Multicomponent Tree Adjoining Grammars

David Weir
Department of EECS
Northwestern University
USA
weir@weir.eecs.nwu.edu

Multicomponent Tree Adjoining Grammars (MCTAG) are intended as a way to extend the domain of locality of Tree Adjoining Grammars. A MCTAG is made up of a set of elementary tree sets and at each step in a derivation all of the trees in a set must be adjoined together. This operation is called multicomponent adjunction. In TAG relationships can be stated between nodes within the same elementary tree. In MCTAG additional expressive power is achieved since relationships can be stated between nodes in different trees that are in the same elementary tree set. Several alternative notions of a MCTAG derivation are possible depending on the size of the domain into which the trees in a set are adjoined during a derivation.

When the domain of multicomponent adjunction is a single elementary tree we show that the system has the same generative capacity as TAG. Many of the cases in which MCTAG have been used to give linguistic analyses assume this version of multicomponent adjunction.

When multicomponent adjunction is local to the nodes of trees in an elementary tree set (we call this local MCTAG) then additional generative power results. We show that the class of string languages generated is larger and depends on the number of trees in the largest elementary tree set. The class of tree sets generated is larger: we show that it is possible to generate tree sets with dependent branches and sets whose path sets are more complex than those of TAG. We have shown (Weir 1988) that this form of MCTAG is weakly equivalent to the Linear Context-Free Rewriting Systems (Vijay-Shanker, Weir and Joshi 1987). One consequence of this is that polynomial recognition of local MCTAL is known to be possible. The derivations of local MCTAG can be represented with trees using a similar approach to that used for TAG. We show that the set of derivation trees for a local MCTAG is a local set, i.e., can be generated by a Context-Free Grammars.

The final case is one in which the domain of multicomponent adjunction is unconstrained (we call this non-local MCTAG). At this point, it is an open question as to how the generative power of this versions of MCTAG relates to the others. We show how the derivations of non-local MCTAG can be naturally represented with acyclic multigraphs.

References

Vijay-Shanker, K., Weir, D. J., and Joshi, A. K., Characterizing structural descriptions produced by various grammatical formalisms. In Proceedings of 25th meeting of Association of Computational Linguistics, July 1987.

Weir, D. J. Characterizing Mildly Context-Sensitive Grammar Formalisms, Ph.D. Thesis, University of Pennsylvania, 1988.