SESSION 4: NATURAL LANGUAGE

Robert C. Moore, Chair

Artificial Intelligence Center
SRI International
Menlo Park, CA 94025

Collectively, the papers in this session are mainly concerned with parsing, semantic interpretation, and inference. Interest in these processes can be motivated if we recognize that the overall goal of the field of NLP is the manipulation of natural language in ways that depend on meaning. Parsing, the recovery of the linguistic or grammatical structure of a natural-language utterance, is of concern to NLP because, in general, the meaning of a natural-language utterance depends on its structure. An example that illustrates this point is the sentence *The man with the umbrella opened the door.* If we tried to process this sentence without paying attention to its grammatical structure, we could easily be misled by the fact that it contains the substring *the umbrella opened.* But this sentence has nothing to do with umbrellas opening. Because of the structure of the sentence, the *umbrella* must be grouped with *with* and not with *opened.*

If the central concern of NLP is the manipulation of natural language in ways that depend on meaning, then semantic interpretation would naturally be expected to play a central role. In practice, semantic interpretation in NLP usually means recovering a representation of the meaning of an utterance that encodes that meaning in a more transparent way than does the utterance itself. How does this contribute to the goal of manipulating language in meaning-dependent ways? We want to have algorithms that manipulate language according to meaning, but meaning is ultimately an abstraction that algorithms can have no direct access to. Algorithms can directly manipulate expressions only according to their structure. Thus we need expressions whose structure corresponds in a very direct way to their meaning. While, as we have argued above, the meaning of a natural-language expression is dependent on its structure, this dependence can be very indirect. By recovering an expression that encodes the meaning of an utterance more directly, we can create modular algorithms that consist of interacting pieces that each look only at a small piece of the structure of the meaning representation. If the pieces of the meaning representation fit together in a natural way that reflects the overall meaning of the utterance, then the algorithms that manipulate them will also be able to fit together in a natural way that reflects the overall meaning of the utterance.

Finally, inference is the pay-off for the previous phases of parsing and semantic interpretation, being the canonical example of a form of manipulation of natural-language that depends on the meaning of utterances. In fact, a strong argument could be made that inference is a process on meanings and not on natural-language expressions per se.

With this as background, we can briefly consider some of the major issues raised by the papers in this session. Some of the most important issues currently being raised about parsing are how complete it needs to be and how complex a structure needs to be recovered for different applications. The paper by McCord takes a fairly traditional view, attempting to recover a complete structural description of any sentence presented to the parser. In the paper by Hobbs, et al., parsing is much more fragmentary, attempting only to recover the structure of pieces of a text that are critical for the particular application. Moreover, the structures recovered in the Hobbs paper are simple enough to be characterized by finite-state automata, while the structures described in the McCord paper are more complex. A second parsing issue, which forms the focus of the McCord paper, is the problem of ambiguity. In parsing, we are given a string of words or tokens, and we have to recover the grammatical structure, but there may be many structures compatible with a given string. McCord, then, addresses the issue of how to find the most likely stucture out of all the ones that are possible.

Issues of semantic interpretation are of greatest concern in the paper by Hwang and Schubert. The type of work reported in this paper can perhaps be best appreciated by keeping in mind some central methodological principles that are ofter used to guide work on semantics. Having such principles is important because of the lack of clear intuititive agreement about the adequacy of semantic representations. Speech recognition, in contrast, is methodologically much simpler than semantics because of the enormous intersubjective agreement as to

what strings of words most speech signals correspond to. While there are particular cases where the proper transcription of a signal can be argued about, in most cases this is simply not a problem. No such intuitive agreement exists in the field of semantics. It is something like speech recognition might be if there were no written languages and no general agreement on segmentation of speech into words.

So, in semantic interpretation, there are two methodological principles that have come to be used as a means of evaluating the adequacy of proposed analyses. The first is that one should be able to give a mathematical, "model theoretic" interpretation of the formal expressions used to represent the meaning of natural-language expressions. This gives a way to decide whether there is really any basis for the claim that the representations in question actually do capture the meaning of the corresponding natural-language. The main alternative seems to be what is sometimes referred to as "pretend it's English" semantics, where one reads the tokens that appear in the representation as if they are English words and sees whether it sounds like it means what is desired—not a very satisfactory state of affairs. A second methodological principle in semantic interpretation is that of compositionality—the slogan being, "the meaning of the whole must be a function of the meaning of the parts." This principle reflects the fact that it is not sufficient just to be able to represent formally the meaning of natural-language expressions; it must be possible to produce them in a systematic way from the natural language. In the Hwang and Schubert paper, the representations used may seem quite complex to someone outside the field, but that complexity is motivated by the need to satisfy these methodological constraints.

In Vilain's paper, the major issue adressed is the tradeoff between expressiveness in a representation formalism and the tractability of the inference problem for that formalism. It is notorious that the more expressive a representation language is, the more computationally complex the inference problem for it is. Vilain looks at whether for a certain type of application, the expressions in the representation language can be limited to a normal form which is known to be computationally tractable.

There are also a number of issues that cut across all phases of processing. One such issue is to what degree systems can be made language and domain independent. The ideal is for the algorithms to be both language and domain independent, with a declaratively specified grammar and lexicon that is language dependent but domain independent, and a final domain-dependent module that interfaces the language processing to the application. The paper by Aone, et al., explores how well this model works in a real multi-lingual data extraction system. A second issue is that of hand coding versus automatic extraction of the knowledge required for NLP systems. Almost all the knowledge embodied in the systems described in this session is hand-coded, while the emphasis in Session 8 is on systems that use methods for automatic extraction. Often this issue is conflated with the issue of whether the knowledge in question is represented by symbolic rules or by numerical parameters such as probabilities, but it is worth pointing out that the paper by Brill in Session 8 uses symbolic rules, but extracts them automatically from a corpus. Finally, several of these papers raise the question of how to evaluate the work reported on. This has come to be recognized as a central methodological issue in the field, and the McCord, Hobbs, and Vilain papers all address the problem in one way or another.