HARRY BUNT AND REINHARD MUSKENS

COMPUTING THE SEMANTIC INFORMATION IN AN UTTERANCE

1. Introduction

To compute the meaning of any given natural language expression is extremely hard. This is partly due to the structural complexity, variability, and flexibility of natural language, but also, and more importantly, to its pervasive ambiiguity . It has been estimated that, due to the fact that individual words as well as their combination in a sentence can usually express a range of semantic concepts and relations, an ordinary sentence of average length can have several million possible meanings. So what do we mean by *'the meaning* of a given natural language expression'?

Language users are hardly ever aware of having to resolve an ambiguity, so in practice the understanding of a natural language expression does not mean choosing 'the right interpretation' among millions of possibilities. The crucial point is, of course, that as a language user we are never confronted with the task of computing the meaning of a sentence in splendid isolation. That happens only in the linguistic literature. In reality, natural language always occurs in a certain context. In a given context we are not talking about just anything, but we have a certain domain of discourse. This means that those word senses can be ruled out that do not belong to this domain, as well as those interpretations of structural ambiguities which express something that would be impossible or highly implausible in the domain.

While the fact that speakers and listeners do not struggle with ambiguity resolution may suggest that context information is sufficient to determine the intended meanings of natural language expressions, it is hard to believe that context information is really sufficient to exclude millions of potential sentence meanings and retain exactly one of them as *the meaning*. It seems more likely that readers and listeners disambiguate meanings to the extent that is required by the circumstances, and that speakers and writers, expecting this of their listeners and readers, formulate their utterances accordingly. In other words, a context comes with certain demands on the precision with which meanings should be computed, and so, in a given context what should be regarded as *the meaning* of a given sentence is something that contains a certain amount of ambiguity or incompleteness. In particular, the context of use determines the appropriate level of granularity for referring expressions, and thus for the permitted vagueness of reference. Frazier and Rayner (1990) report on empirical evidence supporting this intuition for the use of polysemous nouns, whose reference is shown to be often unresolved until information later in a sentence provides disambiguating information, in contrast to the case of homonyms, where an interpreter must make a choice that may turn out to be wrong, leading to garden path sentences. Poesio et al. (2006) provide evidence that under certain conditions a certain amount of 'sloppiness' is permitted in anaphoric reference. A consequence of this view is that a representation of the meaning of a given natural language expression in a given context is not a semantic representation in the classical Montagovian sense, being fully specified, complete, and unambiguous, but is an underspecified semantic representation, that leaves room for ambiguity, vagueness, and incompleteness.

Underspecified semantic representations (USRs) can be regarded not just as imperfect representations of meaning, waiting to become fully specified, but also as a way to salvage the adagio of compositionality: rather than saying that the semantic representation of a sentence represents its meaning, determined by the meanings of its components and its syntactic-semantic composition, we could say that a USR represents the meaning of a sentence *in sofar as determined by the meanings of its components and its syntactic-semantic composition*, or, in other words, a USR captures the semantic information contained in a sentence through its components and its syntactic-semantic composition. In the latter form, the compositionality assumption is no longer a thesis, that can be proved wrong or correct, but simply describes what one does when computing a USR: one computes the semantic information that a given sentence contains, and gives that a formal representation.

A slightly different way of looking at the interpretation of a sentence in context, i.e. of an *utterance*, is to observe that a human interpreter does not so much compute the contextually most appropriate meaning of the utterance, but computes the semantic information that the utterance contributes to the contextual knowledge that the interpreter already has. This view corresponds with the idea that understanding an utterance means trying to integrate the information that it conveys with the rest of one's knowledge, a classical notion in artificial intelligence that is compatible with the approach to utterance meaning currently popular in dialogue studies, generally known as the *information-state update* or *context-change approach* (Smith & van Kuppeveld, 2003 Bunt, 2000; Traum & Larsson, 2003), and to some extent also with the basic ideas of dynamic semantics (Groenendijk & Stokhof, 1985; see also the discussion in Bunt, 1989 and Groenendijk & Stokhof, 1989). If utterance meanings are in general underspecified, as we just argued, then the way in which context models ('information states') are updated by utterance meanings is in general ambiguous, and incomplete... a consequence which researchers who follow this approach have so far not taken on board.

Returning to the view on meaning and underspecification which focuses on computing the semantic information in a given utterance, it may be noted that this view comes very close to that of modern approaches to semantic annotation. Traditionally, annotation is the enrichment of text with notes on some of its properties or background. In computational linguistics annotation has usually taken the form of labelling text elements with certain tags, such as part of speech tags. Semantic annotation is taking a somewhat different turn, where the annotations that are added to a text are supposed to be expressions in a formal language with a well-defined semantics (see Bunt & Romary, 2002; 2004). The reason for this is that, different from other types of information in annotations, semantic annotations are intended to support not only the retrieval of certain text elements, but *reasoning* as well. A clear case is presented by the annotation of the temporal information in a text. If, for instance, a question-answering system is asked

(1) What new products did Microsoft announce in the last quarter?

and the data for providing answers contain a newspaper item, dated 12 May 2006, and stating that:

(2) Microsoft announced its XP follow-up system yesterday at a press meeting in San Francisco.

then the answer to this question should include the new operating system, as a result of reasoning that *yesterday* in this case is the same as May 11, 2006; that May belongs to the second quarter of a year: and that *last quarter* refers to the second quarter of 2006, since the question was asked at a date belonging to the third quarter. It is therefore insufficient to simply tag temporal expressions as being temporal expressions, for instance. Questions that are input to the system must also be time-stamped, and the time stamp should have a well-defined internal structure allowing the recognition of a month and year, which can then be compared to the creation dates of the documents in the database. Moreover, the date referred to by *yesterday* must be annotated not just as a date, but as the date on which the event occurred that is described in the corresponding sentence. Clearly, in order to support the fairly complex inferencing that is needed in such examples, the annotations have to meet syntactic and semantic requirements that go a long way beyond those of just labelling.

The interesting point is here that semantic annotations are developing into formal representations of some of the semantic information contained in the sentences in a text. That makes them formally comparable to underspecified semantic representations, the main difference being that annotations tend to focus on a particular type of semantic information, such as temporal information or semantic roles, whereas USRs are typically intended to capture *all* the semantic information in a sentence.

2. About this book

Following the present introductory chapter, the book continuess with four chapters concerned with aspects of ambiguity, vagueness, and underspecification. The chapter by Massimo Poesio, Uwe Reyle and Rosemary Stevenson, entitled *Justified sloppiness in anaphoric reference*, takes up the issue of the ambiguity that speakers and listeners allow in the meanings of what they say, focusing on the use of anaphoric expressions. They analyze a corpus of spoken dialogues to identify cases in which the addressee of an utterance containing an anaphoric pronoun does not appear to have enough evidence to resolve that pronoun, yet doesn't appear to find the pronominal use infelicitous. The two patterns of anaphoric use that were found to fit these conditions suggest three conditions under which *justified sloppiness* in anaphoric references is not perceived as infelicitous. Preliminary controlled experiments indicate that subjects do find anaphoric pronouns that satisfy the justified sloppiness conditions significantly easier to process than pronouns occurring in minimally different contexts in which these conditions are not satisfied.

In the second chapter of this group, by Aoife Cahill, Mairead McCarthy, Michael Burke, Josef van Genabith and Andy Way, *Deriving*

Quasi-Logical Forms from f-structures for the Penn Treebank, the authors show how the trees in the Penn-II treebank can be associated automatically with simple Quasi-Logical Forms (QLFs). Their approach is based on combining two independent strands of work: the first is the observation that there is a close correspondence between QLFs and LFG's f-structures (van Genabith and Crouch, 1996); the second is the development of an automatic f-structure annotation algorithm for the Penn-II treebank (Cahill et al, 2002a; Cahill et al, 2002b). The approach is compared with that of (Liakata and Pulman, 2002).

In the chapter, *Which underspecification technique for what purpose?*, Harry Bunt examines a number of techniques for underspecification in semantic representations, notably labels and holes, ambiguous constants, metavariables, dominance constraints, radical reification, stores, lists and disjunctions, and *in situ* quantified terms. These techniques are considered for their usefulness in dealing with a variety of linguistic phenomena and cases of incomplete input, which have motivated the use of underspecified semantic representations. It is argued that labels and constraints, and the use of ambiguous constants and variables, have nearly disjoint domains of application, and together cover a wide range of phenomena.

In the last chapter of this group, Alex Lascarides and Nicholas Asher motivate and describe Segmented Discourse Representation Theory (SDRT) as a dynamic semantic theory of discourse interpretation, using rhetorical relations to model the semantics/pragmatics interface. They describe the syntax and dynamic semantics of the SDRT language in which logical forms are represented, a separate but related language in which semantic underspecification is expressed as partial descriptions of logical forms, and a glue logic which uses commonsense reasoning to construct logical forms, relating the semantically underspecified forms that are generated by the grammar to their pragmatically preferred interpretations. The framework is applied to examples involving anaphora and other kinds of semantic ambiguities. Being concerned with the analysis of discourse, this chapter forms a bridge to the next three chapters, that are also concerned with the semantic interpretation of discourse and dialogue.

Raquel Fernández, Jonathan Ginzburg, Howard Gregory and Shalom Lappin present the main features of SHARDS, a semantically-based HPSG approach to the resolution of dialogue fragments. This implemented system interprets short questions ('sluices') and short answers. It provides a procedure for computing the content values of clausal fragments from contextual information contained in a discourse record of previously processed sentences.

Ivana Kruijff-Korbayova and Bonnie Webber describe an investigation into the sensitivity of discourse connectives to the Information Structure (IS) of the utterances they relate, in their chapter *Interpreting concession statements in light of information structure.* They illustrate this with an analysis of connectives signaling concession, distinguishing two senses – *denial of expectation* and *concessive opposition.* Their account thus refines earlier accounts that ignore IS. This work is part of a larger enterprise aimed at understanding what role(s) sentence-level IS plays in the interpretation of larger units of discourse.

Key ingredients in the description of discourse meaning are reference markers: objects in the formal representation that the discourse is about. It is well-known that reference markers are not like first-order variables; the received view is that reference markers are like the variables in imperative programming languages. However, in a computational semantics of discourse that treats reference markers as 'dynamically bound' variables, every noun phrase will get linked to a dynamic variable, so it will give rise to a marker index. In the chapter *Context and the composition of meaning*, Jan van Eijck addresses the question where these indices come from, and how they can be handled when combining (or 'merging') pieces of discourse. He argues that reference markers are better treated as indices into context, and presents a theory of context and context extension based on this view. In context semantics, noun phrases do not come with fixed indices, so the merge problem does not arise. This solves a vexing issue with coordination that causes trouble for all current versions of compositional discourse representation theory.

In the chapter *Meaning, intonation and negation*, Marc Swerts and Emiel Krahmer outline an approach to the study of meaning and intonation. The approach focusses both on what speakers can do, using production experiments, and on what hearers can do, using perception experiments. They show that such an experimental paradigm may yield interesting results from a semantic point of view, discussing the role intonation can play for the interpretation of negation phrases in natural language. Empirical evidence is presented for the existence of a set of prosodic differences between two kinds of negations, descriptive and metalinguistic ones. This distinction has been the subject of considerable debate in presupposition theory and also plays an important role in discussions about the division of labor between semantics and pragmatics. In general, it is argued that intonation gives rise to 'soft constraints', and that an optimality-theoretical framework may be suitable to model the relation between intonation and meaning.

Myroslava Dzikovska, Mary Swift and James Allen in their chapter *Customizing meaning: Building domain-specific semantic representations from a generic lexicon*, argue that language input to practical dialogue systems must be transformed into a semantic representation that is customized for use by the back-end domain reasoners, while at the same time one wants to keep front-end system components as domain-independent as possible for easy portability across multiple domains. They propose a transparent way to achieve domain-specificity from a broad-coverage domain-independent parser. They define a set of mappings from ontologies into domain-specific knowledge representations, and use these mappings both to customize the semantic representations output by the parser for the reasoners, and to specialize the lexicon to the domain – which improves parsing speed and accuracy. This method facilitates instances of semantic type coercion common in many domains by combining lexical representations with domain-specific constraints on interpretation.

The chapter by Aravind Joshi, Laura Kallmeyer and Mariel Romero addresses the problem of formulating constraints for relative quantifier scope, in particular in inverse linking readings where certain scope orders are excluded. They show how to account for such restrictions in the Tree Adjoining Grammar (TAG) framework by adopting a notion of ‘flexible composition’. In the semantics used for TAG they introduce quantifier sets that group quantifiers that are ‘glued’ together in the sense that no other quantifier can scopally intervene between them. The flexible composition approach allows them to obtain the desired quantifier sets and thereby the desired constraints for quantifier scope.

The next three chapters are concerned with the expression of time in natural language. In the first of these, *Serious computing with tense*, Fabrice Nauze and Michiel van Lambalgen describe a comprehensive proposal for dealing with time and events as expressed in natural language. They argue that the simple davidsonian addition of time and event variables to predicates in the representation language is insufficient for reasoning about time and events, and argue that a theory of time and events should have great expressive power and be presentable in axiomatic form, so that it is entirely clear what it predicts and what it doesn’t. They argue that the event calculus that has been developed in robotics (Shanahan, 1997) has all the desired properties. It allows one to formulate a goal and a causal theory of the domain. Based on the causal theory, a plan for reaching that goal can be inferred. In the version of the event calculus proposed by van Lambalgen and Hamm

(2003), the inference mechanism is constraint logic programming with negation as failure. The authors propose to apply this formalism to tense and aspect, since goals seem to play a prominent part there. For example, a profitable way to formulate the meaning of accomplishments is to specify a goal and a causal theory, which together yield a plan which achieves the goal when no unforeseen circumstances occur. This prevents the so-called 'imperfective paradox' to arise in case the goal is for some reason never achieved.

The second chapter in this group, by James Pustejovsky, Robert Knippen, Jessica Litman and Roser Saurí, *Temporal and event information in natural language text*, discusses the role that temporal information plays in natural language text, specifically in the context of question-answering systems. A descriptive framework is defined for examining the temporally sensitive aspects of natural language queries. The properties are investigated that a general specification language would need to have, in order to mark up temporal and event information in text. The language TimeML is presented, a rich specification language for event and temporal expressions in natural language text. The chapter shows the expressiveness of TimeML for a broad range of syntactic and semantic contexts, and demonstrates how it can play an important part in the development of more robust question-answering systems.

In the third of this group of chapters, *Finite-state descriptions for temporal semantics*, Tim Fernando outlines finite-state descriptions for temporal semantics through which to distinguish 'soft' inferences, reflecting manners of conceptualization, from more robust semantic entailments defined over models. Fernando argues that just what descriptions are built (before being interpreted model-theoretically) and how they are grounded in models of reality, explains upon examination why some inferences are soft and others are robust.

The next two chapters are both concerned with semantic aspects of language generation. Claire Gardent and Kristina Striegnitz in their chapter *Generating bridging definite descriptions* focus on the role that knowledge based reasoning plays in the generation of definite descriptions. Specifically, they propose an extension of Dale and Reiter's incremental algorithm which covers not only directly anaphoric descriptions, but also indirect and associative anaphora. Starting from a formalism independent algorithm, they further show how this algorithm can be implemented using description logic.

Kees van Deemter and Emiel Krahmer explore in their chapter *Graphs and booleans: On the generation of referring expressions* how a graph-theoretical perspective may be brought to bear on the generation of *complex* referring expressions. The motivation for this exploration was that, if each of these types of referring expressions can be addressed using one and the same formalism, then this will make it easier to compare and, ultimately, to combine them into one unified algorithm. They sketch how relations, vague properties, and Boolean operators have been tackled by earlier algorithms, and ask how these algorithms can be formalised using a graph-theoretical approach. it is shown that most of the existing algorithms carry over without difficulty, through the technique of making implicit properties explicit in the knowledge base. However, in the case of one algorithm (which focusses on the generation of Boolean descriptions that also contain relational properties), this strategy turns out to be problematic. For this case, a new algorithm is presented, based on partitioning the target set, which can be implemented in a graph-theoretical formalism without difficulty.

The last two chapters of the book return to issues of underspecification and interpretation in context. In the first of these, *Efficient computation of overlay for multiple inheritance hierarchies in discourse modeling*, Jan Alexandersson and Tilman Becker note that default reasoning has been shown to be a convenient means for interpreting user utterances in context and introduce ‘overlay’, the combination of default unification and a scoring function where the latter is used for computing the degree of similarity between new and old information. In this work they continue their efforts for default unification of typed feature structures by giving an efficient algorithm for multiple inheritance hierarchies. The main contribution of this chapter is that, contrary to previous suggestions, most of the computation can be done on the type hierarchy. The scoring function is adapted accordingly.

The final chapter, by Dick Crouch, Anette Frank and Josef van Genabith is concerned with an application of underspecified semantic representations, namely ambiguity-preserving machine translation. In earlier work (van Genabith *et al.*, 1998), the authors developed an approach where transfer takes place on the glue language meaning constructors of (Dalrymple *et al.*, 1996); unfortunately, that approach was unable to deal with structural misalignment problems, such as embedded head switching, in a satisfactory way. This chapter proposes the use of a fragment of linear logic as a transfer formalism, and shows how it provides a more general and satisfactory solution to these problems.

References

Bunt, H.: 1989. Dynamic Interpretation in Text and Dialogue. In H. Bouma and B. Elsendoorn, editors *Working Models of Human Perception*. New York: Academic Press. pp. 81–150.

Bunt, H.: 2000. Dialogue pragmatics and context specification. In H. Bunt and W. Black, editors *Abduction, Belief and Context in Dialogue. Studies in Computational Pragmatics*. Benjamins, Amsterdam. pp. 419–455.

Cahill, A., M. McCarthy, J. van Genabith and A. Way: 2002a. Automatic Annotation of the Penn Treebank with LFG F-Structure Information. In *Proc.LREC Workshop on Linguistic Knowledge Acquisition and Representation,*

Cahill, A., M. McCarthy, J. van Genabith and A. Way: 2002b. Evaluating Automatic F-Structure Annotation for the Penn-II Treebank in *Proceedings of the Treebanks and Linguistic Theories Workshop*, Sozopol, Bulgaria.

Dalrymple, M., Lamping, J., Pereira, F.C.N. and Saraswat, V.: 1996. Quantification, anaphora, and intensionality. *Journal of Logic, Language and Information* 6(3) 219–273.

Frazier, L. and K. Rayner: 1990. Taking on Semantic Commitments: Processing Multiple Meanings vs. Multiple Senses. *Journal of Memory and Language* 29, 181–200.

Genabith, J. van, and D. Crouch: 1996. Direct and Underspecified Interpretations of LFG f-Structures. In *Proceedings COLING 96*, Copenhagen, Denmark, pp. 262–267.

Genabith, J. van, Frank, A. and Dorna, M.: 1998. Transfer constructors. In Butt, M. and King, T. H., editors, *Proceedings of the LFG'98, Conference Brisbane, Australia*. CSLI Publications, pp. 190–205.

Groenendijk, J., and M. Stokhof: 1985. Dynamic Predicate Logic. ITLI Report, Amsterdam: ITLI.

Groenendijk, J., and M. Stokhof: 1989. Context and Information in Dynamic Semantics. In H. Bouma and B. Elsendoorn, editors *Working Models of Human Perception*. New York: Academic Press. pp. 457–486.

Liakata, M. and S. Pulman: 2002. *From trees to predicate-argument structures.* COLING'02, Proceedings of the Conference, Taipei.

Lambalgen, M. van, and F. Hamm: 2003. Moschovakis' notion of meaning as applied to linguistics. In M. Baaz and J. Krajicek, editors, *Logic Colloquium '01.* ASL Lecture Notes in Logic, Wellesley, MA.: A.K. Peters.

Shanahan, M.P.: 1997. *Solving the frame problem.* The M.I.T. Press, Cambridge MA.

Smith, R. and J. van Kuppevelt editors,: 2003. *Current and new Directions in Discourse and Dialogue*. Dordrecht: Kluwer.

Traum, D., and S. Larsson: 2003. The Information State Approach to Dialogue Management. In R. Smith and J. van Kuppevelt, editors,: 2003, *Current and new Directions in Discourse and Dialogue*. Dordrecht: Kluwer, pp. 325–353.