

Semantic interpretation of Dutch spoken dialogue

Jeroen Geertzen

Dept. of Communication & Information Sciences

Tilburg University, The Netherlands

`j.geertzen@uvt.nl`

1 Introduction

Semantic interpretation involves the process of ‘translating’ natural language to a representation of its meaning. It could be understood as the task of mapping syntax to semantics, assuming that the syntactic relationships in an utterance correspond to functional relationships in the meaning representation. Relevant work in this area often uses techniques from machine translation and machine learning in the mapping from natural language to meaning-representation languages (e.g. [9, 7]). These approaches can be robust, and thus would be useful in dealing with large quantities of utterances, but require large amounts of annotated data.

Since the syntax of natural language does not change much from domain to domain, an alternative way is to use the output of a wide-coverage syntactic parser as a basis for single, multiple, or even open-domain language processing. To obtain a sufficiently detailed semantic representation, the phrases in the parses should be linked with domain-specific knowledge concepts. In this paper, a syntactic parsing based system for the semantic interpretation of Dutch spoken language is presented and evaluated.

2 Data

Representation of semantic content is often expressed in some form of predicate logic type formula. Examples are varieties of description logics, which extend semantic frames and networks with a formal logic-based semantics that uses predicates. In any case, the semantic representation should ideally be powerful enough to take into account complexities such as negation, quantification, a certain degree of under-specification and (complex) modifiers to be interesting for use in advanced question answering systems and dialogue systems. Moreover, the logical form should be suitable to support feasible reasoning, for which also theorem provers, model builders, and model checkers can be used. Several semantic representations have been proposed that take these aspects into account, such as for example Quasi Logical Forms [1] and Dynamic Predicate Logic [6]. For the approach presented

here, a simplified first order logic is used similar to quasi logical forms. The dialogue data that is used for semantic interpretation consists of recorded interactions with a help desk on how to operate a fax device. Examples of resulting utterances and their corresponding semantic content, expressed by λ -expressions of first-order logic, are illustrated in the following table:

	<i>utterance</i>	<i>semantic content</i>
1	wat moet ik nu doen? (what do I have to do now?)	$\lambda x . \text{next-step}(x)$
2	druk op een toets (press a button)	$\lambda x . \text{press}(x) \wedge \text{button}(x)$
3	druk op de groene toets (press the green button)	$\lambda x . \text{press}(x) \wedge \text{button}(x) \wedge \text{color}(x, \text{'green'})$
4	wat doet de grote knop? (what does the big button do?)	$\lambda x . \text{function}(x) \wedge \text{button}(x) \wedge \text{size}(x, \text{'big'})$

Three types of predicate groups are distinguished: action predicates, element predicates, and property predicates. In the domain of operating a fax device, the predicates and arguments in the logical expressions refer to entities, properties, events, and tasks in the application domain. The application domain of the fax device is complex but small: the domain model consists of 70 entities with at most 10 properties, 72 higher-level actions or tasks, and 45 different settings.

3 Approach

The semantic representation is obtained in two stages. In the first stage, the utterances are syntactically parsed. In the second stage, the most probable derivation obtained in the syntactic parsing is used to construct the semantic representation.

For the syntactic interpretation of the utterances, the Alpino Parser is used [3]. This HPSG-based dependency parser aims to accurately provide full parses of unrestricted Dutch text and is publicly available.¹

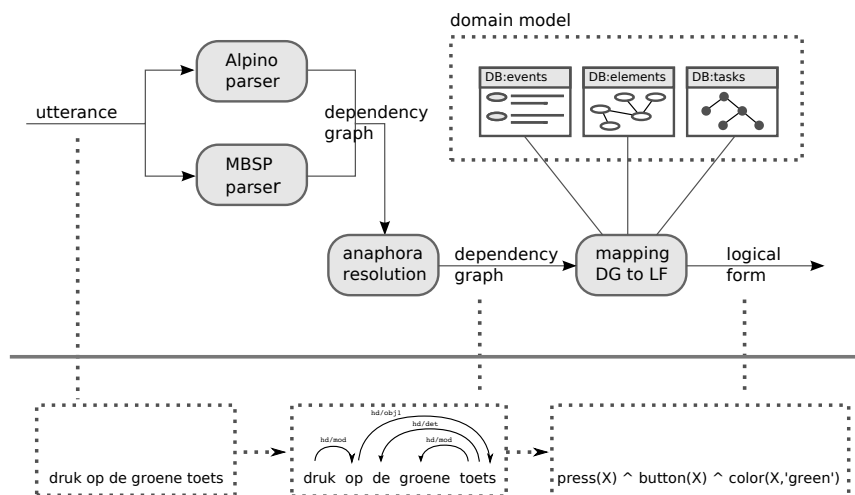
In the context of spoken dialogue processing, a syntactic parser has to deal with fragmented input and many syntactically less well-formed utterances in comparison to text parsing. For this reason, the utterances are additionally parsed with a shallow parser, and the resulting parse is used in case the Alpino parser fails to provide a full parse. As shallow parser, a memory based chunk parser trained for spoken Dutch [4] is employed.

To resolve pronouns, a simple pronoun resolution algorithm has been implemented. This algorithm is similar to the centering algorithm proposed

¹See: <http://www.let.rug.nl/~vannoord/alp/Alpino/>.

in [8]. While processing the utterances, each noun phrase identified is placed on a temporary queue which is pushed on a history stack once the utterance or turn is closed. Upon encountering a pronoun, the first element on the queue that meets gender and number agreement is selected as antecedent. If no candidate is found, the previous queue on the stack is evaluated until an antecedent is found or all queues on the history stack are evaluated.

The semantic representation is constructed by traversing the dependencies in the parse and by mapping words and phrases to domain concepts. These domain concepts are events, elements, and domain tasks stored in a database. This process of semantic interpretation is depicted and exemplified in the following figure:



The approach reported here has several aspects in common with that of Bos [2], who uses a CCG based parser [5] and assigns Discourse Representation Structures (DRSs) to the lexical categories used by the parser after which semantic construction is driven by the syntactic derivation. A notable difference is that Bos first constructs a DRS representation which is subsequently translated into first-order logic. Another difference is that in the approach described in this section, syntactic representations obtained by the wide-coverage dependency parser are complemented with that of a chunk parser, which increases robustness when dealing with fragmented input, common in spoken dialogue.

4 Evaluation

The approach for obtaining semantic representations has been tested on a dataset of 160 utterances and their corresponding semantic content. All utterances are related to the fax domain. The performance on identifying each of the three types of predicates in the semantic representations is specified

in the following table:

	<i>action predicates</i>	<i>element predicates</i>	<i>property predicates</i>	<i>overall</i>
accuracy (%)	92.2	81.4	94.3	88.1

The results show that identification of element predicates is the least successful. Where actions and properties are usually mentioned explicitly, domain elements can be described in various ways. Moreover, in a substantial number of utterances pronouns are used, which cannot always be resolved successfully. Nevertheless, an accuracy of 88.1% is achieved on recovering complete semantic representations.

References

- [1] Hiyan Alshawi. Resolving Quasi Logical Forms. *Computational Linguistics*, 16(3):133–144, 1990.
- [2] Johan Bos. Computational semantics in discourse: Underspecification, resolution, and inference. *Journal of Logic, Language and Information*, 13(2):139–157, 2004.
- [3] Gosse Bouma, Gertjan van Noord, and Robert Malouf. Alpino: Wide-coverage computational analysis of dutch. In *Proceedings of the 11th CLIN Meeting*, pages 45–59. Amsterdam, The Netherlands, 2001.
- [4] Sander Canisius and Antal van den Bosch. A memory-based shallow parser for spoken dutch. In *Proceedings of the 14th CLIN Meeting*, pages 31–45, Antwerp, Belgium, 2003.
- [5] Stephen Clark and James R. Curran. Parsing the WSJ using CCG and log-linear models. In *Proceedings of the 42nd Meeting of the ACL*, pages 103–110, Barcelona, Spain, 2004.
- [6] Jeroen Groenendijk and Martin Stokhof. Dynamic Predicate Logic. *Linguistics and Philosophy*, 14(1):39–100, 1991.
- [7] Rohit J. Kate and Raymond J. Mooney. Using string-kernels for learning semantic parsers. In *Proceedings of the ACL-COLING*, pages 913–920, Sydney, Australia, 2006.
- [8] Joel R. Tetreault. A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics*, 27(4):507–520, 2001.
- [9] Yuk W. Wong and Raymond Mooney. Learning for semantic parsing with statistical machine translation. In *Proceedings of the HLT/NAACL*, pages 439–446, NY, USA, 2006.