PRAGMATICS AND DIALOGUE

GEOFFREY LEECH

MARTIN WEISSER

ABSTRACT

This chapter introduces the linguistic subdiscipline of pragmatics (the investigation of meaning in context) and shows how this is being applied to the development of spoken dialogue systems-currently perhaps the most important applications area for computational pragmatics. Sections 7.1-7.5.5 trace the history of pragmatics from its philosophical roots, and outline some key notions of theoretical pragmatics-speech acts, illocutionary force, the cooperative principle, implicature, relevance. Since pragmatics is concerned with meaning, most of its basic terms are conceptual: intention, belief, inference, and knowledge. Sections 7.6-7.9 turn to the application of pragmatics to dialogue modelling, especially the development of spoken dialogue systems intended to interact with human beings in task-oriented scenarios such as providing travel information. One major illustration of this is the application of speech act theory in the analysis and synthesis of service interactions in terms of dialogue acts (utterance units defined as having a functional role in the dialogue).

7.1 WHAT IS PRAGMATICS?

Thirty years ago, pragmatics was a fledgling branch of linguistics. In the 19705 it gained in importance, and remains an important subdiscipline within linguistics, with its own journals, handbooks, and international association.'

Only recently has pragmatics begun to be a major focus of research in computational linguistics, mainly because of its relevance to the development of spoken dialogue systems (SDSs), that is, computer systems designed to

engage in purposeful dialogues with human beings. This chapter will focus on computational pragmatics in the context of spoken dialogue, although, on a more general level, pragmatics also applies to written language communication: for example, to the disambiguation of meaning (see Chapter 13) and the assignment of reference to personal pronouns such as she and they (see Chapter 14).

Pragmatics is the branch of linguistics which seeks to explain the meaning of linguistic messages in terms of their context of use. It is seen as distinct from semantics, which investigates meaning in a more abstract way, as part of the language system irrespective of wider context. In semantic terms an utterance2 can therefore often be ambiguous, whereas the contextual setting in which the utterance occurs, combined with its intonation, would-in most cases-serve to disambiguate its function.

One way to differentiate pragmatics from semantics is to say that in pragmatics, meaning is a triadic relation: `Sp means x byy'; while in semantics meaning is a dyadic relation: `y means x. This can be illustrated by the frequently quoted example of the utterance `It's cold in here.' If we interpret this utterance on a purely semantic level, it simply states a literal or face-value meaning, i.e. the fact that the temperature in the place where the utterance has occurred is low. However, given the context that there are at least two people in the room at the time of the utterance and that the window is open, the same utterance can additionally take on a different meaning, depending on the context and the speaker's intention. For example, if the social relation between the interlocutors is appropriate, it can take the meaning that Sp wants H to close the window: in effect, it is a request. But the relationship between the two interlocutors is by no means the only personal factor that may influence the interpretation of the utterance. The willingness and ability of the hearer to cooperate with the speaker's request, for example, are amongst many other factors that can affect the meaning. Pragmatics in general is concerned with questions such as:

- What does a listener suppose a speaker to intend to communicate by a given message? And how is this meaning decoded?

- What persons, entities, etc. does the message refer to?

- What background knowledge is needed to understand a given message?

- How do the beliefs of speaker and hearer interact in the interpretation of a given message, or of a given dialogue exchange?

- What is a relevant answer to a given question?

[Pragmatics originated in philosophical thought (e.g. in the work of Charles Morris, J. L. Austin, John Searle, and H. P. Grice)3](#) and may still show a tendency towards academic abstraction which makes it difficult to adapt to concrete computational applications. In the following sections, we will first give a brief overview of some of the theoretical constructs that form the basis of modern-day pragmatics. We will then go onto show how and why computational pragmatics differs from `linguistic' pragmatics, and how pragmatics contributes to the computational analysis of dialogues, with particular respect to SDSs.

7.2 SPEECH ACTS AND ILLOCUTIONARY FORCE

7.2.1 Speech acts

One of the philosophical foundations of pragmatics can be found in the notion of illocutionary acts (often simply called speech acts) as developed by J. L. Austin and J. R. Searle. The idea behind a speech act is that meaning can be explained in terms of action, rather than in terms of concepts like reference and truth conditions. Most philosophical approaches to language since Aristotle had always assumed that to make an utterance is almost by default to state something that can be specified as either true or false. Austin disputes this, saying that

One thing, however, that it will be most dangerous to do, and that we are very prone to do, is to take it that we somehow know that the primary or primitive use of sentences must be, because it ought to be, statemental or constative, in

the philosopher's preferred sense of simply uttering something whose sole pretension is to be true or false and which is not liable to criticism in any other dimension. (Austin 1962:72)

He makes a distinction between the above-mentioned constative utterances and ones that he refers to as performatives, such as `I apologize: utterances that do not state anything about the world, but rather constitute verbal actions. Such utterances may contain an overt performative verb, such as apologize above, or else the performance of an action may remain implicit. For example, a request such as `Could you post this letter?' is an utterance which acts as an attempt to bring about some change through action by the addressee.

According to Austin, such utterances can be characterized in terms of three kinds of verbal act: locution, illocution, and perlocution. The notion of locution here is closest to the literal use of an utterance with a particular sense, whereas illocution relates to what the speaker (Sp) intends to perform, and perlocution relates to what is achieved-including uptake by the hearer (H).4 Let us go back to our earlier example-the utterance `It's cold in here'-to see how we can analyse it according to Austin's principles. The locution is simply the words used to form the utterance and the grammatical form of the utterance expressing a proposition. As for the illocutionary force or intended meaning behind it, we can assume that, given the context of the open window, Sp wants to have H close the window, which would indicate that the illocution or pragmatic function is that of a directive (or request). The perlocutionary effect of the utterance is then twofold, depending on (a) whether H understands the utterances of Sp and (b) if so, whether or not H is actually willing to comply with the request.

This simple example shows the different conceptual levels on which meaningful action works to explain the creating and disambiguation of meaning. It also illustrates the key problem of relating pragmatic (or illocutionary) force to the syntax and semantics of an utterance. As we have just seen, syntactically the utterance appears to be a statement, but illocutionary force and perlocutionary effect realize it as a directive or request, showing an indirect relation between form and function, or between grammar and intended meaning. Perlocutionary effect has on the whole been

neglected in academic pragmatics, since it lies strictly outside the domain of language and its interpretation. In computational pragmatics, however, it cannot be ignored, as it is the key to how one interlocutor responds to another in SDSs.

The perlocutionary component of the utterance also highlights the importance of mental constructs in pragmatics: both Sp and H have certain beliefs that affect their intentions or goals in an exchange, as well as the effect of utterances. We shall see in section 7.8.3 below how this affects issues in computational pragmatics.

Another aspect of Austin's theory is that certain conditions, which he terms felicity conditions, have to be fulfilled every time we perform a verbal action. For example:

(A.1) There must exist an accepted conventional procedure having a certain conventional effect, that procedure to include the uttering of certain words by certain persons in certain circumstances, and further,

(A.2) the particular persons and circumstances in a given case must be appropriate for the invocation of the particular procedure invoked.

(B.1) The procedure must be executed by all participants both correctly and

(B.2) completely. (Austin 1962:15-16)

As we see, Austin's understanding of verbal actions reflects the idea of explicitly achieving those actions according to convention rather than by implication.

In following Austin, Searle formalizes illocutionary acts as `rule-governed intentional behaviour' (Searle 1969:16) and claims that:

the semantic structure of a language may be regarded as a conventionalized realization of a series of sets of underlying constitutive rules , and that speech acts are acts characteristically performed by uttering expressions in accordance with these sets of constitutive rules. (Searle 1969:37)

Four types of rules serve to define different illocutionary acts in different ways (see Searle 1969:57, 62):

(7.1) Propositional act   the propositional content
       Preparatory condition(s) conditions that have to hold in order for the speech act to be
                                 possible, e.g. for something to have happened or be possible
                                 to happen or to be desirable
       Sincerity condition   beliefs or intentions of Sp
       Essential condition   what the speech act 'counts as' in illocutionary terms

He also proposes a typology for speech acts, here summarized according to Searle (1979):

(i) assertives commit Sp to the truth of some proposition (e.g. stating, claiming, reporting, announcing);

(ii) directives count as attempts to bring about some effect through the action of H (e.g. ordering, requesting, demanding, begging);

(iii) commissives commit Sp to some future action (e.g. promising, offering, swearing to do something);

(iv) expressives count as the expression of some psychological state (e.g. thanking, apologizing, congratulating);

(v) declarations are speech acts whose `successful performance ... brings about the correspondence between the propositional content and reality' (e.g. naming a ship, resigning, sentencing, dismissing, excommunicating, christening).

7.3 H.P. GRICE'S COOPERATIVE PRINCIPLE

Another of the major philosophical foundations of computational pragmatics is H. P. Grice's CP (Cooperative Principle), which holds that conversation takes place on the assumption (barring evidence to the contrary) that the interlocutors are being cooperative in contributing to the general goals of the

conversation. The CP can be understood to apply to communication in general. It has four constituent subprin- ciples, which are expressed in the form of maxims to be followed by Sp (the following is a simplification of Grice 1975):

1. Maxim of Quantity (or informativeness): give the right amount of information;

z. Maxim of Quality (or truthfulness): try to make your contribution one that is true;

3. Maxim of Relation(or relevance): be relevant;

4. Maxim of Manner: avoid obscurity or ambiguity; be brief and orderly.

The crux of Grice's explanatory framework is that, since in general we can assume that the CP is being observed, apparent departures from the CP can be accounted for on that basis. An apparent breach of truthfulness, for example, may be due to wilful lying, or to a mistake-or it may be because Sp is trying to get a special point across, e.g. through metaphor or irony. The last case is said to be interpreted by implica- ture-or pragmatic implication (see section 6.4.4). An implicature is weaker than logical implication in that it is defeasible: that is, it can be rejected if other evidence contradicts it. From H's point of view, this is where inference (see section 7.5.4 below) plays a crucial role. Thus if H perceives that Sp is not expressing a literal or face-value meaning in accordance with the CP, H can assume that an alternative interpretation is intended. H therefore attempts to infer (from contextual information, the literal meaning of the utterance, and general principles of communication such as the CP) an interpretation that would make S's utterance rational and cooperative, and thus arrive at a conclusion about what Sp intended to communicate.

Here are examples of the four maxims at work.

1. Maxim of Quantity

If someone says

(7.2) Maggie ate some of the chocolate

it will generally be inferred that the speaker believes that:

(7.3) Maggie did not eat all of the chocolate.

The reasoning is that if Maggie had been noticed eating all the chocolate, the speaker would have been informative enough to say so. Note that (7.2) does not entail (7.3), because it is quite possible to truthfully assert the conjunction of (7.2) and (7.3):

(7.4) Maggie ate some of the chocolate-in fact, she ate all of it.

2. Maxim of Quality

If someone says, talking about an expensive dental treatment,

(7.5) That'll cost the earth

it will generally be assumed that the speaker is not telling the truth (because the proposition (7.5) is not believable). However, the message conveyed will be a proposition closely related to (7.5), in that (7.5) implies it:

(7.6) That'll costa very large amount.

3. Maxim of Relation

In the following exchange,

(7.7) Child: Can I watch TV?
       Parent: It's bath time, Rosy.

the parent's reply apparently does not answer the child's question, and therefore breaks the Maxim of Relation. However, even a child can work out the missing part of the message: `Because it's bath time, there is no time to watch TV, and therefore you cannot.

4. Maxim of Manner

If someone, instead of (7.8), says (7.9):

(7.8) Are you ready?

(7.9) I am asking you whether you are ready or whether you are not ready.

it is obvious that the speaker is not choosing the quickest way of asking for the desired piece of information. This longwindedness will generally be assumed to convey an implicature-probably the implicature that the hearer is being unhelpful in withholding the information concerned.

7.4 COLLABORATION

Grice's idea of cooperation seems to be consistent with more recent research in both conversational analysis (CA) and linguistics that stresses the importance of interlocutors' interactive collaboration in constructing the meaning of exchanges between them (cf. Schegloff 1996; Ono and Thompson 1996). Ono and Thompson demonstrate how this kind of collaboration works even on the level of syntax, by giving examples of participants completing each other's sentences or recovering and repairing them. This clearly does not mean that the CP cannot be broken: there are many occasions where Sp and H do not cooperate in terms of the maxims, and indeed it is arguable that such concepts as informativeness, truthfulness, and relevance are matters of degree, rather than absolute quantities.

7.5 CONCEPTUAL REPRESENTATIONS

From H's point of view, pragmatics deals with the communicative effects that an utterance can have, whether Sp intended them or not. However, communication takes place on the understanding that speaker and hearer share beliefs or assumptions. Therefore one of the key issues relevant to an

understanding of pragmatics is what beliefs or assumptions both Sp and H need to bring into play when producing and interpreting an utterance. Pragmatics requires that propositional attitudes such as `Sp intends x' and `H assumes that y' be represented as part of Sp's or H's meaning.

7.5.1 Intentions or goals

On the part of Sp, there is usually at least one intention or goal behind the production of an utterance. This goal (or set of goals) underlies the illocutionary force of the utterance, for example whether it is intended to inform, to request, to thank, or to offer. However, while the intentions of a speaker may sometimes be relatively easy to understand, in some cases Sp may not manage to convey his or her intentions successfully to the hearer. Sperber and Wilson (1995: 34) cite the following potentially ambiguous example of a dialogue:

(7.10)   *Peter*: Do you want some coffee?
          Mary: Coffee would keep me awake.

As this example is presented, it is not clear whether Mary intends to accept Peter's offer, implying that coffee would enable her to stay awake a little longer, or whether she is refusing his offer because she would have trouble getting to sleep later on. Although this may be a constructed example and in real life Mary could (or would) disambiguate her reply by prefixing it with something like either `Thanks' or `No, thanks, it could conceivably occur in natural language and therefore presents a problem for interpretation. The intended meaning of Sp and the interpretative meaning of H may not correspond: that is, misunderstandings can (and often do) occur.

7.5.2 Beliefs and assumptions

Pragmatics is concerned with interlocutors' beliefs and assumptions about the world, and this includes beliefs about the other interlocutor(s), including their beliefs and intentions. For example, a speaker who makes a request will usually believe that there is a chance that H will comply with the request. In

communication, there are nth order beliefs, just as there are nth order intentions. For example, a second-order belief is normally a belief about someone else's beliefs. A third-order belief can bring in the mutuality of beliefs between Sp and H, and potentially leads to infinite regress, e.g. Sp believes that H believes that Sp believes ... The belief systems attributed to interactants in a dialogue are often complex, and cannot be ignored in computational pragmatics (see Ballim and Wilks 1991).

7 .5.3 Knowledge

Knowledge can be seen as a specially privileged type of belief, a belief that is sanctioned by logic or authority or experience to be a fact. In pragmatic terms, knowledge may be shared by interlocutors, or else may be confined to one interlocutor or the other. Mutual or shared knowledge is often discussed as a key category in explaining communication-it is knowledge which is not only shared by the interlocutors, but known by each interlocutor to be shared by the other interlocutor. However, note that this so-called `knowledge' is fallible, as there is ultimately no guarantee that both interlocutors actually do share the same knowledge. Hence it is better to think of `assumed mutual knowledge' rather than `mutual knowledge: Sperber and Wilson (1995: 17-21, 40-2) think more realistically in terms of `mutually manifest assumptions'which interlocutors share in a`mutual cognitive environment.

7 .5.4 Inference

If intention is the key to meaning from the point of view of Sp, inference is a key concept from the point of view of H, the addressee. Inference is here understood as the use of reasoning to derive a new belief from a set of existing beliefs. Note that reasoning in pragmatics often deviates from classical deductive logic: it is more like the common-sense practical reasoning which human beings use in everyday situations. For example, we may conclude, noticing that the streets are wet, that it has been raining. The classical law of Modus Ponens here appears to apply in the reverse order

(this example and the subsequent discussion are adapted from Bunt and Black 2000: 12-13):

(7.11) If it has (recently) been raining, the street is wet

(7.12) It has (recently) been raining

Therefore

(7.13) The street is wet

In practical reasoning, conclusion (7.13) is our starting point, and in a classically invalid move, we derive from it premise (7.12), which may be characterized as the hypothesis we use to explain what we observe. There could, of course, be other explanations, such as flooding or burst water pipes, but rain is the most obvious one. This kind of reasoning has been formulated in terms of abductive logic (where the reasoner is `permitted to assume additional premisses in order to reach a conclusion deductively'-ibid.). Another way of formulating it is in terms of default logic, using rules of the form: `If p then q, unless there is evidence that not-q' (ibid.). In the case of wet streets, the hypothesis that it is raining is the default assumption we make in normal circumstances. This kind of logic can be readily applied to Grice's implicatures. For example, the assumptions of the CP (see section 7.3), that speakers are being truthful, informative, and relevant, is a useful default assumption which may nevertheless be invalidated by contrary evidence. Thus the addressee, while not having direct access to the intentions of Sp, can infer them from what Sp says, as well as from additional `givens, notably contextual information and general principles of communication, especially the CP.

7 .5.5 Relevance

The third maxim of the CP, `be relevant', although vague in Grice's original formulation, has been formulated in detail and elevated to the major explanatory principle of pragmatics by Sperber and Wilson (1995), for whom it renders the other three maxims of Grice's CP unnecessary. Sperber and

Wilson's Principle of Relevance (1995: 125) explicates the way the interpreter makes sense of what is said, by a trade-off between largeness of contextual effects and smallness of processing effort. By contextual effects is meant some addition to the addressee's set of assumptions (`fresh information') derived from the utterance in relation to its context. By processing effort is meant the amount of mental effort, notably in inference, the addressee has to expend in order to arrive at the interpretation. To revert to our stock example, `It's cold in here' brings no contextual effects if it is interpreted as a mere remark about the low temperature, which is presumably already apparent to the addressee. Hence, in this case, the interpretation as a request to close the window will be more relevant in terms of contextual effects, but will be less relevant in so far as a less direct interpretation costs greater processing effort to figure out the meaning. Perhaps the request interpretation will win out, because the lack of contextual effects does not even justify the small processing effort needed to arrive at the bare assertion about temperature. In Grice's terms, this would be a breach of the Maxim of Quantity (the remark would be uninformative) leading to an implicature (Sp wants the window closed).6

   An alternative formulation of relevance is that of Leech (1983: 94-6), who sees it as the contribution an utterance makes to the (assumed) goals of Sp (whose goals may include helping to satisfy the goals of H). Thus, in this case, H is able to arrive at the request meaning, by hypothesizing that Sp wishes to accomplish a goal: that of raising the temperature.

7.6 DIALOGUE IN COMPUTATIONAL PRAGMATICS

To understand how computational pragmatics relates to dialogue, we have first to ask the question: `What makes computational pragmatics different from "theoretical" pragmatics?' This, in turn, invites the question: `Why is the dialogue between two or more human interlocutors likely to be different from that between humans and computers?'

7.7 ORDINARY DIALOGUE

We may here term `ordinary' dialogue the kind of dialogue that we as humans engage in every day, where two or more people communicate with one another, either face to face, by telephone or even in writing. This is the kind of dialogue that is normally the subject of the study of (linguistic) pragmatics, discourse analysis, and conversational analysis. Ordinary dialogue is essentially unrestricted in the range and complexity of topics and goals addressed, although constrained by such factors as the amount of knowledge or communicative ability the interlocutors bring to it. We can say that there is no restriction on a dialogue's domain (the kind of topic or subject matter it deals with) or activity type (the genre of activity to which the dialogue contributes).

Ordinary dialogue may also be goal oriented, i.e. intended to achieve certain pre determined aims, as in our example `It's cold in here; but most of it comes under the heading of `casual conversation, where goals can be shifting and ill defined. Ordinary dialogue moreover typically involves social interaction, which is very different from the kind of interaction that we normally wish to have with a computer.

7.8 COMPUTATIONAL DIALOGUE

Dialogue involving computers differs in many respects from ordinary dialogue, as it is subject to a number of specific constraints, both of a technical and of an interactional nature. Here we mention two. First of all, almost all human communication with the computer is manifestly task oriented, i.e. goal restricted in seeking to achieve a practical outcome by definable procedures. Secondly, dialogue involving computers is usually highly restricted in domain. This severe domain restriction is not surprising, SDSs being among the most ambitious challenges that face computational linguistics. They integrate most of the components of natural language

processing, including speech recognition, language understanding, information extraction, language generation, and speech synthesis. In addition to these, a dialogue management component is required to interpret the goals of incoming utterances and plan an appropriate response: this is where pragmatics has a key role. Without radical simplifications brought by domain restriction, combining and coordinating all these components would be well beyond current capabilities.

In this chapter, we confine our attention to SDSs which exhibit intelligence in the sense that they involve some kind of pragmatic processing, taking into account the goals and interpretations of utterances. Not all dialogues with computers are of this kind. A well-known exception to the task-driven nature of human-machine dialogue is the conversational system ELIZA (see Weizenbaum 1966). This was built in the 196os to simulate human-computer conversation, and operated by mainly responding to keywords and patterns in the user input to ask seemingly intelligent questions or give non-committal answers like `I'm sorry to hear that XXX' or `Tell me more about your parents. Different implementations of ELIZA can be found at http://128.2. 242.152/afs/cs/project/ai-repository/ai/areas/classics/eliza/o.html. ELIZA was innocent of pragmatics, as of all aspects of linguistic knowledge. More recently, conversational systems have been competitively entered for the Loebner Prize (http://www. loebner.net/Prizef/loebner-prize.html), offered every year to a computer system which is judged to come closest to passing the Turing Test, the test as to whether a computer system's observed (verbal) behaviour is indistinguishable from that of a human being. These, again, do not fall within the definition of SDSs considered here. (A more detailed treatment of SDSs is to be found in Chapter 34.)

7.8.1 Dialogue models and dialogue typology

To establish a typology of the dialogue models we are likely to encounter in computational analysis of dialogue, let us first look at the range of possibilities, i.e. establish what participants can be involved in a task-driven dialogue and in what way. The case that is most similar to the kind of ordinary pragmatics we discussed above is that of human-human dialogue, which can occur in two forms:

(a) non-machine-mediated: ordinary every-day human dialogue that is analysed using the computer.

(b) machine-mediated: a special type of dialogue between two or more people, which is monitored by the computer, so that the computer can offer assistance where the participants have trouble communicating in a lingua franca.

Type (a) is the kind of dialogue that computational linguists, as well as other linguists and conversational analysts, may analyse by extracting and modelling (aspects of) dialogue behaviour. For developing SDSs, such dialogue data may be recorded, transcribed, and analysed in order to provide the basis for a predictive model of human dialogue behaviour in task-oriented domains. For example, researchers may collect a dialogue corpus (see Chapter 24) of data from telephone call centres providing a public service, such as airline or train information and booking services, with a view to building an automated system to perform the same service. It is evident, however, that the human dialogue data so collected will differ in some respects from user dialogue behaviour when communicating with a computer system. For example, the data are likely to contain more socially appreciative utterances (such as `Uh that's wonderful' and `Okay thank you very much') than would occur in corresponding human-computer dialogue.

In type (b), the computer is used only in order to assist human-human communication to achieve a problem-solving task. A good example of this type of dialogue is the German VERBMOBIL system. Machine-mediated dialogue resembles non-mediated human-human dialogue in the way it is processed by computer (e.g. keeping track of keywords Sps use), yet in other ways it resembles our second main category, that of human-machine dialogue.

Human-machine dialogue is any kind of dialogue where a user communicates with a computer interface in order to achieve a set of aims. There are essentially two different kinds of human-machine dialogue:

(a) simulated: both participants are human, but one pretends to be a computer system. The computer interface is a `sham.

(b) non-simulated: genuine interaction between human and computer.

Type (a) is normally set up to investigate the behaviour of a user towards what he or she assumes to be a computer. This is an important means of user modelling since the behaviour of humans supposedly communicating with machines can be the best basis for human-computer dialogue modelling during system design. As such simulations recall the Wizard of Oz's deception in the Walt Disney movie, they are normally called Wizard of Oz (WOZ) experiments (see Gibbon, Moore, and Winski 1998: 581).

Type (b) is obviously the kind of human-machine dialogue which results from the implementation of fully-fledged SDSs. At present, such systems are relatively speaking in their infancy, but many are being developed as research prototypes, and a few have been commercially implemented. McTear (1999) provides an informative survey of the current state of the art, including working systems. As he explains (1999: 8), this technology is becoming important as `it enables casual and naive users to interact with complex computer applications in a natural way using speech:

Having established which combinations are possible, we can now look at the different types of dialogue we are likely to encounter. We have already noted that, unlike ordinary dialogue, computational dialogue is so far only possible in restricted domains, i.e. with clearly delimited topics and vocabulary. So far, the domains covered in computational dialogue (with sample systems) include:

(a) travel information (SUNDIAL, ATIS, Philips Automatic Train Timetable Information System)

(b) transport (TRAINS)

(c) business appointments (VERBMOBIL)

(d) access to on-line information (SUN SpeechActs)

(e) repair and assembly (Circuit-Fix-It Shop)

Other domains under development include telebanking, directory enquiry services, and computer operating systems. The domain of a dialogue heavily influences the kind of vocabulary and background information the computer has to understand. For example, a travel information system needs to `know' a large number of names of locations, whereas a telebanking application will have to'be aware' of financial matters such as currencies, balances, and statements. A stored knowledge base will normally contain precise information about these specialized topics. Since existing computational SDSs are also task driven, the system is expected to perform one or more tasks and has to have some knowledge of how specific tasks are commonly performed. Systems are in general also applications oriented, i.e. they are not just there for the user to be able to have a conversation, but are designed to form a part of a specific application. This means that, to be commercially viable, they have to achieve a high level of accuracy in decoding and interpreting utterances, and in giving error-free responses.

The specific tasks to be performed by a system are closely bound to the domain in which they occur, e.g.:

(a) Negotiating appointments and travel planning (VERBMOBIL)

(b) Answering airline/travel enquiries (SUNDIAL, ATIS, Philips Automatic Train Timetable Information System)

(c) Developing plans for moving trains and cargo (TRAINS)

Additionally, SDSs can be categorized according to activity types:

(a) cooperative negotiation (VERBMOBIL)

(b) cooperative problem solving (TRAINS, Circuit-Fix-It Shop)

(c) information extraction (SUNDIAL, ATIS, Philips Automatic Train Timetable Information System)

Type (a) can normally occur only in systems where there are at least two human interlocutors and the computer present, although appointments can also

be made by one human participant who gains access to a scheduler via a system such as the SUN SpeechActs system (see Martin et al. 1996). Types (b) and (c) are more typical of (single) human-computer interaction.

7.8.2 From speech acts to dialogue acts

It is not part of the task of this chapter to detail the various components of speech and language processing required for the decoding and encoding of components of a human-machine dialogue. They are dealt with in other chapters, for example Chapter i6 (speech recognition), Chapter 12 (parsing), Chapter 15 (natural language generation), and Chapter 17 (speech synthesis). Of primary relevance to this chapter, however, are the pragmatic levels of interpretation and generation of utterances, which maybe roughly thought of as the computational linguist's adaptation of Searle's speech act theory. In computational pragmatics, however, the term more often used nowadays for such speech acts as REQUEST and ASSERTION is dialogue acts. Utterance interpretation, in terms of the identification of such pragmatic categories with their domain-associated content, is the key to the linguistic interpretation of incoming messages from the human users. Dialogue act interpretation, in such terms, has to depend substantially on the information derived from lower decoding levels, including phonic (including prosodic), lexical and syntactic decoding, although contextual information, in terms of what dialogue acts have preceded, also plays a major role.

   We are here dealing with an area of computational linguistics which is still under development. Although dialogue acts are already being used in a number of systems with a certain degree of success, attempts are still being made to classify them to a higher degree, and to standardize them across a variety of different domains. Much effort is currently going into the compilation and annotation of corpora of dialogues (see Leech et al. 2000), so that these can act as training data for the development of automated systems. A common practice in dialogue research is to give parts of corpora, such as the corpora developed for the ATIS or TRAINS systems, to naive or expert subjects, who are then asked to segment them according to the functions of their individual parts (cf. Nakatani, Grosz, and Hirschberg 1995; Passonneau and Litman 1996; Carletta et al. 1997). The resulting decomposition into

utterances by a combination of structural and functional criteria (Leech et al. 20oo)' determines both relations between the individual parts of the dialogue and their functional content. This can then lead to the development of improved dialogue (and dialogue management) models. The functional units- i.e. dialogue acts-differ from the speech acts occurring in everyday conversation, in that their scope is defined and potentially limited by the domain that they occur in, as well as their task orientation. Thus a model can often be built for a specific kind of dialogue with the help of relatively simple techniques such as the identification of keywords and phrases, and observing under which conditions and where within the task performance they have been used.

Some of the seminal work on dialogue acts has since 1996 been done by members of the Discourse Resource Initiative (DRI). The DRI consists of researchers in the fields of discourse and dialogue who have met at annual workshops in order to discuss their research and perform annotation exercises in order to develop and test new coding schemes (cf. Allen and Core 1997). An abridged annotation example from a recent workshop, held in May 1998 in Chiba, Japan, can be seen below:

1	A: so	ASSERT(?), DIRECTIVE, COMMISSIVE
2	we should meet again	ASSERT(?), DIRECTIVE, COMMISSIVE
3	how 'bout	DIRECTIVE, COMMISSIVE
4	how 'bout next week	DIRECTIVE, COMMISSIVE
5	what day are good for you	ABANDONED, INFO-REQUEST
6	what days are good for you	ABANDONED, INFO-REQUEST
7	B: actually next week I am on vacation	ASSERT, REJECT(3, 4), ANSWER(5, 6)
8	A: gosh	ACKNOWLEDGE(7), EXCLAIM
9	I guess we will have to meet the week after that	ASSERT, DIRECTIVE, COMMISSIVE, ACCEPT(7)
10	how 'bout Monday	DIRECTIVE, COMMISSIVE
11	B: Monday the tenth	INFO-REQUEST(?)
12	A: uh-huh	ANSWER(11)
13	B: well unfortunately my vacation runs through the fourteenth and I have plane tickets	ASSERT, REJECT(10–11)
14	I was planning on being on a beach in Acapulco about that point	ASSERT, REJECT(10–11), EXPLANATION(13)
15	A: well	? ACKNOWLEDGE(13–14)
16	when are you getting back	INFO-REQUEST

In this extract, some of the labels have been expanded to make them more intelligible: e.g. INFO-REQ has become INFO-REQUEST. Some of the categories are clearly related to those of Searle (see section 7.2.1 above).

Dialogue acts can in principle be differentiated according to whether they actually contribute to the task itself or whether they serve a task management role, although this is not always an easy distinction to make. For example, stating or requesting new information is normally a direct contribution towards the performance of task goals, whereas clarifications, backchannels, or repairs can be seen as contributions towards the maintenance and management of the task. Other dialogue acts occur during certain phases of the dialogue, e.g. greetings at the beginning and closures towards the end. There is hence a need to recognize higher units of dialogue, to which dialogue acts

contribute. In the VERBMOBIL scheme, a distinction is made between the following phases (Alexandersson et al. 1997: io):

1. H -Hello

2. O - Opening

3. N - Negotiation

4. C - Closing

5. G -Goodbye

Dialogue acts of greeting and introduce are to be expected only in phases (1) and 5), initiate in (2) and accept, reject or request in (3) and (4).

From the system's point of view, the ongoing structure of the dialogue, in terms of dialogue acts or higher units, has to be monitored and controlled by the dialogue manager, to which we now turn.

7.8.3 Dialogue management models

As an SDS is subject to a large number of constraints, attempts have to be made to control the dialogue between system and user in as tight a way as possible, to enable the system to perform its tasks within those constraints and to pre-empt any misunderstandings.

In order to perform a specific task, it is not enough for either the system or the user to have access only to a kind of knowledge base of domain knowledge. Just as in the development of human conversation, the knowledge and intentions of both user and system need to be constantly augmented, i.e. a dynamic context knowledge and ongoing intentional structure (cf. Grosz and Sidner 1986:187) need to be created. To keep track of these is the responsibility of the dialogue manager. Dialogue management models are based on the notion of a cooperative achievement of the task, and are of three

main varieties: dialogue grammars, plan-based approaches, and approaches based on the joint action model.

Dialogue grammars are the oldest and simplest form of dialogue management. They assume that the task has a fixed structure of finite states representing dialogue acts (cf. Cohen 1998 and McTear 1999), and are usually arranged according to the conception of adjacency pairs (Sacks 1967-72) postulated in conversation analysis, e.g. questions followed by answers, etc. However, because of their relatively inflexible structure and the need for all structural options to be `hard-coded, they are only suitable for small-scale systems and are rarely used these days. One additional problem with dialogue grammars is that the initiative rests solely with the system, i.e. the user is constrained in what he or she can say or has to say at any given time.

Plan-based systems are a more flexible way of dealing with the flexibility required of modern SDSs. The following description is from Litman and Allen (1990:371):

Every plan has a header, a parameterized action description that names the plan. The parameters of a plan are the parameters in the header.... Decompositions enable hierarchical planning. Although a plan maybe usefully thought of as a single action at the level of description of the header, the plan may be decomposed into primitive (that is, executable) actions and other abstract action descriptions (that is, other plans). Such decompositions may be sequences of actions, sequences of subgoals to be achieved, or a mixture of both.... Also associated with each plan is a set of applicability conditions called constraints.... A library of plan schemas will be used to represent knowledge about typical speaker tasks. Plan instantiations are formed from such general schemas by giving values to the schema parameters.

Litman and Allen also make a distinction between domain plans, i.e. global domaindependent task plans, and discourse plans, which are domain-independent `metaplans' that regulate the general flow of any dialogue (cf. the task management functions of dialogue acts mentioned above in section 7.8.2).

Approaches based on the joint action model (cf. Cohen 1998) are a relatively recent development. Even more than plan-based systems, they stress

the collaborative effort participants engage in to achieve their aims. Like plan-based systems, they belong to the realm of mixed-initiative systems, where either the system or the user can take the initiative at any given time.

7.9 CONCLUSION

In spite of the inherent problems and complexities of the SDSs, intensive research and development in the area will doubtless lead to substantial advances in the next few years. Returning to the difference between academic and computational prag matics, we ask how far academic approaches are now being reflected in the evolution of SDSs. Simpler approaches, emphasizing dialogue grammar, draw most clearly on rule-based conceptions of pragmatics, notably the speech act theory of Searle, which ironically lends itself more to the controlled nature of task-driven systems than to most ordinary dialogue. As the versatility of computational dialogue models increases, we are seeing a greater influence of theoretical approaches which emphasize the rational, cooperative basis of human-machine dialogue, with their philosophical roots in Grice's CP and related theory. As human-machine dialogue takes on more of the flexible characteristics of ordinary dialogue, the relevance of insights from academic pragmatics is likely to increase.

FURTHER READING AND RELEVANT RESOURCES

The classical texts Austin (1962), Searle (1969,1980) and Grice (1975) are relatively easy and stimulating to read. On relevance, Sperber and Wilson (1995) has also attained classic status, although more demanding. In computational pragmatics, Ballim and Wilks (1991) deals with belief, and Bunt and Black (2000) with pragmatic reasoning. On SDSs, McTear (1999) and Leech et al. (2000) give surveys of the fast developing research and development scene.

Although there are no websites that specifically deal with the topic of computational pragmatics as a whole, below is a list of sites that provide comprehensive information on individual aspects, such as dialogue coding etc., involved in the study of computational pragmatics. These sites also include pointers to many other relevant sites.

Discourse Resource Initiative: http://wwwgeorgetown.edu/luperfoy/Discourse- Treebank/dri-home.html. General discourse research and annotation with pointers to their annual workshop pages.

DAMSL (Dialog Act Markup in Several Layers): http://www.cs.rochester.edu/ research/trains/annotation/RevisedManual/RevisedManual.html. An annotation scheme for dialogues.

EAGLES WP4 homepage: http://www.ling.lancs.ac.uk/eagles/. Survey and guidelines for the representation and annotation of dialogues.

MATE (MultilevelAnnotation,Tools Engineering) project: http://mate.mip.ou.dk/. Survey and development of dialogue annotation schemes and tools.

TRINDI: http://www.ling.gu.se/research/projects/trindi/. Building a computational model of information revision in task-oriented and instructional dialogues and instructional texts.

VERBMOBIL project: http://www.dfki.uni-sb.de/verbmobil. Large-scale dialogue annotation and translation project.

REFERENCES

Alexandersson, J., B. Buschbeck-Wolf, T. Fujinami, E. Maier, N. Reithinger, B. Schmitz, and M. Siegel. 1997. Dialogue acts in VERBMOBIL-2. VM-Report 204 DFKI GmbH, Stuhlsatzen- hausweg 3,66123 Saarbri cken.

Allen, J. and M. Core. 1997. Draft of DAMSL: DialogActMarkup in Several Layers. Available from: http://www.cs.rochester.edu/u/trains/annotation/RevisedManual/ RevisedManual.html.

Austin, J. L. 1962. How to Do Things with Words. Oxford: Oxford University Press.

Ballim, A. and Y. Wilks. 1991. Artificial Believers: The Ascription of Belief. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bernsen, N. 0., H. Dybkjaer, and L. Dybkjaer. 1997. `Elements of speech interaction: Proceedings of the 3rd Spoken Language Dialogue and Discourse Workshop (Vienna), 28-45. http: //www.mip.ou.dk/nis/publications/papers/hcm-paper/index.htm.

Bunt, H. 1989. `Information dialogues as communicative action in relation to partner modelling and information processing. In M. Taylor, E Neel, and D. Bouwhuis (eds.), The Structure of Multimodal Dialogue. Amsterdam: North Holland Publishing Company, 47-71.

-and W. Black (eds.). 2001. Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics. Amsterdam: Benjamins.

Carletta, J., N. Dahlback, N. Reithinger, and M. Walker. 1997. Standards for dialogue coding in natural language processing. Seminar No. 9706, Report No. 167, Schlof3 Dagstuhl, internationales Begegnungs- and Forschungszentrum fur Informatik.

Cohen, P. 1998. 'Dialogue modeling: In Cole et al. (1998).

-J. Morgan, and M. Pollack (eds.). 1990. Intentions in Communication. Cambridge, Mass.: MIT Press.

Cole, R., J. Mariani, H. Uszkoreit, A. Zaenen, and V. Zue (eds.). 1998. Survey of the State of the Art in Human Language Technology. Cambridge:

Cambridge University Press. On-line and postscript versions from: http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html.

Gibbon, D., R. Moore, and R. Winski. 1998. Handbook of Standards and Resources for Spoken Language Systems. Berlin: Mouton de Gruyter.

Grice, H.P.1975.`Logic and conversation' In P. Cole and J. L. Morgan (eds.), Syntax and Semantics, iii: Speech Acts. New York: Academic Press, 41-58.

Grosz, B. and C. Sidner. 1986.Attention, intentions, and the structure of discourse. Computational Linguistics, 12(3), 175-204.

Hirschberg, J. and C. J. Nakatani.1994• A corpus-based study of repair cues in spontaneous speech. Journal of the Acoustical Society of America, 3 (1995)1603-16.

Hovy, E. H. and D. R. Scott. 1996. Computational and Conversational Discourse: Burning Issues-An Interdisciplinary Approach. NATO ASI Series, Series F: Computer and System Sciences, 151. Berlin: Springer-Verlag.

Leech, G. 1983. Principles of Pragmatics. London: Longman.

-and J. Thomas. 199o.'Language, meaning and context: pragmatics: In N.E. Collinge (ed.), An Encyclopaedia of Language. London: Routledge.

M. Weisser, M. Grice, and A. Wilson. 2000. `Survey and guidelines for the representation and annotation of dialogues: In D. Gibbon, I. Mertins, and R. Moore (eds.), Handbook of Multimodal and Spoken Dialogue Systems. Resources, Terminology and Product Evaluation. Boston: Kluwer Academic Publishers, 1-1o1. First pub. 1998.

Levinson, S. 1983. Pragmatics. Cambridge: Cambridge University Press.

Litman, D. and J. Allen. 199o.`Discourse processing and commonsense plans: In Cohen, Morgan, and Pollack (1990).

McTear, M. F. 1999. Spoken dialogue technology: enabling the user interface. http://www. infj. ulst.ac.uk/-cbdg23/survey/spoken-dialogue-technology.html.

Martin, P., F. Crabbe, S. Adams, E. Baatz, and N. Yankelovich.1996. `Speech acts: a spoken language framework: IEEE Computer, July, 33-40.

Nakatani, C. J., B. J. Grosz, and J. Hirschberg. 1995. `Discourse structure in spoken language: studies on speech corpora. Working Notes of the AAAI-95 Spring Symposium in Palo Alto, CA, on Empirical Methods in Discourse Interpretation, 106-12.

Ono, T. and S. Thompson. 1993•`Interaction and syntax in the structure of conversational discourse: collaboration, overlap, and syntactic dissociation'. In Hovy and Scott (1996). First pub. 1993.

Passonneau, R. and D. Litman. 1996. `Empirical analysis of three dimensions of spoken discourse: segmentation, coherence, and linguistic devices: In Hovy and Scott (1996). First pub. 1993.

Sacks, H. 1967-72. Unpublished lecture notes. University of California.

Schegloff, E. 1996. `Issues of relevance for discourse analysis: contingency in action, interaction, and co-participant context: In Hovy and Scott (1996). First pub. 1993.

Searle, J. R. 1969. Speech Acts: An Essay in the Philosophy of Language. Cambridge: Cambridge University Press.

--1980. Expression and Meaning. Cambridge: Cambridge University Press.

Sperber, D. and D. Wilson. 1995. Relevance: Communication and Cognition, 2nd edn. Oxford: Blackwell.

Thomas, J. 1995. Meaning in Interaction: An Introduction to Pragmatics. London: Longman.

Weizenbaum, J. 1966. `ELIZA-a computer program for the study of natural language communication between man and machine: CACM, 9,36-43.

Worm, K. and C. Rupp. 1998. Towards robust understanding of speech by combination of partial analyses. VM-Report 222, Universitat des Saarlandes, Saarbriicken.