Semi Cognitive Approach to RTE-Using FrameNet for Semantic Clustering

Himanshu Shivhare
Parul Nath
Anusha Jain
Where it got started...

- Two years ago...
- Research papers by Iftene, LCC etc.
- Analysis of approaches – interesting analysis
Graph depicts the relative low scores in Information Extraction Subtask.
The systems focus on the *face value* of the sentence and thus scores were low in IE subtask, which required *indirect inference*.

**Conclusion:** The context inference is important.
Example

RTE-1 Development Set:
<pair id="58" value="TRUE" task="IR">
T: Iraqi militants said Sunday they would behead Kim Sun-II, a 33-year-old translator, within 24 hours unless plans to dispatch thousands of South Korean troops to Iraq were abandoned.

H: Translator kidnapped in Iraq.

Entailment pair would not have been resolved by a system which can not view *kidnapping* and *beheading* under the same context of say danger.
Various Techniques

- Various approaches we analyzed
  - N-gram
  - Edit Distance with Dependency Trees
- All above had the same problems:
  - lack of semantics
  - lack of determinism
- For example (RTE-3)

**Text:** Claude Chabrol divorced Agnes, his first wife, to marry the actress Staphane Audran. His third wife is Aurore Paquiss.

**Hypothesis:** Aurore Paquiss married Chabrol.
FrameNet corpus is not new to RTE Fraternity. *(RTE 2 – Burchardt et al 2006)*

But, the usage highlighted the Frame Semantics: local roles, participants and also Frame Hierarchies.
A word \((W)\) may have different senses 
\(\{W_{s1}, W_{s2}\ldots W_{sn}\}\)

The word will *ideally* belong to \(n\) frames.

Each different sense of the word will make it belong to a different frame depicting that sense.
Although two words may not be synonyms or directly related, one of them (W1) might belong to a frame F1 due to sense $W1_{s1}$ in a particular context.

Other word, W2 may have a sense $W2_{s2}$, so that $W1_{s1} \iff W2_{s2}$
Both these words will be clustered under Frame F1, which will ease the semantic mapping.

The improved semantic mapping can be a powerful measure to improve the entailment deduction.
Outline

- Introduction
- System Architecture
  - Syntactic Analysis Module
  - Semantic Analysis Module
- Results & Evaluations
  - Error Analysis and Perspectives
- Conclusion
System Architecture

Preprocessing (XML + Co-referenc Resolution)

Syntactic Matching

Semantic Matching

Heuristics Module

Entailment Engine

Annotated T-H pair

In-House NER

Stanford DependencyParser

Verb Ocean

WordNet

Accronyms

Shalmanesar

Framenet

Corpus

T-H Pair

In house NER

Yes/No
We extract the ‘Syntactic Roles’ of entities from Stanford dependencies.

The dependency tree overlap measures the syntactic similarity between the t-h pairs.

This match gives equal importance to all the relations.
Example:

* Hypothesis: The Irish Republican Army is a Catholic paramilitary group.
Text I: Answering questions in parliament, Ahern told MPs that the meetings with Adams, whose party is the political wing of the IRA, Northern Ireland's main Catholic paramilitary group, had merely been to maintain dialogue.

Many a times, the semantic equivalence is not reflected in the syntactic structure of two sentences.
Instead of typical specific rules which feature in traditional rule based system, Sangyan uses GAH rules (Generically Applicable Heuristics derived from Dependencies).

GAH are templates which can be applied to extract meaning from a wide range of constructs.
<table>
<thead>
<tr>
<th>T-H pairs (RTE-6 Development Set)</th>
<th>Meaning</th>
<th>Syntactic Role</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis:</strong> Gerry Adams is the leader of Sinn Fein.</td>
<td>Gerry Adams &quot;is-a&quot; leader of Sinn Fein</td>
<td>nsubj(leader, Adams)</td>
</tr>
<tr>
<td><strong>Text I:</strong> London and Dublin are awaiting for an Irish Republican Army response to a call from <strong>Gerry Adams, leader of the group's Sinn Fein political wing,</strong> for an end to violence.</td>
<td>Gerry Adams &quot;is-a&quot; leader of the group's Sinn Fein political wing</td>
<td>appos(Adams, leader )</td>
</tr>
<tr>
<td><strong>Text II:</strong> Irish Prime Minister Bertie Ahern admitted on Tuesday that he had held a series of private one-on-one meetings on the Northern Ireland peace process with <strong>Sinn Fein leader Gerry Adams,</strong> but denied they had been secret in any way.</td>
<td>Gerry Adams &quot;is-a&quot; leader of Sinn Fein</td>
<td>nn(Adams, leader)</td>
</tr>
</tbody>
</table>
The biggest challenge for rules is their specific applicability.

* Ex: Gerry Adams is the leader of Sinn Fein.
* ...Gerry Adams, leader of the group's Sinn Fein political wing...

It takes one GAH to tackle both the cases, but, two rules to do the same.
Outline

- Introduction
- System Architecture
  * Syntactic Analysis Module
  * Semantic Analysis Module
- Results & Evaluations
  * Error Analysis and Perspectives
- Conclusion
Semantic Module

- Semantic module handles semantic variability
  - Semantic module tries establish semantic relatedness between the TH pair
    For example:
    Text: *Satomi Mitarai died of blood loss.*
    Hypothesis: *Satomi Mitarai bled to death.*

- FrameNet frames provide an excellent way of representing the context and semantics.
  - Abstract over the various syntactical level problems like noun verb alternation etc
  - We use a combination of Frame Match and Lexical Overlap and not frame elements

- In this example, the frame of both “died” and “death” is “Death”
We use a combination of Frame Match and Lexical Overlap

For example (RTE -6 )

Text : The Patriot Act, passed by Congress a few weeks after the Sept. 11, 2001, terror attacks, gave federal law enforcement officials broader powers of surveillance and prosecution against suspected terrorists, their financiers and their sympathizers.


The frame “Giving” matches in both text and hypothesis and there is a high lexical overlap between the text and hypothesis.
Certain frames are equivalent in particular context.

For example:

**Text:** South Korea has **lifted** a five-year ban on beef imports from the US, despite growing public protests prompted by fears of mad cow disease.

**Hypothesis:** South Korea **removes** a US beef ban.

We get frames *Removing → remove*  
*Body_movement → lifted*

We identified both manually and automatically some such frames and added the functionality to equate them during the match process.

This helps in widening the match criterion of the frame match.
Text: Bahrain's king has appointed a Jewish woman as the country's envoy to the United States. Bahrain's king has appointed a Jewish woman as the country's envoy to the United States.

Hypothesis: Bahrain names a Jewish ambassador.

Equivalent frame: Being_named ⇒ name
Change_of_of_leadership --> appoint
FrameNet is a useful resource

But, **Lack of Coverage** is a problem

* Although FrameNet contains more than 6500 fully annotated lexical units and associated frames,
* not sufficient for many real world scenarios such as RTE-6.
* Also, it takes expert human supervision for the addition of new frames, which is a time consuming task.
* For our system handling coverage become easy as we were not using frame elements
* Added many frames both manually and automatically
Outline

- Introduction
- System Architecture
  - Syntactic Analysis Module
  - Semantic Analysis Module
- Results & Evaluations
  - Error Analysis and Perspectives
- Conclusion
# Report Card Time!!

## RTE 6 – Main Task

<table>
<thead>
<tr>
<th>Development Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>Precision</td>
</tr>
<tr>
<td>Recall</td>
<td>Recall</td>
</tr>
<tr>
<td>F-Measure</td>
<td>F-Measure</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>Recall</td>
<td>F-Measure</td>
<td>Precision</td>
<td>Recall</td>
<td>F-Measure</td>
</tr>
<tr>
<td>30.64</td>
<td>33.0</td>
<td>31.78</td>
<td>21.66</td>
<td>46.03</td>
<td>29.46</td>
</tr>
</tbody>
</table>

© NEC HCL ST
Introspect!! Failure Analysis

- **Median Result**
  - presents immense opportunities to improve

- **Major working area**: 
  - FrameNet module
    - Requirement of more data for FrameNet equivalence
    - Coverage: Need more frames
  - Lexical Module
    - Intelligent match
  - Syntactic Module
    - Give weight age to more important relation
  - Training the system for setting appropriate Threshold
Problem Area: FrameNet

- More frames for frame equivalence required
- **Text**: Her interpreter, Allan Enwiyah, 32, was shot dead and his body abandoned nearby by the kidnappers, while her **driver got away**.
  **Hypothesis**: Jill Carroll’s **driver escaped**

* the frame of “escape” is “Departing” and that of “get away” is “Evading”.
**Text:** Vice President Dick Cheney accidentally sprayed a companion with birdshot while hunting quail on a private Texas ranch, injuring the man in the face, neck and chest, the vice-president’s office confirmed yesterday after a Texas newspaper reported the incident.

**Hypothesis:** Harry M. Whittington is Vice President Dick Cheney’s hunting companion.

Despite the **low frame match**, the true entailment is deduced due to the high lexical overlap.
Text: The White House on Monday carefully distanced itself from Vice President Dick Cheney’s delayed notification about his accidental shooting of a hunting companion. **Hypothesis**: Harry M. Whittington is Vice President Dick Cheney’s hunting companion.

Probabilistic nature of the syntactic module of our system, which causes the system to give many false positives.
Outline

- Introduction
- System Architecture
  - Semantic Analysis Module
  - Syntactic Analysis Module
- Results & Evaluations
  - Error Analysis and Perspectives
- Conclusions
Future looks Meaningful

- While syntactic analysis based systems start hitting the wall pretty soon, Adding semantic analysis module definitely provides a new dimension of growth and improvisation.
- Scope was adding new frames is immense
- Exploring semantic roles for matching
- Problem areas are very solvable
Thank You

- Questions/Answers

- Graphics by UVassociates
  www.uvassociates.in