Synchronous Grammars and Transducers: Good News and Bad News

Stuart M. Shieber
School of Engineering and Applied Sciences
Harvard University
Cambridge MA 02138
USA
shieber@seas.harvard.edu

Much of the activity in linguistics, especially computational linguistics, can be thought of as characterizing not languages simpliciter but relations among languages. Formal systems for characterizing language relations have a long history with two primary branches, based respectively on tree transducers and synchronous grammars. Both have seen increasing use in recent work, especially in machine translation. Indeed, evidence from millennia of experience with bilingual dictionaries argues for synchronous grammars as an appropriate substrate for statistical machine translation systems.

On the positive side, some new results have integrated the two branches through the formal-language-theoretic construct of the bimorphism. I will present some background on this integration, and briefly describe two applications of synchronous grammars: to tree-adjoining grammar semantics and to syntax-aware statistical machine translation.

On the negative side, algorithms for making use of these formalisms are computationally complex, perhaps prohibitively so. I will close with a plea for novel research by the parsing technology community in making the systems practical.

Proceedings of the 10th Conference on Parsing Technologies, page 93,
Prague, Czech Republic, June 2007.