Research Statement
Samuel R. Bowman

My research aims to discover the degree to which artificial neural network models can learn to produce rich semantic representations for sentences, and aims to use that knowledge to advance the state of the art in machine learning techniques for language understanding. My dissertation pursues this research program through experiments on both new and existing neural network models over narrowly targeted artificial language data, existing corpora of English, and newly collected corpora of English.

In order to study the ability of a neural network to develop good semantic representations, it is first necessary to choose a concrete task that instantiates that ability. I argue that the task of natural language inference—also called recognizing textual entailment—is ideal for this purpose. In this setting, a machine learning system is adequate if it can learn representations for sentences that allow it to judge whether any given sentence contradicts or logically follows from any other given sentence. In a typical instance, a model would be asked to label the pair

(1) a. a young man without protective gear is on a bike jumping over piles of sand
b. a cyclist with no helmet is navigating obstacles with his bicycle

as an entailment, rather than as a contradiction or a logically independent example. Since the task is framed as a simple decision problem over sentence pairs, it fits easily with standard machine learning methods. However, a model can only succeed on this task if it grapples with the full complexity of compositional semantics, including quantification, coreference, pragmatic inference, and many other phenomena. My experimental work on neural networks is centered on this task.

Neural network models have recently become the most effective tools for a range of hard applied natural language processing problems, including translation, sentiment analysis, and text generation. These models succeed in large part because they can learn and use their own continuous numeric representational systems for sentence meaning. However, their representations have no direct correspondence with the logic-based representations typically used in linguistic semantics. These models’ successes in learning to solve semantically difficult problems signal that they are a potentially valuable object of study for semantics, and drawing insights from semantics to improve these models could yield substantial progress across applied language understanding tasks. My research pursues these goals.

My work on inference in artificial languages with neural network models (Bowman et al., 2015b,c,d, and the second chapter of my dissertation) centers on three experiments. Together, they are designed to identify whether any fundamental obstacles prevent existing model architectures from developing usable representations. Each experiment isolates one prerequisite behavior for language understanding: building a lexicon, extending it into a compositional grammar, and understanding inferences based on quantification and lexical entailment. I find that neural network models can learn all three behaviors and that tree-structured neural networks, which are explicitly designed to instantiate the semantic principle of compositionality, are especially effective. In light of these successes, developing an understanding of when and why these models work requires a deeper investigation into how they learn from real English.
My next thread of research, exemplified by Bowman et al. (2015a, §2) and the third chapter of my dissertation, centers on public data and introduces a corpus that makes it possible to extend these artificial language results to English. While there are corpora of natural language inference data available, none of them contain more than a few thousand examples. Since neural network models make up for their lack of prior knowledge by demanding vast amounts of training data, they cannot learn well in this context. To make more ambitious learning experiments possible, I introduce the Stanford Natural Language Inference (SNLI) corpus: 570,000 sentence pairs constructed by hand and labeled for natural language inference through a novel crowdsourcing task based on visual scenes.

My work on both artificial and natural language data creation is closely tied to my work on developing modeling techniques for sentence understanding. In Bowman et al. (2015a, §3–4) and in the fourth chapter of my dissertation, I present a set of models for natural language inference trained on SNLI. I find that while some neural network models are competitive with the state of the art in conventional non-neural inference modeling, the compositional tree-structured models that perform best in the artificial data setting cannot scale to the size and difficulty of SNLI. I address this by proposing a novel model (in ongoing work) that adapts ideas from the parsing literature to incorporate syntactic information into the semantic composition process while circumventing the engineering bottlenecks seen in existing syntactically-savvy models.

Training a truly general purpose sentence encoding model will likely require several orders of magnitude more data than is presently available—far more than could realistically be annotated manually. At Google Brain this summer, I worked to develop new unsupervised learning methods for sentence understanding. This line of work will enable the models that I study to learn from the full volume of text available on the internet. This work has so far yielded a joint paper, now under review for ICLR 2016 (Bowman et al., 2015e).

I have additional research experience in formal phonology. My Master’s thesis at the University of Chicago concerned vowel harmony in the context of computational simulations of Optimality Theory, and two subsequent papers at Stanford developed my interest in constraint-based grammar treatments of vowel harmony, first with existing data from Hungarian and Seto (Finno-Ugrian, Estonia) and then with original field data from Kazakh.

In future work, I plan to develop the tools necessary to build neural network models for sentence encoding that capture a diverse array of linguistic phenomena. To better understand the obstacles to building such models, I will develop new corpora that supplement SNLI and allow for the evaluation of models on specific hard semantic problems, including those involving time, beliefs, and the resolution of lexical or syntactic ambiguities. In addition, I will pursue multitask learning techniques that will allow the models that I study to learn from a diverse range of semantic modalities. These techniques could, for example, combine the disparate aspects of meaning highlighted by translation corpora and by sentiment analysis corpora, and could also combine knowledge from human-annotated data like entailment pairs with knowledge inferred directly from naturally occurring text.

In a complementary line of work, I aim to better understand the syntax–semantics interface through the lens of machine learning. In particular, I will investigate how and when existing syntactic formalisms can support neural networks in learning usable semantic representations, and whether neural network models that are trained to induce their own compositional structures yield structures that are interpretable in semantic or syntactic terms.
References


