Sonderforschungsbereich 378
Ressourcenadaptive kognitive Prozesse

KI-Labor am Lehrstuhl für Informatik IV

Leitung: Prof. Dr. W. Wahlster

READY

Universität des Saarlandes
FB 14 Informatik IV
Postfach 151150
D-66041 Saarbrücken
Germany

Tel. 0681 / 302-2363

Bericht Nr. 133

How to Juggle Discourse Obligations

Anthony Jameson, Thomas Weis

Oktober 1996

Sonderforschungsbereich 378
Ressourcenadaptive
Kognitive Prozesse

ISSN 0944-7814

133

Proceedings of the Symposium on Conceptual and Semantic Knowledge in Language Generation, Heidelberg, Germany, 15–17 November 1995.

How to Juggle Discourse Obligations*

Anthony Jameson and Thomas Weis
Department of Computer Science, University of Saarbrücken
P.O. Box 151150, D-66041 Saarbrücken, Federal Republic of Germany
{jameson, tweis}@cs.uni-sb.de

Abstract

In the context of dialog, one challenge for natural language generation is that of dealing with *discourse obligations* such as those created by questions, promises, and even silences. The present paper presents new methods for (a) representing discourse obligations, (b) determining which ones are created by a given dialog move, (c) determining how the status of an existing obligation is changed by a new move, and (d) taking existing and resulting obligations into account in the generation of new moves. The task of choosing a dialog move (which may consist of a signal for the user to produce the next utterance) is conceptualized in terms of the maximization of expected utility. This quantitative approach makes it possible to do justice to observations such as the following: (a) Discourse obligations lead to trade-offs with the system's other dialog goals. (b) Obligations differ in terms of their importance and the time within which they are to be fulfilled. (c) The severity of the consequences of nonfulfillment of an obligation depends on the relationship between the obligated speaker's moves and the total set of existing obligations.

Introduction

Issues Concerning Discourse Obligations

Traum and Allen (1994) argued convincingly that the concept of *discourse obligations* accounts for many phenomena of natural language dialog better than the more usual analyses in terms of, for example, recognition of intentions:

> The model we propose is that an agent's behavior is determined by a number of factors, including that agent's current goals in the domain, and a set of obligations that are induced by a set of social conventions. When planning, an agent considers both its goals and obligations in order to determine an action that addresses both to the extent possible. When prior intentions and obligations conflict, an agent generally will delay pursuit of its intentions in order to satisfy the obligations, although the agent may behave otherwise at the cost of violating its obligations. At any given time, an agent may have many obligations and many different goals, and planning involves a complex tradeoff between these different factors. (p. 2)

As Traum and Allen indicated, the need for the concept of *obligation* becomes especially apparent in dialog situations in which the participants do not have common goals.

Consider, for example, the following moment in a dialog, in which the user $\mathcal{U}$ is a potential used car buyer and the system $\mathcal{S}$ has the role of a used car seller:

---

*This research was supported by the German Science Foundation (DFG) in its Special Collaborative Research Program on Artificial Intelligence and Knowledge Based Systems SFB 314, Project N1, PRACMA. In the project READY of SFB 378, to be supported by the DFG from January 1996, variants of the methods presented here will be applied in a dialog context characterized by resource limitations of the participants.

▷ $\mathcal{U}$: `For me, safety is the most important thing.`
▷ $\mathcal{S}$: [Wonders what to say about the car under discussion.]

Most likely the seller will choose to make some statement about the car's safety-related features; but why? If we assume that $\mathcal{S}$ shares $\mathcal{U}$'s goal of obtaining information about the car that is important to $\mathcal{U}$, then this question is easy to answer: $\mathcal{U}$'s utterance has revealed a way in which $\mathcal{S}$ can help $\mathcal{U}$ to achieve this goal. But what if $\mathcal{S}$ just wants to sell the car and $\mathcal{S}$ knows that the car rates poorly in terms of safety-related features? Then $\mathcal{S}$'s dialog goals would dictate avoiding this topic. Yet even such an $\mathcal{S}$ will not in general simply ignore $\mathcal{U}$'s utterance; she[1] will recognize that it creates a discourse obligation that she can flout only at her own peril. Instead of either ignoring the obligation or fulfilling it completely, she may try to find a satisfactory trade-off between her desire to obey dialog conventions and her desire to sell her car.

A similar tension can arise even if $\mathcal{S}$ is cooperatively motivated. Suppose, for example, that $\mathcal{S}$ is aware of properties of the car that are far more important to $\mathcal{U}$ than its safety-related features—for example, that it is in such poor repair that it is unlikely to go another 10 miles before collapsing. In spite of having $\mathcal{U}$'s interests at heart, $\mathcal{S}$ will be tempted to ignore $\mathcal{U}$'s previous utterance, so that she can proceed directly to the really important information; but here again, she will feel somewhat constrained by the obligation created by that utterance.

We will focus on noncooperative situations in this paper, because there the role of obligations is easier to see; but the considerations will also have relevance for cooperative dialogs.

Although Traum and Allen's (1994) conception of discourse obligations is general enough to cover the type of case just discussed, their realization of it in the context of the TRAINS system is evidently more limited (see also Allen et al., 1995). In particular, they say little about cases in which the system handles trade-offs between the desire to fulfill obligations and the desire to achieve the system's goals. As the TRAINS system has a cooperative motivation, it relatively seldom has a reason to try to avoid fulfilling discourse obligations.

The present contribution can be seen as a fuller working out of Traum and Allen's (1994) general conceptualization (although most of the work was done independently). We will address the following general issues (as well as a number of more specific ones):

1. How are discourse obligations created by dialog moves?
   The general way in which this happens has been discussed by previous researchers, but our account will add some new points.
2. How are discourse obligations fulfilled by subsequent dialog moves—or, more generally, how do dialog moves affect the status of existing obligations?
   Our account differs from previous ones in that it presents a more differentiated picture of the possible relationships between a dialog move and an existing obligation.
3. In what ways do discourse obligations constrain the generation of dialog moves?
   This issue is related to the previous one in that a speaker, when selecting a dialog move, must take into account the effect that it would have on existing obligations. But now the question must be addressed: How do these anticipated effects figure in the evaluation of a potential move? Our approach is characterized by a quantitative assessment of various factors that determine the desirability of a potential dialog move. The sort of trade-off between factors that is mentioned in the quoted passage of Traum and Allen (1994) can be handled especially naturally within such a framework.

Example Domain and System

These issues will be discussed with reference to methods that are, at the time of this writing, being integrated into the dialog system PRACMA (see, e.g., Jameson et al., 1995; Ndiaye & Jameson, 1996). They have been implemented and refined within a smaller prototype—to be called MICROPRACMA—in

[1] For clarity, we arbitrarily use feminine pronouns to refer to the seller in this dialog situation—even as simulated by the system—and masculine pronouns for the buyer.

which almost all of PRACMA's modules are replaced by simple emulators. Except where otherwise stated, all references to “the system” in the following will concern MICROPRACMA.

Both PRACMA and MICROPRACMA model dialogs in which a seller is trying to sell her used car to a potential buyer. The goals of the two dialog participants in this example domain conflict in part: The buyer wants to get the best possible information on which to base a decision about the car, whereas the seller would like to sell the car, whether or not it is really suitable for the buyer.

PRACMA and MICROPRACMA are able to take the role of either the seller or the buyer within their dialog situation. In the examples discussed in this paper, MICROPRACMA will be taking the role of the seller.

Structure of the Paper

So that not all complications have to be dealt with at the same time, the following five major sections will consider a sequence of five increasingly complex dialog situations. Each section will include the following subsections:

1. *Definition of the Situation:* a statement of the assumptions made about each situation.
2. *Theoretical Considerations:* a discussion of the considerations that need to be taken into account by the design of a system that is to act appropriately in this situation.
3. *Realization in the System:* a description of how these considerations have been taken into account in MICROPRACMA.
4. *Resulting Behavior:* a discussion, based on the performance of MICROPRACMA, of typical dialog phenomena that result if the strategies in question are applied.

Situation 1: Dialog Unconstrained by Discourse Obligations

Definition of the Situation

In this first situation, $\mathcal{U}$'s behavior is assumed to be such that it creates no discourse obligations for $\mathcal{S}$: In particular, $\mathcal{U}$ listens largely passively to what $\mathcal{S}$ has to say and answers $\mathcal{S}$'s questions directly. Although $\mathcal{S}$'s moves create obligations for $\mathcal{U}$, $\mathcal{S}$ does not yet take these obligations into account explicitly.

This situation serves as a baseline with which the later situations, which involve discourse obligations, can be compared. It allows us to explain the system's move selection criteria that do not refer to obligations; these criteria will later be applied in conjunction with criteria that do refer to obligations.

We also assume in all situations in this paper that the seller does not engage in complex dialog planning but rather performs basically reactive *dialog management* (cf. Traum & Allen, 1994). Most of the ideas introduced are also relevant in a dialog planning context, though some aspects of the move-selection strategies would have to be replaced.[2]

Finally, to postpone the discussion of turn taking issues until a more appropriate point (Situation 5), we assume for now that each dialog participant generates exactly one utterance when it's his or her turn and then surrenders the turn to the other participant.

Theoretical Considerations

In this situation, two important types of dialog move for the seller (or system) $\mathcal{S}$ are the following:

1. Moves that present information about a given car.
   Given its overall dialog goal, $\mathcal{S}$ should aim mainly to choose moves that will create a favorable evaluation of the car by $\mathcal{U}$.

[2]One example of a dialog plan in the car-selling domain is the trick whereby the seller persuades the buyer to agree to buy a given car at a reduced price, only to report subsequently that the manager has refused to authorize the price reduction.

2. Moves that elicit information from $\mathcal{U}$ concerning $\mathcal{U}$'s evaluation criteria and knowledge.
   This type of information will put $\mathcal{S}$ in a better position to choose moves of the first type. The best moves are the ones that will reduce $\mathcal{S}$'s uncertainty about $\mathcal{U}$ to the greatest extent. The utility of a move like this is in general greatest if it is made early in the dialog, because $\mathcal{S}$ will then have the most opportunity to exploit the information obtained.

The general point is that the benefits associated with each potential move should be evaluated on some sort of quantitative scale; only then is it possible for $\mathcal{S}$, for example, to judge whether it is more worthwhile to ask $\mathcal{U}$ what his profession is or how many children he has.

Realization in the System

The particular techniques that the larger system PRACMA uses to arrive at quantitative evaluations of possible dialog moves are discussed by Jameson et al. (1995).[3] For the present purposes, the only important point is that the techniques yield quantitative assessments of the expected benefits of possible dialog moves. MICROPRACMA simply uses prestored numbers that correspond to the results that would be computed dynamically in PRACMA.

For MICROPRACMA the number of possible dialog moves at any given time is finite, and the possible moves are generated with simple template-based methods. The move selection problem thus reduces in this baseline situation to the problem of choosing from a list of possible moves the move with the highest assessed benefits.

Resulting Behavior

$\mathcal{S}$ tends to ask questions of $\mathcal{U}$ that are likely to yield the most useful information, and $\mathcal{S}$ sends to generate descriptive statements that are likely to make the most favorable impressions. As the benefits of the first type of move are assigned their highest weight early in the dialog, this type of move tends to come early. In the full PRACMA system, $\mathcal{S}$'s utterances become increasingly well tailored to $\mathcal{U}$ as more and more evidence concerning $\mathcal{U}$'s evaluation standards and other characteristics becomes available.

Situation 2: Fulfilling or Rejecting Discourse Obligations

Definition of the Situation

In this situation, it is possible for $\mathcal{U}$ to make dialog moves that create discourse obligations for $\mathcal{S}$. In addition, $\mathcal{S}$ may incur obligations that are created by $\mathcal{S}$'s own moves, as when $\mathcal{S}$ promises to supply some information to $\mathcal{U}$. Cases in which two or more obligations for $\mathcal{S}$ exist simultaneously will not be considered until Situation 3; similarly, for now we only consider moves by $\mathcal{S}$ that either directly fulfill or violate its obligations. The obligations which arise for $\mathcal{U}$ are not yet taken into account explicitly by $\mathcal{S}$.

Theoretical Considerations

First it is necessary to define a concept of a *discourse obligation* as an abstract entity with the following properties, among others:

- the nature of the dialog move(s) that $\mathcal{S}$ is obligated to perform; and
- the severity of the penalty for $\mathcal{S}$ if the obligation is not fulfilled.

[3]The techniques are based on multiattribute utility theory (von Winterfeldt & Edwards, 1986) as an underlying framework, and they use Bayesian networks (Pearl, 1988) to manage the large amount of uncertainty involved and to assess the value of information that might be obtained from $\mathcal{U}$. The same Bayesian networks are used for the interpretation of evidence in $\mathcal{U}$'s dialog behavior.

The concept of a *penalty* is employed here because of the analogy with everyday situations in which a person is threatened with concrete sanctions if he or she fails to fulfill a particular obligation (e.g., the obligation to return a library book by a given date). In the context of a dialog, the negative consequences of failure to fulfill a discourse obligation are of course much less tangible and measurable; they include consequences such as irritation and negative judgments on the part of the dialog partner, as well as the unease that people themselves tend to feel when they deviate visibly from broadly accepted conventions. For simplicity of exposition we will write as if $\mathcal{S}$ had to pay some single fine whose size corresponded to the sum of all of these negative consequences.

The fact that such penalties can differ significantly in their size can be seen if we compare moves that create basically the same obligation but with a different degree of bindingness, such as the following two alternative moves:

▷ $\mathcal{U}$: `I consider fuel consumption to be an important criterion.`

▷ $\mathcal{U}'$: `What's this car's fuel consumption?`

In each of these cases, if $\mathcal{S}$ responds with a dialog move that in no way addresses the issue of fuel consumption, $\mathcal{S}$ will have violated an obligation created by $\mathcal{U}$'s move; but the violation is considerably more serious in the second case, and it will be associated with a higher penalty.

$\mathcal{S}$ has to have knowledge about the obligations created by moves of various types, and $\mathcal{S}$ must apply this knowledge after each dialog move so as to take note of any new obligations that have arisen for $\mathcal{S}$.

Equally essential for $\mathcal{S}$ is knowledge of what conditions a future dialog move has to satisfy if it is to count as fulfilling a given obligation.

Finally, obligations would have little significance if they weren't taken into account in the generation of dialog moves: When choosing a move while faced with an obligation $\mathcal{O}$, $\mathcal{S}$ should consider not only the expected benefit of each possible move $\mathcal{M}$ (as in the previous situation) but also whether each possible $\mathcal{M}$ would fulfill or violate $\mathcal{O}$. In the latter case, the violation of $\mathcal{O}$ should be viewed as a drawback of $\mathcal{M}$, one whose weight depends on the size of the penalty associated with $\mathcal{O}$.

Realization in the System

The implementation of MICROPRACMA (and of the relevant parts of PRACMA) is object-oriented (with COMMONLISP / CLOS as the language of implementation). Most of the aspects of the model to be discussed are realized through the definition of classes and associated methods.

Entities

A number of types of dialog move are distinguished, each type corresponding to a subclass of the class Move. In Figure 1 (some aspects of which will only be discussed in later sections), the left-hand and right-hand columns include examples of move classes. Most of these are domain-dependent; this fact reflects the fact that on the pragmatic level being considered here, the function of an utterance like "I'm especially concerned about safety" is partly specific to the nature of the discourse that is taking place. Each instance of such a class represents a number of properties of a move, such as the speaker, the formulation used, and the content conveyed.

The other major class of entities is Obligation. This class also subsumes a number of subclasses, some of which are shown in the middle column of Figure 1. Each subclass corresponds to a type of requirement that the obligation imposes. The following are the main properties that are associated with each instance $\mathcal{O}$ of a subclass of Obligation. (Properties and concepts that only become relevant in Situation 3 are marked with *.)

1. The dialog move that gave rise to $\mathcal{O}$;
   The content of this move largely determines the content of $\mathcal{O}$.
2. the current *status* of $\mathcal{O}$;

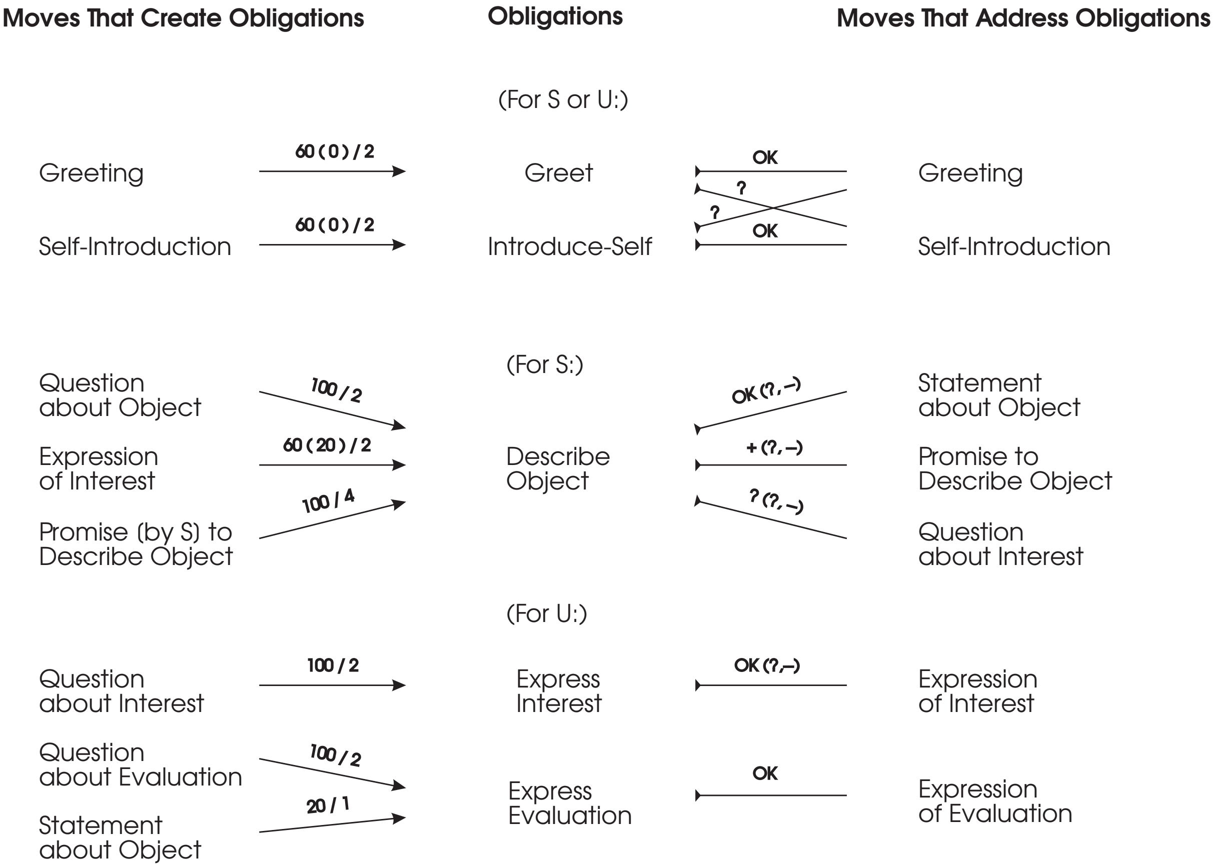

Figure 1: Summary of some methods for the creation and updating of discourse obligations.

Each arrow on the left-hand side means that a move of the type specified creates an obligation of the type specified; in the notation p/d, p refers to the total penalty associated with the obligation and e refers to the length of its expiration period (in moves); in $p_1(p_2)/e$, p_2 is the penalty that applies if the obligation-creating move is not spontaneous (cf. Situation 4). Each arrow in the right-hand side of the figure means that a move of the type specified (in the rightmost column) changes the status of the specified obligation to one of the statuses listed above the arrow (OK = Fulfilled, + = Accepted, ? = Possibly Accepted, − = Rejected). Where more than one status is specified, the first one applies if the content of the move is directly relevant to the content of the obligation; if the move is only partly relevant, or if it is irrelevant, the first or second of the statuses in parentheses apply, respectively. Any move of a type not linked to an obligation by such an arrow changes the obligation's status to Rejected (though this status may subsequently be changed to Delayed, as will be discussed in connection with Situation 3).

The possible statuses are Fresh, Fulfilled, Rejected, *Accepted, *Possibly Accepted, and *Delayed. The meaning of the first three is as follows:

- Fresh means that $\mathcal{O}$ has just been established and that $\mathcal{S}$ has not yet had a chance to address it.
- Fulfilled means that $\mathcal{S}$ has satisfied the requirement imposed by $\mathcal{O}$.
- Rejected means that $\mathcal{S}$ has made it clear that she is not going to fulfill $\mathcal{O}$.

3. the *total penalty* that will be incurred if the status of $\mathcal{O}$ is changed to Rejected;
4. * the *expiration period*, i.e. the number of dialog moves during which the obligated party can fulfill $\mathcal{O}$.

Methods for Creating and Updating Obligations

$\mathcal{S}$'s knowledge about the obligations created by particular moves essentially takes the form of rules, as is the case with Traum and Allen's (1994) model. The rules are implemented procedurally with CLOS methods, each of which is associated with a subclass of moves: For each such subclass, there is a method that determines what obligations, if any, an instance of that subclass creates, and what penalty and expiration period are associated with the created obligation. The arrows on the left in Figure 1 give some examples of such rules. (The numbers appearing in parentheses above some of the arrows reflect the context-dependence of the corresponding rules, which will be discussed in Situation 4.) The two uppermost rules correspond to the relatively domain-independent obligations to respond to a greeting with a greeting and to a statement like “My name is X” with a similar statement of one's own. The next two rules create obligations for the seller in the used-car domain: A question about the object under discussion (e.g., “What's the car's fuel consumption?” or “What safety features does the car have?”) creates a strong obligation to make some relevant descriptive statement about the object; an expression of interest like “I'm concerned about safety features” creates a weaker obligation of the same sort. The next rule illustrates how a dialog move of $\mathcal{S}$ (e.g., “I'll tell you about the car's safety features”) can create an obligation for $\mathcal{S}$ herself. (The lowest third of the figure will be discussed in connection with Situation 4.)

The most straightforward way to represent knowledge about how a given obligation can be fulfilled would be with a list of one or more characterizations of possible moves that could fulfill it (cf., e.g., Traum & Allen, 1994, p. 3). A more general approach is to associate with each subclass of obligation a method that determines, for a given obligation $\mathcal{O}$ and a given dialog move $\mathcal{M}$, the effect that $\mathcal{M}$ has on the status (and other properties) of $\mathcal{O}$. (The advantages of this approach will become clearer in Situation 3, in which a larger number of possible statuses for an obligation are considered.) Each arrow on the right-hand side of Figure 1 with the label “OK” corresponds to a method which specifies that a move of a given class changes the status of an obligation of a particular class to Fulfilled. This change can be subject to particular conditions; for example, a move of the class Statement About Object only fulfills an obligation of the class Describe Object if the move is directly relevant to the topic of the move that gave rise to the obligation. In the following exchange, for example, $\mathcal{S}$'s move changes the status of the Describe Object obligation to Rejected:

▷ $\mathcal{U}$: `Does the car have white-walled tires?`
▷ $\mathcal{S}$: `The car has air conditioning.`

Taking Obligations into Account in Move Selection

In this simplified situation, it is straightforward for $\mathcal{S}$ to take any obligation it faces into account when deciding what to say: The evaluation of a possible move $\mathcal{M}$ considers not only its expected benefits (as in the previous situation) but also—if $\mathcal{M}$ is a move that would change the status of an obligation to Rejected—the penalty that $\mathcal{S}$ would have to pay.

Resulting Behavior

In this situation, the system behavior that results from our conceptualization of obligations is still largely unremarkable: When $\mathcal{S}$ faces an obligation, she is inclined to fulfill it, because otherwise she has to pay the penalty. But there is one nontrivial consequence, which results from the quantitative distinction of different penalties for obligations and different benefits of moves: $\mathcal{S}$ will sometimes flout an obligation if the penalty is relatively low and/or the desirability for $\mathcal{S}$ of making some move that is inconsistent with the obligation is especially high. In such cases $\mathcal{S}$ acts like a person who chooses to pay the fine for keeping a library book past its due date so as to be able to use the book for an important purpose.

Situation 3: More Complex Effects of Moves on Obligations

Definition of the Situation

We now take into account the fact that not all discourse obligations are necessarily fulfilled or rejected immediately after their creation. If $\mathcal{S}$ has some time during which she can address the obligation, her moves during that period can have other, intermediate effects on the status of the obligation.

Theoretical Considerations

Expiration Periods for Obligations

As was suggested earlier, some discourse obligations must be viewed as specifying a relatively long time during which they can be fulfilled (cf. the example expiration periods specified by the obligation-creation arrows in the left-hand side of Figure 1).

The conventions that determine these expiration periods depend on the type of dialog involved. For example, in the context of a courtroom cross-examination, a question presumably obligates the witness to answer without any intervening dialog move, whereas in a casual conversation a question may unobjectionably be answered several moves after it was asked.

One issue that arises is that of why discourse obligations in conversation tend to be seen as having a limited expiration period at all, rather than simply being added without deadline to a general to-do list for the obligated party. There appear to be two general reasons (which will be discussed for concreteness with reference to the roles $\mathcal{S}$ and $\mathcal{U}$):

1. $\mathcal{S}$'s fulfillment of an obligation is often of use to $\mathcal{U}$ only if it occurs within a certain period of time.
   For example, if $\mathcal{U}$ has asked a question, he may need to use the answer as the basis for his selection of further dialog moves.
2. Less obviously, the resource limitations that typify everyday dialog—in particular, the inability to use external storage—generally make it difficult for dialog participants to maintain large to-do lists of obligations.[4]
   The convention that discourse obligations expire sooner or later ensures in two ways that the number of obligations faced by $\mathcal{S}$ will not get very large:
   - Even if $\mathcal{U}$ insists on piling obligations onto $\mathcal{S}$ faster than $\mathcal{S}$ can fulfill them, some of these obligations will expire unfulfilled, so that the number that $\mathcal{S}$ has to deal with will remain small.
   - Recognition of this fact encourages $\mathcal{U}$ not to overload $\mathcal{S}$ with obligations in the first place, as this practice is counterproductive for $\mathcal{U}$.
     This second consideration becomes especially relevant in Situation 5, in which the number of dialog moves per turn is variable; it will be discussed further in connection with that situation.

Intermediate Statuses for Obligations

Before the expiration period has elapsed, a move $\mathcal{M}$ by $\mathcal{S}$ can have one of the following three additional relationships to the obligation in question (cf. the examples in the right-hand side of Figure 1):

1. $\mathcal{M}$ may indicate to $\mathcal{U}$ that $\mathcal{S}$ intends to fulfill $\mathcal{O}$, even though $\mathcal{S}$ has not yet done so.
   A clear example of such a case is one where $\mathcal{S}$ announces the intention to fulfill $\mathcal{O}$.
   This case should not be handled in the same way as a direct fulfillment of $\mathcal{O}$; after all, $\mathcal{S}$ may never actually get around to fulfilling $\mathcal{O}$. On the other hand, the penalty paid by $\mathcal{S}$ will be lower (for the time being) than in the case where $\mathcal{S}$ rejects $\mathcal{O}$ outright. It seems in general reasonable to view $\mathcal{S}$ as paying some intermediate penalty which is positively related to the importance of $\mathcal{O}$ and inversely related to the length of $\mathcal{O}$'s expiration period.
2. $\mathcal{M}$ may leave it unclear to $\mathcal{U}$ whether $\mathcal{S}$ will get around to fulfilling $\mathcal{O}$ before $\mathcal{O}$ expires.
   Consider, for example, the following exchange:
   - ▷ $\mathcal{U}$: `I consider fuel consumption to be an important criterion.`
   - ▷ $\mathcal{S}$: `The car has quite a powerful motor.`

   Unlike the utterance "The car has a great hifi set", this response by $\mathcal{S}$ leaves open the possibility that $\mathcal{S}$ will soon proceed to say something about the car's fuel consumption, having now made an indirectly relevant preliminary remark.

[4] Note, by contrast, that in contexts such as legal debate it is not unusual for one party to raise a large number of issues that will have to be addressed at some point by the other party. This behavior is accepted by both sides because the resources required for the necessary obligation bookkeeping are available.

$\mathcal{S}$ should, however, be viewed as paying some penalty for not addressing the obligation directly, if only because she has kept $\mathcal{U}$ in the dark for a while with regard to the prospect that the obligation will be fulfilled.

3. If there exists simultaneously with $\mathcal{O}$ another obligation $\mathcal{O}'$, $\mathcal{M}$ may fulfill $\mathcal{O}'$ but not $\mathcal{O}$.
   Here is an example:
   - ▷ $\mathcal{U}$: `Reliability is the most important factor for me.`
   - ▷ $\mathcal{S}$: `I'll tell you something about the car's reliability.`
   - ▷ $\mathcal{U}$: `Does the car have white-walled tires?`
   - ▷ $\mathcal{S}$: `It doesn't have white-walled tires.`

   At the point where $\mathcal{S}$ has to choose her second move, $\mathcal{S}$ has accumulated two obligations: to say something about the car's reliability and to say whether it has white-walled tires. $\mathcal{S}$'s second move fulfills only the second obligation, but it does not signify a rejection of the first one. $\mathcal{S}$ should not be seen as paying a penalty for neglecting the first obligation; the existence of the second obligation provides a good excuse.

Realization in the System

As was indicated above, the idea that obligations have different expiration periods is represented simply in the present prototype through the association with each obligation $\mathcal{O}$ of a given number of moves within which $\mathcal{O}$ is expected to be fulfilled (cf. the numbers above the arrows in the left-hand side of Figure 1).

The three intermediate statuses Accepted, Possibly Accepted, and Delayed correspond to the three cases mentioned in the previous subsection. Whereas obligations that have been fulfilled or rejected are removed immediately from the table of $\mathcal{S}$'s obligations, as long as an obligation has one of the three intermediate statuses it remains in the table of obligations for $\mathcal{S}$ until its expiration period has elapsed.

The main differences among the consequences of these three statuses concern the penalty that is incurred by a move that gives an existing obligation the status in question. The exact formulas and numbers currently used in the prototype are highly tentative and somewhat arbitrary, but we present them here for concreteness and as a stimulus for discussion:

1. If a move $\mathcal{M}$ changes the status of an obligation $\mathcal{O}$ to Possibly Accepted, the penalty paid by $\mathcal{S}$ is the Penalty per Move, i.e., the quotient p/e, where p is the total penalty associated with $\mathcal{O}$ and e is the length of its expiration period.
   The rationale is that if $\mathcal{S}$ keeps leaving it unclear whether she will fulfill $\mathcal{O}$ in spite of having a chance to do so, then by the time $\mathcal{O}$ has expired the total penalty should be the same as if $\mathcal{S}$ had rejected $\mathcal{O}$.[5]
2. If $\mathcal{M}$ changes the status of $\mathcal{O}$ to Accepted, only half of the Penalty per Move is paid.
   As argued above, the penalty in this case should be lower than in the first case. This does not mean that $\mathcal{S}$ can in general escape an obligation fairly easily just by accepting it and then not fulfilling it: The dialog move by $\mathcal{S}$ that signals acceptance of $\mathcal{O}$ at the same time creates a further obligation with basically the same content as that of $\mathcal{O}$. So such a move is comparable to the tactic of taking up a new loan in order to pay off an existing one. (The question of how such a move may benefit $\mathcal{S}$ is discussed in the next subsection.)
3. If $\mathcal{M}$ fulfills an alternative obligation $\mathcal{O}'$ and therefore changes the status of $\mathcal{O}$ to Delayed, $\mathcal{S}$ pays no penalty at all—provided that $\mathcal{O}'$ was a more urgent obligation than $\mathcal{O}$, that is, if the Penalty per Move for $\mathcal{O}'$ was greater than that for $\mathcal{O}$. If $\mathcal{O}'$ was *less* urgent, the difference between the two Penalty per Moves is paid.
   The rationale here is that the fulfillment of another obligation which is at least as urgent as $\mathcal{O}$ constitutes a good excuse for not fulfilling $\mathcal{O}$. On the other hand, it is less acceptable for $\mathcal{S}$ to neglect an urgent obligation in order to fulfill a trivial one.

[5]This argument is actually strictly applicable only in Situation 5, in which $\mathcal{S}$ can be given the opportunity to keep talking throughout the expiration period of an obligation.

Resulting Behavior

Now that the possible responses of $\mathcal{S}$ to a discourse obligation have been expanded, when will $\mathcal{S}$ actually choose to make use of the new options?

When two obligations that are associated with similar Penalty per Moves exist, $\mathcal{S}$ can basically choose which one to fulfill. When anticipating the effect of each move $\mathcal{M}$ on its obligations, $\mathcal{S}$ notes that any move that fulfills one of the two obligations can be made without penalty; the choice therefore depends on the expected benefits for $\mathcal{S}$ herself of each such move.

If only one obligation $\mathcal{O}$ exists and the moves that fulfill $\mathcal{O}$ happen to be unattractive to $\mathcal{S}$, $\mathcal{S}$ may decide not to reject $\mathcal{O}$ outright but rather to choose a move that changes $\mathcal{O}$'s status to Accepted or Possibly Accepted. (A familiar example of the latter tactic in everyday dialog is an attempt to change the subject inconspicuously when asked a question that one would prefer not to answer.) If $\mathcal{S}$ is lucky, $\mathcal{U}$ will proceed to create new obligations for $\mathcal{S}$—or (in Situation 5) simply to generate more than the necessary number of moves himself—in which case $\mathcal{O}$ may expire without $\mathcal{S}$ having to pay the full penalty that would be associated with its rejection. But this tactic may also backfire: If $\mathcal{U}$ refuses to allow $\mathcal{O}$ to expire, $\mathcal{S}$ will still end up having to fulfill or reject $\mathcal{O}$, and the penalty she paid for the attempt to avoid $\mathcal{O}$ will have been paid in vain.

The possibilities just mentioned show that, in order to juggle obligations optimally, $\mathcal{S}$ would actually have to anticipate how $\mathcal{U}$ will respond to a move that creates an intermediate status for an obligation. For example, suppose that $\mathcal{U}$ has just asked an obviously important question and that $\mathcal{S}$ anticipates that $\mathcal{U}$ will keep repeating the question until he gets a direct answer. In this case, $\mathcal{S}$ should not even try in the first place to avoid answering the question. The explicit anticipation of the user's next move has not been realized in MICROPRACMA, but it is handled in a general way within the full PRACMA system with the technique of *global anticipation feedback* (Ndiaye & Jameson, 1996). Accordingly, integration of MICROPRACMA's new techniques into the PRACMA system will allow the use of this strategy in connection with obligation management to be explored.

By not anticipating $\mathcal{U}$'s next moves, in effect MICROPRACMA embodies a short-sighted approach to the problem of escaping from unwanted obligations: All other things being equal, it chooses the move that yields the lowest *immediate* penalty. Whether this move turns out to be the best one for $\mathcal{S}$ in the long run depends partly on chance and partly on the dialog skills of $\mathcal{U}$. This short-sighted approach may be typical of human dialog behavior in cases where anticipation of the dialog partner's moves is either too error-prone or too resource-consuming.

Situation 4: Taking Into Account the Obligations of Both Dialog Participants

Definition of the Situation

$\mathcal{S}$ now takes into account the fact that $\mathcal{U}$, as well as $\mathcal{S}$, faces discourse obligations that arise during the dialog.

Theoretical Considerations

The discourse obligations faced by $\mathcal{U}$ are basically the same in nature as those faced by $\mathcal{S}$. They are created mainly by $\mathcal{S}$'s moves and to a lesser degree by those of $\mathcal{U}$ himself. Moreover, since the moves that create obligations are perceived by both dialog participants, it is not fundamentally more difficult for $\mathcal{S}$ to keep track of $\mathcal{U}$'s obligations than it is for $\mathcal{S}$ to keep track of her own.

But why is it in fact necessary for $\mathcal{S}$ to monitor $\mathcal{U}$'s obligations? There are three main reasons:

1. When $\mathcal{S}$ considers a potential move $\mathcal{M}$ that would create an obligation $\mathcal{O}$ for $\mathcal{U}$, $\mathcal{S}$ should take into account the *cost* to herself of imposing $\mathcal{O}$ on $\mathcal{U}$.
   A speaker cannot in general afford to impose arbitrarily many obligations on her dialog partner. Examples illustrating this point will be given in the next subsection.

2. An obligation for $\mathcal{U}$ which is created by $\mathcal{S}$'s move $\mathcal{M}$ will influence $\mathcal{U}$'s future dialog moves and so influence the ultimate desirability of $\mathcal{M}$ for $\mathcal{S}$.
   Consequently, if $\mathcal{S}$ assesses the expected benefits of its own moves by anticipating $\mathcal{U}$'s subsequent moves, $\mathcal{S}$ must take $\mathcal{U}$'s new obligations into account when performing the anticipation.
3. $\mathcal{S}$'s interpretation of a move by $\mathcal{U}$ must take into account the extent to which $\mathcal{U}$'s selection of this move was influenced by $\mathcal{U}$'s obligations.
   More specifically, there is an important distinction between a *spontaneous* move by $\mathcal{U}$—i.e. one that did not fulfill any obligation of $\mathcal{U}$—and a move by $\mathcal{U}$ that was influenced by an obligation. For example, the utterance by $\mathcal{U}$ "My name is Bill" creates an obligation for $\mathcal{S}$ to introduce herself—but this is only fully the case if $\mathcal{U}$'s move is spontaneous; not, for example, if it is a response to $\mathcal{S}$'s utterance "My name is Sue".

Realization in the System

MICROPRACMA maintains two tables of obligations, which represent the obligations of $\mathcal{S}$ and $\mathcal{U}$, respectively. The procedures for processing obligations are basically the same whether it is those of $\mathcal{S}$ or those of $\mathcal{U}$ that are involved. In fact, in the implementation, exactly the same procedures are used, with a parameter specifying whose obligations are involved. The bottom part of Figure 1 shows some examples of obligations that can only be faced by the user $\mathcal{U}$ in the example domain (assuming, as we have in all of our examples, that the user takes the role of the buyer).

The three ways (discussed more abstractly in the previous subsection) for $\mathcal{S}$ to take $\mathcal{U}$'s obligations into account are realized as follows:

1. The obligation-creation methods illustrated in the left-hand side of Figure 1 also specify for each obligation $\mathcal{O}$ that is created the estimated cost of imposing it on the dialog partner (these estimates are not shown in the figure). Such an estimate is based on two factors:
   - How disagreeable the fulfillment of $\mathcal{O}$ is likely to be for $\mathcal{U}$.
     This factor can often only be roughly estimated by $\mathcal{S}$, but it should be taken into account. This factor presumably explains, for example, why a salesperson will hesitate to ask a customer about his income (valuable as this information would be) whereas a question about the customer's place of residence is unproblematic.
   - The penalty that is associated with nonfulfillment of the obligation.
     An obligation that is associated with a high penalty constrains $\mathcal{U}$'s behavior more strongly than does a weaker obligation; it should therefore be imposed only when $\mathcal{S}$ has a good reason to do so. For example, $\mathcal{S}$ may hesitate to ask $\mathcal{U}$ directly about his evaluative reaction to a statement about the car under discussion (e.g., with the question "What do you think of that [= the fact that the car gets 9 miles per gallon]?"). But $\mathcal{S}$ may be willing to impose a weaker obligation of the same sort on $\mathcal{U}$ simply by making the statement and then giving $\mathcal{U}$ the chance to make the next move (cf. the illustration of the obligation class Express Evaluation at the bottom of Figure 1).
2. As was noted above, MICROPRACMA does not explicitly anticipate $\mathcal{U}$'s next move when selecting a move of its own. Such an anticipation is not in general necessary for the determination of a rough heuristic evaluation of a possible move. For example, the computation of the expected utility of a question by $\mathcal{S}$ can be based on the assumption that the question is most likely to be followed by an answer. The fact that such an answer is made likely by a newly created obligation—and the fact that the answer will not necessarily actually be given by $\mathcal{U}$—need not be taken into account explicitly by $\mathcal{S}$ when it computes the expected utility of a question.[6]
3. When $\mathcal{S}$ has to respond to a move $\mathcal{M}$ made by $\mathcal{U}$, one of the first things that $\mathcal{S}$ does is to note $\mathcal{M}$'s impact on $\mathcal{U}$'s obligations (as well as its impact on $\mathcal{S}$'s own obligations). When doing this, $\mathcal{S}$ notes in particular whether $\mathcal{M}$ fulfilled any obligation that was faced by $\mathcal{U}$. If this is not the case, $\mathcal{M}$ is

[6] For a discussion of the advantages and limitations of the anticipation of the user's subsequent moves, see Ndiaye and Jameson (1996).

classified as being *spontaneous*; this classification is taken into account by some of the rules that determine which obligations (if any) a move creates for $\mathcal{S}$. For example, an Expression of Interest by $\mathcal{U}$ (e.g., "Comfort is my highest priority") imposes a stronger obligation on $\mathcal{S}$ if it is spontaneous (cf. Figure 1).

Resulting Behavior

The main consequence of adding the features discussed in this section to MICROPRACMA is to prevent $\mathcal{S}$ from exhibiting certain maladaptive tendencies that it displays in Situations 2 and 3:

1. $\mathcal{S}$ now responds more appropriately to nonspontaneous moves by $\mathcal{U}$.
   For example, exchanges such as the following one no longer occur:
   - ▷ $\mathcal{S}$: `Hello.`
   - ▷ $\mathcal{U}$: `Hello.`
   - ▷ $\mathcal{S}$: `Hello.`
2. $\mathcal{S}$ is less inclined to impose a large number of strong obligations on $\mathcal{U}$.
   To do so would be appropriate only for a dialog participant (such as a police interrogator) whose dominant role allowed her to attach negligible weight to the factor *cost of imposing obligations*.

Situation 5: Taking Turns Shorter or Longer Than One Utterance

Definition of the Situation

Up to now we have been assuming that $\mathcal{S}$ and $\mathcal{U}$ both adhere to an unnaturally rigid system of turn taking: Each participant, after producing a single one-sentence dialog move, allowed the other participant to take the next turn. In general, of course, turns can vary greatly in terms of length. While we will not attempt here to model all of the ways in which turn taking is accomplished (see, e.g., Sacks et al., 1974; O'Connell et al., 1990; Oreström, 1983), we now consider the following possibilities: A speaker who has the turn can keep it as long as he or she likes, producing a sequence of moves that ends with a release of the turn.

Theoretical Considerations

Relationships of Turn Release to Obligations

In this situation there is a new dialog move of the type Release Turn. Furthermore, in the special case where a participant releases the turn before saying anything at all (except perhaps for a back-channel item, not modeled explicitly here), it is convenient to view the participant as having performed the move Pass Up Turn (cf. Traum & Hinkelman, 1992).

One issue that arises is: How do these two new move types affect the status of a speaker's existing obligations? (Without loss of generality, we consider for concreteness the case where the speaker is the system $\mathcal{S}$.)

Pass Up Turn appears to signal the rejection of any obligation $\mathcal{O}$ that $\mathcal{S}$ faces at the moment. After all, $\mathcal{S}$ has just been given a chance to fulfill $\mathcal{O}$, and she has neither done so nor made any other move that could lead to an intermediate status for $\mathcal{O}$ (cf. Situation 3). Note, for example, that it is only in extreme cases that people remain completely silent when they are greeted or asked a question and then given the turn—however disagreeable it might be for them to give a response.

Release Turn, by contrast, seems in itself to have no effect on the status of existing obligations. Consider, for example, the following exchange:

- ▷ $\mathcal{U}$: `I consider reliability extremely important.`
- ▷ $\mathcal{U}$: `I consider comfort important.`

▷ $\mathcal{U}$: [Release Turn]
▷ $\mathcal{S}$: The car just recently passed its inspection.
▷ $\mathcal{S}$: [Release Turn]

When $\mathcal{S}$ releases her turn here, she has not yet addressed her obligation to say something about the attributes of the car that are related to comfort. Yet her failure to do so, even though she could have, doesn't seem to constitute a rejection of the obligation. She is simply giving $\mathcal{U}$ a chance to say something, which he might well want to do at this point. It is only if $\mathcal{U}$ passes up his new turn and $\mathcal{S}$ continues to remain silent that her silence will signify rejection of the remaining obligation.

Assessing the Benefits of Turn Release

A second issue is: What criteria should $\mathcal{S}$ apply when evaluating the benefits associated with the moves Pass Up Turn and Release Turn? In other words, when should a speaker decide that it's in her best interests just to keep quiet and let her dialog partner speak? This may seem like an elementary question; but it appears to have attracted little attention in spite of the vast amount of research that has been done on turn taking.

One approach is for $\mathcal{S}$ to anticipate what $\mathcal{U}$ will say if given the turn now and to evaluate $\mathcal{U}$'s anticipated move according to $\mathcal{S}$'s own criteria: If $\mathcal{U}$'s anticipated move seems more desirable than the best one that $\mathcal{S}$ could produce, this fact should incline $\mathcal{S}$ to release the turn. Once again, the reader is referred to the paper by Ndiaye and Jameson (1996) for a discussion of this type of approach.

A simpler strategy focuses on one important consequence of a decision by $\mathcal{S}$ to release the turn: the consequence that $\mathcal{U}$ is given a chance to address any existing discourse obligations that he might have. By contrast, as long as $\mathcal{S}$ continues to speak, $\mathcal{U}$'s obligations will be expiring. On the whole, it is in $\mathcal{S}$'s interest to allow $\mathcal{U}$ to fulfill $\mathcal{U}$'s obligations; indeed, $\mathcal{S}$ will often have intentionally created these obligations in the expectation of ultimate benefits. Accordingly, the total weight of $\mathcal{U}$'s current obligations can be used as a simple index of the desirability for $\mathcal{S}$ of giving up the turn.

Realization in the System

The main change to MICROPRACMA required for this situation is the introduction of the two new special moves Release Turn and Pass Up Turn. Both of these moves signal to the top-level control algorithm that the next move is to be selected by the other party. When $\mathcal{S}$ is selecting a move, Pass Up Turn is one of the candidate moves if $\mathcal{S}$ has just received the turn, whereas Release Turn is among the possible moves otherwise. In connection with the tracking of the expiration of obligations, these two moves are viewed as taking no time.

The effects of these moves on existing discourse obligations are taken into account by straightforward changes to the methods for updating existing obligations.

The above-mentioned strategy for assessing the benefits of these moves is realized as follows: $\mathcal{S}$ computes the sum of the Penalty per Moves of all of $\mathcal{U}$'s current obligations, and this quantity serves as an index of the total weight of these obligations, which in turn determines the assessed benefit of releasing the turn. The weight that this index has in the evaluation of $\mathcal{S}$'s possible moves is a parameter that can take various values, which lead to different dialog styles as noted in the next subsection.

Resulting Behavior

If MICROPRACMA is run with a low value for the parameter just mentioned, $\mathcal{S}$ starts each dialog with a long turn in which (among other things) it asks every question and makes every statement that promises to yield any benefit at all. The human user cannot keep track of all of the obligations that these dialog moves create for him. And in any case, by the time he has a chance to say something he no longer feels bound by the obligations, except perhaps for the very most recent ones.

If this parameter is set to a very high value, $\mathcal{S}$ releases the turn as soon as $\mathcal{U}$ has incurred even a weak obligation. For example, every time $\mathcal{S}$ makes a statement about a particular car, $\mathcal{S}$ releases the turn so as to give $\mathcal{U}$ a chance to express an evaluative reaction. Turns of $\mathcal{S}$ that consist of several moves can still occur if the moves before the last one preceding Release Turn don't create any obligations for $\mathcal{U}$.

If the parameter is set to an intermediate value, $\mathcal{S}$ will occasionally create more than one obligation for $\mathcal{U}$ within a turn, sometimes even letting the first obligation expire (e.g., asking a sequence of two questions, of which the second effectively replaces the first).

Concluding Remarks

The topic of discourse obligations has attracted less attention than it deserves from researchers interested in the pragmatic aspects of natural language generation in dialog. The conceptualization presented here appears to capture some fundamental facts about the ways in which dialog obligations are created, modified, and taken into account. Though we will continue to develop this conceptualization and to apply it within a different domain (Wahlster et al., 1995), we hope that other researchers will feel stimulated to propose alternatives concerning either the details or the basic assumptions of the conceptualization.

Acknowledgments

The authors are indebted to Joep Simons for fruitful discussions about the topic of this paper. Insightful comments by participants in the Symposium "Conceptual and Semantic Knowledge in Language Generation" (Heidelberg, November 1995) influenced both the content of this manuscript and the development of the model that it presents.

References

Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. H., Kato, T., Light, M., Martin, N. G., Miller, B. W., Poesio, M., & Traum, D. R. (1995). The TRAINS project: A case study in building a conversational planning agent. *Journal of Experimental and Theoretical AI.*

Jameson, A., Schäfer, R., Simons, J., & Weis, T. (1995). Adaptive provision of evaluation-oriented information: Tasks and techniques. In C. S. Mellish (Ed.), *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* (p. 1886–1893). San Mateo, CA: Morgan Kaufmann.

Ndiaye, A., & Jameson, A. (1996). Predictive role taking in dialog: Global anticipation feedback based on transmutability. In S. Carberry & I. Zukerman (Eds.), *Proceedings of the Fifth International Conference on User Modeling* (p. 137–144). Boston, MA: User Modeling, Inc.

O'Connell, D. C., Kowal, S., & Kaltenbacher, E. (1990). Turn-taking: A critical analysis of the research tradition. *Journal of Psycholinguistic Research, 19,* 345–373.

Oreström, B. (1983). *Turn-taking in English conversation*. Lund, Sweden: Liber.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufmann.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking for conversation. *Language, 50,* 696–735.

Traum, D. R., & Allen, J. F. (1994). Discourse obligations in dialogue processing. *Proceedings of the Thirty-Second Annual Meeting of the Association for Computational Linguistics*, Las Cruces, NM, 1–8.

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

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge, England: Cambridge University Press.

Wahlster, W., Jameson, A., Ndiaye, A., Schäfer, R., & Weis, T. (1995). Ressourcenadaptive Dialogführung: Ein interdisziplinärer Forschungsansatz [Resource-adaptive dialog: An interdisciplinary perspective]. *Künstliche Intelligenz, 9*(6), 17–21.