Applicative Structures and Immediate Discourse in the Turkish Discourse Bank

İşin Demirşahin, Adnan Öztürel, Cem Bozşahin and Deniz Zeyrek

Cognitive Science Department, The Informatics Institute, Middle East Technical University (METU/ODTÜ), Ankara

ii.metu.edu.tr

LAW VII & ID, Sofia, ACL 51

Aug 9, 2013
Nature of Discourse Representation

- Entity chains (Knott et al. 2001)
- A single tree of discourse relations (Rhetorical Structure Theory, RST, Mann and Thompson, 1988)
- Successive trees of varying sizes connected and occasionally intertwined at the peripheries (Hobbs, 1985)
- Adjoined trees (Discourse-Lexicalized Tree Adjoining Grammar, D-LTAG, Webber, 2004)
- Directed acyclic graphs (Lee et al., 2006, 2008)
- Chain graphs (Wolf and Gibson, 2005)
- S/DRT: inferential
Sentence-level structures require more than context-free power, but not to the extent of dealing with general graphs, or with strings that grow out of constant control.

Shieber (1985); Joshi (1985)

Do we need something above context-freeness for discourse?

Can we decide by looking at ways of combination? application, composition and their associated semantics
How far do discourse structures deviate from applicative semantics?

**Applicative structures**: binary operations on data
e.g. a connective’s meaning depending only on two arguments.

A system is applicative if it makes use of function application only, not graph reduction or general function composition.

Function application can be linked to applicative structures by currying.

Do we need that in discourse annotation?

If more complex structures are found, we must go above applicative semantics, to do function compositions and graph reductions, which are known to require more computational power.
The level of discourse which can be studied “just above” syntax. Joshi (2011)

Before (or without) reaching for higher-level mechanisms such as goals, motives.

**Argument-taking, semantics, heads, nature of embedding**
● Look for signs of structures in corpus over and above trees

● Tree-violating configurations  (Lee et al., 2006, Aktaş et al., 2010)

● in terms of:

  surface configuration

  semantics (application v. composition, connective v. anaphor)
Turkish Discourse Bank (TDB): 400,000+ words, various genres (Zeyrek and Webber, 2008; Zeyrek et al., 2010)

Annotated in the style of PDTB (Prasad et al., 2008).

- CONN: the connective that sets up the discourse relation (but, and, instead)
- Arg2: the clause to which the CONN syntactically belongs.
- Arg1: the other argument.

Non-structural relations are annotated as such (discourse adverbials, phrasal expressions)

They take one argument; the other is inferential.

Abstract Objects (Asher, 1993)

“And then suddenly there is a sound. Foot steps. Someone has an errand to run, they walk hurriedly or run. THEN the girl stiffens suddenly. The boy stiffens, too; and every run has a hidden wish.”
file name–browser index: 00001131-2&3

Vazgeçmek kolaydı, ertelemek de. AMA tırmanmaya başlandı mı bitirilmeli! Çünkü her seferinde acımasız bir geriye dönüş vardı.

“It was easy to give up, or to postpone. BUT once you start climbing you have to go all the way! Because there was a cruel comeback everytime.”
<table>
<thead>
<tr>
<th>Configuration</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Embedding</td>
<td>695</td>
<td>27.28</td>
</tr>
<tr>
<td>Nested Relations</td>
<td>138</td>
<td>5.42</td>
</tr>
<tr>
<td><strong>Total Non-violating Configurations</strong></td>
<td>833</td>
<td>32.69</td>
</tr>
<tr>
<td>Shared Argument</td>
<td>489</td>
<td>19.19</td>
</tr>
<tr>
<td>Prop. Cont. Argument</td>
<td>194</td>
<td>7.61</td>
</tr>
<tr>
<td>Prop. Cont. Relation</td>
<td>1018</td>
<td>39.95</td>
</tr>
<tr>
<td>Pure Crossing</td>
<td>2</td>
<td>0.08</td>
</tr>
<tr>
<td>Partial Overlap</td>
<td>12</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Total Violating Configurations</strong></td>
<td>1715</td>
<td>67.31</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2548</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Independent relations

```
Rel1
  |------ Rel1
  |   Arg1
  |       Rel1
  |       Conn
  |       Arg2

Rel2
  |------ Rel2
  |   Arg2
  |       Rel2
  |       Conn
  |       Arg1
```
Full embedding

![Diagram showing relationships between Rel1 and Rel2]
Rel1’s Arg2 contains Rel2, and more material
Properly contained argument
Partial overlap
Crossing

Diagram: Rel1 Arg1, Rel2 Arg1, Rel1 Conn, Rel1 Arg2, Rel2 Conn, Rel2 Arg2
Double subordinator

Subordinator (*için* ‘because’) is repeated. Without repetition, full embedding.
file name–browser index: 20510000-31,32&34

Ceza, Telekom’un iki farklı internet alt yapısı pazarında tekel konumunu kötüye kullandığı İÇİN ve uydu istasyonu işletmeciliği pazarında artık tekel hakkı kalmadığı halde rakiplerinin faaliyetlerini zorlaştırdığı İÇİN verildi.

“The penalty was given BECAUSE Telekom abused its monopoly status in the two different internet infrastructure markets and BECAUSE it caused difficulties with its rivals’ activities although it did not have a monopoly status in the satellite management market anymore.”
Single subordinator (hypothetical)

Apparent tree-violation if we take one *because* and one *and* relation:

This was not available to annotators. ML can help here.
Wrap
Double wrap
## Challenges to surface applicativity

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Structural</th>
<th>Anaphoric</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared Argument</td>
<td>158</td>
<td>331</td>
<td>489</td>
</tr>
<tr>
<td></td>
<td>32.31%</td>
<td>67.69%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Prop. Cont. Arg.</td>
<td>65</td>
<td>129</td>
<td>194</td>
</tr>
<tr>
<td></td>
<td>33.51%</td>
<td>66.49%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Prop. Cont. Rel.</td>
<td>547</td>
<td>471</td>
<td>1018</td>
</tr>
<tr>
<td></td>
<td>53.73%</td>
<td>46.27%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Pure Crossing</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>50.00%</td>
<td>50.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Partial Overlap</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>75.00%</td>
<td>25.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Total</td>
<td>780</td>
<td>935</td>
<td>1715</td>
</tr>
<tr>
<td></td>
<td>45.48%</td>
<td>54.52%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
Earlier example has two connectives

file name–browser index: 00010111-54&55

Sonra ansızın sesler gelir. Ayak sesleri. [Birilerinin ya işi vardır, aceleyle yürürler, ya koşarlar.] O ZAMAN kız katılaşır ansızın. Oğlan da katılaşır VE [her koşunun gizli bir isteği var.]

“And then suddenly there is a sound. Foot steps. [Someone has an errand to run, they walk hurriedly or run.] THEN the girl stiffens suddenly. The boy stiffens, too; AND [every run has a hidden wish.]”
Crossing

- Rel1
  - Arg1
- Rel2
  - Arg1
- Rel1
  - Conn
- Rel1
  - Arg2
- Rel2
  - Conn
- Rel2
  - Arg2
There is no relation between the two discourse relations.
No need to compose over trees.
Violations seem to have applicative semantics (function application)
• Turkish DB: public at http://medid.ii.metu.edu.tr

• Tree violations in TDB
  - annotation differences
  - some can be reanalyzed as wrap, which has applicative semantics
  - anaphoric relations
  - connect at same spot at one end

• They seem to be apparent violations

• We have yet to find genuine cases of tree composition over connectives’ meaning
References


