Relevance of Communicative acts

Robert van Rooy
ILLC/University of Amsterdam

Abstract

Why do we speak? Because we want to influence each other's behavior. The *relevance* of a speech act can measure its usefulness. In this paper I argue that (i) the relevance of a speech act depends on the 'language game' one is involved in; (ii) notions of relevance can be defined using decision, information and game theory, and can be used for linguistic applications; and (iii) the strategic considerations of participants in a conversation deserve our attention, especially when we consider mixed-motive games of imperfect information, for instance, to establish the common ground.

1 Introduction

Consider the following example explicitly discussed by Grice (1975) where A is standing by an obviously immobilized car and is approached by B, after which the following exchange takes place:

(1) A: I am out of petrol.

B: There is a garage around the corner.

A can conclude from B's remark that B thinks, or thinks it possible, that the garage is open, and has petrol to sell. The reason behind this *conversational implicature*, according to Grice, is that it is understood between A and B that their communicative acts are *relevant*, and without the implicature B's communicative act would not be so.

The concept of *relevance* is one of the key maxims in Grice's (1975) *pragmatic* account of the 'logic of conversation'. Like the maxims of Quantity, Quality and Manner, also the maxim of Relation (relevance) is used to account for the intuition that more can be inferred from the use of a sentence than what follows from its *conventional meaning.*

Meaning is standardly equated with truth conditions, and a central assumption is that meaning can be determined independently of use. Slowly but surely, it has become increasingly clear that this central assumption is problematic: what is said by a sentence is normally underspecified by its conventional meaning, and depends to a large degree on how it is used. This means that a pragmatic notion like *relevance* is crucial not only to determine what has been implicitly conveyed by a sentence, but also for what has been explicitly said. According to Sperber & Wilson's (1986) Relevance Theory, what is said and what is conveyed both depend on the common assumption that speakers are relevance *optimizers.*

Although relevance seems of great importance, Grice's disparangly vague formulation of the relevance maxim, Be Relevant!, is rather disappointing. Grice fully acknowledges this and indicates some of the difficulties that a better formulation ought to overcome:

> Though the maxim itself is terse, its formulation conceals a number of problems that exercise me a good deal: questions about what different kinds and focuses of relevance there may be, how these shift in the course of a talk exchange, how to allow for the fact that subjects of conversation are legitimately changed, and so on. I find the treatment of such questions exceedingly difficult, and I hope to revert to them in later work. (Grice, 1975, p. 46)

Unfortunately, Grice never came back to it, nor had Sperber & Wilson (1986) a serious attempt to formalize the notion of relevance, although for them the notion became even more important than it was already for Grice (1975). Recently, however, several authors (e.g. Groenendijk, 1999) proposed to explain the relevance of assertions more formally in terms of being a *direct answer* to explicit or implicit questions salient in the discourse. But this proposal has its limitations: (i) it is not general enough, for it seems that an assertion can be relevant, although it doesn't give a direct answer to a salient question; (ii) it can only be applied to *assertions*, and not to other kinds of speech acts like *questions*, *threats* and *acknowledgements*; and (iii) the proposed *qualitative* notion of relevance is just a yes-or-no issue, although, intuitively, relevance comes in *degrees* and thus asks for a *quantitative* analysis. Moreover, the analyses are based on the Gricean assumption that conversation always involves *cooperative* information exchange, but this is only one type of communicative action.

Within argumentation theory (e.g. Walton, 1989), or rhetorics, different kinds of conversations are distinguished: from Inquiries to debates, and from games of persuasion to interactions for problem solving. This suggests, just like Grice's quote above, that the relevance of one's communicative act depends on what kind of conversation one is engaged in: in different kinds of conversations we use language for a different purpose.

Although language seems to be a multi-purpose instrument, the purpose to influence other's behavior seems to be basic. It seems natural to suggest that you want to influence other's people behavior because you believe that when they would change their behavior it would be better for you. According to Game and Decision Theory, the theories of rational human behavior, agents choose their actions by trying to *maximize* their expected *payoffs*, which in turn are determined by their beliefs and preferences. If we want to describe the purpose of language use as influencing other's behavior, it seems we need to make use of these theories of rational behavior.

The purpose of language use depends on the kind of conversation one is engaged in, i.e. on how certain conversational parameters are set. Speaking with Wittgenstein, we can think of different kinds of conversational settings as different kinds of *language games* played between the participants: (i) *Problem solving games* are games of coordination where the goal is to come to a good decision which *action* to perform; (ii) *Inquiries* are alike, except that now only truth of propositions is at stake; no physical actions are involved; (iii) *Debates* might be described as zero-sum games between two participants about which propositions a third party should accept; and (iv) games of *persuasion* can be thought of as games where the preferences are neither strictly opposed, nor strictly in line with each other, but rather mixed, and where the decision problems involve actions to perform.

The purpose of this paper is twofold: First, I would like to show how quantitative notions of relevance can be defined using standard methods of decision and game theory, thereby focussing on the fact that the relevance of particular speech acts depends on the conversational settings; Second, I would like to illustrate *how* these quantitative notions of relevance can be used to explain some linguistic facts.

2 Purely cooperative problem solving games

2.1 Relevance in cooperative problem solving games

In Savage's (1954) classical formulation of Bayesian decision theory, a distinction is made between states of the world, acts, and consequences; states of the world together with acts determine the consequences, each act-world pair has exactly one consequence, and the consequence of an act includes all features that are relevant to the decision maker's values. If we assume that the utility of doing action a in world w is $U(w, a)$, we can say that the *expected utility* of action a, $EU(a)$, with respect to probability function P is

$$EU(a) = \sum_w P(w) \times U(w, a).$$

Let us now assume that the agent faces a *decision problem*, i.e. he wonders which of the alternative actions in A he should choose. A decision problem of an agent can be modeled as a triple, $\langle P, U, A\rangle$, containing (i) the agents probability function, P, (ii) his utility function, U, and (iii) the alternative actions he considers, A.

Suppose now that the agent learns a new proposition C, modeled by conditionalization of the probability function. For each action a_i, its conditional expected utility with respect to new proposition C, $EU(a_i, C)$ is

$$EU(a_i, C) = \sum_w P(w/C) \times U(a_i, w)$$

When our agent learns proposition C, she will of course choose that action in A which maximizes the above value. Then we can say that the utility value of making an informed decision conditional on learning C, $UVL(\text{Learn } C, \text{ choose later})$, is the expected utility conditional on C of the action that has highest expected utility:

$$UVL(\text{Learn } C, \text{ choose later}) = max_i EU(a_i, C)$$

In terms of this notion we can determine the value, or *relevance*, of the assertion C. Referring to a^* as the action that has the highest expected utility according to the original decision problem, $\langle P, U, A\rangle$, i.e. $max_i EU(a_i) = EU(a^*)$, we can determine the utility value or relevance of the *assertion* C, $UV(C)$, as follows:

$$\begin{aligned} UV(C) &= UVL(\text{Learn } C, \text{ choose later}) - UVL(\text{Learn } C, \text{ still do } a^*) \\ &= max_i EU(a_i, C) - EU(a^*, C) \end{aligned}$$

This value, which in statistical decision theory (cf. Raiffa & Schlaifer, 1961) is known as the *value of sample information*, $VSI(C)$, can obviously never be negative. It has a positive utility value only in case it influences the action that the agent will perform. Notice that this condition is not almost trivially met, unless *mixed* actions are allowed. For an early use of this notion of relevance to account for particularized conversational implicatures as described in the introduction, see Parikh (1992).

In terms of the utility value of assertions/answers, we can now determine the utility values of *questions*. Suppose that question Q is represented by the partition $\{q_1, ..., q_n\}$, then we can determine the *expected* utility value of a question, $EUV(Q)$, as the *average* utility value of the possible answers:

$$EUV(Q) = \sum_{q \in Q} P(q) \times UV(q)$$

This value, which in statistical decision theory is known as the *expected value of sample information*, $EVSI$, will never be negative. It will be 0 only in case no answer to the question would have the result that the agent will change his mind about which action to perform, i.e. for each answer q it will be the case that $max_i EU(q, a_i) = EU(q, a^*)$. In these circumstances the question really seems irrelevant, and it thus seems natural to say that question Q is *relevant* just in case $EUV(Q) > 0$. It should be obvious that this measure function also totally orders all questions with respect to their expected utility value.

With respect to a specific decision problem we can compare several questions w.r.t. their usefulness. Can we say something more general about it? According to Groenendijk & Stokhof's

(1984) partition semantics for questions, Q_1 is better than question Q_2 in case the partition Q_1 is *finer* than the partition Q_2, $Q_1 \sqsubseteq Q_2$: $\forall X \in Q_1 : \ \exists Y \in Q_2 : X \subseteq Y$. Denoting by $EUV_{DP}(Q)$ the expected utility value of Q with respect to decision problem DP, Marschak & Radner (1972) have proved the following strong, but also very appealing theorem:[1]

$$\textbf{Theorem} \qquad Q_1 \sqsubseteq Q_2 \quad \text{iff} \quad \forall DP : \ EUV_{DP}(Q_1) \geq EUV_{DP}(Q_2)$$

The 'only if' part is natural, and already implicitly assumed by Savage (1954) and Raiffa & Schlaifer (1961). It shows that it is never irrational (if collecting evidence is cost free) trying to get more information to solve one's decision problem.

The 'if' part is more surprising, and it suggests that the *semantic* entailment relation between questions is an *abstraction* from the more *pragmatic* usefulness relation of questions. The proof is based on the idea that when two partitions are qualitatively incomparable, one can always find a pair of decision problems such that the first partition has a higher expected utility value than the second one according to one decision problem, and a lower expected utility value than the second one according to the other decision problem.

2.2 Linguistic applications: Fixing domain of quantification

The analysis of the previous section to determine the utility of questions was based on the assumption that questions should be represented as *partitions*. As noted above, this corresponds with Groenendijk & Stokhof's (1984) analysis of interrogative sentences. Assuming that the meaning of an interrogative is the set of possible answers that resolves the question, they argue that these possible answers to not only *yes/no*-questions, but also to *wh*-questions are *exhaustive*, and thus mutually disjoint. If $[[A]](w, g)$ denotes the truth value of A with respect to w and g, the meaning of the interrogative represented by $?\vec{x}A$ is defined as the partition $\{\{v \in W | \ \forall \vec{d} \in D^n : \ [[A]](v, g[^{\vec{x}}/_{\vec{d}}]) = [[A]](w, g[^{\vec{x}}/_{\vec{d}}]) \ \}| \ w \in W\}$, where $\vec{e}$ is an n-ary sequence.

In Groenendijk & Stokhof's question semantics a fixed, *context-independent*, domain, D, is assumed over which the *wh*-phrases range. As a consequence, it follows that the partition induced and the semantic notion of a resolving answer are context-independent, too. Ginzburg (1995) and Aloni (to appear) have recently argued, however, that the notion of *resolvedness* is sensitive to the *goals* of the questioner. Adopting the standard assumption that the meaning of a question should be thought of the set of (meanings of) answers that would resolve the question, it follows that the meaning of the question is dependent on the goals of the questioner, too. In van Rooy (1999, 2001) I show how a similar idea can be made precise; the idea that the concept of a resolving answer, and thus the meaning of a *wh*-question, depends on the relevant *decision problem* the questioner faces.

Suppose, for example, that you ask me *Where do you live?* Ginzburg rightly argues that depending on your goal, in some contexts this question might be resolved by an answer like *In Amsterdam*, while in other contexts I should give a more specific, or fine-grained, answer and say something like *In Amsterdam East*.

The most obvious way to account for this fact is to say that the interpretation of a *wh*-interrogative might be *underspecified* by its conventional meaning due to the fact that the *domain* over which the *wh*-phrases range is *context-dependent*. This context-dependence is standardly accounted for by assuming that the relevant domain for each question is simply anaphorically given as a separate feature of the context. But now we can go one step further, and thus give a more *explanatory* analysis of this context-dependence: select the relevant domain by means of the decision problem at stake.

Thus, the idea is that the level of fine-grainedness depends on the decision problem; if we see each other in Germany, and you consider visiting me or not, it would be useless to answer your question *Where do you live?* by saying *In Europe*, while I presumably also need not give my precise

[1] They also relate this theorem with what is known as Blackwell's theorem for comparing information structure.

address. If I meet you in Amsterdam and you want to visit me, however, I should. In abstract, the selected level of precision will be the least one of those levels for which the value of the question asked would be maximal.[2]

3 Inquiries

3.1 Relevance in Inquiries

Inquiries, intuitively, are special cases of cooperative problem-solving dialogues in which only the truth of propositions is at stake. It seems natural to use Shannon's (1949) *information*, or communication, theory to determine the usefulness of speech acts in such conversations. A decision problem, in this theory, can be modeled as the decision which of a set of mutually exclusive hypotheses to accept, and the *entropy* of this decision problem, or question, measures then the *uncertainty* of this decision.

Given a probability function P, we can define the *entropy* of decision problem Q as follows:

$$E(Q) \quad = \quad \sum_{q \in Q} P(q) \times -log_2 P(q)$$

When our agent learns proposition A, we can determine the entropy of decision problem Q *conditional* on learning C, $E_C(Q)$, as follows:

$$E_C(Q) \quad = \quad \sum_{q \in Q} P(q/C) \times -Log_2 P(q/C)$$

In terms of this notion we can now define what might be called the *Relevance* of proposition A, with respect to partition Q, $R_Q(C)$, as the reduction of entropy, or uncertainty, of Q when C is learned:[3]

$$R_Q(C) \quad = \quad E(Q) - E_C(Q)$$

The *entropy of Q conditional on another question Q'*, $E_{Q'}(Q)$, can then be defined as the *average* entropy of Q conditional on learning an answer to question Q':

$$E_{Q'}(Q) \quad = \quad \textstyle\sum_{q' \in Q'} P(q') \times E_{q'}(Q)$$

Shannon (1948) has proven already that for any two partitions of the same state space, for instance our questions Q and Q', it holds that $E_{Q'}(Q) \leq E(Q)$ (Shannon's inequality).

Now we can also determine the relevance of question Q' with respect to question Q. The obvious way to do this is to define it as the *average* reduction of entropy of Q when an answer to Q' is learned:

$$R_Q(Q') \quad = \quad \sum_{q' \in Q'} P(q') \times R_Q(q')$$

Due to Shannon's inequality, it should be obvious that this notion, which in information theory is standardly known as the *mutual information* or *rate of actual transmission* between Q and Q' (see Blahut (1987), for instance), can never be negative.

[2] I should note that domain selection is not only governed by relevance; maximizing relevance should rather by balanced by minimizing *effort* (see Aloni, this volume, for a related, but more qualitative analysis of domain restriction.)

[3] This notion was used by Lindley (1956) already to measure the informational value of a particular result of an experiment.

There seems to be a crucial distinction between the information- and decision-theoretic analyses of relevance, though: whereas in the decision-theoretic analysis we look only at the *optimal* action to do, in the information theoretic analysis we take also all the sub-optimal hypotheses into account, and concentrate on the *whole* probability *distribution*. Can we still think of the one as a special case of the other?

Yes, we can, if we think of individual actions as actions that look at the whole probability distribution. An action should not be thought of simply as a choice of a single hypothesis, but rather as a *choice* of a probability *distribution* over the hypotheses, a *mixed* act.

Bernardo (1979) has show that (i) under some natural conditions the *logarithmic* function is the only function that describes the utility when only truth is at stake; (ii) in that case, relevance as expected reduction of uncertainty (entropy) is, indeed, a special case of the relevance of questions as described in the decision theoretic framework.

3.2 Linguistic applications: generalized conversational implicatures

Groenendijk & Stokhof (1984) argue that answers to questions should normally be interpreted *exhaustively*. Thus, if A answers by saying *Some men* to Q's question *Who are sick?*, we should interpret this normally as the claim that some men are sick and that *noone else* is. Although this strengthening of the answer should intuitively be explained pragmatically, they regret that they don't see how this can be done in terms of Grice's famous maxims of conversation. However, by using the maxim of *maximal relevance* in the information theoretic sense, rather than the Gricean maxims, I indicate in van Rooy (ms) that we can account for this pragmatic strengthening in an intuitively satisfying way: the exhaustive reading of the answer is not only more relevant than the non-exhaustive reading, exhaustification is also the natural strengthening operator that is crucially related with the decision problem/question involved. Perhaps the most interesting consequent of this analysis is that it accounts in a uniform way for all kinds of generalized conversational implicatures (e.g. scalar implicatures: some men are sick $\Rightarrow$ not all men are sick; conditional perfection: B, *if* A $\Rightarrow$ B *iff* A) for which the standard analysis needs various Gricean maxims, and that it does so without the need of *cancellation* of implicature by relevance.[4]

4 Strictly opposing games

Notice that the notion of relevance described in the previous sections presupposes that the game being played is one of pure cooperation: the relevance of the assertion is the relevance for the agent who faces the decision problem and learns the content of the assertion. This can only be the same as for the one who makes the assertion when the latter wants to cooperate with the former. This is in line with Grice's (1975) view on communication, but doesn't cover all kinds of conversations.

4.1 Relevance in strictly opposing games

Ducrot (1973) argued that by making assertions we always want to argue for particular hypotheses, and analyzed linguistic expressions like *but* and *even* in terms of their argumentation orientation. More recently, Merin (1999) proposes to characterize the contexts in which such expressions can be used appropriately in terms of their *argumentative force*, and proposes to implement this argumentative view on language use by means of probability theory. Suppose that an agent wants to argue for hypothesis h, and thus against hypothesis $\neg h$, and that the relevant information state, i.e. the common ground, is represented by probability function P. Notice that h is not independent of proposition C iff learning C increases the probability of h, $P(h/C) > P(h)$. We might say that in this case C is *positively relevant* with respect to h. If $P(h/C) < P(h)$, C would be *negatively*

[4]For a related analysis of particularized conversational implicatures using game theory, see Parikh (1992).

relevant. So we can define the *argumentative force* of proposition C with respect to hypothesis h, $AF_h(C)$, as follows:

$$AF_h(C) \quad \stackrel{def}{=} \quad P(h/C) - P(h)$$

Assuming that an agent wants to argue for proposition h, we can now order other propositions linearly in terms of their argumentative force with respect to h. Thus, we can say that C is a better argument for h than D is iff $AF_h(C) > AF_h(D)$.[5] Notice that $AF_h(C)$ is positive exactly when $AF_{\neg h}(C)$ is negative, which suggests that this is the right notion of relevance in games where the participants argue for diametrically opposing hypotheses.

4.2 Linguistic applications: characterizing licencing contexts

As noted above, Ducrot (1973) and Merin (1999) proposed their notion of argumentative force to account, for instance, for the appropriate use of a linguistic expression like *even* and *but*: we can only say *A and <u>even</u> B* iff claiming A has less argumentative force as claiming B, and *A <u>but</u> B* iff the argumentative forces of A and B are opposing. A potentially even more exciting use of their notion involves the within linguistics widely discussed *polarity items*. Polarity items come in two sorts: *Negative* ones, called NPIs, like *any*, *ever* and *lift a finger*, and *Positive* ones, called PPIs, like *some*, *tons* and *rather*. The linguistically interesting fact about these items is that they can be used appropriately only in certain kinds of contexts. The challenge is to characterize these contexts, and to do so in an explanatory way. According to the traditional *syntactic* approach, NPIs are licenced within the scope of a *negation*, while PPIs are not (where a star, '*', indicates bad use):

(2) a. The dean *didn't* sign *any* of the letters before she left.

b. *The dean signed *any* of the letters before she left.

(3) a. John has *tons* of money.

b. John *doens't* have *tons* of money.

[o.k. as a denial of (a) or with contrastive focus on *tons*]

Because NPIs can also be used appropriately in the antecedent of a conditional,

(4) If *anyone* notices *anything* unusual, it should be reported to the campus police.

Ladusaw (1979) proposed to *semanticize* the analysis by hypothesizing that NPIs and PPIs can be used appropriately only when they occur in more general *downwards* and *upwards entailing* contexts, respectively. A downward/upward-entailing context for α, i.e. an expression $X\alpha Y$, is defined as a context where replacing α with a semantically weaker/stronger constituent β yields a stronger/weaker expression $X\beta Y$. Not only negated sentences, but also antecedents of (standard) conditionals are examples of downward entailing contexts.

Unfortunately, also this purely semantic proposal meets counterexamples. First, because NPIs are also licenced in non-downward entailing contexts (cf. Linebarger):

(5) a. Did Mary *ever* lift *a finger* to help you?

b. John kept writing papers long after he had *any* reason to believe that they would be accepted.

[5] This definition is not exactly the same as the one used by Merin (1999); he in fact uses Good's (1950) function that measures the *weigh of evidence*, a function that is continuously monotone increasing with respect to $AF_h(\cdot)$.

Crucial in these problematic examples is that they share a certain *rhetorical* element: Krifka (1995) hypothesizes that the NPIs in (5a) are used to lower the threshold for a positive answer, showing that the questioner is certain that the answer would be negative, which is what he wants to argue for; in (5b) the speaker wants to strengthen his argument for his claim that John kept writing papers, by claiming that he did so even after he didn't believe anymore that they would be accepted.

This crucial argumentative element is even clearer in some other problematic examples for Ladusaw's proposal: it cannot explain why NPIs are licenced in antecedents of conditional threats, and not promises, while the reverse holds for PPIs (cf. Lakoff, 1970):

(6) a. If you eat *any* LOXO, I'll {batter you/??give you whatever you like}

b. If you eat *some* LOXO, I'll {?batter you/give you whatever you like}

Merin (1994) suggests to account for the licencing of negative and positive polarity items in (6a) and (6b) in terms of negative and positive argumentative force of the antecedents. Suppose that the topic of conversation between speaker and addressee is whether or not the latter will eat LOXO, H. In (6a) the speaker disprefers H and thus wants to argue for $\neg H$. This means that $AF_{\neg H}(H) < 0$ and thus that the antecedent forms a negative argumentative context that licences NPIs. In (6b) the speaker prefers H, and the antecedent thus has a positive argumentative force. For this reason it licences the PPI *some*.

It remains to be seen whether this suggested analysis can be extended: can we analyze most downward entailing contexts as contexts with a negative argumentative force? Merin (1994) suggests a positive answer for standard conditionals. Looking at a *concession* as a type-act with a *negative* argumentative force, $AF_h(\cdot) < 0$, we might think of a standard conditional of the form *If A, then B* as a concession-claim pair. Notice that on the basis of the fact that $AF_h(C)$ is negative exactly when $AF_h(\neg C)$ is positive, such an analysis would also be able to account for the default flip-flop behavior of licencing operators:

(7) a. I would *rather* be in Rome.

b. ??I wouldn't *rather* be in Rome.

c. There isn't any conference participant who wouldn't *rather* be in Rome.

5 Mixed-motive games

Note that we have dealt until now with very specific kinds of games only: purely competitive and purely cooperative games. By looking at these particular kinds of games, we didn't need to make serious use of game theory: to describe the utility, or relevance, of speech acts one needs to look only at one individual decision maker. Another limitations of the above analyses is that it is assumed that the games being played are games of *complete information*: It is assumed that it is commonly known what the preferences are of the participants in a conversation.

To analyze *persuasion* games, however, we have to look at things from a more general perspective; we have to give up the above limitations. Games in which the players' preferences among the outcomes are neither identical, as they are in pure coordination games, nor diametrically opposed, as they are in zero-sum games, are called *mixed motive* games. Games where the beliefs and preferences are not common knowledge are called games of *incomplete* or *imperfect* information.[6] In the remaining part of this paper we will make a start to determine the relevance of different kinds of speech acts in mixed-motive games of imperfect information.

[6]In game theory normally a distinction is made between 'incomplete' and 'imperfect' information. In this paper I will ignore this difference, though.

5.1 Strategic games of perfect information

Where Decision Theory is the theory of individual decision making, Game Theory is the theory of decision in which decision-makers interact. The theory is based on the assumption that each decision-maker, the players, takes into account her knowledge or expectations of the *other* decision-makers' behavior.

The simplest way to look at games is to think of them from a *strategic* point of view. A strategic game, or a game in strategic form, is a model $\langle N, (A_i), (U_i) \rangle$ of interactive decision making in which each player i (element of N) chooses her plan of action (element of A_i) once and for all, and is uninformed, at the time of her choice, of the other players' choices.

What actions will be chosen by the players in a strategic game? This depends on the players' preferences (U_i) over the simultaneous choices of the players. The well established answer is that they play their part of a Nash equilibrium.

A Nash equilibrium of a strategic game $\langle N, (A_i), (U_i) \rangle$ is a profile $a^* \in A$ $(= \times_i A_i)$ of actions with the property that for every player $i \in N$ we have

$$U_i(a^*_{-i}, a^*_i) \geq U_i(a^*_{-i}, a_i) \text{ for all } a_i \in A_i$$

where profile $(a_{-i}, b) = (a_1, ..., a_{i-1}, b, a_{i+1}, ..., a_n)$.

Thus for a^* to be a Nash equilibrium it must be that no player i has an action yielding an outcome that he prefers to that generated when he chooses a^*_i, given that every other player j chooses his equilibrium action a_j. Briefly, no player can profitably deviate, given the actions of the other players.

Two examples: Consider the following two games. In both games $N = \{1, 2\}$, and $A_1 = \{a, b\}$, while $A_2 = \{c, d\}$. In both cases, it is optimal for player 1 (the row-player) to play a when player 2 (the column-player) plays c, and b when 2 plays d. The difference, however, is that in the first game player 2 strictly prefers c to d, while in the second game he strictly prefers d to c. It is easy to check that both games have exactly one Nash equilibrium, but that the equilibria are not the same: In the first game it is the profile (a, c), while in the second it is (b, d). Thus, knowledge of the preferences of *one* player can be crucial to determine for *both* players which action to choose. The games and the Nash equilibria can be easily illustrated by the following matrices:

Game 1:

	c	d
a	4,2	0,0
b	0,4	2,2

Game 2:

	c	d
a	4,0	0,2
b	0,2	2,4

These games have only one Nash equilibrium and it is clear for each agent what to do. In such games, communication seems useless. But now consider games with more Nash equilibria, like the following:

Game 3:

	c	d
a	1,2	0,1
b	0,1	2,3

Ordinal:

	c		d
a	•	←	
	↑		↓
b		→	•

Game 4:

	c	d
a	4,2	0,1
b	1,1	2,4

It is intuitively clear why communication can be useful in game 3: Although profile (a, c) is a Nash equilibrium, they both would gain by playing the other one. Communication can have the result that the Pareto optimal solution will be chosen.[7] In case two NE solutions exist, they can agree which one to play. Game 4 also has 2 equilibria, but here it is not so clear how communication can effect a single solution.[8]

[7] A Pareto optimal solution is the choice of a profile that is optimal for all players looked upon as individual decision makers.

[8] Things change if we allow for *mixed* strategies. In that case the agents can agree to play a so-called *correlated equilibrium* from which none of them has an incentive to deviate (cf. Myerson, 1994).

5.2 Games of imperfect information: Bayesian games

In the analysis of strategic games it is normally understood that every player knows the game she is playing. In particular, each player knows the payoffs of the profiles of each player. This suggests that each player has to know what the state of nature is. But this is not really necessary in order to let a Nash equilibrium be an appropriate solution concept. Suppose, for instance, that game 1 is being played in w_1, while game 2 is played in w_2. Suppose, moreover, that it is unknown to both players what the actual world is: both are commonly known to be equally likely to be played: $P(w_1) = P(w_2) = \frac{1}{2}$. What counts in such a situation are not the actual payoffs in one particular game, but rather the *expected* payoffs. The payoff of profile (b, d), for instance, is $(\frac{1}{2} \times 2 + \frac{1}{2} \times 2, \frac{1}{2} \times 2 + \frac{1}{2} \times 4) = (2, 3)$. In general, the preference relation $\geq_i$ between profiles for each agent i is based on their expected utilities. The game that is being played can then be pictured as follows:

Cardinal:

	c	d
a	4,1	0,1
b	0,3	2,3

Ordinal:

	c		d
a	•	↔	
	↑		↓
b		↔	•

The ingredients of a Bayesian game (cf. Osborne & Rubinstein) are the same as of a normal strategic game, except that we have added now a set of worlds, Ω, plus a probability distribution over these worlds. The payoffs of the profiles are thought of as expected utilities, or *lotteries*. Notice that not only the payoffs of this game are different from the original games, also the expected plays are different: instead of one Nash equilibrium profile we have now two of them.

In the case above we could reduce things to 1 matrix because for both players we assumed the same probability function. But it is easy to see that we could do the same even when the agents assign *different* probabilities to the worlds; for instance, although agent 1 thinks that both worlds are equally likely, the other thinks that the probability distribution of w_1 and w_2 are 0.3 and 0.7, respectively. As long as it is *common knowledge* that the probability distributions of the worlds are like this, both agents could determine the expected utilities of each profile for each agent, and thus also the preference relations and Nash equilibria.

Until now we have assumed that it is common knowledge what (i) the preferences are of each agent in each world, and (ii) the beliefs are of the agents. Now we are going to look at games where these assumptions are given up.

6 Relevance in mixed motive games

6.1 Assertions

Suppose we are in a situation where player 1 knows what the state of nature is, but player 2 does not. Player 2 believes that either w_1 or w_2 might be the actual state, and thinks that w_2 is more likely to be the case than w_1, $P_2(w_1) < P_2(w_2)$. Suppose, moreover, that w_1 is the actual world, and thus that $P_1(w_1) = 1$. The games played in w_1 and w_2 are like before. In this situation player 2 considers it better to play his part of the Nash equilibrium of the second game, (b, d), because this also will be the NE of the Bayesian game. Because player 1 not only knows what the actual world is, but also what player 2's probability function looks like, we can predict that he will also play his part of the Nash equilibrium of w_2. This is unfortunate for him, however, because, given that he knows that w_1 is the actual world, he prefers the Nash equilibrium profile of the game in w_1 to be played. Can he not influence player 2 to play the NE profile of w_1? Yes, he can! By making it common ground that w_1 is the actual world. This is how assertions in mixed-motive games become relevant.

In the above example we have assumed that it is useful for agent 1 to assert what he knows: that w_1 is the actual state of affairs. If agent 2 believes what is asserted, and will act accordingly, this will in these situations be useful for agent 1. In fact, we can even *measure* the usefulness of the assertion (assuming that he will be believed). The *relevance* or *usefulness* of the assertion for agent 1 is the difference between the expected payoff for him in the game where both agents know that w_1 is the actual state of affairs, and the expected payoff in the original game, i.e., the game played according to the original probability functions. It is assumed here that the expected payoff of a game for agent i is the payoff for i of the Nash equilibrium of that game.

This notion of relevance can be generalized to all kinds of assertions. Let C be the content of an assertion. Then we can determine the relevance for speaker i to assert C as

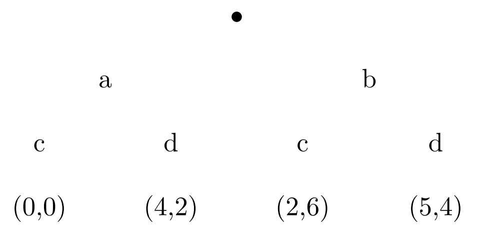
$$UV_i(C) \quad = \quad EU_i(a^n/G_C) - EU_i(a^n/G)$$

In this formula a^n/G denotes the Nash equilibrium profile of game G, where G itself is a Bayesian game of the form $\langle N, \Omega, (A_i), (P_i), (U_i)\rangle$. The expected utility for agent 1 of profile a, $EU_1(a)$, is of course defined by $\sum_{w\in\Omega} P_1(w) \times U_1(a, w)$, where U_i is a function from profiles and worlds to numbers. We will say that G_C is the same game as G except that in this new game it is common knowledge that C is the case, i.e. the probability functions of both agents are, and known to be, conditionalized by C.[9]

6.2 Threats and promises

Although we started out with assertions, when we want to analyze the relevance of speech acts in mixed-motive games, the first communicative acts that really come to mind are promises and threats.

Consider the following *extensive* game, where player 1 has to choose between a and b, and player 2 between c and d afterwards. Notice that in the *strategic* formulation of the game on the right-hand side, player 2 has not 2, but rather 4 strategies: his strategy cd, for instance, means that he chooses c in case a is played, and d in case of b.



Strategic:

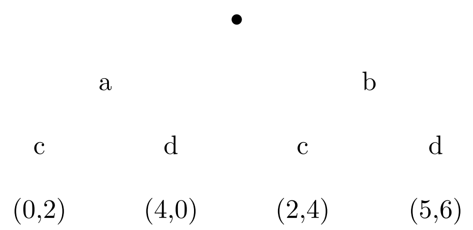
	cc	cd	dc	dd
a	0,0	0,0	4,2	4,2
b	2,6	5,4	2,6	5,4

This game has two Nash equilibria, $\langle a, dc\rangle$ and $\langle b, cc\rangle$, but only the former one is subgame perfect. However, player 2 prefers player 1 to play b. He can now try to 'convince' the first player to do so by *promising* to play d, if player 1 plays b. When player 2 fulfills his promise, it would indeed be rational for the first player to play b: it would give him 5 utils if player 2 plays rational, instead of the expected 4. Alternatively, player 2 can *threaten* to play c, if player 1 plays a. Also in this case it seems that player 1 is better off by playing b: it would give him a payoff that is always higher than the 0 utils he would get by playing a if player 2 caries out his threat.

Unfortunately for player 2, the first player has no reason to take either the promise or the threat seriously: he knows that when player 2 is a utility-optimizer he neither will fulfill his promise nor his threat in case the conditions are met. Both are *incredible*.

[9]What if there is more than one Nash equilibrium? If we stick to our strategic form games, one simple proposal might be to say that if there are n Nash equilibria, it can be expected that the chance that a particular Nash equilibrium profile will obtain is $\frac{1}{n}$.

The above discussion was based on the assumption that the game was one of complete information: in particular that player 1 knows the preferences of player 2. But now suppose that player 1 thinks that either the above game is played, in w_1, or the one described on the left-hand side below, in w_2, and both with equal probability:



Strategic:

	cc	cd	dc	dd
a	0,1	0,1	4,1	4,1
b	2,5	5,5	2,5	5,5

As far as player 1 is concerned, he and player 2 are playing a game of imperfect information and its strategic form is given at the right-hand side. This game has 4 Nash equilibria: $\langle a, dc\rangle$, $\langle b, cc\rangle$, $\langle b, cd\rangle$, and $\langle b, dd\rangle$, and all of them happen to be subgame-perfect too. In this game player 2 seems indifferent between playing c and d, and both choices seems to be equally likely according to player 1. But this suggests that the latter will play a, because this has on average a higher payoff (4 against $3\frac{1}{2}$).[10] Player 2, however, knows his own preferences, and knows that the first game is being played. But he also knows what player 1 knows, and thus expects player 1 to play a. This is undesirable for him, and he wants to change the behavior of the first player.

Suppose that in this situation he makes the conditional promise or threat described above: 'I will play d, if you play b' and 'I will play c, if you play a'. Should player 1 take the promise and threat seriously now and play b instead of a? We have seen already that in the first situation the promise and threat are *incredible*. In the second situation (w_2), however, they are not: the second player has no reason not to carry out his 'obligations', and if he does player 1 would be better off. Notice that also in the game of imperfect information the promise and threat make sense. As a result, he might well be tempted to play b instead of a: the promise and threat are *credible*. So, in this case these speech acts can actually influence the choice of the other participant, and are thus *relevant*. Because it is actually the first game that is being played, and known to be played by player 2, the speech acts have a *positive* relevance of 4 utils. Player 2's resulting payoff will not be 2 utils (because 1 plays a and 2 d), but rather 6 utils (because 1 plays b and 2 c). Poor player 1: due to his imperfect information he can be manipulated, and loses 2 utils.

6.3 Negotiating common ground

In the sections 2, 3 and 5.1 we have seen that it might be good to share information because it makes available better potential agreement. But discussing the relevance of assertions we saw that revealing all information to an opponent is not always the most advantagegeous policy (see the immense literature on signalling in economics, e.g. Crawford & Sobel). Afterwards we considered the relevance of threats and promises in extensive games, and saw that although sometimes these speech acts have to be taken seriously, at other times they can be rejected. So, both speaker and hearer have choices to make, guided by strategic consideration. But on closer investigation it turns out that something similar happens with assertions. The reason is that what becomes common ground is crucial for the payoffs of both agents.

The importance of the common ground is well established within pragmatic analyses of natural language. Both speaker and hearer rely on it: the speaker for his choice of words; the hearer for how she should interpret the message of the speaker. The ease of their communication depends to a

[10]Despite the fact that average payoffs don't play a role in determining the NEs of Bayesian games.

large degree on the amount of common ground between speaker and hearer. Despite the interest of linguists in the common ground, another important reason why common grounds are so important for interactive behavior has been widely neglected here: that the actual common ground is crucial to determine which actions the agents will perform. That is, linguists have widely neglected the fact that agents *prefer* certain common grounds above others, because the expected play (the Nash equilibrium profile) depends on the actual common ground.

Looking back at the definition of an assertion's utility of section 3.1 we see that in principle they can have a *negative* value for the speaker. Moreover, the utility does not only concern the speaker, it also gives a (positive or negative) utility for the *hearer*. The payoffs of the Nash equilibrium of the new game can be both higher and lower than the NE of the old game, and this holds not only for the speaker, but also for the recipient. This gives obviously rise to new kinds of *strategic considerations*: If I have information that would be relevant for the payoffs if it would become common knowledge, it is only rational to communicate this when the utility value of asserting it would be positive for me. But also the recipient has real strategic power in such a game: if accepting an assertion C would result in a Nash equilibrium profile being played that has for him a lower payoff, he might well do *as if* he does not accept the assertion, such that C does not become common knowledge, and the old game is still being played. Notice that in this case even if the recipient does believe the assertion, and would conditionalize his beliefs, he should still play the game with respect to his old beliefs, because that is what the speaker expects (given the fact that the recipient did not accept by means of *acknowledgement* and the common ground remains the same).

The one who makes the assertion has lots of possible actions to consider: he can claim whatever he wants. The recipient has even more strategies: for each assertion made, he can either (pretend to) accept or reject it. From a theoretical perspective we are not so much interested in what the speaker will actually say, but rather in the question what he could say (i.e. *offer* to make common ground) that has a good chance to become common ground (*accepted* by the recipient). The result of this 'negotiation' can be analyzed as an equilibrium of a *bargaining game* between agents.

References

[1] Aloni, M. (to appear), 'Questions under cover', in D. Palmer-Plummer, et al (eds.) *Proceedings of LLC 8*, CSLI, Stanford, CA.

[2] Bernardo, J.M. (1979), 'Expected information as expected utility', *Ann. Statist.* **7**: 686-690.

[3] Blahut, R. E. (1987), *Principles and Practice of Information Theory*, Addison Wesley, Massachusetts.

[4] Crawford, V.P. & J. Sobel (1982), 'Strategic information transmission', *Econometrica*, **50**: 1431-1451.

[5] Ducrot, O. (1973), *La peuvre et le dire*, Paris: Mame.

[6] Ginzburg, J. (1995), "Resolving Questions, I", *Linguistics and Philosophy*, **18**: 459-527.

[7] Grice, H. P. (1975), 'Logic and Conversation', In: P. Cole & Morgan (eds.), *Syntax and Semantics 3: Speech Acts*, New York: Academic Press.

[8] Groenendijk, J. & M. Stokhof (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, Amsterdam.

[9] Groenendijk, J. (1999), 'The logic of interrogation', In: T. Matthews & D.L. Strolowitch (eds.), *The Proceedings of the Ninth Conference on Semantics and Linguistic Theory*, CLC Publications, Stanford.

[10] Good, L.J. (1950), *Probability and the Weighing of Evidence*, Griffin, London.

[11] Harsanyi, J.C. (1977), *Rational behavior and bargaining equilibrium in games and social situations*, Cambridge University Press, Cambridge.

[12] Krifka, M. (1995), 'The semantics and pragmatics of polarity items', *Linguistic Analysis*, **25**: 209-257.

[13] Lakoff, R. (1970), 'Some reasons why there can't be any some-any rule', *Language*, 45, 608-615.

[14] Lindley, D. V. (1956), "On a measure of information provided by an experiment", *Ann. Math. Stat.*, 29, pp. 986-1005.

[15] Linebarger (1991), 'Negative polarity and grammatical representation', *Linguistics and Philosophy*, **10**: 325-87.

[16] Marschak, J. & M. Radner (1972), *Economic Theory of Teams*, Yale University Press, New Haven.

[17] Merin, A. (1994), 'Algebra of elementary social acts', In: S. L. Tsohatzidis (ed.), *Foundations of Speech Act Theory*, London: Routledge.

[18] Merin, A. (1999), 'Information, relevance, and social decisionmaking', In: L. Moss et al. (eds.), *Logic, Language, and Computation, Vol. 2*, Stanford.

[19] Myerson, R.B. (1994), 'Communication, correlated equilibria and incentive compatability', In: R. J. Aumann & S. Hart (eds.), *Handbook of Game Theory, Vol. 2*, Elsevier.

[20] Osborne, M.J. & A. Rubinstein (1994), *A Course in Game Theory*, MIT Press, Cambridge.

[21] Parikh, P. (1992), 'A game-theoretical account of implicature', in Y. Vardi (ed.), *Proceedings of the fourth TARK conference*, Monterey, California.

[22] Raiffa, H. & R. Schlaifer (1961), *Applied Statistical Decision Theory*, MIT Press, Cambridge, MA.

[23] Rooy, R. van (1999), 'Questioning to resolve decision problems', In P. Dekker (ed.), *Proceedings of the Twelfth Amsterdam Colloquium*, ILLC, Amsterdam.

[24] Rooy, R. van (2001), 'The pragmatics of mention-some readings of questions', in H. Bunt et al. (eds.), *Proceedings of IWCS-4*, Tilburg.

[25] Rooy, R. van (ms), *Relevant Implicatures*, University of Amsterdam.

[26] Savage, L.J. (1954), *The Foundations of Statistics*, New York: Wiley.

[27] Shannon, C. (1948), 'The mathematical theory of communication', *Bell System Technical Journal*, **27**.

[28] Sperber, D. & D. Wilson (1986), *Relevance; Communication and Cognition*, Blackwell, Oxford.

[29] Walton, D.N. (1989), *Informal Logic. A Handbook for Critical Argumentation*, Cambridge University Press, Cambridge.

[30] Wittgenstein, L. (1958), *Philosophical Investigations*, trans. G.E.M. Anscombe, Oxford, Blackwell.