I give definitions of Gricean communication, speaker meaning, and addressee interpretation. In order to do this, I first develop a game-theoretic model of communication which provides a set of sufficient conditions for communication. I then abstract from these to derive necessary and sufficient conditions. And finally, I base my account of speaker meaning and addressee interpretation, two symmetric notions, on my account of communication.

Language is a special kind of tool. It is in fact a complex social institution. All social institutions are of course tools that enable us to organize different aspects of social life. It seems plausible that the primary function of language is communication. Indeed, it is possible to see language, in particular, meanings, as arising from the interactions of a group of agents. This is how all social institutions emerge. Language is no different, except that it arises from the *communicative* interactions of agents. Lewis (1969) and Schiffer (1972), for example, have argued for such a view. This makes communication a key concept in any account of language. Grice's (1957, 1969) ideas on nonnatural meaning provide the best starting point for constructing a model and definition of communication.

I will first describe the basic ideas that undergird communication. These are the ideas of meaning and content and certain aspects of game theory and strategic inference. I then introduce our two main characters $\mathcal{A}$ and $\mathcal{B}$ and try to fix the parameters of the problem of communication in somewhat precise terms. Greater precision will come as we move along, but this will serve as a starting point.

Having done this, I present a range of types of examples that my theory will be able to account for. This list is far from exhaustive, but it provides an initial indication of the scope of the model. I then take up one of these and analyze it. This analysis yields a set of sufficient conditions for communication. As I do this, I urge the reader to keep the other examples in

Linguistics and Philosophy **23:** 185–212, 2000.

mind, because the analysis applies equally to them. And perhaps readers can test the model with their own examples.

2. Meaning and Content

Most languages are situated. This makes it possible for different propositions to be communicated in different circumstances with the same sentence. For example, an appropriate utterance of “It’s 4 pm” on different days results in quite different propositions being expressed, that it is 4 pm on the day of utterance. The situation theory of Barwise and Perry (1983) makes this context-dependence an integral part of utterances. Once we allow situations a role in the determination of content, it becomes clear that there are certain aspects of utterances that are constant across utterance situations and there are others that vary from one situation to another. One of the most salient linguistic constants is the meaning of a sentence. This is different from its content in an utterance, which varies from situation to situation.[1] Meaning is the collection of possible contents of a sentence. If $\mathcal{A}$ utters a sentence in a situation, it allows $\mathcal{B}$ to disambiguate it and choose one or more propositions from this collection as its content. We could in fact write a simple schematic equation of the form “(meaning of) sentence $\oplus$ discourse situation (or the situation of utterance) = content.”

Part of the task of a theory of communication is to explain how this equation comes about. If a language is given, its meanings are given. Then the problem is to get from meaning to content via the discourse situation. Solving this problem in a completely general way turns out to be an extremely difficult task. This is what we will set out to do. To carry it out, we will need some tools.

3. Game Theory

I bring to this problem the powerful ideas of game theory developed by von Neumann, Nash, Arrow, Debreu, Aumann and other game theorists and economists. The ideas of rational agency, strategic interaction, and equilibrium developed in this tradition provide the framework we need to solve this problem. They allow us to extract one more salient constant (like meaning) from the discourse situation and refine the schematic equation

[1] Many writers use “meaning” to refer to what I am calling content. This is also the colloquial use. Unfortunately, there are two different concepts we need to talk about, so we need two different terms. I also use meaning to refer to the related meaning function.

above to "agent architecture ⊕ sentence meaning ⊕ situation of utterance = content."

The idea of rational agency tells us, via its axioms (see Myerson, 1995), how a rational agent chooses an action from a set of actions. In our case, these actions are utterances and interpretations. The agent has a preference ordering over these actions and chooses actions in accordance with these preferences. This preference ordering can be translated into a numerical scale. Each action results in a payoff that can be measured on a utility scale. The agent then chooses the action with the highest utility. A slight wrinkle is introduced when we consider uncertainty. If payoffs are uncertain, as they often are (consider buying a lottery ticket as an action), then the agent assigns a probability distribution to the possible outcomes, and a payoff to each outcome. In this case, the agent chooses the action with the highest *expected* utility.

An agent can no longer do this quite so simply when there are other rational agents around. The actions of other agents also affect the first agent's payoffs. That is, payoffs are functions of everyone's actions. The idea of strategic interaction tells us how a rational agent takes into account another rational agent's possible actions before choosing his best option. In our case, this means how $\mathcal{A}$ and $\mathcal{B}$ take each other's possible actions into account before choosing their utterance and interpretation. Taking another agent's actions into account involves considering not only his options but also his knowledge and beliefs, especially his shared knowledge (with the other agent) of the situation. This is a generalization of the first idea to a multiperson situation. I call the reasoning of agents in a game *strategic inference*. We could of course consider more than two agents if we wanted. But we will stick to two agents to keep the logic simple and clear.

The idea of equilibrium comes from physics. In the context of game theory, it tells us when the combination of choices by two or more agents is in balance. No agent has an incentive to change his action. There are other possible conditions on equilibrium (and this has been an area of research in game theory), but the basic idea is that optimality in the single-person case gives way to equilibrium in the multiperson case.

Grice, and subsequently, Strawson (1964), and especially Schiffer (1972) have shown how communication involves extremely complex interactions between speaker and addressee. These interactions are precisely strategic interactions. However, game theory as currently formulated does not provide a ready-made tool to model communication. It is necessary to develop its insights from first principles. We will also need to generalize the framework of game theory itself.[2]

Apart from the obvious benefits of formalization, why do we need game theory? Isn't it possible to improve upon what Grice, Strawson, and Schiffer did using their methods? I think not. If we are really to give an account of natural language communication, we have to take into account its situatedness explicitly. This means we need to be able to disambiguate between multiple contents. I claim this is not possible without mathematics, because probabilities are involved, and this leads to a vastly more complex structure than it is convenient or possible to handle in natural language. Besides, once one employs the relevant mathematics, things actually become simpler, and this is one of the obvious benefits of formalization. Another is precision. A third is the possibility of defining concepts like communication and deriving their properties rigorously. Note that the game theory I introduce doesn't simply extend the analyses of communication and speaker meaning to cases where there is ambiguity. It provides a new and, I think, more correct *definition* of these concepts. It thereby provides better prospects for *reducing* them to the more basic concepts of intentions, beliefs, and knowledge.

This brings us to the parameters of our situation.

4. THE SITUATION

We already have two rational agents $\mathcal{A}$ and $\mathcal{B}$. What do we need to know about them to get started? $\mathcal{A}$ and $\mathcal{B}$ have common knowledge[3] of their rationality.

Next, we have $\mathcal{A}$ uttering an indicative sentence φ assertively in discourse situation d to convey some information p to $\mathcal{B}$. $\mathcal{B}$ attempts to interpret $\mathcal{A}$'s utterance in d. When $\mathcal{A}$ utters φ, $\mathcal{B}$ uses his knowledge of the language to get at its meaning $m(\varphi)$, which is the collection of possible contents of the utterance.

---

[2] A game is a structure where all agents have common knowledge of this structure. I generalize this notion of a game to what I call a *strategic interaction* where agents no longer have common knowledge of the structure. More about this later.

[3] Common knowledge between $\mathcal{A}$ and $\mathcal{B}$ of a fact f is the requirement that $\mathcal{A}$ knows f, $\mathcal{B}$ knows f, $\mathcal{A}$ knows $\mathcal{B}$ knows f, $\mathcal{B}$ knows $\mathcal{A}$ knows f, and so on, ad infinitum. This concept was first introduced by Lewis (1969) and Schiffer (1972). It has since become a staple of game theory.

Some aspects of the utterance will be public.[4] The agent architecture is public before the utterance and the sentence uttered will be publicly available to both agents after the utterance. The meaning of the sentence, being a linguistic constant, will also be assumed to be public. Other aspects will, in general, be private, like the beliefs and intentions of the speaker and addressee.

Our initial problem is to spell out sufficient conditions for $\mathcal{A}$ to *communicate* p to $\mathcal{B}$ by uttering φ in d.

5. Strategic Inference

Sometimes it is helpful to embed a problem in a larger problem either to get a better perspective on it or to solve the larger problem as a way of solving the smaller problem. We will do it for the first reason, to get a better perspective on communication.

We can embed communication in the larger picture of information flow developed by Dretske (1981), Barwise and Perry (1983), and Barwise (1997).

Reality can be viewed as consisting of situations linked by constraints. It is the constraint between two situations that makes one situation carry information about (naturally or nonnaturally mean, in Grice's sense) another situation. A smoky situation involves a situation with fire in it. This is the constraint we describe when we say "Smoke means fire." This is an instance of natural meaning. An utterance situation with the sentence "There is a fire" also involves a situation with fire in it. This is the constraint we describe when we say the speaker means something is on fire. This is an instance of nonnatural meaning. In the first case we would write $s_1 \implies s_2$ and in the second, $u \implies s_2$. An agent who perceives the first situation (either smoke or the utterance) and who knows the relevant constraint (either natural or nonnatural) can infer the existence of the second situation s_2.

Though the two constraints are quite different, I will argue later that the terms "natural" and "nonnatural" are perhaps not the best way to capture this distinction. The distinction originates with the classical distinction between "natural" and "conventional" but Grice introducd the term "nonnatural" to accommodate nonconventional transfers of information that are not natural, like nonconventional gestures, drawings, sounds, and the like.

[4] Public knowledge is interchangeable with common knowledge, more or less. See Barwise (unpublished).

Given a group of agents, or distributed system as computer scientists call it, there will be all kinds of information flows. A communication is a special type of information flow between agents. Indeed, it is the type of flow that language makes relatively easy to accomplish, but that is not exclusive to language.

In what way is communication special? While smoke indicates fire, it doesn't *communicate* fire.[5] Why not? Because it doesn't have an *intention* to do so. Why is an intention required? Because our intuitive notion of communication is that it is something only agents can do. This certainly rules out all inanimate objects, except maybe sufficiently sophisticated computers.[6] What about insects like bees, however? We do say that bees communicate even though they don't have intentions.[7] I suppose we have to admit two differing intuitions here. One intuition is that communication is the mere transmission of information, the other is that it is something agentive and more complex. The problem with the first notion is that the intuitive distinction between animate and inanimate transmission[8] also collapses, and all information flows become communicative. Besides, there is the intuition that human communication is different from mere information flow. How do we do it justice? By bringing in intentions to start with.

Grice brought in a lot more conditions as counterexamples to proposed definitions piled up, but the starting point was the requirement that the speaker have an intention to convey the relevant information. One important condition Grice introduced was that this intention be recognized by the addressee. This was required because if $\mathcal{A}$ were to leave a sign (e.g. someone's, say $\mathcal{C}$'s, handkerchief) for $\mathcal{B}$ at the scene of a crime to indicate that $\mathcal{C}$ had been there, $\mathcal{B}$ may not be able to infer that $\mathcal{A}$ had intended to put it there. Intuitively, this is not a case of "full" communication. Something is missing, and this, Grice suggested, is the recognition of $\mathcal{A}$'s intention. Grice, and Strawson and Schiffer after him, developed this line of reasoning considerably, adding more conditions to the definition of communication.

I will sidestep this reasoning involving definitions and counterexamples, and jump directly to building a model and definition of communication. What we need for the moment from the foregoing is that communication involves both the speaker and addressee jointly inferring

[5] That is, it doesn't communicate that there is a situation with fire in it.

[6] Though here it may be the programmer's intentions that are relevant, which allows us to impute intentions to computers. But more on this later.

[7] This itself is perhaps a moot point.

[8] Bees occupy a middle ground.

various things about each other. I will call this joint two-sided inference a *strategic inference*.

My basic insight is that all intended information flows between agents involve a strategic interaction between them. When the strategic interaction is common knowledge between the agents, that is, when it is a *game* (with a unique solution), the flow will be communicative. Roughly then, $\mathcal{A}$ communicates to $\mathcal{B}$ just in case there is a game between $\mathcal{A}$ and $\mathcal{B}$. It is this insight I will make precise in my definition of communication.

I argue this by first developing a detailed account of one strategic inference. In my view, every utterance involves many separate acts and corresponding strategic inferences. For example, communication typically involves a referential act. Figuring out the reference will then involve a strategic inference. In general, each bit of information communicated will require its own strategic inference. So any complete utterance involves a system of simultaneous strategic inferences. These inferences have to be simultaneous because, in general, they codetermine each other. An utterance of "Mary had a little lamb" will require inferring the designata of each of the five words in the sentence, not to mention its internal structure. Only then is it possible to get at the content of the utterance. No individual word has any priority in this determination. That is, there may be interactions among the various strategic inferences. And the embedding circumstances play a vital role in each inference. Mathematically, this amounts to a system of simultaneous equations.[9]

To keep things simple, I will focus on just one strategic inference in isolation. I will assume $\mathcal{B}$ has the partial information obtained from all the other inferences. $\mathcal{B}$'s problem is then to use this partial information together with the utterance situation to get to the intended content.

Consider as an example the sentence "Every ten minutes a man gets mugged in New York." This is a familiar type of ambiguity, typically viewed as an ambiguity between two possible quantifier orderings. One reason for this type of choice is that it is widespread in language and in the literature. There are other ambiguities too in this sentence. For example, "minutes" is ambiguous between the temporal meaning and the minutes of a meeting. New York is ambiguous between the city and the state. "Every ten minutes" is also vague because it usually indicates "about every ten minutes." But I will consider only the quantifier orderings.

A successful strategic inference requires a number of assumptions involving rationality, the agents' intentions, and their knowledge and beliefs. These assumptions will be our sufficient conditions. An important consequence of the analysis is that the content communicated depends not

[9] It is possible to deal with all the inferences as one big inference, of course.

only on what was uttered but also, crucially, on what the speaker *might* have uttered but chose not to, and on their shared information about these choices.

I will build up a strategic inference step by step from the discourse situation d. This will make the role of the assumptions clearer and suggest ways in which the construction can be generalized or modified to include other complexities. The constructed structure $g(\varphi)$ turns out to be a new kind of game. I call it a game of partial information. The content communicated will then be given by the Pareto-undominated Nash equilibrium of the game.[10]

6. Some More Examples First

Lest the reader think the model applies only to the example above or only to this type of ambiguity involving quantifier orderings, I consider in this section many different types of examples of resolution and ambiguity to which the model applies.

1. I'm going to the bank (lexical ambiguity)
2. He saw her duck (structural ambiguity)
3. It is 4 (indexical resolution)
4. He is eating (pronominal resolution)
5. Bill said to Bob that he would join him today (double anaphora)
6. The book is highly original (noun phrase resolution)

Most of the choices for interpretation in the above examples are fairly obvious. Once again, there are multiple ambiguities and resolution problems in each. I have identified which problem I'm considering in parentheses. My list is far from exhaustive. Indeed, the game-theoretic model applies to any and every type of communication, including visual, gestural, aural, and even olfactory and tactile ambiguities. It applies to all actions.

The reader should keep these other examples in mind as we proceed with the main example. This will make it easy to see how the model can be adapted to these other examples and indeed, to any example of communication.

[10] I explain these terms below.

7. The Main Example

Suppose $\mathcal{A}$, after having picked up the information in a recent newsletter of The Muggees Association of New York (i.e. M.A.N.Y.), says to $\mathcal{B}$ in situation d:

"Every ten minutes a man gets mugged in New York." (φ)

$\mathcal{A}$ could mean either that some person or other gets mugged every ten minutes (call this p) or that a particular man gets mugged every ten minutes (call this p'). It must be the situation d that enables $\mathcal{B}$ to disambiguate φ.

Though both interpretations are possible in different circumstances, it seems plausible to say that in d $\mathcal{B}$ would infer p as the intended content. In fact, we could say that $\mathcal{A}$ *communicates* p to $\mathcal{B}$.

I will make two sets of assumptions to explain this disambiguation and its communication. The first set applies to all situations of interest, more or less. They have to do with the architecture (or "nature") of communicating agents generally. The second involves more specific circumstantial assumptions, pertaining to the discourse situation d.

For the first set, called the Background Assumptions, we assume that both $\mathcal{A}$ and $\mathcal{B}$ are rational agents. (Grice also assumes rationality, but not in its choice-theoretic form.) Moreover, this is common knowledge. This is important because the agents would act differently if they didn't know they shared a common architecture.[11]

$\mathcal{L}$ is a shared language. m is its meaning function. It is a mapping from sentences to propositions. I said earlier that meaning is constant across situations. That is why this assumption is in the background.

The Circumstantial Assumptions contain, in this particular example, the assumption that $\mathcal{A}$ has the intention to convey[12] p to $\mathcal{B}$.

Next, $\mathcal{A}$ utters φ publicly. After all, the process has to get off the ground.

$\mathcal{B}$ must have a corresponding intention to intepret φ. Without it, he will not play his interpretive part.

[11] They would have to consider alternative architectures and so on.

[12] We cannot say "intends to communicate" here because that would imply a circularity later when we define communication. The word "convey" just means "transfer." This is in fact the kind of simple intention we have when we communicate.

TABLE I

Summary of Assumptions

Background Assumptions

1. $\mathcal{A}$, $\mathcal{B}$ are rational.
2. $\mathcal{L}$ is a shared language.
3. m is a function from $\mathcal{L}$ to the power set of the collection of propositions. I call it the meaning function of $\mathcal{L}$ or just the meaning of $\mathcal{L}$.
4. The above assumptions are common knowledge between $\mathcal{A}$ and $\mathcal{B}$.

Circumstantial Assumptions

1. $\mathcal{A}$ intends to convey p.
2. $\mathcal{A}$ utters φ.
3. $\mathcal{B}$ intends to interpret φ.
4. $\mathcal{B}$ receives and interprets φ.
5. $m(\varphi) = \{p, p'\}$.
6. p' is relatively unlikely.
7. Expressing p, p' unambiguously takes greater effort than expressing them ambiguously.
8. All of the above except (1) and (3) are common knowledge.

$\mathcal{B}$ must also receive and interpret the utterance and this must be public.[13] Without publicity, p won't become public at the end of the communication.[14]

$$m(\varphi) = \{p, p'\}.$$

Also, p' is relatively unlikely, and expressing p unambiguously takes greater effort than expressing it ambiguously. The meaning and use of these assumptions will become clearer as we proceed.

Except for $\mathcal{A}$'s and $\mathcal{B}$'s intentions, the assumptions are common knowledge between $\mathcal{A}$ and $\mathcal{B}$.

[13] These four assumptions (speaker and addressee intention and utterance and reception/interpretation replace Grice's principle of cooperation. That is, if agents act in the right way, communication can occur, but if they don't, communication can't occur. There is nothing that forces them to cooperate, as Grice required.

[14] Usually, a copresent addressee responds with movements of the head (nodding) and eye contact, indicating that he is attending to the conversation.

The two sets of assumptions taken together will be called the $\mathcal{BC}$ assumptions. The Background Assumptions hold in the background situation B and the Circumstantial Assumptions hold, of course, in d. B is a part of d.

$\mathcal{A}$ and $\mathcal{B}$ need not be persons. They can be suitably equipped artificial agents.

My claim then is that if all the $\mathcal{BC}$ conditions above are satisfied $\mathcal{A}$ will *communicate* p to $\mathcal{B}$.

8. Games Rational Agents Play

We are now ready to build a game-theoretic model. There are many different types of games depending upon the application we have in mind. Unfortunately, none of these is directly suitable for us. The closest type is games of incomplete information.[15]

I will keep this model in mind, but adapt it to our purposes. The new type of game we get is something I call a game of *partial* information. To construct this, I will start more or less from scratch.

There are different ways to interpret game-theoretic models. I think it is unrealistic for many applications to imagine that agents play games explicitly. We certainly don't seem to be playing games explicitly when we communicate.

It seems better to view the game as a model of a class of constraints that captures the underlying logic of communication. Modus ponens captures the logic of a deductive inference without implying anything about how agents actually act when they arrive at a warranted conclusion. The game I will construct describes a valid strategic inference without implying anything about how agents arrive at the correct interpretation of an utterance.

Despite this, it is convenient as a façon de parler to talk as if agents are actually performing the relevant steps of a strategic inference. This makes the model more intuitive.

9. The Model

We have the background B and discourse situation d. B contains the background assumptions and d contains the circumstantial assumptions. We want to derive the equation "agent architecture $\oplus$ meaning $\oplus$ $d = p$," and

[15] See Myerson (1995).

justify the claim above. The first two components of the equation are the elements of B.

I will start with $\mathcal{A}$'s intention because that comes first. $\mathcal{A}$ intends to convey p to $\mathcal{B}$. There are many ways $\mathcal{A}$ could do this, but the most convenient (i.e. efficient) is to use $\mathcal{L}$. $\mathcal{L}$ has many sentences that will do the job. One may be φ. I say "may be" because $\mathcal{A}$ doesn't know at the outset if the ambiguity can be resolved. Another is μ which is "Every ten minutes some man or other gets mugged in New York." μ is unambiguous (with respect to quantifier orderings) so that $m(\mu) = \{p\}$. In fact, any unambiguous sentence will do, but we will stick with μ.

So $\mathcal{A}$ has two possible actions. We can collect these to form his choice set $C(p)$.[16]

Since $\mathcal{A}$ is rational, he will evaluate the consequences of both actions and then choose. I will first consider φ.

10. THE LOCAL GAME

We know from the assumptions that $\mathcal{B}$ receives φ if uttered. He then forms an intention to interpret it. This puts him in a predicament because φ is ambiguous. After all, $m(\varphi) = \{p, p'\}$ by assumption. At this stage, $\mathcal{B}$ has no way to choose either content. This situation can be modelled by the tree (actually forest, because it is two simple trees) in Figure 1.

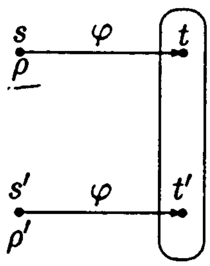

Figure 1. Stage One of Local Game $g(\varphi)$

The two initial nodes s and s' represent $\mathcal{A}$'s intention to convey p or p'. s and s' are situations containing the relevant intentions. $\mathcal{A}$ knows he is in s and not in s'. That is, he knows his own intention. $\mathcal{B}$ knows only that

[16] We could consider more actions, but why complicate things unnecessarily? In any case, we can already write down his entire choice set in symbols. It is $C(p) = \{m^{-1}(P) | P \in m(\mathcal{L})$ and $p \in P\}$.

either s or s' could be factual. In fact, he knows this only after $\mathcal{A}$ utters φ, something that will be important later.[17]

The two branches emanating from s and s' are the same action of uttering φ. φ is a possible action for both intentions.

If $\mathcal{A}$ utters φ in either situation, he moves to a new situation t or t'. Once again, $\mathcal{A}$ can distinguish between t and t', but $\mathcal{B}$ can't. $\mathcal{B}$'s epistemic indigence is represented by an oval and $\{t, t'\}$ is called an information set.

All this is common knowledge because they started with common knowledge of $m(\varphi)$ and we assumed the utterance and its reception are public.

One thing remains: the two ρs. Since it is common knowledge that p is more likely than p', the agents can take it to be common knowledge that $\mathcal{A}$ probably intends to convey p rather than p'.[18] We can, for specificity, take $\rho = 0.9$ and $\rho' = 0.1$. That is, $\mathcal{A}$ and $\mathcal{B}$ assess the same probabilities. This is not necessary of course, but I will assume it for convenience. These probabilities can be objective or subjective, in general. But they will usually be subjective because objective information will often not be available.

$\mathcal{B}$ has to choose an interpretation at this point. He has the same two choices at t and t', either p or p', as shown in Figure 2.

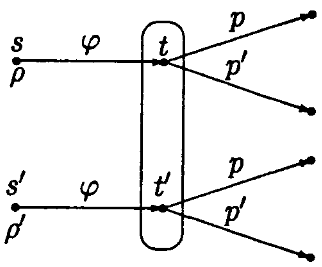

Figure 2. Stage Two of Local Game $g(\varphi)$

Knowing that t or t' is factual allows $\mathcal{B}$ to infer that s or s' is factual. This is the same as saying that he infers $\mathcal{A}$'s possible intentions. Importantly, $\mathcal{A}$ does not have to *intend* this recognition of $\mathcal{A}$'s primary intention. It just happens as a logical consequence of rationality. After all, $\mathcal{B}$ wants to interpret $\mathcal{A}$'s utterance, and to do this he needs to figure out his preferences. This requires him to infer $\mathcal{A}$'s intentions.

---

[17] This is what makes the game we are constructing a game of partial rather than incomplete information.

[18] In general, there is a difference between these two probabilistic situations, and it is only the second, involving $\mathcal{A}$'s intention, that matters. In the absence of any special information, one situation does inform the other, as above.

If he is in t, p is the preferred choice, and if he is in t', p' is the preferred choice. Unlike standard game theory however, these preferences are not externally given, but derived internally from the relevant intentions. They are endogenous, not exogenous. This requires $\mathcal{B}$ to recognize $\mathcal{A}$'s possible intentions, as required by Grice, but in a generalized setting where there is ambiguity. Later, we will see that this is not really required, and that payoffs can be exogenously given.

At this stage, it is common knowledge that if $\mathcal{B}$ is in t, the intended and preferred interpretation is p, and if he is t', it is p'. It is common knowledge because the agents started with common knowledge and everything that followed is a consequence of this common knowledge.

Unfortunately, $\mathcal{B}$ has as yet no clue about his location. He knows only that he is either in t or t'.

Because $\mathcal{B}$ cannot tell where he is in the tree, it is not clear how he should choose: the optimal action is different at t and t'. $\mathcal{A}$ is probably conveying p, and this is common knowledge between them. But this additional information does not enable $\mathcal{B}$ to eliminate the uncertainty involved.

$\mathcal{A}$ knows of course that it is t that results from his utterance of φ. But $\mathcal{A}$ also knows that for $\mathcal{B}$ t and t' are in the same information set. In fact, it is easy to see that the $\mathcal{B}C$ assumptions imply that the information represented by the tree above becomes common knowledge between them once $\mathcal{A}$ utters φ.

As it stands, the preference ordering above needs to be strengthened into a von Neumann–Morgenstern (N–M) utility function. Each interpretation can then be assigned a numerical value.[19]

To keep things simple, I will assume that $v(s, \varphi, p) = v(s', \varphi, p') > v(s, \varphi, p') = v(s', \varphi, p)$, where v is the payoff function. Information carries the same utility and misinformation the same disutility. To be specific, I will assign the two numbers 10 and -10 respectively.

In general, there will be two payoff functions $v^{\mathcal{A}}$ and $v^{\mathcal{B}}$ for $\mathcal{A}$ and $\mathcal{B}$. I will assume that $v^{\mathcal{A}} = v^{\mathcal{B}} = v$.

It is quite possible to entertain different numbers and different payoffs for $\mathcal{A}$ and $\mathcal{B}$. We will be doing so later. Right now it would just complicate things to consider all these possibilities.[20]

Figure 3 makes $\mathcal{B}$'s dilemma clear. He cannot tell where he is and the payoffs are symmetric and conflicting. If he chose p, he would get 10, if

[19] All (positive) linear transformations of the payoff function are considered equivalent.

[20] Communication is usually embedded in other actions. It is possible to derive a numerical value of information from these embedding actions. If we do this, we do not need to assume arbitrary values as I have done above. But this valuation procedure may not be as general.

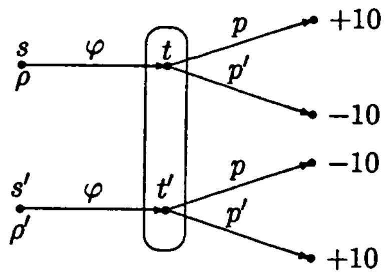

Figure 3. Stage Three of Local Game $g(\varphi)$

he was in t. But if t' were factual he would end up with -10. And the same problem crops up with a choice of p'. If he were to choose one or the other randomly, say by tossing a (fair) coin, he would get an expected payoff of 0, certainly much lower than his maximum possible payoff!

If there was nothing else to the discourse situation, it would not be possible for $\mathcal{B}$ to disambiguate φ in d. $\mathcal{A}$ and $\mathcal{B}$ need to compare this ambiguous utterance against an unambiguous one, to ensure that it is more efficient. In other words, they need to consider sentences μ and μ' where μ' is something like "Every ten minutes a particular man gets mugged in New York". $m(\mu) = \{p\}$ and $m(\mu') = \{p'\}$.

This enables $\mathcal{B}$ to construct the model of their interaction in Figure 4, once $\mathcal{A}$ has said φ.

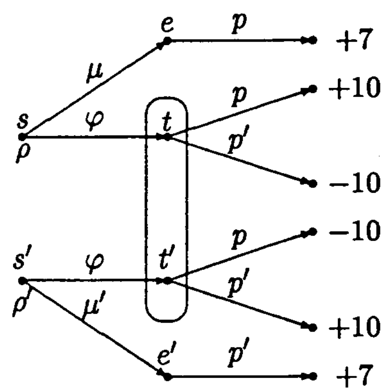

Figure 4. Local Game $g(\varphi)$

$\mathcal{B}$'s only choice of interpretation at e is p and at e' is p'. This is because μ and μ' are unambiguous.

How do we assign payoffs to the new paths? The only difference between $v(s, \varphi, p)$ and $v(s, \mu, p)$ lies in the costs involved. μ and μ' are longer than φ and so costlier.[21] So $v(s, \mu, p) < v(s, \varphi, p)$. Also, the additional cost involved is small relative to the difference between information and misinformation. I will take this payoff to be 7. The same considerations apply to the path with μ' and we set $v(s', \mu', p') = 7$.

This is $\mathcal{B}$'s model of their interaction. But because it is constructed from their common knowledge the entire model is in fact available to both $\mathcal{A}$ and $\mathcal{B}$ (to $\mathcal{B}$ only after φ has been uttered). This makes the model itself common knowledge (also only after φ has been uttered). This structure, denoted by $g(\varphi)$, is a new type of game that I will call a game of partial information. I will also call it a local game because it will turn out to be part of a larger structure called a global game. $g(\varphi)$ is then a local game of partial information.

We began this discussion to look at the possible consequences of $\mathcal{A}$'s uttering φ in d. We now have part of the answer to this question. Upon uttering φ, (the information contained in) $g(\varphi)$ becomes common knowledge between $\mathcal{A}$ and $\mathcal{B}$.

$g(\varphi)$ is the choice situation $\mathcal{B}$ faces, and so is the situation $\mathcal{A}$ has to consider before he actually chooses φ. Is there enough information in $g(\varphi)$ for $\mathcal{B}$ to eliminate t' as a possible location and so be able to choose p? This is $\mathcal{B}$'s problem. But it is also therefore $\mathcal{A}$'s problem, because $\mathcal{A}$'s optimal choice will depend in part on whether $\mathcal{B}$ has enough information to solve this problem. This is what makes their interaction strategic, since each agent has to consider the other's situation.

I will show in the next section that it is possible for $\mathcal{B}$ to eliminate t'. The argument for this is intricate and it seems best to discuss it after a complete account of the choice structure. For now assume that $\mathcal{B}$ is able to eliminate t' and choose p as his preferred interpretation. Since $g(\varphi)$ is common knowledge between them, it seems reasonable to assume that $\mathcal{A}$ has access to this reasoning. This enables $\mathcal{A}$ to anticipate $\mathcal{B}$'s choice of p. As a result, both receive a payoff of 10. We will say that the value of $g(\varphi)$, $v[g(\varphi)]$, is 10.

$\mathcal{A}$'s choice structure is now easy to see. For every sentence ψ in $C(p)$ there is a corresponding local game $g(\psi)$. For example, $g(\mu)$ is the trivial game in Figure 5.

$g(\mu)$ clearly has a value of 7.

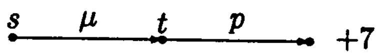

Figure 5. Local Game $g(\mu)$

[21] In general, costs will depend upon a number of factors including length and grammatical structure.

11. The Global Game

$\mathcal{A}$ has to choose the sentence ψ that yields the highest value $v[g(\psi)]$. This problem is represented in Figure 6.

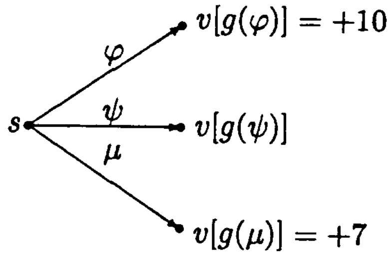

Figure 6. Global Game $G(p)$

If A considers only φ and μ, his optimal action is to utter φ. I will call this larger structure (in which the various local games are embedded) the global game (of partial information) and denote this by $G(p)$.[22] I have implicitly assumed that every local game has a value and this is something I justify later.

But we now have a complete picture of the interaction that makes it possible for $\mathcal{A}$ to communicate p to $\mathcal{B}$.

We have come full circle. I started by considering $\mathcal{A}$'s intention to convey p. That took us on a long path, along which we constructed $g(\varphi)$, and which ended with why $\mathcal{A}$ should utter φ and $\mathcal{B}$ should intepret it as conveying p. It is this dual, "two-sided" interaction, or strategic rationality, that makes communication possible. Grice and Schiffer focus on speaker meaning and so miss out on the strategic aspect of communication, and also, incidentally, of speaker meaning, as we will see. It is to make this strategic dimension clear that I have started with communication rather than speaker meaning.

I should point out an important asymmetry in $G(p)$ however. $\mathcal{B}$ gets to consider only one local game, the one constructed from $\mathcal{A}$'s optimal choice. $\mathcal{A}$, on the other hand, has to consider all the local games issuing from his choice set and then choose the best one. It is not possible or necessary for $\mathcal{B}$ to consider $C(p)$.

It is now time to justify the two provisional statements made earlier. The first concerns the reasoning that $\mathcal{B}$ can employ to eliminate t' upon

[22] This structure can be thought of as a two-stage game for $\mathcal{A}$ and a one-stage game for $\mathcal{B}$. But it isn't quite this either. That is why I have given it a new name, a game of partial information.

receiving φ. The second has to do with the existence of a value for every $g(\psi)$ that $\mathcal{A}$ might consider.

12. Solving Games Rational Agents Play

Before turning to solving $g(\varphi)$ it is worth making a general distinction in the context of our game-theoretic analysis. It is important to distinguish between the model $g(\varphi)$, the different sorts of interactions $g(\varphi)$ could be a model of, and consequently the different ways in which the model could be analysed. A similar distinction is explicitly made by Aumann (1985) and also by Kreps (1985). I have already made it implicitly by separating the model from its analysis.

Solving a game involves finding a pair of "strategies" (one for each player in a two player game) that is in some sense optimal. It appears that there are many different ways to solve the same abstract game, each way being more or less appropriate depending on the particular interpretation we give to the game. A solution concept that may seem appealing in an economic or political context may not be as appealing in a discourse situation; even two different economic contexts or two different discourse situations may provide different grounds for accepting or rejecting a proposed solution.

One persistent problem with many solution concepts is the existence of multiple solutions.[23] This multiplicity is troublesome in most situations because if two or more strategy pairs are optimal then players may not know which strategy to play and choosing different or nonmatching optima may result in a suboptimal outcome.

To take Schelling's (1960) example, if two people have to meet in New York and are not in a position to communicate with each other, any place in the city would do as long as they both choose the same spot. This is a (coordination) game with multiple "solutions". But such solutions are obviously not particularly helpful in prescribing a course of action. Of course, in situations like these, in the absence of other relevant information (perhaps both players are natives and this is common knowledge between them making Grand Central Station a "salient" spot; or it is common knowledge that they are both tourists which might make the Empire State Building "focal") one should not expect unique solutions. In fact, Lewis (1969) uses this nonuniqueness as a necessary condition in the definition of conventions. It is the existence of multiple rational ways of doing something that

[23] This is an important part of the reason why so many solution concepts have been investigated.

makes it worthwhile for the players involved to agree on a convention. But it turns out that most solution concepts also allow many unintuitive and implausible solutions to slip past their restrictions, at least under some interpretations of the abstract game under consideration. Exactly which (and how many) solutions are intuitively warranted in a game seems to depend on other features of the particular context being modelled.[24]

I made the distinction above to keep open the possibility of using different solution concepts for the same local game $g(\varphi)$. The solution concept I use here combines one of the more popular solution concepts called Nash equilibrium with the idea of Pareto dominance.

To spell out the concept of a Nash equilibrium we first need to say what a strategy is. A strategy prescribes actions for a player in all possible situations where he has to act. It is essentially a function from the set of all the decision nodes of a player to a set of actions. For example, the function $\{(s, \varphi), (s', \varphi)\}$ is one of $\mathcal{A}$'s strategies in the game $g(\varphi)$. $\mathcal{A}$ has only two possible choice situations and a strategy specifies his choices in both of them. Obviously, $\mathcal{A}$ has exactly four strategies in this game. It is important to note that a strategy for $\mathcal{A}$ involves a specification of what he would do in s' even though he knows that s' isn't factual. This is necessary because $\mathcal{B}$ needs to consider what $\mathcal{A}$ might do in s' and so $\mathcal{A}$ needs to consider what $\mathcal{B}$ might do if he takes into account the possibility that $\mathcal{A}$ might be choosing an action in s'.

$\mathcal{B}$'s strategies involve a slight complication and with it a small refinement of the rough definition of strategy above. $\mathcal{B}$ has four choice situations to consider t, t' and e, e'. Since t and t' belong to the same information set, $\mathcal{B}$ cannot distinguish between the two and so $\mathcal{B}$'s choices at t and t' have to be constrained to be the same. That is, the correct domain for the strategy of a player is not the set of all decision nodes but the set of all information sets. In $\mathcal{A}$'s case, the domain of a strategy will contain the singleton information sets $\{s\}$ and $\{s'\}$.

$\mathcal{A}$ has the following four strategies:

1. $s \mapsto \phi,\ \ s' \mapsto \mu' \ = \langle \phi, \mu' \rangle$
2. $s \mapsto \phi,\ \ s' \mapsto \phi \ = \langle \phi, \phi \rangle$
3. $s \mapsto \mu,\ \ s' \mapsto \phi \ = \langle \mu, \phi \rangle$
4. $s \mapsto \mu,\ \ s' \mapsto \mu' \ = \langle \mu, \mu' \rangle$

And $\mathcal{B}$ has the following two strategies:

24 In many discourse situations nonuniqueness has in fact a different sort of interpretation, making possible the extraction of information not otherwise available. If a local game has multiple solutions and if $\mathcal{A}$ chooses to play it then some sort of ambiguity is left unresolved in the communication and this may convey to the addressee that the ambiguity was intentional for some reason or other (punning, for example).

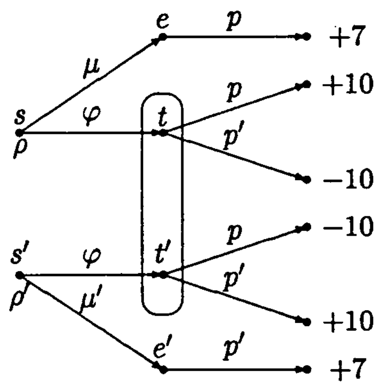

Figure 7. Local Game $g(\varphi)$

1. $e \mapsto p,\ \ t, t' \mapsto p,\ \ e' \mapsto p'\ \ = \langle p, p, p' \rangle$
2. $e \mapsto p,\ \ t, t' \mapsto p',\ \ e' \mapsto p'\ \ = \langle p, p', p' \rangle$

We can simplify the notation if we specify the information sets of a player in some predefined order, say, from top to bottom with respect to the tree. We can further simplify the representation of $\mathcal{B}$'s strategies by explicitly mentioning only those decisions that represent a "real" choice. Thus, $\mathcal{B}$'s choices are constrained to be p and p' at e and e' respectively, so we need not mention these explicitly. Note that we have already made an implicit simplification in our specification of the strategy functions above. The values of the two functions have been represented as either sentences or propositions. They are actually the corresponding actions, either of uttering the sentence in question or of interpreting the uttered sentence as expressing the relevant proposition.

If a is a strategy of $\mathcal{A}$'s and b of $\mathcal{B}$'s then the pair $\langle a, b \rangle$ is called a joint strategy. This gives exactly eight strategies in the game $g(\varphi)$ above. These constitute the strategy space. Intuitively, the unique solution of this game is $\langle \varphi, \mu', p \rangle$.

What I have defined above is the concept of a *pure* strategy. In general, players can *mix* strategies by randomizing on their pure strategy sets. The strategy space then is the set of ordered pairs of probability distributions, one for each player. The argument below can be easily extended to this larger space so I will restrict my remarks to pure strategies. 'Strategy' will henceforth mean pure strategy.

A strategy is a Nash equilibrium if no player has an incentive to deviate unilaterally from this strategy. Unilateral deviation by a player is deviation keeping the strategies of other players fixed. Consider $\langle \mu, \mu', p \rangle$. $\mathcal{B}$ certainly has no incentive to deviate from p to p'—it makes no difference which of the two he chooses because $\mathcal{A}$'s strategy doesn't allow the information set $\{t, t'\}$ to become factual. But if $\mathcal{A}$ deviates unilaterally to $\langle \varphi, \mu' \rangle$ then he certainly does better. This eliminates $\langle \mu, \mu', p \rangle$. A quick run through the strategy space will show that only $\langle \varphi, \mu', p \rangle$ and $\langle \mu, \varphi, p' \rangle$ are Nash equilibria. Call them N_1 and N_2.

This is probably the most widely used solution concept in the theory of (noncooperative) games. It is worth pointing out though that this criterion cannot be directly deduced from the axioms of utility theory that characterize the behaviour of individual rational agents (see Bernheim (1984), Brandenburger and Dekel (1985)). Its plausibility lies in its being a necessary condition for rationality if it is already assumed a priori by the players that some rational prescription for action exists in the game.

I will use the Nash criterion without further justification here. We still have to face the fact that there are two Nash equilibria N_1 and N_2 only one of which is intuitively plausible. We need further conditions to differentiate between N_1 and N_2.

To solve this multiple equilibrium problem I will use the idea of Pareto-dominance. It says simply that of two strategies in a game, if one results in higher payoffs for all players concerned, the other can be eliminated. Though this appears to make perfect intuitive sense there is a problem with it because it implicitly assumes some degree of correlated action (deviation) among players, something that requires additional assumptions to be warranted in a noncooperative game. In fact, there is often a conflict between the Nash criterion and the Pareto dominance criterion (as evinced, for example, by the Prisoner's Dilemma).

I will use Pareto dominance as a second-order criterion. First, we determine the set of Nash equilibria. Then we apply the Pareto criterion to this set. This ensures that all solutions satisfy the important Nash property that there is no incentive to deviate. That after all is what justifies calling it an "equilibrium" strategy. And this second-order way of using it to eliminate counterintuitive Nash equilibria is easier to justify.

Applying the Pareto criterion to the Nash set ousts N_2. The expected payoff from N_1 to both players is $0.9(10) + 0.1(7) = 9.7$. The expected payoff to both players from N_2 is $0.9(7) + 0.1(10) = 7.3$. This implies that N_1 Pareto-dominates N_2 and that both players can with certainty choose N_1. Since s is factual, this results in $\mathcal{A}$ saying φ and in $\mathcal{B}$ choosing p rather than p'. (Note that the optimal expected payoff is 9.7, much greater than what is obtainable by tossing a coin (i.e. 0) and foregoing strategic reasoning, at least so long as its costs are ignored.)

This completes our discussion of how the game $g(\varphi)$ is solved. A strategy that satisfies this solution concept is called a Pareto–Nash equilibrium. N_1 turns out to be the unique Pareto–Nash equilibrium of $g(\varphi)$ and so its value $v[g(\varphi)]$ is 10, as I had asserted above. This is one more place where the difference between games of incomplete information and games of partial information shows up. Not only do they differ in their qualitative structure but also in certain quantitative aspects. The relevant value of the local game for $\mathcal{A}$ is 10, not 9.7, because $\mathcal{A}$ knows which situation he is in. However, this value is derived from the solution to the local game where the expected value is what counts.

Note that if we assume equal instead of skewed probabilities ("A comet appears every ten years") we are unable to eliminate the second solution N_2 and this squares with our intuition as well. Both interpretations seem equally plausible in a general context. In this case the optimal solution would seem to be to spell out the content literally by using μ as φ remains ambiguous. This is interesting because it shows how to justify the use of a more elaborate expression to avoid an ineliminable ambiguity. Also, if we have skewed probabilities as above, but s' is factual rather than s, then again the optimal strategy is to spell things out by using μ' instead of μ.

How do we define the value of a game that has multiple equilibria? In the usual way, as the expected value of the set of multiple values. However, it is unclear what distribution we should use in evaluating expected values. We will assume that, in the absence of any further information each equilibrium strategy is equally likely. (This equiprobable criterion is known to have many weaknesses, but it is perhaps less objectionable in such higher-order contexts.) In the case of the comet, where s and s' have the same likelihood of occurrence, $\mathcal{A}$ should assign equal probabilities to both solutions in the absence of any information about a preference that $\mathcal{B}$ might have for one or the other. Thus, if $\mathcal{B}$ plays N_1 they get 10 and if he plays N_2 they get -10, and the expected value of this set of two values with respect to a uniform distribution is just 0. This is the value that $\mathcal{A}$ should consider in making his optimal decision in the global game.

We need to make certain that every game $g(\psi)$ that $\mathcal{A}$ might consider does in fact have a value. This is guaranteed to us by a theorem of Nash's (1951). Every game has at least one Nash equilibrium in the larger space of mixed strategies. In fact, it is easy to show that every game of pure coordination (games in which players have identical payoff functions) has an equilibrium in pure strategies. In either case, this guarantees in turn that every $G(p)$ has a solution. (Actually, the step to the existence of a solution for every $G(p)$ isn't quite so immediate. It requires a consistency condition between the local and global games to be satisfied. This can be found in Parikh (1990).)

This completes my analysis of why $\mathcal{A}$ chooses to say φ and correspondingly of how $\mathcal{B}$ comes to choose the right interpretation p in the discourse situation d.

13. DEFINITIONS

We have now discussed more or less completely one set of sufficient conditions for communication. How do we go from here to a set of necessary and sufficient conditions?

The trick is to abstract and generalize from the BC assumptions by dropping unnecessary assumptions.[25] I have prepared the way for this by constructing $g(\varphi)$. What we do is drop the assumptions about cooperation, probability and payoffs that enabled us to construct $g(\varphi)$ and require instead that we merely have a game with a unique solution. The details of the game and its solution process could be anything at all.

Let me state it without more ado. We start by noting the existence of a certain function m (with domain the set of actions and range the power set of the set of propositions) which establishes a connection between an utterance of φ and a set of propositions P. This function m is the *meaning* of the utterance.

DEFINITION 1. $\mathcal{A}$ communicates p to $\mathcal{B}$ by producing φ iff

1. $\mathcal{A}$ intends to convey p to $\mathcal{B}$.
2. $\mathcal{A}$ utters φ.
3. (2) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.
4. $\mathcal{B}$ intends to interpret φ.
5. $\mathcal{B}$ interprets φ.
6. (5) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.
7. $p \in m(\varphi)$.
8. (7) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.
9. $g(\varphi)$ has the unique solution $\langle \varphi, p \rangle$ for $\mathcal{A}$ and $\mathcal{B}$.

A few comments are in order. First, I need to say what $g(\varphi)$ is precisely, but it is beyond the scope of this paper. Essentially, as I said above, I drop the assumptions about probabilities and payoffs in Table 1, and allow these

[25] This is exactly like dropping extraneous facts about Euclidean distance to get a definition of metric space or similar abstractions in mathematics.

to be given circumstantially in the most general form as functions. My only requirement is that they yield the desired solution $\langle\varphi, p\rangle$.[26]

φ is now no longer restricted to sentences of a language. It can be any utterance, a gesture, showing a photograph, a pictorial action, Herod bringing in St. John the Baptist's head on a charger, saying "grrr" and the like. In other words, φ includes "natural" and "nonnatural" actions and both can communicate in the right circumstances. The right circumstances primarily include common knowledge of $m(\varphi)$, the meaning of the utterance. The actions can exploit all sorts of connections with the meaning, whether it is resemblance or something else. All that is required is that there be some connection (i.e. some function m) between the utterance and its meaning. Schiffer (1972), whose definition appears to be closest to mine in terms of the examples it includes, also requires a relation R between φ and p. This is a departure from and generalization of Grice's definitions, which required a recognition of speaker intention. This has now been generalized to recognition of $m(\varphi)$ as the medium through which p is identified. In the cases where there is no ambiguity (see Figure 5), no recognition of intention is involved, just a recognition of m. The intention is available to be recognized, of course, but actually recognizing it involves an avoidable cost. Interestingly, when there is ambiguity, recognition of intention appears to be required in some cases at least, but we have dropped the assumption of cooperation to allow a much wider range of noncooperative interactions.

[26] Though I have not made this explicit in my definition, it is necessary to assume that the probabilities and payoffs, whatever they are, are common knowledge. This is standardly assumed in game theory. This makes $g(\varphi)$ common knowledge and the solution common knowledge. This option will cover some situations but far from all.

If we don't wish to assume common knowledge of the probabilities and payoffs and therefore of the game, then we must assume there are two games, one for the speaker and one for the addressee, such that their common solution is $\langle\varphi, p\rangle$. This is in fact the usual situation, where speaker and addressee do not share much knowledge of their strategic interaction. However, they do have common knowledge of the game tree and the information sets. We would thus generalize the last clause of the definition to read:

9′. $g_{\mathcal{A}}(\varphi)$, $g_{\mathcal{B}}(\varphi)$ have the unique solution $\langle\varphi, p\rangle$ for $\mathcal{A}$ and $\mathcal{B}$.

Note that these two games are defined to have a shared tree structure and information sets that are derived from the first six conditions above. Only the probabilities and payoffs are different. In other words, these structures are not completely general strategic interactions, but something in between a game with full common knowledge and a completely general strategic interaction. This is really the formulation we need, but since tools to cope with strategic interactions are scarce, I have chosen the stronger formulation above. This more general formulation reflects the point made above that speakers cannot directly perceive the addressee's payoff maximization and interpretation. It has to be inferred from subsequent actions and that too, only probabilistically much of the time. And vice-versa.

In such cases, the payoffs are just given as common knowledge, or arise in some way other than through the recognition of intention.

As Strawson (1964) has pointed out, there is a certain symmetry between speaker and addressee, and as should be obvious, this is perfectly captured by the game. That is why communication is a *joint* act, like, for example the joint act of two or more people pushing a cart uphill. This suggests that the concept of a joint act may also be defined in terms of games, with individual acts being present, common knowledge of individual acts, and perhaps an n-person game with an appropriate solution. In other words, such a definition would be a further generalization of the definition of communication that I have presented here.

Computers don't have intentions or beliefs when they convey or interpret information and the definition bars computers from communication on this ground. It may be possible to interpret the condition of an intention to convey or interpret more liberally. While intentions and beliefs are required at least for human communication, some other equivalent architecture may operate in other domains. I will leave the question open.

I mentioned vagueness earlier and we would have to extend the definition to have two meanings, $m_{\mathcal{A}}$ and $m_{\mathcal{B}}$, one for the speaker and an overlapping one for the addressee. There would be no common knowledge of m and we would have different game-theoretic structures $g_{\mathcal{A}}$ and $g_{\mathcal{B}}$ as we did before. This suggests yet another generalization of our model. All that would be required for communication in such a case is that there would be two solutions of the strategic interaction $\langle \varphi, p_{\mathcal{A}} \rangle$ and $\langle \varphi, p_{\mathcal{B}} \rangle$ with the property that $p_{\mathcal{A}}$ and $p_{\mathcal{B}}$ had a sufficient overlap. The criterion for how much overlap is sufficient could be utility-based, as suggested by Rohit Parikh (1994). Vagueness is rife in natural languages and so is vague communication, so this is an important generalization to consider. It means that communication isn't quite a game, but is instead a wider thing called a strategic interaction. We have now two reasons to go beyond games to strategic interactions. Not only that, my solution suggests that all communication involves two propositions, one for the speaker ($p_{\mathcal{A}}$) and one for the addressee ($p_{\mathcal{B}}$). These two propositions may coincide in the case of nonvague communication, but in general will merely overlap for vague communication.

We can now see why "natural" and "nonnatural" are perhaps not the best terms to distinguish between types of information flow. Both natural and nonnatural features are involved in the definition of communication for one thing. And more importantly, communication covers cases of natural utterances as well as nonnatural utterances. We may as well use "communicative” and “noncommunicative”, or “mean$_C$” and “mean$_{NC}$” instead, now that we know what these terms are.

Indeed, we can define “means$_C$” in the following way:[27]

DEFINITION 2. $\mathcal{A}$ means$_C$ p by producing φ iff there is an agent $\mathcal{B}$ such that

1. $\mathcal{A}$ intends to convey p to $\mathcal{B}$.
2. $\mathcal{A}$ utters φ.
3. $\mathcal{A}$ believes that (2) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.
4. $p \in m(\varphi)m$.
5. $\mathcal{A}$ believes that (4) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.
6. $\mathcal{A}$ believes that $g(\varphi)$ has the unique solution $\langle \varphi, p \rangle$ for $\mathcal{A}$ and $\mathcal{B}$.[28]

Note that “convey” merely signifies a transfer of information. We can give a corresponding definition for the concept that we might call “interpretation$_C$” on the side of the addressee.

DEFINITION 3. $\mathcal{B}$ interprets$_C$ φ as conveying p iff there is an agent $\mathcal{A}$ such that

1. $\mathcal{A}$ utters φ.
2. $\mathcal{B}$ believes that (1) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.
3. $\mathcal{B}$ intends to interpret φ.
4. $\mathcal{B}$ interprets φ.
5. $p \in m(\varphi)$.
6. $\mathcal{B}$ believes that (5) is common knowledge between $\mathcal{A}$ and $\mathcal{B}$.

---

[27] I include Schiffer’s (1972) definition for reference:

$\mathcal{A}$ meant that p by uttering x iff $\mathcal{A}$ uttered x intending thereby to realize a certain state of affairs E which is such that E’s obtaining is sufficient for $\mathcal{A}$ and a certain audience $\mathcal{B}$ mutually knowing that E obtains and that E is conclusive evidence that $\mathcal{A}$ uttered x with the primary intention

1. that there be some ρ such that $\mathcal{A}$’s utterance of x causes in $\mathcal{B}$ the activated belief that $p/\rho(t)$;
   and intending
2. satisfaction of (1) to be achieved, at least in part, by virtue of $\mathcal{B}$’s belief that x is related in a certain way R to the belief that p;
3. to realize E.

[28] We would have to replace this with a more general clause if we were considering strategic interactions rather than games. This clause is as follows:

6′. $g_{\mathcal{A}}(\varphi)$ has the solution $\langle \varphi, p \rangle$ and $\mathcal{A}$ believes that $g_{\mathcal{B}}(\varphi)$ also has the solution $\langle \varphi, p \rangle$.

7. $\mathcal{B}$ believes that $g(\varphi)$ has the unique solution $\langle \varphi, p \rangle$ for $\mathcal{A}$ and $\mathcal{B}$.

Note that "interpret" is different from "interpret$_C$." The first word merely signifies $\mathcal{B}$'s attempt to figure out what $\mathcal{A}$'s utterance means without implying the full apparatus required by the technical term "interpret$_C$."

Note that $\mathcal{A}$ (and $\mathcal{B}$) have only the simple intention to convey (or receive) rather than the extremely complex intentions proposed by Grice, Strawson, and Schiffer. If my definitions work, this is obviously a distinct advantage. My point is that there is a kind of division of labor between the speaker and the game and the ambient game does much of the work. It may be possible to argue that the somewhat complex beliefs my definitions require operate like situated beliefs, in that they do not need to be explicit. As Perry (1986) has argued, the circumstantial nature of action does not require an agent to have every relevant belief explicitly present in its mind. We do not need to consider gravity each time we reach for a glass.

It is interesting to examine the connection between these concepts. We can more or less immediately deduce the following facts:

1. Communication implies meaning$_C$.
2. Communication implies interpretation$_C$.
3. Meaning$_C$ and interpretation$_C$ do not imply communication.

To see how we get the first two, note that common knowledge of a fact implies belief that there is common knowledge of the fact. (In fact, it implies knowledge of common knowledge of the fact.) This is an interesting property of common knowledge. A solution of a game is common knowledge and as such implies belief.

For the third, note that belief that there is common knowlege does not imply common knowledge. This is an obvious property of belief.

These results are just as they should be, given our intuitive understanding of them.

I should emphasize that these definitions are intended to be quite general and to capture the intuitive concept of communication we all share. They are not stipulative definitions. The reader is invited to test them against his or her favorite examples.

14. CONCLUSION

In this paper, I have shown how communication can be modelled by games and strategic interactions. Perhaps the most important consequence of this model is that it suggests necessary and sufficient conditions for communication, meaning, and interpretation.

REFERENCES

Aumann, Robert J.: 1985, 'What is Game Theory Trying to Accomplish?', in Kenneth Arrow and Seppo Honkapohja (eds.), *Frontiers of Economics*, Blackwell, Oxford.
Barwise, John: 'Modeling Shared Information', unpublished manuscript.
Barwise, John and John Perry: 1983, *Situations and Attitudes*, MIT Press, Cambridge.
Barwise, John and Jerry Seligman: 1997, *Information Flow*, Oxford University Press, London.
Bernheim, B.D.: 1984, 'Rationalizeable Strategic Behaviour', *Econometrica* **52**, 1007–1028.
Brandenburger, A. and E. Dekel: 1985, 'Rationalizeability and Correlated Equilibria', Unpublished preprint, Stanford University and Harvard University.
Dretske, Fred I.: 1981, *Knowledge and the Flow of Information*, MIT Press, Cambridge.
Grice, H. P.: 1957, 'Meaning', *Philosophical Review* **66**, 377–388.
Grice, H. P.: 1969, 'Utterer's Meaning and Intentions', *Philosophical Rview* **78**, 147–177.
Kreps, David: 1985, *Signalling Games and Stable Equilibria*, Mimeo, Stanford University.
Lewis, David: 1969, *Convention*, Harvard University Press, Cambridge.
Myerson, Roger: 1995, *Game Theory: Analysis of Conflict*, Harvard University Press, Boston.
Nash, J.: 1951, 'Non-cooperative Games', *Annals of Mathematics* **54**, 286–295.
Parikh, Prashant: 1990, 'Situations, Games and Ambiguity', in Robin Cooper, Kuniaki Kukai, and John Perry (eds.), *Situation Theory and Its Applications, I*. CSLI Publications, Stanford.
Parikh, Rohit: 1994, 'Vaguness and Utility: The Semantics of Common Nouns', *Linguistics and Philosophy* **17**, 521–535.
Perry, John: 1986, 'Circumstantial Attitudes and Benevolent Cognition', Technical Report No. CSLI-86-53, Stanford University.
Schelling, T. C.: 1960, *The Strategy of Conflict*, Harvard University Press, Cambridge.
Schiffer, Stephen: 1972, *Meaning*, Oxford University Press, Cambridge.
Strawson, P. F.: 1964, 'Intention and Convention in Speech Acts', *Philosophical Review* **73**, 439–460.

Visiting Scholar
Institute for Research in Cognitive Science
University of Pennsylvania
USA