Game theory and communication

1 Introduction

This paper looks at recent attempts to shed light on communication using game theory. It is divided into three parts. First, motivations for a game-theoretic approach to communication are briefly investigated. In the second part, one of the most fully developed game-theoretic accounts of communication is examined: Prashant Parikh's post-Gricean utterance-by-utterance account (Parikh 1990, 1991, 2001). Doubts are raised about some of the aspects of Parikh's treatment and suggestions are made for refinements of cost factors to improve predictive power. A more fundamental problem is that the model drops a Gricean constraint on inference in communication. I argue that this leaves it without an account of the content of implicatures. Some comparisons are made with relevance theory (Sperber and Wilson, 1986/95), a non-game-theoretic utterance-by-utterance account of communication, which retains a form of the Gricean constraint.
In a final section, I look at the broader prospects for using standard (that is, non-evolutionary) game theory to capture the link between rationality and communication. The idea of using game theory to find constraints on language is briefly sketched, and I mention some general doubts about the formalisation of rationality in standard game theory.

1.1 Why employ game theory in an account of communication?

Game theory is concerned with strategic interactions, that is, situations where two or more players (or agents) have to make decisions and the way things turn out for each player may depend on the other player or players' choices.
Superficially, at least, human communication looks like this sort of situation[1], since a speaker makes an utterance and a hearer tries[2] to interpret it. We say that the hearer *tries* to interpret the utterance because a great deal can be meant by the speaker which is not encoded in the linguistic form (the lexical items and the syntactic structure) of the phrase or sentence uttered. The hearer must, at least, choose a meaning for ambiguous expressions, assign reference to indexical elements such as pronouns, decide on reference class and scope for quantifiers, work out for some lexical items how loosely or narrowly they are being used and recover intended implicit meaning ('implicatures').
How successful the speaker and hearer each are seems to depend on choices made by the other: if the interpretation the speaker intended is (close enough to) the one the hearer works out, then communication is successful, otherwise there is miscommunication. This apparent degree of shared interest has led to the suggestion that communication be modelled as a coordination game, that is, a game where the players' payoffs are aligned. (e.g. Lewis, 1969; Parikh, 1991).

1.2 Game theory, communication and rationality

A central attraction of game-theoretic models of communication is that they might help to elucidate the link between rationality and human communication (if there is one). Grice was convinced that principles governing communication would fall out from general assumptions about rationality and the cooperative nature of communication (perhaps when taken together with other assumptions about human beings or the communicative situation).

> "I am… enough of a rationalist to want to find a basis that underlies these facts [about people behaving as the conversational principle and maxims say]… . I would like to think of the standard type of conversational practice not merely as something that all or most do IN FACT follow but as something that it is REASONABLE for us to follow, that we SHOULD NOT abandon"(Grice, 1975, p. 48, his emphases.)

Standard game theory has certain assumptions about common knowledge and rationality (CKR) built in, and an account of communication in terms of standard game theory would inherit these assumptions. This approach promises, therefore, to make the link between rationality and communication clearer and more precise. On the other hand, a game-theoretic model of communication would also inherit any empirical and theoretical disadvantages of the particular formalisation of rationality adopted by standard game theory.

There are at least two ways of making the link between game theory and communication (van Rooy, 2004, p. 494): either rationality considerations are taken to apply to languages or directly to the communicative situation. Lewis' work and recent work by van Rooy (2003, 2004) is in the former tradition; Parikh's model (Parikh, 1991, 2000, 2001) takes the second approach[3], looking at how communication is achieved between a speaker and a hearer with the language as a given. In section 2, I examine Parikh's model, arguably the most developed attempt at either of the approaches. I will compare this model with the account of communication given by relevance theory, with the aim of assessing how well Parikh's account compares with modern cognitive pragmatics in its explanation of retrieval of explicit and implicit meaning.

1.3 Cooperative games

Situations where the interests of the players are lined up are modelled as cooperative or coordination games.

Consider Figure 1:

		player 2		
		Trafalgar Square	**UCL front quad**	**…**
player 1	**Trafalgar Square**	20, 20	-10, -10	**…**
	UCL front quad	-20, -10	10, 20	**…**
	…	**…**	**…**	**…**

Figure 1 Meeting game

Here both players want to meet, so the outcomes where they make the same choice have high payoffs for both. The outcomes where they do not meet have negative payoffs to reflect the effort involved in making the journey. Also, player 1 has an aversion to UCL, so for him the payoffs at UCL are lower than elsewhere. If this were not the case, this would be a purely cooperative game, or, equivalently, a pure coordination game. As we will see, Parikh models communication as either cooperative or nearly so.

In the simple game where Trafalgar Square (t) and UCL front quad (u) are the only choices, ⟨t, t⟩ and ⟨u, u⟩ are both 'Nash equilibria': situations where neither player will be better off if he changes his choice unilaterally. In coordination games, there are typically multiple Nash equilibria 'along the diagonal' and game-theoretic accounts of communication have had to focus on criteria for a unique solution.

Games can also be played sequentially: first one player takes a turn, then the other, *with the second player knowing what the first player has chosen*. In this case, the games can be represented as trees. Figure 2 is a tree for a meeting game with two players and two choices each.

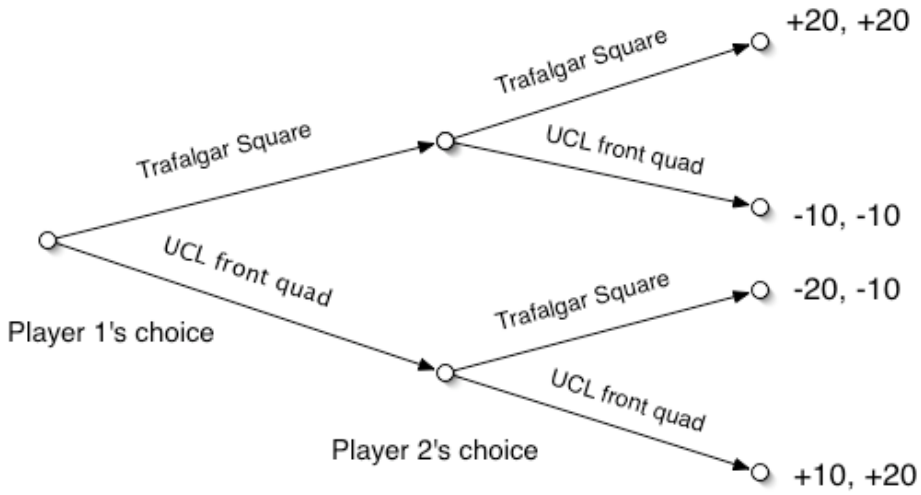

Figure 2 Sequential meeting game with two choices and two players

Here player 1 gets to move first. How should he choose his move? He looks at what player 2 will do in each sub-tree, that is, in each situation that player 2 could be in. Here it is simple: we assume that player 2 chooses so as to maximise his payoff. It follows that if player 1 chooses t then player 2 will choose t; if player 1 chooses u then player 2 will choose u (maximising his pay-off to +20 versus −10 in each case). The other outcomes can be ruled out. So player 1's choice is between ⟨t, t⟩ (the top branch), and ⟨u, u⟩ (the bottom branch). He prefers the pay-off of +20 for ⟨t, t⟩ to the +10 he would get for ⟨u, u⟩ so he will choose t and player 2 will then choose t.
The games Parikh uses are also sequential games: a speaker chooses an utterance, and a hearer, knowing what has been uttered but not the intended interpretation, chooses an interpretation.

2 Parikh's model

In this section of the paper I introduce Parikh's game-theoretic model of communication, giving a critical commentary.
Consider a situation where a speaker makes an utterance and a hearer tries to interpret it. Parikh starts by examining cases where the sentence uttered has two possible meanings. This could be due to lexical or structural ambiguity, to the need to assign reference or to a purely pragmatic availability of two readings for the sentence.[4] Parikh gives an example in which the speaker utters (1):

(1) φ: Every ten minutes a man gets mugged in New York

According to Parikh, this has the two possible meanings:

(2) p: Every ten minutes someone or other gets mugged in New York.[5]

(3) p': Some particular man gets mugged every ten minutes in New York.

As Parikh acknowledges, there are other parts of the meaning of (1) which the hearer must resolve: 'New York' might mean the state or the city, and 'ten minutes' is clearly being used somewhat loosely. The aim is to show first how the model of communication resolves one uncertainty at a time, then allow for its extension to cover realistic cases where several aspects of the meaning must be fixed simultaneously.

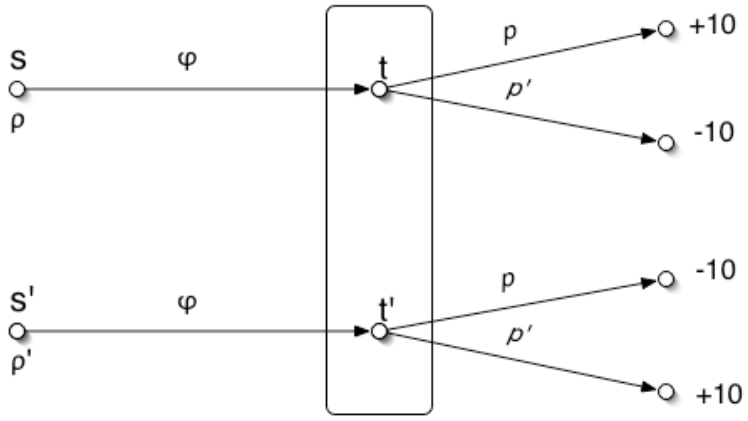

Figure 3 Part of local game for utterance of φ (Parikh, 2001, p. 30)

Consider figure 3. There are two initial situations, s and s'. s is the situation in which the speaker intends to convey p; s' is the situation in which she intends to convey p'. In either case, she may utter φ. After the utterance there are also two situations, t and t', where t is the situation where the speaker means to convey p and has uttered φ and t' is the situation where she means to convey p' and has uttered φ. The speaker, knowing what she wants to convey, knows which situation she is in. The hearer does not. His uncertainty is represented by the box around t and t'. He assesses the probability of s as ρ and the probability of s' as ρ'. He then chooses an interpretation—either p or p'. Note that his preferred choice depends on information he does not have. If he is in t he prefers to play p; if he is in t', he prefers p'. This is reflected in the payoffs. For the moment, Parikh assumes that successful communication is worth +10 to each player and unsuccessful communication is worth −10. This is based on two assumptions, both of which he later relaxes: first, that all information has the same value; secondly that the information is worth the same to both players.

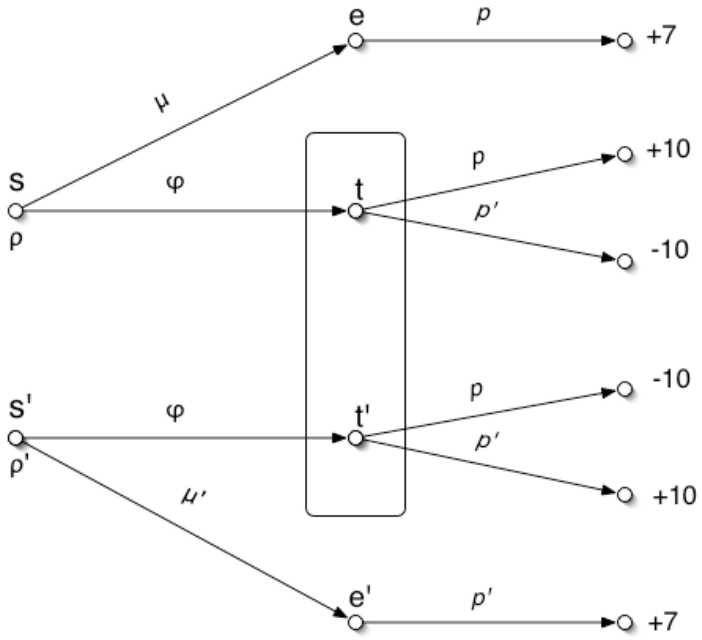

Figure 4 Local game with alternative utterances (Parikh, 2001, p. 31)

According to Parikh, successful communication depends on consideration of alternative utterances. In figure 4 the alternative utterances μ and μ' have been included. μ unambiguously means p and μ' unambiguously means p'. Therefore if the speaker utters μ, the hearer knows he is in situation e and will choose interpretation p. Similarly an utterance of μ' leads the hearer to choose p'. The payoffs are worked out by assigning +10 as before for successful communication, and -3 for the extra effort involved in the production and comprehension of these longer and linguistically more complex utterances. Parikh does not give details of the way the linguistic complexity translates into effort. Later we will see that he allows the cost of constructing or processing a mental representation to come in also as a negative factor in payoffs.
Figure 4, then, is the hearer's model of the interaction, which he is able to construct when φ is uttered. The speaker can also construct this model, since it is based on shared knowledge. If there is a unique solution to the game, therefore, it will be known to both players. Successful communication using φ will be possible. I return to the solution of the local game in section 2.1 below.

If the speaker knows the payoff of the local game she knows the payoff from uttering φ. She compares this with uttering μ and other alternatives as shown in the 'global game', figure 5.

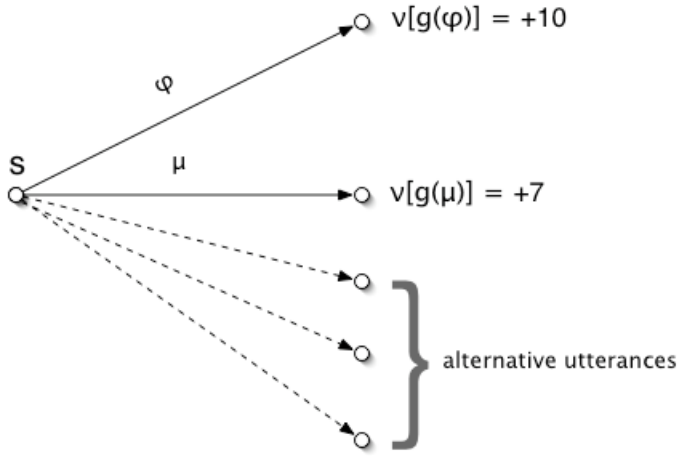

Figure 5 The speaker's choice of utterances (after Parikh, 2001, p. 32)

As shown, φ has a higher payoff than μ, so the speaker should choose it as long as it also has a higher payoff than other alternative utterances.

2.1 Solutions for the game

Returning to the problem of solutions for the local game, consider the different possible strategies for both players. A strategy is a specification of the choices a player will make at all decision nodes. Here, the non-trivial choices are for the speaker at s and at s' and for the hearer at t or t' (the hearer's choice is constrained to be the same at t and t', since he cannot know which of them he is at). This gives eight different strategies, two of which are Nash equilibria, that is, solutions where neither player can do better by changing his or her choice unilaterally.

What A will say in s	What A will say in s'	What B will choose in (t, t')	Nash equilibrium?	If not Nash equilibrium, why not?
φ	φ	p	no	A should defect to μ' in s'
φ	μ'	p	yes	-
φ	φ	p'	no	A should defect to μ' in s'
φ	μ'	p'	no	A should defect to μ in s; B should defect to p
μ	φ	p	no	A should defect to μ' in s'; B should defect to p'
μ	μ'	p	no	A should defect to φ in s
μ	φ	p'	yes	-
μ	μ'	p'	no	A should defect to φ in s'

The first Nash equilibrium (N1) is the intuitively correct solution: if the speaker wants to mean p—the more likely meaning—then she says φ, the shorter but ambiguous utterance; if she wants to mean p'—the less likely meaning—then she says μ', the longer but unambiguous utterance; and the hearer correctly interprets the ambiguous utterance φ as having the more likely meaning, p.
The second Nash equilibrium (N2) is a kind of inverse of the intuitively correct solution: if A wants to mean p—the more likely meaning—then she says μ, the longer but unambiguous utterance; if she wants to mean p'—the less likely meaning—then she says φ, the shorter but ambiguous utterance; and B correctly interprets the ambiguous utterance φ as having the less likely meaning, p.
It is clearly good that the other strategies are ruled out, since they describe even stranger arrangements than the second Nash equilibrium. However it is also necessary for Parikh to rule out the equilibrium N2 (although there is no miscommunication in N2), leaving only the intuitively correct solution, N1.
Note that this problem of multiple Nash equilibria comes with the decision to model communication as a coordination game. As noted above, coordination games generally have multiple Nash equilibria. This is exemplified by figure 3: the cells on the diagonal will be preferable to both players to the off-diagonal outcomes that can be reached by unilateral change of choice. Any game-theoretic approach to communication which models the communicative situation as a coordination game must resolve this problem.
Parikh's proposal here is to bring in another solution concept, Pareto-dominance, which is defined as follows: some solution A Pareto-dominates a solution B iff solution A has a better payoff for at least one player than solution B and the payoff of A is not less than the payoff of B for any player. In general, Pareto-dominance can pick out different solutions from Nash equilibrium[6]. To avoid this, Parikh applies it as a secondary criterion to choose between Nash equilibria.[7] In this case, this works out as choosing the Nash equilibrium with the highest expected payoff. Using expected utility as the measure of the worth of an outcome, as is usual in game theory, the expected payoff for an outcome = payoff x probability of outcome. (Thus, for example, a rational agent should prefer a 50% chance of €100 to €49 for sure, neglecting risk avoidance.)
Thus, setting $\rho = 0.9$ and $\rho' = 0.1$, we have the following payoffs for the Nash equilibria:

payoff of 'correct' solution, N1 = $10 \times \rho + 7 \times \rho' = 9.7$

payoff of 'incorrect' solution, N2 = $7 \times \rho + 10 \times \rho' = 7.3$

The intuitively correct solution Pareto-dominates the other solution, which can therefore be eliminated. Generally, N1 Pareto-dominates N2 in this game, iff

$$\upsilon(\varphi \text{ in } s, p).\rho + \upsilon(\mu' \text{ in } s', p').\rho' > \upsilon(\mu \text{ in } s, p).\rho + \upsilon(\varphi \text{ in } s', p').\rho'$$

where υ(strategy) is the payoff for that strategy.
This is a criterion for successful communication in this game, therefore.

2.2 Where do the probabilities come from?

The subjective probabilities assigned by the hearer to the situations s and s' are crucial in determining whether successful communication occurs in the model. A number of questions might be asked about these probabilities, particularly where they come from, that is, how do the hearer and the speaker arrive at them? (Note that the speaker must assign at least similar probabilities to the ones assessed by the hearer for communication to be successful.)
Parikh's answer *for the example given*[8] is that the probability, ρ, of the speaker's being in situation s, that is, wanting to convey p, is related to the probability that p is true, although with a complication:

> "Since it is common knowledge that p is more likely than p', we can take it as common knowledge that A [the speaker] probably intends to convey p rather than p'... In general, there is a difference between these two probabilistic situations, and it is only the second, involving A's intention, that matters. In the absence of any *special* information, one situation does inform the other..." (Parikh, 2000, p 197, my emphasis.)

Three important questions are raised by this formulation. First, what is 'special' information and how do we know when it comes into play? Secondly, what can be said about the general case, given that this formulation is only intended to apply to the example given? Thirdly, if the relative probabilities of the propositions, p and p' sometimes tell us whether ρ > ρ', as Parikh claims they do in the example, why could Parikh not generally use (some transformation of) the probabilities of p and p' for ρ and ρ'?
The answer to the first question is given in a parallel quotation from an earlier paper: "Note that in general, there is a big difference between the likelihood of a proposition's being true and the likelihood of an agent intending to communicate that proposition. *It is the absence of further relevant information that justifies identifying the two possibilities***.**" (Parikh, 1991, p 482, my emphasis.)
No characterization of relevance is given, so it appears that an intuitive notion of relevance is playing a crucial role here.
Secondly, Parikh says about the general case only that the probabilities will "usually be subjective because objective information will not be available. In general, the initial probabilities are a result of the prior beliefs and goals of speaker and addressee, based either on prior discourse or actions, or just the relevant background beliefs and goals of each agent." (2001, p. 28) This is too vague to be useful in making predictions, which is problematic given that sometimes a small change in the probabilities could lead to a radical change in the predicted interpretation. The problem is not that the numbers involved are not specified by the theory, and have to be put in by hand, rather that the list of factors involved in determining the probabilities is qualitatively vague.
Given these difficulties, one might ask why Parikh needs to go beyond supposing that the correct probabilities come from the probabilities of the propositions. That is, does he really need to suppose that there will be 'special' cases and that the general case is even more complicated? In the next section I draw on work by Wilson and Matsui to show why it would not be generally correct to use the probabilities of the propositions as the probabilities of the situations s and s'. I also propose an alternative, which is not (yet) quantitatively precise in every case but is qualitatively specific about the factors involved.

2.3 Truth, relevance and disambiguation

Wilson and Matsui (1998) examine accounts of pragmatic processes such as disambiguation which use rules such as 'The correct interpretation is the one which is most likely to be true.' Accounts of

this type are called truth-based. These accounts often make incorrect predictions: in general the correct interpretation need not be the one most likely to be true, as examples such as (4) show:

(4) John wrote a letter.

(a) John wrote a letter of the alphabet

(b) John wrote a letter of correspondence

(Wilson & Matsui, 1998, p. 24)

The lexical item 'letter' is ambiguous, so that (4) could mean either 4(a) or 4(b). The disambiguation naturally arrived at (at least in non-biasing contexts) is 4(b), but this cannot be more likely to be true than 4(a) since 4(a) is entailed by 4(b): anyone writing a letter of correspondence must be writing letters of the alphabet[9]

It is examples of this kind which rule out just using the probabilities of p and p' for ρ and ρ' in Parikh's model. Would it be a good move to say that *normally* a truth-based approach is followed but that it needs to be modified in certain cases? Symmetrical examples like (5) and (6) suggest that it is not.

(5) Mary is frightened of *dogs*. (ambiguous between male dogs and dogs in general)

(6) Mary is frightened of *cats*. (ambiguous between cats in general (lions, domestic cats, tigers etc.) and domestic cats)

(Wilson, lectures at UCL. See also Sperber & Wilson, 1986/95, p. 168)

In (5) the intuitively correct reading has the more general sense of the term, *dog in general*. Parikh could deal with examples like this in the same way as example (1). In (6), on the other hand, the intuitively correct reading has the more specific sense of the ambiguous term, *domestic cat*, and this reading is surely less likely to be true, given that big cats are more frightening than tabbies. In this case, therefore, Parikh would presumably say that relevant information somehow takes precedence over what is known about the probabilities of the propositions so that the probability that the speaker wants to convey that Mary is frightened of domestic cats is higher than the probability that she means that Mary is frightened of cats in general.

There is a better solution available, however. In both cases, the intuitively correct meaning is the one with the more accessible of the two readings of the terms. *Accessibility* is a psycholinguistic concept. Empirical work in psycholinguistics aims to determine what makes a lexical item or a sense of a lexical item more accessible on one occasion or another. Some current theories propose that the frequency of use of the sense and the recency of its use are the key factors in accessibility of different senses in disambiguation. In a 'neutral' context there are no recent uses to consider, so accessibility for a reader encountering (5) and (6) as they are presented here would depend only on the frequency of the senses. Relying only on this cue for disambiguation would get both examples right: the more common meanings are *domestic cat* and *dog in general*.[10]

These kinds of examples are not particularly hard to find. Parikh's own examples raise the same questions. Recall that the explanation given of why we arrive at reading (2) for an utterance of (1) is that it is common knowledge that the alternative reading—in which it is a particular man who gets mugged—is less likely to be true.

According to Parikh, the ambiguity in (7) is ineliminable (in some contexts only, I assume—Parikh has "in a general context" (2001, p.41)):

(7) A comet appears every ten years. (Parikh, 2001, p. 41)

The claim is that here $\rho = \rho' = 0.5$, so the expected payoffs for N1 and N2 are equal and the model predicts that the utterance should be unresolvably ambiguous. Note that in this case, Parikh has set

the probabilities ρ = ρ' by hand, even though it is clear that the more general meaning ('Some comet or other appears…) is entailed by and therefore at least as likely to be true as the other ('A particular comet appears…'), that is, the probability of p ≥ the probability of p'.
In example (8), where the more specific reading is the intuitively correct one (in 'neutral' contexts), Parikh would have to appeal to special information to set ρ ≤ ρ'. Here, as previously, the probability of p ≥ the probability of p', that is, the same entailment between the propositions expressed by the general and the specific reading holds and the specific reading cannot be more likely to be true than the general one.

(8) A comet appears every 76 years.

Perhaps a better move for a supporter of Parikh's model would be to say that the probabilities ρ and ρ' generally reflect the psycholinguistic activation and accessibility data (rather than sometimes relating them to the probabilities of the propositions the speaker wants to convey and sometimes to "the prior beliefs of speaker and addressee" (op. cit.) in general). This would allow the model to make correct predictions at least in cases like (4),(5) and (6) without appeal to special information. Another way of arriving at the same result would be to do away with the prior probabilities and allow activation to be reflected in the payoffs as cost or effort, where lower activation of a sense would translate into greater cost. This would amount to adopting part (a) of the relevance theory definition of effort (see section 2.6 below.)

2.4 Implicatures

Parikh (2001, pp 80f.) considers a case where there is an utterance φ, which could have an implicature but need not. The meaning conveyed without the implicature is l; the meaning conveyed with the implicature is p. For example:

(9) φ: "It's 4pm."

l: The time now is 4pm.
p: The time now is 4pm. + It's time to go to the talk.

As before, it is assumed that information is valuable and misinformation has a negative value (of −2 in this case).
It is also assumed that φ is uttered in a context where the speaker knows that the hearer wants to go to a talk at 4pm. In this case p is worth more to the hearer than l, because the extra information would help him in a decision he has to make. In other cases of communication involving implicatures, p may be worth more than l for other reasons (I comment on this in section 2.6 below). Here the values of p and l are set to 9 and 4 respectively for both speaker and hearer.
On the effort side, it is assumed that there are processing costs for the hearer and for the speaker: they are greater for the speaker because she has to produce the utterance. Parikh also assumes that arriving at p is more costly than arriving at l only, since "p has to be contextually inferred from l… [with] additional processing involved." (Parikh, 2001, p. 81)
Then we have the game in figure 5. As in the game for explicit meaning, alternative unambiguous utterances are considered, here μ, which unambiguously means p, and silence, which is taken to convey no information, perhaps problematically, since silence is often communicative, in fact, and might have been so here.
Provisionally accepting all of these assumptions and taking the probabilities of s and s' to be negligibly different, we have as a unique Pareto-Nash equilibrium the strategy in which the speaker utters φ if she wants to convey p and silence if she wants to convey an 'empty interpretation' and in either case is correctly understood by the hearer. How adequate is this kind of account of

implicatures? In the next section, I consider some questions which any account of implicatures should answer.

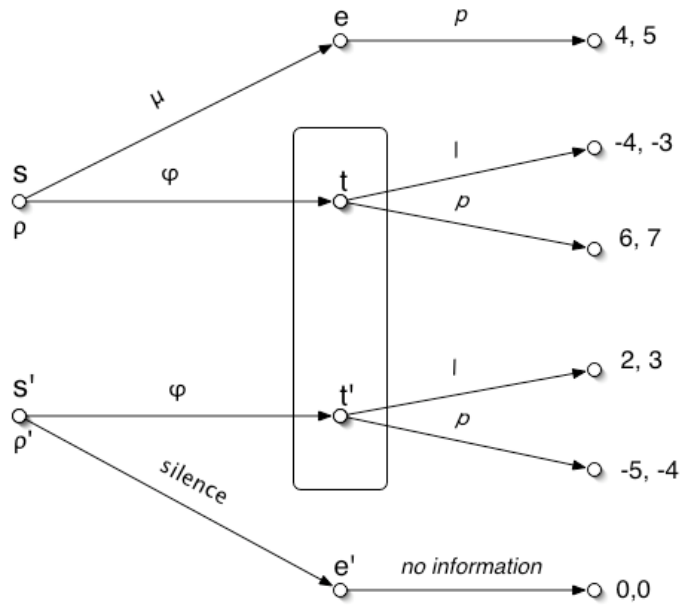

Figure 6 Local game for implicatures (Parikh, 2001, p. 82)

2.5 Questions about implicatures

A crucial question for an account of implicatures is what the search space for implicatures is. In Parikh's model no constraints are placed on p except that it and l are both meanings of φ within the language shared by the speaker and the hearer. This applies to implicit meaning as well as explicit meaning. Formally, this is expressed as a meaning function that maps from utterances onto (multiple) meanings, both explicit and implicit.

If current arguments are correct that free contextual enrichment can affect the proposition expressed (Carston, 2002, p. 40), then it is a mistake to say that all the possible explicit meanings, p, p', p"… are given by the language. Leaving this debate aside, though, for implicatures it seems much less plausible that a language would specify all possible meanings in all possible contexts. One reason that this approach might seem right for disambiguation and some other aspects of explicit meaning is that there really is a small set of possibilities for the disambiguation that are context-independent, so that 'bank' for example, either means a financial institution or the side of a river. Thus it is not unreasonable to try to model disambiguation as a choice between options specified by the language, where context affects the payoff and probabilities of the options, but does not affect which options exist. But this reasoning does not carry over to implicit meaning. For implicatures, the set of possible meanings would itself depend heavily on context, as we see in Parikh's example, where we are told that the speaker and hearer both know that there is a talk at 4 o'clock and that both are interested in attending. This information should constrain the possible implicatures. Parikh's model does not explain how.

In adopting this formulation, Parikh abandons the Gricean idea that implicatures should follow logically from the combination of explicit content and some assumptions. That proposal is desirable because it vastly reduces the search space for implicatures, given that there are many propositions which do not follow from the proposition expressed plus a limited set of assumptions.

The way that Grice's insight is developed in relevance theory makes this particularly clear: implicated conclusions are logically warranted by explicit meaning together with implicated contextual assumptions. A relevance theory treatment of example (9) claims that the implicature 'It's time to go to the talk' is only available as part of an interpretation if it follows logically from the explicit meaning plus contextual assumptions such as (10) which must be manifest in the context.

(10) Getting to the talk will take at least five minutes;
It is better to arrive no more than five minutes past the hour for a talk; etc.

In contrast, in Parikh's model, context places no restriction on the meanings that are possible for an utterance, rather, it is only used to guide the selection between possible meanings. Perhaps the reason for this is that in Parikh's model only possible meanings of the utterance are represented and not other propositions which are involved in comprehension. In particular, he does not mention contextual assumptions, so he cannot show how explicit meaning logically constrains implicatures. To restate the point: getting disambiguation right can depend on getting the right context. With implicit meaning this is even more clearly important, because there is an open-ended number of candidates to consider for implicatures. That makes it an even more attractive possibility to have a way of getting to the right interpretation that starts from the context in finding candidates to evaluate.

These considerations are linked to the next question I want to consider: How is the context for interpretation selected? I agree with Sperber and Wilson (1986/95, pp. 137–42) that context selection is a very serious problem for pragmatics, if not *the* problem. Along with many other models of communication, Parikh's seems to me to be making use of intuitions of relevance at this point, at least in the examples given. In the example Parikh chooses p (minus l) as a candidate for an implicature. Intuitively this is a good candidate in this example: it would make sense in the context which is assumed. These intuitive assumptions seem to be taking the place of an account of the way that the hearer decides in what context he should process the utterance to make the intended sense of it. Without wishing to go too deeply into this question here, I want to stress that this is a non-trivial problem. For example, a speaker, Sherlock Holmes, may utter a sentence where reference assignment or disambiguation is needed, such as "He went to the bank," and his hearer, Dr. Watson, may be simply unable to work out which person and what kind of bank are involved because although he and Holmes are in the same physical environment, perhaps even attending to the same aspects of it, and have heard the same prior discourse, still he and Holmes are in different *contexts*. This is because Watson has not made the same inferences as Holmes and thus starts with different information and—importantly—because he is simply unable to make those inferences and therefore cannot get himself into the same context as Holmes.[11] (In relevance theoretic terminology, different propositions are manifest to him.)

A good account of communication need not be a general theory of contexts (whatever that might be), but it must have something to say about the way in which a hearer, presented with an utterance, searches for a context which the speaker took to be accessible to him and in which the utterance can be interpreted so as to make the kind of sense the speaker intended. A Parikhian account of communication might try to answer this need by drawing on work on presupposition and accommodation (e.g. Stalnaker 1973, Lewis 1979)[12].

Note that the contextual assumptions involved do not need to be things that the hearer believed or had even thought about before the utterance, as example (11) shows:

(11) Mary and Peter are looking at a scene which has many features, including what she knows to be a distant church.

Mary: I've been inside that church. (Sperber & Wilson, 1986/95, p. 43)

Mary "does not stop to ask herself whether [Peter] has noticed the building… All she needs is reasonable confidence that he will be able to identify the building as a church when required to… it might only be on the strength of her utterance that it becomes manifest to him that the building is a church." (ibid., pp. 43 f.)
Sperber and Wilson develop an account, including the notion of mutual manifestness, which explains examples like these. Accounts of communication within game theory also need to deal with context selection.

A further problem for modelling any open-ended inference process is proposing a stopping rule that works. My final question in this section, then, is how is the search for implicatures stopped?
Parikh gives an example of the way his model stops further implicatures from being generated. I would like to suggest that to adopt his answer is effectively to adopt a principle of maximal relevance.
The example considered (Parikh, 2001, pp. 84f) models a situation where there are three possible interpretations to be considered if the speaker utters φ: l, p and q. l and p are as before; q includes p plus some extra information. The result is a more complicated game with a solution consistent with the previous one—the speaker, wishing to convey p, utters φ and the hearer interprets it as p—as long as the value of the extra information, l, is less than the effort that would have to be expended to process it. This follows because choosing p will be the equilibrium strategy for the hearer (after φ is uttered) just when the payoff for choosing q rather than p is lower because the cost of the new information outweighs its value. If q is p plus something informative but irrelevant, as in Parikh's example, "let's take the route by the lake", then as Parikh puts it: "Suppose there is some other proposition, q, ... that is more informative than p. It will certainly have positive value but it is easy to see that it cannot have more value than p [in the context discussed]. ... Moreover it is reasonable to assume that the greater the information, the more costly it is to process." (2001, p. 84).
Effectively, then, the stopping criterion that emerges from this model is: generate implicatures until the effort involved in doing so is less than the value of the information they contain. This is superficially close to the communicative principle of relevance, discussed below. If Parikh's model needs this principle and relevance theory can provide an account of communication with a similar (but different) principle and little other machinery then Parikh's account seems to suffer from relative lack of economy. Note that Parikh's stopping rule differs significantly, however, from the relevance-theoretic communication procedure in that it seems to look for maximal rather than optimal relevance. If so, the two different principles would make some different predictions. I discuss this in the next section, which gives details of the relevance theory comprehension procedure and compares relevance theory and Parikh's model at several points.

2.6 Further comparisons with relevance theory

Relevance theory is a theory of cognition. (Sperber & Wilson, 1986/95; Wilson & Sperber, 2002) It claims that human cognition tends to be geared to the maximization of relevance. (This is the cognitive principle of relevance (1986/95, pp. 260f).) Relevance is defined in terms of cognitive effects and processing effort (2002, p. 252):

(12) *Relevance* (1986/95, p. 125; Wilson & Matsui, 1998, p.16)

a. The greater the cognitive effects, the greater the relevance;
b. The smaller the effort needed to achieve those effects, the greater the relevance.

Cognitive effects occur when new information interacts with existing contextual assumptions in one of three ways:

(13) *Cognitive effects* (Wilson & Matsui, 1998, p.16)

a. Strengthening an existing assumption;
b. Contradicting and eliminating an existing assumption; or
c. Combining with an existing assumption to yield contextual implications.

(14) *Processing effort* is affected by: (Wilson & Matsui, 1998, p.16)

a. the form in which the information is presented
b. the accessibility of the context.

In the special case of ostensive inferential communication, the speaker, by making an utterance, is making an offer of information. This raises the expectation that the utterance will be optimally relevant:

(15) *Optimal relevance* (2002, p. 256)

An utterance is optimally relevant to an addressee iff:
a. it is relevant enough to be worth the addressee's processing effort;
b. it is the most relevant one compatible with the speaker's abilities and preferences.

This expectation is spelled out in the Communicative Principle of Relevance (2002, p. 256):

(16) Every utterance communicates a presumption of its own optimal relevance.

This implies the relevance-theoretic comprehension procedure (2002, p. 259):

(17) a. consider interpretations in their order of accessibility (i.e. follow a path of least effort);
b. stop when the expected level of relevance is achieved.

Compare this with Parikh's account of communication. Parikh also factors in effects and effort, but he is less specific about what may contribute to them.

Effects

Effects enter the calculation as the 'value of information' which contributes positively to payoffs. As mentioned, an initial, provisional assumption is that all information is equally valuable; in more complex examples Parikh shows how information can be assigned a value by considering its worth in a game modelling a decision that (the speaker knows) the hearer may make. As, I would claim, with Lewis' work, (e.g. 1969, Ch 4)[13] this risks blurring the distinction between the illocutionary and perlocutionary aspects of meaning. Parikh, perhaps anticipating that there is an issue here, separates implicatures into two types: type I, where an utterance has a direct effect on the hearer's behaviour, and type II, where an utterance affects only the hearer's thoughts, initially at least. In the second case Parikh says, "this type of implicatures can be modeled in more or less the same way except that we need to consider preferences for information directly, rather than via direct action." (2001, p. 86) Parikh has not specified how the value of the information in type II cases is arrived at. In general, I claim, Parikh's model does not specify in the kind of detail that relevance theory does—in (13)—the ways in which information can be valuable to the hearer. Someone wanting to make more explicit predictions with the model would have to be more explicit about the ways that information is valuable.

Before moving on to Parikh's proposals about effort, I want to dwell on his division of implicatures into two types, which I think is undesirable for two reasons. First, it seems that explicit meaning as well as implicatures sometimes leads more directly to action than at other times. Should there also be two categories of explicit meaning? Secondly, the first type of implicature seems redundant. Presumably all utterances, including ones that lead fairly immediately to decisions and actions, have their effects on the hearer by affecting his thoughts. (Or at least, in cases where this is not true we

would not want to call what has happened communication.) So all implicatures will belong to Parikh's second type.

Effort

I have already commented that Parikh allows as effort factors only linguistic complexity (with the metric unspecified) and, later, the cost of representing or processing an implicature mentally. I have argued that in order to account for examples such as (4) to (9), the model needs to incorporate at least (14)a) from relevance theory, so that effort reflects accessibility factors connected with linguistic items in the utterance[14].

There is another issue connected with effort which presents central problems for a game-theoretic account of communication. In game theory the payoffs do not include the cost of constructing the representation of the game, understanding it and finding a solution or solutions. Parikh's model sets up all of the possible meanings of the utterance in parallel and therefore shares a problem with truth-based approaches to pragmatics. As Wilson and Matsui (1998) point out, in order to find which interpretation is most likely to be true, all possible interpretations must be considered. This makes these approaches psychologically implausible. Parikh's model has this problem doubled or quadrupled. First, both the speaker and the hearer must consider all possible interpretations; secondly, for every interpretation an unambiguous alternative utterance must be found.[15]

Considering all possible interpretations is not a trivial matter even for explicit meaning. Parikh's example (1) has at least eight different possible readings since there are at least three degrees of freedom, given that the meaning of 'New York' and the precision or otherwise of 'ten minutes' as well as the scope of the quantifiers are underdetermined by the linguistic form. For implicatures there does not seem to be any principled reason why there should be a determinate number at all. At the least there must be a huge finite number of interpretations that any given utterance could have. In fact this kind of indeterminacy arguably gets into the explicit meaning as well, as for example in (1) where resolution of the meaning of 'ten minutes' is more a matter of finding a degree of precision on a continuum than of choosing among a limited number of possibilities. Contrast Parikh's model with the relevance theoretic comprehension procedure which:

> "integrates effort and effect in the following way. It claims that the hearer is entitled to expect at least enough cognitive effects to make the utterance worth his attention, that the processing effort is the effort needed to achieve these effects, and that the hearer is entitled to accept the first interpretation that satisfies his expectation of relevance." (Wilson and Matsui, 1998, p. 18)

So the hearer works through interpretations in order of (decreasing) accessibility until one of them has cognitive effects worth the processing effort—in which case this will be the interpretation understood (or until the pragmatic faculty exceeds the amount of effort it can spend on this occasion and gives up—in which case no interpretation will be arrived at). In practice this means that generally not very many interpretations will need to be considered; often, as in (5) and (6), the most accessible interpretation will be the right one. Thus while any interpretation might be considered, combinatorial explosion is avoided. The relevance theoretic comprehension procedure is simple and seems computationally tractable: it is a fast and frugal heuristic (Wilson & Sperber, 2002/2004, p. 276, citing Gigerenzer et al. 1999).

Parikh can claim that his model does not need to answer the charge of psychological implausibility since it may be that the model does not describe cognitive structures or processes. He says, "It seems better to view the game as a model of a class of constraints that capture the underlying logic of communication… . The game …describes a valid inference without saying anything about how agents arrive at the correct interpretation of an utterance." (Parikh, 2001, p. 25) I think this misses something important. Certainly, comprehension might be carried out by a heuristic which arrives at

the same interpretations as the model (at least often enough). But if the model correctly describes the logic of the situation then it implies that the mental processes involved in communication must be sufficiently sophisticated to grasp the situation correctly in some way. So the more complicated the model of the situation, the more mysterious the success of the heuristic and the more difficult it would seem to give an account of the workings of that heuristic.[16] In other words, the model does have implications for 'how agents arrive at the correct interpretation' in that it specifies at least the nature and complexity of the problem that they have to solve. Arguably, an account of communication which shows that communication is complicated without making suggestions about how people manage to understand each other is lacking in a crucial respect.[17]

A stopping rule for implicatures

Recall that in section 2.5 above I argued that Parikh's method of stopping in implicature derivations effectively amounted to something very like a principle of relevance, that is, something like: keep generating implicatures as long as the value of the information in an implicature is greater than the effort costs associated with it. This seems to be a principle of maximal relevance, since it makes search continue while there is any more value to be obtained that is worth the effort used in obtaining it. In contrast, the principle of communicative relevance entails that search will only continue until an optimally relevant interpretation is reached. The difference is that relevance theory claims that hearers look for an interpretation such that the utterance is "the most relevant one *compatible with the speaker's abilities and preferences*." (my emphasis) A theory which claims that hearers look for maximal relevance, ignoring these provisos, predicts certain implicatures that apparently do not arise. Carston discusses a number of cases like this in a paper on so-called 'scalar implicatures' (Carston, 1998) including example (18) (her example (71), taken from Green (1995, 96–97):

(18) B: Are some of your friends Buddhist?
A: Yes, some of them are.

Theories which claim hearers look for maximal relevance, including Parikh's model apparently, predict here that A will be taken to implicate that not all of her friends are Buddhist, since

> "in the context that Green sketches, it is evident that there is a more relevant response that A could have given, concerning whether all or most of her friends are Buddhist; this would have more contextual effects for the hearer (B) and would cost him negligible further processing effort. Since A has chosen not to utter this, doesn't it follow that she must be communicating that *only some* (that is, not all or most) of her friends are Buddhist?" (Carston, 1998, p. 33).

On the other hand, relevance theory correctly predicts that this implicature will not arise if it is manifest that the speaker was not willing to make a stronger statement.

> "Green's context makes it plain that while the speaker has the *ability* to make the stronger statement, she *prefers* not to (she is afraid of being considered a Buddhist-groupie) and the hearer is aware of this. Hence the relevance principle correctly predicts that the speaker is not implicating that not all of her friends are Buddhist and that the hearer recovers no such assumption as part of what is communicated."(ibid.)[18]

However, Parikh's model may not be in as much trouble with this kind of example as other frameworks. Many neo-Gricean and post-Gricean accounts of communication assume as a foundational principle that communication is cooperative. As a consequence they are simply unable to give an account of utterances where the speaker will not cooperate. Parikh's model does not assume a Cooperative Principle, so, at least in principle, it could make correct predictions in these cases, given assumptions about the way the speaker's preferences affect payoffs. As far as I can see,

a defender of Parikh's model would have to write in a proviso like the second half of the second clause of optimal relevance, so that the value of extra implicatures would be zero, no matter how useful the information might be to the speaker, if the speaker manifestly was not willing or able to communicate them.

The asymmetry between speaker and hearer

According to relevance theory, when hearers try to find the interpretation of an utterance intended by the speaker there does not have to be any cooperation between the speaker and hearer, except that the speaker wants to be understood and the hearer to understand (Sperber & Wilson, 1986/95, p. 268). The speaker knows that the hearer is built so as to take the first interpretation that is optimally relevant as the correct one, so she has to make sure that her utterance will lead the hearer to entertain this interpretation before any other which would be relevant enough to stop the search. Parikh's model has a similar asymmetry between the speaker and the hearer: both must consider the local game which determines the interpretation of an utterance but only the speaker needs to consider the global game, choosing the utterance which has the highest payoff given an intended interpretation.

There is a difference, however. According to relevance theory the interpretation must be optimally relevant for the hearer but not, in general, for the speaker. The constraint from the speaker's point of view is to produce an utterance which is optimally relevant to the hearer, compatible with the speaker's abilities and preferences. In contrast, in Parikh's model the solution must be optimal for both speaker and hearer for successful communication. There seems to be something problematic about this, since the reasons why the solution will be optimal will generally be different for the speaker and the hearer. Which interpretation will be optimal for the hearer depends on the worth to him of the information he can derive from it. For the speaker the optimal solution is simply the one in which the hearer arrives at the interpretation the speaker intended (or something close enough to it). So the payoffs for the speaker and the hearer will not generally be the same. Parikh's model allows for this in principle, but there still seems to be a worry here, since the model predicts that miscommunication will occur if the payoffs come apart too far. A defender of the model would need to show that this does not generally happen in ordinary cases where, as we have seen, the interests of the speaker and hearer are different.

This is one aspect of a more general worry about the model, since Parikh allows that a number of factors—the effort and effect factors in the payoffs, the probabilities and even the set of meanings for an utterance—can be different for the speaker and hearer. Any of these might come apart, perhaps leading to miscommunication. To consider just the effort factor, this will generally be very different for the speaker and the hearer even though the narrowly psycholinguistic costs of an utterance are often assumed to be the same. Other costs may well be different since the tasks involved are different. The speaker knows (roughly) what she wants to mean, and has to work out what utterance will direct the hearer to this interpretation; the hearer knows what has been uttered and has to work out what was meant, including implicatures—which may require considerable inference.

3 Rationality and game theory

3.1 The other type of model

Do the problems with Parikh's model carry over to other game-theoretic approaches to communication? In section 1.2 I noted that van Rooy (2004) applies game theory to communication in a different way from Parikh, taking rationality and economy considerations to apply to language rather than utterances in context. This approach faces the familiar problem of multiple equilibria.

Van Rooy (2004) looks at a situation where there are two different possible meanings, one more salient than the other, or 'unmarked' in Horn's terminology (Horn, 1984) and two utterance-types, one less linguistically complex than the other (also 'unmarked'). Van Rooy, like Parikh, finds two Nash equilibria. In his model these represent possible linguistic conventions. N1, the intuitively correct solution, is the convention corresponding to Horn's 'division of pragmatic labour': unmarked utterances carry unmarked meaning; marked utterances carry marked meaning. The other Nash equilibrium, N2, is the 'anti-Horn' case: unmarked utterances carry marked meaning; marked utterances carry unmarked meaning. Van Rooy rejects Pareto-dominance as a secondary criterion for equilibria, so both N1 and N2 are solutions. He suggests that it may be possible to eliminate N2 by showing that only N1 is evolutionarily stable(p. 515).[19] Thus the existence of the two equilibria could be reconciled with the non-existence of anti-Horn communities. Perhaps this approach will work. It is worth noting, though, that the Horn generalisation, to the degree that it is true, falls out naturally from the communicative principle of relevance: if a speaker puts a hearer to more trouble than she might have (for example by using a 'marked' utterance), then the hearer is entitled to more cognitive effects (in other words, a 'marked' interpretation).

I think that there are also more general problems with the approach which applies game theory to language rather than to particular utterances. It seems to conflate language with communication, as for example when van Rooy writes that: "Speakers obey Horn's rule because they use a conventional language that, perhaps due to evolutionary forces, is designed to minimize the average effort of speakers and hearers." This is in direct opposition to the Chomskyan view that languages are not tools for communication. Famous evidence is the existence of grammatical sentences that are unusable for communication. Further, a great deal of work in psychology suggests that pragmatic abilities are associated with theory of mind, a separate ability from language competence. (See Happé, 1993 and other references in Wilson and Sperber, 2000, fn. 40, p. 275)

3.2 Doubts about standard game theory's formalisation of rationality

In section 1.2, I mentioned that empirical and theoretical questions have been raised about the way that *standard* (as opposed to *evolutionary)* game theory formalizes rationality. Colman (2003) provides a summary. Briefly, in some games[20], actual players consistently do better than theoretical agents who maximise utility as game theory suggests they rationally should. (p. 149) Examination of the strategy required of a player in the same game "raises a suspicion that the CKR [Common Knowledge and Rationality] assumptions may be incoherent" (ibid.). A (standard) game-theoretic account of human communication would inherit these problems. Until there is greater understanding of these issues, perhaps within psychological game theory (ibid.) it will be unclear whether a standard game-theoretic account of communication successfully links communication with reasonable assumptions about rationality, which, after all, was the main aim of applying standard game theory to the communicative situation.

One assumption made by standard game theory, that the players have common knowledge of the game and of the rationality of the other player, may prove particularly problematic for game-theoretic accounts of communication. Sperber and Wilson discuss reasons why common knowledge of *context* is not a reasonable assumption for a pragmatic theory (1986/95, pp. 15–21). There need not be a direct clash with the position in game theory (pace Sally, 2003, p. 1234), but it may be that game-theoretic accounts of communication also need to build in common knowledge of (some) features of the context. Further, some of Sperber and Wilson's arguments may be effective against common knowledge of the game and of rationality assumptions, in which case any treatment of communication employing standard game theory would be rendered doubtful. Unlike Sally, I see no reason for a presumption in favour of CKR and game theory in this area, certainly not before there

is a fairly successful game-theoretic account of communication—whose success might then be taken as corroboration of its assumptions.

4 Conclusions

I have tried to explain and explore Parikh's interesting game-theoretic model of communication in order to illuminate the possibilities that game theory offers for understanding communication. I have raised empirical and theoretical doubts about this model and, briefly, about van Rooy's alternative account, and mentioned some questions about the foundations of standard game theory. The criticisms of Parikh's model I have presented fall into six categories. I have argued that in some cases Parikh's game-theoretic approach seems to get the data wrong. Secondly, it relies heavily on prior probabilities, which are numbers that the theory does not generate or constrain. Thirdly, I have tried to show that it lacks predictive power, and that some of the predictions it does generate rely on notions of relevance or appropriateness that have not been fully spelled out, thus assuming part of what a theory of communication should predict. A fourth line of criticism is that Parikh's account leaves it unclear how communication is achieved, given the huge complexity of the communicative situation in this theory.

Most importantly perhaps, I have argued that Parikh's model abandons an essential Gricean insight, that explicit meaning plus assumptions logically warrant implicatures, leaving implicature-content unconstrained.

---

[1] See for example Parikh (1991, p. 473) and Sally (2003, p. 1223) for statements of this intuition.

[2] Use of the word 'tries' is not meant to imply that the processes involved need be conscious or available to introspection, either during processing or subsequently. The same caveat, standard in cognitive science, applies to other verbs I have used in describing the hearer's task, including 'choose', 'decide', 'work out' and others.

[3] Although Parikh thinks that both approaches are applicable. In his talk at GDP 2003 he gave an example of a game-tree for the Horn case considered as a rule at the level of language.

[4] Parikh regards all of these cases as cases of ambiguity, not restricting the term, as is more usual in linguistics, to cases where two or more sets of linguistic items or linguistic structures correspond to the same phonetic form. I will use the term 'ambiguous' only in its narrow sense, since *disambiguation* marks a theoretically important category which is in contrast with reference assignment and pragmatic enrichment at least. Note that it makes no difference to Parikh's account whether or not (1) is structurally ambiguous, corresponding to two representations which differ in quantifier scope, for example, given that it has the two readings given in (2) and (3).

[5] Although Parikh presents this sentence as unambiguous, it has the same two readings as (1). It might seem problematic for Parikh that it is often difficult to find unambiguous alternatives, since in his examples successful communication depends on finding unambiguous utterances, but he has shown (at GDP 2003) that ambiguous alternatives will also do.

[6] Although not in *pure* coordination games (thanks to a reviewer for pointing this out).

[7] Van Rooy has made some game-theory-internal criticisms of this solution concept (van Rooy, 2004, p. 506), saying that use of Pareto dominance is normally motivated by pre-talk communication. 'cheap talk' (communication with no costs) before the game. Naturally, pre-talk communication is not an option for Parikh on pain of an infinite regress, since communication is what is to be explained. (See Parikh, 2001, p. 36 fn. 2.) It is true that Pareto-dominant Nash equilibria are not necessarily the ones that are chosen: in a coordination game with many Nash equilibria, all of which have equal payoffs except for one which is Pareto-dominated by all of the others, that one is a focal point and is likely to be chosen, despite its lower payoff. An example is meeting at UCL in the game in figure 3.

Another problem for Pareto-Nash equilibrium as a solution concept might arise from recent research (mentioned by Sally, 2003, pp. 1229 f.) showing that players often prefer 'risk-dominant' (actually risk-avoiding) strategies to Pareto-dominant strategies.

[8] My thanks to Prashant Parikh (p.c.) for pointing out to me that the formulation is specific to the example. Of the general case he writes, "My view is that all the features of the situation have to be taken into account in determining the initial probabilities - these may have very little or nothing at all to do with the likelihood of the proposition itself." (p.c.)

[9] Abstracting away from non-alphabetic writing systems, of course.

[10] As Wilson (p.c.) points out, "Sperber & Wilson's notion of relevance sheds some light on why those particular senses of 'letter', 'dog' and 'cat' are the most accessible ones. The fact that someone wrote a letter in the 'letter of the alphabet'

sense would very rarely be relevant enough (achieve enough effects, at low enough effort) to be worth mentioning, so 'wrote a letter' will rarely be used in that sense, and this explains its infrequency of use. Similarly, narrowing 'dog' to 'male dog' would rarely make enough difference to relevance to be worthwhile. By contrast, 'cat' in the general sense covers such very different sub-cases that it may make a huge difference to relevance to know which sub-case is involved. So 'accessibility' isn't a magic potion but something that is (a) empirically testable and (b) often theoretically predictable."

[11] See *The Blue Carbuncle*, for example:

"I can see nothing," said I, handing [a hat] back to my friend.

"On the contrary, Watson, you can see everything. You fail, however, to reason from what you see. You are too timid in drawing your inferences."(Conan Doyle, 1897/1957, p. 125)

[12] Thanks to a reviewer for pointing this out.

[13] Although the distinction is arguably clearer in Ch 5. (Thanks to a reviewer for pointing this out to me.)

[14] I do not have enough space here to show that (14)b is also necessary.

[15] I do not want to deny that consideration of alternative utterances often plays a role in hearers' recovery of meaning (and speakers' choices of wording). It seems that it would be good for a theory only to take alternative utterances into account where they are easily accessible and to use at most the fact that they are ***not*** easily accessible in other cases. This is how relevance theory deals with these cases (eg Sperber & Wilson, 1986/95, pp. 200-1.)

[16] This argument is a distant relative of an argument used recently in minimalist syntax. (For example by Tanya Reinhart at ESPP 2003, commenting on Chomsky, 2001.) We cannot ignore effort considerations in our account of syntactic competence, according to the argument, because the parser's performance must match the competence and, on the face of it, the more complicated are the representations the competence generates, the more difficult is the job of the parser.

[17] One response to this criticism might be that only pragmatics needs to say anything about *how* communication is achieved. Parikh explicitly denies that his account is a pragmatic theory: "I see this whole book as a part of semantics, not pragmatics, because rational agency is part of a properly situated semantics." (2001, p. 141) But both pragmatic theories and Parikh's model are in the business of explaining communication. If a particular pragmatic theory can say *what* is communicated and *how* the hearer recovers the intended meaning and Parikh's model only has an answer to the first question, then *ceteris paribus* the pragmatic theory is to be preferred.

[18] Van Rooy (pc) has suggested that this example should be accounted for in terms of the information requirement imposed by the word 'some' in the question, following work by Groenendijk & Stokhof (1984), van Rooy & Schultz (forthcoming) and others. One problem for this type of solution is that a small change in context could lead to the speaker meaning "Some but not all of my friends…" using the same form of words in response to the same question, so it seems that the wording of the question is not the crucial factor here.

[19] Similarly, Asher et al (1999) attempt to derive a Gricean maxim of truthfulness using evolutionary game theory. Of course, evolutionary approaches also have the advantage that they do not assume CKR.

[20] See Colman's (2003) discussion of 'Centipede'

References

Asher, N., Sher, I., & Williams, M. (1999) Game-theoretical foundations for Gricean constraints. In *Proceedings of the Thirteenth Amsterdam Colloquium, ILLC, Amsterdam.*

Carston, R. (1998) Information, Relevance and Scalar Implicature. In Carston and Uchida, eds. Relevance Theory: Applications and Implications. Amsterdam: John Benjamins: 179–236.

Carston, R. (2002) Thoughts and Utterances: The Pragmatics of Explicit Communication. London: Blackwell.

Chomsky, N. (2001) Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics 20*. Cambridge, Mass.: MITWPL.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *The Behavioral and Brain Sciences,* 26: 139–15

Conan Doyle, A. (1892/1957) The Blue Carbuncle. In *The Adventures of Sherlock Holmes*. References are to the 1957 edition. London: John Murray.

Gigerenzer, G., Todd, P. & the ABC Research Group. (1999) Simple Heuristics that Make us Smart. Oxford: Oxford University Press.

Grice, H. P. (1975). Logic and Conversation. In Cole, P. & Morgan, J. (1975) *Syntax and Semantics of Speech Acts*. New York: Academic Press: 41–58. Reprinted in Grice, H. P. (1989) *Studies in the Way of Words.* Cambridge, MA: Harvard University Press: 22–40. References here are by page numbers in Cole & Morgan.

Groenendijk, J. & Stokhof, M. (1984). Studies on the Semantics of Questions and the Pragmatics of Answers, dissertation, University of Amsterdam.

Happé, F. (1993) Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition* 48.2: 101–19.

Horn, L. (1984), 'Towards a new taxonomy of pragmatic inference: Q-based and R-based implicature'. In: D. Schiffrin (ed.), Meaning, Form, and Use in Context: Linguistic Applications, GURT84, 11–42, Washington; Georgetown University Press.

Lewis, D. (1969). *Convention.* Cambridge, MA: Harvard University Press.

Lewis, D. (1979), "Scorekeeping in a Language Game," in R. Bäuerle et al. (eds.) *Semantics from Different Points of View*, Berlin: Springer, 172-187. Also in *Journal of Philosophical Logic 8*: 339-359.

Parikh, P. (1991). Communication and strategic inference. *Linguistics and Philosophy*, *14*, 473–513.

Parikh, P. (2000). Communication, meaning and interpretation. *Linguistics and Philosophy*, *23*, 185–212.

Parikh, P. (2001). *The Use of Language*. Stanford, California: CSLI Publications.

van Rooy, R. (2003). Being polite is a handicap: Towards a game-theoretical analysis of polite linguistic behaviour. In the proceedings of TARK 9.

van Rooy, R. (2004). Signaling games select Horn strategies. *Linguistics and Philosophy*, 2004, 27: 493–527.

van Rooy, R & Schulz, K. (forthcoming). Pragmatic Meaning and Non-monotonic Reasoning: The Case of Exhaustive Interpretation. (Available from http://turing.wins.uva.nl/~vanrooy/papers.html).

Sally, D. (2003). Risky speech: behavioral game theory and pragmatics. *Journal of Pragmatics, 35,* 1223–1245.

Searle, J. (1969). *Speech Acts*. Cambridge University Press: Cambridge.

Stalnaker, R. (1973), "Presuppositions," Journal of Philosophical Logic 2: 447-457.

Sperber, D., & Wilson, D. (1986/95). *Relevance: Communication and Cognition (2nd edition).* Oxford: Blackwell.

Sperber, D., & Wilson, D. (1997/98). The Mapping Between the Mental and the Public Lexicon. *UCL Working Papers in Linguistics, 9*: 107–25. Reprinted in Carruthers, P. & Boucher, J. (eds.) (1998) *Language and Thought: Interdisciplinary Themes.* Cambridge: Cambridge University Press: 184–200.

Wilson, D & Matsui, T. (1998) Recent Approaches to Bridging: Truth, coherence, relevance. *UCL Working Papers in Linguistics* 10: 173–20

Wilson, D. and Sperber, D. (2002/2004). Relevance theory. In G. Ward and L. Horn (eds), (2004) *Handbook of Pragmatics.* Oxford: Blackwell: 607–632 (References here are to the longer version, Wilson, D. and Sperber, D. (2002) Relevance theory In *UCL Working Papers in Linguistics* 14: 249–287, also available online at http://www.phon.ucl.ac.uk/home/deirdre/home.html).lin