Pair Annotation: Adaption of Pair Programming to Corpus Annotation

Işın Demirşahin, İhsan Yalçinkaya and Deniz Zeyrek
Middle East Technical University, Ankara, Turkey
disin@metu.edu.tr, yalcinkaya.ihsan@gmail.com, dezeyrek@metu.edu.tr
Turkish Discourse Bank (TDB)

- 500000-word-subcorpus of Metu Turkish Corpus (Say et al., 2002)
- Annotated for discourse relations in the style of Penn Discourse Tree Bank (Prasad et al., 2008)
  - Discourse connectives
  - Arguments -> Abstract objects (Asher, 1993)
Turkish Discourse Bank (TDB)

- Discourse Annotation Tool for Turkish (Aktaş, et al., 2010)
- Version 1.0:
  - 147 connectives, 8482 relations
  - http://medid.ii.metu.edu.tr
Annotation Procedure in the TDB

- Independent Annotation

1. Search for tokens
2. Eliminate non-discourse connectives
3. Annotate spans
   (Conn, Arg1, Arg2, Mod, Shared, Supps)
Initial Annotation Procedure in the TDB

- Agreement procedure

[Diagram showing the steps:
  1. Locating disagreements
  2. Group discussion
  3. Adjudication]
Initial Annotation Procedure in the TDB

Annotation guidelines

Independent annotations

Agreement procedure

Modification of guidelines

Proof procedure

Gold Standard
Common Divergences among Independent Annotations

- Physical errors
  - Word boundaries
  - Spaces and punctuation
  - Pairs of punctuation marks

“Gidelim” dedim, ama dışarıdan yeni sesler geldi.

“Let’s go” I said, but new voices came from outside.

Başka kimse olmadığınıdan iki kadının da yüzü açıktdı.

Since there was no one else, the faces of both women were unveiled.
Common Divergences among Independent Annotations

- Ambiguous cases
  - Same relation, with some differences
  - Different relations

<table>
<thead>
<tr>
<th>Salt gülmek için gelmişlerdi.</th>
<th>[Salt] gülmek için gelmişlerdi.</th>
</tr>
</thead>
<tbody>
<tr>
<td>They came to just laugh.</td>
<td>They came [only] to laugh.</td>
</tr>
</tbody>
</table>


The first ceremony for Usumi was arranged in front of the Association of the Journalists of Turkey (TGC). Orhan Erinç, the chairman the TGC, emphasized that Usumi would be missed. Then, the Usumi’s funeral was moved to the Sultan Ahmed Mosque.
Common Divergences among Independent Annotations

- Missing/invalid annotations
  - Simple overlook
  - Disagreement about the discourse function

Yarın demedim, ancak hafta başına hazır olur.
I didn’t say it would be tomorrow, it won’t be ready until the first day of the week.
I didn’t say it would be tomorrow, but it will be ready on the first day of the week.

Kızın saçları siyah ve kıvırcıktı.
The girl’s hair was black and curly.
Pair Programming

- Two programmers work as one (Williams, et al, 2000)
  - Driver
  - Navigator
- Joint ownership
  - Credit
  - Responsibility
- The roles are switched periodically
## Pair Programming

<table>
<thead>
<tr>
<th>PROS</th>
<th>CONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>More efficient use of coding hours</td>
<td>Increased programming hours (15% - 60%)</td>
</tr>
<tr>
<td>Increased performance (Pair Jelling)</td>
<td></td>
</tr>
<tr>
<td>Decreased communication time between units</td>
<td></td>
</tr>
<tr>
<td>Faster production</td>
<td>Increased man-hours</td>
</tr>
</tbody>
</table>
Pair Annotation and Pair Programming

• Similarities
  • The driver and the navigator
  • Switching roles

• Difference: Mixed approach
  • One independent annotator
  • One pair of annotators
Observed Benefits of Pair Annotation

- Higher inter-annotator agreement
  - Less physical errors
- Faster annotation of ambiguous cases
  - Increased discussion during annotation -> decreased discussion during agreement
  - Utilization of notes field
- Increased motivation
Agreement Statistics

- Independent annotation vs. Pair annotation

<table>
<thead>
<tr>
<th>Connective</th>
<th>IA – Arg1</th>
<th>PA – Arg1</th>
<th>IA – Arg2</th>
<th>PA – Arg2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ama</em> ‘but’</td>
<td>0.832</td>
<td>0.956</td>
<td>0.901</td>
<td>0.969</td>
</tr>
<tr>
<td><em>sonra</em> ‘after, later’</td>
<td>0.820</td>
<td>0.889</td>
<td>0.902</td>
<td>0.953</td>
</tr>
<tr>
<td><em>ve</em> ‘and’</td>
<td>0.692</td>
<td>0.945</td>
<td>0.791</td>
<td>0.964</td>
</tr>
<tr>
<td><em>ya da</em> ‘or’</td>
<td>0.843</td>
<td>0.939</td>
<td>0.974</td>
<td>0.973</td>
</tr>
</tbody>
</table>

- Pair-wise averaged K for 3 individual annotators in IA – individual against individual VS K for pair vs. individual in PA
### Agreement Statistics

- **Individual annotator vs Pair annotator**

<table>
<thead>
<tr>
<th>Connective</th>
<th>Ind – Arg1</th>
<th>Pair – Arg1</th>
<th>Ind – Arg2</th>
<th>Pair – Arg2</th>
</tr>
</thead>
<tbody>
<tr>
<td>aslında ‘actually’</td>
<td>0.766</td>
<td>0.937</td>
<td>0.889</td>
<td>0.984</td>
</tr>
<tr>
<td>halde ‘despite’</td>
<td>0.834</td>
<td>0.973</td>
<td>0.898</td>
<td>1.000</td>
</tr>
<tr>
<td>nedeniyle ‘because of’</td>
<td>0.905</td>
<td>0.937</td>
<td>0.984</td>
<td>0.984</td>
</tr>
<tr>
<td>nedenle ‘for this reason’</td>
<td>0.952</td>
<td>1.000</td>
<td>0.987</td>
<td>1.000</td>
</tr>
<tr>
<td>ötürü ‘due to’</td>
<td>1.000</td>
<td>1.000</td>
<td>0.907</td>
<td>0.953</td>
</tr>
<tr>
<td>yüzden ‘because of this’</td>
<td>0.916</td>
<td>0.992</td>
<td>0.983</td>
<td>1.000</td>
</tr>
</tbody>
</table>

- **Individual annotator vs. gold standard K VS Pair annotator vs. gold standard K**
Issues in Pair Annotation

- Less sets of independent annotations
- Concern about navigator’s contributions
  - Switching roles
  - Periodical feedback
- Biased annotations
  - Keeping record of initial intuitions and disagreement within pair
- Biased agreement process
  - Treating the pair as a single entity
- All solutions are project specific
## Pair Annotation

<table>
<thead>
<tr>
<th><strong>PROS</strong></th>
<th><strong>CONS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant increase of inter-annotator agreements</td>
<td>Possibility that the pair’s annotation could be biased</td>
</tr>
<tr>
<td>Significant increase of annotator vs gold standard agreements</td>
<td>Possibility that the pair could dominate the agreement process</td>
</tr>
<tr>
<td>Faster agreement process</td>
<td>One less set of annotations</td>
</tr>
</tbody>
</table>
Future Work

- Quantitative evaluation of the increase in annotation hours
- Inter-annotator agreement with different pairs
- Inter-annotator agreement between two pairs of annotators.
Acknowledgements

- We gratefully acknowledge the support of the Turkish Scientific and Technological Research Council of Turkey (TUBITAK) and METU Scientific Research Fund (no. BAP-07-04-2011-005) while developing the Turkish Discourse Bank.
- We thank the Asian Federation of Natural Language Processing (AFNLP) and the National Science Foundation (NSF) for supporting İşın Demirşahin with student grants for travel, registration and accommodation.
References

Additional Reference

Initial Annotation Procedure in the TDB

- One search token -> Multiple connectives
- \textit{Halde}
  - \textit{-Diğl halde} ‘although, despite’
  - \textit{o halde} ‘then, in that case’
  - \textit{şu halde} ‘in the current situation, in this specific case’
  - \textit{aksi halde}, ‘otherwise’
  - \textit{bir halde} ‘in such a manner that’
Initial Annotation Procedure in the TDB

**Doğu Beyazıt'da gecelediğimiz** halde bir dünya şaheseri olan İshak Paşa medresesini göremeden Ankara'ya döndük.

Although *we spent the night in Doğu Beyazıt, we returned to Ankara without seeing the İshak Paşa Medresseh, which is a masterpiece.*

Feyzi Bey böyle bir durumda mebusluktan istifa edeceğini, aksi halde [de] **Falih Rıfkı Bey'in istifa etmesi gerektiğini** belirtmiş.

Mr. Fevzi stated that *in such a situation he would resign from parliament membership, otherwise Mr. Falih Rıfkı would have to resign.*

O gün akşama kadar ne yapacağımı bilmez bir halde dolaştım evin içinde.

I walked around the house till evening that day, not knowing what to do.