Towards Resolving Referring Expressions by Implicitly Activated Referents in Practical Dialogue Systems*

Norbert Pfleger and **Jan Alexandersson**
DFKI GmbH
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany
{pfleger,janal}@dfki.de

Abstract

We present an extension to a comprehensive context model that has been successfully employed in a number of practical conversational dialogue systems. The model supports the task of multimodal fusion as well as that of reference resolution in a uniform manner. Our extension consists of integrating implicitly mentioned concepts into the context model and we show how they serve as candidates for reference resolution.

1 Introduction

The basis for any spoken interaction between two or more interlocutors is common ground. The common ground of two people is the sum of their mutual beliefs about the conversation and the surrounding world. As indicated in (Clark, 1996), when strangers meet they start building up presuppositions about the background knowledge of the other, thereby mutually updating and extending their common ground. Consequently, many referring expressions are only meaningful in the particular context in which they are uttered.

When we consider the course of interactions, it becomes clear that for virtually every contribution the context is extended with more or less related concepts. For a hearer, the process of relating new information to concepts already *known* is vital to the comprehension of a discourse. The basic mechanisms of this process—called *bridging*—is discussed in (Clark, 1977). In addition to direct references, he identifies, for instance, *indirect* references by association. Consider the following example taken from (Clark, 1977):

> “`I looked into the room. The ceiling was very high.`”

During the processing of the first sentence, a hearer with profound knowledge about rooms will most likely not only activate the concept *room* per se but also associated concepts (e. g., *ceiling*, *wallpaper*, etc.). The degree of activation is influenced by factors such as the hearer’s personal interests, recency of the acquired knowledge etc. But also the situational context (current location, time, weather conditions, etc.) is vital. For instance, for a mobile dialogue system in a tourist scenario it will be necessary to activate—and deactivate—buildings and streets while the user is moving around.

The aim of our work is the incorporation of these findings into a module for reference resolution for a multimodal conversational agent. Here, we are focusing in particular on the correct interpretation of named entities and definite noun phrases whose referents have not been explicitly mentioned but are part of the implicit context. Key to our approach is the integration of a long term memory (LTM) modeling the complete knowledge of an agent. Next to this LTM is a working memory (WM) that realizes a comprehensive context model. However, as we will argue in this paper, some processes in the human LTM have direct impact on the organization and structure of the contextual model. To this end, we integrate a structure resembling the human long term memory into our discourse model. The LTM represents the complete knowledge a discourse participant of a particular social role and status is supposed to *know*.

The paper is organized as follows: In the next

*This research is funded by the German Ministry of Research and Technology (BMBF) under grant 01IMD01A and by the EU 6th Framework Program under grant FP6-506811 (AMI). The responsibility lies with the authors.

brandial 2006: Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue, pages 2–9.

section we will give a brief overview of the system within which this approach is being developed. Then we describe our context model in section 3. In section 4 we will detail how references to implicitly activated concepts are resolved in this framework.

2 System Context

In this paper, we use the Question-Answering system SMARTWEB (Reithinger et al., 2005). However, our model including the approach described here is used in other systems as well. SMARTWEB is a mobile, multimodal interface to the semantic web. The user is able to ask open-domain questions to the server-based dialog system via a smartphone. Primary input modalities are speech and pen input that operate in a continuous recognition mode. There are also a camera based on-/off-view detection and a prosody-based on-/off-talk detection that prevent the system from processing user contributions that were not intended to address the system. Figure 1 depicts an example configuration of the system's graphical user interface.

Figure 1: Graphical user interface of the SMARTWEB system.

An important aspect of SMARTWEB is its context-aware processing strategy. All recognized user actions are processed with respect to their situational and discourse context. A user is thus not required to pose separate and unconnected questions. In fact, they might refer directly to the situation, e. g., , *"How do I get to Berlin from here?"*, where *here* is resolved via GPS information, or to previous questions (e. g., *"And in 2002?"* in the context of a previously posed question *"Who won the Fifa World Cup in 1990?"*). The interpretation of user contributions with respect to their context is performed by a component called *Fusion and Discourse Engine*—FADE (Pfleger, 2005). FADE integrates the verbal and nonverbal user contributions into a coherent multimodal representation and enriches it with contextual information, e. g., resolution of referring and elliptical expressions.

The basic architecture of FADE consists of two interweaved processing layers (see figure 2): (i) a production rule system—PATE—that is responsible for the reactive interpretation of perceived monomodal events, and (ii) a discourse modeler—DiM—that is responsible for maintaining a coherent representation of the ongoing discourse and for the resolution of referring and elliptical expressions. Both processing layers operate on the same working memory. This paper deals with processes that are related to the discourse modeler of FADE, for a comprehensive overview of FADE see (Pfleger, 2005).

2.1 Knowledge Representation

All knowledge in the SMARTWEB system is represented by means of a single system-wide ontology *SWINTO*. This ontology integrates central concepts of SUMO (Niles and Pease, 2001) and DOLCE (Gangemi et al., 2002) and covers a set of sub-ontologies like a sportevent ontology, a navigation ontology, a media representation ontology, a linguistic ontology, a discourse ontology, etc. All data exchanged within SMARTWEB's dialogue component is represented by means of RDF (resource description framework; see `http://www.w3.org/RDF/`) instances of concepts of the SWINTO ontology.

In this paper we focus on the sportevent sub-ontology. This ontology covers a great number of instances describing the Fifa World Cups since 1954: knowledge about all players and countries that are somehow related to the World Cups, all the games including events like goals, yellow and red cards, etc.

2.2 Sample Dialogues

We will discuss two sample interactions with the SMARTWEB system: The first example illustrates how our approach deals with implicit knowledge while the second shows how the physical context influences the activation process and thus the resolution of referring expressions.

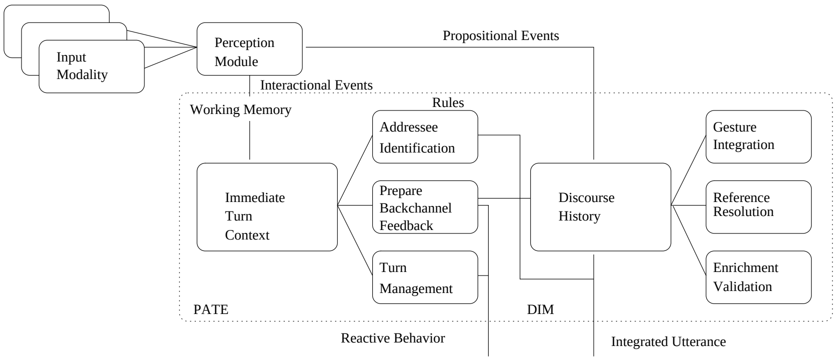

Figure 2: A schematic overview of FADE and its basic functionality.

In example 1 the user's first question sets the context for the interpretation of the second one. A possible context-free interpretation of the second contribution is *"How many goals did Ziege score in his entire career?"*. However, in the given context, the intended interpretation is rather *"How many goals did Ziege score in this particular game?"*.

(1) **User:** *"Wie ist das Spiel Deutschland gegen die USA bei der WM 2002 ausgegangen?"*
(*"How did the game Germany against USA during the World Cup 2002 end?"*)

System: *1:0.*

User: *Wieviele Tore hat Ziege geschossen?*
(*"How often did Ziege score?"*)

The intended interpretation is settled by the fact that the football player "Ziege" was a member of the German team and participated in that particular game. Thus, what we need is to establish a link between the instance representing the FOOTBALLPLAYER Ziege, which is embedded in an under-specified FIELDMATCHFOOTBALLPLAYER and the already activated instance of the class FIELDMATCHFOOTBALLPLAYER of the game mentioned in the previous user turn (see instance *Ziege_1149* in figure 3). This link can only be established if not only the game itself is activated but also related concepts like all the players that participated in that game, the location where the game took place, etc. are activated as well.

Example 2 shows the need for making concepts accessible not only through relatedness to verbally mentioned concepts but also to graphical—or even physical—objects.

(2) **User:** *"Wer war im Finale der WM 2002?"*
(*"Who was in the World Cup final in 2002?"*)

System: [Displays pictures of the two finalists France and Brazil]
(The user selects a picture of the Brazilian team and looks at it)

User: *"Wie heißt der Torwart?"*
(*"What's the goalkeeper's name?"*)

The user refers with *the goalkeeper* to a person that has not been explicitly introduced into the discourse yet but is visible on a picture of the Brazilian team (which is in the current focus of attention of the user). Again, for the resolution of the referring expression we need access to concepts that are somehow related to the pictures.

3 A Comprehensive Context Model

The architecture of our context model is centered around the idea of two interweaved memory representations: (i) a working memory (WM) where the actual processing of contributions takes place and (ii) a long term memory (LTM) that serves as the central knowledge repository of the system. Vital for the work described here is that every concept has a certain amount of activation in the codomain $[0, 1]$ defining its accessibility.

The LTM and WM are separated by a threshold (see figure 3). All instances whose activation is below the threshold belong to the LTM and vice versa. Thus, the LTM comprises all the instance knowledge of the system that is not directly accessible and the WM comprises all those concepts that have been explicitly and implicitly mentioned in the previous discourse. An increase in activation is not only triggered by verbal reference to a knowledge base entry but also by nonverbal ac-

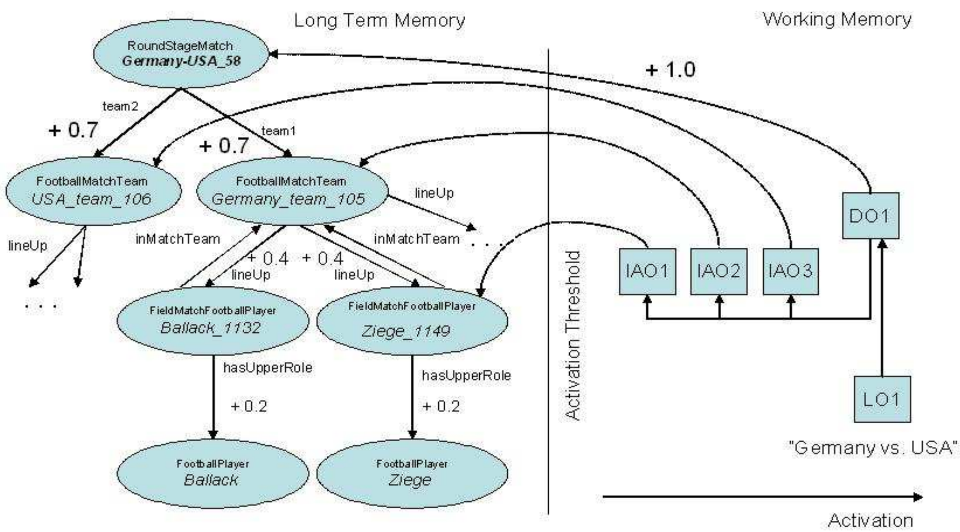

Figure 3: Basic organization of the context model.

tions (e. g., pointing and iconic gestures, etc.) or by physical presence.

3.1 Representation of Knowledge Chunks

All data of the SMARTWEB system is represented by means of instances of ontological concepts. This representation, however, is not sufficient for our purpose as we need a more expressive representation in order to represent the activation of the individual instances. To this end we take a slightly different view on the data and convert every incoming structure into a typed feature structure (TFS; see (Carpenter, 1992)) like representation. Basically, this extended TFS-based representation has the same expressive power as the RDF instances but supports activation values that are directly associated with an individual instance and supports unification and overlay (Alexandersson and Becker, 2003; Pfleger et al., 2002). For a more detailed description of this extended TFS implementation see (Kempe, 2004).

3.2 Organization of the Working Memory

Following the *three-tiered discourse representation* of (LuperFoy, 1991), our discourse modeler comprises two main layers: (i) a modality object layer—extending its linguistic layer and (ii) a concept layer—extending its discourse layer. The third layer—the knowledge base or belief system—corresponds in our approach to the LTM.

3.2.1 Modality Layer

The objects of the modality layer provide information about the surface realizations of objects at the concept layer that have been introduced into the discourse. Thus, modality objects describe the circumstances that caused the increase in activation of their corresponding concept object. The modality layer consists of three classes of objects reflecting the modality by which the corresponding working memory element was referenced: (i) *linguistic actions*, (ii) *nonverbal actions*, and (iii) *physical events*.

Linguistic Actions Linguistic actions resemble the linguistic objects of (LuperFoy, 1992). They comprise information about the surface realization of a concept like lexical information (the lemma used to reference a concept), syntactical information (e. g., number, gender or case), its realization time, or the type of reference (e. g., definite/indefinite, deictic/anaphoric/partial anaphoric). Each Linguistic Action is linked to exactly one instance of the knowledge base and when this link is established, the referenced object's activation is increased. Linguistic Actions are of particular importance for the resolution of referring expressions as they provide the linguistic information needed to identify co-references on the linguistic level.

Nonverbal Actions Nonverbal Actions represent the nonverbal behavior of the interlocutors that contribute to the propositional content of the utterance (e. g., pointing gestures, iconic gestures, emblematic gestures, but also gaze behavior, or drawings, etc). Nonverbal Actions comprise information about the type of nonverbal action, its start and end time. Nonverbal Actions facilitate the resolution of deictic expressions (e. g., *"What's the name of that [pointing gesture] player?"*).

Physical Events Physical Events describe the appearance or disappearance of objects in the physical environment. They comprise information about the type of the event, when it happened, and about the spatial properties of that object (including its relative position to other objects in the scene).

3.2.2 Concept Layer

Objects at the concept layer provide the link to the concepts of the LTM. Each object at the concept layer (if completely disambiguated) represents a unique instance of a concept of the LTM whose activation value exceeds the threshold. We distinguish three types of objects at the concept layer: (i) Discourse Objects, (ii) Implicitly Activated Objects, and (iii) Physical Objects.

Discourse Objects Discourse Objects are containers for concepts that were directly mentioned during the preceeding discourse. They comprise a unified representation of the semantic information gathered so far. In case a Discourse Object is completely resolved, its unified representation is replaced by a link to the corresponding concept of the LTM. Additionally, it contains a set of links to objects at the modality layer; every time a Discourse Object is mentioned, a new link is added.

Implicitly Activated Objects Implicitly Activated Objects are objects that are related to a Discourse Object. If a Discourse Object accesses a concept in the LTM, the activation of concepts related to it is increased by a dynamic factor which depends on the activation of the superordinated concept and the *strength* of the relation between them. The spreading of activation is a recursive process (see section 3.3).

Implicitly Activated Objects may appear in the WM when their corresponding Discourse Object appears. This happens in case their activation exceeds the threshold. Also, the activation of Implicitly Activated Objects decreases faster than that of Discourse Objects. Consequently they are only accessible for a short time.

Physical Objects Physical Objects represent objects that can be perceived in the visual environment. If a Physical Object is explicitly activated through the mentioning of a Discourse Object, it can serve as a referent for a referring expression. Physical Objects are not only part of the discourse layer but they are also part of a superordinate structure representing the complete physical surroundings by modeling the relations between the physical objects located in a scene (e. g., the grey building is on the left of the blue building, etc.). In our implementation, we treat objects displayed on the screen as Physical Objects.

3.3 Organization of the Long Term Memory

As in the human LTM, the objects represented in our LTM are connected through relations. These relations have also a strength associated with them that defines the proximity between the two connected objects. The left part of figure 3 depicts a small excerpt of such a semantic network. At the bottom of this figure there is a concept representing the German football player *Michael Ballack*. This concept exhibits several connections with other concepts like teammates, or the German national team that participated in the game against the USA in the 2002 World Cup.

3.3.1 Lookup and Retrieval of Concepts of the LTM

The concepts of the LTM are stored directly in the type hierarchy, i. e., every type also provides a storage position for its concepts. This means a concept of Type *A* is stored directly at that type and can be retrieved in turn via that type. Thus, lookup and retrieval of concepts of the LTM is comparatively cheap since the type of the search pattern already restricts the search space to a restricted subset of the complete knowledge base. The actual matching is done by unifying the search pattern with the individual concepts stored for that type. Note that since the lookup is typically based on an under-specified concept, it might return more than one match. In such a case the reference resolution algorithm must deal with this ambiguity.

3.3.2 Activation Propagation

An important aspect of our LTM is that every object has an activation value defining its accessibility. The higher the activation value, the easier it is to access the object (i. e., to retrieve the object from the complete set of knowledge). To account for the activation of neighboring concepts that can be observed in human interactions, the activation of a knowledge chunk is passed on to its associated chunks by a process called *spreading activation*. Spreading activation doesn't only mean that each connected object receives part of the activation of its neighbors but also that it spreads its own activation on to its own neighbors.

An important aspect of a spreading activation model is that activation may spread not only to directly related concepts but also from those concepts to concepts further away in the memory network. This is called the *multi-step* assumption as opposed to the *one-step* assumption that predicts spreading activation only between directly related concepts. In an experimental study, (Sharifian and Samani, 1997) found evidence that also supports the multi-step assumption and the assumption that the activation reduces as it traverses intermediate concepts.

Another aspect of a spreading activation network is the amount of activation that is passed on from one concept to another. Typically, this is controlled by means of strength values associated with the connections (slots) between concepts.

Both the reduction of activation passed from one node to the next and the strength values of connections are important features that influence the behavior of such a network. Therefore, we are currently in the process of developing an empirical method to gather this information for specific domains (see section 5).

3.3.3 Activation Decay

In order to reflect the processes of the human memory, the activation of a concept fades out in time. This means that the longer a concept has not been referenced, the lower its activation will be. Eventually, when the concept's activation is below the threshold, the corresponding object at the concept layer of the working memory will disappear so that the object is no longer directly accessible. However, the activation of an instance will never get below its basic activation.

The three different objects at the concept layer exhibit different intensities in activation decay. Implicitly Activated Objects show the most rapid decay, followed by Discourse Objects. The activation of a Physical Object normally remains on a level that is above the threshold where objects disappear from the working memory.

3.3.4 Current Settings

In our current implementation, the basic activation is $BA = 0.2$. An explicitly mentioned concept receives an increase in activation of 0.7—the connection strength. Our spreading activation algorithm uses this number for multiplying the activation of the related concepts and stops when the result is below the threshold. The threshold is currently set to 0.4. Note, that some concepts might receive an activation between the basic activation and the threshold, in which case they remain slightly more activated in the LTM. Eventually, they might, due to repeated mentioning of related concepts, qualify for the WM. Clearly, these numbers are nothing but heuristics; but as indicated in section 6 we strive for more natural and elaborated numbers.

In the current implementation, the search for referents in the WM is cheap since there are—in our experiments so far—well below hundred. For lookups in the LTM, even in the case of 100k concepts, we have a response time of less than 250 ms.

4 Activation-Based Reference Resolution

Our reference resolution approach differs from standard approaches for reference resolution in one major aspect, see, e. g., (Jurafsky and Martin, 2000) for an overview. A standard reference resolution algorithm initially computes a candidate list of potential antecedents. However, since our context model is self-organizing with regard to the activation of the concepts, there is no need to compute this list because it is always accessible and ordered. Thus, our algorithm takes the existing list and tries to narrow it down by using the linguistic features of the referring expression thereby looking for compatible semantic representations. Note that due to the decay in activation, the concept layer comprises only those candidates that have either been mentioned recently or multiple times. Since we focus here on the resolution of named entities and NPs that might refer to implicitly activated objects, the description of the algorithm will be focused on the resolution of these references.

Given such a reference, the algorithm traverses the objects of the Concept Layer several times until a match is found. In the first run it assumes a reference to an explicitly mentioned discourse object or a physical object. Only if there was no matching object (i. e., an object whose syntactic information is compatible and whose semantic representation unifies with the referring expression) it starts a second run through the list, now focusing on Implicitly Activated Objects. This search continues until the algorithm encounters an Implicitly Activated Object whose semantic representation is unifiable with the one of the referring expression. Finally, if no matching object has been found, there is a third run assuming a discourse reference which we will not discuss here due to spatial restrictions.

4.1 Revisiting Example 1

In the second contribution, the user mentions the named entity *Ziege* that had not been introduced until then. However, the previous reference to the game Germany against USA activated the German team of that game which in turn activated all players that were members of that team. Among these activated players is also our target referent, an instance describing the FIELDMATCHFOOTBALLPLAYER "Ziege" (as depicted in figure 3).

Because of the intra-sentential context of the second user utterance (*"How often did Ziege* **score***?"*), the speech analysis component of SMARTWEB will come up with an interpretation for the named entity "Ziege" where the instance of FOOTBALLPLAYER is enclosed by an unresolved instance of a FIELDMATCHFOOTBALLPLAYER (see the figure 4). Given this interpretation of the named entity, the reference resolution algorithm of FADE fails to find a matching Discourse Object or Physical Object in the Concept Layer. In the second run FADE encounters the implicitly activated instance of the FIELDMATCHFOOTBALLPLAYER *Ziege_1149* that is unifiable with the semantic interpretation of the named entity.

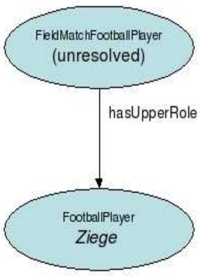

Figure 4: Analysis result for the named entity *Ziege*.

5 Discussion

The application of associative networks and spreading activation for the identification of named entities or the resolution of lexical ambiguity has a long research tradition in psychology and artificial intelligence, see e. g., (Kintsch, 1988; Hirst, 1988). However, in practical dialogue systems this type of contextual information has, to our knowledge, not been applied yet.

Indeed, there exists a number of comprehensive models for the resolution of referring expressions, e. g., (LuperFoy, 1991; LuperFoy, 1992; Allen et al., 2000; Allen et al., 2001). However, all of these models lack the inclusion of implicitly activated concepts into their model of the ongoing discourse.

Our model is best compared to that of (Allen et al., 2000; Allen et al., 2001). There, an architecture for implementing interactive conversational agents is described. We relate our work to their discourse and reference module. In their discourse module, five types of information are present: (i) salient entities, (ii) preceeding utterance, (iii) turn status, (iv) discourse history, (v) discourse obligations.

In contrast to the TRIPS architecture, in our module the salient entities (i) are extended with the activated referents which allow for an interpretation already in the discourse module. Moreover, our context model (see section 3) includes a rich discourse history ((iv), (ii)) where preceeding utterances—along with information such as speech act—is just one part. To a certain extent, their model is able to deal with implicitly mentioned concepts too, but whereas our model utilizes the ontology, their model relies on the plans.

6 Conclusion and Future Work

We have presented a cognitively motivated comprehensive discourse model that mimics the behavior of humans by means of a Long Term Memory (LTM) and a Working Memory (WM). We have shown how *explicitly* mentioned objects are activated and how their activation exceeds a threshold—the edge of consciousness—transfered from the LTM to the WM. Once a concept is in the WM, its activation decreases as time goes by until its activation falls below the threshold and then vanishes from there again. Focus, of this paper has been to show that by using spreading activation, the activation of *implicitly* mentioned concepts increases and when their activation exceeds the threshold that they are transferred into the WM. This enables the interpretation of natural utterances as humans produce them.

6.1 Future Work

As pointed out, our current implementation of the spreading activation process is based on handmade numbers and this, of course, is not feasible in the long run. We therefore recently started to work on an empirical method for measuring the strength of connections between concepts for a given domain. Currently, we aim at a combined experiment that will provide not only information about the strength of connections but also about the frequency measures for particular instances that can be used to compute the basic activation of instances. The ultimate goal is to define a set of experiments and post-processing steps so that we will be able to automatically extract the connection strengths between related concepts.

Moreover, we are currently investigating to what extent it is possible to apply some kind of online-learning functionality for adjusting the connection strength and the introduction of completely new connections/associations between previously unrelated concepts. If, for example, in the course of interactions two unrelated concepts *A* and *Z* appear frequently in the same context, these two concepts will be connected. This means that the mentioning of concept *A* will in the future activate concept *Z*.

Future work will also include the incorporation of implicitly activated concepts that are part of plans or scripts.

References

Jan Alexandersson and Tilman Becker. 2003. The formal foundations underlying overlay. In *Proceedings of the Fifth International Workshop on Computational Semantics (IWCS-5)*, pages 22–36, Tilburg, The Netherlands.

James Allen, Donna Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent. 2000. An architecture for a generic dialogue shell. *Natural Language Engineering special issue on Best Practices in Spoken Language Dialogue Systems Engineering*, 6(3), December.

James Allen, George Ferguson, and Amanda Stent. 2001. An architecture for more realistic conversational systems. In *Proceedings of Intelligent User Interfaces 2001 (IUI-01)*, Santa Fe, NM, January 14–17.

Bob Carpenter. 1992. *The logic of typed feature structures*. Cambridge University Press, Cambridge.

Herbert H. Clark. 1977. Bridging. In Philip N. Johnson-Laird and Peter C. Wason, editors, *Thinking: Readings in Cognitive Science*, pages 169–174. Cambridge University Press, Cambridge, England.

Herbert H. Clark. 1996. *Using language*. The Press Syndicate of the University of Cambridge.

Aldo Gangemi, Nicola Guarino, Claudio Masolo, Alessandro Oltramari, and Luc Schneider. 2002. Sweetening ontologies with DOLCE. In *13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02)*, volume 2473 of *Lecture Notes in Computer Science*, pages 166–182, Siguenza, Spain, Oct. 1–4.

Graeme Hirst. 1988. Resolving lexical ambiguity computationally with spreading activaiton and Polaroid Words. In *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence*, pages 73–107. Morgan Kaufmann Publishers, San Mateo, CA.

Daniel Jurafsky and James H. Martin. 2000. *Speech and Lanuage Processing*. Prentice Hall.

Benjamin Kempe, 2004. *PATE – a production rule system based on activation and typed feature structure elements*. http://www.dfki.de/˜kempe/pate.pdf.

Walter Kintsch. 1988. The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. *Psychological Review*, 95(2):163–182.

Susann LuperFoy. 1991. *Discourse Pegs: A Computational Analysis of Context-Dependent Referring Expressions*. Ph.D. thesis, University of Texas at Austin, December.

Susan LuperFoy. 1992. The Representation of Multimodal User Interface Dialogues Using Discourse Pegs. In *Proceedings of the* 30^{th} *Annual Meeting of the Association for Computational Linguistics - (ACL-92)*, pages 22–31, Newark, USA.

Ian Niles and Adam Pease. 2001. Towards a Standard Upper Ontology. In Chris Welty and Barry Smith, editors, *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)*.

Norbert Pfleger, Jan Alexandersson, and Tilman Becker. 2002. Scoring functions for overlay and their application in discourse processing. In *KONVENS-02*, Saarbrücken, September – October.

Norbert Pfleger. 2005. Fade - an integrated approach to multimodal fusion and discourse processing. In *Proceedings of the Doctoral Spotlight Session of the International Conference on Multimodal Interfaces (ICMI'05)*, pages 17–21, Trento, Italy.

Norbert Reithinger, Simon Bergweiler, Ralf Engel, Gerd Herzog, Norbert Pfleger, Massimo Romanelli, and Daniel Sonntag. 2005. A look under the hood: design and development of the first smartweb system demonstrator. In *ICMI '05: Proceedings of the 7th international conference on Multimodal interfaces*, pages 159–166, New York, NY, USA. ACM Press.

F. Sharifian and R. Samani. 1997. Hierarchical spreading of activation. In F. Sharifian, editor, *Proceedings of the Conference on Language, Cognition, and Interpretation*, pages 1–10. IAU Press.