Hypergraph Unification-Based Parsing for Incremental Speech Processing

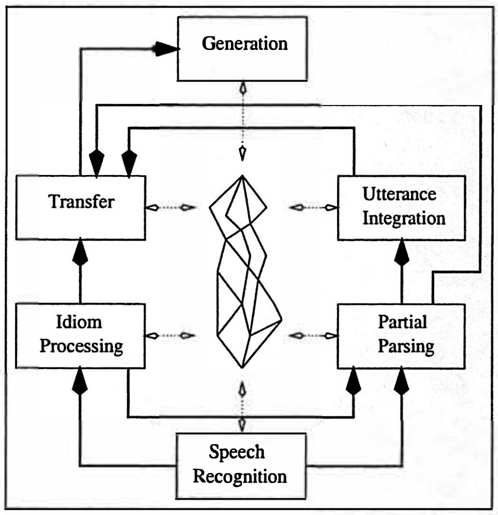
Jan W. Amtrup
Computing Research Laboratory
New Mexico State University
email: jamtrup@crl.nmsu.edu

Parsing word graphs is hard. Even though they provide an extremely compact way of representing recognition results (Aubert and Ney, 1995), they still consist of a large number of edges. The ambiguity on the lowest level (words of an utterance) can be high, depending on the experimental setting for the speech recognition system. Parsing word graphs that were produced using an incremental speech recognition system is even harder, since they can be ten times larger than their non-incremental counterparts. However, for highly sophisticated and natural applications, incremental recognition is a sine qua non.

The complexity of incremental word graphs can be partly overcome by converting them to hypergraphs. Here, edges may have several start and end vertices. Following Weber (1995), we call these sets *families of edges.*

Word graphs and their generalizations as hypergraphs are instances of a chart-like structure (Kay, 1980). Similar to standard charts, hypergraphs constructed by inserting word hypotheses are directed, acyclic, rooted graphs. Thus, the hypergraph resulting from this process can be viewed as the lowest level of representation for chart-based speech processing — it relates to the graph that contains only preterminal edges.

We are using a hypergraph as underlying data structure in the speech translation system MILC (*M*achine *I*nterpreting by *L*ayered *C*harts) (Amtrup, 1999). This system was designed as a prototype to demonstrate the feasibility of incremental speech translation and as an experimental platform for the research into the interaction between different processing stages. The figure shows an outline of the architecture. MILC uses typed feature structures with appropriateness for the representation of all linguistic structures in the system. In order to be applicable for a distributed system, we developed an array-based encoding for feature structures (Carpenter and Qu, 1995). This kind of representation allows for an easy exchange of feature structures across machine boundaries.



The structural analysis of hypergraphs constructed incrementally while the parsing processes are already running poses some interesting questions with regard to the consistency of the overall operation of the system. Within such a system, the quality (score) of a piece of information may change over time. This is partly due to further incoming recognition results, and partly due to penalties inflicted upon certain hypotheses by individual parsing processes.

The structural analysis within MILC consists of three components with different purposes and descriptive power. First, idioms are recognized. In our case, idioms are fixed expressions without any inflectional or order variation. This stage performs a fast incremental graph search by intersecting the input hypergraph with a graph describing the known idioms and compounds.

The second component performs an incremental left-to-right parse of an utterance, with an emphasis on the syntactic structure of limited size constituents (Abney, 1991).

The following integration component is implemented as a bidirectional island-parser, capable of extending hypotheses to the left. This is useful, e.g., if the subcategorization frame of a verb is known late due to a verb-last position in an utterance. In our experiments, employing a second parser for incorporating complements to the left resulted in a speedup of ten as compared to a strict left-to-right approach. The principle question with this method, however, is how much of the incrementality is lost using islands due to delayed attachment of elements to the left. In practice, this delay involves only a few calls to the fundamental chart parsing rule, mainly to incorporate complements into verb phrases. In theory, the worst case would arise if the island was the rightmost element of every rule. Suppose that a rule has n elements on the right hand side, then $n - 1$ more elements would have to be incorporated than with the left-to-right approach (the standard approach has to incorporate these as well, but does that before the last element is reached). Thus, at any given point in the chart, the additional effort is bound by the number of elements on the right hand side of a rule, which in turn is bound by the number of words to the left.

The combination of three methods for structurally analyzing an input utterance yields a reasonably high performance, in our tests six-fold real time for a machine translation of spontaneous speech input.

References

Steven Abney. 1991. Parsing By Chunks. . In Robert Berwick, Steven Abney, and Carol Tenny, editors, *Principle-Based Parsing.* Kluwer Academic Publishers, Dordrecht.

Jan W. Amtrup. 1999. *Incremental Speech Translation.* Number 1735 in Lecture Notes in Artificial Intelligence. Springer Verlag, Berlin, Heidelberg, New York.

Xavier Aubert and Hermann Ney. 1995. Large Vocabulary Continuous Speech Recognition Using Word Graphs. In *ICASSP 95.*

Bob Carpenter and Yan Qu. 1995. An Abstract Machine for Attribute-Value Logics. In *Proceedings of the 4^{th} International Workshop on Parsing Technologies (IWPT95)*, pages 59–70, Prague. Charles University.

Martin Kay. 1980. Algorithmic Schemata and Data Structures in Syntactic Processing. Technical Report CSL-80-12, Xerox Palo Alto Research Center, Palo Alto, CA.

Hans Weber. 1995. *LR-inkrementelles, Probabilistisches Chartparsing von Worthypothesengraphen mit Unifikationsgrammatiken: Eine Enge Kopplung von Suche und Analyse.* Ph.D. thesis, Universität Hamburg.