Invited Speaker

Yejin Choi
University of Washington

Sketch-to-Text Generation: Toward Contextual, Creative, and Coherent Composition

Abstract

The need for natural language generation (NLG) arises in diverse, multimodal contexts: ranging from describing stories captured in a photograph, to instructing how to prepare a dish using a given set of ingredients, and to composing a sonnet for a given topic phrase. One common challenge among these types of NLG tasks is that the generation model often needs to work with relatively loose semantic correspondence between the input prompt and the desired output text. For example, an image caption that appeals to readers may require pragmatic interpretation of the scene beyond the literal content of the image. Similarly, composing a new recipe requires working out detailed how-to instructions that are not directly specified by the given set of ingredient names.

In this talk, I will discuss our recent approaches to generating contextual, creative, and coherent text given a relatively lean and noisy input prompt with respect to three NLG tasks: (1) creative image captioning, (2) recipe composition, and (3) sonnet composition. A recurring theme is that our models learn most of the end-to-end mappings between the input and the output directly from data without requiring manual annotations for intermediate meaning representations. I will conclude the talk by discussing the strengths and the limitations of these types of data-driven approaches and point to avenues for future research.