Title: What Can Computers Do for Us? Grammar Engineering for Linguistic Hypothesis Testing, Linguistic Typology, and Language Documentation

Abstract:
Grammar engineering is the process of encoding formal grammars in
machine-readable form, so that the computer can do the tedious work of
verifying analyses against data. In this talk, I will give an overview
of two long-standing projects which aim to facilitate the use of grammar
engineering for linguistic research: The Grammar Matrix and AGGREGATION
Projects. The Grammar Matrix (Bender et al 2002, 2010) is an open-source
toolkit for helping create implemented precision grammars based on a
shared core grammar and a series of typologically informed 'libraries'.
The Grammar Matrix itself provides an interesting test-bed for
typological generalizations, as each new library must be interoperable
with existing ones. The Grammar Matrix solicits a linguistic description
through a web questionnaire and then outputs a grammar to spec. The
AGGREGATION project (Bender et al 2013, Howell et al 2017, Zamaraeva et
al 2019, Howell forthcoming) is exploring methods for automatically
answering the Grammar Matrix questionnaire on the basis of collections
of interlinear glossed text produced by linguists working in the field.
In the short-term this project provides useful feedback to the linguist
about patterns in their data, facilitating the language documentation
effort. In the long-term, our goal is to be able to create implemented
grammars which can be used to parse interestingly large fragments of the
languages at hand. These grammars should be useful for both linguistic
research and ultimately language technology.