Contrasting the Interaction Structure of an Email and a Telephone Corpus: A Machine Learning Approach to Annotation of Dialogue Function Units

Jun Hu
Rebecca Passonneau
Owen Rambow
Columbia University – CCLS
rambow@ccls.columbia.edu
Interaction Structure

Origins in face-to-face spoken interaction

- Adjacency Pairs (Sacks, Schegloff, Jefferson 74)
- Speech Acts (Austen)

```
RequestInformation          Inform
```
```
RequestAction               Commit
```

- A: Will you marry me?
- B: <silence>
Email Example

• From: Melinda Solata
  To: Kevin Glick, Gio diStefano
  When did the package to DX go out? I am worried about them, they have hinted to me that they may switch providers. Who is in charge of that account?
  Melinda

• From: Kevin Glick
  To: Melinda Solata
  The package went out last week, I think.

• From Gio diStefano
  To: Melinda Solata, Kevin Glick
  I am not sure who is in charge of the account, perhaps Ram. The package to DX went out yesterday.
  GdS
Broader Notion of Interactive Communications: General Issues

Issue: underlying commonalities versus distributional differences

• Multiparty interactions
• No physical co-presence
• Asynchronous communication
• Non-linear discourse structure (tree, dag)
• Non-adjacent adjacency pairs
• Unanswered questions (or other openers)
• Answers to unasked questions (or other “volunteered” closings)
Outline

• Annotation Scheme: Dialog Function Units (DFU)
  – Dialog Acts
  – Links
  – Segments
• Corpora
• Automatic Prediction
Annotation Scheme: Dialog Acts (DAs)

- **Inventory: Reduced**
  - Goal: study interaction comparatively
  - Avoid massive data skew (and domain-specific labels)
  - Easier to generalize
  - Segmentation not given but determined by dialog act; therefore: too fine-grained DAs mean segments that are too small

- **Our set (n=8):** Request-Information, Inform, Request-Action, Commit, Conventional, Perform, Backchannel, Other
Annotation Scheme: Links

• Related segments are:
  – Not necessarily adjacent
  – Not necessarily predicted by dialog acts
  – Not necessarily complete
  – Don’t know where responding act will be
  – 1-to-many links
  – Discourse flow not necessarily linear: can be tree or dag

• Need explicit links between segments
  – Link: between request for information or action and its response
  – Secondary link: between any other DA pair
  – Dangling link
Annotation Scheme: Segmentation

- Dialog Function Units: defined **functionally**
- Longest segment which has a single discourse purpose:
  - Same DA tag
  - Same link structure
- Note: subsequent responses can alter segmentation!
- Also content requirement: generally same topic
Corpus 1: Loqui
Telephone Conversation

• Recorded phone conversations at New York City's Heiskel Library
• Dialogs: 175 collected, 82 transcribed, 48 annotated
• Annotated dialogues pertain to one or more book requests by customers
• Annotators worked from combination of transcription and audio.
Corpus 2: Enron Corporate Email

- 122 email threads of the Enron email corpus (with missing messages restored)
- Mostly information exchange, scheduling meetings, and solving problems, also purely social emails.
Corpora: Procedure

- Annotation developed looking at Enron, Switchboard, Dover Trial court transcripts
- Six people have been trained to annotate using our guidelines
- Most annotation used in experiments done by single annotator
- Guidelines have been under development and may be revised again (and data re-annotated)
- Have no inter-annotator agreement data for now
## Corpus: Size and Dialog Acts

<table>
<thead>
<tr>
<th>Dialog Act Labels</th>
<th>Tel.</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>21,097</td>
<td>17,924</td>
</tr>
<tr>
<td>Segments</td>
<td>3,845</td>
<td>1,400</td>
</tr>
<tr>
<td>Inform</td>
<td>50%</td>
<td>61%</td>
</tr>
<tr>
<td>Request-Information</td>
<td>20%</td>
<td>11%</td>
</tr>
<tr>
<td>Request-Action</td>
<td>1%</td>
<td>3%</td>
</tr>
<tr>
<td>Commit</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>Conventional</td>
<td>7%</td>
<td>25%</td>
</tr>
<tr>
<td>Backchannel</td>
<td>13%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Tel.</td>
<td>Enron</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Segments by Link</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part of a Paired Links</td>
<td>32%</td>
<td>14%</td>
</tr>
<tr>
<td>Links by Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Link</td>
<td>54%</td>
<td>28%</td>
</tr>
<tr>
<td>Secondary Link</td>
<td>39%</td>
<td>37%</td>
</tr>
<tr>
<td>Dangling</td>
<td>7%</td>
<td>33%</td>
</tr>
</tbody>
</table>
Outline

• Annotation Scheme: Dialog Function Units
• Corpora
• Automatic Prediction
Automatic Tagging

• 2 Tasks
  – Dialog Act Tagging: Choose DA for given segment
  – Link Choice: decide if two DFUs with DA tags are linked

• 3 Methods
  – Baseline
    • DA tagging: majority baseline
    • Link prediction: next plausible DFU (Serious baseline!)
  – Regular SVM: binary classifier (extended to n-ary)
    • Yamcha
  – Structured SVM: chooses among structure
Automatic Tagging:
Corpora

- 5-fold cross-validation
- Telephone: 3845 segments
- Email: 1400 segments
Structured SVM

- Learns discriminant function $F$: inputs $x \rightarrow \mathbb{R}$
- DA tagging:
  - Input = structure of segments
  - Output = sequence of DA tags
- Link prediction:
  - Input = structure of segments + (predicted) DA tags + link consideration space
  - At most 1 link starts or ends in any given segment
  - No crossing links
  - Output = chosen links
  - Constraints on link consideration space limit accuracy of prediction!
Conclusions:
• Task harder for Telephone than for Email
• Regular SVM does as well as Structured SVM; no surprise
### Results for Link Prediction
(Remall, Precision, F-Measure)

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Regular SVM</th>
<th>Structured SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>P</td>
<td>F</td>
</tr>
<tr>
<td>Tel.</td>
<td>30</td>
<td>56</td>
<td>39</td>
</tr>
<tr>
<td>Email</td>
<td>17</td>
<td>40</td>
<td>24</td>
</tr>
</tbody>
</table>

**Conclusions:**
- Task harder for Email than for Telephone
- Structured SVM paying off for Email corpus only!
Future Work

• Continue annotating Telephone and Email
• Start annotating web forum or similar
• Redo link prediction without restrictions (only one link per segment, no crossing links)
## Corpus: Links

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Tel.</th>
<th>Enron</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segments by Link</strong></td>
<td>Part of a Paired Links</td>
<td>32%</td>
<td>14%</td>
</tr>
<tr>
<td><strong>Links by Type</strong></td>
<td>Link</td>
<td>54%</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>Secondary Link</td>
<td>39%</td>
<td>37%</td>
</tr>
<tr>
<td></td>
<td>Dangling</td>
<td>7%</td>
<td>33%</td>
</tr>
<tr>
<td><strong>Links By Dialog Act</strong></td>
<td>Inform</td>
<td>50%</td>
<td>61%</td>
</tr>
<tr>
<td></td>
<td>Request-Information</td>
<td>20%</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>Request-Action</td>
<td>1%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>Commit</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>7%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Backchannel</td>
<td>13%</td>
<td>0%</td>
</tr>
</tbody>
</table>