I think your paper could benefit from more big-picture contextualizing. Your project has a real cutting-edge-neat-o feel. I can imagine hearing about it on NPR’s Science Friday segment. The beginning of your introduction says something like, “Everyone knows Shakespeare is great, so we subjected one of his plays to transfer entropy analysis.” But why do that?

Later in the Discussion/Summary section you hint at the answer: maybe transfer entropy captures something about the semantic information readers pick up from the words in a play. Then you have the hypothesis that the TE measures are below random because the narrative device of “suspense” relies on the withholding of information. That’s a nice idea. One argument against a hypothesis like that might be that, yes, suspense does rely on the withholding of information, but not the withholding of information in general. Rather, suspense relies on the withholding of a few key pieces of information. A storyteller needs to give an audience lots of information for the audience to recognize something is at stake, but only needs to withhold a relatively small amount of information to make the audience feel suspense (e.g. *Pride and Prejudice*, will Elizabeth end up with Mr. Darcy? We need lots of info to even understand the question and then even more information to make us care. Jane Austen only has to withhold the one key bit of information (yes/no) to create suspense).

But then I also remember you saying that Shakespeare’s plays are driven by the ignorance of his characters. The plays are huge misunderstandings, basically. So does TE capture something about the lack of information characters have about each other’s future states? Or does it capture something about the lack of information readers have about the future states of characters? Or both? Or neither? I say don’t be afraid to get a little deeper into some of the interesting problems surrounding what you are measuring and how it might connect to MAAN as a meaningful piece of art. A project like yours cries out for such speculation.

Getting down to more technical stuff, there were two points in your article where I lost the thread of what was going on. I have left comments marking both in my annotated version of your article.

First, in the “Model Description” section, I wasn’t sure what role the edge weights played in your TE analysis. I also thought you were also a bit vague about how those edge weights were determined. I mean, you made your network Boolean by assigning each character to a node and giving those nodes a value of 1 for “present” and 0 for “absent”, right? I don’t see how edge weight would effect whether or not a character was present or absent, except that the existence of an edge in general would indicate that the source node character was present since a character must appear in a sub-scene to say anything. More detail in that section would help.

Second, I didn’t understand this sentence in your “Discussion” section: “Thus, the characters appear to contain more information about their future in their own pasts than if their pasts were randomized, and this has the effect of decreasing the influence of other characters relative to the random associations.” Maybe that’s because I don’t have a deep understanding of how TE works.

The edits I made to your article are, of course, just suggestions. I edit a lot of papers for my TA position, so now I just instinctively change things as I read.

Glad I was assigned your project for review,

Kelle