The multimodality of diagrams - insights from linguistics for computational processing?

Tuomo Hiippala
University of Helsinki

2/22/2019 at 10:30 am
CoRE A 301

Abstract

Within many fields, including linguistics and computer science, there is a growing interest in multimodality, or how communication relies on combinations of different forms of expression. In this presentation, I focus on the multimodality of diagrams, which combine natural language with illustrations and diagrammatic elements such as arrows and lines, and set up relations between these elements that need to be resolved during interpretation. Due to their multimodal characteristics and inherent variation, the computational processing of diagrams presents a formidable challenge (Haehn et al. 2019).

One dataset developed for research on computational tasks such as diagram understanding and visual question answering is the Allen Institute for Artificial Intelligence Diagrams (AI2D) dataset (Kembhavi et al. 2016). AI2D contains nearly 5000 diagrams from school textbooks with crowd-sourced annotations for diagram constituents and the relationships that hold between them. Each diagram is represented using graph with diagram constituents as nodes and relations as edges.

To better account for the complex multimodal structure of diagrams, I am currently working on a drop-in replacement for the AI2D annotation, which builds on linguistically-informed approaches to the multimodality of diagrams (Alikhani & Stone 2018; Hiippala & Orekhova 2018). Instead of representing the diagrams using a single graph, AI2D-RST adopts a stand-off approach by splitting the description into three separate graphs, which account for (1) hierarchical organisation of content, (2) connections set up by lines and arrows, and (3) discourse structure using Rhetorical Structure Theory (RST). I argue that such fine-grained descriptions are necessary for improve our understanding of diagrams, and can also help to advance their computational processing, which has been shown to benefit from information on diagram structure (Kim et al. 2018).

Bio

Tuomo Hiippala (PhD 2014) is Assistant Professor of English Language and Digital Humanities at the University of Helsinki, Finland.

References: Alikhani & Stone (2018) Arrows are the verbs of diagrams. In Proc. COLING 2018.

Haehn et al. (2019) Evaluating graphical perception with CNNs. IEEE TVCG 25(1).

Hiippala & Orekhova (2018) Enhancing the AI2D Diagrams dataset using Rhetorical Structure Theory. In Proc. LREC 2018.

Kembhavi et al. (2016) A diagram is worth a dozen images. In Proc. ECCV 2016.

Kim et al. (2018) Dynamic graph generation: generating relational knowledge from diagrams. In Proc. CVPR 2018.

Faculty Host: Matthew Stone