FASTUS: Extracting Information from Natural-Language Texts

Presented by Warren Wong

Overview

• Motivations for FASTUS
• General Architecture
• Example run through
• Extensibility

Motivations

Want to Extract Information from Text

Context Free System
• Slow
• No need to understand all constructs

Example:
TACITUS
• Text Understanding System
• 36 hours to process 100 messages

Motivations

Finite State System
• Fast (12 minutes for 100 messages)
• Less processing than TACITUS but more relevant processing
• Uses less memory in general

FASTUS (Finite State Automaton Text Understanding System)
Example Run Through
Bridgestone Sports Co. said Friday it has set up a joint venture in Taiwan with a local concern and a Japanese trading house to produce golf clubs to be shipped to Japan.

The joint venture, Bridgestone Sports Taiwan Co., capitalized at 20 million new Taiwan dollars, will start production in January 1990 with production of 20,000 iron and "metal wood" clubs a month.
Bridgestone Sports Co. said Friday it has set up a joint venture in Taiwan with a local concern and a Japanese trading house to produce golf clubs to be shipped to Japan.

The joint venture, Bridgestone Sports Taiwan Co., capitalized at 20 million new Taiwan dollars, will start production in January 1990 with production of 20,000 iron and "metal wood" clubs a month.
Stage 3: Complex Phrases

- Attachment of appositives
- Construction of Measure Phrases
- Prepositional attachment of “for” and “of”
- Noun group conjunction

Create relations from complex noun groups

Example Run Through

Bridgestone Sports Co. said Friday it has set up a joint venture in Taiwan with a local concern and a Japanese trading house to produce golf clubs to be shipped to Japan.

The joint venture, Bridgestone Sports Taiwan Co., capitalized at 20 million new Taiwan dollars, will start production in January 1990 with production of 20,000 iron and "metal wood" clubs a month.
Example Run Through

Stage 4: Domain Events

Try to match phrases to the schema

- {Company/ies} {Set-up} {Joint-Venture} with {Company/ies}
- {Produce} {Product}
- {Company} {Capitalized} at {Currency}
- {Company} {Start} {Activity} in/on {Date}
Example Run Through
Stage 5: Merging Structures

So far, 1 or more relation for each sentence. Relationships spanning multiple sentences?
Coreference?
<table>
<thead>
<tr>
<th>Performance</th>
<th>Extensibility?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1st blind test</strong></td>
<td><strong>2nd blind test</strong></td>
</tr>
<tr>
<td>Recall</td>
<td>44%</td>
</tr>
<tr>
<td>Precision</td>
<td>55%</td>
</tr>
<tr>
<td>F-Score</td>
<td>48.9</td>
</tr>
<tr>
<td>GE-Recall</td>
<td>62%</td>
</tr>
<tr>
<td>GE-Precision</td>
<td>53%</td>
</tr>
<tr>
<td>Human</td>
<td>65-80%</td>
</tr>
</tbody>
</table>