NLP
Discourse Analysis

Discourse Parsing
## Sample Rhetorical Relations

<table>
<thead>
<tr>
<th>Relation</th>
<th>Nucleus</th>
<th>Satellite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antithesis</td>
<td>ideas favored by the author</td>
<td>ideas disfavored by the author</td>
</tr>
<tr>
<td>Background</td>
<td>text whose understanding is being facilitated</td>
<td>text for facilitating understanding</td>
</tr>
<tr>
<td>Concession</td>
<td>situation affirmed by author</td>
<td>situation which is apparently inconsistent but also affirmed by author</td>
</tr>
<tr>
<td>Elaboration</td>
<td>basic information</td>
<td>additional information</td>
</tr>
<tr>
<td>Purpose</td>
<td>an intended situation</td>
<td>the intent behind the situation</td>
</tr>
<tr>
<td>Restatement</td>
<td>a situation</td>
<td>a reexpression of the situation</td>
</tr>
<tr>
<td>Summary</td>
<td>text</td>
<td>a short summary of that text</td>
</tr>
</tbody>
</table>
1) Title: Bouquets in a basket – with living flowers
2) There is a gardening revolution going on.
3) People are planting flower baskets with living plants,
4) mixing many types in one container for a full summer of floral beauty.
5) To create your own "Victorian" bouquet of flowers,
6) choose varying shapes, sizes and forms, besides a variety of complementary colors.
7) Plants that grow tall should be surrounded by smaller ones and filled with others that tumble over the side of a hanging basket.
8) Leaf textures and colors will also be important.
9) There is the silver–white foliage of dusty miller, the feathery threads of lotus vine floating down from above, the deep greens, or chartreuse, even the widely varied foliage colors of the coleus.

Christian Science Monitor, April, 1983

from Mann/Matthiessen/Thompson
1) Title: 
Bouquets in a 
basket - with 
living flowers 

Preparation 

2) There is a 
gardening 
revolution going 
on. 

Elaboration 

3) People are 
planting flower 
baskets with 
living plants. 

Elaboration 

4) Mixing many 
types in one 
container for a 
full summer of 
floral beauty. 

Elaboration 

5) To create 
your own 
"Victorian" 
bouquet of 
flowers, 

Purpose 

6) Choose 
varying shapes, 
sizes and forms, 
besides a 
variety of 
complementary 
colors. 

Elaboration 

7) Plants that 
grow tall should 
be surrounded 
by smaller ones 
and filled with 
others that 
tumble over the 
side of a 
hanging basket. 

Elaboration 

8) Leaf textures 
and colors will 
also be 
important. 

Elaboration 

9) There is the 
silver-white 
foliage of dusty 
miller, the 
feathery threads 
of lotus vine 
floating down 
from above, the 
deep greens, or 
chartreuse, even 
the widely varied 
foliage colors of 
the coleus.
Discourse Parsing

- Four RST relations: contrast, cause–explanation–evidence, condition, elaboration + non–relation
- Up to 4M automatically labeled examples per relation
- Naïve Bayes
- Word co–occurrence features

[Marcu and Echihabi 2002]
Cross-document structure (CST)

<table>
<thead>
<tr>
<th>Number</th>
<th>Relationship type</th>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identity</td>
<td>Any</td>
<td>The same text appears in more than one location</td>
</tr>
<tr>
<td>2</td>
<td>Equivalence (paraphrasing)</td>
<td>S, D</td>
<td>Two text spans have the same information content</td>
</tr>
<tr>
<td>3</td>
<td>Translation</td>
<td>P, S</td>
<td>Same information content in different languages</td>
</tr>
<tr>
<td>4</td>
<td>Subsumption</td>
<td>S, D</td>
<td>One sentence contains more information than another</td>
</tr>
<tr>
<td>5</td>
<td>Contradiction</td>
<td>S, D</td>
<td>Conflicting information</td>
</tr>
<tr>
<td>6</td>
<td>Historical background</td>
<td>S</td>
<td>Information that puts current information in context</td>
</tr>
<tr>
<td>7</td>
<td>Cross-reference</td>
<td>P</td>
<td>The same entity is mentioned</td>
</tr>
<tr>
<td>8</td>
<td>Citation</td>
<td>S, D</td>
<td>One sentence cites another document</td>
</tr>
<tr>
<td>9</td>
<td>Modality</td>
<td>S</td>
<td>Qualified version of a sentence</td>
</tr>
<tr>
<td>10</td>
<td>Attribution</td>
<td>S</td>
<td>One sentence repeats the information of another while adding an attribution</td>
</tr>
<tr>
<td>11</td>
<td>Summary</td>
<td>S, D</td>
<td>Similar to Summary in RST: one sentence summarizes another</td>
</tr>
</tbody>
</table>

S=Sentence, P=Paragraph, D=document
## Cross-document structure (CST)

<table>
<thead>
<tr>
<th>Number</th>
<th>Relationship type</th>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Follow-up</td>
<td>S</td>
<td>Additional information which reflects facts that have happened since the previous account</td>
</tr>
<tr>
<td>13</td>
<td>Elaboration</td>
<td>S</td>
<td>Additional information that wasn’t included in the last account</td>
</tr>
<tr>
<td>14</td>
<td>Indirect speech</td>
<td>S</td>
<td>Shift from direct to indirect speech or vice-versa</td>
</tr>
<tr>
<td>15</td>
<td>Refinement</td>
<td>S</td>
<td>Additional information that is</td>
</tr>
<tr>
<td>16</td>
<td>Agreement</td>
<td>S</td>
<td>One source expresses agreement with another</td>
</tr>
<tr>
<td>17</td>
<td>Judgement</td>
<td>S</td>
<td>A qualified account of a fact</td>
</tr>
<tr>
<td>18</td>
<td>Fulfilment</td>
<td>S</td>
<td>A prediction turned true</td>
</tr>
<tr>
<td>19</td>
<td>Description</td>
<td>S</td>
<td>Insertion of a description</td>
</tr>
<tr>
<td>20</td>
<td>Reader profile</td>
<td>S</td>
<td>Style and background-specific change</td>
</tr>
<tr>
<td>21</td>
<td>Contrast</td>
<td>S</td>
<td>Contrasting two accounts of facts</td>
</tr>
<tr>
<td>22</td>
<td>Parallel</td>
<td>S</td>
<td>Comparing two accounts of facts</td>
</tr>
<tr>
<td>23</td>
<td>Generalization</td>
<td>S</td>
<td>Generalization</td>
</tr>
<tr>
<td>24</td>
<td>Change of perspective</td>
<td>S,D</td>
<td>The same source presents a fact in a different light</td>
</tr>
</tbody>
</table>
Argumentative Zoning

- **Aim**
  - research goal of the paper
- **Textual**
  - statements about section structure
- **Own**
  - description of the authors’ work (methodology, results, discussion)
- **Background**
  - generally accepted scientific background
- **Contrast**
  - comparison with other work
- **Basis**
  - statements of agreement with other work
- **Other**
  - description of other researchers’ work

[Teufel and Moens 2002]
NLP