Seth Kulick

Statement of Research Interests

Dissertation

My dissertation work was an interdisciplinary project in computer science and linguistics. The two major claims made in this work are that:

- There is a correlation between desirable formal properties of grammars and constraints on non-local dependencies in syntax. That is, for a certain family of formal systems, those that retain desirable formal properties also derive linguistic constraints on movement.

- The different syntactic constraints on different types of non-local dependencies do not need to be independently stated; rather, they all all follow from more general properties of how syntactic units of a very small size are combined, with their differences following from the details of constraints on these syntactic units.

These claims were made in the context of an investigation of reduced constructions in Romance (e.g., clitic climbing) and German (e.g., long distance scrambling) from the perspective of Tree Adjoining Grammar (TAG). TAG has been studied for many years, and is of interest for both linguistic reasons, for the way in which it enforces locality violations, and for computational reasons, for its nice mathematical and parsing properties. However, attempts to handle reduced constructions have led to revised systems that lose some of these attractive linguistic and computational properties.

In my work I proposed a revised system to handle these problematic cases without losing the desirable properties. I showed that by expanding the data set and also by examining closer some of the core cases previously applied to TAG, the problems caused by reduced constructions can be seen as just another case of a structural problem already existing in TAG. This led to a unified solution that was able to retain the desirable properties. Within this unified solution I presented arguments for the two claims listed above.

Future Goals

- Following my dissertation work, I plan to continue the exploration of constrained mathematical formalisms for natural language analysis and the interaction between formal properties and linguistic constraints. There are many syntactic constructions which pose a challenge to the claims I have made. Other researchers have proposed analyses for these constructions, such as German verb clusters or cases of “remnant movement” in German, Dutch, and Hungarian, that cannot be implemented within this modified version of TAG. How serious a matter this is must be investigated. It is possible that it may lead to changes in the formal system with undesirable formal and linguistic properties, or it may be the case that there are alternative analyses that can be utilized.

- My co-authors and I have explored recasting categorial grammar as a hybrid logic, in which a derivation requires two separate logics. The resulting system has attractive computational properties, such as the elimination or simplification of the structural modalities that are common in much contemporary work in “multimodal” categorial grammar, with a resulting improved parsing problem for such grammars. There are many areas to further pursue in this work, such as a more rigorous specification of the two logics, and improving the parsing algorithm in light of problematic sentence constructions such as coordination.

- Along with some colleagues in the Computer Science department at the University of Pennsylvania, I have been co-supervising some projects by undergraduates on information extraction and clustering
of verbs into syntactic/semantic classes. This has led to an increased interest on my part in issues of lexical semantics and how clustering techniques can approximate or improve upon the verb classes outlined by linguists such as Beth Levin. I am currently investigating different clustering techniques for this purpose.

- I would like to continue research that I have recently begun into the computation of semantic representations in the context of systems like TAG. There is no level of Logical Form in TAG. Since the elementary units of the grammar are small pieces of phrase structure trees rather than individual lexical items, the semantic representation is computed during the course of the derivation based on the semantic contributions of each of the elementary trees. This is similar in concept to some of the early ideas in generative grammar of using T-markers for semantic interpretation, Bach’s work on combining Montague Grammar and Transformational Grammar, and more recent work by Higginbotham, but with some important differences, in my view, due to the character of the elementary trees. I have therefore started examining some of the classic arguments for LF in this context, such as antecedent-contained deletion. I suspect that in this context many of the arguments for a level of LF lose their force.

- I would like to expand my research to the area of psycholinguistics and language processing by exploring the issue of incremental parsing with a TAG grammar, and how results from psycholinguistics research can provide a needed perspective to this work. Previous work on this topic has used the original definition of the TAG formalism, which is inadequate for many constructions which can be studied from a psycholinguistic perspective, and it is an open question as to how to extend previous TAG-based work on psycholinguistics with a modified form of TAG.