**Final Project Proposal**

For my final project, I would like to explore the relatively young field of *computational semantics*, whose goal is to “find techniques for automatically constructing semantic representations for human language.” The standard approach in this discipline is to translate natural language into some formal language, for example, Higher-Order Logic (HOL) with a typed lambda calculus. Underlying this approach are the formal notions of *constituency* (modular phrase structure) and *compositionality* (meanings combine via function application from adjoined constituents). I am fairly familiar with these concepts from my undergraduate studies in formal linguistics, where we discussed the work of Richard Montague and others (with a heavy emphasis on quantification).

In this project, my aim will be to study how formal semantics can be applied to the problem of machine translation. In many ways, this could potentially be easier than other problems in formal semantics such as automated inference, which requires an accurate treatment of modals and belief contexts, which are very complicated topics. If our goal is just to translate, however, we can leave certain philosophical questions unresolved. In the semantics literature, much is written about translation from English into formal logic, but what is virtually absent is any discussion about translation from logic into English, or into other languages. I hope to address this topic by experimenting with some toy models of language, e.g. semantically-annotated CFGs, and to see if I can translate statements of logic into English and into the other language I am familiar with, Japanese. (I might try French too, if I can cajole a friend to help). I do not expect I will uncover techniques that will improve the status of machine translation over existing techniques, but I want to see if I can make a proof-of-concept for this “reverse translation” from logic to language. I will mostly make use of Python’s Natural Language Toolkit (NLTK) for my experiments, but I will also look into current research efforts involving formal semantics, such as Abstract Meaning Representation (AMR).

In my final report, I will give an overview of my experiments and any conclusions I draw from them. If the most I come up with is, “Translation using formal semantics is extremely difficult,” then my fallback plan will be to do a survey of the formal semantics literature that I read, but viewing it from the angle of machine translation. That is, I will try to address the question: What kinds of cases arise in natural language that cannot be translated effectively with syntactic/statistical methods alone?

References (and the list is growing):

Heim & Kratzer, *Semantics in Generative Grammar*, 1998.

Bos & Blackburn, *Computational Semantics*, 2005.

Beaver & Frazee, *Semantics*, 2013.

*Abstract Meaning Representation*. amr.isi.edu