The ABC of Computational Pragmatics

Harry Bunt and Bill Black

1 Introduction: general and computational pragmatics

Pragmatics is the branch of linguistics that studies the relations between linguistic phenomena and aspects of the context of language use. To understand these relations is of crucial importance in many areas of theoretical, computational, and applied linguistics.

In theoretical linguistics, the analysis of such phenomena as anaphora, deixis, and tense requires taking properties into account of the context in which expressions exhibiting these phenomena can be used. Utterances which are context-dependent in such ways are called *indexical*. Bar-Hillel (1954) has argued that indexicality is pervasive in natural language, and speculated that more than 90% of all declarative utterances are indexical.

Indexical expressions encode information about aspects of context: about objects introduced earlier in the discourse, about objects that form part of the physical and perceptual context, or about the (relative) time of speaking. The information partially encoded in the linguistic expression must become fully specified relative to an actual context in order for the expression to have a determinate meaning. The same goes for expressions that carry presuppositions and conversational implicatures, where the relevant context dimension is formed by the speaker's assumptions about the sharing of beliefs and knowledge with the hearer. Expressions in which these phenomena occur can thus be understood only through the relations between linguistic aspects and aspects of context.

The observation that linguistic form may encode context information is in fact not restricted to a small number of aspects of linguistic form, but is a very widespread phenomenon. For instance, the use of definite noun phrases may encode a speaker's assumption about beliefs that are shared with the hearer (presuppositions), and word order and intonation often encode the speaker's intention to structure the information he conveys into a part that is believed to be new for the hearer and a part that is assumed to be known (given-new distinctions), or which part of the information the speaker intends the hearer to focus on (topic-focus distinctions). The linguistic end of pragmatic relations thus involves not only lexical but also syntactic, prosodic, and semantic properties, and in fact it is hard to imagine any linguistic property which may *not* be relevant to take into account.

An important contribution of pragmatics (or of language philosophy, to be more precise, cf. Austin 1962) is the insight that many utterances are best viewed as *actions*. By speaking, we not only convey information, but also *request, assert, order, accuse, deny, threaten, apologise* and so on, thereby indicating the *function* of the information in the context.[1] Languages have distinct sentence structures that can be used to encode some of these functional distinctions explicitly: the *indicative, interrogative* and *imperative* moods, and a speaker can use *speech act verbs* for this purpose within the utterance (e.g. *I'm warning you...; I promise...*). Despite the facilities languages provide for encoding the intended function of the utterance, the use of *indirect* speech acts is very common, where a speaker intends his utterance to have a function beyond what he encodes explicitly. The context is again indispensable to the hearer's ability to interpret the speaker's intended meaning in such cases.

Pragmatics also contributes to the related discipline of sociolinguistics, and its application. Where pragmatics studies how relations between utterances and contexts govern speakers' abilities to interpret and plan utterances, sociolinguistics is concerned with the ways in which social relations constrain communicative situations (and in which patterns of communication construct and maintain social relations, cf. Berger and Luckmann 1966). The pragmatic perspective is essential if sociolinguistic hypotheses are to be adequately grounded in linguistics at an explanatory level. Much sociolinguistic research is empirically-based, and uses rather ad-hoc notations and concepts in the transcription and analysis of linguistic data. A clarification of what constitutes context, and of the relation of utterances to context, supplied by research in pragmatics, may provide much better analytic tools for the study of language in social contexts.

Although 'applied linguistics' is potentially much broader, this field is most heavily engaged with the improvement of methods of teaching and

[1] In a dialogue, the participants perform their communicative actions collaboratively, and may sometimes be considered to perform *joint*, rather than individual actions. See e.g Clark (1996) and Grosz and Kraus (1996).

learning. On one side, the learning of language is an obvious application of linguistics, but on the other, considerable effort is devoted to the study of linguistic interaction as the indispensable medium of learning in all subjects. Nearly all approaches to language instruction place considerable emphasis on the learning of contextually appropriate utterances. Modern learning materials are typically organised not by grammatical phenomena, but via a classification of contexts and types of speech act. Teachers and textbook writers often advertise their methods as a 'communicative approach', in contrast to traditional grammar-based approaches.

Much empirical work on the study of linguistic interaction (e.g. Sinclair and Coulthard 1975) has been conducted in classrooms, and the sociolinguistic study of communicative competence (Hymes 1972) is based on the appropriateness of utterances to the context of their use. These concepts provide much of the underpinning of curricula in native language development. Both activities are applications of pragmatic and sociolinguistic theories much more than they are applications of the syntactic and semantic analysis of language.

Computational pragmatics is concerned with the same relations between utterances and context that are the concern of sociolinguistics and applied linguistics, but from an explicitly computational point of view. This implies in the first place a concern for how to compute the relations between linguistic aspects and context aspects. There are, evidently two sides to this. On the one hand, given a linguistic expression, the question is how to effectively 'decode' those aspects of it that encode context information, i.e. how to compute the relevant properties of the context. This side of the relation is in focus when the meanings of linguistic expressions are to be computed. On the other hand, when we consider language generation, where the task is to construct a linguistic expression that encodes the context information that the speaker (or writer) wants to convey, the question is how to compute the relevant properties of the linguistic expression to be generated given the relevant properties of the context.

Studying the relations between linguistic phenomena and context properties from a computational point of view means that these relations are considered not in an abstract way, but in terms of effectively building explicit representations at one end of the relation from explicit representations at the other end. This raises questions concerning the details of *representations* of the relevant properties at either side of the relation, and these questions are more difficult than it may seem at first.

On the linguistic side, studies in syntactic structure, morphology, phonology, and formal semantics have resulted in a rich variety of representational formalisms that have been proposed to represent linguistic phenomena of all kinds. For many of the context-related linguistic phenomena mentioned above (anaphora, deixis, tense, presuppositions) well-defined representation techniques have been developed, although there is no general agreement on what is the best way to represent the various kinds of linguistic properties that utterances may have, and for some types of properties the search for appropriate representations is currently an active area of research (e.g., semantically relevant prosodic properties, or the representation of ambiguity and vagueness in 'underspecified' logical forms).

The context side of pragmatic relations is much less understood, has not been the subject of a great deal of systematic investigation, and does not enjoy the availability of a wealth of representational formalisms. Investigation of the context side should first of all address the question which kinds of information should be taken into account in context representations. In the above examples of context-dependent linguistic phenomena we have mentioned the following kinds of context information:

- discourse context: objects that have been introduced in the preceding discourse;
- physical and perceptual context: objects that are known to be present or visible (or audible, or..) in the speaker' and hearer's environment; events and actions perceivable in that environment;
- spatiotemporal information: (relative) time and place of speaking;
- attitudinal context: what a speaker believes, intends, knows, fears,...; what he believes a hearer knows, believes, intends, fears,...

The representation of all these kinds of context information in such a way that effective computations can be applied to them and can be used to generate them, is a formidable task, which has so far been accomplished only in part. Moreover, this inventory of aspects of context, relevant for pragmatic relations, is not exhaustive.

General pragmatics being a branch of linguistics, it is no surprise that pragmaticians take an interest not only in relations between linguistic phenomena and context aspects, but also in the linguistic phenomena themselves. Pragmatic studies have in fact looked more at linguistic phenomena related to aspects of language use as such, rather than at systematic properties of the *relations* between these phenomena and context, and they have paid even less attention to the analysis of the context side of these relations *per se*. The analysis of context and the design of context representations has been undertaken only in recent years, primarily in artificial intelligence (see McCarthy 1993; van Deemter and Peters 1996; Iwanska and Zadrozny 1997; McCarthy and Buvač 1994, and the Context'97 and '99 conference proceedings), and not always in ways that are directly useful to the study of language.

In practice, research in pragmatics is concerned with a rather bewildering variety of topics, including implicatures, politeness, coherence, cooperation, presuppositions, salience, interpretation in context, speech acts, common ground, and discourse structure. This diversity has played a part in creating an image of pragmatics as a hardly coherent field of study. Work in computational pragmatics much of the time takes place within the context of projects that aim at the development of language understanding systems or dialogue systems. The design of, for instance, a sophisticated natural language dialogue system forces one to make decisions on how to deal not only with direct questions and answers, but also with indirect speech acts, with implicatures and presuppositions, with salience for contextual interpretation of referential expressions, with cooperativeness in response formulation, and in fact with most if not all of the topics mentioned above. Moreover, dealing with these topics in an operational way requires the application of techniques for contextual knowledge representation, modelling of plans and intentions, and reasoning, thereby effectively connecting the various topics in a coherent computational framework. Work in computational pragmatics is therefore beneficial for general pragmatics both in contributing to the clarification and further developments of theoretical concepts, and for establishing an coherent framework in which these concepts are connected.

Computational pragmatics is a new, emerging subfield of computational linguistics that does not only study linguistic phenomena requiring a contextual explanation, but also investigates the relations between linguistic phenomena and aspects of context with an eye to effective computability, paying equal attention to the analysis and representation of the information at either end of these relations.

2 Approaches and methods in computational pragmatics

2.1 Understanding and inference

Almost from the earliest days of natural language processing, certainly as practised by those who preferred to describe their disciplinary allegiance as artificial intelligence (AI) rather than computational linguistics, the importance of inference in language understanding was acknowledged and operationalised. The main characteristic of the inferences studied has been that they rely on premisses which are not part of the content of the utterance itself. Much of this early AI work, particularly that conducted and inspired by Schank, was concerned with the general background knowledge that speaker and hearer must share so that the one can infer what is implied by the other.

Textbook introductions to pragmatics (Blakemore 1992; Green 1989; Leech 1983; Levinson 1983; Mey 1993) are united in the pre-eminence they give to the view that understanding utterances is achieved by *inference* as much as by shared linguistic knowledge. The source for this perspective on language understanding is usually traced either to Grice (1975) or to Sperber and Wilson (1986). Grice is best known for the notion that successful communication relies on an assumption of cooperation or common purpose, and that this assumption helps to supply specific premises used in the unreliable form of inference he termed 'implicature'. Sperber and Wilson dispute Grice's reliance on cooperation but share with him the emphasis on implicature as a vehicle for the communication of intentions. In this literature, the study of implicature is invoked primarily to explain indirect speech acts; how, for example, a statement can communicate a request (1) or a yes/no question a request for specific information (2).

(1) There's a howling gale in here! = Shut the door/window.

(2) Do you know when the Edinburgh train arrives? = When does the train from Edinburgh arrive?

On the inferential view of communication (contrasted with the *code* or *conduit* model by Sperber and Wilson 1986), these and indeed any utterances or non-linguistic acts function to give evidence of the speaker's intention, and it is primarily intentions that are communicated.

Intentions like the above are not often made explicit in the corresponding utterances. For one thing, being too direct often offends against considerations of politeness or face (Brown and Levinson 1987). Instead, the process of *implicating* one's intentions relies on the hearer being able to add appropriate additional assumptions. For example, (1) should enable most understanders of English to supply additional premises and conclusions such as those in (3).

(3) The speaker has brought a gale to my attention.
The speaker has some attitude to the gale.

Gales, winds, etc. lower the apparent temperature by the wind chill factor.
The gale is making the speaker cold and therefore uncomfortable.
There's an open window.
Open windows admit draughts, winds, gales to rooms.
The open window is probably responsible for the gale the speaker is feeling.
I am nearer the open window than the speaker and could easily close it.
The speaker wants me to close the open window.

A defence against the radical inferential approach to pragmatics may be that utterances like the above are in fact *conventionally* encoded. Whilst a line of reasoning like the above could be invoked as each utterance is being comprehended, it does not need to be on most occasions; conventional indirect forms are remembered as conversational gambits that are directly matched to the situation as needed. However, the context can always change the meaning, or may be needed to disambiguate a form that may conventionally encode several possible intentions. Consider (4), where the second sentence in (a) is a request and the identical sentence in (b) counts as an offer.

(4) a. Where do I sign? Do you have a pen?
b. Please sign here. Do you have a pen?

Grice offered an explanation of how implicatures succeed in conveying intentions through what he called the *cooperative principle* (CP). He illustrated this by reference to example (5),[2] in which A and B are strangers.

(5) A: I'm out of gas.
B: There's a garage around the corner.

One can say that A's utterance conveys that he would like B to assist in procuring gasoline for him by conventional indirectness. However, according to Grice, A can infer under the CP that B believes that the garage around the corner will be open for the sale of gas by the time A will be able to reach it. Had B any reason to believe that the garage was closed, didn't sell gas, etc., he would under Grice's Maxim of Quantity have had to add something like *but I don't think you'll be able to get any there just now,* or not

[2]The lexis in this example has been modified to transatlantic to facilitate the next example.

mentioned the garage at all. The Maxim of Relation (or Relevance Principle in Sperber and Wilson's terminology) thus helps to constrain which out of all the possible premisses the hearer could supply using his general knowledge, (s)he will add as implicated premiss(es) to take together with the explicit utterance content in order to derive an implicated conclusion.

Sperber and Wilson (1986) also point out that logically prior to the process of implicature which Grice described is one of *explicature* or *enrichment* of utterance content. Sense disambiguation is one aspect of this process, and appears to need the same sort of inference as the implicatures needed to infer speaker's intention. Such inferences are just as defeasible as implicatures - consider (6) as a continuation of B's utterance in (5), which could lead A to revise the notion that B meant a gas station.

(6) If you go inside, you'll find a can.

The elaboration of *can* to *can containing gasoline* is also necessary to give (6) an interpretation.

This kind of analysis leads to a questioning (see also Wilensky et al. 1988) of a simple model of utterance interpretation in which first one establishes the 'literal meaning' of an utterance, and then deploys the mechanisms of implicature to uncover the 'non-literal meaning' or the 'interpretation' of an utterance. We will return to this point in section 2.4.

Independently from the philosophy literature, researchers in natural language processing have long been concerned with various aspects of the problem of determining what is meant by what is said. The part of the problem which Sperber and Wilson have classified as explicature has received the most attention, and manifests itself in the natural language processing (NLP) literature as ambiguity resolution. Examples of the kinds of ambiguity that may have to be resolved are grouping or attachment ambiguity (7); noun-compound structure (8) or interpretation (9); sense ambiguity (10); and reference ambiguity (11).

(7) (I (bought ((a flat) in London)))
(I ((bought (a car)) in London))

(8) (Dynamic ((random access) memory))

(9) The Passport Office = the office that issues passports
The Boston Office = the office located in Boston

(10) I keep my lawn mower in the garage.
I'm taking my car to the garage to have it serviced.

(11) If the baby doesn't thrive on cows' milk, boil it.

Examples (7) to (11) all show that the understanding of utterances depends of knowledge of how things are in the real world - flats are immobile whereas cars are mobile; random access is a way of organising the storage of information, whereas to say that memory is for access is uninformative, and so on.

Often, in both the pragmatics and discourse semantics literature and in NLP, utterance interpretation is discussed as though all discourses were monologues. However, spoken utterances most often occur in dialogues. In these situations, utterances not only encode intentions, they also respond to previous utterances and invoke further responses. Words like *yes* have few felicitous uses except in dialogue when following an utterance by an interlocutor, which strongly constrains how it will be interpreted. One-word utterances like this can be seen as one end of a spectrum of explicitness. Elliptic constructions, where one or more elements in a later part of a sentence have been left out and should be inferred from an earlier part of the sentence, are mostly studied in monologue texts. Such constructions abound in dialogues, where the source of the omitted material is often in the preceding utterance from the other participant, as in: *John and Mary* in reply to *Who's coming tonight?* or as in *I believe at four fifteen* in reply to *Do you know when the next train to Tilburg leaves?*

We assume that any NLP system capable of interpreting the intentions behind utterances will need to be able to understand them relative to an evolving context, and that this will involve inference of a non-deductive, indeed an *abductive* nature.

2.2 Abduction

The New Shorter Oxford Dictionary holds that the usage of abduction to denote a syllogistic argument with the minor premiss unspecified goes back three centuries. However, the term seems to have fallen out of use until discovered by Peirce, and to have been further neglected until quite recently, as documented by Neal and others in this volume. As we will see from the discussion in the chapters by Neal and by Oberlander and Lascarides in this volume, there are various ways to define abduction, but the common characteristic is that the inference mechanism is permitted to assume additional premisses in order to reach a conclusion deductively.

Abductive reasoning seems to be something that people very commonly do. When we look out of the window and see that the street is wet, we conclude that it is (or has just been) raining. What we seem to be doing here is applying Modus Ponens in reverse order:

$$(12) \qquad \frac{\begin{array}{l}\text{(1) If it rains, the street is wet}\\ \text{(2) It rains}\end{array}}{\text{(3) The street is wet}}$$

While ordinary deductive reasoning allows us to conclude (3) from the premises (1) and (2), when we reason abductively we observe that (3) is the case and, knowing that (1) is a general truth, we conclude (2). This is not a logically valid conclusion, since it may be the case that the street is wet because a street cleaning car has just passed, or because there is a problem with the sewerage system. What we in fact do in this form of reasoning is use deductive patterns to generate a hypothesis that explains an observed fact (one therefore also speaks of 'hypothetical reasoning'). We commonly do this in language understanding; for instance, conversational implicatures arise when we generate an explanation that enables us to assign a coherent and/or cooperative interpretation to an utterance, as in example (5). Because the abductive generation of explanations is not logically valid, conversational implicatures are defeasible.

Abduction is both intuitively and formally related to *default logic*. Default logic, proposed first by Reiter (1980), is concerned with drawing conclusions from incomplete evidence. More specifically, default logic uses inference rules of the form *If p then q, unless there is evidence that not-q.* Perrault (1989) has proposed to use default logic for specifying the impact of a speech act. For instance, when a speaker uses a declarative sentence with propositional content p, the hearer by default concludes that p. This is achieved by combining two default rules, one saying that when S performs a declarative speech act with propositional content p then (by default) S believes that p, and one saying that if H believes that S believes that p then (by default) H believes that p.

(13) **Declarative rule** $DO(S, act(p)) \Rightarrow B_S p$

(14) **Belief-transfer rule** $B_H B_S p \Rightarrow B_H p$

Beun (1989) has applied default logic for the recognition of speech acts, for instance explaining how a declarative utterance like *You're going home tonight.* may be recognised as a question rather than an assertion, an issue which is discussed further in Beun's chapter in this volume.

Like abductive reasoning, default reasoning produces conclusions that may not be valid. Abduction and default logic are formally related in that abduction can be used as an inference method in simple default theories (those that enjoy the property of semi-monotonicity); see Brewka et al. (1997), ch. 5.

2.2.1 Abduction in NLP

One of the simplest ways to use abduction in natural language processing is in parsing with an incomplete lexicon. Suppose we seek to parse the phrase *the driblet* containing a rare word not in the lexicon. The grammar contains a rule $np \rightarrow det\ n$ and the lexicon has the entry $det \rightarrow the$, but no entry $\_Cat \rightarrow driblet$. Such a situation can be remedied to give a syntactic analysis of `np(det(the),n(driblet))` by a rule that says $n \rightarrow \_Word$ if $\_Word$ is not already listed as a n. Since parsing with rules and a lexicon of ground terms as premisses amounts to constructing a deductive proof of phrasehood, parsing with the same rules and an infinitely extendable lexicon corresponds to constructing an abductive proof of phrasehood.

Hobbs, Stickel, Appelt, and Martin (1993) use (15) as an expository example of several problems that can be solved in the interpretation-as-abduction paradigm, which they characterise as (16).

(15) The Boston office called.

(16) Prove the logical form of the sentence,
together with the constraints that predicates impose
on their arguments,
allowing for coercions,
merging redundancies where possible,
making assumptions where necessary.

Boston office is an example of a noun compound, whose interpretation requires the relation between the second and the first noun to be determined from among *located_in, owned_by, used_for*, etc. The definite reference of *the Boston office* also needs to be resolved, as does the metonymic relation between the office and the human caller. Just to take the noun compound interpretation problem, the syntax/semantics rule accepting a $\_n_1 \frown \_n_2$ sequence assigns a semantic representation of the form $\_n'_1(\_x)$ & $\_n'_2(\_y)$ & $nn(\_y, \_x)$,[3] in this case instantiated as $boston'(\_x)$ & $office'(\_y)$ & $nn_(y, \_x)$. The knowledge base against which this interpretation takes place

[3] In these examples, variables are prefixed by an underscore _. Any symbol (excluding operators) not prefixed by an underscore is either a predicate or an individual constant, according to context. l' denotes the predicate corresponding to the lexeme l, but where a predicate and any corresponding lexeme are spelt differently, there is no prime.

contains among other facts $office'(O), boston'(B)$ and $located\_in(O, B)$. It also contains the generalisation $located\_in(\_x, \_y) \rightarrow nn(\_x, \_y)$ (as well as others with the same right hand side). If one follows (16) taking this knowledge base, a deductive proof is in fact possible without making any additional assumptions. However, this is something of an oversimplification: since the description corresponding to $office'(\_y)$ was syntactically marked as definite, any arbitrary instantiation of y would yield an alternative proof or interpretation, since there may be other offices than O located in Boston, and other known offices with different relations to Boston than location. Hobbs, Stickel, Appelt, and Martin (1993) go on to point out that in their approach of *weighted abduction*, it is relatively costly to assume facts corresponding to the definite NPs in the sentence. In their approach, differential costs propagated through the proof trees determine which of the provable interpretations will be favoured.

An alternative way to implement abduction, not dissimilar in spirit to Hobbs et al's heuristic of making the fewest assumptions, is that used by Guessoum and Gallagher (see their chapter in this volume, and Guessoum and Lloyd 1990, 1991). In their approach, derived from the theory of deductive databases, generalisations about the world are represented as *integrity constraints*. Their paradigm for language processing is to seek to add new propositions derived from the utterance into an integrity-preserving knowledge base management system or belief revision system (cf Gärdenfors 1992) and to make the fewest/cheapest additional assumptions (or retractions) needed to restore integrity. One might, for example, have an integrity constraint $nn(\_x, \_y) \& \neg rel(\_x, \_y) \rightarrow inconsistent$, as well as the two assertions $office'(O)$,$boston'(B)$, $located\_in(x, y) \rightarrow rel(x, y)$ and $located\_in(O, B)$. Asserting $nn(O_1, B)$ into the database results in a need to prove $rel(O_1, B)$ in order to restore integrity. One way to prove this is simply to assume it directly. Alternatively, $rel(O_1, B)$ is provable if we can assume $O = O_1$, and hence prove $located\_in(O_1, B)$. Whichever of $O = O_1$ and $rel(O_1, B)$ is assumable more cheaply will be the preferred way to restore integrity. As with Hobbs et al's approach, it is up to the knowledge engineer to ensure that the costs are assigned to assumable predicates appropriately.

2.3 Reasoning about beliefs

We earlier noted some characteristics of the pragmatic perspective on language use: linguistic behaviour is action; communication is the activity of getting intentions across; speakers give evidence of their intentions, relying on their interlocutors to be able to combine that evidence with background knowledge as well as the specific context of the interaction. Both speakers and hearers are primarily working with beliefs about the other in communicating intentions. Searle claims (Searle 1969) that a speech act of assertion is not felicitous unless the speaker has reason to believe that the hearer does not already believe the proposition to be true; similarly a warning is not necessary if the speaker believes the hearer already knows of the circumstance and the danger that it poses. The indirect speech act exemplified by (2) above apparently requires reasoning by the hearer that the speaker did not only have curiosity about the hearer's mental state, but had some intention that could be served by her/himself knowing the time of the train.

Given the centrality of reasoning about beliefs, a computational pragmatics system will require an inference engine that can reason about beliefs. In the previous section, we discussed abductive reasoning without discussing the specific problems of reasoning about anything beyond classical predicate logic (or even simpler systems like Horn clause logic or Prolog). We cannot reason about beliefs without a more expressive logic than that, because *believes* is not an ordinary predicate, but an operator that induces an 'opaque context'. Whereas in predicate logic we can substitute equivalent expressions and preserve truth, it does not follow from *Believes(Oedipus, married(Oedipus, Iocasta))* and *Iocasta = Oedipus' mother* that *Believes(Oedipus, married(Oedipus, Oedipus' mother))*, because the belief operator is opaque to truth-preserving substitutions. The more expressive logic that we need to reason about beliefs includes axioms and rules of inference in addition to those of classical predicate logic, such as the ones in (17). Knowing something can be distinguished from belief by the axiom of knowledge (17b), which has no counterpart for belief. Modelling a rational agent, we have to decide which additional axioms and inference rules express its powers of inference, and several have been proposed from time to time. Axiom (17c), for example, empowers the agent to apply Modus Ponens within his knowledge space; axiom (17d) makes an agent's knowledge a subsets of his beliefs, and rule (17e) empowers an agent to know everything that can be proved in the logic.

(17) a. $\mathbf{B}_a\phi \rightarrow \mathbf{B}_a\mathbf{B}_a\phi$

b. $\mathbf{K}_a\phi \rightarrow \phi$

c. $\mathbf{K}_a(\phi \rightarrow \psi) \rightarrow (\mathbf{K}_a\phi \rightarrow \mathbf{K}_a\psi)$

d. $\mathbf{K}_a\phi \rightarrow \mathbf{B}_a\phi$

e. $if \vdash \phi\ then\ \vdash \mathbf{K}_a\phi$

Axiom (a), being recursive, generates an infinite set of beliefs. (Moreover, the ϕ can be an expression denoting a belief of someone else, so it is easy to produce expressions like $\mathbf{B}_a\mathbf{B}_b\mathbf{B}_a\mathbf{B}_b\mathbf{B}_a\mathbf{B}_b\mathbf{B}_a\mathbf{B}_b\ldots$). This implies that any mechanism we employ to reason about beliefs will be constrained not to infer all the consequences of a belief, but rather to investigate whether a specific belief is true.

The axiom of distribution (17c), together with (17d) and the axioms of predicate logic with equality, would ensure that if Oedipus beliefs include both that *married(Oedipus,Iocasta)* and that *Iocasta = Oedipus' mother*, he would also believe *married(Oedipus,Oedipus' mother)*.

The axioms of a deductive logic of belief or knowledge do not on their own help much with the reasoning about an interlocutor's belief that we do in linguistic communication, although they can help avoid error that might arise from treating *believes* as an ordinary predicate. We need to be able to make assumptions about the beliefs of others in the absence of any direct evidence. We need, for example, to be able to assume by default that another member of the same linguistic community has access to most of the same general knowledge that we have, but perhaps not all of the specific local, personal, professional and interest-related things that we know. We need to be able to abduce from an utterance when a speaker apparently does not know something and believes that we may know it. Building a model of a rational agent, we also need to have a mechanism for taking on beliefs that we believe others have when we see no reason not to accept them.

A rational agent must reason not only about beliefs, but also about intentions; his own and those of others. Being able to analyse and reason about intentions is crucial both for the understanding of utterances in dialogue (see Allen and Perrault 1980; Grosz and Sidner 1986) and for the generation of actions, including communicative actions. This involves also reasoning about sequences of actions (i.e. *plans*) that will bring about those intentions (as well as any harmful side-effects that some alternative actions or plans might have). This kind of reasoning requires to be supported with an appropriate formal system, such as *situation calculus* (McCarthy and Hayes 1969). Furthermore, in interpreting the behaviour of others, unlike planning our own actions, the operators of the calculus must be used abductively and not deductively: that is to say, utterances give *clues* to intentions, but no more.

Wilks and collaborators have developed a system for the computational modelling of the beliefs and other attitudes of the participants in a dialogue, based on the idea of 'belief spaces'. The set of beliefs of an agent A is partitioned into subsets of beliefs about the world, beliefs about another agent B's beliefs, beliefs about B's goals, and beliefs about B's intentions. The subset of A's beliefs about B's beliefs is further partitioned into A's beliefs about B's beliefs about the world, A's beliefs about B's beliefs about A's beliefs, and so on (see Ballim and Wilks 1991; Wilks, Barnden and Ballim 1991; Lee and Wilks 1996). They claim that their system, with relatively shallow nestings of belief spaces, is a computationally adequate basis for interpreting and generating speech acts. The idea of belief spaces is well known from artificial intelligence (see e.g. Moore 1973; Cohen 1978; Bunt 1994a; Allen 1995; Shapiro 1993) and is intuitively simple and attractive; however, the correct representation of disjunctive beliefs and negative beliefs in terms of belief spaces is technically quite complex (see Moore 1980, Borghuis 1994, and Bunt, this volume). Wilks et al. present their 'ViewGen' system as a computationally-based alternative to logic-based approaches, restricting themselves to relatively simple forms of belief and reasoning about belief.

2.4 Context

One strand of enquiry in computational pragmatics is to seek to apply the principles of rational agenthood to the modelling of a (computer-based) dialogue participant, where a rational communicative agent is endowed not only with certain private knowledge and the logic of belief, but is considered to also assume a great deal of common knowledge/beliefs with an interlocutor, and to be able to update beliefs about the interlocutor's intentions and beliefs as a dialogue progresses.

In dialogues that take place in service encounters, the client has a preconceived idea of the kinds of thing the service agent knows, or s/he wouldn't even begin the dialogue. S/he is also aware that it is the role of the service agent to provide information and possibly to effect transactions on behalf of the enquirer/customer. For his/her part, the service agent will expect the customer to have some intention which can be put into effect by the service (e.g. a plan to travel somewhere). This kind of dialogue takes place, then, in a situation which is very far from a 'null context', in which the knowledge and intentions of the parties are predictable in general terms from their roles, but which each party will try to make specific by their initial moves in the dialogue. Like all beliefs, even the beliefs in this background context may turn out to be wrong (the agent doesn't know about other companies' services, isn't familiar with the geographical situation, etc.), but broadly speaking the knowledge of roles etc. provides the static context in which a dialogue takes place.

A dynamic context is immediately set up by any communicative act, which (normatively and logically) both directs and constrains what either party can do next, and provides a resource of mutually available knowledge that can be exploited in planning and interpreting successive utterances. Suppose an air travel agent opens a dialogue with *Where would you like to go to?* If the customer says the name of a place, e.g. *Chicago*, the agent is entitled to form a belief $\mathbf{B}_{me}\mathbf{W}_{he}flyto(he, here, chicago, \_)$; the context set up by the question enables the agent to interpret the elliptical answer as an assertion that the client wants to go to Chicago. Moreover, the static institutional and role context allows *travel_to* to be specialised to *flyto*, since the agent is there to act for an airline, and by default believes that the client knows this.

The dynamic context contains some kind of trace of previous utterances. This trace is partly at the surface level, since the act of saying something changes the context in that this something has (just) been said now, and dialogue participants can explicitly refer to this, as in *Did you say "Thursday"?* or in *Could you please repeat that?* (Note also that repeating something does not leave the context unchanged, as something has then been said twice.) If this trace is purely at the surface level, however, the utterances in it would have to be (re-)analysed before any reasoning could take place, so it is going to be more efficient to also record semantic/logical representations. The context also contains recollections of the things that have been talked about. For example, if a travel agent has just quoted details of a specific flight, the enquirer can use anaphora in seeking further information, e.g. *Does **it** stop at New York?* since the flight is the most salient thing in the context for both of them. If there were other 'discourse objects' equally accessible, we could bring in shared world knowledge, such as 'stopping is a thing that flights sometimes do' to help choose between them (as would be the case in deciding in this example that 'it' refers to the flight, not to its price). Note that discourse objects do not necessarily correspond to objects in the real world, as the following example illustrates:

(18) C: Could you please book me on the first available KLM flight to Beijing, and tell me how much **it** costs?
I: I'm sorry Sir, there are no KLM flights to Beijing.

The second sentence makes clear that the **it** toward the end of the first sentence does not refer to any real-world object; it refers to the discourse object *the first available KLM flight to Beijing.*

Reference resolution, i.e. finding the intended referent of a referring expression, is one of the aspects of utterance interpretation that crucially relies on an adequate representation of the context and on the ability to reason about the suitability of candidate referents. In face-to-face human dialogue, speakers and hearers not only use the linguistic context, formed by what has been said before, but also the visual and the semantic context, integrating linguistic and nonlinguistic information; see Beun and Cremers (1998).

More generally, the notion of context for language understanding and natural language communication involves a variety of information types, including most importantly the participants' states of knowledge, belief and intention (and possibly other attitudes, depending on the kind of interactive situation); that what has been said or written before; the domain of discourse; the objects that are visible (or touchable, or ...), and the state of processing of each of the interacting partners (are they listening; do they understand; do they agree,...). The chapter by Bunt in this volume presents an analysis of the various kinds of information that form a useful notion of 'context' for the purposes of language understanding and language generation in dialogue.

2.5 Context and communicative action

Noticing that dialogue participants tend to be relatively well-behaved, and to some extent predictable, linguists have been tempted to look at the regular patterns of utterances in dialogues and hence to classify utterances as types of *speech act* or *dialogue act*.[4] We may expect to find that greetings and valedictions tend to get paired up with similar utterances, and that in the body of a dialogue we get lots of initiative–response pairs such as question–answer. It is also noticeable that if an expected response does not occur, a sub-dialogue often appears to take place, and when completed, the original initiative–response pair resumes. Observations like these have suggested to some researchers that context is structured in terms of a 'grammar' of dialogue, which is capable of being modelled on phrase structure grammar or finite state grammar (Sinclair and Coulthard 1975, Roulet 1985). If we base a representation of context on this idea, we could say that if a rule for an initiative-response exchange has been fired top down, and an initiative has just been uttered, the active rule contains a prediction of a response, which can be used in the interpretation of the next utterance when it comes.

[4]The term *dialogue act* goes back originally to Bunt (1979). It has been used in the literature both loosely in the sense of 'speech act, used in a dialogue' (as for instance in Maier 1994 and in Schmitz and Quantz 1994), but also in the more specialised sense of context-changing operation discussed by Bunt (1995) and in chapter 3 in this volume.

One criticism of this approach is that such rules only take utterance functions into account, not utterance content. They say such things as that a THANKING requires an ACKNOWLEDGEMENT to follow, and a QUESTION requires and ANSWER; they are unable to capture semantic requirements such as that an answer should supply information that corresponds to meeting (or at least being relevant for reducing) the information need expressed by the QUESTION. Another criticism is that the rules demanding the occurrence of a certain type of dialogue act sometimes fail. The following dialogue fragment, based on an example from (Luzzati 1995), illustrates this:

(19)

1. U: I am looking for a garage.
2. M: Where?
3. U: Is there a hotel in Martinville?
4. M: Sorry?
5. U: I need to find a hotel where I can wait until my car is repaired.
6. M: No, there is no hotel in Martinville.
7. U: Thank you, goodbye.
8. M: Goodbye.

M's first question in this dialogue never gets an answer, and neither does the indirect request for information expressed by U's first utterance (and U's expression of gratitude in 7 does not get acknowledged either). What happens is that, when M does not immediately get an answer to his question 2, instead of bothering to pursue it he decides that another issue, namely U's question 3 is more important. This being understood by M after utterance 5, and moreover the answer 6 obviating the question 1, neither participant bothers dealing any further with the pending questions 1 and 2. At the end of the dialogue, M's valediction in response to U's valediction also seems to make an acknowledgement of U's expression of thanks unnecessary. This illustrates that dialogue grammar rules, requiring questions to be answered and thanks to be acknowledged, are descriptively inadequate, even when we forget about the requirement of propositional content relations. They are clearly also analytically inadequate, in having no explanatory power whatsoever.

The rigidity of a prescriptive dialogue grammar can be overcome by making such a grammar probabilistic, but this would introduce the fundamentally suspicious idea that the choice of one's communicative actions in a dialogue is made statistically depending on what was said before, rather than determined by intentions and other contextual considerations. The Verbmobil project makes use of a probabilistic grammar of dialogue act types (see Alexandersson, Maier and Reithinger 1995), but merely as a heuristic device to help the automatic recognition of dialogue acts in spoken dialogues for the purpose of the computer translation of dialogue utterances.

Sophisticated dialogue systems may be based on rules that specify the dialogue acts that are appropriate, given the properties of the current context (see e.g. Piwek, 1998). This is where both the practical and the theoretical importance of dialogue acts lie: not in their use in dialogue grammars, but as context-changing operations. On this view, a dialogue act of a certain type (i.e., with a certain communicative function) has a certain type of effect on the dynamic context; this approach is elaborated in Dynamic Interpretation Theory; see Bunt's chapter in this volume, and see the discussion of pragmatics-based language processing systems in section 3.

2.6 Cooperation and rationality

We have noted above that the use of context information in language understanding and generation processes requires the application of abduction-type forms of (hypothetical) reasoning. Among the various kinds of context information that speakers and listeners use, their beliefs about each other's beliefs play a central role, since the effects of communicative actions are always achieved via the changes they may bring about in the listeners' beliefs about the speaker. It therefore seems appropriate to call *abduction, belief* and *context* the *ABC* of computational pragmatics.

These three *ABC* concepts together form a clear, coherent picture: we perform hypothetical reasoning when we combine beliefs and other context information with information extracted from utterances, adding hypotheses where needed to obtain coherent interpretations and consistent new belief states. There is another *C*, however, which we have also seen to be very important in pragmatics: that of *cooperation*. How does cooperation fit into this picture? And we may ask the same question about another concept that is often claimed to be fundamental in pragmatics: *rationality*. Action-based approaches to language, such as Speech Act Theory, Communicative Activity Analysis (see Allwood 1994 and in this volume), and Dynamic Interpretation Theory (see Bunt 1990 and in this volume) assume that communication is a form of rational activity, where the participants use communicative acts to achieve communicative as well as noncommunicative goals in a rational way. The place of rationality and cooperation in the *ABC* picture follows from the *coherence* and *consistency* that the inferential processes aim to establish when they combine utterance information with context information.

In language understanding, the combination of utterance information and context information appears to serve two purposes:

1. to supplement utterance information, which is more often than not insufficient to construct a coherent interpretation of the utterance;
2. to specialise or shift (in the case of indirect speech acts, of irony, of metaphoric use,..) an interpretation of an utterance to the one that the speaker intended.

For instance, in example (4a) above: *Where do I sign? Do you have a pen?* context information is used for the first purpose when interpreting the indexicals *I* and *you*, and for the second purpose when specialising the meaning of *Where* to *At which place in this document* rather than, for instance, to *In which room*, and when interpreting *pen* as writing implement, rather than as enclosure for domestic animals; context information is also used for the second purpose when shifting the interpretation of *Do you have a pen?* from a yes/no-question to a request. Grice (1975) has suggested that the use of context for computing intended interpretations and implicatures relies on first constructing a *literal meaning,* with its ambiguities resolved and its referents specified. The cooperation principle is then applied to this literal meaning in combination with contextual information such as the (mutual) beliefs and intentions of speaker and addressee to compute an *intended meaning.*

Upon closer inspection, the distinction between these two purposes turns out to be problematic, however, as is the very notion of 'literal meaning' (see also Gibbs 1989; Searle 1978). Consider, for instance, the interpretation of an utterance containing an anaphoric pronoun. In order to construct a coherent interpretation for the utterance, an antecedent must be chosen. Being an aspect of referential interpretation, this would appear to be an activity serving the first purpose, and in Grice's terminology to form part of constructing a literal meaning. Yet the choice of an antecedent clearly depends on what antecedent the speaker is believed to have intended. It makes no sense to assign an antecedent to an anaphoric pronoun without taking into account what the speaker is likely to have intended; the whole point of anaphora resolution is to find the *intended* antecedent. For anaphoric expressions, a literal meaning as distinguished from the intended meaning does not seem to make a great deal of sense.

Similarly for ambiguous and vague expressions. To understand the first utterance in *Where do I sign? Do you have a pen?*, the word *Where* must be interpreted. At this point, answers such as *At the bottom of this document* and *In the chancellor's office* are both possible. Again, it would make no sense to first attempt to choose between such interpretations without taking into account what the speaker was likely to mean; the point of disambiguation is to determine the *intended* meaning. (If speaker intention-independent information were available to allow this determination, the expression would not be ambiguous in this respect.) 'Literal meanings' of ambiguous expressions, which would apply out of context, or in a hypothetical 'null context', again do not seem to make sense.

Constructing a coherent interpretation of an utterance thus amounts to using contextual information for obtaining an interpretation that is not only *internally* coherent, in that all the elements of utterance information are semantically linked to concepts in the interpreter's knowledge and with each other, but also *contextually* coherent, in that they are linked with elements in the local context, including the interpreter's beliefs about the speaker's intentions. Rationality and cooperativeness play their part as properties of the way an agent reasons when trying to establish internal and contextual coherence, in the following way.

Both rationality and cooperativeness are thus properties of the way an agent acts. Rationality is an 'inward-looking' property, relating an agent's actions to his goals and intentions in a manner that is both instrumental and efficient: a rational agent chooses his actions in such a way that they lead toward his goals and do so at minimal 'cost'. (Where 'cost' is anything that is unpleasant for the agent to spend, such as effort, or time. For an analysis of rationality see Allwood 1976.) The above example *Do you have a pen?* illustrates this: after *Where do I sign?* the speaker is understood as having the intention to sign a document; the second sentence is interpreted as requesting a writing implement because that would be instrumental to this intention. The inferential processes combining utterance and context information may also be assumed to be rational; in fact, rationality might seem to be an *inherent* property of inference processes, but when one considers the 'inference engines' that have been built in artificial intelligence, it is clear that this assumption is by no means obvious. In fact, one of the main difficulties in building such engines is to develop strategies that guide the engine so as to only make relevant inference steps.

Cooperativeness is 'outward-looking' in that it relates an agent's actions to the assumed goals and intentions of his dialogue partner. This plays an obviously important role in the generation of communicative actions that constitute 'cooperative' behaviour. In order to be able to act cooperatively for a partner B, an agent A must have beliefs about B's goals; such beliefs typically derive from A's interpretation of B's behaviour. In order to correctly induce B's goals and beliefs from his communicative behaviour, this interpretation process must be based on the assumption of B acting rationally and cooperatively.

Together, rationality and cooperativeness thus provide the goals, beliefs and intentions that the inference processes aim to connect to utterances during utterance interpretation. On the generation side, cooperativeness and rationality likewise provide the properties of contexts that communicative actions aim to achieve, and constrain the choice of communicative actions to realise that. Rationality and cooperativeness thus guide a dialogue participant's reasoning and planning processes in interpretation and generation, and in this way fit into the ABC picture outlined above.

2.7 Underspecification and contextual interpretation

A language interpretation or dialogue system works with inputs which are concrete realizations of a sentence or other linguistic construct, with physical properties such as intonation, temporal structure and loudness in the case of a spoken utterance, and with lay-out and punctuation marks in the case of a written utterance. These physical properties contribute to the meaning of an utterance. An utterance, moreover, appears in a certain context: it has a speaker, a time and a place of occurrence in the case of a spoken utterance, and in the case of a textual utterance it has an author, and also a temporal and spatial context of occurrence. These and other contextual properties also contribute to the meaning that an utterance has.

For computing the meaning(s) of an utterance, we thus have three information sources:

1. linguistic information: the semantic information carried by the component words and phrases and by the syntactic composition; anything that goes into the meanings of the sentence that is uttered;

2. physical utterance information: the information carried by the physical properties of the speech signal or the textual form of the utterance;

3. context information: the information that the speaker and hearers have about the domain of discourse; their (knowledge of their) respective goals and purposes; the spatio-temporal properties of the situation in which the utterance occurs, etc.

An issue of central importance in the design of a language understanding system is how these information sources are combined; in particular, how the semantic machinery for compositional computation of sentence meanings may be invoked. For an utterance u of a sentence S, the computation of the sentence meaning of S contributes directly to the computation of the meaning of u only when the processing of the linguistic information of u is a separate stage. Assuming that physical utterance information and context information, being rather different types of information, are taken into account in separate stages as well, the most obvious organization of such a process would be as follows.

(20)
1. Compute the possible meanings of S from the linguistic utterance information.
2. Apply physical utterance information to filter out unintended readings of S and to add pragmatic meaning aspects.
3. Use context information to select the most plausible and relevant readings of S, allowed by and enriched in step 2.

Given the huge numbers of meanings that have to be considered by sentence meaning computation processes, due to the overwhelming ambiguity that sentences display when considered in isolation (see Bunt and Muskens 1999), this seems the most inefficient way possible to organize an utterance meaning computing process. To obtain a more efficient process, one idea is to take nonlinguistic information into account at an earlier stage.

One way to achieve this is to use knowledge of the domain of discourse to restrict the possible meanings of lexical items to those that belong to the domain of discourse. This can be an off-line preprocessing step applied to the dictionary, that can be helpful to reduce ambiguity at the lexical level. For ambiguity at the phrasal level, a possibility that is commonly used in language processing systems is to use domain information to rule out certain combinations of lexical meanings. This can be done by means of semantic checks that are either interleaved with the construction of sentence meanings, or in a filtering step afterwards, and is essentially a way to implement 'selection restrictions'. When semantic checks are interleaved with the linguistic analysis, then formally there is no separate stage of sentence meaning construction, but in reality the sentence meaning construction operations are clearly distinct from the semantic checks, so the sentence meaning construction process can be made to play a significant role. Although the use of domain information may reduce the number of meanings that are constructed, in general these approaches still result in the generation of large sets of meanings, most of which have to be discarded through more sophisticated ways of taking context information into account, since finding the intended meaning of an utterance is more a matter of taking into account what is contextually coherent, plausible and relevant than of deciding what is semantically possible.

To resolve this situation, two approaches seem possible:

- Apply context information interleaved with linguistic information and block the generation of sentence meanings at an early stage.
- Do not generate a large number of detailed alternative sentence meanings, which exhaust all the meanings that an utterance of the sentence could possibly have, but use the linguistic utterance information to generate the *constraints* on the meanings that an utterance of the sentence may have.

The first of these approaches is extremely difficult in practice, since the application of context information involves complex reasoning with pieces of information which during the linguistic information processing are only partly available. This poses problems both for the reasoning involved, and for the organization of interleaved linguistic information processing and reasoning. The second approach has been developed in computational semantics under the name of *underspecified semantic representation*, and may be considered as a breaktrough in the direction of effective utterance meaning computation.

While traditional semantic representations are expressions in a formal language that represent fully disambiguated sentence meanings, an underspecified semantic representation is a representation of the semantic information expressed in the sentence, in such a way that anything which cannot be resolved by the linguistic information alone, is left unspecified. Underspecified semantic representations allow for a *monotone process* of utterance meaning computation, i.e., instead of the generate-and-test approach corresponding to the various variants of the (2.7) organization, context information can be used to *elaborate* underspecified representations. Bringing context information to bear on the interpretation process now means that some or all of the points where a USR is underspecificied are resolved.

The idea of underspecified semantic representations is often ascribed to quasi-logical forms developed by Alshawi in the Core Language Engine project (Alshawi 1992). However, the idea has a longer history. In the PHLIQA project during the seventies the idea was developed of building semantic representations that leave all lexical ambiguity and vagueness unresolved the idea being that precise lexical interpretation is only possible on the basis of contextual information (see Medema et al. 1975; Bronnenberg et al. 1979; Bunt 1981). Alshawi's quasi-logical forms have been well publicized and have been very influential. Recently other, more sophisticated underspecification techniques have been defined in Discourse Representation Theory (Reyle 1983), in the TENDUM dialogue project (Bunt et al. 1984); in the multimodal DENK project (Kievit 1998), and in the Verbmobil translation project (Bos et al. 1994; Pinkal 1999).

Underspecified semantic representations form a powerful concept that is promising for developing effective processes of computing contextual utterance meanings, that construct a bridge between computational pragmatics and computational semantics. In practice, the development of this concept, and the investigation of the questions concerning its theoretical status, its precise definition, and how it can be used in processing, is mostly considered to belong to the field of computational semantics rather than to that of computational pragmatics.

3 Pragmatics in language understanding systems

Since the 1960s, attempts have been made to build systems with the ability to understand natural language to a nontrivial extent. The language processing in first-generation language understanding systems (as well as in more recent systems that expressly aim at a very restricted form of understanding) draws primarily on lexical and syntactic resources and is aimed at the recognition of terms and patterns that have a special significance for the application at hand.[5] Semantics in these systems is mostly restricted to the specification of the application-specific significance of words and phrases, or to incorporating domain-specific categories as elements in a 'semantic grammar.'[6] Pragmatics, in the sense of taking context into account in a principled way, plays virtually no role at all in these systems.

[5]Restricted forms of language understanding can be useful for a variety of computer applications, such as special-purpose user interfaces (data base querying, information retrieval, voice command systems,..), linguistically-supported text processing, or translation aids.

[6]Semantic grammars have been developed in the PLANES system (Waltz, Finin, Green, Conrad, Goodman and Hadden 1978) and in the LIFER/LADDER system (Hendrix, Sacerdoti, Sagalowicz and Slocum 1978). For other early systems aiming at restricted but potentially useful forms of language understanding see e.g. Harris (1978) and Thompson and Thompson (1978).

3.1 First-generation systems

Attempts to build full-fledged language understanding systems started with Winograd's SHRDLU system (Winograd 1972) and Woods' LUNAR system (Woods et al. 1972). These pioneering systems were based on the use of combinations of syntactic and semantic rules, allowing to parse a sentence into its components and to compute the semantic consequences of the syntactic structure. The LUNAR system did not incorporate any pragmatics, aiming only at answering isolated data base queries; the SHRDLU system did incorporate some dialogue facilities and context-dependent interpretation and generation functions, but only in a fragmented and *ad hoc* way. Continued research on computer understanding of natural language has since the late seventies resulted in a number of experimental systems with increasing capabilities, though until recently taking context information into account only in the following limited ways:

- using knowledge of the application domain in the form of domain-specific lexica and/or syntactic categories;
- using knowledge of the tasks of the application in a crude determination of the communicative functions of utterances (e.g. whether an utterance is a question or a confirmation);
- using dynamic information about the history of the dialogue and about the state of the current task for keeping track of discourse referents, for anaphora resolution, and for planning the next system action.

Systems that use contextual information only in these limited ways are bound to be very restricted in their understanding capabilities, as is obvious from the pervasive influence of dynamic as well as static context on the interpretation and generation of utterances in concrete communicative situations as discussed above. One of the ways in which these limitations show up is that natural language systems are notoriously fragile, failing when confronted with imperfect input, or input which has not been anticipated by the system designer. While natural language is an extremely powerful means of communication between partners that share a rich background of contextual knowledge, natural language utterances are hopelessly ambiguous, vague, and informationally incomplete for an interpreter without such a background. For an interpreter with poor context information, the interpretation of utterances is bound to break down very easily.

A system that does not break down in unexpected or problematic situations, but that behaves appropriately in all conditions, is called *robust*.

In order to increase the robustness of his system, the designer often adds special rules to the grammar or relaxes some of the constraints in grammar rules and lexical items to cover cases that were unanticipated before (allowing for instance certain frequently occurring typing errors or grammatical sloppiness). In addition, techniques are often applied that may be classified as 'low-level': automatic spelling correction, on-line lexical acquisition, semantic grammars, and so on. These low-level techniques miss the heart of the problem, which is to react appropriately in a context created or updated by the fact that a user has said or typed something at a keyboard with the express purpose of communicating a message that is meaningful in the current context. A pragmatics-based approach, by contrast, takes as its point of departure that the user has the intention of communicating a meaningful message, and aims at reconstructing this message by considering the signals that the user sends within the current context.

In the rest of this section we describe a number of experimental language understanding systems that have marked advances in the use of contextual information and pragmatic knowledge in natural language processing. We will see that, in the course of time, an increasingly prominent place is given to contextual information as a knowledge source in interpretation and generation, and to pragmatic concepts and principles in dialogue management.

3.2 PHLIQA, TENDUM, SPICOS, and CLE

In the PHLIQA question-answering system (Bronnenberg et al. 1979), a level of analysis was distinguished with semantic representations in a formal language where the terms correspond to syntactically differentiated word senses, and another level with terms denoting concepts in a discourse domain. The terms at the first level preserve the ambiguity and vagueness of the corresponding words. The analysis at this 'context-independent' level represents only the semantic consequences of syntactic structure and of the use of function words. At the second level, ambiguity and vagueness are resolved relative to the granularity of the domain model. The first level can be said to be 'context-independent' and the second to be 'context-dependent', if 'context' is taken in the sense of domain of discourse.

The PHLIQA system pioneered the use underspecified sematic representations, using so-called *metavariables*. Metavariables are a kind of variable terms in representations that can be instantiated in different ways by unambiguous subexpressions, and can be considered as formal placeholders for semantic choices that still have to be made. This technique was developed for dealing with lexical ambiguity and vagueness.

The TENDUM dialogue system (Bunt et al. 1984) went a step further in recognising that the resolution of *structural* ambiguity in general also requires domain knowledge. This applies for instance to deciding whether a quantification is to be taken collectively or distributively, and to whether a noun is to be taken as a count noun (like *rope* in *a short rope*) or as a mass noun (as in *a metre of rope*). Representations were designed that leave quantifier distributivity and mass/count distinctions underspecified, by using metavariables to represent these choices.

The SPICOS system, a joint Philips-Siemens research experiment and a precursor of the Verbmobil project (see below), re-implemented many of the TENDUM techniques for semantic analysis in the context of a document retrieval application with a German speech front-end (see van Deemter et al. 1985; Thurmair 1989), thereby showing the generality of these techniques.

Unlike the systems mentioned above, the Core Language Engine (CLE; Alshawi 1992) was not aimed at performing a specific interactive task like data base question answering, or document retrieval, but was intended as a general-purpose natural language front-end customisable to a variety of applications. The CLE provides a translation from English sentences to underspecified representations of their propositional content into *'Quasi-Logical Forms'* (QLFS). These structures are intended to describe context-independent meaning aspects; the design of QLFS was focused specifically on the representation of unscoped quantifiers. The QLF notation (Alshawi 1992) was indeed *quasi*-logical in the sense that the notation, though looking more or less like a formal language, had no semantic definition, unlike the formal representations used in the PHLIQA and TENDUM systems (see Landsbergen and Scha 1979; Bunt 1985). A definition of the semantics of QLFs has been given later in terms of their possible unambiguous instantiations (Alshawi and Crouch 1992).

3.3 TRAINS and Verbmobil

The TENDUM system was also innovative in attempting to deal not only with truth-conditional aspects of meaning but also with 'pragmatic' meaning aspects by treating dialogue utterances as communicative acts, which have *communicative functions* that capture the speaker's intentions and associated beliefs and other attitudes. The analysis of user utterances thus includes identifying goals and beliefs of various sorts that are attributed to the user, and which together constitute a dynamic *user model*. Assigning an interpretation to a user utterance in this system depends not only on the global context formed by the discourse domain, but also on the more 'local' context formed by the current user model. Consider, for instance, the following fragment taken from a recorded telephone dialogue:

(21)
1. U: What are the daily departure times of flights to Munich?
2. S: That's 7.45, 8.30, 9.50, 14.30, 18.25, and 20.30.
3. U: On Sunday too.
4. S: On Sunday too.

Utterance 4 in this fragment is a near-perfect copy of utterance 3, also in its intonation; yet it clearly has a different function. We understand utterance 3 as a verification and 4 as a confirmation because of the goals and beliefs we attribute to the participants at the relevant points in the dialogue.

For a dialogue system to play the role of S in this fragment, the understanding of the user's utterance 3 thus requires the system to consult the user model built up at that point. In TENDUM, the user model containing the user's beliefs and intentions is the backbone of the *local* context model, i.e. of the model of that contextual information that changes through the dialogue (as opposed to the *global* context, the contextual information that remains unaffected by the dialogue). Local and global context information are used in the system both to interpret the user's utterances (and are crucial in cases like utterance 3 above and for the recognition of indirect speech acts) and to generate the system's next dialogue contribution.

Other dialogue systems, such as SUNDIAL (Peckham 1993; Eckert and McGlashan 1993) and TRAINS (Allen et al. 1995) have also used combinations of local and global context representations, focusing mostly on the representations of current beliefs and intentions, and using speech act concepts both to interpret user utterances and to generate system contributions. The TRAINS system, for example, is an intelligent planning assistant that helps the user by interacting in natural language to construct plans for using a railroad freight system. It takes a speech act theoretical approach to dialogue and manages a dialogue by taking into account various types of local and global context information. Similar to the systems mentioned above, a user utterance in TRAINS is first assigned a context-independent, underspecified meaning representation, called an 'Unscoped Logical Form'. In order to operate with these USRs, the system uses the following context information: (1) beliefs shared by the user and the system; (2) communicative intentions that the system has, in the form of intended speech acts plus the reason for performing the act; (3) a segmentation of the dialogue as a stack of 'discourse units'; (4) knowledge about plans for using the freight system; (5) a set of 'conversational obligations' (for instance, a question generates an obligation to answer).

The system uses this context information firstly to interpret underspecified logical forms, e.g. making use of the (in-)accessibility of discourse units to interpret anaphora, and secondly for managing the dialogue, using e.g. the shared beliefs and the current conversational obligations to detect the communicative function of a utterance and to compute the system's next contribution to the dialogue (see further Traum 1994, chapter 6). The representation of shared beliefs in the TRAINS system and the way this representation is constructed, represents some of the most advanced computational work on *grounding*, i.e. the establishment of a common ground between dialogue participants (see Clark and Schaefer 1989, Traum and Hinkelman 1992, and Traum 1994).

The German mega-project Verbmobil has as its long-term goal the development of a prototype system for the translation of spoken dialogue utterances to support two persons with different native languages (Japanese and German) who conduct a business conversation in English, but who may also use their native language and have their utterances translated on line, on demand (see Wahlster, 1993, for more details about the objectives of the project). The system, of which a first prototype was completed in 1997 that supports a dialogue with the goal of finding a date for a business meeting, therefore has two processing modes:

- *Deep processing* when one of the dialogue participants requests a translation. In this mode, the input goes through phases of speech recognition and syntactic and semantic analysis for the source language; dialogue processing; transfer; and generation and speech synthesis for the target language.

- *Shallow processing* when the dialogue participants interact without requesting a translation. In order to follow the dialogue superficially, a key word spotter examines the input for cue words which are characteristic for certain dialogue steps.

In the 'deep processing' mode, the Verbmobil system interprets the input as a speech act (or more accurately, as a 'dialogue act' in the sense of Dynamic Interpretation Theory; see Jekat et al. 1995 and Bunt, this volume) and uses contextual information both for utterance interpretation and for dialogue management. In the final Verbmobil prototype, planned for the year 2000, deep, shallow and statistical processing modes are integrated in a sophisticated fashion and can support each other.

A key issue in the Verbmobil project is to develop robust processing methods that can cope with unreliable and incomplete input as is typical for spoken language systems. One important means to achieve this is the availability of contextual information. This information is used e.g. to disambiguate translational equivalents during transfer, to resolve anaphoric expressions, to control lexical variation in the generation of target expressions, and to recognise indirect speech acts.

Local context information in the Verbmobil system has two main parts, chronologically and conceptually ordered, respectively:

1. a 'sequence memory' reflecting the sequential order in which the utterances and the related dialogue acts have occurred in the dialogue;
2. a 'thematic structure', representing the propositional contents that have been communicated and their status in the dialogue, i.e. who they have been proposed by and whether they have already been accepted or rejected.

The two components are closely intertwined to permit easy access of all information for every utterance of the dialogue (see further Maier 1995; 1996).

Different from dialogue systems, dialogue management in Verbmobil is used not for generating system contributions, since the system does not participate in the dialogue as such, but in order to predict follow-up speech acts in dialogue processing. The dialogue management component has a hybrid architecture containing both statistical and rule-based methods, both using dialogue acts as basic units of processing (see Alexandersson et al. 1995; 1998; Reithinger and Maier 1995).

As already mentioned, the Verbmobil project has also developed techniques for underspecified semantic representations, motivated primarily by the observation that ambiguities in the source language often do not have to be resolved (should not even be resolved, sometimes), in order to be translated adequately into the target language (see Bos et al. 1994; Copestake et al. 1996).

3.4 PLUS and DenK

The Esprit project PLUS ('Pragmatics-based Language Understanding System') started from the point of view that the robustness of language understanding systems could benefit greatly from taking contextual information into account in order to deal with linguistic imperfections, unanticipated constructions, unknown or misspelled words, etc. This view was incorporated in a design where a logical form language was defined for underspecified representation (ULF, Geurts and Rentier 1993) and an HPSG grammar was developed producing such forms as intermediate semantic representations (Verlinden, 1998). Abduction was used to bring contextual information to bear in the interpretation of ULFs in order to produce representations that are free of ambiguity or vagueness, relative to the domain of discourse (Kievit, 1998). This design has been developed and implemented in a prototype system.

Several components of the PLUS system have been reused in the DenK system (Bunt et al. 1998), which was built with the aim to construct a generic cooperative multimodal user interface, combining natural language with other modalities, in particular with direct manipulation of objects in a visual representation of the domain of discourse. The DenK system uses a refined version of the ULF language (Kievit 1996) for the intermediate semantic representation of English dialogue utterances, computed by an HPSG parser, and uses a context-driven interpretation process resulting in disambiguated representations in the logically powerful formalism of *type theory*. Type theory has been originally developed for foundational research in mathematics, and has also been applied in the theory of programming languages and in the semantics of natural language (see De Bruijn 1980; Barendregt 1991; Ahn and Kolb 1990; Ranta 1994). Research in the DenK project has shown the great potential of type-theoretical representations as formalisations of context information, provided that the representation structures and proof methods of standard type theory are extended in order to deal with the beliefs and intentions of agents participating in a dialogue (see Borghuis 1994; Ahn and Borghuis 1998). Type theory is used not only for the disambiguated semantic representation of the utterances exchanged by the user and the system, but also for representing the system's knowledge of the discourse domain, and for representing the beliefs that the system believes to be shared with the user.

The DenK system incorporates the ideas of Dynamic Interpretation Theory, using dialogue acts, defined as context-changing operations, as the functional units in terms of which pragmatic principles in dialogue behaviour are defined, the implementation of which forms the heart of the system's dialogue management component (Piwek 1998).

The PLUS project has been exceptional in not just accepting that a language understanding system should take the pragmatic aspects of language understanding seriously, and make use of contextual information in principled ways, but going a step further and taking the stand that language understanding should be *driven by pragmatics*. By this we mean that syntactic, semantic and pragmatic analysis cooperate not in the traditional way, where pragmatic analysis is applied as a filter on what syntactic and semantic analysis have produced, but in such a way that pragmatics actively influences how syntactic and semantic analysis should operate and what kind of information they should deliver.

A pragmatics-based language understanding system is robust for the same fundamental reason why human interlocutors are 'robust' in the sense that they have little difficulty in understanding each other's utterances also when these are incomplete, ambiguous, grammatically imperfect, or corrupted by noise. Human interlocutors have this ability not because they employ special rules for ungrammatical sentences, or spelling correction algorithms, but because they apply pragmatic knowledge to supplement or correct information coming from syntactic and semantic processing. Context information allows one to add missing elements, to infer the intended interpretation of an ambiguous utterance, and to check, clarify or correct results from syntactic-semantic processing when they don't fit in the current context. The principal characteristic of a pragmatics-based system would be robustness and flexibility in a wide range of situations: the system should be capable of handling unexpected input (such as extra-grammatical sentences, 'elliptical' fragments, misspellings, unknown proper names and other unknown words,..) but also of dealing with more complex issues such as deriving conversational implicatures, determination of relevance, deriving speaker beliefs, and other context-related issues needed to allow a real dialogue with the user.

The keynote of the PLUS project was to achieve robustness and flexibility in human-computer dialogue by treating dialogue as a communicative activity whose essential characteristic is to convey a meaning that is both appropriate and relevant contextually. In order to demonstrate the usefulness of this approach, a prototype interactive Yellow Pages information service was implemented.

4 About this book

The chapters in this book all relate to some aspects of the role of abduction, belief, and context in computational pragmatic work, and have been clustered in three groups, in an attempt to do justice to the particular way in which they deal with these concepts.

Allwood presents an approach to communication and pragmatics called

Communicative Activity Analysis that he has been developing since the mid 70s. After a critical review of speech act theory, of conversational analysis, of relevance theory, and of the work of Grice and Clark, his chapter describes the main concepts and features of his approach. Communicative Activity Analysis views communication as action, which is seen as constituted by a combination of behavioral form, intention, context, and result. Every utterance in a conversation is assumed to have a functional structure with three components: (1) functions obligated by the preceding discourse; (2) functions obligating for the succeeding discourse; (3) 'optional' functions, which are neither obligated not obligating. Obligations are analysed as deriving either from general rational and ethical requirements on communication, from the requirements to manage the interaction, or from interaction between the goals of communicative acts and the embedding activity context. Concerning the latter source of obligations, it is claimed that communicative intentionality has an *expressive* aspect, which is to express a certain attitude, and an *evocative* aspect, which is to evoke a certain reaction from the addressee. The various kinds of obligations created by communicative acts and by the embedding activity context, in particular the pairing of obligating and obligated aspects of communicative acts, are assumed to be responsible for the dependencies and regularities that may be observed in dialogues. Communicative Activity Analysis offers a rich conceptual framework for the analysis of human dialogue, and has been influential in the PLUS dialogue project.

Bunt's chapter is concerned with the conceptual and computational modelling of local dialogue contexts, i.e., of the contextual information that is created and changed through communication. To this end he applies his *Dynamic Interpretation Theory* (DIT), which views utterances as intended to change the local context in certain ways. A particular way to change the context, using the propositional information of the utterance, is called a 'communicative function'; the combination of such a function and a propositional content is called a *dialogue act.* For the major categories of communicative functions distinguished in DIT, a conceptual analysis is given of the kind of information addressed by a dialogue act and the intended local context change. This leads to a conceptual structure of local dialogue context consisting of six information types. The logical properties of these information types are studied, with particular attention to their logical complexity, their use in inferential processing, their time dependence, their depth of recursion, and mutual dependencies between them. This gives rise to the suggestion that a computational model of local dialogue context is best organised into four components, one of which acts as a dialogue memory. Of the other three components, a central role is played by a dialogue agent's beliefs and intentions regarding the task that motivates the dialogue and regarding the dialogue partner. This information is highly recursive and logically complex. The chapter is concluded by considering two nonstandard formalisms which seem particularly interesting for representing this kind of information: constructive type theory and modular partial models. DIT has played a major role in the dialogue modelling and context representation in the PLUS dialogue project, and has been the basis of the TENDUM and DenK systems.

Sabah's chapter examines four fundamentally different approaches to the study of language and communication for their consequences for the architecture of language processing systems, concentrating mainly on how and when to use contextual and other pragmatic knowledge for automatic text understanding. He argues that the control mechanisms traditionally used in language understanding systems are inadequate, and that new mechanisms need to be developed that allow flexibility and reflectivity (i.e., the system can reason about its own control structure). He argues in favour of the use of *multi-expert systems* for maintaining modularity without introducing artificial ambiguities and proposes additional concepts for making such systems reflective. These considerations have led to the design of the CARAMEL-1 language processing system. To control all the processes participating in a large modular architecture in a way that allows reflectivity, a new memory model is introduced called the *Sketchboard.* This model is applied to various context-dependent perception and interpretation tasks. Sabah discusses the embedding of heterogeneous modules in a plausible cognitive architecture and how this forms the basis of a new system, CARAMEL-2.

Taylor and Waugh's chapter presents a framework for communication analysis called the *Layered Protocol* theory, that has been developed by Taylor and co-workers over the years. LP theory takes its inspiration partly from the modelling of communication in computer networks, where the use of layered protocols has become standard. These standards operate according to a 'coding-decoding' model of communication, assuming that the coding transformations applied by one site are inverted at the receiving end. This model is too simple for human communication, as e.g. Sabah argues in his chapter, but LP theory applies its basic ingredients in a model of human communication, assuming that acting and interpreting in the real world is done in a series of layers of abstraction, where each level performs coding and decoding operations in a way that may be described by a protocol. A message may be communicated at any level of abstraction as a 'virtual' message, its reality being implemented at a lower, supporting level, in which the messages are characteristically different in kind. The way a dialogue participant may make progress toward a communicative goal is structured in the form of three basic statements relating to the state of the recipient's understanding of the message. Their beliefs about the basic statements, and about each other's beliefs, determine what each does, which allows a 'General Protocol Grammar' to be defined for a dialogue, which applies at every layer of the LP model. These ideas are illustrated by analysing two sample dialogues from Wizard-of-Oz simulations of the interaction between an information service and a client.

Redeker in her chapter argues that textual coherence and conversational coherence are not as incommensurable as much of the traditional research on those discourse types might suggest. On the basis of current developments in discourse theory and extensive corpus-analytic studies of monologue and dialogue discourse, she develops a framework in which coherence is viewed as consisting of three parallel components: ideational (semantic) structure, rhetorical structure, and sequential (or segment) structure. It is claimed that the ideational, rhetorical, and sequential structures are in principle isomorphic and can for descriptive purposes be conflated into one hierarchical structure of the discourse at hand. The framework, called the *Parallel-Components Model*, is intended to accommodate both monologue and dialogue structures, although admittedly a lot of work still needs to be done in order to provide a satisfactory account of, for instance, sequential relations in dialogue. The model allows predictions about the use of discourse operators, and predicts complementarity in the lexical marking of ideational and pragmatic links, which is shown to hold in spoken narrative and in newspaper discourse.

Carter's chapter argues for the importance of the notion of focus in resolving referential and other ambiguities in discourse. He discusses the distinction between global and local focus and the three related problems of how to structure context representations in which focus plays a dominant part, how to determine the referents of referring expressions as they are encountered, and how the context representation should be updated in response to processed referring expressions. A comparative review is presented of several implemented mechanisms for focus tracking and anaphor resolution, and these are evaluated for their performance and theoretical insight.

Ramsay's chapter reviews the use of speech act theory in AI natural language systems and argues that attempts to embody speech acts as simple STRIPS-like planning operators must fail. The difficulties for such an implementation of speech acts arise partly from the fact that there is no reliable way of associating speech act types and linguistic surface forms, and partly from properties of epistemic logic. For instance, an interrogative sentence like *Do you know what time it is?* can be used as a request to tell the time or as a reproach for being late. Since the sentence itself is ambiguous in this respect, in order to identify which action has been performed speaker and hearer must already have the additional information (like: hearer is late; both speaker and hearer know this, etc.) available which would be conveyed by saying what action is performed. Nothing can therefore be gained by positing an extra layer of analysis where utterances are characterised as particular types of speech acts. This point is in essence the same as Bunt's argument, made in chapter 3, to the effect that indirect speech acts do not fit into a context-change theory of meaning; it only makes sense to use speech act types that correspond to observable utterance features.

Beun's work emerges from his empirical analysis of a corpus of spoken information dialogues. The chapter in this volume is concerned with the utterances expressed in the declarative form which function as questions, and how they are reliably interpreted correctly by dialogue participants, whether or not they are marked by intonation as questions. The corpus study revealed a heavy preponderance of declarative forms used when the purpose of the question was verification of something already stated or implicated by the hearer. This was verified experimentally in experiments where subjects were asked to predict continuations of selected dialogues and to express their own confidence in the propositional content conveyed in the earlier part of a dialogue. This work makes a valuable contribution to the empirical study and the understanding of how the choice of particular indirect speech acts is determined by the context.

Thijsse's chapter deals with one important aspect of the precise and formal description of the knowledge used and conveyed in conversations, namely the epistemic force of declarative utterances: What is known by the speaker when h/she is uttering an assertion? More precisely, when speaker S utters a sentence with propositional content p, meant to inform a hearer that p, what does this tell us about S's state of information? It seems impossible to say *p, but I don't believe that p* (Moore's paradox), which suggests that when uttering p, S should be attributed the belief that p. But this does not account for the fact that it seems equally strange to say *p, but I don't know whether p* (Hintikka's example), which suggests that S should be attributed the *knowledge* that p. On the basis of a detailed analysis of such puzzles, Thijsse proposes to attribute to S the *belief that he knows that p*. He shows that this accounts for both Moore's and Hintikka's paradox, and provides a logical reconstruction of Grice's quality maxims. This reconstruction presupposes a distinction between pragmatics and semantics, for which it is argued that independent motivation can be found.

Meyer's chapter discusses the merits of a conceptual modelling approach to the representation of the static and dynamic knowledge that a pragmatics-based language understanding system relies on and updates. CML (Conceptual Modelling Language) uses the basic constructs of objects, labelled attributes and inheritance associated with the frame approach to knowledge representation, now more familiar as the object-oriented paradigm. Many examples are given in illustration of the use of CML-like languages, and Meyer argues in support of claims that this kind of language is suited to the representation of the attitudes of agents; that it is as expressive as more standard logical notations; that its inherent modularity makes models encoded in it extensible and maintainable; that making situations into objects provides a way of representing evolving restricted contexts in which only the most recent and salient discourse acts and events are preferentially accessible to determine interpretation and reaction.

Neal's chapter is a brief introduction to and discussion of the logical nature of abductive reasoning. He compares and contrast the accounts given by Peirce, credited with introducing the term, Charniak and McDermott, who use it extensively in their discussion of how expert systems may be constructed that help find causal explanations of observable symptoms, and finally by Hobbs et al's 'interpretation as abduction' account of natural language utterance interpretation. He finds that they all mis-describe the logical structure of abductive reasoning and proposes that Kenney's schemata for practical reasoning is a better reference point than these authors, as an explanation of the reasoning patterns they actually employ.

Oberlander and Lascarides discuss the interpretation and planning of utterances with implied discourse relations between them, for example: *Max fell. John pushed him.* In their terms, this discourse is *laconic* because whilst there is no discourse relational connective like *because*, the hearer can reliably infer such a relation between the two eventualities. Amongst other things, discourse understanding includes the enrichment of the literal meaning of the two sentences with a discourse relation that renders the utterance reliably coherent. Planning discourses involves being able to predict when this kind of inexplicitness will work. In their chapter, Oberlander and Lascarides claim that defeasible deduction is more suitable to interpretation whilst abduction has a role in planning. Their defeasible deduction reasoning schema has a rule (the *'penguin principle'*) that prefers the defeasible implication that has the more specific antecedent in case of a conflict. The *explanation* discourse relation being more specific in this way than *narra-*

tion, and applicable here because the two eventualities have a common key event, it is inferred in the above example. In planning the same text, the explanation and key event rules are applied abductively to build up the discourse structure and to select the linguistic expressions that can realise it. Proposed realisations are then subjected to a test for the reliability with which the discourse structure will be inferrable by the hearer.

Guessoum and Gallagher, having elsewhere proposed consistent knowledge base update as a mechanism for abduction, and also considered its application to anaphor resolution (Guessoum, Black, Gallagher and Wachtel 1993) here make the interesting proposal that the same paradigm can be applied to the problem of strategic control in a dialogue system. Checking the integrity of a database receiving updates depends on knowledge of the inconsistency of certain combinations of assertions. Other chapter authors (esp. Allwood and Bunt) have investigated the way in which in a dialogue an initiating act may create a strong normative pressure for a certain type of response. Gallagher and Guessoum model this by treating the violation of such pressure to respond as an anomaly. If the initiating act is inserted into a KB in which there is an anomaly constraint against an initiative of this type not being met with an appropriate response, the mechanism (intended for logical constraint application) will return a transaction to produce the response as a way of restoring equilibrium.

Hinkelman and Spackman's chapter, finally, presents several aspects of speech act recognition as carried out by an implemented natural language understanding system. The first part of this process relies on an HPSG grammar that correlates speech acts with partial linguistic descriptions. They argue that certain configurations of feature structures (with disjunction and negation) are expressively equivalent to implications in predicate logic, and depending on the way in which they are initially instantiated, the unification operation has the effect of deductive or abductive inference. The next part of the chapter describes a simplified technical solution to the representation of mutual beliefs and intentions, using 'corporate agents'. Individuals as well as groups can form 'corporations', any set of agents with a common purpose, including perhaps even strangers who meet together for a dialogue. Dialogue contribution types (as distinct from speech acts) are definable in terms of the membership of the corporation and of other contributions. For example, the agreement of the corporation z to the proposition p is definable as the conjunction of the inform contribution by a sub-agent x to z or p and the acknowledgement of p by z and the membership of x in z. Speech act recognition and planning can use these rules reversibly, either abductively or inductively. The final part of the chapter discusses an implemented system which conducts dialogues about appointments, in which these ideas are applied.

References

Ahn, R. and T. Borghuis (1998). Communication modelling and context-dependent interpretation: an integrated approach. In *Proceedings of MUNDIAL'98, Formal Semantics and Pragmatics of Dialogue*.

Ahn, R. and H. Kolb (1990). Discourse representation meets constructive mathematics. In L. Kálmán and l. Pólos (eds.), *Papers from the Second Symposium on Logic and Language.* Budapest: Akademiai Kiadó, 105–124.

Alexandersson, J., E. Maier and N. Reithinger, J. (1995). A robust and efficient three-layered dialogue component for a speech-to-speech translation system. *Proceedings of the 7th. Conference of the European Chapter of the ACL, Dublin* 188–193.

Alexandersson, J., B. Buschbeck, T. Fujinami, M. Kipp, S. Koch, E. Maier, N. Reithinger, B. Schmitz and M. Siegel (1998). Dialogue acts in VERBMOBIL-2, Second Edition. Verbmobil Report No. 226, Saarbrücken: DFKI.

Allen, J., L. Schubert, G. Ferguson, P. Heeman, C. H. Hwang, T. Kato, M. Light, N. Martin, B. Miller, M. Poesio and D. Traum (1995). The TRAINS project: a case study in building a conversational planning agent. *Journal of Experimental and Theoretical Artificial Intelligence 7,* 7–48.

Allen, J. (1995). Natural Language Understanding. Redwood, CA: Benjamin/Cummings.

Allen, J. and C.R. Perrault (1978). Analyzing intention in utterances. *Artificial Intelligence 15,* 143–178.

Allwood, J. (1976). *Linguistic communication as action and cooperation.* Gothenburg Monographs in Linguistics 2. Göteborg University, Department of Linguistics.

Allwood, J. (1994). Obligations and options in dialogue. *THINK Quarterly 3*(1), 9–18.

Allwood, J. (1999). An Activity based Approach to Pragmatics. *This volume, 49–80.*

Alshawi, H. (1992). *The Core Language Engine*. Boston: MIT Press.

Alshawi, H. and D. Crouch (1992). Monotonic Semantic Interpretation. *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, 32–39.

Austin, J. (1962). *How to Do Things with Words.* London: Oxford University Press.

Ballim, A. and Y. Wilks (1991). *Artificial Believers.* Hillsdale, NJ: Erlbaum Associates.

Barendregt, H. (1991). Introduction to generalized type systems. *Journal of Functional Programming 1*(2).

Bar-Hillel, Y (1954). Logical Syntax and semantics. *Language 30,* 230–237. reprinted in Y. Bar-Hillel *Language and Information* Reading, Mass: Addison-Wesley.

Berger, J. and T. Luckmann (1966). *The Social Construction of Reality: a treatise in the sociology of knowledge.* New York: Doubleday, and London: Allen Lane (the Penguin Press) 1967.

Beun, R.J. (1989). *The recognition of declarative questions in infromation dialogues.* Ph.D thesis, Tilburg University.

Beun, R.J. (1998). Object reference in a shared task domain. In *Proceedings of Second International Conference on Cooperative Multimodal Communication (CMC/98),* 28–43.

Blakemore, D. (1992). *Understanding Utterances.* Oxford: Blackwell.

Borghuis, T. (1994). *Coming to terms with modal logic.* Ph.D. thesis, Eindhoven University of Technology, Eindhoven.

Bos, J., E. Mastenbroek, S. McGlashan, S. Millies, and M. Pinkal (1994). A compositional DRS-based formalism for NLP-applications. In *Proceedings of International Workshop on Computational Semantics*, Tilburg.

Brewka, G., J. Dix and K. Konolige (eds.) (1997). *Nonmonotonic reasoning: an overview.* Stanford, CA: Center for the Study of Language and Information.

Brown, P. and S. Levinson. (1987). *Politeness: some universals in language use.* Cambridge: Cambridge University Press.

Bronnenberg, W., Bunt, H., Landsbergen, J., Medema, P., Scha, R. and Schoenmakers, W. (1979). The Question Answering System PHLIQA1. In L. Bolc (ed.), *Natural language question answering systems.* London: MacMillan and München: Hanser.

Bruijn, N.G. de (1980). A survey of the Project Automath. In J. Seldin and J. Hindley (eds.), *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalisms.* New York: Academic Press.

Bunt, H. (1979). Conversational principles in question-answer dialogus. In D. Krallmann (ed.), *Zur Theory der Frage.* Essen: Narr Verlag, 119–141.

Bunt, H. (1981). *The formal semantics of mass terms.* Ph.D. thesis, University of Amsterdam.

Bunt, H. (1985). *Mass terms and model-theoretic semantics.* Cambridge, UK: Cambridge University Press.

Bunt, H. (1989). Information dialogues as communicative action in relation to partner modeling and information processing. In M. Taylor, F. Néel and D. Bouwhuis (eds.), *The structure of multimodal dialogue.* Amsterdam: North-Holland Elsevier, 47–73.

Bunt, H. (1990). DIT – Dynamic Interpretation in Text and Dialogue. In In L. Kálmán and l. Pólos (eds.), *Papers from the Second Symposium on Logic and Language*. Budapest: Akademiai Kiadó, 67–104.

Bunt, H. (1994a). Representing knowledge extracted from natural language dialogue. In *Proceedings of PRICAI'94, International Workshop on knowledge engineering and applications,* H. Bunt and Z. Shi editors, Beijing, 1–22.

Bunt, H. (1994b). Context and dialogue control. *THINK Quarterly 3*(1), 19–31.

Bunt, H. (1995). Dynamic interpretation and dialogue theory. In M. Taylor, F. Néel and D. Bouwhuis (eds.), *The structure of multimodal dialogue*, Volume 2. Amsterdam/Philadelphia: John Benjamins 2000, 139–166.

Bunt, H. (1999). Dialogue pragmatics and context specification. *This volume, 81–150.*

Bunt, H., R. Ahn, R. J. Beun, T. Borghuis, and C. van Overveld (1998). Multimodal Cooperation with the DenK System. In H. Bunt, R. J. Beun, and T. Borghuis (eds.), *Multimodal Human-Computer Communication.* Berlin: Springer, 39–67.

Bunt, H., R. J. Beun, F. Dols, J. v.d. Linden, and G. Schwartzenberg (1984). The theoretical foundations of the TENDUM system. In *IPO Annual Progress Report 19.* Eindhoven: IPO, 105–113.

Bunt, H. and R. Muskens (1999). Computational Semantics. In H. Bunt and R. Muskens (eds.), *Computing Meaning, Vol 1.* Dordrecht: Kluwer, 1–32.

Clark, H. (1996). *Using language.* Cambridge, UK: Cambridge University Press.

Clark, H. and E. Schaefer (1981). Contributing to discourse. *Cognitive Science 13,* 259–294.

Context'97: *Proceedings of the International and Interdisciplinary Conference on Modeling and Using Context.* Rio de Janeiro, February 1997.
A selection of revised papers also published as: P. Brézillon, M. Cavalcanti and R. Nossum (eds.) *Formal Aspects in Context,* Dordrecht: Kluwer 2000.

Context'99: P. Bouquet, L. Serafini, P. Brézillon, M. Beneceretti and F. Castellani (eds.) *Modeling and Using Context. Proceedings of the Second International and Interdisciplinary Conference CONTEXT'99*, Trento, September 1999.

Cohen, P. (1978). *On knowing what to say: planning speech acts.* Ph.D. thesis and Technical Report 118, Department of Computer Science, University of Toronto.

Copestake, A., D. Flickinger and I. Sag (1996). Minimal Recursion Semantics: An Introduction.
`ftp://csli-ftp.stanford.edu/linguistics/sag/mrs.ps.gz`.

Deemter, K. van and S. Peters (eds.) (1996). Semantic Ambiguity and Underspecification. Stanford, CA: CSLI Publications.

Deemter, K. van, G. Brockhoff, H. Bunt, M. Meya and J. de Vet (1985). From TENDUM to SPICOS, or: How flexible is the TENDUM approach to question answering? *IPO Annual Progress Report 20,* 83–90.

Eckert, W. and S. McGlashan (1993). Managing spoken dialogues for information services. *Proceedings of Eurospeech'93*, 1653–1656.

Gärdenfors, P. (ed.) (1992). *Belief Revision.* Cambridge, UK: Cambridge University Press.

Geurts, B. and G. Rentier (1993). Quasi Logical Form in PLUS. Esprit project PLUS report. Tilburg: Institute for Language Technology and Artificial Intelligence (ITK).

Gibbs, R. Jr. (1989). Understanding and literal meaning. *Cognitive Science 13,* 243-251.

Green, G. (1989). *Pragmatics and Natural Language Understanding* Hillsdale, NJ: Lawrence Erlbaum.

Grice, H.P. (1975). Logic and Conversation (from the Willaim James Lectures, Harvard University, 1967). In P. Cole and J. Morgan (eds.). *Syntax and Semantics 3: Speech Acts* New York: Academic Press.

Grosz, B. and S. Kraus (1996). Collaborative plans for complex group action. *Artificial Intelligence 86 (2),* 269–357.

Grosz, B. and Sidner, C. (1986). Attention, intention and the structure of discourse. *Computational Linguistics 12,* 175–204.

Guessoum, A., W. Black, J. Gallagher and T. Wachtel (1993). Abduction for Pronoun Resolution. *Proc. International Conference on Logic Programming, Workshop on Abduction.* Budapest.

Guessoum, A. and J. Lloyd (1990). Updating knowledge bases. *New Generation Computing 8*(1), 71–89.

Guessoum, A. and J. Lloyd (1991). Updating knowledge bases ii. *New Generation Computing 10*(1), 73–100.

Harris, L. (1978). The ROBOT system: Natural language processing applied to database query. Report TR87-1, Mathematics Department, Dartmouth College, Hanover NH.

Hobbs, J., M. Stickel, D. Appelt, and P. Martin (1993). Interpretation as abduction. *Artificial Intelligence 63*, 69–142.

Hendrix, G., E. Sacerdoti, D. Sagalowicz and J. Slocum (1978). Developing a natural language interface to complex data. *ACM TODS 3,* 105–147.

Hymes, D. (1972). On Communicative Competence. In J. Pride and J. Holmes (eds.) *Sociolinguistics*. Harmondsworth, U.K.: Penguin.

Iwanska, L. and W. Zadrozny (1997). Introduction to the special issue on context. *Computational Intelligence 13 (3),* 301–308.

Jekat, S., A. Klein, E. Maier, I. Maleck, M. Mast, and J. Quantz (1995). Dialogue acts in Verbmobil. Verbmobil Report No. 65, Saarbrücken: DFKI.

Kievit, L. (1996). Underspecified Logical Form. DenK Report 9627. Tilburg: Institute for Language Technology and Artificial Intelligence (ITK).

Kievit, L. (1998). *Context-driven natural language interpretation.* Ph.D thesis, Tilburg University.

Lee, M. and Y. Wilks (1996). An ascription-based approach to speech acts. In *Proceedings of 1996 International Conference on Computational Linguistics COLING-96.*

Leech, G. (1983). *The Principles of Pragmatics.* London: Longman.

Levinson, S. (1983). *Pragmatics.* Cambridge, UK: Cambridge University Press.

Luzzati, J. (2000). A Dynamic Dialogue Model for Human-Machine Communication. In M. Taylor, F. Néel and D. Bouwhuis (eds.), *The structure of multimodal dialogue*, Volume 2. Amsterdam/Philadelphia: John Benjamins, 207–222.

Landsbergen, J. and R. Scha (1976). Formal languages for semantic representation. In: S. Allén and J. Petöfi eds. *Aspects of automatized text processing.* Hamburg: Buske.

Maier, E. (1994). Dialogmodellierung in Verbmobil – Festlegung der Sprechhandlungen für den Demonstrator. Verbmobil Report No. 31, Saarbrücken: DFKI.

Maier, E. (1995). A Multi-Dimensional Representation of Context in a Speech Translation System – A Practical Approach. Verbmobil Report No. 64, Saarbrücken: DFKI.

Maier, E. (1996). Context Construction as a Subtask of Dialogue Processing - the VERBMOBIL Case. *Proceedings of TWLT 11 Workshop: Dialogue Management in Natural Language Systems,* 113–122.

McCarthy, J. (1993). Notes on formalizing context. In *Proceedings of 1993 International Joint Conference on Artificial Intellligence.*

McCarthy, J. and P. Hayes (1969). Some philosophical problems from the standpoint of artificial intelligence. In: B. Meltzer and D. Michie (eds.) *Machine Intelligence 4.* Edinburgh: Edinburgh University Press, 463–502.

McCarthy, J. and S. Buvač (1994). Formalizing context (expanded notes). Technical Note STAN-CS-TN 94-13, Stanford University.

Mey, J. (1993). *Pragmatics: an Introduction.* Oxford: Blackwell.

Medema, P., W. Bronnenberg, H. Bunt, J. Landsbergen, R. Scha, W. Schoenmakers and E. van Utteren(1975). PHLIQA 1: Multilevel semantics in question answering. *American Journal of Computational Linguistics microfiche 32.*

Moore, R. (1973). D-SCRIPT: A computational theory of description. In *Proceedings of 1977 International Joint Conference on Artificial Intelligence,* 223–227.

Moore, R. (1980). Reasoning about knowledge and action. SRI Artificial Intelligence Center Technical Note 191, Menlo Park, CA. Extended version in

J. Hobbs and R. Moore (eds.) *Formal theories of the commonsense world,* Norwood, NJ: Ablex.

Peckham, J. (1993). A new generation of spoken dialogue systems: results and lessons from the SUNDIAL project. *Proceedings of Eurospeech'93,* 33–40.

Perrault, R. (1989). An application of default logic to speech act theory, In P. Cohen, J. Morgan and M. Pollack (eds.) *Intentions and Communication.* Cambridge, MA: MIT Press.

Pinkal, M. (1999). On semantic underspecification. In H. Bunt and R. Muskens (eds.), *Computing Meaning, Vol. 1.* Dordrecht: Kluwer, 33–56.

Piwek, P. (1998). *Logic, Information and Conversation.* Ph.D. thesis, Eindhoven University of Technology.

Ranta, A. (1994). *Type-theoretical grammar.* Oxford: Clarendon.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence 13,* 81–132.

Reithinger, N. and E. Maier (1995). Utilizing Statistical Speech Act Processing in VERBMOBIL. *Proceedings of the 33rd. Conference of the Association for Computational Linguistics,* Cambridge (MA), 188–193.

Reyle, U. (1993). Dealing with ambiguities by underspecification: construction, representation and deduction. *Journal of Semantics 10,* 123–179.

Roulet, E. (1985). *L'articulation du discours en Français contemporain.* Bern: Peter Lang.

Schmitz, B. and J. Quantz (1995). Dialogue acts in automatic dialogue interpretation. *Proceedings of 6th Int. Conference on Theoretical and Methodological Issues in Machine Translation,* Leuven, 33–47.

Searle, J. (1969). *Speech Act Theory.* Cambridge, UK: Cambridge University Press.

Searle, J. (1978). Literal meaning. *Erkenntnis 13*, 207–224.

Shapiro, S. (1993). Belief spaces as sets of propositions. *Journal of Experimental and Theoretical Artificial Intelligence 5, 2-3*, 225–235.

Sinclair, J and M. Coulthard (1975). *Towards an Analysis of Discourse*. London: Oxford University Press.

Sperber, D. and D. Wilson (1986). *Relevance*. Oxford: Blackwell.

Thompson, F. and B. Thompson (1978). Rapidly extensible natural language. *Proceedings of ACM Annual Conference*, Washington DC, 178–182.

Thurmair, G. (1988). Semantic Processing in Speech Understanding. In: H. Niemann, M. Lang and G. Sagerer (eds.) *Recent Advances in Speech Understanding and Dialog Systems.* Berlin: Springer, 397–420.

Traum, D. (1994). *A Computational Theory of Grounding in Natural Language Conversation.* Ph.D. thesis, University of Rochester, NY.

Traum, D. and E. Hinkelman (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence 8 (2),* 575–599.

Verlinden, M. (1998). A constraint-based grammar for dialogue utterances. Ph.D. thesis, Tilburg University.

Wahlster, W. (1993). Verbmobil – Translation of face-to-face dialogues. In *Proceedings of 4th Machine Translation Summit,* Kobe.

Wilks, Y., J. Barnden and A. Ballim (1991). Belief ascription, metaphor and intensional identity. *Cognitive Science 15*,1.

Waltz, D., T. Finin, F. Green, F. Conrad, B. Goodman and G. Hadden (1978). The PLANES system: Natural language access to a large database. Report T-34, Univ. of Illinois, Urbana Ill.

Wilensky, R., D.N. Chin, M. Luria, J. Martin, J. Mayfield and D. Wu (1988). The Berkeley Unix Consultant Project. *Computational Linguistics* **14**(4), 35–84.

Winograd, D. (1972). *Understanding Natural Language.* New York: Academic Press.

Woods, W., R. Kaplan and B. Nash-Webber (1972). The lunar sciences and natural language information system: Final report. BBN Report 2378, Cambridge, MA: Bolt, Beranek and Newman.