Invited paper:

Conversational Games, Belief Revision and Bayesian Networks

Stephen G. Pulman*

Abstract

The paper uses a simple and abstract characterization of dialogue in terms of mental state changes of dialogue participants to raise three fundamental questions for any theory of dialogue. It goes on to discuss currently popular accounts of dialogue with respect to these three questions. Next, the notion of 'conversational game' is revisited within a probabilistic and decision theoretic framework, and it is argued that such an interpretation is plausible both intuitively and as the basis for computational implementation. An illustrated sketch of a proposed implementation using Bayesian networks is described.

Three Questions for Dialogue

A simple, rather abstract description of a canonical dialogue is that it consists of a sequence of utterances with a corresponding sequence of mental states of the participants in the dialogue. Person A has a sequence of mental states $S_{A1} \ldots S_{An+1}$ and person B also has a sequence $S_{B1} \ldots S_{Bn+1}$. Connecting these two sequences is a third sequence, the sequence of utterances. U_{A1} is produced by A in state A1, U_{B2} is produced by B in B2 and so on. Furthermore, A's state S_{A2} and B's state S_{B2} are, at least partially, determined by the utterance U_{A1} which precedes them. The utterances change the mental states of the participants to the point where no further communication is regarded by them as necessary: the goals of the conversation, whatever they were, have been achieved as far as is possible. This is represented by the diagram in figure 1.

Even this simple picture reveals that there are several large questions to be answered in order to be in a position to build a machine capable of playing the part of A or B:

(i) what are mental states?

(ii) how do they change?

(iii) how do utterances connect with them and change them?

*SRI International Cambridge Computer Science Research Centre and University of Cambridge Computer Laboratory.

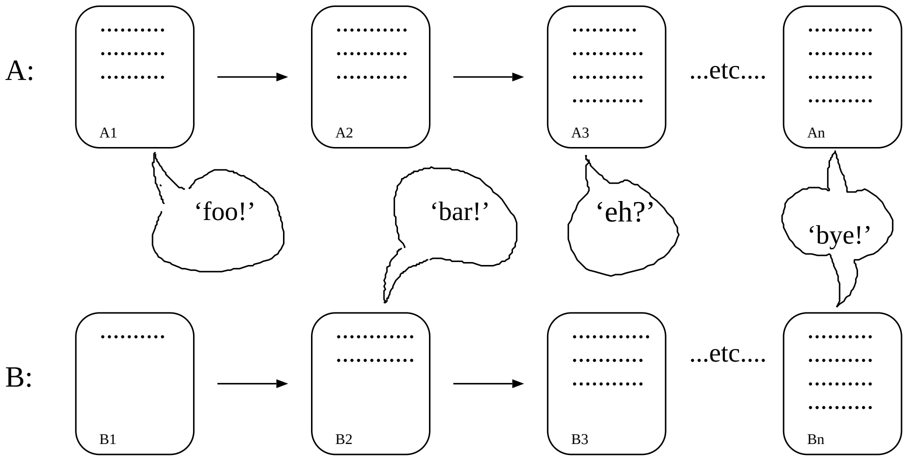

Figure 1: Two-person Dialogue

1 The BDI tradition

Insofar as the current literature on computational models of dialogue has a received wisdom on the answers to these questions, it is probably that given by the 'BDI' model of rational agency, as described for example in Cohen, Morgan, and Pollack (1990). The answer to the first question is that mental states are, or can be modelled as, sets of sentences in some logic, expressing the Beliefs, Desires, and Intentions of an agent (see Cohen and Levesque (1990)). Various axioms connect the having of desires and intentions with the performance of actions, some of which are linguistic actions. A rational agent, given an initial mental state, will reason as to the best course of action so as to fulfil the highest priority desires. Conversation proceeds via the performance of these linguistic actions. Part of the reasoning involves a model of the mental state of the other participants, and inferences about what their goals and intentions might be, based on the observed linguistic acts they carry out.

For a partial answer to the second question, how do mental states change, if mental states are modelled as sets of sentences in some logic, then it is appropriate to turn to the belief revision literature: e.g. Gärdenfors (1988), Galliers (1990). Belief revision is modelled via the addition or subtraction of propositions (if expressed on closures of belief bases, i.e. the deductive closure of some set of axioms) or of sentences (if expressed on belief bases themselves), operations which are required to preserve consistency. It is in the latter sentential form in which belief revision has to be implemented for the purposes of computational dialogue modelling, of course. A simple approach to belief revision within this framework would posit two basic operations, given a set of sentences Δ representing the existing mental state, and a sentence α, which is some component to be added or removed as the result of processing an utterance.

Subtraction:

If Δ does not entail α then Δ'=Δ;

Else, find some β in Δ such that $\Delta - \beta$ does not entail α,
and Δ' = $\Delta - \beta$

In many cases β and α will be the same, or α will follow directly from β perhaps in conjunction with some other sentences which taken alone do not entail α. Of course β may be a conjunction of several different sentences.

Addition:

If Δ does not entail $\neg\alpha$, then Δ' = $\Delta + \alpha$;

Else, find some β such that $\Delta - \beta$ does not entail $\neg\alpha$,
and $\Delta' = (\Delta - \beta)+\alpha$

It is worth noting that we need not just belief revision, but also revision of desires, and intentions. Conflicting goals and incompatible intentions are drivers of conversational processes just as much as detection of mismatches in beliefs. It is also worth pointing out that the mechanisms presupposed in belief revision, like detection of inconsistency or conflict, are required in some form for approaches that do not necessarily describe themselves as doing belief revision. Any approach to dialogue needs to be able to tell when an answer to a question is a plausible and appropriate one; when two goals cannot both be simultaneously achieved; or when some piece of information is implied by what is mutually known and therefore need not be explicitly repeated. Any formal mechanism that achieves this is addressing the problem of belief revision.

Let us turn now to the answer given to question (iii), how do utterances relate to, and change, mental states? The BDI answer to this question is essentially that derived from the speech act literature, as presented by Cohen and Perrault (1979) and Perrault and Allen (1980). Characterising an utterance as a particular type of speech act enables it to be related to properties of the speaker's mental state, by linguistic and other conventions governing that type of act. These conventions ('felicity conditions' in the original formulation) represent necessary and sufficient conditions for the performance of a genuine instance of a particular kind of speech act as in Searle (1969), and those conditions are at least in part conditions on the speaker's mental state, requiring the speaker to have the right kind of beliefs, desires, and intentions. Thus a hearer can make inferences about the speaker's mental state once an utterance has been recognised as instantiating a particular kind of speech act.

Given background axioms of 'rational agency' characterising the behaviour of an ideally cooperative rational hearer, the BDI approach also has an account of how an utterance can change the mental states of the participants in a dialogue. As an illustration of the general approach, a typical story about how a request can lead to a change of mental state and a consequent action on the part of a hearer will go something like this. We assume that the speech act conditions, and the rational agency axioms are characterised along roughly these lines:

```
Request Precondition:
IF Speaker wants A
   AND Speaker believes Hearer can do A
   AND ... etc.
THEN Speaker requests Hearer to do A

Request Postcondition:
IF Speaker requests Hearer to do A
   AND ... etc.
THEN Hearer believes Speaker wants A
```

Axioms of 'rational behaviour':

```
Cooperativity:
  IF Hearer believes Speaker wants A
     AND ... etc.
  THEN Hearer wants A

Desire leads to Action:
  IF X wants A
     AND X can do A
  THEN X does A
```

Now a typical piece of reasoning that could lead a Speaker to make a request in order to achieve some desire might proceed as follows:

Speaker requests Hearer to do A	
∴ Hearer believes Speaker wants A	(Request Postcondition)
∴ Hearer wants to do A	(Cooperativity)
∴ Hearer does A	(Desire leads to Action)

That is, the Speaker desires that A be done and he reasons that by issuing a request he will start the above chain of events that results in A being done. This reasoning is typically done by backward chaining from the goal state, but that is really an implementational issue that does not affect the logic.

1.1 Some problems for the BDI tradition

The BDI tradition has led to many theoretical insights into the nature and functioning of dialogue, and there are several very impressive implemented systems based on versions of the approach: for example, those described by Allen, Miller, Ringger, and Sikorski (1996) or Sadek, Ferrieux, Cozannet, Bretier, Panaget, and Simonin (1996). Nevertheless there are several areas where the theoretical content is unclear or questionable, and there are many aspects of the theory which do not seem likely to yield satisfactory large scale computational implementations. We turn now to discussion of some of these problems.

The basic propositional attitudes countenanced by the BDI tradition are those from which it derives its acronym: belief, desire, and intention. However, since the earliest formal work on dialogue it has been recognised that many of the propositions that correspond to utterances in a dialogue do not fall easily into these three categories. Hamblin (1971) pointed out (p 36ff) that many sentences correspond to propositions that are not (yet, anyway) believed by the participants. He introduces the notion of a commitment, which is not necessarily a belief (though it may become one) but a function purely of what has been said. Speakers are generally committed to a statement if they make it, or agree to one made by someone else, or if it clearly follows from something else to which they are committed. In particular commitments may be later *retracted* but not *denied*.

In the recent literature other closely related terms have been used. Traum uses the term 'proposal' in Traum and Hinkelman (1992) and the idea of propositions that are being 'grounded' but not yet agreed appears in Clark and Schaefer (1989). For example, in the following dialogue (from the 'Autoroute' corpus described by Moore and Browning (1992)), between a 'wizard' pretending to be a route-planning system, and a caller, the proposition 'caller wants to go to Edwinstowe' cannot be said to be a belief of the wizard until at least step 4, rather than step 2, where the proposition is 'in the air'. (We assume throughout that what is happening is that at step 3 the wizard is not sure she has heard correctly. At step 6 the system she is operating has reported that there is more than one Edwinstowe).

```
1. w: Where would you like to go?
2. c: Edwinstowe
3. w: Edwinstowe?
4. c: Yes
5. w: Please wait
6. w: Is that Edwinstowe in Nottingham?
7. c: Yes
```

More recently, several authors, like Traum and Allen (1994) and Bunt (1997), have pointed to the need to also recognise a category of 'obligations' or 'social commitments' which arise from linguistic and social conventions. If someone asks you a question, you are, as a reasonable member of the same language community, thereby placed under some kind of obligation to respond.

Many other types of phenomena that are encountered in real dialogues seem to resist an easy classification into one of the three propositional attitudes countenanced by the approach. These include what Bunt calls 'dialogue control' phenomena: utterances (feedback, acknowledgements, pause-fillers, etc.) whose function is to maintain the dialogue and coordinate the participants, rather than to directly express beliefs, desires, or intentions.

These observations do not threaten the central role of beliefs, desires and intentions, of course, but they do indicate that as an empirically adequate account of what actually goes on in dialogues the BDI approach needs considerable supplementation and extension. The notion of 'mental state' provided by the theory is too simple to explain everything that happens in a natural dialogue.

Let us turn now to the question of change of mental state, and the belief revision framework assumed implicitly or explicitly by BDI approaches.

The classical belief revision framework (and associated approaches such as dynamic logic: Groenendijk and Stokhof (1991), Jaspars (1996)), while giving a clear logical theory of change of information state, present many problems when large scale practical implementations are contemplated. As is well known, a simple method of belief revision like that sketched above is very highly non-deterministic. Even given such simple choices for the existing set of beliefs Δ and a candidate for addition or subtraction α as:

$$\Delta = \{a, a \rightarrow b\}, \alpha = b \text{ (Subtraction) or } \neg b \text{ (Addition)}$$

there will be a choice about which β to remove. Practical belief revision requires us to assume some priority ordering on sentences in a belief base, such that given several candidates for elimination, the one which is 'cheapest' in terms of some overall score will be given up. This priority ordering usually corresponds to an intuitive notion like 'strength of belief' or 'degree of commitment'. Deciding on the adjustment that makes the least overall change required to preserve consistency can be a computationally intensive operation. Note that any such system of weighting is not part of the logic itself and so some separate mechanism is required to make sure that the weighting scheme itself observes reasonable properties.

Implementing classical belief revision of course requires us to be able to detect inconsistency, and thus some kind of classical negation is necessary in our logics. It would be impossible to do belief revision on sets of pure Horn clauses, for example. But this means that we have problems, not just with efficiency, but also with 'logical omniscience'. If the logic is strong enough to detect inconsistencies between complex beliefs, it is likely also to make the contents of a belief state imply logical consequences of basic beliefs that are actually beyond human ability to compute.

For both of these reasons it is desirable for an implementation also to model something like 'focus of attention' or 'salience' of sentences in the mental state, so that reasoning can be restricted to relevant subsets of sentences, and conclusions can be limited to those that are humanly processable. However, all the obvious ways of achieving this notion (e.g. limiting chains of inference to a certain depth) compromise completeness and (global) consistency, as discussed in Konolige (1986). Since these are not properties that characterise human reasoning, especially in dialogue, this may actually turn out to be an advantage to us, but nevertheless it is not easy to see how to achieve exactly the right kind of restrictions without unwanted negative effects.

Lastly, but by no means least, there is the fact that the classical approach to belief revision requires us to axiomatise the relevant properties of the domain in order to be able to track what follows from what. As anyone who has ever tried to carry out such an exercise in knowledge representation will confirm, this is an exceedingly difficult undertaking, especially when classical first order logic is the representation language. It soon becomes obvious why all the textbook examples are simple blocks worlds, or equally well structured and clean domains. Anything else is generally just too messy and hard, and the resulting axiom set is always very fragile and incomplete in its coverage.

Turning now to the third of our questions, how to connect utterances with mental state, we also find problems with the logical reconstruction of speech act theory that is needed within the BDI framework. For example, many people, not least the original proponents of the theory, have commented on the implausibility of the 'Cooperativity' axiom (and its analogues for other speech acts). There are actually two problems: firstly, it does not allow for the case where the hearer might not want to cooperate, or where external circumstances may bring about conflicting goals if he does: see Galliers (1990). It can also be the case that a hearer might be cooperative in some respects but not others. To some extent this can be alleviated by introducing some notion of defaults (although how to square this with the requirements of classical belief revision is not obvious).

Secondly, and more seriously for the interpretation of the BDI account as a contribution to a theory of *dialogue*, is the fact that these axioms are, in the theory, the only way of achieving 'uptake' of a speech act; that is, of creating a link between an utterance by a speaker, and subsequent modification of the hearer's beliefs or intentions concerning anything other than the speaker's mental states. In many respects, the original speech act theory is rather solipsistic or one-sided: it deals with the conditions for the successful performance of some act by a speaker, but has virtually nothing to say about what happens next, or in fact about anything outside the speaker's head. For example, as far as speech act theory proper is concerned, it is largely unexplained why a request is typically met either with an acceptance or a refusal, or why a question is typically met with an answer rather than (say) a request or another question. In the speech act literature, and in the BDI tradition derived from it, there are no dialogue units larger than a single utterance: a response to a request, or an answer to a question, cannot within the theory be distinguished from a conversation-initiating declarative.

Also completely unexplained, even with the appropriate axioms in place, is why there is a pressure on a hearer to respond somehow to an utterance even if he is not in a position to respond appropriately to it. Requests which are not going to be complied with are still acknowledged; questions that cannot or will not be answered still evoke some kind of explanation or diversion. Complete silence is not an option, although it is not easy to see how that option would conflict with anything in speech act theory.

2 Responses

There have been broadly two types of response to this problem. (Actually, only one is a direct response; the other is more of a parallel development that can also be seen as offering a solution). Traum and Allen (1994) propose the addition of a new mechanism to a speech act-based approach, namely 'discourse obligations'. A discourse obligation is a linguistically based social convention having the effect that when a particular speech act is recognised by a hearer, the hearer incurs an obligation to respond in an appropriate way:

Speech Event	Discourse Obligation
S request A	H accept or reject A
S ask whether P	H say whether or not P
Utterance failure	H repair utterance
etc.	

Thus we now have what might be called a BDIO model: a new propositional attitude is added. However, the notion of a 'speech event' is now much wider than that of a speech act: although the latter, in their original formulation at least, e.g. in Searle (1969), included some acts that one might think of as dialogue control acts rather than as BDI related. Nevertheless, even in the most ambitious formulations, there was no speech act of utterance failure.

However, while this formulation begins to describe the conventional association between questions and replies, requests and acknowledgements, and so on, it does not fully capture the nature of the more general social pressure to respond that is characteristic of normal dialogues. For example, in cases where a politician is asked an awkward question in an interview, he will usually fail to obey the specific question-related discourse obligation described above, but he cannot just remain silent. What he will typically do is talk about something else that he hopes will be taken as a relevant response, but which does not actually constitute an answer. It seems plausible that there are at least two types of obligation involved in discourse: those which are associated with particular speech acts or utterance types (e.g. that a yes/no question demands the answer yes or no), as described by Traum and Allen, and those which are more general social and communicative obligations, not specific to particular constructs, and concerned with the maintenance of communication norms.

The second line of work which can be seen as addressing this particular defect of speech act theory is the 'Conversational Games' tradition: Power (1979), Houghton (1986), Kowtko, Isard, and Doherty (1992), Reithinger and Maier (1996). More of a descriptive framework than a theory, this tradition posits a set of 'conversational games' or 'dialogue games' each consisting of a set of moves, where an utterance may realise one or more moves. The important thing is that the games encompass both partners in dialogue: for example, a yes/no game consists of a yes/no question along with its yes/no reply. Thus the conventional link between utterance type and response type is achieved by making the unit of discourse something that by definition is not restricted to a single utterance. This may not be a very sophisticated theoretical innovation, but it at least describes the facts correctly.

Some conversational games postulated by Kowtko, Isard, and Doherty (1992) on the basis of study of the Edinburgh 'map task' corpus are: Instruction, Confirmation, Question-YN, Question-WH, Explanation, Alignment. The moves can be broken into two categories:

Initiating Moves:

```
Instruct      (provides instruction)
Check         (elicits confirmation of known information)
Query-yn      (asks yes-no question for unknown information)
Query-wh      (asks wh-question for unknown information)
Explain       (Gives unelicited description)
Align         (Checks alignment of position in task)
```

Response and feedback moves:

```
Clarify       (clarifies or rephrases given information)
Reply-y       (responds affirmatively)
Reply-n       (negatively)
Reply-wh      (Respond with requested information)
```

`Acknowledge   (acknowledge and request continuation)`
`Ready         (Indicates intention to begin a new game)`

In principle the framework of conversational games can easily cover those utterance types that do not fit happily into a pure speech act framework, recognising that the function of some of these is to provide information about the state of the dialogue (e.g. alignment - making sure both partners know where they are in the dialogue) and to increase the degree of confirmation about some piece of information.

A more refined characterisation of these 'dialogue control' acts is given by Bunt (1997). He distinguishes different aspects of context: semantic, cognitive, physical, social, and linguistic, with different types of dialogue act for each. Dialogue acts are acts which change one or more aspects of the context.

Conversational Games are a useful descriptive framework. But as a theoretical contribution to the understanding of dialogue they have remained somewhat weak. Firstly, it is not clear how they differ from the BDI framework in the way they try to establish a link between utterances and mental states. From the perspective of speech act theorists, conversational games look like a hard-wiring of some of the patterns of inference that they derive from first principles.

Secondly, the theory seems very unconstrained. For example, is there a satisfactory answer to the questions of how many games there are, how they vary according to the type of dialogue, and what constraints there are upon possible games? These are the kinds of question routinely asked of every other level of linguistic formalism. For example, in the Verbmobil system as described in Reithinger and Maier (1996), games of much finer level of detail than in the Map Task are envisaged: e.g. 'arranging a time', or 'confirming a date'. These are justified along exactly the same lines as those developed for the Map task, namely, intuitive agreement that a certain level of commonality exists between different utterance/context pairs. But it is clearly not very much further down this route before there is a distinct game for practically every utterance. We would therefore like some theoretical grounding to establish what granularity is characteristic of useful games.

To illustrate these issues, consider the question: What distinguishes a move from a game? One cannot simply identify games with standardised sequences of moves, although this is at first sight a tempting idea (and explicitly proposed in Houghton (1986)). For example, one might think that a WH query game should consist of a WH-query move followed by a WH-reply move. But if this were the case then we would have to say that the WH query game in the dialogue fragment we saw earlier would be over after turn 2, whereas intuitively one would want to say that it was only completed after the two checking games.

		Move	**Game**
1. w: `Where would you like to go?`		query whq	WH
2. c: `Edwinstowe`		reply whq	
3. w: `Edwinstowe?`		check	CHK
4. c: `Yes`		clarify	
5. w: `Please wait`	(time management)	align/acknowledge	
6. w: `Is that Edwinstowe in Nottingham?`		query ynq/check?	CHK
7. c: `Yes`		reply yes/clarify	

Ian Lewin has suggested (p.c.) that in general we should single out those (sequences of) dialogue acts that serve to change the status of propositions currently under discussion from 'proposals' to 'agreed commitments'. The boundaries marked by these transitions do seem, at least in these types of dialogue, to correspond to natural divisions in a dialogue. Thus although there is a context change between each of the utterances above, and there are three games played, there is only one significant change to the agreed commitments of the participants. Utterances 4-7 serve to check and ground the information introduced by 1 and 2, and so although they do change the linguistic and other aspects of the context, there is a good sense in which they have a different status. Lewin points out that it is plausible, for example, that to the extent that dialogues are consciously or unconsciously planned, the units of planning are those represented by the acquisition of agreed propositions rather than the units that correspond to the conversational games like 'checking' or 'acknowledgement'. It is not plausible to assume that such moves are planned: rather, they arise as an immediate response to the current state of the dialogue.

3 Conversational Games Reconstrued

Let us reconsider what a notion of conversational game might tell us about the answers to the three questions with which we began our investigation. In particular, we will explore a somewhat different, and in some ways more traditional, interpretation of the notion of a 'game'.

We consider (task-oriented) dialogues to be a kind of game whose goal is to achieve the purposes of the dialogue (e.g. booking an airline ticket, planning a car journey) usually as quickly and economically as possible. A suitable example game to explain the analogy might be a card game like bridge. The players are in the position that a certain amount of information about the hand that the other player has is overtly available via the content of utterances, but the rest has to be inferred on the basis of bid behaviour and knowledge about cards. Some good reasoning or lucky guesses may lead to a speedy conclusion of the game. But a bad guess might put one at a disadvantage. So each move has to be made with an eye to its possible positive or negative effects. In formal decision theory, the effects are of course called 'utilities', and each move is calculated (if the player is rational) to maximise utilities. Moves are seldom made simply in response to the previous move by the opponent (although sometimes this is necessary, as when to move a king out of check) but are more often part of a longer range strategy.

Pursuing the analogy at a more detailed level, then, our conversational game framework requires at least the following components:

(i) move interpretation: when a player puts down some cards, we use that information to work out what other cards the player may have, or may want. The conversational game analogue of this to the classification of an utterance as a realisation of one or more conversational moves. Classifying an utterance as a move is making one hypothesis about the speaker's mental state. Equally, one may make further hypotheses about what it is reasonable to think led to that particular move being made.

(ii) tactics: planning the next move in response both to the immediate situation but also the longer range strategy. In some cases the immediate situation may be the most important factor, as when moving a piece to avoid capture, or requesting clarification when an utterance has not been recognised with sufficient confidence, or when it presents the belief revision component with an apparent contradiction. But if things are going according to plan, the next move is both an appropriate response to the previous one, and a step forward in the overall plan.

(iii) strategy: planning the next game or sequence of games to be played in order to win the dialogue. Strategy needs to be continually re-evaluated as new information is obtained.

(iv) a fourth but vital component is that knowledge of the domain which supports the various types of reasoning in (i-iii).

As the computational underpinning of all four components we intend to explore the use of Bayesian Networks, as developed by Pearl (1988), and described in Neapolitan (1990), a formalism which is becoming widely used in the AI community for knowledge representation, causal reasoning, belief revision, and decision theoretic reasoning. There is little doubt that, modulo some important provisos below, this formalism can provide a plausible platform for component (iv) and so in the remainder of this section we concentrate on (i)-(iii).

4 What is a Bayesian Network?

Given a probability space of events, E, a 'propositional variable' is a function from E to a finite subset of E of mutually exclusive and exhaustive events. Given a propositional variable A, let $a_1 ... a_n$ be the set of possible values of A. We write $P(A = a_i)$ as $P(a_i)$ and an expression like $P(A \mid B) = P(A)$ is a shorthand for the expressions $P(a_i \mid b_j) = P(a_i)$ for all i and j. Given a set of propositional variables A, B, C, ... we can define a joint probability distribution on them such that:

$$\sum_{ijk...} P(a_i, b_j, c_k, ...) = 1$$

Given a set of such variables, $\{X_1 ... X_n\}$, the 'marginal probability' of any subset of them, say $X_i, ..., X_j$, relative to this joint probability distribution is defined as:

$$P(X_i, ..., X_j) = \sum_{k \neq i...j} P(X_1 ... X_n)$$

A Bayesian or causal network is a set of propositional variables, associated with vertices in a directed acyclic graph, where there are conditional (in)dependencies between some of the variables, as reflected pictorially in the associated graph. More formally, a DAG consisting of vertices/variables V and edges E, with an associated joint distribution P, constitutes a Bayesian network under conditions below.

First, given a variable v which is a member of V, let *c(v)* ('causes of v') be the set of v's parents, let *d(v)* be the set of v's descendants, and let *a(v)* be $V - (d(v) \bigcup v)$, that is, all the variables except v and v's descendants.

Let W be any subset of *a(v)*. W and v are conditionally independent given *c(v)*, where $P(c(v)) \neq 0$, under three conditions:

if $P(v \mid c(v)) = 0$
- because nothing further (in particular W) can affect v

if $P(W \mid c(v)) = 0$
- because nothing further (in particular v) can affect W

if $P(v \mid W \bigcup c(v)) = P(v \mid c(v))$
- which can be verified by calculation

If every subset of W is conditionally independent of v given $c(v)$ then the DAG is a Bayesian network.

The thing to notice about conditional independency is that although some independencies will be permanent because of the configuration of the network, and will not be affected by instantiation of variables (i.e. when it is known which value of the propositional variable = 1), some variables will become independent of each other only when some intervening variable has been instantiated.

The conditional independencies in a network can be exploited to reduce the amount of computation involved in working out joint probabilities. Take for example, a network of the form

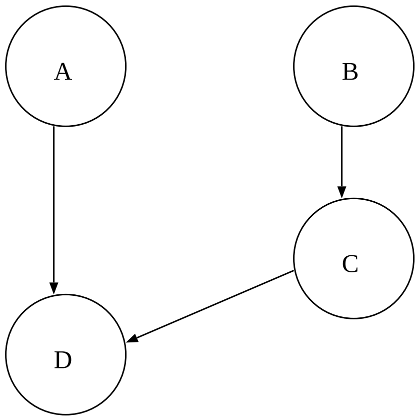

Figure 2: bayesian net

The 'chain rule' of probability theory tells us that the joint probability can be calculated from conditional probabilities thus:

$$P(A, B, C, D) = P(A \mid B, C, D) \times P(B \mid C, D) \times P(C \mid D) \times P(D)$$

In order to fully exploit this equivalence we must reorder the variables so as to reflect the structure of the DAG.

$$P(D, C, A, B) = P(D \mid A, C, B) \times P(C \mid A, B) \times P(A \mid B) \times P(B)$$

Now we can use the structure of the DAG to determine the conditional independencies: A and B have no parents so they are not dependent on any other variable: thus $P(A \mid B) = P(A)$. Variable C is dependent only on the value of B so $P(C \mid A, B) = P(C \mid B)$. Variable D is dependent only on the value of A and of C (since any effect of B has to be via C) and so $P(D \mid C, A, B) = P(D \mid C, A)$.

$$P(D, C, A, B) = P(D \mid A, C) \times P(C \mid B) \times P(A) \times P(B)$$

For larger networks, this simplification avoids many unnecessary computations.

Now the general version of the chain rule for Bayesian networks can be written:

$$\prod_i P(v_i \mid c(v_i)) where P(c(v_i)) \neq 0$$

This enables us to compute the joint distribution from the conditional probabilities. We can also compute the conditional probabilities given the joint distribution, using the chain rule in the other direction:

$$P(A \mid B, C, D) = \frac{P(A, B, C, D)}{P(B, C, D)}$$

and so on.

When a variable (usually one with no parents or no children) is instantiated, i.e. when we know which of its values is the observed one, the probabilities in the network have to be updated. This is done by propagation from the instantiated variable. The probability of each variable V can be calculated by combining the evidence for V from the nodes above it in the network, and those from below: let the evidence from the parents of V be E_c and the evidence from the daughters of V be E_d. Then:

$$P(V \mid E_c, E_d) = \frac{P(E_d \mid V) \times P(V \mid E_c)}{\alpha}$$

where α is a normalising constant.

The precise algorithm for computing these quantities and for propagating their effects throughout the relevant portions of the network is very complex, since a node may have many parents and many children and allowance has to be made for the mutual effect of new information on any of these. The algorithm assumes that networks are 'singly connected' i.e. that for any pair of nodes there is only one path that can be found between them (ignoring directions on arcs). This is not a limitation in principle because any multiply connected graph can be transformed to a singly connected one, although at some resulting computational cost.

The original version of the algorithm can be found in Pearl (1988); a very detailed tutorial description can be found in Neapolitan (1990); and a simplified version restricted to tree-shaped networks is given in Shoham (1994).

One serious restriction that Bayesian networks impose is that the variables are propositional: i.e. they have only a finite number of atomic values. This means, in effect, that quantificational reasoning, or reasoning that depends on the internal structure of propositions, is not directly possibly. However, the networks can be very large: many applications use networks of tens of thousands of nodes each with a large number of values, and so for many practical purposes this restriction does not begin to bite. Propositions of the form 'pred(A,B)' can be modelled as a node 'pred' with values in the product A $\times$ B of all relevant A and B values. Implementational devices can be used to keep this kind of thing manageable. An alternative is to generalise a potentially infinite number of propositions to a 'proposition type' which stands for all of them, if the differences between tokens are not important.

5 Bayesian networks for move recognition

When a hearer categorises an utterance as realising a conversational move there are presumably several factors taken into account in making this decision. Firstly, the linguistic form and content of the utterance is important: for example, it is very unusual for an utterance of 'no' to be interpreted as realising a 'reply-y' move, (although possible if enough intonational cues are given to signal a non-literal interpretation). Secondly, the previous few moves, or perhaps the recognition of the game currently being played has to play an important part. (In some approaches it is the only factor taken into account: see Reithinger and Maier (1996)). Thus an utterance of 'OK' might be interpreted as a 'reply-y' move if the previous move was a 'query-yn', but if the previous game has been seen to be completed it is more likely to be a 'ready (for a new game)' move. Thirdly, knowledge about the speaker's mental state is relevant: if the hearer knows that the speaker should know, or has at least been told, P, then an utterance which looks superficially like a 'query-wh' or 'query-yn' move is probably more likely to be a 'check'. If it is categorised as a check then that hypothesis in turn would weaken the likelihood that the speaker is certain of P: you don't check things you are certain of.

We can illustrate this with an example from the 'Autoroute' domain described in Lewin, Russell, Carter, Browning, Ponting, and Pulman (1993) and Lewin and Pulman (1995). In this domain a person interacts with a system to plan an automobile route between places within the UK. The relevant parameters are start and end of journey, with optional information like type of car, stops on the way, whether to optimise for speed or distance, avoid or follow motorways etc. The system engages in a dialogue to instantiate as many of these parameters as possible and then sends the information to a commercial PC package (described in (NextBase 1991)) which calculates the optimal route.

We can encode the observations described earlier into a network representing the influence of these factors on the recognition of conversational moves. Assume that there are only a finite number of types of proposition $P_1...P_n$ which can arise in our domain (which will usually be the case for the kind of simple task-oriented dialogues we are considering, even though there may be an infinite number of ways of expressing them). They will be simplified representations of the propositional content of actual utterances, for example:

destination=cambridge; no; ok; origin=what; etc.

We will assume the variables and values described below (these are just for illustration: in reality, determining the precise form of the network can only be done in conjunction with a close analysis of the corpus dialogues).

Pr: Previous-move = query-yn(P_i),reply-n,....

C: Content and form = positive,negative,ynq(P_i),whq(P_i),dcl(P_i)

K: S-knows-P = yes,no,maybe

M: Current-move = query-yn(P_i),reply-n,....

We also want the results of a particular move classification to feed into an updated model of the speaker's current beliefs. This can be achieved by using the move categorisation network to provide evidence that instantiates a value of a variable in another network. We indicate this in the diagram below by a dotted line connection between node K and an independent subnetwork representing the hearer's beliefs about the speaker's beliefs.

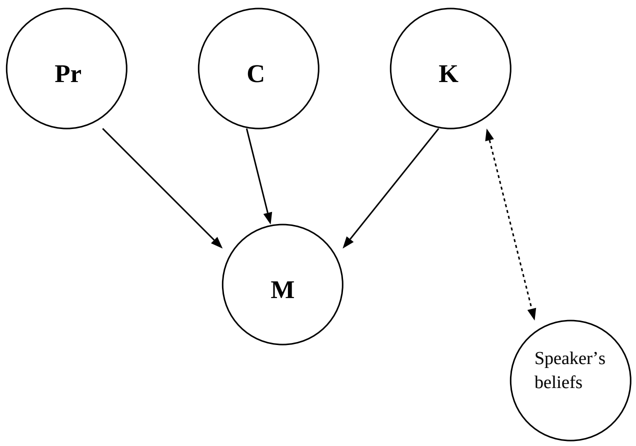

Figure 3: Bayesian net for move recognition

Having decided on the structure of the network, we need to assign *a priori* probabilities to the various values of the variables. This should be done on the basis of statistics derived from annotated corpora, although in many applications estimates of probabilities derived from experts have proved to be quite accurate. In our example, we will assume that the top three nodes may have a fairly uniform initial *a priori* distribution on them, reflecting the fact that in the absence of any evidence, there is no previous move more likely than any other, no proposition more salient than any other, and no hypothesis about the other's beliefs more detailed than any other. However, we can provide some conditional probabilities which express *a priori* dependencies between particular values of M and its parents, e.g.

```
P(M=query-yn(Q)|C=ynq(Q),K=yes,...)          = very low
```

The probability that a yes-no question about Q realises a 'query-yn' move when the other is believed to already know the answer to Q is very low. You don't ask questions about things you already know. (Notice that there is a clear connection here with the notion of preconditions for speech acts. The analogous link between preconditions and moves is reflected in the assignment of probabilities).

```
P(M=query-yn(Q)|C=ynq(Q),K=no,...)           = high
```

A yes-no question is more likely to realise a query move if the speaker is believed not to know the answer already.

```
P(M=check(Q)|C=ynq(Q),K=no,...)                 = low
P(M=check(Q)|C=ynq(Q),K=maybe,...)              = a bit higher
P(M=check(Q)|C=ynq(Q),K=maybe,Pr=reply-wh(R))   = pretty high
... etc.
```

A yes-no question is most likely to be expressing a checking move if the speaker may not know the answer and the previous move was a reply to a question.

```
P(M=ready|Pr=query-yn,C=ok,...) = very low
```

The probability that 'yes', or 'ok' realises a ready move when the previous move was a query is rather low.

```
P(M=ready|Pr=reply-n,C=ok,...) = quite high
```

The probability that 'yes', or 'ok' realises a ready move when the previous move was one which can close a game is quite high.

On the assumption that we have a complete and plausible set of probabilities like this we can give a hypothetical illustration of how such a network might be used in the first few turns of our illustrative sequence.

The basic cycle (from the point of view of one person, here the user) is:

1. instantiate C (and Pr and K - from the speaker's belief network- if possible), update probabilities.

2 find the value of 'move' that maximises $P(M = move \mid Pr, C, K)$

3. instantiate M to 'move', propagate revised probabilities.

4. find value of 'k' that maximises $P(K = k \mid Pr, C, M)$, and feed into speaker's beliefs sub-network.

5. re-initialise main network, and go to 1.

Of course, we also need to provide for the user making their own move, and updating other networks as well.

We can illustrate with our earlier example dialogue:

1. w: Where would you like to go?

We assume for illustration that this is the opening move and so Pr is not instantiated. C is instantiated as 'whq(destination)'. K is not yet instantiated. We will further assume that given the *a priori* probabilities the most likely move for for a wh-question under these circumstances is a wh-query, and so that is the answer at step 2. We now set the value of M to 'wh-query', and propagate the resulting probability changes. Given our estimated conditional probabilities and the new instantiated nodes the value for k that maximises $P(K = k \mid Pr, C = whq, M = wh - query)$ will be 'no', and so the proposition that, at that stage in the dialogue, the speaker does not know the destination is added to the record of beliefs built up in the subnetwork.

The user then plans and executes his own move (exactly how this is done we will return to below):

2. c: Edwinstowe (reply-whq)

Back comes the reply:

3. w: Edwinstowe?

This time round the cycle, Pr='reply-whq', C=ynq(destination=Edwinstowe), and K is 'yes' for the proposition 'destination=Edwinstowe'. This latter value we assume to be a consequence of the user's previous reply: normally you would expect someone to know something they have just been told. This information can be recovered from the 'speaker's belief' subnetwork.

We assume that given the probabilities above, the most likely move assignment for this yes-no question is as a 'check' rather than a genuine question. So we now instantiate M for this value. Recalculating probabilities the most probable value for K with respect to these instantiations should now be 'maybe' rather than 'yes' and this can be used to update the record of speaker beliefs being built up.

5.1 Choosing the next move

Bayesian networks can be extended so as to represent information not only about probabilities, but also utilities attached to the consequences of particular actions. This enables the integration of reasoning about the probability of an effect along with the desirability of that effect. Utilities can be combined and propagated by essentially the same algorithm as is used for probabilities (Neapolitan (1990) esp. Chapter 9).

Bayesian networks extended in this way are usually referred to as 'causal influence diagrams'. To the set of nodes representing propositional variables we add one or more 'decision' nodes, representing a choice about whether or not to perform an action, and exactly one 'value' node, where all the utilities associated with the different actions are represented. If there is more than one decision node, later ones must be dependent on earlier ones.

As an illustration, we will take a network representing the decision whether to start a new game, or to check the previous move. There are consequences associated with these choices, and also there are several factors which we want to influence the choice that is made. The consequences of choosing to check will typically be that the overall dialogue will take longer. However, there is a lesser risk of a misunderstanding or an error causing problems later on. Going on to a new game will typically speed up the dialogue, but if the previous piece of information has not been properly 'grounded' then it may turn out to be insufficient to proceed at some later stage. For example, in our illustration, the wizard might have got the wrong 'Edwinstowe', leading to either an inaccurate route, or a later repeat of most of the dialogue. If speed is important, it might be preferable to move to a new game as soon as possible provided there is reasonable confidence that understanding has been achieved, whereas if accuracy was preferred to speed, frequent checking moves and a more cautious dialogue style would be called for.

The decisions have consequences, but the consequences might also be dependent on other causal factors. For example, if the environment is a noisy one, or the speech recogniser is unreliable, it may be that frequent checking will lead to a better accuracy/speed ratio than a less cautious strategy. Thus a decision will be a calculation based on the likelihood of the effects given the prevailing circumstances, and the utilities associated with those different possible outcomes.

We can illustrate this with the following partial network for deciding which move to make next. In this network we have to choose to check the last move, or start a new game. The causal consequences of this decision are represented by a 'speed' node, saying whether the dialogue is likely to be completed quickly or not, and an 'accurate' node, saying how likely it is that the route given is actually the one asked for.

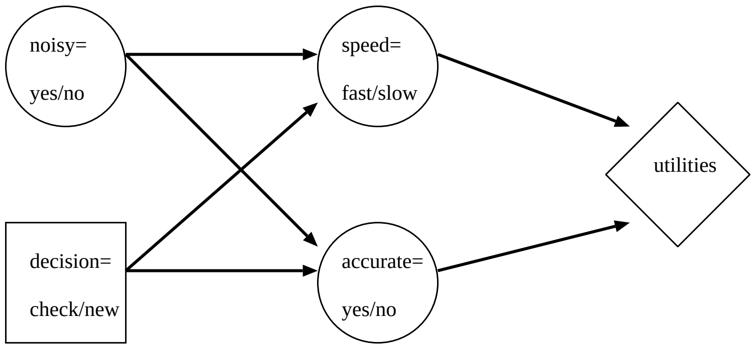

Figure 4: Bayesian Net for Move Choice

The round nodes are propositional variables, as before. (In the 'causal influence diagram' literature they are referred to as 'chance' nodes). The square one is a decision node, representing the choice of possible actions. Although in our example this is not the case, chance nodes can have arcs to decision nodes.

The value node is represented as a diamond. The value node can be regarded as a propositional variable which contains one utility value for each possible combination of its parent nodes. These utilities can be computed from those assigned to the parents, or assigned directly. (The fact that there is only one such node makes the DAG look as if is no longer singly connected. But provided that the subgraph consisting of the chance nodes remains singly connected that does not matter, since the value node does not affect any of the probabilities).

We assign probabilities to chance nodes as before. The probabilities depend both on parent chance nodes, and on whether a particular decision is taken. Thus, for example, the assignment of probabilities to the 'speed' node will depend on what decision was taken, and on how reliable the communication channel is. The precise values we assign to these probabilities do not matter for the sake of the illustration, but we would want the probabilities and the utilities to obey the following constraints:

`P(fast|check,noisy) > P(fast|check,~noisy)`

A check is more likely to lead to an overall speedup in a noisy environment.

`P(accurate|check,noisy) > P(accurate|newgame,noisy)` A check is more likely to maximise accuracy in a noisy environment than moving on to a new game.

`U(fast,accurate) > .... > U(slow,~accurate)` We prefer fast, accurate dialogues. Slow inaccurate ones are of course the worst of all worlds.

Given a network like this with utilities and probabilities assigned, we can calculate for any action its expected utility with respect to the current instantiations of chance nodes. Let a be an action (i.e. a value of the decision node), and let the the current instantiations of chance nodes be represented by e (=evidence). The value node will describe the utility of the causal effects $\mathrm{C}_{1\ldots n}$ of each action, where the notation for such a utility measure is $u(\mathrm{c}_i)$:

$$U(a) = \prod_i u(c_i) \times P(c_i \mid a, e)$$

Now we choose the action that has the maximum overall utility under the circumstances, execute the conversational move corresponding to it, and update variables, etc. We would hope that in our example scenario, given the circumstance that the reliability of the communication channel is low, and that the utility that is to be maximised is accuracy, then the decision that would score the highest would be to do a checking move rather than begin a new game.

5.2 Higher level planning

So far we have seen how it is possible to recognise utterances as realising particular conversational moves, and how to select the maximally useful next move, while updating and combining information of several different sorts. Using Bayesian networks, augmented with utility calculations, offers the promise of being able to model locally rational conversational behaviour in a way that has not so far proved possible in practice on a large scale for the traditional BDI-based systems. However, while a system based upon the components we have sketched so far would be a satisfactory 'reactive' system, we have not yet shown how to reproduce the higher levels of strategic planning that are one of the strong points of the traditional architectures.

However, at least as far as relatively simple task-oriented dialogues of the Autoroute, Verbmobil, or ATIS types are concerned, it seems quite possible to extend this scheme to completely replace the traditional types of planning that most dialogue systems rely on for their overall strategy. The analogy here is with the use of decision networks in expert systems, particularly medical diagnosis systems. Here the diagnosis does not necessarily proceed by going through some fixed sequence of questions; rather, the most informative next question is chosen dynamically by testing to see which propositional variable it would be most useful to know the value of. For example, knowing the age of a patient is a very important piece of information, even if not directly relevant to a diagnosis. If the patient is a child, questions about level of alcohol intake are unlikely, even in these times, to yield much diagnostically relevant information. Thus the utility of asking a question about age may be high in terms of speedy diagnosis, even though the answer itself may not be directly relevant.

In the case of our Autoroute domain, we might have variables corresponding to the main parameters of an Autoroute query: start, destination, car type, etc. It is difficult to think of an assignment of utilities that is not rather trivial: for example, we clearly need to know the start and the destination, and so the utilities associated with those variables should be higher than those of e.g. car type. Also, of course, the utility of asking questions about the values of variables that are already instantiated is likely to be very low. However, we might complicate the picture by making the utility of some variables dependent on the values of others: for example, if we know that the user has a fast car, then it is probably less important to ask whether he is interested in a scenic route for his journey. If the user wants to avoid motorways, then he is probably not interested in the fastest as opposed to the shortest journey.

Given a decision network having the general form of those above, and encoding these specific dependencies and utilities, it is possible to decide which propositional variable should be sampled next by the following means.

1. Given a variable with m values: $\mathrm{V}_{1\ldots m}$, then for action a, evidence e, causal effects $\mathrm{C}_{1\ldots n}$ of a, we can calculate the utility of an action with respect to the value of a variable by the following expression:

$$U(a \mid V_i) = \prod_{j=1..n} u(C_j \mid a, V_i) \times P(C_j \mid a, e, V_i)$$

This expression is related to that used earlier for calculating the utility of a move: the difference is that there, the relevant variable, V, was assumed to be instantiated already.

2. Now we can define the utility for each value of V as:

$$U(V_i) = max_a U(a \mid V_i)$$

This expression tells us the maximum utility that can theoretically be derived from

this value of the variable.
3. Now the overall utility of querying V can be computed by summing the product of the utilities and the likelihood of realising them under the current set of evidential instantiations:

$$\prod_{i=1\ldots m} P(V_i \mid e) \times U(V_i)$$

Performing these computations for all uninstantiated variables will allow the most useful one to be questioned next: the variable that has the highest overall possible utility in the current circumstances is a rational choice for the subject of the next question. Of course, in a large network these calculations might be rather expensive: a practical method might involve some kind of stochastic sampling of variables rather than an exhaustive comparison.

Related Work

The notion of 'language game', 'conversational game' or 'dialogue game' has a long history in 20th century philosophy of language, starting with Wittgenstein. Games interpreted in a decision-theoretic way have also been used within philosophy of language, notably by Hintikka, although within computational linguistics this line of enquiry is probably best known, in one version at least, through the work of Carlson (1983). However, the most direct inspiration for the approach described here is a paper by Gamback, Rayner, and Pell (1991), in which they describe a hybrid rule-based/neural network approach to the pragmatics micro-world of bidding in bridge, in which bids are seen as various kinds of simple speech act.

Bayesian Networks have been used in natural language processing for story understanding (see, for example, Charniak and Goldman (1991)) and word-sense disambiguation. They have also been used by Araki, Kawahara, and Doshita (1995) for dialogue understanding, although in a somewhat different way than envisaged here. The system they describe uses two networks: one 'Conversational Space' network is responsible for hypothesising the interpretation of an utterance and the associated speaker intention. It combines syntactic, semantic, and discourse structure information into a single network, which is constructed dynamically for each new utterance. The second network ('Problem Solving Space') encodes a model of the task domain and is responsible for plan recognition, and for selecting the appropriate type of response. Other mechanisms (e.g. utterance type trigrams) are also used, and 'mental state' is modelled separately, apparently not by a Bayesian network. Explicit utilities and the framework of causal influence diagrams are not used.

Conclusions

We began with three questions that should be answered by any satisfactory computational theory of dialogue. It is worth spelling out the kinds of answers that are given to these questions by the framework we have sketched.

(i) what are mental states? - in the Bayesian network approach, mental states are represented by sets of propositions linked by causal (or logical) relations, with a probability distribution on them that respects these causal relationships. There is a straightforward interpretation of these networks as networks of beliefs, and indeed that is how they were originally envisaged in Pearl (1988). When Bayesian networks are augmented with the apparatus of decision and value nodes, and utilities, then it is plausible to think of them as modelling some aspects of desire and and possibly intention, although the correspondence is not exact. This type of Bayesian reasoning is less powerful than that assumed in classical belief revision or associated frameworks like dynamic logic. Quantificational reasoning, beliefs about beliefs, etc. can only be handled to the extent that they can be 'compiled out' to a propositional format. However, classical BDI implementations have not been able to actually make use of this extra power on a large scale yet and so it remains to be seen whether this is a serious practical constraint.

(ii) how do they change? - states change by the instantiation of nodes representing new evidence, and the consequent updating of probabilities. Conflict between beliefs or intentions in the presence of new input is not modelled explicitly, but can be associated with large differences between *a priori* probabilities and values derived from new evidence. Evidence from multiple sources can be combined unproblematically.

(iii) how do utterances connect with them and change them? - the connection between utterance types and mental states is conventionalised via conversational games, and this conventional connection is encoded in the structure of the relevant networks for move recognition and response. Many of the insights of speech act theory and the BDI tradition are retained and encoded in this way, although their interpretation is now partly probabilistic rather than strictly logical.

Clearly, there is great deal of work to be done before the preceding ideas can be implemented and tested in detail. However, we regard this as a promising perspective from which to approach the problem of building dialogue understanding systems. The Bayesian network architecture seems to provide the right combination of rule based and statistical methods. We can retain what is intuitively correct about the BDI tradition, while overcoming the difficulties and fragilities associated with strictly axiomatic systems.

One obvious question of course, is: where do the networks and their associated probabilities come from? Although it is possible in principle to learn the structure of a Bayesian net from examples, we feel that it is more productive at least in the short term to think of their basic structure as reflecting (corpus-guided) linguistic descriptions of conversational game and move structure. However, the probabilities associated with the nodes in a network should reflect observed properties in a relevant corpus, and it is quite plausible to think of these as being automatically trained from an annotated corpus.

Acknowledgements

Most of the work described here was carried out under a contract to SRI from the Speech Research Unit, DERA Malvern, Contract CSM3/002, and the current paper is a revised and extended version of part of the final report for that project. As well as at CLIN 96 in Eindhoven, versions of the paper have been given to the Speech Research Unit, DERA Malvern; and the Human Communication Research Centre, Edinburgh. I am grateful for the comments received on all these occasions.

I would also like to acknowledge the many valuable inputs to this work by Ian Lewin.

References

Allen, J., B. Miller, E. Ringger, and T. Sikorski (1996). A robust system for natural spoken dialogue. In *Proceedings of 34th ACL Santa Cruz*, pp. 62–70.

Araki, M., T. Kawahara, and S. Doshita (1995). Cooperative spoken dialogue model using bayesian network and event hierarchy. In *Proceedings of ESCA Workshop on Spoken Dialogue Systems, Denmark*, pp. 177–180.

Bunt, H. (1997). Dynamic interpretation and dialogue theory. In M. Taylor, D. Bouwhuis, and F. Neel (Eds.), *The structure of multi-modal dialogue*, Volume 2. Amsterdam John Benjamins.

Carlson, L. (1983). *Dialogue Games*. Dordrecht: D. Reidel Publishing Co.

Charniak, E. and R. Goldman (1991). A probabilistic model of plan recognition. In *Proceedings of Ninth National Conference on AI*, pp. 160–165. AAAI.

Clark, H. and E. Schaefer (1989). Contributing to discourse. *Cognitive Science 13*, 259–294.

Cohen, P. and H. Levesque (1990). Rational interaction as the basis for communication. In P. Cohen, J. Morgan, and M. Pollack (Eds.), *Intentions in Communication*. MIT Press.

Cohen, P., J. Morgan, and M. Pollack (1990). *Intentions in Communication*. MIT Press.

Cohen, P. and C. Perrault (1979). Elements of a plan-based theory of speech acts. *Cognitive Science 3*(3), 177–212.

Galliers, J. (1990). Belief revision and a theory of communication. Technical Report Technical Report 193, University of Cambridge Computer Laboratory.

Gamback, B., M. Rayner, and B. Pell (1991). Pragmatic reasoning in bridge. SRI Cambridge technical report CRC-030, http://www.cam.sri.com/.

Gärdenfors, P. (1988). *Knowledge in flux: modeling the dynamics of epistemic states*. MIT Press.

Groenendijk, J. and M. Stokhof (1991). Dynamic predicate logic. *Linguistics and Philosophy 14*, 39–100.

Hamblin, C. (1971). Mathematical models of discourse. *Theoria 37*, 130–155.

Houghton, G. (1986). *The Production of Language in Dialogue.* Ph. D. thesis, University of Sussex.

Jaspars, J. (1996). A modal unification of dynamic theories. Building the Framework: FraCas Deliverable D15.

Konolige, K. (1986). *A Deduction Model of Belief.* London: Pitman.

Kowtko, J., S. Isard, and G. Doherty (1992). Conversational games within dialogue. HCRC research paper RP-31.

Lewin, I. and S. Pulman (1995). Inference in the resolution of ellipsis. In *Proceedings of ESCA Workshop on Spoken Dialogue Systems, Vigso, Denmark*, pp. 53–56.

Lewin, I., M. Russell, D. Carter, S. Browning, K. Ponting, and S. Pulman (1993). A speech-based route enquiry system built from general-purpose components. In *Eurospeech '93: Proceedings of the 3rd European Conference on speech communication and technology*, Volume 3, pp. 2047–2050.

Moore, R. and S. Browning (1992). Results of an exercise to collect 'genuine' spoken enquiries using woz techniques. In *Proceedings of the Institute of Acoustics 14 6*, pp. 613–620.

Neapolitan, R. (1990). *Probabilistic reasoning in expert systems: theory and algorithms.* Wiley.

NextBase (1991). Autoroute plus user guide. NextBase Limited, Headline House, Chaucer Road, Ashford, Middlesex, England.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufman.

Perrault, C. and J. Allen (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics 6*(3-4), 167–182.

Power, R. (1979). The organization of purposeful dialogues. *Linguistics 17*, 107–152.

Reithinger, N. and E. Maier (1996). Utilizing statistical dialogue act processing in verbmobil. In *Proceedings of 33rd ACL, Cambridge Mass.*, pp. 116–121.

Sadek, D., A. Ferrieux, A. Cozannet, P. Bretier, F. Panaget, and J. Simonin (1996). Effective human computer cooperative spoken dialogue. In *Proceedings ICSLP 96 Conference in Spoken Language Processing, Philadelphia*, pp. 546–549.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language.* Cambridge Univeristy Press.

Shoham, Y. (1994). *Artificial Intelligence Techniques in Prolog.* San Francico: Morgan Kaufman.

Traum, D. and J. Allen (1994). Discourse obligations in dialogue processing. In *Proceedings of the 32nd ACL, Las Cruces, New Mexico*, pp. 1–8.

Traum, D. and E. Hinkelman (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence 8*(3).