Towards Automatic Scoring of Cloze Items by Selecting Low-Ambiguity Contexts

Tobias Horsmann & Torsten Zesch

Language Technology Lab
University of Duisburg-Essen
Germany
Fill-In-The-Blank-Tests (Cloze tests)

“He lost his _____ yesterday”
Fill-In-The-Blank-Tests (Cloze tests)

“He lost his _____ yesterday”
Fill-In-The-Blank-Tests (Cloze tests)

Solution: Automatic scoring
Requires: Knowing all valid answers

“He lost his ___ yesterday”

house
wife
car
job
hand
keys
glasses
wallet
soul
phone
marbles
confidence
hope
will
future
degree
friend
head
book
pen
rubber
hat
ring
Fill-In-The-Blank-Tests (Cloze tests)

Solution: Automatic scoring
Requires: Knowing all valid answers

“He lost his ___ yesterday”

Goal: Find low ambiguity contexts
(for nouns)
Reduced Ambiguity

An ambiguity reduced context

“The disease usually starts in the wrists, hands or feet, and can spread to other joints and other _____ of the body”

Predictable answers

<table>
<thead>
<tr>
<th>parts</th>
<th>regions</th>
<th>limbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>organs</td>
<td>areas</td>
<td></td>
</tr>
</tbody>
</table>

Ambiguity reduction in the example by
• Ngram word co-occurrence → “parts of the body”
• Hearst pattern (W such as X, Y or Z) → “body parts such as wrists, hands or feet”
Evaluation

• Compilation of a cloze test of detected contexts

• Target word deleted

• Two groups of sentences
  • Top-ranked: many evidences for low ambiguity
  • Bottom-ranked: few evidences for low ambiguity

• 30 native speakers of English were asked to restore the deleted word
Overview Results

- Ambiguity benchmark: Frequency of target word as answer
Thank you for your attention

Any questions?