Nina Dethlefs

Title: Using linguistic prior knowledge in deep learning

Abstract:

Deep learning is the main technique behind recent breakthroughs in artificial intelligence, such as self-driving cars, computer vision or applications in natural language including machine translation and personal assistants such as Siri, Alexa or Cortana. Deep learning has been shown to beat more traditional methods on a number of common benchmark tasks for which large amounts of training data are available. However, neural networks struggle with domains for which no such data is available, or datasets are small, noisy, or depend on linguistic and social contexts. This talk will explore the idea of using linguistic prior knowledge in systems that are otherwise trained with deep learning. The particular example I will discuss is from natural language generation (NLG), i.e. computer systems that generate natural language from data or other non-linguistic input representations.

Data-driven NLG systems that are trained from labelled datasets are often domain-specific in their annotation and in their mapping from semantic input representations to lexical-syntactic outputs. As a result, trained systems often generalise poorly across domains. For example, a system trained to generate route instructions in urban environments might subsequently fail to generate scene descriptions, even if the linguistic constructions used in both domains are similar. In this talk, I will discuss the problem of domain adaptation for natural language generation. I will show how linguistic knowledge from a source domain, for which labelled data is available, can be adapted to a target domain by reusing training data across domains. As a key to this, I propose to employ abstrac meaning representations (AMRs) as a common semantic representation across domains. AMRs are linguistically motivated meaning representations that focus less on the syntactic properties of language and more on semantic nuances, capturing e.g. the differences between verbs of sensing, acting, etc. and can therefore provide a general-purpose linguistic framework to underpin machine learning.