

Visualizing links between sentences in asynchronous conversations: a new feature for ConVis

CPSC 547 2015W1 project update and draft of Previous Work
Jordon Johnson (jordon@cs.ubc.ca)

1. Project Update

There have been two major obstacles to the project so far. The first is an illness that affected me for over a week. As a result, I am in the process of catching up in multiple different areas, which makes catching up in my project difficult.

The second significant obstacle is the combination of my lack of experience with the languages used in this project and the lack of commenting used in the existing ConVis code. While there is some commenting, it is insufficient for someone who is not already familiar with the work at a code level; and so understanding the existing code is more difficult than anticipated. Fortunately, the original author of the ConVis software is available from time to time so I can ask him questions.

I am continuing to read the code in order to understand where the changes would need to be made to implement the proposed visualization. I have implemented the first milestone (ensuring my datasets are in a format readable by ConVis) and am in the process of testing/troubleshooting that implementation. I have also reconsidered my milestones. The second milestone in the proposal was to implement a toggle that would allow a user to switch between the “commenter view” and the “links view”; while this will need to happen in order to integrate the proposed visualization into ConVis, it is not necessary in order to implement the visualization itself. I am therefore dropping that milestone and will proceed directly to drawing the curves encoding the links.

2. Previous Work

This work extends the functionality of ConVis [Hoque and Carenini 2014], a state-of-the-art tool for visualizing asynchronous conversations that encodes conversation topics, post sentiment, and user activity. ConVis does not yet support the visualization of links between individual sentences.

The OnForumS project [Kabadjov et al. 15] is an endeavour to combine summarization, argument mining and sentiment mining into a single shared task, of which a portion is creating algorithms to find links between sentences. Exploratory research into this task prompted the idea for the proposed visualization.

To the author's knowledge, there has not yet been any work done in visualizing rhetorical relationships between sentences in asynchronous conversations. While much of the existing work has as its scope individual documents (this work treats an asynchronous conversation as a single document), a significant amount deals with collections of documents, as made evident by recent surveys of visual text analytics [Alencar et al. 2012]. In addition, the focus in visual text analytics seems largely to be on either individual words (for example, term frequencies [Rohrer et al. 1998] or relationships between words [van Ham et al. 2009]) or topics [Liu et al. 12].

The proposed visualization idiom is based on the arc diagram, which dates back to the 1960s as a candidate to minimize edge crossings [Saaty 1964, Nicholson 1968]. In an arc diagram, the nodes are

placed on a line, and edges between them are represented by semicircles above and below the line. The temporal nature of asynchronous conversations lends itself to this idiom, though the proposed visualization is slightly different in some respects (as will be discussed in Section **XX**).