

Discourse and Rhetorical Relations

March 7, 2025

- Definition of Discourse
- Penn Discourse Treebank
- PDTB and RST
- Discogem
- AART
- Preliminary results

Definition of Discourse

- **Discourse:** Beyond a list of sentences, texts exhibit complex global structures.
 - Sections, subsections, and hierarchical organization collectively shape overall meaning.
 - Meaning emerges from how parts of a text relate, not just from standalone propositions.
- **Discourse Parsing:** Assigns an analysis of these structures within a theoretical framework, identifying:
 - Types of semantic/pragmatic meanings (e.g. CAUSE, CONCESSION)
 - How sections of a document combine hierarchically

- **Knott (1996)**: Base relations on **cue phrases** (conjunctions, adverbials, etc.).
- **Webber et al. (2003)**: Distinguish between coordinating/subordinating connectives vs. discourse adverbials (treated as anaphors).
- **Penn Discourse Treebank**:
 - Largest corpus for connectives and discourse relations.
 - Emphasizes explicit vs. implicit connectives.

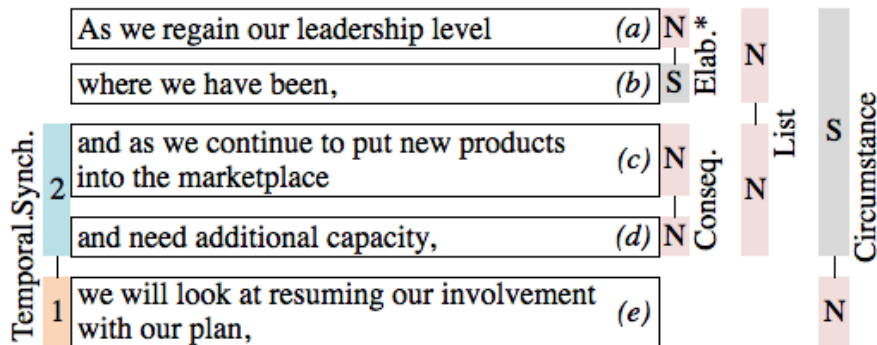
- **Common Ground:** Both view texts as connected by coherent relations beyond sentence boundaries.
- **PDTB (Penn Discourse TreeBank):**
 - *Shallow* approach labeling discourse relations with explicit/implicit connectives.
 - Minimal spans as arguments (clauses/sentences).
 - Relations often revolve around connectives (e.g. *because*, *although*).
- **RST (Rhetorical Structure Theory):**
 - Proposes a hierarchical tree structure (Mann & Thompson 1988).
 - Relations include causal, contrastive, elaborative, etc.
 - **Nuclearity:** Distinguishes main (nucleus) vs. supporting (satellite) spans.

PDTB and RST

PDTB

Text

RST



Discogem + Discogem 2.0 (Demberg, Scholman)

It is likely that the strike could last even longer even though the company and union resumed negotiations since earlier this week. The union will meet with the management again until a conclusion can be drawn.

Drag and drop the word / phrase that links the highlighted texts:

also

or

in other words

for example

this illustrates that

thereby

as if

rather than

instead

other than that

except

(no direct relation)

because

as a result

in that case

if

if not

unless

to ensure that

so that

even though

nonetheless

on the other hand

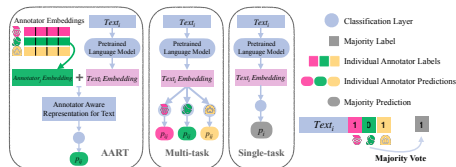
finally

after on

at the same time

- Only underspecified.
- Crowd-sourcing connectives, post-hoc translation to PDTB labels.
- **5618** items in **English**, 2588, 2628, and 2000 in German, French and Czech.
- Cohen's kappa 0.55: moderate agreement.

- "Capturing Perspectives of Crowdsourced Annotators in Subjective Learning Tasks"
- Learning of annotator representations.
- Only tested on binary tasks.
- Combined loss: cross-entropy + contrastive loss for ann. embeddings.



Preliminary results

- Level-3 relations (28-wise)
- 1) Prioritize embedding contrastive loss.
- 2) Prioritize cross-entropy.
- Comparison: 23% on level-2 relations (11-wise) in baseline systems.

Task	Method	Avg Acc	Avg F1
Perspectivist	AART-1 RoBERTa	25.4	8.64
Perspectivist	AART-2 RoBERTa	31	15
Majority label	RoBERTa	33.9	13.46