Domain-Dependent Content Models

- Capture topics and their distribution
- Are based on pattern matching techniques
  - Motifs of semantic units
  - Distributional model
- Useful in generation and summarization

Domain-Dependent Rhetorical Model

Domain: Scientific Articles

- Humans exhibit high agreement on the annotation scheme
- The scheme covers only a small fraction of discourse relations

Domain-Independent Rhetorical Model

- Model elements:
  - Binary Relations
  - Compositionality Principle
- Requirements:
  - Stability and Reproducibility of an Annotation Scheme
  - Expressive Power of a Model
Informational Structure

- How many different coherence relations are there?
- Are different taxonomies of coherence relations compatible with each other?
- Some real-time evidence for validity of some coherence relations: pronoun experiments (difference cause-effect/=resemblance)

Example of Coherence Relation (1)

Causal relations: Cause-Effect

\[
\text{cause} \quad \rightarrow \quad \text{effect}
\]

John is dishonest because he is a politician

Coherence Relations: Historic Perspective

Aristotle
(4th cent. BC)

Boccaccio
(14th cent.)

Hume
(18th cent.)

Example of Coherence Relation (2)

Causal relations: Violated-Expectations

John is honest although he is a politician.

\[
\sim
\]

John is dishonest
Example of Coherence Relation (3)

**Causal relations:** Condition

If someone is a politician he is dishonest.

---

Example of Coherence Relation (4)

**Resemblance relations:** Parallel

John organized rallies for Gore, and Fred distributed pamphlets for him.

---

Example of Coherence Relation (5)

**Resemblance relations:** Contrast

John supported Gore, and Fred cheered for Bush.

---

Example of Coherence Relation (6)

**Elaborations relations:**

John supported Gore, and Fred cheered for Bush.
How many coherence relations?

- Some accounts of coherence assume 2, other more than 400 coherence relations
- Hovy & Maier 1995: taxonomies with more relations represent subtypes of taxonomies with fewer relations
  - cause-effect → volitional, non-volitional

Find Coherence Relations

Consider this extract from “The Kreutzer Sonata” by L. Tolstoy

(A) It is amazing how complete is the delusion that beauty is goodness.
(B) A handsome woman talks nonsense, you listen and hear not nonsense but cleverness.
(C) She says and does horrid things, and you see only charm.
(D) And if a handsome woman does not say stupid or horrid things, you at once persuade yourself that she is wonderfully clever and moral.

Problem: Ambiguity

(Mann & Thompson: 1988, Matthessen & Thompson: 1988)

- Developed in the framework of natural language generation
- Aims to describe “building blocks” of text structure
  - Nucleus vs Satellites
  - Binary Relations between Discourse Units
- Compositionality principle defines how to build a tree from binary relations
No matter how much one wants to stay a non-smoker, the truth is that the pressure to smoke in junior high is greater than it will be any other time of one’s life. We know that 3,000 teens start smoking each day, although it is a fact that 90% of them once thought that smoking was something that they’ll never do.

**Binary Relations**

- (JUSTIFICATION, A, B)
- (JUSTIFICATION, D, B)
- (EVIDENCE, C, B)
- (CONCESSION, C, D)
- (RESTATEMENT, D, A)

**Relations**

<table>
<thead>
<tr>
<th>Relation</th>
<th>Nucleus</th>
<th>Satellite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>text whose understanding is being facilitated</td>
<td>text whose understanding is being facilitated</td>
</tr>
<tr>
<td>Elaboration</td>
<td>basic information</td>
<td>additional information</td>
</tr>
<tr>
<td>Preparation</td>
<td>text to be presented</td>
<td>text which prepares the reader to expect and interpret the text to be presented</td>
</tr>
</tbody>
</table>
Compositionality

Whenever two large text spans are connected through a rhetorical relation, that rhetorical relation holds between the most important parts of the constituent spans.

Marcu (1997): used constraint-satisfaction approach to build discourse trees given a set of binary relations.
Wolf (2004): tree structure is not an adequate representation of discourse structure.

Automatic Computation of RST Relations

(Marcu, 1997)

- Aggregate discourse relations to a few stable groups: (contrast, elaboration, condition, cause-explanation-evidence)
- Establish deterministic correspondence between cue phrases and discourse relations:
  - { But, However } → Contrast
  - { In addition, Moreover } → Elaboration

Accuracy

- Compared against manually constructed trees
- Tested against human-constructed trees
- Automatically constructed trees exhibit high similarity with human-constructed trees
- However, see (Marcu&Echihabi, 2002) CONTRAST vs ELABORATION: only 61 from 238 have a discourse marker (26%)
Surface cues for discourse relations:

I like vegetables, but I hate tomatoes.

Method

- Assume that certain markers unambiguously predict discourse relations
- Create Cartesian product of words located on two sides of a discourse marker
- For each pair of words, compute its likelihood to predict a discourse relation

\[
\arg\max_{r_k} P(r_k|(s_1, s_2)) = \arg\max_{r_k} P((s_1, s_2)|r_k) \times P(r_k)
\]

where \(s_i\) is a discourse clause, \(w_i\) is a word and \(r_k\) is a discourse relation

\[
P((s_1, s_2)|r_k) = \prod_{i,j \in s_1, s_2} P((w_i, w_j)|r_k)
\]

Evaluation

- Training data:
  - Raw 1 billion words corpus (41,147,805 sents)
  - BLIPP parsed corpus (1,796,386 sents)
- The system can compute accurately some relations (see handout)
- The size and the quality of the training data matters a lot