Towards Learning Dialogue Structures from Speech Data and Domain Knowledge: Challenges to Conceptual Clustering using Multiple and Complex Knowledge Source

JENS-UWE MÖLLER
Natural Language Systems Division,
Dept. of Computer Science, Univ. of Hamburg,
✉ Vogt-Koelln-Str. 30, D-22527 Hamburg, Germany,
☎ ++49 40 54715 - 516 / Fax: - 515, jum@informatik.uni-hamburg.de

Abstract

This paper introduces an engineering-oriented approach towards dialogue modelling. While dialogue models in existing dialogue systems usually are manually coded, or at least the data on which they are based is manually labeled, we investigate the possibility of learning dialogue models from a large set of example dialogues using real data in a spoken dialogue system. We assume the dialogue system to be front end to a knowledge based system (kbs) . In this scenario, different modules like the kbs itself, lexicon, word recogniser, parser etc. provide information that may contribute to the recognition of dialogue acts and to the generation of system's reaction. Information in our scenario may be ambiguous, probalistic or have a complex structure, and our learning algorithm must be able to process it. Implications on the learning algorithm from these demands are discussed.

Unlike other approaches that use machine learning for learning classification rules for a set of predefined classes from some discourse theory, we want to learn an application specific classification schema from speech data using conceptual clustering. The reason for this is that all theories have their strengths and weaknesses, and combining all of them seems to be overkill for the limited variability of dialogues in a specific application (the engineering viewpoint). Additionally, we expect some new insight to the importance of different discourse phenomena and perhaps results impose some changes to existing theories.

1. Introduction

In this introduction we motivate our approach, define a learning task and choose an appropriate learning algorithm. Section 2 investigates different knowledge sources in an spoken dialogue system and section 3 points out their implications on the learning algorithm. Finally there are some concluding remarks.

Dialogue modelling for natuaral language interfaces has become a topic of substantial interest during the last decade. It combines work on discourse analysis and speech acts with domain tasks and human computer interaction. All of the work on dialogue presumes the existence of a set of units, often called dialogue acts which are building blocks for a dialogue model.
Some work tries to define a set of dialoue acts from a theory of action, contributing to models of mutual knowledge [Coh87, Coh90, Faw89]. Bunt [Bun94] distinguishes dialogue acts contributing to different dimensions of context, like linguistic, semantic,

physical, social and cognitive context. Others dealing with dialogues pay more attention to exchange structures [Wac86, Jön91], whereas some theories made for text structuring, like the rhetorical structure theory (RST) [Man88] do not cover interaction [Moo92]. Many researchers argued that the domain and task structure has an important impact on the discourse structure and thus integrated them into their theories and systems [e.g. Woo84, Gro86, Lam91]. There is some work on integrating different structures into a unique framework [Faw92, Dar94]. The recognition of discourse structures from linguistic structures again is a proper subject of research [Hir93, Lit94, Kno91].
Two approaches for representing dialogue or discourse structure are competing: plan-based models [All82, Lit85, Car91, LaMöl91, Ram91a, You94, Chu94] and discourse grammars [Bil91, Faw89, Bun89, Ste93].
Our prior work has contributed to the engineering of dialogue structures starting out from a knowledge based system [Möl90, Möl92a, Möl95a]. Other engineering-oriented work contributes to dialogue acquisition techniques [Jön93].

1.1. Some problems

When modelling a dialogue for the use in an application's natural language interface, we have experienced, that all these theories provide some means for structuring a dialogue, but none of them well fits to what is needed in a specific application environment [c.f. Jön93]. All theories fail to explain some phenomena, and we have shown in prior work, that there are even domain-specific structures constituting these kinds of domain-specific dialogues [Möl92b, Ram91b]. Most existing dialogue system engineers thus develop their own set of dialogue acts [c.f. Mai94].
Additionally, labeling dialogues is a very time-consuming work, it requires substantial expertise in linguistics and domain knowledge and it usually does not provide reproduceable results, because the classification of a dialogue into units offered by those theories is sometimes ambiguous [c.f. Hir92] and depends on understanding and interpretation by the labeling person.
A further throw-back for the integration of NL frontends into application systems is that the acquisition of appropriate dialogues itself is a complex and long lasting task. Either natural occuring dialogues have to be recorded, transcribed (lots of work!) and afterwards be restricted to a set of possible interactions which are supported by the application system. Alternatively, one may adopt some Wizard-of-Oz experiments to acquire dialogues [Jön93], but this work is restricted to written language. Transcriptions of phonetic and prosodic data are far beyond the possible effort in a dialogue engineering task, thus there is a loss of certain information in analysing transcribed dialogues.

1.2. Steps towards a solution

Notably when thinking about modelling a dialogue for a spoken dialog frontend, we should consider all those information sources which are present in speech processing system.
As a first test environment we used data from the VERBMOBIL project and ascribed rhetorical relations to textual transcriptions. It turned out, that they are not always applicable, notably for interaction phenomena, e.g. when motivating turn shift etc. We compared applicable relations with prosodic transcriptions and discovered correlations

between RST relations and prosodic events and topic shifts, thus supporting the results given in [Hir92].

The question raised, whether it makes sense to develop a new theory that integrates different structures. Our previous work, where we have been confronted with domain-specific structures were one objection to this idea.

Another objection to a new theory was our goal of making dialogue modeling engineerable, i.e. to minimise the effort to build a dialogue model. Litman has successfully shown, that machine learning makes it easy to induce classification rules for cue words. Thus, easy use allows to process larger amount of data which raises accuracy compared to manually abstraction of rules [Lit94, c.f. also Sie94]. Similar results have been reported on discourse level information extraction [Leh94]. One possible approach for building a dialogue model under the objective of efficiency is to use similar supervised machine learning techniques to classify data from dialogues into classes each of which corresponds to some previously known unit from one of the dialogue or discourse structuring theories.

Even if supervised learning would lead to good recognition of units from existing discourse or dialogue theories, we would be confronted with the problem of combining different theories in a way, that we are able to build up a dialogue model which considers interaction aspects as well as discourse structure. As we are directly interested in classes of these combinations[1], we persue another approach.

We will try to abstract new domain and task specific classes of dialogue units directly from data, that are available in a speech processing system. To do this, we need some conceptual clustering algorithm that does unsupervised learning of dialogue units. First tests with the COBWEB algorithm [Fis87] applied to the above mentioned transcribed data from VERBMOBIL were promising and resulted in some clusters according to RST relations.

However the learning environment, i.e. the analysis of available data and their preparation as well as their integration into the learning algorithm was not well designed. The design of the learning task and related problems for the learning algorithm will be the topic of this paper.

1.3. Definition of the learning task

From spoken dialogues we would like to learn dialogue models. A first not very striking idea is to present entire dialogues as cases to a learning algorithm. But being able to classify entire dialogues will not help us building a dialogue system, in which we have to decide, what to answer to a question etc. We should look for smaller units, that might correspond in some way to dialogue acts. A very natural smaller unit is the dialogue turn, i.e. the utterance of one speaker before the turn changes to the dialogue partner. Our task is therefore to learn dialogue acts from turns and to assemble them to a dialogue model later on. For the reasons pointed out above, we do not want to use dialogue acts or some other units from an existing structuring theory and use them for learning a classification tree. Instead we would like to use learning algorithm for inventing new classes of dialogue structuring units and how these classes could be characterised in terms of data that are available in a spoken dialogue system.

---

[1] We intend to follow the suggestion of Litman: "Another advantage of the machine learning approach is that the ease of inducing ruleset from many different sets of features supports an *exploration of comparative utility of different knowledge sources*"

1.4. Requirements to and choice of a learning algorithm

A very nice overview to classify learning algorithm has been given by Fisher[Fis85]. Supervised learning – as it is done by most learning algorithm – is out of the question, because – again – we deliberately will not rely on an existing theory of dialogue structure. Our learning problem is that of aggregating turns or with other words, we do concept formation for new classes of dialogue structuring units.

As we know from many parts of natural language processing different information channels (like topicalisation (syntax) and accent (prosody)) are often combined to express things. This pleads against using monothetic learning algorithms, because they would classify along a single attribute. Instead, we have to use one of the rare polythetic learning algorithms.

Another classifying feature for learning algorithms is the kind of the result, that we achieve. Does learning result in a set of mutually exclusive clusters, a classification tree or a set of clusters which might overlap? At a first glance, we would prefer a learning algorithm that produces a hierarchy, because we will get information about the most important classifying attributes. This may lead to some new insight in dialogue modelling. On a second glance, when thinking about different dialogue structuring factors, it might be useful to have multiple hierarchies with overlapping clusters.

There are not many unsupervised learning algorithms working polythetic. The demand for a hierarchy and directly inspectable classifying attributes rules out some neural network algorithms that do unsupervised learning like Kohonen networks [Rit89, Koh84], competitive learning [Rum80, Gro87] or ART [Car88, Car90]. One of the best known symbolic machine learning algorithm, that fulfills our criteria is COBWEB [Fis87] or its decendent CLASSIT [Gen89], respectively. Thus a COBWEB-like algorithm is the basis for our further work. There are some improvements concerning overlapping clusters and multiple hierarchies, which will be considered [Mar94]. We shall now have a look at the data from which we want to learn.

2. Available data in a speech dialogue system

This sections introduces different data sources in a speech dialogue system that might provide useful information for our task. All data that are available from other modules like syntax, semantics, prosody, word recognizer, lexical semantics and domain knowledge should be considered and their possible contribution will be investigated.

2.1. Using domain knowledge

The structure of the task and of the domain knowledge are major coherence creating factors in goal-oriented dialogue [Möl90, Ram91b, Möl95a]. Knowing these structures is useful for the recognition of a dialogue act as it strictly reduces the search space.

In contrast to the dialogue models that we persue, domain-independent approaches to dialogue modelling tend to be less specific and therefore their recognition guarantees less accuracy. Domain-independent dialogue act recogniser can exclusively be based on linguistic phenomena and do not consider the specificity of a domain-specific sublanguage [Gri86, Leh86].

A domain and task model is an integral part of most kbs. No additional modelling has to be done to use knowledge about the domain and the task for the dialogue module, because it is already available in the system environment.
Task knowledge is typically represented using goals and plans, domain knowledge in natural language systems as well as in some other knowledge based systems is represented by terminological logics. A system that clearly distinguishes different knowledge level and allows different well investigated representations is KARL [Ang91, see also Fen94].

2.2. Data from the word recogniser

A first step in processing written language is to extract words from character strings and to access a lexicon. When processing speech, a word recogniser analyses a signal using pattern recognition techniques and provides word hypotheses or graphs of word hypotheses. Word hypotheses are probalistic and may be ambiguous, but they are certainly one of the basic information source even for recognising dialogue acts. Figure 1 shows the 'best path'-output of a word recogniser system [Alt95] with a logarithmic acoustic score and time intervals that are retrained by forced alignment.

310	460	ja	-76.611984
460	800	prima	-82.753677
800	1010	dann	-87.347595
1010	1230	lassen	-78.228905
1230	1380	Sie	-74.399940
1380	1470	uns	-83.025497
1470	1590	doch	-80.284233
1590	1650	noch	-87.236458
1650	1780	einen	-75.157867
1780	2060	Termin	-83.896736
2060	2580	ausmachen	-80.108749
2580	2760	wann	-81.232422
2760	2860	w"are	-87.095200
2860	2970	es	-78.521584
2970	3210	Ihnen	-74.976280
3210	3330	denn	-84.806023
3330	3700	recht	-70.472916

Figure 1: Output of a word recogniser

2.3. Prosodic data

Besides word recognition, prosodic features like accent or pause can be deduced from the speech signal, too [Nöt91]. Prosody also yields a sentence melody, but this is of less importance in our context. By intonation accents are set and thus important information is stressed and less important things are put to background [Bat93]. Semantic units are often marked by pauses at the beginning and at the end (see Figure 2). These properties are supposed to be clues to understanding of spoken language, whereas in written language properties like correct syntax and more elaborate formultion have to compensate the function of prosody.
In Figure 2 you see the result of a manual prosodic transcription following [Pie80, Uhm91, Sil92, Fér93, Rey94]. In contrast to those certain symbolic data, the output of a accent

recognising system (see Figure 3, [Str94]) provides time intervalls in which an accent occurs with some confidence value.

```
<P> ja prima     dann lassen Sie uns   doch noch einen Termin
        L+H* H-H%      H*           L-                  H*
             B3                     B2
        PA             NA                               PA

ausmachen     <P> wann w"ar's Ihnen denn recht     <P> <#Klicken> <P>
          L-L%                 H*         L*  H-H%
           B3                                      B3
                                  NA          PA   ?
```

Figure 2: Result of a manual prosodic transcription

```
 590  --    690 ms accent, length  100 ms, confidence 0.471084
 1010  --   1130 ms accent, length  120 ms, confidence 0.650081
 1330  --   1430 ms accent, length  100 ms, confidence 0.561492
 1690  --   1780 ms accent, length   90 ms, confidence 0.330556
 1840  --   2070 ms accent, length  230 ms, confidence 0.477964
 2300  --   2470 ms accent, length  170 ms, confidence 0.400369
 2650  --   2750 ms accent, length  100 ms, confidence 0.442033
 2780  --   2890 ms accent, length  110 ms, confidence 0.551336
 2960  --   3490 ms accent, length  530 ms, confidence 0.410003
```

Figure 3: Output of an accent recognizer[2].

2.4. Lexical semantics

As most natural language processing systems, we assume having a lexicon. Of course it should include phonological and syntactical properties for word recognition as well as for a syntactical parser. Additionally our lexicon contains lexical semantics (like in [Sow92]) and it makes sense to suppose that there is a consistent correspondence between lexical semantics and concepts in the domain model. Lexical semantics is an important information source for recognising dialogue acts. Figure 4 shows a canonical graph which is a conceptual graph contraining the structure of a concept, here *appointment*. Different relations connect concepts like *time*, *date* and *place* to the appointment concept. The agents of an appointment are at least two persons, expressed as a set of persons with a quantity of at least two.

```
[APPOINTMENT]-
     (AGNT)->[PERSON: {*}]->(QTY)->[NUMBER:@>=2]
     (PTIM)->[TIME]
     (PDAY)->[DATE]
     (LOC)->[PLACE]
```

Figure 4: A canonical graph for an appointment

[2] H* is a high level accent, L* is a low level accent, % marks the end of a phrase, B3 are phrase boundaries, PA are phrase accents and NA are secondary accents

2.5. Syntax

Spoken language usually does not consist of well formed sentences [Web95], thus a speech dialogue system cannot rely on syntactically sound sentences. Nevertheless, smaller units like nominal phrase may contribute a little. Compared to dialogue systems processing written natural language, syntax is less important for spoken dialogue systems [Sch92, Sch93].

2.6. Semantics

As far as parts of a sentence have been syntactically analysed, lexical semantics could be joined to larger units (c.f. [Sow88]). By joining more and more units we approach to understanding an utterance.

3. Complex data and their treatment within a machine learning environment

Most machine learning work contributes to algorithms that work with simple nominal attributes to characterise a case or event. Every attribute may have exactly one value out of a countable and finite set. Some work contributes to numerical attributes [Mer93]. As we have seen above, data in a speech dialogue system are more complex and current algorithms will have to be modified to deal with that complexity. In this section we will address data characteristics that are problematic to deal within machine learning and indicate their integration into a learning algorithm.

If we look at the list of useful data, we have to realise, that especially those data stemming from signal processing modules do not provide certain values but probalistic information and these data describe some events that are assigned to time intervals. Sometimes data are not only probalistic but they also allow ambiguity, i.e. different possible solutions at the same time.

Data stemming from higher processing modules usually have a complex structure, like parse trees for syntax or semantic networks for semantics.

3.1. Learning from nominal attributes

This is the well known case in most machine learning algorithm. COBWEB incrementally builds a classification tree and the distiction between subnodes are made from a function over all attributes (polythetic). COBWEB might use some pruning function and we may assume a past-performance prediction.

3.2. Learning from probalistic data

In a spoken dialogue system we do not deal with certain data as we have shown in Section 2. We will now address the question, what it means for a learning algorithm if we present a case which has an associated confidence level. I would supply the idea, that a case or event presented to a machine learning algorithm with one attribute having a probability value should not be treated as an entire case, that has to be classified, but depending on the probability value a half or a quarter case, or whatever value is given. This has some influence on the classification function as well as on the past-performance prediction.

Within COBWEB's past-performance prediction "for each attribute and node, a count is maintained of the number of times the attribute was correctly predicted at the node (i.e., correct-at-node counts) during training."[Fis89] We have to change the number of times to the sum of the confidence values for all those cases, where the attribute was correctly predicted.

3.3. Learning with ambiguous data

Ambiguous data means, that one case's attribute may have any yet unknown value out of a set of values. Classification of such a case may end up in different pathes through the classification tree, depending on the value. If the ambiguity is expressed with some probabilites for each value, we could just compile out the ambiguity creating a set of more or less probable cases and present them to the learning algorithm.
Proceeding this way raises the question, whether ambiguous data should be split all over the classification tree. Ambiguity stems from some regular processing of speech, and the same rules will always produce the same ambiguity from the same data. Thus splitting them up to different branches will not improve our classification system, as it reduces prediction. A more promising approach is that propagated in the system OLOC [Mar94] supplying multiple overlapping hierarchies.

3.4. Learning from attribute streams

A further very important feature of processing speech data is, that the discription of a turn consist of a list of events (words, prosodic events etc.) which are ordered by time. Before investigating simultaneous events and time intervals, we will look at the consequences of having an open ended list of equal attributes (e.g. words) and an order over these attributes. At this point we are not interested in learning from streams and predicting a future event, but we have to face the task to transform these data to a representation usable by the learning algorithm. Multiple entries of one attribute can be handled by COBWEB, as it yet stores probabilities of different attribute values in the same way. The order of the events in a data stream could be represeted by a binary relation *successor*, which might be associative or show some degrading associativity. To keep the attribute space tiny, we will first try to get along with exclusively adding a successor relation between two subsequent events of an attribute stream.

3.5. Learning from data with associated time intervals

Having data that are true only in an interval of time is very common in linear tasks like speech processing. If we would like to be able to consider correlations between simultaneous events as some information to learn, we need some kind of time logic, e. g. like that for time intervals [All83]. The problem here again is, that when introducing the whole set of time relations, the attribute space will grow enormously when all relations between all attribute events will be compiled out. We would thus restrict ourself to regard simultaneous, i.e. overlapping events, only.

3.6. Learning from structured data

Structured data that we are confronted with from knowledge representation for domain knowledge and semantics can be assumed being some kind of terminological logics, which are widely used in natural language processing. Even inference steps can be declaratively described using the same representation formalism [Möl95, Kes95]. We will rely on Conceptual Graphs as a representation formalism, which do not make a distinction between terminilogical and assertional knowledge. There is some work that addresses learning of terminological descriptions from case descriptions [Kie94, Coh94], but our problem is the other way around. We have to deal with terminological descriptions as attributes and we have to evaluate correlation on these descriptions. When using Conceptual Graphs, there are fast algorithms to evaluate subsumption between two graphs [Ell93]. Subsumption could be seen as a simple correlation criterion over structured data as it comprises subsumption of individuals to classes as well as the subsumption of more specific individual graph to more general ones. Up to now, we do not judge an exact measure of correlation being useful, like having a function, that counts the number of different concepts, because this would imply semantic correspondence being countable.

4. Summary

Our approach to learning dialogue models is conducted by the idea of dialogue engineering in the context of a spoken dialogue system. In contrast to all other recent applications of machine learning to discourse or dialogue phenomena [Hir92, Lit94, Sie94, Leh94], we will not rely on hand-labeled linguistic data, but only on data which can be automatically generated in such a system. This has two positive effects, first to get rid of the enourmous effort to label dialogues, and as a consequence of this second it is possible to use very large amounts of data for learning. For domain and task knowledge we assume the existence of an appropriate background knowledge based system.

A second difference to all other known approaches using machine learning for discourse or dialogue modelling, is that we will not rely on some given discourse theory with a given set of classes of dialogue acts or whatever units. Instead, we will use machine learning to discover – may be domain-specific – classes from a set of example dialogues. These classes are based on data that are available in a dialogue system environment, and thus our dialogue model is directly applicable to this system. Additionally, we expect some insight on the importance of different dialogue and discourse structures and their cooperation.

We first presented an analysis of requirements to the overall task and to the learning algorithm, then we analysed available data and their implications and requirements on the learning algorithm. From first tests there is some evidence for a success of this approach, but an implementation and validation is subject of near-future work.

References

All82. Allen, J.F., Frische, A.M., and Litman, D.J. ARGOT: the Rochester Dialog System. In *Proc of Annual Meeting of the American Association of Artificial Intelligence*, 1982, pp. 66-70.

All83. Allen, J.F. Maintaining knowledge about temporal intervals. *Communications of the ACM* , 26 (11) (1983), 832-843.

Alt95. Althoff, F., Carson-Berndsen, J., Drexel, G., Gibbon, D., Hübener, K., Jost, U., Kirchhoff, K., Pampel, M., Petzold, A., and Strom, V. BELLEx3+1: Linguistische Worterkennung unter Berücksichtigung der Prosodie. Tech. Rept. TD 22-95, Verbmobil, 1995.

Ang91. Angele, J., Fensel, D., Landes, D., and Studer, R. KARL: An executable language for the conceptual model. In *6th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop,* Banff, Alberta, 1991, pp. 1.1-20.

Bat93. Batliner, A., Kompe, R., Kießling, A., Nöth, E., Niemann, H., and Kilian, U. The prosodic marking of accents and phrase boundaries: Expectations and results. In *NATO ASI: New Advances and Trends in Speech Recognition and Coding,* Rubio, A.J., Bubion, Granada, 1993, pp. 89-92.

Bil91. Bilange, E. A Task Independent Oral Dialogue Model. In *Proc. European Chapter of the ACL,* 1991, pp. 83-88.

Bun89. Bunt, H. Information Dialogues as Commmunicative Actions in Relation to Partner Modelling and Information Processing. In *The Structure of Multimodal Dialogue,* North-Holland, Amsterdam, 1989, pp. 47-73.

Bun94. Bunt, H. Context and Dialogue Control. *Think* , May 94 (1994), 19-31.

Car88. Carpenter, G.A. and Grossberg, S. The ART fo Adaptive Pattern Recognition by a Self-Organzing Neural Network. *IEEE Computer* , 3/1988 (1988).

Car90. Carpenter, G.A. and Grossberg, S. Adaptive Resonance Theory: Neural Network Architecture for Self-Organizing Pattern Recognition. In *Parallel Processing in Neural Systems and Computers (10th Cybernetics DGK),* North Holland, Amsterdam, 1990, pp. 343-349.

Car91. Carberry, S. *Plan recognition in Natural Language Dialogue,* MIT Press, Cambridge, Mass. (1991).

Chu94. Chu-Carroll, J. and Carberry, S. A Plan-Based Model for Response Generation in Collaborative Task Oriented Dialogues. In *Proc. Annual Meeting of the American Association for Artificial Intelligence,* Seattle, 1994.

Coh87. Cohen, P.R. and Levesque, H.J. Rational Interaction as the Basis for Communication. Tech. Rept. CSLI-87-89, Center for the Study of Language and Information - Report, 1987.

Coh90. Cohen, P.R. and Levesque, H.J. Intention Is Choice with Commitment. *Artificial Intelligence* , 42 (1990), 213-261.

Coh94. Cohen, W.W. and Hirsch, H. The Learnability of Description Logics with Equality Constrains. *Machine Learning* , 17-2/3 (1994), 169-199.

Dar94. Daradoumis, T. Building an RST-Based Multi-Level Dialogue Context and Structure. In *Proceedings of the 10th Conference on Artificial Intelligence for Applications,* 1994, pp. 465-466.

Ell93. Ellis, G. Efficient retrieval from hierarchies of objects using lattice operations. In *Conceptual Graphs for Knowledge Representation-1. Int Conf Conceptual Structures,* Springer-Verlag, Berlin NewYork London, 1993, pp. 274-293.

Fér93. Féry, C. *German Intonational Patterns,* Niemeyer, Tübingen , Linguistische Arbeiten(1993).

Faw89. Fawcett, R.P. and Taylor, M.M. A Generative Grammar for Local Discourse Structure. In *The Structure of Multimodal Dialogue,* North-Holland, Amsterdam, 1989.

Faw92. Fawcett, R.P. and Davies, B. Monologue as a Turn in Interactive Discourse: Towards an Integration of Exchange Structure and Rhetorical Structure Theory. In *6th International Workshop on Natural Language Generation, 1992.,* Dale, Rösner, H., Trento, Italy, 1992.

Fen94. Fensel, D. and Harmelen, F. A Comparison of Languages which Operationalize and Formalise KADS Models of Expertise. *The Knowledge Engineering Review* , 8 (2) (1994), 49.

Fis85. Fisher, D.H. and Langley, P. Approaches to Conceptual Clustering. In *Proc. of the Ninth International Joint Conference on Artificial Intellingence,* 1985, pp. 691-697.

Fis87. Fisher, D.H. Knowledge Acquisition via Incremental Conceptual Clustering. *Machine Learning* , 2 (1987), 139-172.

Fis89. Fisher, D.H. Noise-Tolerant Conceptual Clustering. In *11th International Join Conference on Artificial Intelligence,* N.S.S., N.S.S., Morgan Kaufmann, Detroit,Michigan,USA, 1989, pp. 825-830.

Gen89. Gennari, J.H., Langley, P., and Fisher, D.H. Models of Incremental Concept Formation. *Artificial Intelligence* , 40 (1989), 11-61.

Gri86. Grishman, R. and Kittredge, R. *Analysing language in restricted domains,* Lawrence Erlbaum Associates, Hillsdale, NJ (1986).

Gro86. Grosz, B.J. and Sidner, C.L. Attention, Intentions, and the Structure of Discourse. *Computational Linguistics* , 12 (1986), 175-204.

Gro87. Grossberg, S. Competitive Learning: From Adaptive Activation to Adaptive Resonance. *Cognitive Science* , 17 (1987), 23-63.

Hir92. Hirschberg, J. and Grosz, B.J. Intonational Features of Local and Global Discourse Structure. In *Proc. Speech and Natural Language Workshop, Harriman, N.Y.,* Morgan Kaufmann Publishers, San Mateo, CA, 1992, pp. 441-446.

Hir93. Hirschberg, J. and Litman, D.J. Empirical Studies on the Disambiguation of Cue Phrases. *Computational Linguistics* , 19/3 (1993), 501-530.

Jön91. Jönsson, A. A Dialogue Manager Using Initiative-Response Units and Distributed Control. In *Proc. European Chapter of the ACL,* 1991, pp. 233-238.

Jön93. Jönsson, A. A Method for Development of Dialogue Managers for Natural Languages. In *Proc. of Annual Meeting of the American Association of Artificial Intelligence,* 1993, pp. 190-195.

Kes95. Kessel, T., Medeiros, H., Rousselot, F., and Keith, B. Some Useful Enhancements of Kl-One Languages to become Modeling Languages. In *Modelling Languages for Knowledge-Based Systems,* Vrije Universiteit Amsterdam, Amsterdam, 1995.

Kie94. Kietz, J.U. and Morik, K. A Polynomial Approach to the Constructive Induction of Structural Knowledge. *Machine Learning* , 14-2 (1994), 193-217.

Kno91. Knott, A. *New Strategies and Constraints in RST-based Text Planning.* MSc thesis, Dept. of Artificial Intelligence, Univerity of Edinburgh (1991), 81.

Koh84. Kohonen, T. *Self-Organization and Associative Memory,* Springer Verlag, Berlin , Springer Series in Information Sciences(1984).

Lam91. Lambert, L. A Tripartite Plan-Based Model of Dialogue. In *Proc. 29th Annual Meeting of the ACL,* 1991, pp. 47-54.

Leh86. Lehrberger, J. Sublanguage Analysis. In *Analyzing Language in Restricted Domains: Sublanguage Description and Processing,* Grishman, R.; Kittredge, R., Lawrence Erlbaum Associates, Hillsdale, NJ, 1986, pp. 19-38.

Leh94. Lehnert, W. and Soderland, S. Corpus-Driven Knowledge Acquisition for Discourse Analysis. In *Proc. Annual Meeting of the American Association for Artificial Intelligence,* Seattle, 1994, pp. 827-832.

Lit85. Litman, D.J. *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues*, Ph.D. dissertation, ??, 1985.

Lit94. Litman, D.J. Classifying Cue Phrases in Text and Speech Using Machine Learning. In *Proc. Annual Meeting of the American Association for Artificial Intelligence,* Seattle, 1994, pp. 806-813.

Möl90. Möller, J.U. Fachdialogwissen als Mittler zwischen Fachwissen und Dialogwissen. In *GWAI-90 14th German Workshop on Artificial Intelligence, Eringerfeld,* Marburger, Heinz, Eringerfeld, 1990, pp. 58-64.

Möl92a. Möller, J.U. Leitfaden zur Modellierung von Fachdialogen zwischen Nutzern und System. Tech. Rept. Memo 211/92, Berichte des Fachbereichs Informatik, Universität Hamburg, 1992.

Möl92b. Möller, J.U. Fachspezifische Implikatoren für die Planung und Generierung von Inhalt und Form in Dialogen. In *Proc. GWAI-92 Workshop Bewertungskriterien zur Auswahl zwischen Alternativen in der natürlichsprachlichen Generierung* 16th German Workshop on Artificial Intelligence, unpublished, 1992.

Möl95. Möller, J.U. Operationalisation of KADS Modules by Using Conceptual Graph Modules. In *Knowledge Acquisition for Knowledge Based Systems Workshop,* Banff, 1995.

Möl95a. Möller, J.U. Domain-Related Focus-Shifting Constraints in Dialogues with Knowledge Based Systems. In *Trends in Natural Language Generation: an Artificial Intelligence Perspective,* Adorni, Giovanni; Zock, M., Springer Verlag, LNAI, 1995, to appear.

Mai94. Maier, E. Dialogmodellierung in VERBMOBIL - Festlegung der Sprechhandlungen für den Demonstrator. Tech. Rept. Memo 31, Verbmobil, 1994.

Man88. Mann, W.C. and Thompson, S.A. Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text* , 8,3 (1988), 243-281.

Mar94. Martin, J. and Billman, D.O. Acquiring and Combining Overlapping Concepts. *Machine Learning* , 16 (1994), 121-155.

Mer93. Merckt, T. Decision Trees in Numerical Attribute Spaces. In *Proc. of the 13th International Joint Conference on Artificial Intelligence,* 1993, pp. 1016-1021.

Moo92. Moore, J.D. and Pollack, M. A Problem for RST: The Need for Multi-Level Discourse Analysis. *Computational Linguistics* , 18/4 (1992), 537-544.

Nöt91. Nöth, E., Batliner, A., Kuhn, T., and Stallwitz, G. Intensity as a predictor of focal accent. In *ICPhS91,* Aix-en-Provence, 1991.

Pie80. Pierrehumbert, J. *The Phonology and Phonetics of English Intonation,* Ph.D. dissertation, MIT, Cambridge, MA., 1980.

Ram91a. Ramshaw, L.A. A Three-Level Model for Plan Exploration. In *Proc. 29th Annual Meeting of the ACL,* 1991, pp. 39-46.

Ram91b. Rambow, O., Kittredge, R., and Korelsky, T. On the Need for Domain Communication Knowledge. *Computational Intelligence* , 7(4) Dec. 91 (1991).

Rey94. Reyelt, M. Untersuchung zur Konsistenz prosodischer Etikettierungen. Tech. Rept. TD 03-94, Verbmobil, 1994.

Rit89. Ritter, H. and Kohonen, T. Self-Organizing Semantic Maps. *Biological Cybernetics* , 61 (1989), 241-254.

Rum80. Rumelhart, D.E. and Zipser, D. Feature Discovery by Competitive Learning. In *Parallel Distributed Processing: Explorations in the Mircostructure of Cognition,* Rumelhart, D. E.; McClelland, J.L., MIT Press, Cambridge, 1980, pp. 151-193.

Sch92. Schröder, M. Knowledge Based Analysis of Radiology Reports Using Conceptual Graphs. In *Conceptual Structures: Theory and Implementation - Proc 7th WS, Las Cruces, NM,* Pfeiffer, H. D.; Nagle, T.E., Springer Verlag, Lecture Notes in Artificial Intelligence, Berlin, 1992, pp. 293-302.

Sch93. Schröder, M. *Erwartungsgestützte Analyse medizinischer Befundungstexte - Ein wissensbasiertes Modell zur Sprachverarbeitung,* Ph.D. dissertation, infix, St. Augustin, 1993.

Sie94. Siegel, E.V. and McKeown, K.R. Emergent Linguistic Rules from Inducing Decision Trees: Disambiguating Discourse Clue Words. In *Proc. Annual Meeting of the American Association for Artificial Intelligence,* Seattle, 1994, pp. 820-826.

Sil92. Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., and Hirschberg, J. Tobi: A standard for labeling english prosody. In *International Conference on Spoken Language Processing,* 1992.

Sow88. Sowa, J.F. Using a Lexicon of Canonical Graphs in a Semantic Interpreter. In *Relational Models of the Lexikon,* Evens, M., Cambridge University Press, 1988.

Sow92. Sowa, J.F. Lexical Structures and Conceptual Structures. In *Semantics in the Lexicon,* Pustejovsky, J., Kluwer Publishers, 1992.

Ste93. Stein, A. and Thiel, U. A Conversational Model of Multimodal Interaction in Information Systems. In *Proc. of Annual Meeting of the American Association of Artificial Intelligence,* 1993, pp. 283-288.

Str94. Strom, V. Die Prosodiekomponente in INTARC I.2. Satzmodusbestimmung aus der F0. Tech. Rept. TD 6-95, Verbmobil, 1994.

Uhm91. Uhmann, S. *Fokusphonologie: eine Analyse deutscher Intonationskonturen im Rahmen der nichtlinearen Phonologie,* Niemeyer, Tübingen , Linguistische Arbeiten(1991).

Wac86. Wachtel, T. Pragmatic sensitivity in NL interfaces and the structure of conversation. In *11th Int. Conf. on Computational Linguistics, Bonn (COLING86),* 1986, pp. 35-41.

Web95. Weber, V. and Wermter, S. Towards Learning Semantics of Spontaneous Dialog Utterances in a Hybrid Framework. In *AISB95,* 1995.

Woo84. Woolf, B. and McDonald, D.D. Context-dependent Transitions in Tutoring Discourse. In *Proc. Annual Meeting of the American Association of AI,* 1984, pp. 355-361.

You94. Young, M. and Moore, J.D. Does Discourse Planning Require a Special-Purpose Planner? In *Proc. Annual Meeting of the American Association for Artificial Intelligence,* Seattle, 1994.