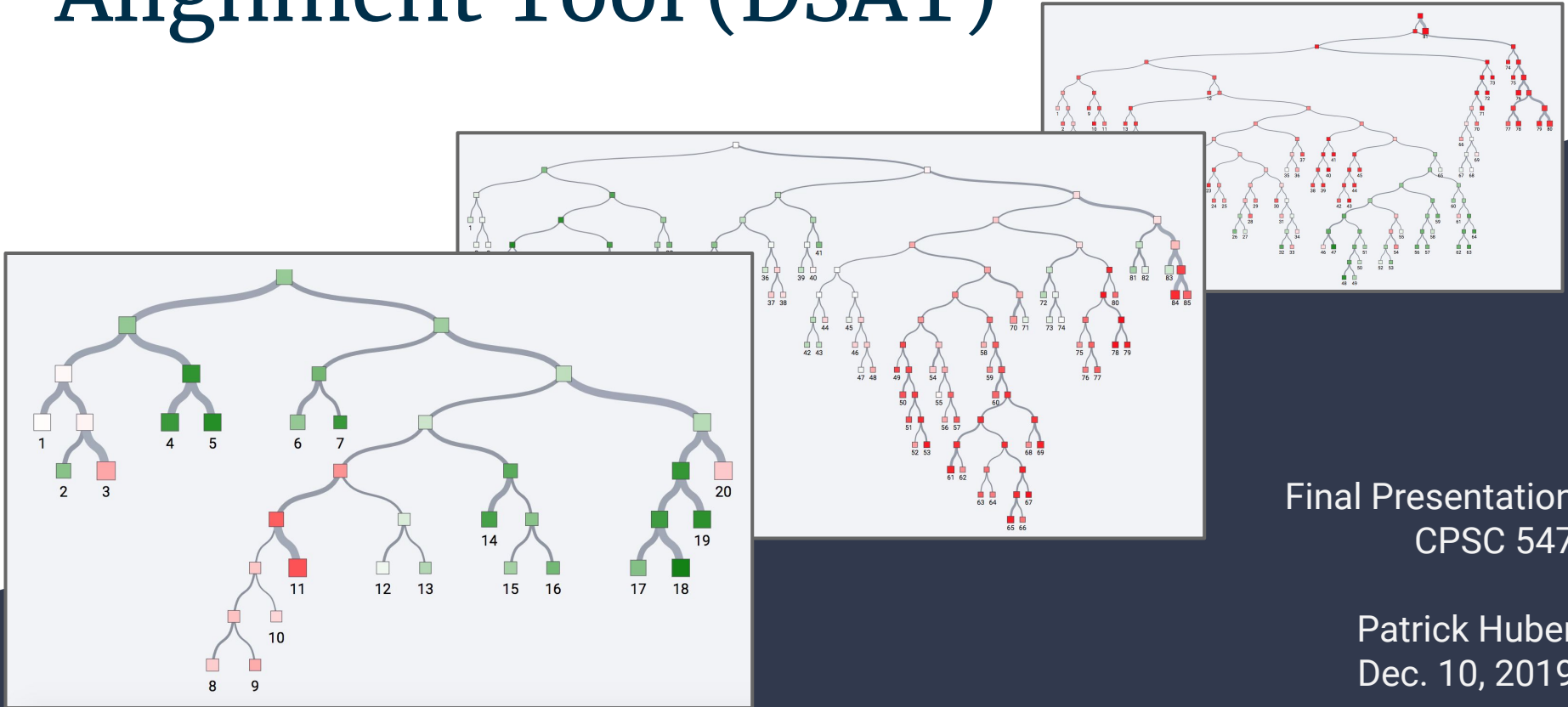


Discourse - Sentiment Alignment Tool (DSAT)



Final Presentation
CPSC 547

Patrick Huber
Dec. 10, 2019

Introduction

Discourse Parsing

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

Introduction

Discourse Parsing

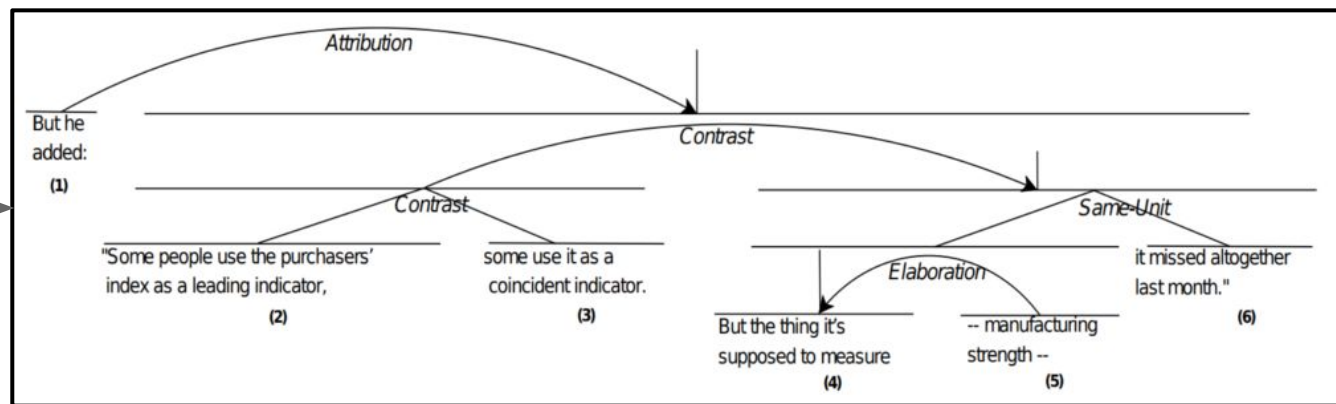
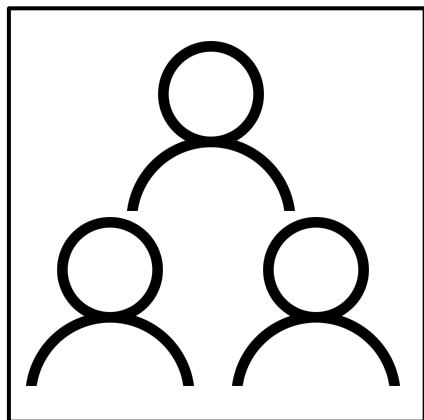
In the Past:

- Human-annotated gold-standard discourse trees

Recently:

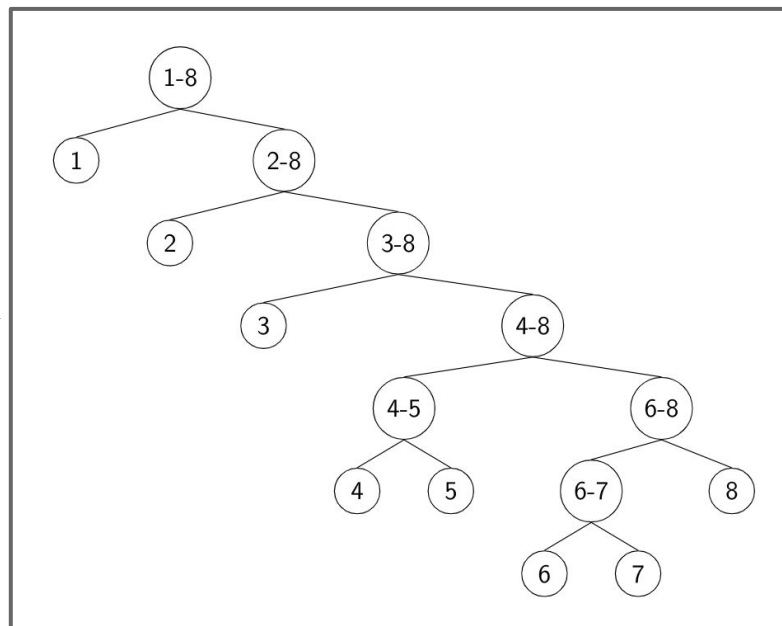
- Use large datasets without gold-standard trees
- Automatically infer discourse trees
- Using distant supervision from sentiment data

A human-annotated Discourse Tree



An automatically-generated Discourse Tree

[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)



Problem & Objective

Problem:

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

Objective:

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

Data

```
( Root (span 1 22)
  ( Nucleus (span 1 10) (rel2par Topic-Shift)
    ( Nucleus (span 1 3) (rel2par span)
      ( Satellite (leaf 1) (rel2par attribution) (text _!Moody's Investors Service said_!) )
      ( Nucleus (span 2 3) (rel2par span)
        ( Nucleus (leaf 2) (rel2par span) (text _!it reduced its rating on $165 million of subordinated debt of this Beverly Hills, Calif., thrift,_!) )
        ( Satellite (leaf 3) (rel2par result) (text _!citing turmoil in the market for low-grade, high-yield securities._!) )
      )
    )
  ( Satellite (span 4 10) (rel2par summary-n)
    ( Nucleus (span 4 9) (rel2par span)
      ( Nucleus (span 4 7) (rel2par span)
        ( Nucleus (span 4 6) (rel2par span)
          ( Satellite (leaf 4) (rel2par attribution) (text _!The agency said_!) )
          ( Nucleus (span 5 6) (rel2par span)
            ( Nucleus (leaf 5) (rel2par List) (text _!it reduced its rating on the thrift's subordinated debt to B-2 from Ba-2_!) )
            ( Nucleus (leaf 6) (rel2par List) (text _!and will keep the debt under review for possible further downgrade._!) )
          )
        )
      )
    )
  )
)
```

"sentiment": 0.2515886536341643,

"sentiment": 0.24531555738486907,

"sentiment": 0.24325398682811736,

"sentiment": 0.2344575598835945,

"attention": 0.9277132898569107,

"attention": 0.8705874979496002,

"attention": 0.8900138735771179,

"attention": 0.8881056308746338,

Demo

Design Decisions & Idioms

Design Decisions:

- 3-column design
 - Left: Document selection
 - Center: Vertical tree layout
 - Right: Textual representation
- Restricted navigation/zooming
- Dynamic scrolling
- Relative/Absolute sentiment scaling
- Hierarchical dynamic highlighting

Idioms Used:

- Visual Encoding:
 - Node-link diagram
 - Topological structure important
 - Less than 150 nodes
 - Node ordering → Text
- Interaction Idiom:
 - Bidirectional linking
 - Text → Leaf-node
 - Subtree → Connected text spans

Analysis – What / How / Why

What?

Data:

Textual representation of trees annotated with sentiment and importance (attention)

Derived:

- Spatial tree representation
- Sentiment annotation as diverging, sequential color-scale
- Importance as node & link size

Shown:

- Tree + Text of a single document

How?

Executed:

Flat data converted in hierarchical tree structure

Shown:

- Linked discourse tree / text
- Sentiment & importance on every node
- Textual index at leaf-nodes
- Restricted navigation on complete subtrees

Analysis – What / How / Why

Why?

Important:

- Correct alignment of tree / sentiment
not easy to find without spacial
encoding of tree
- Repetition / Biases / Misalignment can
be graphically explored
- Dataset-level evaluations not sufficient
(but very common)

Conclusion & Future Work

Conclusion:

- Visualization helped confirm hypothesis
- Especially useful for mixed reviews
- Positive feedback from user in the domain

Future Work:

- Augment document selection
- Restricted local Pan & Zoom
- Overlay multiple trees
- Collapsible layout

Thank You

Questions?