

Where is natural language understanding?

Toward context-dependent utterance interpretation

John Bryant*, **Nancy Chang***, **Robert Porzel**** and **Keith Sanders***

* International Computer Science Institute ** European Media Laboratory
1947 Center St., Suite 600 Schloß-Wolfsbrunnenweg 33
Berkeley, CA 94704-1198 69118 Heidelberg
U.S.A Germany

{jbryant,nchang,ksanders}@icsi.berkeley.edu porzel@eml.villa.-bosch.de

In the tourism domain, a simple question such as "Where is the castle?" may be interpreted solely as a request for the castle's location. More often, however, such a question is used to ask for directions to the castle. A felicitous response to such a request may depend not only on the questioner's current location but also on other contextual features, such as the weather, traffic conditions, mode of transportation, and time. The need to integrate such diverse sources of information is characteristic of natural-language interactions: In interpreting utterances and formulating responses, listeners habitually rely on contextual features including the previous discourse, the situation, common ground, and shared world knowledge.

Handling implicit and contextual information poses major challenges to current parsing and dialogue management systems (Porzel & Strube 2000). We argue that an embodied approach to language understanding (Feldman et al. 2000) can help address these challenges. Linguistic representations based on an extension of construction grammar (Goldberg 1995; Bergen and Chang. 2000) provide a link between formal (morphosyntactic) structures and embodied (sensorimotor) representations of meaning and context. Parses based on these representations therefore provide rich semantic information that can in turn be used for metaphorical inference (Narayanan 1999) and disambiguation (Narayanan and Jurafsky 1998).

We describe an application that provides tourist information (e.g. spatial instructions and descriptions) in response to natural language input. Successful interpretation of naturally occurring utterances hinges on the use of semantically rich representations that integrate the following sources of information:

- linguistic knowledge
- ontology
- user model
- situation model

Our system uses probabilistic graphical models to represent and integrate these diverse knowledge sources by taking individual expert nodes, which summarize the contribution of each knowledge source, and combining them in a dynamic decision process. The output, called a rich action description, combines a number of discrete decisions and can either be handed to an action planning module. Alternatively, it can be used to specify parameters for a dynamic simulation of the appropriate situation, resulting in additional inferences (Narayanan

1997). The resulting implementation is designed as a module in a multi-modal dialogue system called SmartKom (www.smartkom.org), allowing our approach to be evaluated against the conventional techniques currently employed by the system.

References

B. Bergen and N. Chang (2002). Simulation-Based Language Understanding in Embodied Construction Grammar. In J. Ostman (Ed.), *Construction Grammar(s): Cognitive and Cross-language dimensions*.

J. Feldman, S. Narayanan, and G. Lakoff (2000). Active Schemas for Generation, Recognition, and Understanding of Action, International Conference on Mirror Neurons and the Evolution of Language and Cognition, Dalmshurst, July 5-9.

A. Goldberg (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

S. Narayanan. Reasoning about Events (1999). Proceedings of the National Conference on Artificial Intelligence. AAAI-99. Orlando, Florida.

S. Narayanan (1997). KARMA: Knowledge-Based Active Representations For Metaphor and Aspect. Ph.D. Dissertation, Computer Science Division, University of California, Berkeley.

S. Narayanan and D. Jurafsky (1998). Bayesian Models of Human Sentence Processing. Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society, COGSCI-98, Madison.

R. Porzel and M. Strube (2000). Towards Context Adaptive Natural Language Processing. In Proceedings of the International Symposium: Computational Linguistics for the New Millennium, University of Heidelberg.