Discourse-Sentiment Alignment Tool (DSAT)

Patrick Huber
Nov 19, 2019

Intro

Discourse Parsing:

- Crucial task within the area of NLP
- Enhances many downstream applications
  - Sentiment analysis
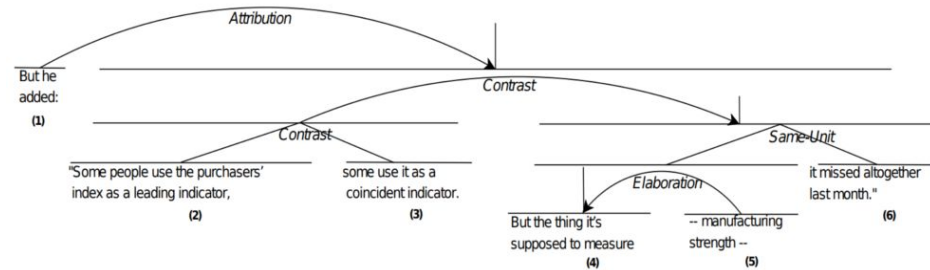 - Summarization
  - Question answering

Goal:

- Reveal the underlying structure of coherent text (a discourse)
  - Complete documents
  - Multiple sentences

Intro

Normally:

- Human-annotated gold-standard discourse trees

In my research:

- Use large datasets without gold-standard trees
- Automatically infer discourse trees
- Using distant supervision from sentiment data

Problem & Objective

Problem:

- Fully automated generation
- No human-in-the-loop
- Existing tools limited to comparisons against gold-standard

Objective:

- Create InfoVis system, which generates insights into the alignment of discourse trees and sentiment

Data

Yelp '13 review corpus

- ~336,000 reviews
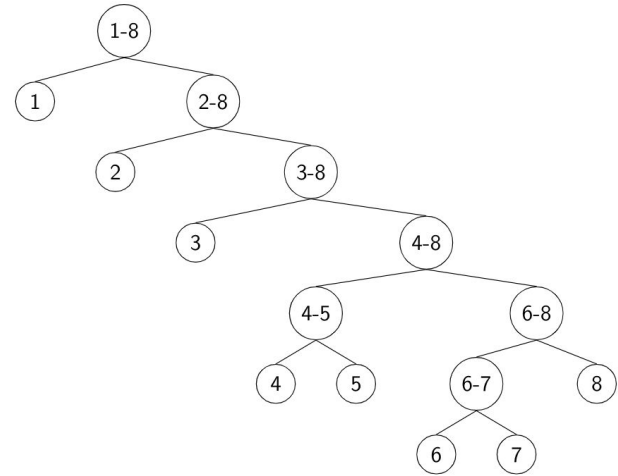- 1-5 star rating per review

Star Rating

+

Customer Review
#1223

Data

[What happened to Dunkin' Donuts?] (1)
[Holy crap does this place suck.] (2)
[The donuts are stale and taste weirdly like chemicals.] (3)
[I can not recommend anything] (4)
[except that you drive five minutes to Bosa Donuts on McDowell.] (5)
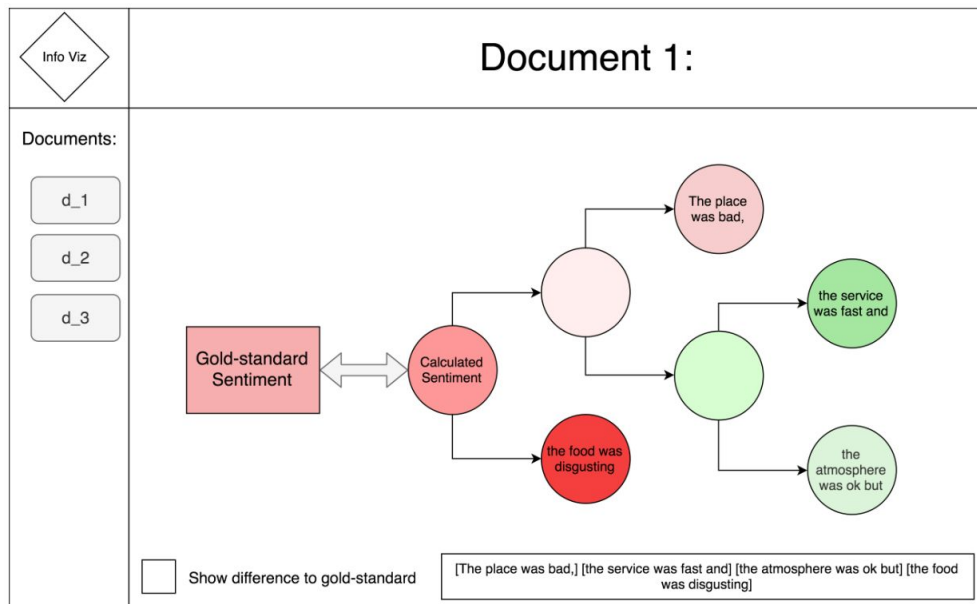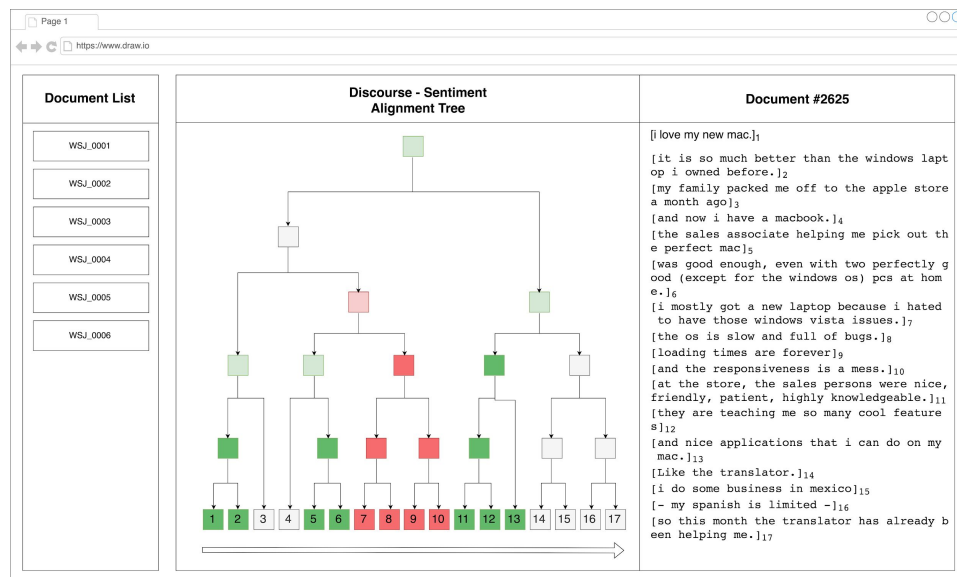[Great donuts] (6)
[and locally owned.] (7)
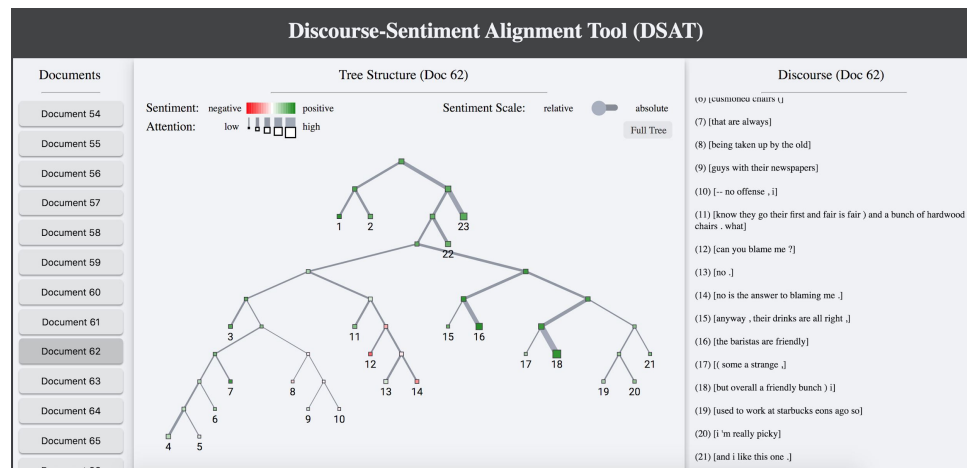[Support local.] (8)

The Visualization Stage 1

The Visualization Stage 2

The Visualization Stage 3 (Now)

Demo