A survey of knowledge sources in dialogue systems

Annika Flycht-Eriksson
Department of Computer and Information Science
Linköping University, SE-581 83, LINKÖPING, SWEDEN
annfl@ida.liu.se

Abstract

Dialogue systems utilise a variety of knowledge sources and models. However, there is a confusion concerning the purposes and contributions of specific models and the relationships among them. In this paper we present a study of different dialogue systems and the knowledge sources and models they use. The models are characterised in terms of what knowledge they contain and the roles of various models and the relations between them are discussed. Implications for development of dialogue systems are also presented.

1 Introduction

Today much research is carried out within the area of natural language interfaces and dialogue systems, and more and more systems are becoming publicly available. Most of the systems provide information retrieval services or assistance in solving a specific task. The systems differ in complexity due to the domain and the approach taken in the design. Some systems are very knowledge intensive and contain several interacting knowledge sources and models, others rely on much simpler models and procedures. The variety of dialogue system architectures that incorporate various models, has led to confusion when it comes to the purpose and contributions of a specific model. The relations between various knowledge sources and models are also diffuse. In order to clarify the situation, we will examine the following issues:

- What characterise the knowledge sources commonly used in dialogue systems?
- What role do the different knowledge sources and models have?
- What are the relations between the different knowledge sources and models?
- What are the implications for the design of dialogue systems?

We have restricted the survey of knowledge sources to dialogue, discourse, task, domain and user models. Although there exist other models as well, e.g. models used for argumentation [Zukerman and McConachy, 1993] and intelligent tutoring [Rangemalm, 1996], this selection of models reflect the most common types of knowledge sources utilised in dialogue systems.

Dialogue systems are developed for various reasons, e.g. to be used commercially or for research purposes. They can be classified along a number of dimensions such as grammar-based, plan-based, general purpose, domain specific, multi-modal, information retrieval and task planning. We will present a study of how knowledge sources and models are utilised in seven different dialogue systems: GALAXY, LINLIN, RAILTEL, SUNDIAL, TRAINS, VERBMOBIL, and WAXHOLM. The purpose is not to classify the systems, but to illustrate the knowledge and reasoning that is used in dialogue systems of different types.

The GALAXY system, is a distributed multi-modal multi-domain simple-service system that provides users with information about for example travel and weather [Goddeau *et al.*, 1994; Seneff *et al.*, 1998]. The LINLIN system is also a dialogue system used for information retrieval [Ahrenberg *et al.*, 1990], which has been customised for various domains, e.g. information of second-hand cars, charter trips to the Greek archipelago, and more recently timetable information for bus traffic [Flycht-Eriksson and Jönsson, 1998].

The RAILTEL system [Bennacef *et al.*, 1996] and the SUNDIAL system [Bilange, 1991], both give information over the telephone about a specific domain, railway transportation and air traffic information respectively. These systems differ in one aspect, the RAILTEL system is an example of a practical system that is in use while SUNDIAL is a large research project. The WAXHOLM system provides users with information about boat traffic in the Stockholm archipelago. It has much in common with RAILTEL, it is also a domain specific spoken dialogue system originating from the speech research community. However, WAXHOLM also allows for limited multi-modal interaction [Carlson and Hunnicut, 1996].

The TRAINS system is used for mixed-initiative collaborative planning in the transport domain [Allen *et al.*, 1995]. It differs from the other systems presented above in many respects. First of all, it is task-oriented, giving the user advice on how to perform a real world task, while the other systems can be described as simple-service systems [Hayes and Reddy, 1983] that retrieve and present information requested by the user. Furthermore, the TRAINS system is plan-based while most of the other systems are grammar-based. The VERBMOBIL system is also very different from the other systems as it is not a traditional dialogue system. The system does not participate actively in the dialogue instead it listens in on two persons having a conversation in a language that is not their mother tongue, i.e. a German and a Japanese speaking English. The system's task is to monitor the dialogue and to give the users translations when needed [Alexandersson *et al.*, 1994].

2 Characteristics of knowledge sources and models

As stated above dialogue systems utilise various knowledge sources and models. In this section the models are characterised in terms of which knowledge they represent in the dialogue systems.

2.1 Discourse and dialogue models

A discourse is a sequence of connected utterances, for instance a text, a dialogue, or a mixture of both. Characteristic features of a discourse is that it can be divided into segments, there exist co-references between segments, and it is coherent [Grosz, 1995].

In dialogue systems two aspects of the dialogue need to be modelled, a generic description of how the dialogue can be constructed, and a representation of the resulting dialogue. We will refer to the first as the dialogue model and the second as the dialogue history. The dialogue model is often closely connected to and dependent on the information represented in the dialogue history.

Dialogue models

The general information about the construction of a dialogue, held by the dialogue model, is often based on a representation of relations between the constituents of a dialogue. This knowledge is used to control the dialogue, i.e. to decide what action to take in a certain situation.

Two different approaches to dialogue modelling commonly discussed are dialogue grammars and plan-based models of dialogue (e.g. [Cohen, 1995; Jönsson, 1993]). Dialogue grammars have a rather long history and are based on the notion of adjacency pairs. Adjacency pairs express the fact that speech acts typically form a regular sequence such as a question followed by an answer. Rules in the dialogue grammar capture the sequential and hierarchical constraints of dialogues in the same way grammar rules describe the structure of a sentence.

Plan-based models go beyond the utterance and try to model the speakers intentions and goals. Communication becomes a part of the speakers over-all behaviour. Elements in these models are plans, actions, mental states and mechanisms for recognising a specific plan and for reasoning about the speakers beliefs, intentions, and actions.

Four kinds of dialogue models, which incorporate the distinction between grammar-based and plan-based dialogue modelling, can be identified among the presented systems. RAILTEL, SUNDIAL, LINLIN, and WAXHOLM have a dialogue model consisting of a *dialogue grammar* that models the dialogue's hierarchical structure.

In RAILTEL the dialogue is divided into three phases each consisting of a number of subdialogues. It is modelled as a hierarchical structure of subdialogues and dialogue acts, represented by a set of rules in a dialogue grammar [Bennacef *et al.*, 1996]. The same approach is taken by SUNDIAL [Bilange, 1991] and LINLIN [Jönsson, 1997], which model the dialogue using a dialogue grammar and speech acts.

The dialogue model in WAXHOLM is based on the idea that the dialogue should be described as a grammar and at the same time be probabilistic [Carlson *et al.*, 1994]. WAXHOLM thus has a dialogue grammar represented as finite state machines but also a *statistical* component for topic prediction which is part of the dialogue model.

A combination of approaches is also utilised in VERBMOBIL. The dialogue module in VERBMOBIL consists of three submodules: a statistical module, a finite state machine, and a dialogue planner. The interaction is represented in a dialogue model using dialogue acts that can be combined in a limited number of sequences. The statistical module can, given a dialogue state, predict all possible successive dialogue acts and their likelihood. The finite state machine has a similar task, it checks whether a dialogue act is consistent with the underlying dialogue model and which subsequent dialogue acts are possible. The dialogue planner is the most sophisticated of the three submodules. Plan operators are utilised to plan the dialogue and handle different phases like negotiations, clarifications and repairs. A plan is built up hierarchically with the leaves in the hierarchy being dialogue acts. This approach resembles those using a dialogue grammar [Alexandersson *et al.*, 1994].

TRAINS takes a distinct *plan-based* approach to dialogue modelling. The system is modelled as a conversational agent having goals, intentions, plans, and obligations. Answers to user utterances are planned by the system in order to reach it's goals and at the same time fulfil it's obligations [Traum, 1996].

The fourth way to model dialogue is to do it more *procedurally* as in GALAXY where the dialogue is controlled by a stepwise process. The process involves interpretation and disambiguation of user utterances, clarification request when information is missing, database queries, and updates of the dialogue history [Seneff *et al.*, 1996].

Dialogue histories

The dialogue history in a dialogue system represents the state of the dialogue, that is, what has been talked about and what is being talked about at the moment. It is used for dialogue control, disambiguation of context dependent utterances, and context sensitive interpretation, e.g. reference resolution and handling of ellipses.

The representations of the dialogue history range from complex hierarchical structures containing a variety of information to much simpler sequential representations. For example LINLIN, VERBMOBIL and SUNDIAL captures several different levels of information.

In LINLIN a dialogue tree consisting of dialogue objects is constructed during the interaction [Jönsson, 1997]. The tree has three levels, which correspond to the whole dialogue, discourse segments called initiative response units, and speech acts. A dialogue object contains situation parameters and content parameters. The first corresponds to the illocutionary type and the latter models the focus, or the attentional state using Grosz and Sidner's terminology [Grosz and Sidner, 1986].

The representation of the dialogue history in SUNDIAL is distributed over several models. A tree is used to represent the dialogue structure. It resembles the dialogue tree in LINLIN but has one more intermediate level to represent a common topic [Bilange, 1991]. An interpretation is called a belief and a sequence of beliefs which corresponds to the user's utterances are represented in a contextual model [McGlashan *et al.*, 1992]. It is used to make context sensitive semantic interpretations of user utterances. Changes in the contextual model are reflected in a flag *status* that indicates how the contextual model has changed, the value repeat, for example, means that nothing new has been contributed. This information can be used by the dialogue module to reason about the function of an utterance, e.g. if it is a confirmation or a new request [Heisterkamp *et al.*, 1992].

The dialogue planner in VERBMOBIL contains a dialogue memory representing intentional, thematic and referential information. Intentional information corresponds to the dialogue acts performed by the participants and are structured in a tree-like representation. The thematic information refers to relevant information in utterances while referential information is the lexical realisation of utterances. A representation of the proceeding dialogue is constructed dynamically in the dialogue memory during the dialogue [Alexandersson *et al.*, 1994].

Some of the other systems focus only on the objects that has been mentioned, that is the attentional level of discourse structure. GALAXY, RAILTEL, and WAXHOLM all maintain a history of semantic frames that are used to represent the objects and relations in focus. The discourse module in GALAXY maintains a history table of referable objects to facilitate the interpretation process. Items and requests are represented internally as semantic frames, and the history table consists of a sequence of frames [Seneff *et al.*, 1996]. In RAILTEL the context is represented by a dialogue history and a generation history that contain the semantic frame representation of the user's and the system's previous utterances respectively [Bennacef *et al.*, 1996]. The utterances are stored sequentially within these. WAXHOLM utilises a dialogue history based on the semantic frame representation and the updates of this representation [Carlson *et al.*, 1994].

The dialogue history in TRAINS lies somewhere in between the complex multi-level and simple frame-based representations, as it models both attentional and intentional features. The discourse is modelled as a stack of discourse units (DUs) that represent utterances and the corresponding speech acts. In the model the initiator and state of DUs are also recorded. Two other contextual features that are represented in the system are the *turn* and the *local initiative* each of which are said to be held by one of the dialogue participants. The turn indicates who is speaking now, while the local initiative tells which speaker is obliged to speak next [Traum, 1996].

The variety of ways to model discourse information has also led to a variety of names, e.g. discourse memory, discourse history, dialogue memory, dialogue history, history table, and context model. We propose that to avoid confusion only the term dialogue history should be used. Dialogue histories could then be described as partial or full depending on the number of information levels. The degree of structure could also be used to further describe the model.

2.2 Domain models

Domain models hold knowledge of the world that is talked about. Information from the domain model is primarily used to guide the semantic interpretation of user's utterances; to find the relevant items and relations that are discussed, to supply default values, etc. The knowledge represented in a domain model is often coupled to the background system, e.g. a database system. In such cases the domain knowledge is used to map information in the user's utterances to concepts suitable for database search.

The amount of domain knowledge needed in a dialogue system differ depending on the domain and the system's task. This means that the domain model can range from a rather simple conceptual model to a full-fledged world model capable of complex reasoning. Dahlbäck and Jönsson [1997] makes a distinction between domain models and conceptual models. The domain model represents the structure of the world and usually comprises a subset of general world knowledge, while the conceptual model represents the conceptual relationships between the objects in the domain. They claim that often it is enough with one of the models but in a few cases both are necessary. A study of the systems presented in this paper confirms this claim.

RAILTEL, SUNDIAL and VERBMOBIL utilise a conceptual model and some simple inference rules. The semantic frames used in the RAILTEL system contains the relevant domain concepts and serve as a simple conceptual model. The domain knowledge also consists of two kinds of rules: *Default value rules* supply default values, e.g. the current or next month for a departure date; *Interpretative rules* transform vague qualitative values into more precise quantitative values used by the system, for example map the concept "morning" onto the precise interval "between 6 am and 12 noon" [Bennacef *et al.*, 1996].

The domain knowledge in the SUNDIAL system is distributed. One module contains a hierarchy of surface-oriented concepts while another module contains a hierarchy of task-oriented concepts. The surface-oriented concepts, for example "at noon", are mapped onto task-related concepts, "at twelve o'clock". The domain knowledge thus consist of the two concept hierarchies and inference rules used to map one to the other [Heisterkamp *et al.*, 1992].

Domain knowledge in VERBMOBIL is represented in a conceptual hierarchy that represents relations between different categories, entities, and situations. Situations are entities determined by spatial and temporal features like year, month, week, day, location, etc. The hierarchical structure of the model makes it suitable for decisions about possible references. It also allows inheritance, if a temporal object is available for scheduling the superordinate object is regarded as free too, similarly, if an object is not available none of the subordinated objects are either [Quantz *et al.*, 1994].

TRAINS on the other hand has a more complex domain/world model. The domain plan reasoner in the system is responsible for the representation of and reasoning about the domain. It maintains a representation of the state of the world, gives new suggestions about routes, and makes adjustments to the current plan given new constraints. Inspection of plans that have been suggested by the user is also done to check if there are any unknown conditions that has to be considered [Ferguson *et al.*, 1996a].

In GALAXY and LINLIN conceptual and domain models are combined for some domains. The domain knowledge in the GALAXY system can be found in two places, declarative tables in the discourse module, and in domain servers that contain the application data. The tables describe the various possible semantic classes a value can have and some relations between them, and can be seen as a conceptual model. The domain servers can contain more specific domain knowledge necessary to interpret the semantic frame [Seneff *et al.*, 1996].

In the application areas of information retrieval on second hand cars and charter trips the domain model in LINLIN is a rather simple representation of relations between domain objects [Jönsson, 1993]. However, the time-table domain requires more complex domain knowledge about the geographic and temporal aspects of bus traffic. This knowledge is modelled in a spatial and a temporal reasoner which are co-ordinated by a component called the knowledge co-ordinator [Dahlbäck *et al.*, 1999]. LINLIN is also extended with a conceptual model for this domain [Dahlbäck and Jönsson, 1997].

The WAXHOLM system contains no explicit domain model, domain knowledge is instead incorporated in the lexicon and the grammar. The parser is based on a semantic feature structure with features of two kinds, basic features and function features. The basic features, e.g. boat, place, are organised in a hierarchy and provides simple semantic information about a word. The function features, e.g. departure_time, to_place, does not have the same structure and they are associated with actions rather than objects [Carlson and Hunnicut, 1996].

2.3 Task models

The terms task and task model are often used when describing dialogue systems, but they can refer to very different phenomena. It is important to make a clear distinction between the system's task(s) and the user's task(s). A user task is non-linguistic and takes place in the real world. Models of such tasks involve the user's goals and how they can be achieved. Models of system tasks describe how the system's communicative and other tasks, e.g. database access, are carried out. A system task model can consist of a hierarchical representation of subtasks which captures the task structure. In the case of information retrieval systems it can also be a much simpler specification of the information the system has to collect before a database access.

Dahlbäck [1996] uses the term *dialogue-task distance* to describe the degree of connectedness between the user's non-linguistic task and the dialogue structure. For some types of dialogues it is important that the system knows what task the user is performing or is planning to perform. In these cases the system can often infer the necessary information from the linguistic information. For other types of dialogue, where the dialogue-task distance is long, information about the user's task is unnecessary.

van Loo and Bego [1993] classify dialogue systems based on the relation between the user task and the system tasks. In *task dialogues* the system's role is to guide the user when she is performing a task in the real world. This requires a model of the user task. *Planning dialogue* are similar to task dialogues, the system supports the user in the planning of a task execution, which also requires a model of the user task. *Parameter dialogues* differ from task and planning dialogues as the system can provide the user with information without any knowledge about what real world task it is assisting.

In this survey TRAINS is the only system which is strictly task-oriented and has a model of the user's task. The task in TRAINS is route planning and knowledge about the task, i.e. the planning process, is represented explicitly. The plans that are developed through the interaction with the user are represented as a hierarchy of goals that are expanded into subgoals. The hierarchy is complemented with a linear history of the planning process. For some subproblems specialised domain servers are used [Ferguson *et al.*, 1996b].

The other systems are all some sort of simple-service systems providing users with requested information retrieved from databases, and falls in the category of parameter dialogues. In VERBMOBIL and WAXHOLM the system task knowledge is integrated and lies implicit in the system. GALAXY, RAILTEL and LINLIN have a frame-based specification form that is more or less utilised as a system task model. In GALAXY there is no explicit task model but a frame representation is used to see if all the required slots are filled and in case some are missing ask the user for a clarification [Seneff *et al.*, 1996]. In a very similar way task rules in the grammar of the RAILTEL system are connected to the semantic frame representation of the provided information and subdialogues are initiated if there is some missing information in the frame [Bennacef *et al.*, 1996]. Models of the system tasks are explicit in LINLIN since different information needs has to be handled in the time-table domain. A task model in LINLIN consists of a description of the required information pieces, called an Information Specification Form [Dahlbäck and Jönsson, 1999]. The dialogue manager utilises this information to ask follow up questions if not all the necessary information has been supplied by the user.

SUNDIAL utilises a more sophisticated system task model that represents the task structure and keeps track of the current status of the task. This often means deciding whether the user has given enough information and if not, what has to be further provided. Another role is to handle situations that arise when the provided information is incorrect. In this case the task model is used to relax the parameters instead of returning a negative answer. For example, if the user has requested a flight at noon and none is available, the system tries to find one in the morning instead [McGlashan *et al.*, 1992].

2.4 User models

User models represent the user's goals and plans, capabilities, attitudes, and knowledge or belief. They vary in complexity, ranging from user stereotypes to complete models of the user's knowledge, intentions, attitudes, etc. [Kass and Finin, 1988].

Depending on what kind of information a user model contains it can be used for various purposes. If a user model represents what the user knows about the domain, the system can adapt it's answers so that they are informative and easily understandable. User models are also utilised for dialogue control. They are not that common in dialogue systems today.

The TRAINS system uses a very elaborated user model but also a self model. The user model and self model represent the user's and system's mental state and contain information about the beliefs and proposals about the domain. The beliefs can be private to either the system or the user, or they can be shared by both. Mutual beliefs include aspects of the domain plans that are considered to be agreed on by both system and user. Beliefs the system supposes the user to hold are the domain plans that have been proposed but not acknowledged. The systems private beliefs consists of the domain plans the domain planner has derived but which have not been presented to the user [Traum, 1996].

3 The role of different models

In the survey so far we have tried to characterise the different knowledge sources and models commonly used in dialogue systems. The focus has been on what kind of knowledge they contain, in this section the focus is shifted to how the different models are used, i.e. what their roles are in a system.

3.1 Dialogue models

Dialogue models have the common purpose of describing how the system should respond to user utterances, for example, by accessing a database or asking for clarifications. The different approaches to dialogue modelling, i.e. dialogue grammar versus plan-based, reflect the various kinds of behaviour one wants the system to have and the types of dialogue it should be able to handle. For advisory or task-oriented dialogues a plan-based model might be appropriate, in which the user's intentions and goals are represented. For these types of dialogues the dialogue-task distance is short, i.e. there is a close connection between the task and the language, which means that it is possible to infer the non-linguistic intentions behind an utterance. In simple human-computer interaction for information retrieval the dialogue-task distance is longer. The user task, for example to travel by bus to a wedding the next Sunday, is not necessarily reflected in the dialogue that seemingly only consists of information exchanges. This makes it harder to infer the underlying intentions if one wants to model these in the system. On the other hand that type of information is not necessary for the system to be able to respond appropriately, a dialogue grammar is adequate for this type of interaction [Dahlbäck, 1996].

The kind of dialogue history required in a dialogue system depends on the linguistic phenomena that has to be handled, such as misunderstandings, interruptions, and deictic expressions, and the complexity of the task and domain which is reflected in the interaction. If a dialogue is made up of many small subdialogues a more structured representation may be preferred to a sequential.

3.2 Domain models

A domain model is in many cases necessary to make the dialogue system natural and intuitive to use. The character of the domain knowledge can differ from conceptual to realistic world models. With the help of conceptual domain knowledge the system can interpret and in some way understand user utterances. Domain knowledge related to the world can be used to reason about the domain and in some sense makes the system smarter. For example, in the time-table domain of LINLIN, conceptual knowledge is utilised to reason about the relation between concepts like trips, departure times and departure places while the domain model contains geographical knowledge of the domain which can be used to figure out which is the nearest bus stop to a place mentioned by the user [Flycht-Eriksson and Jönsson, 1998]. Sometimes only one of the two types is necessary, e.g. a dialogue system which access fairly simple databases may only need a conceptual model.

3.3 Task models

The dialogue manager in a dialogue system is responsible for deciding what to do given the user's input and the context. This decision is highly dependent on the status of the user task or subtask. Some possible ways for the dialogue manager to handle this is to use an explicit model of the user task. User task models can vary in complexity depending on the amount of information that has to be exchanged, the structure of the task, and the negotiation necessary. The structure or lack of structure is important. If a task is ill-structured, the system can not predict which kind of information the user will request and when. In a well-structured task with a predefined, at least partially ordered exchange of information, it is much easier for the system to control the interaction. The presence of an explicit model of the system tasks in a dialogue system, separated from the dialogue model, can make the dialogue more fluent and the system more efficient. In an information retrieval system the system task model can assist the system in judging if all the required parameters are present and in case where they are not, determine what to ask the user for. The use of an explicit system task model makes the system more flexible since it's easier to modify or add new tasks.

3.4 User models

Kass and Finin [1988] discuss when user models can be of great assistance and when they can be left out of a system. In simple question-answering systems no user model is really necessary. In co-operative question answering systems user models can be more beneficial. For such systems a simple model of the user's goal can be used but it suffices with a generic model. Co-operative consultation systems on the other hand need an extensive user model.

4 Relations between models

One problem with clearly defining the roles of the knowledge sources and models are that sometimes different models can be used for the same purpose, achieving the same functionality. In this observation lies an implication that sometimes two or more models can be collapsed into one. This question is addressed in this section which discusses the relations between the different models. We focus on the relation between user and discourse models[1], and among dialogue, domain, and task models.

4.1 User models and discourse models

The relation between user models and discourse models has been debated. Opinion range from user models and discourse models being completely separate to being two sides of the same coin [Kobsa, 1988; Schuster, 1988; Cohen, 1988]. The differences in opinion seems to a great extent to depend on how the two types of models are defined. Some researchers focus on the user's goal and compare it to the discourse' intentional level, others look at the user's beliefs and knowledge and compare it to the attentional information in discourse models.

Even though there is no consensus about the relation between user models and discourse models, one possible way to separate them would be to conclude that they have different purposes and different time spans. User models are long term knowledge sources that are used to decide how the system should respond to a user while discourse models only last during a session and are utilised for context sensitive interpretation and generation in the conversation.

4.2 Domain, task and dialogue models

The least discussed and also the most problematic, are the relations between domain, task and dialogue models. To be able to interpret user utterances and decide how to continue a dialogue, the system has to have knowledge about both the domain and the task. Thus, it can be very tempting to integrate this kind of knowledge in the dialogue model. The task and domain knowledge can also get mixed up since the performance of a task often is domain dependent.

In task-oriented dialogue systems the task and domain knowledge easily get intertwined since the user's planning or execution of the task partly constitute the domain. It is however not necessary to separate them as they are used for the same purpose. The domain model can thus be incorporated in the user task model or vice versa. A compromise is otherwise to have a user task model and complement it with a conceptual model.

Another typical example of the integration of domain and task knowledge is the use of the semantic frames, which are present in many simple-service dialogue systems. They represent the relevant domain concepts and are sometimes coupled to rules for interpretation of domain concepts and rules to fill in default values in a frame. Besides the role of conceptual models the semantic frames often serve as system task models since they are used to describe what information has to be gathered by the system before a database access. This multiple use of semantic frames is useful as long as the system task is rather simple and well-structured. If the task gets more complex, separate system task models might be needed.

Dialogue systems for information retrieval often lacks an explicit system task model, in these system this knowledge very easily becomes a part of the dialogue model. This works well in simple domains where the system only performs one or a few similar tasks but if a system is to manage many different tasks explicit task models of the types found in SUNDIAL and LINLIN are required.

It is hard to draw conclusions about the relations between dialogue, domain and task models except that the boundaries between them in some cases are very diffuse. For example, an elaborated user task model can make

[1] In section 2.1 we proposed that the term dialogue history should be used for the representation of the state of the dialogue. Discourse model, is however widely used in the literature to describe this kind of knowledge and so we use this term instead of dialogue history here.

a domain model superfluous, and a system task model may be incorporated in the dialogue model. The decision of whether to combine two models or not must be guided by the situation the system will be used in, but it is important that it is an explicit choice and that the restrictions it imposes are considered.

5 Conclusions

The knowledge sources utilised in dialogue systems are not clearly defined. The mix-up of knowledge sources and models has a number of drawbacks, especially for research systems and systems not primarily developed for limited dialogue to one application. First of all it makes it hard and time consuming to port a system to a new domain or task because changes has to be made in many different parts of the system. Another related problem is that it is difficult to experiment with the systems behaviour, for example different dialogue strategies, since a change of some aspect often implicitly causes other changes. The lack of clear boundaries between models also makes it hard to reuse and incorporate previous work done by others.

To handle these problems the design of a dialogue system should be based on an analysis of the system's intended functionality and dialogue behaviour. This analysis should then be used as the basis for the design of knowledge sources and models needed for the dialogue system. This is only possible if the roles and responsibilities of different models are clearly defined. Our survey is a step towards providing such definitions. The next step is to define a taxonomy of dialogue types which could be utilised to predict what kind of knowledge is required to achieve a particular dialogue behaviour. Here some work has been done by van Loo and Bego [1993] and Dahlbäck [1996] but clearly more work is needed.

6 Acknowledgements

This work is supported by The Swedish Transport & Communications Research Board (KFB). Thanks to Nils Dahlbäck and Arne Jönsson for stimulating discussions and comments on previous versions of this paper.

References

[Ahrenberg *et al.*, 1990] Lars Ahrenberg, Arne Jönsson, and Nils Dahlbäck. Discourse representation and discourse management for natural language interfaces. In *Proceedings of the Second Nordic Conference on Text Comprehension in Man and Machine, Täby, Sweden*, 1990.

[Alexandersson *et al.*, 1994] Jan Alexandersson, Elisabeth Maier, and Norbert Reithinger. A robust and efficient three-layered dialogue component for speech-to-speech translation system. Report 50, December 1994.

[Allen *et al.*, 1995] James Allen, Lenhart Schubert, George Ferguson, Peter Heeman, Chung He Hwang, Tsuneaki Kato, Mark Light, Nathaniel Martin, Bradford Miller, Massimo Poesio, and David Traum. The TRAINS project: a case study in building a conversational planning agent. *Journal of Experimental and Theoretical Artificial Intelligence*, 7:7–48, 1995.

[Bennacef *et al.*, 1996] S. Bennacef, L. Devillers, S. Rosset, and L. Lamel. Dialog in the RAILTEL telephone-based system. In *Proceedings of International Conference on Spoken Language Processing, ICSLP'96*, volume 1, pages 550–553, Philadelphia, USA, October 1996.

[Bilange, 1991] Eric Bilange. A task independent oral dialogue model. In *Proceedings of the Fifth Conference of the European Chapter of the Association for Computational Linguistics, EACL'91*, pages 83–88, Berlin, Germany, 1991.

[Carlson and Hunnicut, 1996] Rolf Carlson and Sheri Hunnicut. Generic and domain-specific aspects of the Waxholm NLP and dialog modules. In *Proceedings of International Conference on Spoken Language Processing, ICSLP'96*, volume 2, pages 677–680, Philadelphia, USA, October 1996.

[Carlson *et al.*, 1994] Rolf Carlson, Sheri Hunnicut, and Joakim Gistafsson. Dialog management in the Waxholm system. In *Papers from the Eighth Swedish Phonetics Conference, Working Papers 43*, pages 46–49, 1994.

[Cohen, 1988] Robin Cohen. On the relationship between user models and discourse models. *Computational Linguistics*, 14(3):88–90, 1988.

[Cohen, 1995] Phil Cohen. *Dialogue Modeling*, chapter 6, pages 234–240. In Varile and Zampolli [1995], 1995.

[Dahlbäck and Jönsson, 1999] Nils Dahlbäck and Arne Jönsson. Knowledge sources in spoken dialogue systems. In *Proceedings of Eurospeech'99*, Budapest, Hungary, 1999.

[Dahlbäck *et al.*, 1999] Nils Dahlbäck, Annika Flycht-Eriksson, Arne Jönsson, and Pernilla Qvarfordt. An architecture for natural dialogue systems. In *Proceedings of ESCA Tutorial and Research Workshop on Interactive Dialogue in Multi-modal Systems*, Kloster Irsee, Germany, 1999.

[Dahlbäck and Jönsson, 1997] Nils Dahlbäck and Arne Jönsson. Integrating domain specific focusing in dialogue models. In *Proceedings of Eurospeech'97*, volume 4, pages 2215–2218, Rhodes, Greece, 1997.

[Dahlbäck, 1996] Nils Dahlbäck. Towards a dialogue taxonomy. In *Proceedings of ECAI'96 Workshop Dialogue Processing in Spoken Language Systems*, pages 28–34, 1996.

[Ferguson *et al.*, 1996a] George Ferguson, James Allen, and Brad Miller. TRAINS-95: Towards a mixed-initiative planning assistant. In *Proceedings of the Third Conference on Artificial Intelligence Planning Systems, AIPS-96*, pages 70–77, 1996.

[Ferguson *et al.*, 1996b] George M. Ferguson, James F. Allen, Brad W. Miller, and Eric K. Ringger. The design and implementation of the TRAINS-96 system: A prototype mixed-initiative planning assistant. TRAINS Technical Note 96-5, October 1996.

[Flycht-Eriksson and Jönsson, 1998] Annika Flycht-Eriksson and Arne Jönsson. A spoken dialogue system utilizing spatial information. In *Proceedings of International Conference on Spoken Language Processing, ICSLP'98*, Sydney, Australia, December 1998.

[Goddeau *et al.*, 1994] David Goddeau, Eric Brill, James Glass, Christine Pao, Michael Philips, Joseph Polifroni, Stephanie Seneff, and Victor Zue. Galaxy: A human-language interface to on-line travel information. In *Proceedings of International Conference on Spoken Language Processing, ICSLP'94*, pages 707–710, Yokohama, Japan, September 1994.

[Grosz and Sidner, 1986] Barbara J. Grosz and Candace L. Sidner. Attention, intention and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[Grosz, 1995] Barbara Grosz. *Discourse and Dialogue*, chapter 6, pages 227–229. In Varile and Zampolli [1995], 1995.

[Hayes and Reddy, 1983] Philip J. Hayes and D. Raj Reddy. Steps toward graceful interaction in spoken and written man-machine communication. *International Journal of Man-Machine Studies*, 19:231–284, 1983.

[Heisterkamp *et al.*, 1992] Paul Heisterkamp, Scott McGlashan, and Nick Youd. Dialogue semantics for a spoken dialogue system. In *Proceedings of the International Conference on Spoken Language Processing, ICSLP'92*, Banff, Canada, 1992.

[Jönsson, 1993] Arne Jönsson. *Dialogue Management for Natural Language Interfaces*. PhD thesis, Linköping University, 1993.

[Jönsson, 1997] Arne Jönsson. A model for habitable and efficient dialogue management for natural language interaction. *Natural Language Engineering*, 3(2/3):103–122, 1997.

[Kass and Finin, 1988] Robert Kass and Tim Finin. Modeling the user in natural language systems. *Computational Linguistics*, 14(3):5–22, 1988.

[Kobsa, 1988] Alfred Kobsa. User models adn discourse models united they stand... *Computational Linguistics*, 14(3):91–94, 1988.

[McGlashan *et al.*, 1992] Scott McGlashan, Norman Fraser, Nigel Gilbert, Eric Bilange, Paul Heisterkamp, and Nick Youd. Dialogue management for telephone information systems. In *Proceedings of the International Conference on Applied Language Processing, ICSLP'92*, Trento, Italy, 1992.

[Quantz *et al.*, 1994] Joachim Quantz, Manfred Gehrke, Uwe Kssner, and Birte Schmitz. The verbmobil domain model version 1.0. Technical Report 29, Technishe Universität Berlin, September 1994.

[Rangemalm, 1996] Eva L. Rangemalm. Student diagnosis in practice; bridging a gap. *User Modelling and User Adapted Interaction*, 2(5):93–116, 1996.

[Schuster, 1988] Ethel Schuster. Establishing the relationship between discourse models and user models. *Computational Linguistics*, 14(3):82–85, 1988.

[Seneff *et al.*, 1996] Stephanie Seneff, David Goddeau, Christine Pao, and Joseph Polifroni. Multimodal discourse modelling in a multi-user multi-domain environment. In *Proceedings of International Conference on Spoken Language Processing, ICSLP'96*, pages 192–195, Philadelphia, USA, October 1996.

[Seneff *et al.*, 1998] Stephanie Seneff, Ed Hurley, Raymond Lau, Christine Pao, Philipp Schmid, and Victor Zue. Galaxyii: A reference architecture for conversational system development. In *Proceedings of International Conference on Spoken Language Processing, ICSLP'98*, volume 3, pages 931–934, Sydney, Australia, December 1998.

[Traum, 1996] David R. Traum. Conversational agency: The TRAINS-93 dialogue manager. In Susann LuperFoy, Anton Nijhholt, and Gert Veldhuijzen van Zanten, editors, *Proceedings of Twente Workshop on Language Technology, TWLT-II*, 1996.

[van Loo and Bego, 1993] W. van Loo and H. Bego. Agent tasks and dialogue management. In *Workshop on Pragmatics in Dialogue, The XIV:th Scandinavian Conference of Linguistics and the VIII:th Conference of Nordic and General Linguistics, Göteborg, Sweden*, 1993.

[Varile and Zampolli, 1995] Giovanni Varile and Antonio Zampolli, editors. *Survey of the State of the Art in Human Language Technology.* URL: http://cslu.cse.ogi.edu/HLTsurvey/, 1995.

[Zukerman and McConachy, 1993] Ingrid Zukerman and Richard McConachy. Consulting a user model to address a user's inferences during content planning. *User Modeling and User-Adapted Interaction*, 3(2):155–85, 1993.