Adj-to-V Paraphrasing in Japanese
Based on Lexical Constraints of Verbs

Atsushi Fujita @ Nagoya Univ.
Naruaki Masuno @ IBM Eng. & Tech.
Satoshi Sato @ Nagoya Univ.
Takehito Utsuro @ Univ. of Tsukuba

Approaches to handle paraphrases
- Collecting synonymous expressions
  - from thesauri, parallel/non-parallel corpus, Web, etc.
  - Lexical paraphrases ("burst into tears" ⇔ "cry")
- Describing transfer rules
  - transformational grammar, synchronous grammar, etc.
  - Syntactic paraphrases (passivization, topicalization)
- Combining transfer rules and lexical resources
  - Lexical relation (Ls in MTT: e.g. v(attractive) = attract)
  - Lexical constrains (aspectual property, agentivitiy, etc.)
  - Compositional paraphrases (alternation, category-shifting)

Phenomena we focus on
- Adj-to-V paraphrasing in Japanese
  - A challenge to inter-categorial paraphrasing
  - Bridging gaps between different syntactic categories
  - Reasonable material for extending lexical semantics
  - Utilizing frameworks for verbs, at first

Distribution of verbal suffixes
- Closed example collection $C_{adj}$
  - 91 Adj tokens
  - 148 candidate Vs
  - Choice points
    - Voice: re/pe
    - Aspect: tel/pe
    - Tense: ru/ta
  - Majority
    - $ta$
    - re-$ru$
  - Distribution of verbal suffixes

<table>
<thead>
<tr>
<th>Verbal suffixes</th>
<th># of Adj-V pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>9</td>
</tr>
<tr>
<td>$ri/u$</td>
<td>5</td>
</tr>
<tr>
<td>$re$</td>
<td>14</td>
</tr>
<tr>
<td>$re$</td>
<td>2</td>
</tr>
<tr>
<td>$ri/ri$</td>
<td>57</td>
</tr>
<tr>
<td>$ri/ta$</td>
<td>2</td>
</tr>
<tr>
<td>$ri$</td>
<td>6</td>
</tr>
<tr>
<td>$ri/ri$</td>
<td>0</td>
</tr>
<tr>
<td>$ri/ta$</td>
<td>4</td>
</tr>
<tr>
<td>$ri/ta$</td>
<td>1</td>
</tr>
<tr>
<td>non-paraphrasable</td>
<td>48</td>
</tr>
</tbody>
</table>

Issue
- Add proper verbal suffixes to ensure equivalency
- What sorts of suffixes are appended?
- What determines them?
Task

Determining verbal suffixes
- Input
  - Pair of Adj (source) and V (target)
  - Head of noun phrase modified by the Adj/V
- Output
  - Verbal suffixes for the V
    - candidates are those in Table 1

Method: verbal features
- Verb generation ⇒ Utilizing lexical constraints
  - 7 linguistic tests derived from LCS (Table 2, p.43) [Kageyama, 1998] [Kato et al., 2005] [Takeuchi et al., 2006]
    - Transitive or intransitive (V)
    - Aspectual property (5 tests: \( V \_p \_t \_p \), \( V \_p \_t \_s \))
    - Agentivity (1 test: \( V \_t \))
- Examples of manual examination result
  - 
  - Necessary conditions for V hiking and re-re are covered
  - But those rules do not give sufficient conditions
  - Others are not well-analyzed yet

Experimental results (Table 4)

<table>
<thead>
<tr>
<th>Verbal suffix</th>
<th>( C_{\text{Recall}} )</th>
<th>( C_{\text{Precision}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>be</td>
<td>Vt 42/64</td>
<td>Vt 42/63</td>
</tr>
<tr>
<td></td>
<td>Vt 3/13</td>
<td>Vt 3/3</td>
</tr>
<tr>
<td></td>
<td>1/12/12</td>
<td>7/15/7/11</td>
</tr>
<tr>
<td>screw</td>
<td>Vt 3/7</td>
<td>Vt 3/3</td>
</tr>
<tr>
<td></td>
<td>Vt 9/2</td>
<td>Vt 9/0</td>
</tr>
<tr>
<td>to tighten-V-Past</td>
<td>Vt 1/2</td>
<td>Vt 1/7</td>
</tr>
<tr>
<td></td>
<td>Vt 2/4</td>
<td>Vt 2/2</td>
</tr>
<tr>
<td>da (ta)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Method: handmade rule-set

- The rule-set consists of 8 rules (Table 3, p.43)
  - Each labels either of verbal suffixes (Table 1)
  - Conditioned with conjunction of feature-value
  - 7 linguistic test results
  - Some clues are employed in addition
    - A: Semantic class of the noun [NUL, 2004]
      - e.g. ngi (screw); material f (1.4151)
    - D: Affix pattern
      - e.g. yuru (be loose) ⇒ yuru (to loosen): A, i-V, mu
    - C: Clause or not
  - Otherwise non- paraphrasable

Conclusion

- What we do:
  - Adj-to-V paraphrasing
    - As a case study of inter-categorial paraphrasing
  - Determining verbal suffixes in Adj-to-V paraphrasing
    - To ensure equivalence in aspectual meaning
    - Aspectual and agentive properties are useful
      - Majority of the test collection is explained (R, P=52.63)
      - (Beats ML and SLM for our small data-set)
    - (Over-generation & SLM-based filtering work well)
- Future work
  - Further experiments on Adj-to-V paraphrasing
    - with a larger dataset
      - Explore lexical properties of adjectives and contextual clues
      - Re-design the task: features → aspect → suffixes
  - Other types of inter-categorial paraphrasing
    - Predicative and adverbial usages of adjectives to verbs
    - V-to-Adj (reverse)
    - Adj-to-N, N-to-Adj, etc.
  - Re-normalization as a problem of lexical choice
    - In what condition, V is preferable to Adj, Vi is to Vt