Natural Language Inference: Overview

Christopher Potts

Stanford Linguistics

CS224u: Natural language understanding
Associated materials

1. Code
   a. nli.py
   b. nli_01_task_and_data.ipynb
   c. nli_02_models.ipynb

2. Homework and bakeoff: hw_wordentail.ipynb

3. Core readings: Bowman et al. 2015; Williams et al. 2018; Nie et al. 2019; Rocktäschel et al. 2016

4. Auxiliary readings: Goldberg 2015; Dagan et al. 2006; MacCartney and Manning 2008; Gururangan et al. 2018
## Simple examples

<table>
<thead>
<tr>
<th>Premise</th>
<th>Relation</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>A turtle danced.</td>
<td>entails</td>
<td>A turtle moved.</td>
</tr>
<tr>
<td>turtle</td>
<td>contradicts</td>
<td>linguist</td>
</tr>
<tr>
<td>Every reptile danced.</td>
<td>neutral</td>
<td>A turtle ate.</td>
</tr>
<tr>
<td>Some turtles walk.</td>
<td>contradicts</td>
<td>No turtles move.</td>
</tr>
<tr>
<td>James Byron Dean refused to move without blue jeans.</td>
<td>entails</td>
<td>James Dean didn’t dance without pants.</td>
</tr>
<tr>
<td>Mitsubishi Motors Corp’s new vehicle sales in the US fell 46 percent in June.</td>
<td>contradicts</td>
<td>Mitsubishi’s sales rose 46 percent.</td>
</tr>
<tr>
<td>Acme Corporation reported that its CEO resigned.</td>
<td>entails</td>
<td>Acme’s CEO resigned.</td>
</tr>
</tbody>
</table>
NLI task formulation

Does the premise justify an inference to the hypothesis?

- Commonsense reasoning, rather than strict logic.
- Focus on local inference steps, rather than long deductive chains.
- Emphasis on variability of linguistic expression.

Perspectives

- Zaenen et al. (2005): Local textual inference: can it be defined or circumscribed?
- Manning (2006): Local textual inference: it’s hard to circumscribe, but you know it when you see it – and NLP needs it.
- Crouch et al. (2006): Circumscribing is not excluding: a reply to Manning.
Connections to other tasks

Dagan et al. (2006)
It seems that major inferences, as needed by multiple applications, can indeed be cast in terms of textual entailment.

[...]

Consequently, we hypothesize that textual entailment recognition is a suitable generic task for evaluating and comparing applied semantic inference models. Eventually, such efforts can promote the development of entailment recognition “engines” which may provide useful generic modules across applications.
Connections to other tasks

<table>
<thead>
<tr>
<th>Task</th>
<th>NLI framing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paraphrase</td>
<td>text $\equiv$ paraphrase</td>
</tr>
<tr>
<td>Summarization</td>
<td>text $\sqsubset$ summary</td>
</tr>
<tr>
<td>Information retrieval</td>
<td>query $\sqsupset$ document</td>
</tr>
<tr>
<td>Question answering</td>
<td>question $\sqsupset$ answer</td>
</tr>
<tr>
<td></td>
<td>$Who$ left? $\Rightarrow$ $Someone$ left</td>
</tr>
<tr>
<td></td>
<td>$Someone$ left $\sqsupset$ Sandy left</td>
</tr>
</tbody>
</table>
Models for NLI

- Logic and theorem proving
- Natural Logic
- Semantic graphs
- Clever hand-built features
- Deep learning (2015)
- N-gram variations

Effectiveness

Depth of representations

- Bos & Markert 2005
- MacCartney 2009
- Hickl et al. 2006; de Marneffe et al. 2006

See the Excitement Open Platform (Pado et al. 2012)

A standard baseline, often very robust!
Models for NLI

- Logic and theorem proving (Bos & Markert 2005)
- Natural Logic
- Semantic graphs
- Clever hand-built features
- N-gram variations
- Deep learning (2017–)

Depth of representations vs. effectiveness:
- Deep, brittle
- Robust, shallow

Standard baseline, often very robust!

See the Excitement Open Platform (Pado et al. 2012)
References


Yoav Goldberg. 2015. A primer on neural network models for natural language processing. Ms., Bar Ilan University.


Christopher D. Manning. 2006. Local textual inference: It’s hard to circumscribe, but you know it when you see it – and NLP needs it. Ms., Stanford University.


