UNCERTAINTY IN THE COMMON GROUND

Marie Nilsenová
ILLC, University of Amsterdam

December 2000

Abstract

This thesis is about the uncertainties that naturally form a part of human communication. Language as a communication medium is inherently ambiguous. It is generally assumed that its use is made possible by the existence of common ground among communicating participants, but the question of how common ground comes about is usually left aside. Descriptions of natural conversations suggest that in view of the noisiness of the channel, agents constantly employ means to ensure that changes in the mutual context took place, i.e., they attend to a "dialogue about the dialogue". A particular set of communicative acts used on this level is called grounding. Certain grounding acts pose difficulties to existing formalizations of rational communicative behavior, because they come out as improper or irrelevant. This is an undesirable effect considering the fact that they belong to the most commonly performed dialogue acts. To solve the difficulty, we propose an extension of two recent decision-theoretic approaches to communication. We take it that communicating agents base their actions on probabilistic beliefs about what forms the set of mutually believed propositions. During the interpretation process, the relevance of each utterance is evaluated against the interpreter's epistemic state, represented in a Bayesian model. Subsequent belief changes can ensue, serving as the basis for planning further communicative actions. The analysis can account for the relevance of different types of acknowledging acts, and potentially for other problematic phenomena on the semantic-pragmatic interface.

Contents

1 Introductory Remarks

When we communicate, we expect from each other a certain amount of preexisting knowledge. Specifically, it is the knowledge of the language that serves as a communication medium, more or less identical perceptual experience of the surrounding environment (possibly enriched with some socially determined information) and knowledge of causal relations that hold in our world. This kind of information is often referred to as the conversational *common ground*. On one hand, speakers are supposed to use it to make their contributions efficient, i.e., such that they fulfill the intended discourse purposes with the least necessary effort. On the other hand, hearers rely on it in the interpretation process, in order to disambiguate speaker's utterances and often to increase their information value beyond its literal meaning (as in the case of implicatures, see [Grice 1989], further discussed below).

Simple examples that illustrate reliance on mutual context in natural dialogues are sentences containing anaphoric expressions. Consider the exchange below:

(1) A: *Look, the man is trying to push that oak down!*
B: *He's just a jogger stretching his calf muscles.*
A: *I saw him yesterday roller–skating.*

Note that the definite description *'the man (trying to push that oak down)'* is in subsequent contributions of both A and B replaced with the pronoun *'he'/'him'*. Sentences containing anaphoric pronouns are potentially ambiguous since in different contexts the pronouns can refer to different individuals. Nevertheless, when B uses the pronoun in (1), presumably to make the production of her utterance least costly in terms of effort, [1] she expects A to

[1]Note that the low costs of using an anaphoric pronoun cannot be directly related to Zipf's universal principle of *Least Human Effort* by assumed **articulatory** costs. As pointed out by [Levinson 1985], pronouns are used also in instances when their usage does not result in any articulatory reduction - cf. (i) and (ii).
(i) Ann came in. *Ann* sat down.
(ii) Ann came in. *She* sat down.
Moreover, there is no detectable general tendency towards 'telegraphic speech', in fact, under many circumstances speakers prefer to speak at length. Levinson proposes instead to account for the lower costs associated with the use of pronouns, in terms of semantics/informativeness.

be able to determine its reference: in this case the currently most salient male individual. Anaphora resolution clearly would not be possible unless the context of the dialogue, or its common ground, was actively used in the communicative situation by all of the communicating agents.

Common ground is also necessary to derive conversational implicatures. Take the well–known example of a philosophy professor writing a letter of recommendation for one of his students [Grice 1989]. Instead of discussing the student's intellectual qualities and work attitude, the professor points out that the man has a wonderful handwriting. The intended effect on the reader is to make her realize that there is not much else the student can be recommended for. Due to politeness phenomena, i.e., in order to save the student's 'face' [Brown & Levinson 1978], the professor restrains himself from saying so openly. Nevertheless, he assumes that the implicature will be derived by the recipient of the letter as intended. In this assumption, the professor relies on the expected mutual knowledge of what forms a relevant content of recommendation letters, as well as principles of human interaction. Derivation of the implicature is then quite straightforward in terms of Grice's conversation principles.

Common ground is used also for more complex inferential processes. Since agents assume that a track record of what has been said in the conversation forms a part of it (this is one of the reasons why [Lewis 1979] refers to common ground as the *conversational scoreboard*), they can draw conclusions about other agents' reasoning. Another well-known example that can illustrate how mutual context serves this particular purpose, is the Muddy Children Puzzle [Fagin et al. 1995]:

> Imagine n children playing together. The mother of these children has told them that if they get dirty there will be severe consequences. So, of course, each child wants to keep clean, but each would love to see the others get dirty. Now it happens during their play that some of the children, say k of them, get mud on their foreheads. Each can see the mud on others but not on his own forehead. So, of course, no one says a thing. Along comes the father, who says, *'At least one of you has mud on your forehead,'* thus expressing a fact known to each of them before he spoke (if $k > 1$). The father then asks the following question, over and

> over: *'Does any of you know whether you have mud on your own forehead?'*

What will happen? Let us consider the situation when there are three children altogether and two of them have mud on their foreheads, i.e. $n=3$ and $k=2$. Any of the children answers *'no'* if he or she is the first, because each can see that there is mud on at least one child's face. But after one of the *muddy* children answers *'no'*, the other muddy child realizes that he must be muddy himself, since otherwise his sibling would have answered *'yes'*. The Muddy Children Puzzle illustrates common knowledge at work. Apart from the fact that everybody has to know that everybody else knows that 'At least one of us has mud on your forehead,' to be able to arrive at the conclusion as above, there are other assumptions that have to be mutually known. In particular, all the children have to know that the others are intelligent (they can make the right inferences), perceptive (they have heard their father and they can see mud if there is some) and truthful. If one of the children were lying, the whole process of reasoning employed by the others would break down.

Suppose someone would ask one of the children, Jim, whether his brother, Jack, always speaks the truth. Jim would say *'no'*, his brother actually lies quite often. But does that mean that in the introduced scenario (all other preconditions satisfied), Jim would not be able to reach the correct conclusion? Not necessarily; even though Jim knows that Jack does not always speak the truth, he could assume that in this particular situation he probably, though not certainly, did. This does not mean, however, that Jim *knows* that Jack is speaking the truth. In other words, Jim participates in a communicative situation that contains a certain degree of uncertainty. The uncertainty can be further increased under possible modifications of the scenario. For example, the father's character could be flawed: as already mentioned, it is crucial for the children that their father's utterance *'At least one of you has mud on his face'* is commonly known - but it also has to be true! Or, there could be a price/lack of punishment for the child that first finds out whether he or she is muddy. In such a case, Jim might consider it more likely that Jack is lying.

What the Muddy Children Puzzle shows is that even in a very simple communicative situation uncertainty abounds and that it comes in degrees. Undoubtedly, presupposed mutual knowledge plays an important role for communicating agents, but it does not need to be, in fact *cannot* be, perfect. For instance, if common knowledge that everybody speaks the truth (and nothing but the truth) were a prerequisite for successful communication, the communicating agents would be aiming at something they would not necessarily be in a position to achieve.

The assumption that agents share information available to their perceptual mechanism in the conversational setting is another potential source of uncertainty. To take a trivial day-to-day example, it cannot be taken for granted that everybody heard or heard correctly what was uttered by other agents. Less trivially, agents may consider as salient different phenomena in general [Sperber & Wilson 1986]. As a consequence, they would create different hypotheses about causal relations between events. In the extreme cases, pathological conditions cannot be excluded. [2]

Language itself as a communication medium is unreliable. While there must be an overlap between the structure of our language and the language of people we communicate with, neither our syntax nor our semantics are completely identical (not even when we happen to be native speakers of the same language). Evidence for this fact ranges from conflicting grammaticality judgements (on a minor scale reflected in varying preferences for parsings of the same sentence), and the mechanism of language change [Lightfoot 1991], to differences in concepts speakers can associate with a particular lexical item. For these reasons, even correctly heard propositions expressed by truth–speaking participants in one communicative situation cannot be considered common knowledge: there is always some degree of doubt whether the speaker really intended her utterance to mean what the audience thinks it meant (and vice versa).

To sum up, the content of the common ground is elusive.[3] Arguably, a model of human communication should take this fact into account. In fact, we will attempt to illustrate on one type of pragmatic phenomena that in order to account for it, it is *necessary* to make uncertainty a part of the human communication model.

[2][Sacks 1995] gives an example of patients with acquired blindness who, nevertheless, answer when asked what they see in front of them on TV.

[3]See also [Traum 1994], who reaches the same conclusion.

In the **second chapter** of the thesis, we discuss some representations of common ground. The literature on this topic is vast and an exhausting overview is beyond the scope of this thesis. Instead, we tried to focus on the most often discussed approaches. The proposals concerning mutual knowledge or information contained in the common ground sometimes limit themselves to the question of properly characterizing the phenomenon, leaving aside the issue of how it actually comes about. This approach can lead to difficulties. As pointed out by [Barwise 1987], any description of common knowledge should be compatible with an explanation of how it arises and is used. In the discussion of this problem we are taking the conclusions of [Barwise 1987] and [Clark 1996] as the basic reference.

In the **third chapter**, we turn to the structure of dialogues, discussing first the speech act theory, followed by more recent approaches that can be found in AI literature. We argue for ascribing a more active role to the hearer or interpreter in natural dialogues and introduce the notion of *grounding*, first described by [Clark & Schaefer 1989]. Specifically, attention will be payed to the most common dialogue act, which is acknowledging. We give examples of various types of acknowledgements and relate them to Clark's [1996] *action ladder*, an entailment relation between types of communication acts.

In the **fourth chapter** we discuss various formal approaches to relevance in dialogues, namely qualitative dynamic semantics [Groenendijk 1999] and the quantitative decision–theoretic approach of [van Rooy 2000 a,b,c]. We argue that in order to account for the facts introduced in the previous chapter, some modifications of the theories need to be made. For qualitative dynamic semantics, the pertinence of acknowledgements in general can be accounted for with some additional assumptions, but it leads to a paradox. A natural extension of the qualitative dynamic approach can be seen in the quantitative theories. [Gmytrasiewicz & Durfee 2000] offer a general way of calculating the relevance of acknowledgement acts, but essentially build on the notion of *noise* without giving it a real content.

Therefore, in the **fifth chapter**, we focus on analyzing the possible sources of noise in human communication. Basing the account on a Bayesian epistemic model, we relate different types of acknowledgements to different kinds of noise. Taking the decision–theoretic approaches introduced in the previous chapter as a starting point, we offer a possible way of calculating the relevance of communicative acts in a general sense, i.e., of both speaking and interpreting.

The proposal is summarized in the **conclusion** to the thesis. We discuss possible future elaborations of the framework in connection to game theory. The concluding part is followed by (a different sort of) *Acknowledgements* and by references to relevant literature.

2 Common Knowledge

'...while common knowledge is widely recognized to be an important phenomenon, there is no agreement as to just what it amounts to. Or rather, as we will see, what agreement there is presupposes a set of simplifying assumptions that are completely unrealistic.' John Barwise

2.1 Introduction

In virtue of the fact that communication takes place in a noisy environment, it is a result of a joint effort to encode and decode the transmitted message in the most efficient way. This is made possible by the use of the conversation's common ground. Each communicative act is both affected by and affects the content of the common ground. It is desirable that any definition of common knowledge be flexible enough to accommodate this particular property, together with a number of other features pointed out below.

The main definitions of common knowledge that have been discussed in the literature are the **iterative**, the **reflexive** and the **shared basis** approach. In this chapter, it will be argued that of these three, the last one appears to be the most suitable for purposes of describing human (language) interaction.

2.2 Coordinating Attack

Consider a simple case in which there are two possible messages, m_1 and m_2, the first of which can have both the content c_1 and c_2, while m_2 can only mean c_2. If we define, following [Hasida, Nagao & Miyata 1995], the encoding E as a binary relation between the set of possible contents and the set of possible messages,

$$E \subseteq C \times M$$

then we get the pairs $\langle m_1, c_1 \rangle \in E$, $\langle m_1, c_2 \rangle \in E$, and $\langle m_2, c_2 \rangle \in E$. Wanting to communicate the content c_2, the speaker can choose between the pairs $\langle m_1, c_2 \rangle \in E$ and $\langle m_2, c_2 \rangle \in E$. It is more rational to choose the unambiguous encoding m_2, rather than m_1, risking that the message will be misinterpreted. In turn, a hearer following this line of reasoning would tend to interpret m_1 as meaning rather c_1 than c_2, coordinating her interpreting behavior with the speaker's. Note that coordination of communication presupposes that agents share a number of beliefs.

In other cases, however, coordination can be harder to achieve. What happens, for example, when there are two possible messages, m_1 and m_2 each with two possible meanings c_1 and c_2? As illustrated by [Hasida, Nagao & Miyata 1995], the tree of inferences performed by the interpreter is in such case (with lack of outside knowledge) potentially infinite. For example, in case the speaker states m_1, the hearer will try to decide whether the communicated content is supposed to be c_1 or c_2. To decide, she would attempt to coordinate her behavior with that of the speaker and try to establish the most optimal encoding of c_1 and c_2, respectively. However, for both of these cases, there are two equally optimal encodings, namely m_1 and m_2. Thus we receive a binary branching infinite tree where each message leads to two possible contents and each content again to two possible messages.

The situation can be compared to the problem of coordinating attack. In this scenario, there are two generals who want to defeat their mutual enemy. If they succeed to attack simultaneously, they will achieve victory. On the other hand, if one of them attacks later, the enemy will beat them both. Aware of this situation, general A sends a messenger to general B, saying: *'I will attack at dawn.'* Unfortunately, the messenger has to carry general A's message through the enemy's area and general A thus cannot be certain that his message really got through. General B, upon receiving the message, realizes that it is not enough that he knows when general A is about to attack; it is also necessary that general A knows that general B knows this. Therefore, he sends the messenger back with the message *'I know that you will attack at dawn.'* General A receives this information - but is it enough? Now A knows that B knows what A wants to do, but B does not know whether A knows. This means that it is still not safe for A to attack. He can send a message back, saying *'I know that you know that I will attack at dawn'*...but obviously, this is still not enough. As long as the information that 'A will attack at dawn', and that both A and B know that, is not simultaneously believed by both generals, they cannot take the decision to attack risk-free.

Communication is a joint activity in a sense similar to the case of attacking generals. Since natural language is inherently ambiguous, the same representation in different contexts conveys different meanings. However, all communicating agents are expected to cooperate with each other ([Grice 1989], [Clark 1996], [Groenendijk 1999], a.o.), either because success of the communication is to their immediate advantage and/or because of the long-term advantage associated with cooperative communicative behavior. [4] The speaker sends out a message that she wants the hearer to interpret in a certain manner, and also the hearer wants to understand what the speaker intended to communicate. Aware of the fact that language is an unreliable medium, the speaker, when planning her utterance, has to design her speech act by taking interpretation into account [Dekker & van Rooy 2000], [Hasida, Nagao & Miyata 1995]. In other words, the speaker has to coordinate her action with the action of the hearer, and vice versa.

Like generals attempting to defeat a mutual enemy, the communicating agents lack the resources to endlessly confirm that a communicated message was received and interpreted in the way intended by the speaker. Therefore, the process of planning and interpreting an utterance has to be done in the most efficient manner, by drawing upon previously established shared knowledge, i.e. by making use of the common ground.

2.3 Common Ground

What is common ground? [Lewis 1979] compares it to a score board in a game of baseball, on which information about the players/teams and their actions in the current game is noted down and continuously updated. This metaphor can be misleading, however. In conversation, common information is not available in the form of a single independently existing board which agents can check any time they need to. Common ground must be an inherent part of the agents' belief states simply because it cannot exist independently of it [Clark 1996]:

[4] It can be disadvantageous to take Grice's principles of cooperative communication behavior (in particular the Maxim of Quality, viz chapter 4) as a given, in view of the fact that they can be disregarded (e.g., in situations when agents attempt to deceive each other, or when they decline to share requested information). An alternative is to analyze the principle as a long–term commitment to collaborate, made by a selfish utility maximizing agent in view of her future expected outcomes. At any point in the interactive process, the agent has the option of defaulting from the team-related commitment, for the sake of a more beneficial individualistic opportunity (see the work of B. Grosz and S. Kraus, a.o. [Grosz & Kraus 2000], for details).

> [W]hen [we] act ”on the basis of our common ground”, we are in fact acting on our individual beliefs or assumptions about what is in our common ground.

This also means, however, that the agents' ideas of what the common ground contains may differ [Stalnaker 1998]. Various kinds of information can be assessed as more or less likely to be contained in it (e.g., what has been said so far in the conversation is more reliable mutual context than what the agents read independently in the newspaper). Yet the propositions contained in the common ground do not neccesarily need to be believed by any of the participants; the agents can be trying to deceit each other or simply establishing as true certain facts only for the purposes of the communicative situation. As discussed by R. Stalnaker [Stalnaker 1978], a proposition is a part of the common ground,

> if the speaker is disposed to act as if he assumes or believes that the proposition is true, and as if he assumes or believes that his audience assumes or believes that it is true as well.

In general, during the process of communication, the content of common ground is subject to constant readjustment by means of explicit updates, as well as presupposition accommodation. The definition of common ground for natural communication has to be flexible enough to account for all the above mentioned properties. In the next subsection, various definitions of common knowledge will be discussed from this perspective.

2.4 Definitions of Common Knowledge

Following the work of [Barwise 1985], [Clark 1996] and [Gerbrandy 1997], we will focus on three definitions of common knowledge. Generally, the most common approach appears to be the iterate one, but there is some disagreement as to who actually takes the position.

(I) **Iterated Propositions**
σ is common knowledge between agents p and q iff:
1. p knows σ,
2. q knows σ,
3. p knows that q knows σ, q knows that p knows σ, p knows that q knows that p knows σ, and so forth, *ad infinitum.*

According to [Barwise 1985], D. Lewis uses the iterative definition of common knowledge to account for the use of conventions [Lewis 1969] (on the other hand, [Clark 1996] and [Gerbrandy 1997] consider the author to adopt the **shared basis** approach, which indeed appears to be closer to Lewis's original text (see below)). According to [Clark 1996], another proponent of the iterate approach is [Sperber & Wilson 1986]. In fact, the authors explicitly argue against this definition of common knowledge on cognitive grounds: human beings obviously lack the capacity to perform the computation required by it. In case of only a finite number of embeddings, the definition proves to be too weak, a fact exemplified by the Coordinating Attack scenario. For natural language, potentially infinite representation is necessary, e.g., for interpretation of definite descriptions, as shown by [Clark & Marshall 1981]. One way of avoiding the difficulty of processing costs is to assume that the definition describes an idealized state of affairs, where in reality the agents only perform the computation up to a certain degree of embeddings. In that case, there has to exist a property which allows the agents to add another layer of embedding in case it is needed. This is the approach to common knowledge referred to as **reflexive** or as the **fixed point**:

(II) **Reflexive**
σ is common knowledge between agents p and q iff *(i)* p and q know σ and that *i*.

As shown by [Barwise 1985], the iterate approach and the reflexive one are not equivalent, if the distinction between *knowing p* and *having the information that p available* – where the first one is stronger than the second – is made. An agent is able to act based on a certain information only if she *knows* she has the information. One example, often discussed in the AI literature, is that of implication. It can be the case that $\alpha \rightarrow \beta$ and that the agent knows that α, and yet does not believe that β because she does not know that one implies the other. In this sense, she has the information that β, but not the knowledge. Another example is that of *de re* belief attributions, originally discussed by [Quine 1956]. If an agent, call him Ralph, sees a suspiciously behaving man with a brown hat, he can arrive at the knowledge that the man is a spy. Now if the man happens to be called Ortcutt but Ralph does not know that, then Ralph has the information that Ortcutt is a spy, but in fact he does not *know* that Ortcutt is a spy.

However, as pointed out by [Clark 1996], the reflexive definition of common knowledge does not make clear on which basis the knowledge is arrived at. The last approach avoids this difficulty, as well as the processing costs associated with the iterated definition in (I) by assuming the existence of a shared basis of beliefs (see, e.g., [Clark & Marshall 1981]).

(III) **Shared Basis**
p and q have common knowledge of θ just in case there is a situation s, such that:
(i) $s \models \theta$,
(ii) $s \models p$ knows s,
(iii) $s \models q$ knows s,
where $s \models \theta$ is a notation for θ is a fact of s.

To use Schiffer's original example, consider a situation s in which p and q are sitting at a table with a burning candle in their view. Then the fact θ that there is a burning candle on the table is common knowledge between p and q, by virtue of the fact that they can see they can both see it. It seems indeed that this is the approach taken by [Lewis 1969], who defines common knowledge as (III'), where δ stands for the commonly known proposition and A refers to the situation that forms shared basis for deriving common knowledge.

(III') **[Lewis 1969]**
It is common knowledge in population P that δ iff some state of affairs A holds such that:
(1) Everyone in P has reason to believe that A holds.
(2) A indicates to everyone in P that everyone in P
has reason to believe that A holds.
(3) A indicates to everyone in P that δ.

What does the word *indicate* mean for Lewis? The interpretation offered by him closely resembles the condition included in the reflexive approach:

> Let us say that A *indicates* to someone x that ___ if and only if, if x had reason to believe that A held, x would thereby have reason to believe that ___. What A indicates to x will depend, therefore, on x's inductive standards and background information.

It seems clear from the citation that the fact that a certain situation provides clues to members of a particular population or participants in a dialogue does not necessarily mean that everybody will successfully arrive at the indicated knowledge. This is one of the features distinguishing between the reflexive definition and the shared basis. Nevertheless, some authors consider the definitions equivalent.

2.5 Shared Basis

According to [Gerbrandy 1997], the reflexive and the shared-basis definitions of common knowledge are equivalent if formalized in the classical possible world semantics (see *ibid* for proof).[5] Gerbrandy expresses the two definitions formally as below. The agent's a information state in a possibility w is modeled by the set of possibilities v that are accessible from w, i.e., the worlds that the agent a considers possible in w (a possibility is a function that maps propositional letters to truth values and information states to agents).

• **Fixed point (Reflexive):**
$w \models C^{fix}\phi$ iff $\forall a \forall v \in w(a) : v \models \phi$ and $v \models C^{fix}\phi$

• **Shared:**
$w \models C^{share}\phi$ iff there is a set of possibilities σ such that:
(1) $v \in \sigma \Rightarrow v \models \phi$
(2) $v \in \sigma \Rightarrow v(a) \subseteq \sigma$ for each a
(3) $w(a) \subseteq \sigma$ for each a

Note, however, that Gerbrandy's formalization does not distinguish between a situation merely *indicating* that a certain proposition holds and the situation/possibility logically *implying* that something is the case. Merging the two cases is not necessarily desirable, for reasons discussed in the *Introduction*: although we can make assumptions about other agents' epistemic states, we cannot assess them with an absolute certainty. As also pointed out by H. Zeevat (p.c.) in Schiffer's 1972] example with a burning candle, which can appear innocent at first, it could happen that one of the participants is on LSD and instead of a candle sees a burning forest. In such a case, one agent would think the mutual context contains a candle on the table, while the other agent would be equally confirmed that there is a forest.

[5] J. Gerbrandy shows that in his formalization, actually all the three definitions of common knowledge come out as equivalent. I will follow others in assuming that the iterative and the reflexive definitions differ for reasons discussed in the previous section.

Furthermore, the fixed–point (but not the shared basis) definition, as formulated above, allows for a case like the following [Clark 1996:96]:

> Suppose [...] that my son and I mutually believe I will be home at six. If I hold this belief because of a note I left him but he didn't read, and he holds it because of a note he left me but I didn't get, we hold our mutual beliefs on different bases, and neither of us is justified in our beliefs.

Finally, note that in order to decide whether something is or is not common ground, the fixed point definition requires the presence of an omniscient being to decide [Clark 1996]. In natural communicative situations, it is more appropriate to refer only to the epistemic states of the communicating agents.

In general, it thus seems that in Gerbrandy's formalization, the fixed–point and the shared–basis definitions come out as equivalent because the formalizations unify a number of properties that are advantageously kept distinct for purposes of describing human communication.

To summarize, the result of the discussion up till now is that there are several reasons to reject the iterative and the reflexive definition of common ground. The processing costs associated with the first, as well as the perspective assumed in the second make them unsuitable for a psychologically realistic model of communication, in which common ground arises and is maintained over an extended period of interaction [Barwise 1985]. Even the shared-basis definition as formulated above, however, does not quite capture all the properties of common ground discussed previously: it is not directly applicable on cases when agents take contradictory information to be common ground, as well as when deceptions and lies are used.[6] In order to accommodate these facts, the definition could be restated in terms individual agents' having access to information about what represents common knowledge among them and other dialogue participants:

[6]Note that under the reflexive definition, conflicting information would not come out as common knowledge. It seems that natural language phenomena such as corrections or presupposition accommodations would thus remain unaccounted for.

(III") **Shared Basis:** For any two agents A and B, A considers ϕ common ground between A and B if for some real world situation σ:

- $Bel_A(Bel_B \sigma)$
- $Bel_A(Bel_B(Bel_A \sigma))$
- $Bel_A(Bel_B(\sigma \rightarrow \phi))$,

where *Bel* stands for "having access to information". While this definition is weaker than the original Lewis's and Barwise's definitions, it can handle the problems pointed out previously. First of all, whether something is considered common ground is not decided by an omniscient third party but by the communicating agent herself. Second, the assumed common ground is based on the shared basis σ. Third, information the agent herself does not believe to be the case – but assumes others believe her to believe that it is the case – can still be treated as common ground. And finally, no infinite processing requirements are necessary.

To illustrate the definition in practice, let us consider again Clark's example of a father and a son who both expect that the father will be at home at six, but they base their expectations on different grounds. With the use of the above definition, we will rightly come to the conclusion that both the father and the son will put ϕ – 'Father is home at six' – into what they think of as their common ground. However, the situation σ is different in the two cases, i.e., ϕ is arrived at using a different assumption about the shared basis. Furthermore, even in a situation where the father knows that his son did not get to read his note, he would still consider ϕ to be in the common ground based on the existence of his son's note.

It should be noted in relation to the initial example of this chapter, i.e., the Coordinating Attack scenario, that the communicating agents can assumme that other agents have access to some information with a various degree of probability. In particular, with each confirmation message providing the shared basis σ, the probability that the generals will be able to coordinate rises. We will come back to the example once we have introduced the probabilistic approach in the chapters to follow.

2.6 Summary

The conclusion that the shared belief definition of common ground is the most appropriate one for natural communication is in accordance with the conclusion of [Clark 1996],[7]. The first author also claims that pieces of common ground can be assigned varying likelihood by the communicating agents. This assumption will become relevant in the fifth chapter of this thesis, in which the actual communication model will be introduced. Before presenting the probabilistic approaches, however, we will show what joint means are used by dialogue participants in order to accommodate new information into the common ground.

[7] As well as [Barwise 1985]

3 Changing the Common Ground

3.1 Introduction

How do mutual beliefs result during a conversation? As discussed in the previous chapter, for a belief to become shared both by the speaker and the hearer, some kind of joint effort is required. Yet in the two most prominent pragmatic descriptions of natural conversations, attention is traditionally given solely to the speaker. Although Austin's original proposal at several places indicates the importance of hearer's uptake, generally in the speech act theory, as well as in context change semantics, the hearer or the interpreter of an utterance is assigned a passive role. However, more than just speaker's effort is necessary for a proposition to become a part of common ground. Corpora studies have shown that speech acts are multi-agent actions [Traum & Allen 1992]: for an utterance to be recognized in a conversation, it is necessary that it first be *grounded* [Clark & Schaefer 1989].

In this chapter, we will first shortly discuss the philosophical background of speech act theory. In the subsequent two sections, we will mention some context–change approaches and show examples of grounding. After that, we will focus on a particular kind of grounding acts, namely acknowledgements.

3.2 Speech Act Theory

Especially in the second part of the 20th century, it has become an established assumption that the meaning of utterances cannot be fully described without attending to their usage [Levinson 1983]. An illustrative example is that of performatives, first introduced by the philosopher J. Austin. Performatives are utterances that *change* the state of affairs, rather than falsely or truthfully expressing something *about* the state of affairs (examples of performatives are acts of promising, betting, greeting, thanking or christening). As discussed by Austin, for a performative act to be successful, a number of **felicity conditions** have to be met. Three general categories of felicity conditions are distinguished:

- (A) There must be a conventional procedure having a conventional effect and the circumstances and persons must be appropriate, as specified in the procedure.

- (B) The procedure must be executed correctly and completely.
- (C) Often, the persons must have the requisite thoughts, feelings and intentions, as specified in the procedure, and if consequent conduct is specified, the relevant parties must do so.

Thus for the act of christening, condition (A) captures the fact that the person performing the christening has to be socially appointed for that role, that it has to happen at a specific location (e.g., a priest may baptize a child in the church or a captain can christian a ship while standing in front of it). Specific formula has to be used, such as *"I name this ship 'Matilda'* and the speaker of the utterance presumably has to have the intention of naming the ship in question (conditions B and C [8]).

The role of the hearer is not explicitly included in the felicity conditions. For some performatives, hearer's *uptake* is required (for example for betting), but in general, whether a performative act is considered successful in the theory depends solely on the speaker and the circumstances of the utterance. Austin discusses briefly the issue of hearer's response [Austin 1976]:

> Naturally sometimes uncertainties about whether anything further is required or not will arise. For example, in formal business acceptance is required, but is this ordinarily so? Similar uncertainty arises if an appointment is made without the consent of the person appointed. The question here is how far can acts be unilateral? Similarly the question arises as to when the act is at an end, what counts as its completion?

The role of the hearer becomes directly relevant in the discussion of *perlocutionary acts*. In saying something, Austin claims, three different types of acts can be performed:

- (i) **locutionary act:** the utterance of a sentence with determinate sense and reference
- (ii) **illocutionary act:** the making of a statement, offer, promise, etc. in uttering a sentence, by virtue of the conventional force associated with it (or with its explicit performative paraphrase) [9]

[8] There most probably exist additional necessary conditions that make the act felicitous.

[9] An explicit paraphrase of a performative would be, e.g., saying *"I hereby promise to do it"* rather than just *"I'll do it"*.

- (iii) **perlocutionary act:** the bringing about of effects on the audience by means of uttering the sentence, such effects being special to the circumstances of utterance.

While the above discussed performatives belong to the category of illocutionary acts, there are a number of utterances that are perlocutionary, in the sense that they require certain effects to be achieved. Typical example is that of imperatives, such as *"Shoot her!"* or *"Take out the garbage!"*.

As Austin points out, it is obvious that the boundary between illocutionary and perlocutionary acts is quite loose once the notion of hearer's uptake is seriously taken into consideration. Nevertheless, in later work within speech act theory, the issue has been neglected [Levinson 1983]. For example, [Searle 1970:25] mentions Austin's notion of perlocutionary acts *in passim*, but otherwise focuses mainly on illocutionary acts. He explicitly argues against merging the categories of illocutionary and perlocutionary acts, as done by P. Grice in his article *'Meaning'* [Grice 1957]. Grice suggests that for a speaker S to mean something by X is to say that S intended the utterance of X to produce some effect in a hearer H by means of the recognition of this intention. However, according to Searle [1970],

> saying something and meaning it is a matter of intending to perform an illocutionary, not necessarily a perlocutionary, act.

Searle's claim is that it is possible to make a statement without intending any effect on the audience, as well as recognize the intention of the speaker without actually satisfying it (e.g., to understand that somebody wants me to believe what he says, while not actually acquiring the belief in question). Understanding, or recognizing the intention of the speaker, according to Searle, does not constitute a perlocutionary effect.

> When I say "Hello" and mean it, I do not necessarily intend to produce or elicit any state or action in my hearer other than the knowledge that he is being greeted. But that knowledge is simply his *understanding* what I said, it is not an additional response or effect.

To summarize, the speech act theory as originally proposed by Austin contains informal apparatus through which hearer's role could be captured in general terms. At least in the elaboration of the theory done by Searle, however, hearer's contribution to the success of a speech act is understood to be quite limited.

3.3 Context Change

In the context change theories of meaning, speech acts are represented as functions from contexts to contexts. While these theories incorporate a great deal of the traditional pragmatics (see [van Rooy & Zeevat, to appear] for a summary), they essentially focus on information exchange among dialogue participants about facts which do not concern the status of what is being said *as such*, i.e., they are not concerned with what could be called 'dialogue about the dialogue'. For example, questions analyzed in context change semantics are of the type *'Are John and Mary sick?'* or *'Who came to the party?'*, rather than *'Did you hear what I just said?'* or *'Do you agree with me?'*. The reason why the dialogue-maintenance utterances are often excluded is presumably the assumption that their semantic and pragmatic effects can be explained by the general analysis. In fact, it is necessary to make some changes in the formalizations in order to account for utterances that concern perlocutionary effects in the dialogue.

3.4 Grounding

In the last decade, the study of human conversations within artificial intelligence research has been approached from a more data–oriented perspective. As a result, a number of preexisting assumptions have been corrected, one of them being the complexity of dialogue structure [Traum 1994], [Clark 1996]. From the point of view of the present discussion, a relevant set of facts concerns the organization of information exchanged in the communicative situations. It turns out that throughout the dialogue, participants constantly make sure that they both understand what is going on on the dialogue's meta-level. Within the study of human-machine interaction for dialogue systems, these actions are as important as the "core" acts (such as giving instructions or requesting information). From a theoretical perspective, the data is of relevance as well, because they can provide insight into the epistemic states of the communicating agents, and their interaction during the exchange.

As reported by [Traum & Heeman 1997], in a corpus of task–oriented spoken dialogue, over 95 % of speaker turns begin with some form of feedback on the previous turn, unless the previous turn consisted of a very simple form of acknowledgement (such as *'ok'*). These results confirm previous findings concerning the structure of dialogues. In particular, it appears that for a proposition to be considered a part of the common ground in the conversation, it first has to go through the process of grounding, i.e., it not only has to be *presented* by the speaker, but also *accepted* by the hearer [Traum & Heeman 1997].

> In the presentation phase, the first speaker specifies the content of his contribution and the partners try to register that content. In the acceptance phase, the contributor and partners try to reach the *grounding criterion*: 'the contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose'.

What are the possible reasons for grounding? Following the work of H. Clark, [Traum & Dillenbourg 1998] list four sources of miscommunication that can arise during communication. The communicative act of the speaker is represented as

$$Done(S, \alpha) \wedge \alpha \rightarrow \mu$$

where S is the speaker, α the uttering act and μ the communicative meaning of α. I.e., we understand α to stand for an uttering act that can have different communicative meanings μ depending on the context of the utterance. In general, miscommunication is defined as

$$Bel(S, \mu) \wedge \neg Bel(H, \mu)$$

Note that this definition is not exact, since the hearer can understand what the speaker said and yet believe the opposite. If it will make their communication smoother, agents can pretend that certain information is a part of their common ground, even though it does not truthfully represent their beliefs. For example, they can decide to refer to a certain individual as *'Roger's wife'* even though they both know that Roger is single and the individual in question is someone else's wife. Similarly, the hearer can hold the believe μ *prior* to the start of the conversation and then during its course misunderstand when the speaker is stating μ. In this case, it is appropriate to talk of miscommunication, yet the condition above is not satisfied. More accurately, we could define miscommunication in terms of the speaker's act as

$$(Done(S, \alpha) \wedge \alpha \rightarrow \mu) \wedge \neg(Bel(H, (Done(S, \alpha) \wedge \alpha \rightarrow \mu))$$

Miscommunication can be caused by any of the following communicative breakdowns:

- (1) α was not accessible to the hearer H (H never perceived any communication to interpret).
- (2) α was accessible to H, but H did not perceive α (and thus formed no beliefs about the communicative intent).
- (3) H perceives α, but does not understand α as meaning μ. If H believes that α had the meaning ν, $Bel(H, \alpha \rightarrow \nu)$, for some $\nu \neq \mu$, then this is the case of true misunderstanding. If there is no ν, then this is the case of non-understanding.
- (4) H believes $\neg\mu$. This is different from (3), since H also believes $Done(S, \alpha) \wedge \alpha \rightarrow \mu$, but disagrees.

The sources of miscommunication correspond to Clark's *joint action ladder* [Clark 1996], which satisfies upward causality. The joint action of conversation can be split into four levels (where the three layers from the bottom correspond to Austin's *phonetic*, *phatic* and *rhetic* acts [Austin 1976] from the point of view of the speaker):

- **Level 1 Execution and attention:** A is proposing joint project w to B, B is considering A's proposal of w.
- **Level 2 Presentation and identification:** A is signaling that p for B, B is recognizing that p from A.
- **Level 3 Signaling and recognition, or meaning and understanding:** A is presenting signal s to B, B is identifying signal s from A.
- **Level 4 Proposal and consideration:** A is executing behavior t for B, B is attending to behavior t from A.

Clark defines the relation between the layers in terms of Upward Completion, a property that implies Downward Evidence.

Upward Completion: In a ladder of actions, it is only possible to complete actions from the bottom level up through any level in the ladder.

Downward Evidence: In a ladder of actions, evidence that one level is complete is also evidence that all levels below it are complete.

If the first step on the action ladder is not satisfied, we can assume that communication as such did not take place. For all the other steps, possible communicative breakdowns can ensue.

In order to reach a state of perceived mutual understanding, the hearer can signal to the speaker that none of the above breakdowns actually happened. This is done by one of the available grounding acts that correspond to the type of miscommunication they are designed to prevent. For example, a hearer can signal that her attention is focused on the speaker by saying *'hmm'*, thus issuing the affirmation that the second step of the ladder was successfully reached. Two steps higher, she can signal her understanding by issuing a relevant response. On speaker's side, level 1 and 2 would represent the locutionary act as such, while the level 3 and 4 would correspond to direct and indirect speech acts, respectively. For example, in a situation where the speaker states

(2) *Would you like to sit down?*

while pointing at a chair, the four levels of actions on the speaker's side would be (see [Clark 1996] for details):

(1) executing the articulation of the sentence and the accompanying movement of her arm,
(2) presenting to the hearer the signal composed of the sentence and the pointing at the chair,
(3) requesting some information from the hearer,
(4) proposing to the hearer to sit down.

On the part of the hearer, her four levels of actions are as follows:

(1) attending to the speaker's articulation of the sentence and the movement of her arm,
(2) identifying the speaker's signal as being composed of certain words and a pointing at a chair,
(3) recognizing speaker's request for information,
(4) considering speaker's proposal to sit down.

In general, grounding acts can be either implicit or explicit; they can be included in communicative acts that present new material, or they can represent grounding acts *sui generis*, such as acknowledgements, repairs and requests for repair.

There exist a number of labelling conventions for dialogue moves. [10] For the present purposes, we can refer to the simple labelling of [Traum & Allen 1992] who propose the following set of grounding acts:

- **Initiate** One of the agents begins on a new conversational unit, the content of which is separate from the previous unit.
- **Continue** Speaker adds related content to the open conversational unit.
- **Acknowledge** Hearer demonstrates or claims understanding of the material presented by the speaker.
- **Repair** Correction of a (potential) misunderstanding of the content of the previous conversation unit.
- **Request Repair** Signal of lack of understanding.
- **Request Acknowledgement** Signal of one agent to the other to acknowledge.
- **Cancel** Abandonment of the current conversation unit, leaving it ungrounded or ungroundable.

As an illustration, consider the following conversation unit, adapted from [Traum & Allen 1992]:

Agent	Utterance	Type of Act
S	`that's four loads of orange juice`	$Initiate_1$
H	`oka.. we..`	Ack_1
S	`of oranges`	$Repair_1$
H	`oranges, yes`	Ack_1
S	`and we have tanker T-two with us, right`	$Initiate_2$
H	`that's right`	Ack_2, $Initiate_3$
S	`okay`	Ack_3

[10] E.g., [Traum & Heeman 1997], the Discourse Resource Initiative, or [Carletta et al. 1997]. See the latter for a summary of older approaches.

3.5 Acknowledging

Although they are subject to cultural differences, and highly dependent on the type of communicative situation acknowledgements (sometimes also referred to as *confirmations* or *acceptances*, viz [Ginzburg 1996]) appear to belong to the most common type of dialogue act. [Traum & Heeman 1997] studied grounding phenomena in a corpus of problem-solving dialogues in which the communicating human agents had no visual contact. In 26 separate dialogues, which ranged in length from 50 to 500 seconds, and were composed of 1366 turn transitions, 51% of the turns started with an explicit acknowledgement.

What are the circumstances under which acknowledgements are given? If the reason hearer acknowledges is her awareness that the environment is noisy, then presumably acknowledgements should always be given. This assumption is made in the proposal of [Clark & Schaefer 1989] who suggest that all speech acts are composed of two phases, the presentation and the acceptance phase:

- **Presentation Phase:** S presents the utterance μ to H on the assumption that H will give positive evidence ϵ that H understands what S means by μ.
- **Acceptance Phase:** H accepts the utterance μ by giving evidence ϵ' that H believes to understand what A means by μ, on the assumption that once S registers the evidence ϵ', she will believe that H understands μ.

Note, however, that if we adopt Clark & Schaefer's proposal, we would be wrongly led to the prediction that every acknowledgement is again acknowledged and so on, *ad infinitum* (every acknowledging act is by its nature also a new contribution). This is obviously not the case, unless we make the default assumption that acknowledgements are *always* given, albeit not explicitly. [11] Then we have merely disguised the original question, though, since now it is not clear why acknowledgements are sometimes explicit and sometimes implicit. Furthermore, as suggested by [Traum 1999], explicitly demonstrating understanding is greater evidence than just acknowledging perception, which

[11] Note, however, that as discussed in the second chapter of this thesis, in a noisy environment, no amount of acknowledging can insure perfect certainty on all sides that the message was accepted.

in turn is stronger than simple continuation by generating a new conversation unit. These facts indicate that there is a decision process involved on the side of the hearer when she selects one or another acknowledging act.

In general, acknowledging acts can be separated into three major categories:

- *explicit* acknowledgements by the use of feedback words [12] (typically words such as *'ok', 'hmm', 'right'*),
- *repetitions* or *paraphrases* of the material presented by the speaker,
- *implicit* acknowledgements, by continued attention, or by generating the next turn.

Examples of the first type are easy to find, but they are not really interesting for the present purposes (though note that depending on the intonation adopted by the speaker, even short feedback words can acquire additional meaning and be interpreted as new dialogue contributions, e.g., as requests for repair or additional information).

An example of acknowledgement by repetition and rephrasing is A's utterance in the dialogue below:

(3) A: *Do you have a stone circle at the bottom?*
B: *No.*
A: *No, you don't.*

Sometimes, only a part of the previous utterance can be acknowledged. For example, it seems possible to acknowledge something as the topic of the discourse unit:

(4) A: *Bill left last night.*
B: *Bill, hmm.*

Implicit acknowledgements can be inferred from the fact that the interpreter generates a new contribution to the dialogue. An answer of any kind to a question posed by the speaker gives evidence that the hearer has heard the question (thus satisfying Level 2 of Clark's action ladder and hence by Downward Evidence also Level 1). More strongly, a *relevant* answer to a

[12] I will disregard non-verbal communicative acts.

question is in the same time an acknowledgement of the fact that the hearer has heard and understood the question, i.e. that Level 3 of the action ladder was satisfied (and, by Downward Evidence, also Level 2 and Level 1).

How can we predict whether in a given communicative situation, an acknowledgement of a certain kind will be given by the hearer? In the past, the useful notion of *grounding criterion* has been employed by [Traum & Dillenbourg 1998], based on the work of H. Clark and associates. The grounding criterion GC of an utterance with the communicative meaning μ, $GC(\mu)$, defines the desirability of a grounding act relative to the task at hand and the communicative setting. For example, information specifically related to the collaborative plans of the agents will have a high grounding criterion, while information that is by its nature transient (it changes before its current value can be of use) will have a low grounding criterion.

[Traum & Dillenbourg 1998] propose to define the utility of a grounding action β for the utterance act with a communicative meaning μ, $U(\beta \rightarrow \mu)$ using the grounding criterion, the groundedness of μ prior to the grounding action β, $G_\beta(\mu)$, the groundedness of μ after the grounding action has been performed, $G(\mu)$, and the cost associated with performing β, $C(\beta)$ as

$$U(\beta \rightarrow \mu) = \frac{GC(\mu) \times (G_\beta(\mu) - G(\mu))}{C(\beta)}$$

As discussed above, the grounding criterion can presumably be established based on the setting of the communicative situation. It is less clear what phenomena affect the value of groundedness prior to the grounding acts and after it has been performed.

3.6 Summary

In this chapter, we have shown that propositions become part of the common ground as a result of speaker's and hearer's joint effort. We have introduced different types of acknowledging acts and related them to Clark's action ladder. Finally, we have raised the question what constitutes the reasoning behind choosing one or another type of acknowledgements and cited the suggestion of D. Traum and P. Dillenbourg. In the fifth chapter, we will argue that the groundedness values are derived based on agents' beliefs about other agents' beliefs relevant for the communicative task. In particular, we will relate the notion of groundedness to that of noise, as proposed by [Gmytrasiewicz & Durfee 2000]. Obviously, a communicative setting in which agents cannot see and hear each other very well or where the communicated message is received with a noticeable delay (see [Clark & Brennan 1991] for details) would lower the level of groundedness prior to the acknowledging act. In most communicative situations, however, noise is rather inherent to the meaning of the utterance.

Before turning to the model itself and the calculation of groundedness, we will discuss previous approaches to relevance of speaking acts in general in the following chapter.

4 Previous Analyses

4.1 Introduction

The notion of relevance has first been discussed in pragmatics by [Grice 1989] in his influential article on rational communication behavior. In this article, Grice postulates general principles of conversation, namely the *Cooperative Principle*, the *Maxim of Quality*, *Maxim of Quantity*, *Maxim of Manner* and *Maxim of Relation*. While the maxims are sometimes considered to be specific sub-principles of the more general Cooperative Principle, this was most likely not the originally intended interpretation. Violation of the Cooperative Principle would, according to Grice, result in a *"succession of disconnected remarks and would not be rational"*. On the other hands, the maxims are often violated in case of conversational implicatures (viz *Introduction*).

- **The Cooperative Principle:** Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purposes or direction of the talk exchange in which you are engaged.
- **Maxim of Quality:** Do not say what you believe to be false or for which you lack adequate evidence.
- **Maxim of Quantity:** Do not make your contribution less or more informative than required.
- **Maxim of Manner:** Be perspicuous (avoid obscurity, ambiguity, be brief and orderly).
- **Maxim of Relation:** Be relevant.

Of the four maxims, the Maxim of Relation is the one that has received most attention in informal approaches to pragmatics, but only recently attempts have been made to express it in a more formal language. While in Grice's original presentation, the maxim is taken to be violable, [Sperber & Wilson 1986] have argued that violation is, in fact, close to impossible. In case a statement is made during the course of the conversation that does not appear to be relevant, the interpreter will rather assume that it *is* relevant, and attempt to derive contextual implications by supplying additional premises. An attested example is cited by [Sperber & Wilson 1986]:

(5) Flag–seller: *Would you like to buy a flag for the Royal National Lifeboat Institution?*
Passer–by: *No thanks, I always spend my holidays with my sister in Birmingham.*

The list of premises necessary to render the response of the passer-by relevant is long. Nevertheless, it is still preferable to supply a long chain of reasoning based on known information, rather than to assume that the passer–by was not cooperative.

(a) Birmingham is inland.
(b) The Royal National Lifeboat Institution is a charity.
(c) Buying a flag is one way of subscribing to a charity.
(d) Someone who spends his holidays inland has no need of the services of the Royal National Lifeboat Institution.
(e) Someone who has no need of the services of a charity cannot be expected to subscribe to that charity.
(f) Hence, the passer–by cannot be expected to subscribe to the Royal National Lifeboat Institution.

Sometimes the chain of inferences is not sufficient in order to see the relevance of a proposition and in such cases, an accommodating context has to be looked for. Consider the examples with temporal anaphora below:

(6) *My car broke down. The sun had set.*
(7) *Max poured himself a cup of coffee. He had entered the room.*
(8) *John took a train from Paris to Istanbul. He likes spinach.*

The sentences above are often cited as examples of incoherent discourses ([Hobbs 1979], [Lascarides & Asher 1994]). The source of incoherence is seen in the fact that the clausal discourse segments in each example cannot be attached to one another with a relevant interpretation. But in a natural setting, regarding a discourse as incoherent/irrelevant is not the most preferred option for the interpreter. Rather than assuming that the speaker is behaving uncooperatively, the hearer looks for a suitable context that would render the examples relevant [Hobbs 1979]:

> Well, maybe the French spinach crop failed and Turkey is the only country...

It seems, though, that considering the Maxim of Relation *a priori* inviolable is too strong. First of all, the interpreter indeed looks for accommodating contexts but accommodation has to be associated with a certain amount of processing costs (see also [Sperber & Wilson 1986], [Pulman 1997]). Second, there are discourses where accommodation hardly seems to be an option. This may be due to the fact that it can be gathered from the simplest consideration of things or because whether or not the earth lay at the center was not what was most important to Copernicus. Notice that there does not appear to be any relevant interpretation of the last two sentences that would make them into a coherent discourse unit – either in the current context or in an accommodated one.

Grice's principles of rational conversation exchange have been formalized in qualitative semantics and we will discuss the approaches of [Zeevat & Scha 1992] and [Groenendijk 1999] in the second section of this chapter. However, already the examples (5)–(8) show that a more fine-grained, quantitative, approach could offer a better approximation to the facts. One type of quantitative approach has been developed in the decision–theoretic framework by [van Rooy 2000a,b,c] and we will briefly summarize his proposal in the third section. In the fourth section, a similar suggestion made by [Gmytrasiewicz & Durfee 1999] will be shown, including the treatment of acknowledgements, as proposed by the authors. We will point out some shortcomings of the models and suggest how these could be solved in the chapter following this one.

4.2 Update Semantics

In update semantics, assertions are interpreted as context updates, where context refers to the participants' common ground. Context is represented as a set of possible worlds. An update results in elimination of those possible worlds from the context set in which a given proposition does not hold. In the tradition established by [Stalnaker 1978], an assertion is considered relevant if and only if it produces an update of the context [Zeevat & Scha 1992], i.e., if it adds new information to a given information state without conflicting with the information that is already there:

- **Consistency:** An assertion ϕ must be consistent, i.e., the information state σ, updated with ϕ, $\sigma[\phi]$ must not equal $\emptyset$.

- **New information:** An assertion ϕ must give new information, i.e., $\sigma[\phi] \neq \sigma$.

Under the original dynamic approach, acknowledging dialogue acts discussed in the previous chapter do not come out as proper context updates. This is because they do not contribute any new information and thus violate the condition above. Consider again the example discussed in chapter 3:

(9) A: *Do you have a stone circle at the bottom?*
B: *No.*
A: *No, you don't.*

If ϕ stands for the proposition 'B does not have a stone circle at the bottom', then the second utterance of A, which is an elided realization of the same proposition, would not come out as proper under the above definitions. An elaboration of the approach has been proposed by H. Zeevat [Zeevat 1997]. We can assume that speaker's utterance does not bring an update with the proposition ϕ expressed by the utterance, but rather with the proposition $B_S(\phi)$ – the speaker believes that ϕ. In order for the proposition to become a part of the common ground, the hearer has to acknowledge it, with the update $B_H(\phi)$. Only by virtue of this acknowledgement, ϕ becomes common ground.

However, now we cannot account for cases when no acknowledgement is given but a proposition is still treated as a part of the common ground (recall the Muddy Children Puzzle from *Introduction*). If we make the claim that acknowledgements in such cases are silent, we are left with the problem from the preceding chapter: why acknowledge explicitly when one can do it with less effort silently.

Another possibility is to analyze acknowledgements as answers to implicit questions in the framework of [Groenendijk 1999]. J. Groenendijk interprets dialogues as a *'game of interrogation'*. All dialogue contributions are subject to the following requirements:

- The interrogator raises issues only by asking non–superfluous questions.

- The witness only makes credible non–redundant statements which address the issues raised by the interrogator.

The requirements above can be straightforwardly linked to Grice's maxims of conversation: the fact that the witness can only make credible statements is related to the Maxim of Quality, that the questions be non–superfluous and the statements non–redundant is the second part of the Maxim of Quantity, and finally the Maxim of Relation is formulated as the condition that the witness must address solely the issues raised by the interrogator. As before, contextual data are possible worlds that are compatible with the data provided by the preceding discourse. Every interrogative introduces a partition of the context and raises an issue – its context change potential is to disconnect certain worlds. Indicatives, on the other hand, provide data – their context change potential is to eliminate possible worlds. They are relevant with respect to a certain question if they are

(i) **consistent**,
(ii) **informative**,
(iii) **licensed by the question**.

A proposition is licensed by a question iff it addresses exclusively the issues raised by the question. Following [van Rooy 2000b] we can define ϕ_Q as the set of cells of the partition Q introduced by the interrogative, that are compatible with the indicative ϕ:

$$\phi_Q \quad = \quad \{q \in Q : q \cap \phi \neq \emptyset\}$$

The indicative ϕ is then licensed iff the condition is satisfied that when it is false in one world of some cell of the partition, it should be false in all other worlds that are in the same cell:

- **(Licensing)** ϕ is licensed by Q iff $\bigcup \phi_Q = \phi$

Obviously, acknowledgements as such do not come out as relevant at the first blush, since they do not resolve any issue raised by the speaker. However, we could take it that the question they answer is *implicit*, the question being whether the hearer has heard and accepted her assertion. As J. Groenendijk suggests,

> the content of any sentence can [...] be taken to consist of a (possibly empty) assertive part, and a (possibly empty) interrogative part.

Like above, the proposition A's second utterance stands for would not be 'B does not have a stone circle at the bottom', but rather something like 'A knows that B does not have a stone circle at the bottom'. Without additional assumptions, however, we would be led to the infinite loop described for the approach of [Clark & Schaeffer 1989]: acknowledgements could again be acknowledged. Further restrictions on the nature of the implicit question acknowledgements answer have to be made also in order to predict different types of acknowledging acts.

Finally, note also that acknowledgements can be made in the form of a question:[13]

(10) A: *Mary is sick.*
B: *Is Mary sick?*
A: *Yes, she ate meat.*

In this example, B's question comes out as impertinent in the *Logic of Interrogation*, since it raises a superfluous issue. In order to account for the exchange in (10), as well as the previous example with a supposed implicit question, we have to account for *why do we ask questions?* The answer provided by [van Rooy 2000a,b,c] is that we do that in order to resolve a decision problem.

4.3 Maximizing Utilities

R. van Rooy proposes an extension of the qualitative approach of [Groenendijk 1999] in a decision–theoretic framework. In decision theory, beliefs and preferences of agents are represented by a *probability function* and a *utility function.*

Typically, a decision situation is represented in a decision matrix where the *probability function* assigns a real number P, such that $0 \leq P \leq 1$, to different events/worlds, the sum of the probabilities adding up to one in case of a finite number of worlds. The *utility function*, in turn, assigns a payoff to each action, given an event/world. The rationally most optimal action an agent wants to perform is an action with the highest *expected value.* The expected utility can be calculated by multiplying the payoff of the action in

[13]Example the courtesy of R. van Rooy.

each world with the probability assigned to that world and summing up the products.

For example, in the simple matrix below, there are two worlds, w_1 and w_2, with the assigned probabilities 3/4 and 1/4, respectively. Conditionalized by two possible actions, they can result in four different kinds of payoffs.

		Action	Action
World	Prob	a_1	a_2
w_1	3/4	4	2
w_2	1/4	3	1

To determine which action to perform, we calculate their expected utilities:

$$EU(a_1) \quad = \quad \sum_w P(w) \times U(w, a_1) = (3/4 \times 4) + (1/4 \times 3) = 15/4$$

$$EU(a_2) \quad = \quad \sum_w P(w) \times U(w, a_2) = (3/4 \times 2) + (1/4 \times 1) = 7/4$$

In general, decision problems are modeled as the triple $\langle P, U, A\rangle$, where P is the probability function, U the utility function and A the set of alternative actions. Choosing an action with the highest expected utility can in turn be represented by a function O from decision problems to sets of actions. In the example above, $O(\langle P, U, A\rangle) = \{a_1\}$.

The utility of an assertion can be calculated in terms of its effect on expected utility of the action with highest expected utility. This can be done by means of conditionalization: accepting an assertion C amounts to conditionalizing of the initial probability function P with C. The expected utility of any action a after conditionalization with C is then

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

The relevance of an assertion can be equated with its utility after the assertion was accepted, i.e., with the difference between performing an action *a'* with the highest utility expected *after* conditionalization of the probabilities and performing the action a with highest utility expected *prior* to the conditionalization:

$$U(C) \quad = \quad max' EU(C, a') - EU(C, a)$$

where

$$a \in O(\langle P, U, A \rangle)$$

In general, an assertion is considered relevant if the set of optimal actions after C was accepted is not the same as the original set of optimal actions (note that the value can never be negative):

- An assertion C is *relevant* iff $O(\langle P, U, A \rangle) \neq O(\langle P_C, U, A \rangle)$

For simplicity, let us assume that in regular dialogues, the decision problems can be restricted to one special kind, namely problems where actions correspond to choosing alternative mutually incompatible hypotheses of a set H. In such cases, decision theoretic utility can be reduced to reduction of uncertainty about which hypothesis is true, which in turn can be reduced to communication theoretic informativity in case H is the finest–grained partition and all elements have equal probability (R. van Rooy, p.c.). The value of an assertion can thus be measured in terms of the reduction of uncertainty about which alternative in H to choose. Following [Bar–Hillel & Carnap 1953], van Rooy defines the informativity of a proposition A, $inf(A)$, as the logarithm with base 2 of $1/P(A)$, the same as the negative logarithm of the probability of A:

$$inf(A) \quad = \quad \log_2(1/P(A)) \quad = \quad -\log_2 P(A)$$

To illustrate the definition at work, consider a situation based on the popular $BBC2$ TV-quiz, "The Weakest Link". In the quiz, there is a team of eight participants who have to answer individual questions. After each round, the team members have to select a person of whom they think he or she is the weakest link in the team, i.e., has gotten the fewest questions right in the past round. The amount of information associated with the decision of the team corresponds to the number of possibilities that it eliminates. According to information theory, a convenient way of counting the possibilities is in terms of binary decisions. In case of eight possible choices, it would take 3 binary decisions to arrive at the final result: we would first select one half of the group of eight, then another half of the group of four and finally one individual out of the remaining two.

When at a later stage of the game the number of participants is reduced to four and again the weakest link has to be selected, the information associated with this decision is 2 – it would take two binary decisions to arrive at the result. The formula that denotes the amount of information $inf(A)$ provided by selecting one of n equally likely possibilities is

$$inf(A) \quad = \quad \log_2 n$$

In our example so far, we assigned equal probability to each possible decision. Before the quiz participants make their decision, however, we could place different bets, or assign different probabilities to one or another candidate being selected. [14] For example, we can assign the probability of 1/2 to Maurice being selected and the probability of 1/14 to any other player being selected. In such a case, the informativity of Maurice being determined as the weakest link is

$$inf(A_M) \quad = \quad \log_2(1/P(A_M)) \quad = \quad \log_2(1/\frac{1}{2}) \quad = \quad 1$$

since the logarithm with base 2 of n is the power to which 2 has to be raised to give n. The informativity of any other team member being selected is much higher

$$inf(A_N) \quad = \quad \log_2(1/\tfrac{1}{14}) \quad \approx 3.8$$

Obviously, the informativity of learning that any other participant was selected over Maurice is higher then in the previous case – it is more *surprising*. For this reason, informativity of a proposition can also be referred to as its *surprisal value*.

Now we are back to the expected utilities discussed previously: we can calculate the average value of learning the answer to who is considered the weakest

[14] Note that the probability cannot be equated with the actual performance of the players. This is because the quiz questions are posed too fast for the participants to estimate the performance of others with absolute certainty.

link in the quiz (where N is the set of propositions corresponding to the set of participants and $EV(A_{WL})$ stands for expected value of learning the answer to the question):

$$EV(A_{WL}) \quad = \quad \sum_{n \in N} P(n) \times inf(n) \quad = \quad (\frac{1}{2} \times 1) + 7(\frac{1}{14} \times 3.8) = 2.4(bits)$$

The average value of information generated at a given source is sometimes also referred to as the *entropy* of the source. Note that when we assumed the possible outcomes of the players' decision, we have partitioned the context in the same way as if we were denoting the meaning of the question *'Who is the weakest link?'* in the sense of [Groenendijk 1999] (see also [Groenendijk & Stokhof 1997]). This is in accordance with the claim of R. van Rooy that decision problems can be thought of as questions and the uncertainty associated with a given problem is the entropy of the corresponding question.

The notion of information value allows us to measure the relevance of answers in a more fine–grained manner. This corresponds to our intuitions that some non–exhaustive answers are more relevant than others. Consider the following example:

(11) A: *Who of John, Mary and Sue are sick?*
B_1: *At least John and Mary are sick.*
B_2: *At least Mary is sick.*
B_3: *Only Mary is sick.*

The last answer is obviously the most relevant one since it eliminates most cells in the partition introduced by the question. Both the answer B_1 and B_2 are less relevant, but of these two, the answer B_2 is understood as the more relevant one. The intuitive ordering relation between the answers can be derived straightforwardly by means of comparing their information value [van Rooy 2000a,b].

In some cases, an answer can be perceived as relevant, even though it does not eliminate any cells from the context partition.

(12) A: *Will John come?*
B: *Most probably not.*

Here, B's answer is informative because it changes the probability distribution of the cells of the partition. To account for this example, we could make use of so–called Jeffrey's conditionalization, which will be discussed in the next chapter.

The fact that interpretation effort, required for processing highly unexpected information, is not made a part of the model can lead to difficulties. Take the following dialogue as an example:

(13) A: *Who can carry the piano?*
B: *Jim, Jack and Fred can carry the piano.*

In the present setting, the question posed by A introduces a partition on the context space. For the sake of simplicity, let us assume that there are only three individuals in the universe of discourse, namely $\{Jim, Jack, Fred\}$ and that the interpreter assigns equal probability to each of the cells in the partition:

Jim	Jim & Jack
Jack	Jim & Fred
Fred	Jack & Fred
Jim, Jack & Fred	nobody

Obviously, B's answer is ambiguous between a collective and a distributive reading. Under the distributive reading, what B is really stating are three propositions, namely 'Jim can carry the piano', 'Jack can carry the piano' and 'Fred can carry the piano', which we will denote P, Q and R, respectively. Under the collective reading, Jim, Jack and Fred carried the piano together, and taking our world knowledge into account, this is the preferred interpretation of the sentence. Note, however, that this reading never arises in the partition above. This is because the cells in the partition are all interpreted *distributively*, and hence the only possible reading of the sentence is the distributive one. There is also no way we could interpret it as referring to the cells containing Jim, Jack and Fred individually, because such an interpretation would be inconsistent under the assumption that each of the three cells is associated with the reading *only Jack.../only Jim.../only Fred.*

One possibility is to interpret the question as introducing a more fine–grained partition in which all possible collective and distributive readings are taken into consideration. With the union of cells associated with the reading *at least* X, rather than *only* X, and a *collective* interpretation, the cell 'Jim, Jack & Fred' would be associated with the collective meaning of the sentence, while the distributive one would be represented by the union of cells 'at least (a group containing) Jack', 'at least Jim' and 'at least Fred'. Since in this case, P, Q and R are independent, they would behave as if conjoined by an informationally additive conjunction, i.e., $inf(A \wedge B) = inf(A) + inf(B)$ [van Rooy 2000a]. However, as long as there is a correlation between the most relevant interpretation of a sentence and the information value of the proposition it expresses, the least probable readings would still come out as the preferred ones. [15]

Note that our calculations were based on the assumption that the question posed by A introduces the finest-grained partition of the context space. But this is not necessarily the case: the asking agent may not want to know *exactly* how the world looks like but rather wants to find out which of her hypotheses about the world are true. If it is not consistent with her beliefs that any of the three individuals could carry the piano alone, her partition would look as below:

Jim & Fred	Jim & Jack
Jack & Fred	Jim, Jack & Fred

In this case, the only possible reading that can be assigned to B's answer is, indeed, the collective one. On this example, we can see the difference between a question introducing a partition and a partition based on the

[15] The same problem reappears in examples of expressions, such as causatives, that usually receive an unmarked interpretation, although a marked one is possible as well. Thus a sentence like *'John stopped the car'* is typically understood as 'John stopped the car in the usual way'. An interpretation where the car was actually stopped in an untypical manner (e.g., by telekinesis) would, however, have a higher informative value (would be more surprising) and is thus predicted by the presently discussed quantitative approach.

decision problem of an agent concerning which of her hypotheses is true. In many cases, these two notions may not coincide.

Going back to our original problem, what is the relevance of information values for acknowledgements? Recall our proposal that acknowledgements be analyzed as answers to an implicit question posed by the speaker. In R. van Rooy's proposal (like in [Groenendijk 1999]), one of the conditions for a question to be considered relevant (with respect to some information state K) is that the partition introduced by the question, Q_K is a non-singleton set. To render the questions for acknowledgements relevant, we will assume that they could be verbalized in terms of Clark's action ladder as follows, corresponding to the three highest levels in the ladder.

$\mathbf{Q_2}$: *Did you hear me talking?* (Level 2)
$\mathbf{Q_3}$: *Did you understand what I said?* (Level 3)
$\mathbf{Q_4}$: *Do you agree with what I said, i.e., can we consider the proposition expressed by my utterance as mutual knowledge?* (Level 4)

Note that the entailment relations between the questions, $Q_4 \models Q_3$, $Q_4 \models Q_2$ and $Q_3 \models Q_2$ correspond to the entailment relation between their answers. Q_4 entails question Q_3 and Q_2, since knowing the true answer to that question means knowing the answer to the other two as well. For positive answers, this fact is defined in Clark's Downward Evidence, and it is also the case for negative answers: disagreeing with what somebody said presupposes perceiving and understanding what was said (at least in the ideal case).

As pointed out by R. van Rooy, the notion of semantic entailment is in accordance with the notion of informativity. Since the informative value of a question can be equated with the average amount of information generated by its complete answers, the question corresponding to Level 4 on the ladder is the most informative one, and hence the most relevant one. We could assume that this is the implicit question to which the hearer answers with an acknowledgement. However, if we take the acknowledging agent to be cooperative, we would predict that she would always acknowledge that Level 4 was satisfied if this were the case. This is not a correct prediction, since it is certainly possible to agree with what someone else is saying without giving one's positive feedback constantly back.

In fact, there are cases when acknowledgements are rather interpreted as referring to lower levels of the ladder, even though interpretation on Level 4 is possible - consider the example below: [16]

(14) A: *I have a headache.*
B: *Yes.*

An alternative solution to why different kinds of acknowledgements are given is to assume that the implicit question does not always have to be the most relevant one – and does not always have to be asked. The acknowledging agent bases her answer on a qualified guess, taking into account her beliefs about the other agent and about the communicative environment. This is, indeed, the perspective assumed by [Gmytrasiewicz & Durfee 2000].

4.4 Reaching the Observation Post

Like R. van Rooy, [Gmytrasiewicz & Durfee 2000] also analyze pragmatic phenomena with the help of decision theory. They assume that communicative acts are primarily based on the 'knowledge base' of an agent, the knowledge base containing information about the agent's preferences, abilities and beliefs about the physical world, as well as beliefs about other agents, their preferences and abilities, their beliefs about the world and about the other agents, their beliefs about others' beliefs and so on. In communicative situations, the knowledge base of an agent is conveniently represented in a pay–off matrix: based on the matrix, speakers evaluate which messages if sent would result in the greatest gain in terms of expected utility associated with the speaker's action. In general, the utility is defined as below:

Utility of Communicative Act: Expected utility of the communicative act M is the difference between the payoff the agent expects before and after executing the act.

When planning a communicative act, the speaker takes into account the reliability of the communication channel. She does so by assigning probability to the other agent properly receiving and understanding the content of her

[16] There is a possible explanation in terms of entailment why Level 4 cannot be the reason for acknowledging in this case. Since it is impossible to *disagree* with someone that s(he) has a headache, hearing and understanding the message presumably entails agreeing with it. Nevertheless, it still has to be accounted for why Level 3 is excluded as a possible interpretation of hearer's response (leaving implicatures aside).

message. If the probability is sufficiently low, it can become disadvantageous to send a message with an initially low utility.

As for acknowledgements themselves, [Gmytrasiewicz & Durfee 2000] define a particular type of communicative acts, namely modeling messages. These are messages that contain some information about the modeling probabilities which represent the speaker's beliefs about other agents in the environment. According to the authors, acknowledging acts are a type of modeling messages: they are sent to confirm the reception of other messages. In a situation where agent R_1 receives a message from R_2, she can send back an acknowledgement with the effect of eliminating the possibility that R_2 does not know whether R_1 has received her message. In principle, acknowledgements of acknowledgements are possible, but as in the case of other communicative acts, they would only be sent if their expected utility is positive:

> If R_1 thinks that R_2 thinks that the original message had a good chance of getting through, then the acknowledgement is not needed; otherwise it is in R_1's best interest to acknowledge.

Note, of course, that since Gmytrasiewicz and Durfee consider communicating agents to be selfish utility maximizers, there has to be a utility assigned by agent R_1 to the fact that agent R_2 *knows* that R_1 has received her message.

A question [Gmytrasiewicz & Durfee 2000] do not answer, however, is how the agent evaluates the reliability of a communication channel. In natural dialogues, what does the notion of noise correspond to?

4.5 Noise

In information theory, noise is associated with disturbances in the channel which in effect create uncertainty about the correspondence between the transmitted message and the received message. In order to illustrate the possible difference, let us imagine a situation in which there is a person watching a TV soap opera and suddenly the electricity is switched off. This happens at a moment when there were three couples, a, b and c, all about to get married. But full of surprises as TV series tend to be, it cannot be predicted *who really got married?* Assigning equal probability to each outcome of the possible eight, the value of the answer A to the question is

3 bits, $inf(A) = 3$. Since the viewer is a devoted fan of the series, he calls up her friend of whom he knows she usually watches as well. It so happens, however, that the viewer's friend only saw a part of the show that night where couple b got married in a ceremony later proclaimed invalid. She reports that at least couple b got married, the actual situation being that it was couple a.

We denote the actual TV situation as s, the situation of the viewer's friend supplying the requested information as r. Following [Dretske 1981], who in turn builds on [Shannon 1948], we will refer to the information available at r independently of s as *noise*: it is a reduction of possibilities at r that is independent of what happened at s. We can measure the dependency between s and r, $inf_s(r)$ and from it the *equivocation*, E, or the amount of information that was lost during the information transfer by the formula

$$E = inf(s) - inf_s(r)$$

We can also calculate the value of noise in the example as

$$noise = inf(r) - inf_s(r)$$

Noise and equivocation will be maximal in cases when there is no relation between the message transmitted by the sender and what the receiver obtains. In case the two situations are identical, i.e., there is a maximum dependence between the event of sending and the event of receiving a message, both equivocation and noise will be zero. In our original example, the $inf_s(r)$ is zero - there is no relation between the information at the source and the information that was available to the recipient. Both equivocation and noise thus have the value 3, since $inf(r)$ and $inf(s)$ is 3 bits for both.

Taking probabilities into account, we can calculate average noise for the event of couple c getting married by the formula

$$N(c) = -\sum P(r_i/c) \times logP(r_i/c)$$

where r_i stands for any cell of the partition in question being reported and $P(r_i/c)$ is the conditionalized probability of r_i being reported given that c was the case.

In the next chapter, we will discuss the unreliability of the communication channel that results from different beliefs of the agents about the world. We will argue that for Level 3 and 4 acknowledgements, this is the source of noise referred to by [Gmytrasiewicz & Durfee 2000].

4.6 Summary

To summarize, in this chapter we have seen a number of pragmatics related data that appear to call for a quantitative analysis, and the quantitative refinement of the original qualitative dynamic frameworks can fruitfully account for many of them. In order to account for circumstances under which acknowledgements are given, Gmytrasiewicz & Durfee refer to the unreliability of the channel. To see what is meant by this claim, we have discussed the notion of *noise* and *equivocation*, as these are understood in communication theory.

5 Communicating with Uncertainty

5.1 Introduction

As pointed out in the first and second chapter, common ground contains information which is assumed to be shared by agents to a varying degree. [Clark 1996] informally introduces probabilities as part of common ground when he states

> People tacitly evaluate shared bases for quality, recognizing that pieces of common ground range in likelihood from 0 to nearly 1.

As for propositions expressed by the agent herself during the course of the conversation, note that these will form a part of the common ground by virtue of the shared basis σ. Recall again the adopted definition of common ground:

(III") **Shared Basis:** For any two agents A and B, A considers ϕ common ground between A and B if for some real world situation σ:

- $Bel_A(Bel_B\sigma)$
- $Bel_A(Bel_B(Bel_A\sigma))$
- $Bel_A(Bel_B(\sigma \to \phi))$,

Performing an act of speaking creates a shared basis σ, from which it can be inferred by the hearer that the speaker attempts for some proposition to become part of the common ground.[17] The speaker, in turn, holds the belief that her conversational commitment was understood as such by the hearer.

The issue that we will discuss in this chapter is under what circumstances propositions become part of common ground and with what degree of certainty. We will analyze the utility of speaking acts and account for when acknowledgements are given by making them dependent on the expected utilities of the speaker, costs associated with speaking and the amount of noise in given communicative situation.

[17]Following the tradition, we can refer to this proposition as the *commitments* of the agent (rather than her belief).

5.2 Modelling Epistemic States

We will make the following assumptions:

Assumption 1: Both speaking and interpreting are communicative acts based on rational decisions of communicative agents.

Assumption 2: Communicative agents make their communicative decisions with the purpose of maximizing the utility of their communicative acts: they compute the expected utility of alternative behaviors and execute the behavior with the highest value.

Assumption 3: Agents desire to have as much information about the actual world as possible, as long as its utility does not exceed the costs of acquiring it.

Furthermore, we take it that agents have beliefs about the "outside" world, beliefs about their preferences and about their possible courses of action. Crucially for interpretation, they also make hypotheses about the belief system of the agents they communicate with. Since the hypothesized belief system of other agents also contains beliefs about the belief system of the original agent, there is a possibility for embeddings of the models. Unlike with the iterated definition of common ground, however, it does not seem unreasonable to assume that this nesting is finite because the amount of information based on which the agent models another agent is finite [Noh & Gmytrasiewicz 1999]. Infinity of the nesting is in no sense a precondition for communication.

The formal framework we will adopt is based on [van Rooy 2000a,b,c, to appear], on the findings of [Shannon 1948], interpreted by [Dretske 1981], and on the standard Bayesian model, as discussed by [Gärdenfors 1988]:

Definition 1 [Models] A model M is a pair $M = \langle \Omega, P \rangle$ where

(i) Ω is a finite non-empty set of possible worlds, and

(ii) P is a probability function defined over the set Ω, such that:

- for any ω, $\omega \in \Omega$, it holds that $O \leq P(\omega) \leq 1$
- $\sum_{\omega \in \Omega} P(\omega) = 1$

A proposition A is represented by a set of worlds and the probability of A is defined as the sum of probabilities in which it holds:

Definition 2 [Propositions] A proposition A is a set of possible worlds, such that $P(A)$, probability of the proposition, is $\sum_{\omega \in A} P(A)$.

It can now be established for any two logically disjoint propositions what their probability will be:

Lemma 1 If A and B are logically disjoint, then $P(A \vee B) = P(A) + P(B)$.

The intuition that there are facts of which the agent has no belief, i.e., propositions stating their truth or falsity are equally likely from the agent's perspective, can be modelled by assuming that the probability of the proposition equals the probability of its negation. For example, for the proposition p 'It is now raining in Montreal', its probability would be equal to the probability of its negation, $\neg p$, i.e. $\sum_{\omega \in p} P(\omega) = \sum_{\omega' \in \neg p} P(\omega')$.

The dynamics of belief change is described in terms of *conditionalization*, where conditionalization of P on A is denoted as P^*_A. Conditionalization is defined as below:

Definition 3 [Conditionalization] Conditionalization of P on A for every proposition B is $P^*_A(B) = P(B/A) = P(A \wedge B)/P(A)$ if $P(A) \neq 0$, otherwise undefined.

We will say that two propositions are independent with respect to the probability function if the following condition holds:

Definition 4 A and B are *independent* with respect to P if $P(A \wedge B) = P(A) \times P(B)$, that is, $P(B/A) = P(B)$, if $P(A) > 0$.

Definition 4 is a special kind of conditionalization which holds for a proposition A if A is fully accepted, i.e., $P'(A) = 1$, where P' is the posterior probability function. A more general kind of conditionalization was defined by R. Jeffrey [Jeffrey 1965]. It represents changes by inputs which could be formulated as 'The probability of A is p', where p is $0 \leq p \leq 1$. It is defined as below, where P is the actual belief state and P' stands for the revised belief state after conditionalization.[18]

Definition 5 For all B, $P'(B) = (p \times P(B/A)) + ((1-p) \times P(B/\neg A))$

As we will see, Jeffrey's conditionalization will be relevant for belief change with propositions the probability of which is lowered by the noise in the communication environment.

5.3 Uncertainty in Communication

Let us first see what the expected utility of a speaking act can be. Consider the following example in which there are four worlds, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and two agents, a_1 and a_2. From the perspective of agent a_1, the four worlds are assigned certain probabilities. Similarly, still from the perspective of agent a_1, the same set of worlds is assigned different probabilities by agent a_2. The table bellow thus represents (i) the beliefs of a_1 about the actual world, and (ii) the beliefs of a_1 about the beliefs of a_2.

	Prob	Prob
World	a_1	a_2
ω_1	1/6	3/5
ω_2	1/2	1/5
ω_3	1/4	3/20
ω_4	1/12	1/20

Take a proposition A, such that $A = \{\omega_1, \omega_2\}$. Then the probability assigned to that proposition by agent a_1 is P_{a_1} and the probability the agent a_1 believes

[18]We disregard causal relations between propositions. As pointed out [Pulman 1997], however, there is a very simple example which shows that causal relations can be assigned probabilities and be subject to revisions. Take K to represent the set of accepted propositions, and A the proposition which is to be processed by the agent, and assume that $K = \{B, B \rightarrow \neg A\}$. Then belief revision can either result in removing B from K or in removing $B \rightarrow \neg A$. Representation of causal relations can be done in a Bayesian or causal network, which consists of a set of propositional variables with conditional (in)dependencies between them.

of agent a_2 to assign to the proposition is P_{a_2}.

$$P_{a_1}(A) = \sum_{\omega \in A} P(\omega) = \frac{1}{6} + \frac{1}{2} = \frac{2}{3}$$

$$P_{a_2}(A) = \sum_{\omega \in A} P(\omega) = \frac{3}{5} + \frac{1}{5} = \frac{4}{5}$$

The amount of information associated with an occurrence of the proposition, i.e., its surprisal value, is for a_1 expressed as $inf_{a_1}(A)$ and, according to the beliefs of agent a_1, for agent a_2 as $inf_{a_2}(A)$.

$$inf_{a_1}(A) = log_2(1/\frac{2}{3}) = log\frac{3}{2} \approx 0.58$$

$$inf_{a_2}(A) = log_2(1/\frac{4}{5}) = log\frac{5}{4} \approx 0.32$$

We can now calculate the dependency between the belief state of agent a_1 and the belief state she ascribes to agent a_2 for proposition A and which we will represent as $inf_d(A)$:

$$inf_d(A) = inf_{a_1}(A) - inf_{a_2}(A)$$

$$inf_d(A) = 0.58 - 0.32 = 0.26$$

The value of $inf_d(A)$ is relevant because it can help us to calculate the noise the agent a_1 will associate with sending the message expressing A to agent a_2. Note that we are not really interested in the *average amount of information* associated with one source or another. This is the difference from using information value of messages/propositions in engineering and using it for the present purposes. For engineering, the calculation of an average information generated by a source is of higher interest than the surprisal value for a particular event. On the other hand, we are rather interested in the average amount of *noise* that is estimated by agent a_1 to influence her communication with agent a_2. Ideally, we would perform the calculation just indicated for proposition A for a representative number of propositions, also evaluating the noise for each case by the formula below (specified for the proposition A):

$$noise(A) \quad = \quad inf_{a_1}(A) \quad - \quad inf_{a_2}(A)$$

With the values of noise for different propositions, we can calculate the value of average noise for x-number of propositions as

$$noise_{average} = (P(A_1) \times noise(A_1)) + ... + (P(A_x) \times noise(A_x))$$

We are also interested in the value of average *equivocation*, the average amount of information lost during the communication. We can compute it in a parallel fashion by summing up the individual values of equivocation of x-number of propositions, weighting them according to their respective probabilities of occurrence.

Why are all the above stated values interesting for an agent that is about to communicate a certain message? From the values of average noise and average equivocation, the speaker can estimate the probable successfulness of the communication. In particular, she can calculate the estimated information value of the uttered proposition in the belief state of the other agent. If due to high equivocation the probability associated with it would be low, the costs of sending the message may be higher than the utility of sending it.

To give an example, take two worlds, ω_1 and ω_2, such that the proposition ϕ 'Agent X knows A' holds in ω_2 but not in ω_1. Let us further assume that the probabilities assigned to the worlds are $P(\omega_1) = 3/4$ and $P(\omega_2) = 1/4$ and that the utilities associated with them are 2 and 8, respectively. Naturally, the utility of a speaking act S will be the difference between the expected utility prior to the speaking act and after it has been performed.

$$U_S \quad = \quad EU_S - EU$$

The expected utility prior to S_a, a speaking act performed by agent a, is calculated straightforwardly as [19]

[19] Note that following [Jeffrey 1965], utility here is defined on worlds, not on world–action pairs as in [van Rooy 2000].

$$EU_{S_a} = \sum_{\omega} P(\omega) \times U(\omega)$$

$$EU = (3/4 \times 2) + (1/4 \times 8) = 7/2$$

What is the expected utility of telling agent b about A? In order to find out, we have to conditionalize the current belief state of the speaker with the probability that agent b will find out about A after the speaking act has been performed. The probability of the proposition ϕ that the agent knows A already is $1/4$, since $P(\phi) = \sum_{\omega \in \phi} P(\omega)$, and let us assume that the probability of the agent finding out about A by virtue of the speaker's act is $1/8$. In our example, this value is fixed arbitrarily. In case the speaker has no beliefs concerning the value, though, it would be based on the average dependency ascribed by the speaker to her belief state and the belief state of the hearer, i.e., on average noise and equivocation associated with the communication.

Recall that for any proposition A

$$inf(A) = log_2(1/P)$$

and the information value the agent b would be assumed to associate with a proposition would be computed by

$$inf_d(X) = inf_a(X) + equivocation_{average}$$

$$inf_b(X) = inf_d(X) + noise_{average}$$

Going back to our original example, by Jeffrey's conditionalization we now get

$$P'(\phi) = 1/8 \times 1 + ((1 - 1/8) \times \frac{1}{4}) = 11/32$$

if we assume that the probability of believing A after finding out that it is the case is one, and the probability of not finding out anything is equal to the prior probability assigned to A. After speaker's act was performed, we thus get $P(\omega_2) = 11/32$ and $P(\omega_1) = 21/32$ and we can calculate the expected utility of such a state of affairs as

$$EU(a_S) \quad = \quad (21/32 \times 2) + (11/32 \times 8) \quad \approx \quad 4$$

The expected utility of the speaker's act is thus

$$U_{S_a} \quad = \quad EU(S_a) - EU \quad \approx \quad 0.5$$

The utility of the act can now be compared to the costs associated with performing the act. Obviously, in case costs exceed the utility, it is not rational for the speaker to say anything.

The rationality behind saying something will in general be calculated as above also for acknowledgements. However, in order to relate the act of acknowledging to the immediately preceding utterance of the other agent, we have to consider also hearer's perspective.

On the side of the hearer, the attempt to maximize her utility in the dialogue situation is in general desire for information (Assumption 3). In cases when she does not ask any question to resolve a particular decision problem, she will interpret propositions with respect to her belief state, i.e., with respect to the set of hypotheses/partitions to which she assigns probabilities. During the interpretation process, she will take into account also her beliefs about the surrounding environment and about the beliefs of the other agent(s). On the basis of these beliefs, she also decides whether or not to supply feedback to speaker's previous utterance.

5.4 Why acknowledge

Based on the Assumption 2, we take it that acknowledgements are only made when the acknowledging act has a positive utility for the acknowledging agent, with the result that the other agent can then be expected to behave in a way that is good for the acknowledger.

As argued in the previous chapter, when an agent decides to acknowledge, she takes into account the *noise* in the communicative situation. Noise can come from different sources and have different effects.[20]

First, it can be due to **external factors** such as the setting of the communication: the need for acknowledging is proven to be higher in conversations that do not take place face–to–face, e.g., on the phone, or take place in environments that contain disturbing elements. The value of noise can be independently established based on the average noise associated with the situation. In the extreme case, if the speaker says something which the hearer cannot hear at all, no information is transmitted and both equivocation and noise are maximal.

The values of average noise and equivocation in this case can be calculated without making reference to the beliefs agents have about the beliefs of other agents and acknowledgements will concern the second level of Clark's action ladder. They can be considered answers to the implicit question *'Did you hear me talking?'* and, in general, the utility of the acknowledgment act can be assumed to correspond to the utility ascribed to the communicative situation as such taking place.

Another type of noise can be associated with the ability of the agents to understand and agree with each other. Here, noise has to be calculated in a more sophisticated manner, since it is based on **internal factors**. In particular, beliefs of agents about the beliefs of others have to be taken into account, in the way indicated in the previous section. Unlike previously, however, the calculation we have made from the perspective of agent a_1 will be embedded in the beliefs of the interpreting agent a_2. In other words, it will be beliefs of a_2 about the beliefs a_1 holds about a_2.

For Level 3, the hearer can base her estimate on what she assumes to be the most likely intended interpretation of the speaker's utterance. Going back to the piano example from chapter 4, trivially, if the hearer believes the speaker believes that the hearer has met Fred, Jim and Jack and is aware of the fact that they are not very strong, she would assign higher probability to

[20]We will disregard situations when an agent acknowledges because she feels invited to do so by the speaker's question intonation, pausing or (non–)verbally indicating a degree of uncertainty concerning the content of her assertion. In these cases, acknowledgements can be explained in the same line as answers to speaker's questions.

the intended meaning being collective, rather than distributive. For general cases, where the hearer cannot make a qualified individual estimate, she could take into account precedent cases of agreement and assumed conformity to a convention [Lewis 1969], to arrive at a certain probability by virtue of the computations described above. In situations where the expected utility of the participants understanding each other is high (for example in case of planning future actions), the hearer may consider acknowledging, e.g., in terms of rephrasing.

Finally, for acknowledgements on Level 4 the situation is similar. Instead of understanding, the hearer now estimates with what probability the speaker assumes her to accept the uttered proposition. Recalling the Coordinating Attack scenario again, we can now predict when one of the coordinating generals will decide not to send an acknowledging messanger back: if after several acknowledgements the probability he assumes the other general assigns to his belief concerning the acknowledged message is sufficiently high. In such a case, the costs of further acknowledging would exceed the utility of the act and acknowledging would not be rational.

Let us now go back to the definition of the utility of a grounding action for some communicative act, as it was proposed by [Traum & Dillenbourg 1998].

$$U(\beta \rightarrow \mu) = \frac{GC(\mu) \times (G_\beta(\mu) - G(\mu))}{C(\beta)}$$

What is the relation between the communication–theoretic notions employed in this chapter and the definition of Traum & Heeman? As a matter of fact, we can now give content in a straightforward manner to the authors' grounding prior and after the grounding action, and to the grounding criterion. The groundedness value of a proposition (i.e., in Traum & Heeman's terminology, the communicative meaning of an utterance act) prior to the grounding action β is simply the probability assumed by an agent a_1 to be assigned by agent a_2 to this proposition. Similarly, the groundedness value of the proposition after the grounding action was performed is the assumed changed probability. The groundedness criterion corresponds to a_1's utility of a_2 believing the proposition, and costs in both cases are the costs associated with speaking.

To illustrate the relation between the two definitions, take a case in which there is only one world ω in which the proposition μ holds (Traum & Heeman do not make use of possible worlds). Then

$$GC(\mu) \times (G_\beta(\mu) - G(\mu)) = GC(\mu) \times (G_\beta(\mu)) - GC(\mu) \times (G(\mu))$$

Since we assume that $(G_\beta(\mu)$ is the same as $P'(\omega)$ (because μ only holds in the world ω) after the speaking/grounding act, $G(\mu)$ is the probability $P(\omega)$ prior to the act, and $GC(\mu)$ is in both cases the utility associated with the world ω, we get

$$U_{S_a} = GC(\mu) \times (G_\beta(\mu) - G(\mu))$$

since

$$U_{S_a} = EU(S_a) - EU = P'(\omega) \times U(\omega) - P(\omega) \times U(\omega)$$

which given the assumptions stated explicitly above means that

$$U_{S_a} = GC(\mu) \times (G_\beta(\mu)) - GC(\mu) \times (G(\mu))$$

assuming that there are no costs associated with S_a. Dividing the value of U_{S_a} by the costs will get us the $U(\beta \to \mu)$, i.e., the utility of a grounding action β compared to associated costs $C(\beta)$.

5.5 Summary

In this chapter, we have shown how agents use their beliefs about other agents' beliefs to estimate the reliability of the communication channel. This can help them to calculate the expected utility of performing a speaking act. In case of acknowledging, the model of speaker's beliefs is embedded into the beliefs of the interpreter. Based on an individual estimate or using the average values of noise and equivocation, she can compute the probability with which it will become assumed by the speaker that she has heard, understood and agreed with the speaker's utterance.

6 Conclusion

6.1 Summary

The results we have arrived at in this thesis can be summarized as follows:

We have argued that the shared basis definition of the common ground is the most appropriate one for purposes of describing natural human communication. Beliefs of agents about what forms a part of their conversational common ground can be assigned varying likelihood. This is influenced by the dialogue contributions of other agents, in particular by hearer's uptake of speaker's utterances.

In this respect, we have focused on grounding dialogue acts, in particular acknowledgements, the relevance of which is difficult to account for in previous formalizations of dialogues. Since the purpose of acknowledgements is to increase the probability of propositions being contained in the common ground, their use can be interpreted in communication theory as an effort to overcome the unreliability of the communication channel. We have analyzed the data from this perspective.

Firstly, we proposed that noise and equivocation, two values that represent the unreliability of the channel, can be caused by both external and internal factors. We attended especially to the latter, more problematic type, and argued that it is caused by agents having different beliefs about the beliefs of other agents. Finally, we showed how the values of average noise and equivocation can be computed based on previous experiences of the communicating participants, and used in their subsequent conversation planning.

There a number of issues related to the present topic that we could only touch upon. They can be seen as possible further extensions of our approach.

6.2 Other Uncertainties

Assume a communicative situation in which the speaker states:

(15) *Elvis Presley is alive.*

There are a number of factors that can influence whether a proposition like (15) will be accepted. Some hearers would be easier persuaded by rational argumentation that can be offered to support the assertion, others by emotional weight of it or by their trust into the agent conveying the utterance. Speakers, in turn, may anticipate these factors and attempt to influence them. At this point, the description of speaker's intentions and the planning of discourse meet.

Like contributions to dialogues, also discourse is rule–governed in the sense that it has to be composed of "appropriate" moves. As indicated in chapter 4, the binding force behind discourse relations (see, [Hobbs 1979], [Mann & Thompson 1988], [Lascarides & Asher 1994], among many others) is also understood in terms of Grice's maxims, the Maxim of Relation in particular. The formalization of the notion for question–answer pairs has been recently extended also on relations among sentences in monologues by [Zeevat, in prep]. Our description of the utility of speaker's act in terms of influencing the belief state of the hearer (given a certain probability that the message will be understood in the intended way) can be understood as another step in the same direction. One extension of the present work would thus be an analysis of discourse relations and discourse planning with speaker's intentions as the starting point [Grosz & Sidner 1986].

Another possible extension is to relate the factors influencing agents' possible conversational moves to game theory. There is a long tradition starting with L. Wittgenstein of analyzing dialogues as games. In a number of linguistic proposals it is employed merely as a useful metaphor (see, a.o., [Lewis 1979], [Kowtko, Isard & Doherty 1992]), but especially in the last years, more formal proposals have been made which relate game–theoretical findings with semantics and pragmatics of natural language ([Dekker & van Rooy 1999], [Groenendijk 1999], [Glazer & Rubinstein 2000]). Based on the present proposal, one could interpret the behavior of communicating agents as a bargaining game, i.e. as a game in which the preferences of players are neither strictly opposed, nor consonant (see [Bartels & Merin 1999] for an implementation in the analysis of English phrasal intonation). There are other potentially useful types of games, such as games of imperfect information and signalling games, which have been proposed for the analysis of linguistic data and could serve for further exploration of the facts.

6.3 Final Remarks

We have attempted to describe a small set of phenomena that are a natural part of human communication. In general, doing that can be compared to an attempt of a rower in Hume's boat, who rows *"in a certain rhythm by a tacit and temporary convention"* [Lewis 1969] and in the same time measures the interval between every stroke, the force exerted to move the oars and the trajectory of their movement. To the extent that making measurements is possible, it becomes a part of the act of moving the boat forward and subject to the same general conditions.

One of the basic assumptions made throughout the thesis is that communicating agents are rational in the sense that they have beliefs about other agents' beliefs and act in accordance with their beliefs. In reality, there is no proof that other agents actually behave in this way and not merely *appear* to behave so. Just like it can be that everyone drives on the right side of the street for no particular reason, it can also be that everyone only appears to communicate rationally. The rules of human communication are a part of the common ground and hence to a certain degree uncertain.

Acknowledgements

I would like to thank to my family and friends. It is not just a matter of politeness for me to say that that without their help, this thesis would never have been written. Many thanks and all my love to my mother and to my sister for their unconditional support, to Øystein without whom I would not have come to Amsterdam, and to Jiřina and Patrick for helping me to get a new home.

ILLC is a great and stimulating environment and when there, I was surrounded by people who were both encouraging and critical. I was lucky to have Robert van Rooy as my supervisor and I am thankful to him for his kindness and patience, for his time and energy, and especially for being an endless source of inspiration and optimism at the frequent occasions when I was lacking both. Many thanks also to Henk Zeevat for his feedback on a substantial part of this thesis.

Finally, my most special thanks to Miguel for technical support.

References

[1] J. Austin, *How To Do Things With Words (The William James Lecture delivered at Harvard University in 1955)*, Oxford University Press, 1976.

[2] Y. Bar–Hillel & R. Carnap, *Semantic Information*, Proceedings of the Symposium on Applications of Communication Theory, Butterworth Scientific Publications, London, 1953.

[3] J. Barwise, *Three Views of Common Knowledge*, ms., CSLI, Stanford University, 1987.

[4] P. Brown & S. Levinson, *Universals in language usage: politeness phenomena*, Questions and Politeness: Strategies in Social Interaction (E. Goody, ed.), Cambridge University Press, 1978, 56-311.

[5] J. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon & A. Anderson, *The Reliability of A Dialogue Structure Coding Scheme*, Computational Linguistics **23:1**, 1997, 13-32.

[6] H. Clark, *Using Language*, Cambridge University Press, 1996.

[7] H. Clark & S. Brennan, *Grounding in Communication*, Perspectives on Socially Shared Cognition (L. Resnick, J. Levine & S. Teasley, eds.), APA, 1991.

[8] H. Clark & C. Marshall, *Definite Reference and Mutual Knowledge*, Elements of Discourse Understanding (A. Joshi, B. Webber & I. Sag, eds.), Cambridge University Press, 1981, 10-63.

[9] H. Clark & E. Schaefer, *Contributing to discourse*, Cognitive Science **13** (1989), 259-294.

[10] P. Dekker & R. van Rooy, *Optimality theory and game theory, some parallels*, Papers on Optimality Theoretic Semantics (H. de Hoop & H. de Swart, eds.), UIL/OTS, Utrecht University, 1999, 22-45.

[11] F.I. Dretske, *Knowledge and the Flow of Information*, MIT Press, Cambridge, MA, 1981.

[12] R. Fagin, J. Halpern, Y. Moses & M. Vardi, *Reasoning about Knowledge*, MIT Press, Cambridge, MA, 1995.

[13] P. Gärdenfors, *Knowledge in Flux*, MIT Press, Cambridge, MA, 1988.

[14] J. Gerbrandy, *Changing the Common Ground*, Proceedings of the Munich Workshop on Formal Semantics and Pragmatics of Dialogue (A. Benz & G. Jäger, eds.), CIS, München, 1997, 40-58.

[15] J. Ginzburg, *Interrogatives: Questions, Facts and Dialogue*, Handbook of Contemporary Semantic Theory (S. Lappin, ed.), 1996, 385-422.

[16] P. Gmytrasiewicz & E. Durfee, *Rational Communication in Multi-agent Environment*, ms., University of Texas and University of Michigan.

[17] H.P. Grice, *Meaning*, Philosophical Review **67**, 1957, 377-388.

[18] J. Groenendijk, *The Logic of Interrogation - Classical version*, The Proceedings of the Ninth Conference on Semantics and Linguistic Theory (T. Matthews & D.L. Strolowitch, eds.), Santa Cruz, CLC Publications, 1999.

[19] J. Groenendijk & M. Stokhof, *Questions*, Handbook of Logic and Language (J. van Benthem & A. ter Meulen, eds.), Elsevier, Amsterdam, 1997, 1055-1124.

[20] B. Grosz & C. Sidner, *Attention, Intention, and the Structure of Discourse*, Computational Linguistics **12:3**, 1986, 175-204.

[21] J.R. Hobbs, *Coherence and Coreference*, Cognitive Science **3**, 1979, 67-90.

[22] J.C. Kowtko, S.D. Isard & G.M. Doherty, *Conversational Games within Dialogue*, HCRC research paper RP-31, 1992.

[23] A. Lascarides & N. Asher, *Temporal Interpretation, Discourse Relations and Common Sense Entailment*, Linguistics and Philosophy **16**, 1994, 437-493.

[24] S. Levinson, *Pragmatics*, Cambridge University Press, 1983.

[25] S. Levinson, *Minimization and Conversational Inference*, The Pragmatic Perspective: Selected Papers from the 1985 International Pragmatics Conference (M. Papi & J. Verscheuren, eds.), John Benjamin, Amsterdam/Philadelphia, 61-129.

[26] D. Lewis, *Convention, A Philosophical Study*, Harvard University Press, Cambridge, MA, 1969.

[27] D. Lewis, *Scorekeeping in a language game*, Journal of Philosophical Logic **8**, 1979, 339-359.

[28] D. Lightfoot, *How to Set Parameters*, MIT Press, Cambridge MA,1991.

[29] W.C. Mann & S. Thompson, *Rhetorical Structure Theory: Toward a functional theory of text organization*, Text **8:3**, 1988, 243-281.

[30] A. Merin & C. Bartels, *Toward a formal semantics of English phrasal intonation*, ms, 1999.

[31] S. Noh & P. Gmytrasiewicz, *Towards flexible multi–agent decision making under time pressure*, Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, 1999, 492-498.

[32] M. Poesio & D. Traum, *Towards an Axiomatization of Dialogue Acts*, Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues (J. Hulstijn & A. Nijholt, eds.), Enschede, Universiteit Twente, Faculteit Informatica, 1998, 207-222.

[33] S.G. Pulman, *Conversational Games, Belief Revision and Bayesian Networks*, ms., SRI Cambridge, 1997.

[34] W.V. Quine, *Quantifiers and Propositional Attitudes*, The Journal of Philosophy **53**, 1956, 177-187.

[35] R. van Rooy (a), *Comparing Questions and Answers: A bit of Logic, a bit of Language, and some bits of Information*, 2000.

[36] R. van Rooy (b), *Strategic Questioning and Answering*, 2000.

[37] R. van Rooy (c), *Decision Problems in Pragmatics*, Proceedings of Gotalog, Fourth Workshop on the Semantics and Pragmatics of Dialogue (M. Poesio & D. Traum, eds.), Gothenburg, 2000.

[38] R. van Rooy (to appear), *Context Change and Attitudes*, Elsevier Publishers.

[39] R. van Rooy & H. Zeevat, *Introduction*, Journal of Semantics, to appear.

[40] O. Sacks, *An Anthropologist on Mars*, Alfred A. Knopf, New York, 1995.

[41] S. Schiffer, *Meaning*, Clarendon Press, Oxford, 1972.

[42] J. Searle, *Speech Acts*, Cambridge University Press, 1970.

[43] C.E. Shannon, *A Mathematical Theory of Communication*, corrected version, originally published in The Bell System Technical Journal, **27**, 1948, 379-423, 623-656.

[44] R. Stalnaker, *Assertion*, Syntax and Semantics 9: Pragmatics (P. Cole, ed.), Academic Press, New York, 1978, 315-332.

[45] R. Stalnaker, *On the Representation of Context*, Journal of Logic, Language and Information, **7:1**, 1998, 3-19.

[46] D. Traum, *A Computational Theory of Grounding in Natural Language Conversations*, PhD thesis, University of Rochester, 1994.

[47] D. Traum, *Computational Models of Grounding in Collaborative Systems*, ms., University of Maryland, 1999.

[48] D. Traum & J. Allen, *A "Speech Acts" Approach to Grounding in Conversation*, Proceedings of the International Conference on Spoken Language Processing (1992), 137-140.

[49] D. Traum & P. Dillenbourg, *Towards a normative model of grounding in collaboration*, Proceedings of the ESSLLI'98 workshop on Mutual Knowledge, Common Ground and Public Information, 1998.

[50] D. Traum & P. Heeman, *Utterance Units and Grounding in Spoken Dialogue*, Dialogue Processing in Spoken Language Systems - ECAI-96 Workshop, Lecture Notes in Artificial Intelligence (E. Maier, M. Mast & S. Luperfoy, eds.), Springer-Verlag, Heidelberg, 1997, 125-140.

[51] H. Zeevat, *The Common Ground as a Dialogue Parameter,* Proceedings of the Munich Workshop on Formal Semantics and Pragmatics of Dialogue (A. Benz & G. Jäger, eds.), CIS München, 1997, 195-214.

[52] H. Zeevat, *Discourse Relations and Tense*, in preparation.

[53] H. Zeevat & R. Scha, *Integrating Pragmatics into Update Semantics*, Communication from an AI perspective (A. Ortony, J. Slack & O. Stock, eds.), Springer-Verlag, 1992, 17-33.