Investigation of annotator’s behaviour using eye-tracking data

Ryu Iida, Koh Mitsuda, Takenobu Tokunaga
Tokyo Institute of Technology, Japan

LAW VII & ID (August 9, 2013)
Manual annotation: essential for ML-based approaches in various NLP tasks

- **Shallow processing tasks**: POS tagging, NP chunking
  - ML-based approaches have been largely successful
  - Surface information (e.g. word and POS) can be easily introduced as useful features

- **Deeper processing tasks**: coreference resolution, discourse parsing
  - Deeper linguistic knowledge has been integrated
    - WordNet, linguistic theories (e.g. Centering Theory)
  - There is still room for further improvement
Cognitive science approach based on annotator’s behaviour

- Look into human behaviour during annotation
  - Elicit useful information for NLP tasks requiring deeper linguistic knowledge
  - Focus on annotator eye gaze during annotation

- Developments in eye-tracking technology
  - Eye gaze data has been widely used
  - Psycholinguistics & problem solving (Duchowski, 2002)
  - Tomanek et al. (2010): utilised eye-tracking data to evaluate the degree of difficulty in annotating named entities
Aim

- Design experimental setting for collecting annotator’s behaviour (annotation events & eye gaze) during annotation

- Investigate annotator’s behaviour to elicit useful information in an NLP task
  - Annotating predicate-argument relations in Japanese
  - Moderately difficult annotation task due to the existence of zero-anaphora
    - Meaningful eye movement may be observed
Outline

1. Motivation of analysing annotation behaviour

2. Task setting of annotating predicate-argument relations in Japanese and data collection including annotation behaviour

3. Manual investigation using collected data
Annotation task: annotating obligatory arguments (subj, obj, iobj) of predicates in a text

- Segments of predicates and candidate arguments are pre-annotated automatically

Tom went to a park.

φ(he) met John there.
Annotation tool: modified version of Slate (Kaplan et al. 2012)
Recorded annotation events

- Record seven event types together with occurring time of each event and its related segments

<table>
<thead>
<tr>
<th>Event label</th>
<th>Description</th>
<th>predID</th>
<th>argID</th>
<th>linkID</th>
<th>link type</th>
</tr>
</thead>
<tbody>
<tr>
<td>create_link_start</td>
<td>creating a link starts</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>create_link_end</td>
<td>creating a link ends</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>select_link</td>
<td>a link is selected</td>
<td></td>
<td></td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>delete_link</td>
<td>a link is deleted</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>select_segment</td>
<td>a relation type is selected</td>
<td>✔</td>
<td>or</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>annotation_start</td>
<td>annotating a text starts</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>annotation_end</td>
<td>annotating a text ends</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>
Annotation environment

- **Equipment**
  - Eye-tracker: Tobii-T60
    - size: 1,280x1,024
  - Chin rest
  - Keyboard
    - select link type: \( ga \) (subj), \( o \) (obj), \( ni \) (iobj)
  - Mouse
    - create link between a predicate and its argument
Experimental settings

- Recruited three annotators
  - Experience in annotating predicate-argument relations
- Data: 43 articles in BCCWJ PB-corpus (Maekawa et al. 2010)
- Texts were truncated to about 1,000 characters to fit onto the screen to prevent scrolling
Annotation results done by three human annotators

<table>
<thead>
<tr>
<th>case</th>
<th>annotator A</th>
<th>annotator B</th>
<th>annotator C</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>3,353</td>
<td>3,764</td>
<td>3,462</td>
<td>10,579</td>
</tr>
<tr>
<td>selected</td>
<td>1,776</td>
<td>1,430</td>
<td>1,795</td>
<td>5,001</td>
</tr>
<tr>
<td>ga (subj)</td>
<td>1,170</td>
<td>904</td>
<td>1,105</td>
<td>3,179</td>
</tr>
<tr>
<td>o (obj)</td>
<td>383</td>
<td>298</td>
<td>421</td>
<td>1,102</td>
</tr>
<tr>
<td>ni (iobj)</td>
<td>223</td>
<td>228</td>
<td>269</td>
<td>720</td>
</tr>
</tbody>
</table>

- Our analysis requires an annotator’s fixation on segments of both a predicate and its argument. → available instances for analysis were reduced.
Outline

1. Motivation of analysing annotation behaviour

2. Task setting of annotating predicate-argument relations in Japanese and data collection including annotation behaviour

3. Manual investigation using collected data
Division of annotation process

- Divided into three stages (Russo & Leclerc (1994))
  - first fixation on target predicate
  - first fixation on linked argument

- reads a given text and understands its context
- searches for an argument of a target predicate
- looks around the context in order to confirm the pred-arg relation
Division of annotation process

- Divided into three sub-processes (Russo & Leclerc (1994))
  - first fixation on target predicate
  - first fixation on linked argument
  - evaluation

- Most informative for extracting useful features
- Analysing annotator eye gaze during this stage could reveal useful information for predicate-argument analysis
- Insufficient to regard only fixated arguments during this stage (annotator captures an overview of the current problem during the orientation stage)
Division of annotation process

- Divided into three sub-processes (Russo & Leclerc, 1994)
  - First fixation on target predicate
  - First fixation on linked argument
  - Verification

- Probable argument has been already determined and its validity confirmed by investigating its competitors
- Considered competitors are explicitly fixated during this stage
- Possible to analyse annotator’s behaviour during this stage based on eye gaze → concentrated on the analysis of the verification stage
Two viewpoints for investigation

1. Types of eye movement of annotator in verification stage

2. Distance of a target predicate and its argument in terms of character-based distance
1. Eye movement in verification stage

- **Concentrated**: after the first fixation of the argument annotated earlier, the fixations are concentrated onto it and the target predicate.

- **Distracted**: fixates on the competitors.
2. Distance of a predicate and its argument

- Hypothesis: annotator’s behaviour depends on the distance between predicate-argument
- Classified into the either Near and Far type
Investigation from three aspects

1. Predicate-argument distance and argument case

2. Effect of pre-annotated links

3. Specificity of arguments and dispersal of fixations
1. Distance of predicate-argument relations and their case

- Annotator changes her/his behaviour with regard to the case of the argument

<table>
<thead>
<tr>
<th></th>
<th>Near</th>
<th>Far</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ga (subj)</td>
<td>2,201 (0.44)</td>
<td>978 (0.90)</td>
<td>3,179 (0.64)</td>
</tr>
<tr>
<td>o (obj)</td>
<td>1,042 (0.34)</td>
<td>60 (0.05)</td>
<td>1,102 (0.22)</td>
</tr>
<tr>
<td>ni (iobj)</td>
<td>662 (0.22)</td>
<td>58 (0.05)</td>
<td>720 (0.14)</td>
</tr>
</tbody>
</table>

- 90% of Far class
  - *ga* arguments are often omitted to make ellipses
  - *o* and *ni* arguments less frequently appear as Far instances because they are rarely omitted

- Each case requires individual specific treatment in a model of predicate-argument analysis
1. Distance of predicate-argument and their case (Cont’d)

- Concentrated/Distracted distinction impacts on Near/Far distinction?

<table>
<thead>
<tr>
<th></th>
<th>Near-Concentrated</th>
<th>Near-Distracted</th>
<th>Far-Concentrated</th>
<th>Far-Distracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>ga (subj)</td>
<td>0.40</td>
<td>0.47</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>o (obj), ni (iobj)</td>
<td>0.60</td>
<td>0.53</td>
<td>0.08</td>
<td>0.10</td>
</tr>
</tbody>
</table>

- Concentrated/distracted distinction does not impact the distribution of the argument types
  - Even if an argument appears far from its predicate, the verification is completed without seeing any competitors
2. Effect of pre-annotated links

- In the situation of annotating $A$ for $P$, 6 links $SL$ have already been annotated.
  - These links make the argument visually or cognitively salient in annotator’s short-term memory.
Relationship between #already-existing links and #dwell on competitors

- Only Far instances
- Peaks around the intersection of instances with the fewest #links and dwells on competitors
- Lower #links
  - Mostly symmetrical relation
- Higher #links
  - Symmetry brakes
- Visual and cognitive salience reduces annotators’ cognitive load
  → efficiently confirming correct arguments
3. Relationship of specificity of arguments and dispersal of eye gaze

- Specific problem of our annotation setting
  - Only head of NP is pre-annotated as a segment in our annotation setting
  - e.g. *Benkyo-suru* to study *koto* -ing (to study / studying)

- Head noun of an argument does not always have enough information
  - Inspecting a whole NP including its modifiers is necessary to verify the validity of the NP for an argument
Empirical investigation about dispersal of eye gaze: head of NP

- Annotated arguments which have any NP modifiers are classified into ...
  - (a) fixations remain within the region of the argument NP
  - (b) fixations go out of the region

<table>
<thead>
<tr>
<th></th>
<th>(a) within NP</th>
<th>(b) out of NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentrated</td>
<td>1,190</td>
<td>-</td>
</tr>
<tr>
<td>Distracted</td>
<td>242</td>
<td>839</td>
</tr>
</tbody>
</table>

- 22% of Distracted arguments (242 instances) with any modifiers remain within NP region
- Need to treat candidate argument depending on if they have modifier or not
  - In addition to the head of NP, we should introduce information on modifiers into ML algorithms as features
Aim: analysis of annotator’s behaviour during her/his annotation for eliciting useful information for NLP tasks

Conducted an experiment for collecting three annotators’ eye gaze and annotation events during annotation of predicate-argument relations in Japanese texts

Analysed from three aspects:
- Relationship of predicate-argument distances and argument cases
- Effect of already-existing links
- Specificity of arguments and dispersal of eye gaze
Future work

- Further investigation of the collected data
  - Use of mining techniques for finding unknown but useful information may be advantageous
  - Employ mining techniques for finding useful gaze patterns for NLP tasks

- Current work: limited to the analysis of the verification stage of annotation
  → the orientation and evaluation stages include important clues for examining human behaviour during annotation