Does Logic-based Reasoning Work for Dutch?

Lasha Abzianidze    Konstantinos Kogkalidis

Utrecht Institute of Linguistics OTS, Utrecht University

16-06-2021 @ NaLoMa
Logic-based approach to NLI

Raw sentences in 🇳🇱

Alpino & Neural Proof Nets

ILL_{◊,□} derivations

Conversion rules

\(\lambda\rightarrow\) terms

Open Dutch WordNet

LangPro

Relation label
The linear $\lambda$-calculus

Words are assigned ILL types, inductively defined as: $\mathcal{T} := a \mid t_1 \rightarrow t_2 \mid \diamond \delta t \mid \Box \alpha t$

where

- $a$ an atom, from a finite set $A$:
  - $np, s_{main}, s_{sub}, pron, ...$
- $t_1 \rightarrow t_2$ a linear function that consumes $t_1$ to produce $t_2$
  - $np \rightarrow s_{main}, np \rightarrow (np \rightarrow s_{main}), (np \rightarrow s_{sub}) \rightarrow (np \rightarrow np), ...$

Syntactic derivations $\equiv$ proofs $^c_{hc}$ functional programs:

$$\tau := c^t \mid (\tau_1^{t_1 \rightarrow t_2} \tau_2^{t_1})^{t_2} \mid (\lambda x^{t_1} . \tau_2)^{t_1 \rightarrow t_2} \mid \delta \square \tau \mid \delta \square \tau \mid \delta \diamond \tau \mid \delta \diamond \tau$$

A boy plays:

$$\text{play}^{\diamond su np \rightarrow s_{main}} (su^{\diamond} (\text{det}^{\square} a ^{\square \text{det} (n \rightarrow np)} \text{boy}^n))$$
Proof Nets

Proof net

A polarized forest (proof frame) and a bijection between $+$ and $-$ (axiom links)

\[
\begin{array}{cccc}
  n & np & & n \\
\end{array}
\]

\[
\begin{array}{cccc}
  n & n & n & n \\
\end{array}
\]

\[
\begin{array}{cccc}
  - & - & & - \\
\end{array}
\]

\[
\begin{array}{cccc}
  n & n & n & n \\
\end{array}
\]

A very easy example
Neural Proof Nets: from sentences to $\lambda$-terms

Supertagging

From sentences to proof frames with seq2seq transduction

Proving

From proof frames to axiom links with Sinkhorn-Knopp

\[ n_1 \quad n_{p2} \quad n_3 \quad n_4 \quad n_5 \quad n_6 \quad n_7 \quad n_8 \quad n_9 \]

a very easy example
Other ingredients: Alpino & ODWN

Alpino:
- Stochastic Attribute Value Grammar (HPSG) for Dutch
- Builds dependency structures
- Used for pre-processing Dutch treebanks

Open Dutch WordNet:
- 52K synsets (vs 117K Princeton WN)
- Derived from the Dutch lexical semantic database Cornetto
- We converted ODWN into the prolog format
- Used relations: synonymy, hypernymy, antonymy, similarity, and derivation.
Natural Tableau: proving and learning

1. a hedgehog \( (\lambda x. a \text{ boy } (\lambda y. \text{ by } y \text{ cradle } x)) \): T
2. a person \( (\lambda x. a \text{ (small animal)} (\lambda y. \text{ hold } y x)) \): F

3. hedgehog: \([h]: T\)
4. a boy \( (\lambda y. \text{ by } y \text{ cradle } h) \): T
5. boy: \([b]: T\)
6. by \( b \) cradle: \([h]: T\)
7. cradle: \([h, b]: T\)
8. person: \([b]: F\)
9. a (small animal) \( (\lambda y. \text{ hold } y b) \): F
10. small animal: \([h]: F\)
11. hold: \([h, b]: F\)

\( \times \)

OPEN

- Natural logic + semantic tableau
- Uniform to Ent./Cont.
- Prove with refutation
- Uniform to premise number
- \( \approx \) Syntactic trees
- Native higher-order logic
- Decomposing meaning
- Learning as abduction: hedgehog \( \sqsubseteq \) small animal
Syntactic $\lambda$-terms to $\lambda$-logical forms

\[
\text{play}^{\text{su\,np}} \overset{\text{smain}}{\rightarrow} \left( \text{su}^{\diamond} \left( \left( \text{det}^{\square} \text{a}^{\text{det}(\text{n-np})} \text{boy}^{\text{n}} \right) \right) \right) \quad \leadsto \quad \text{play}^{\text{np},s} (\text{a}^{\text{n-np}} \text{boy}^{\text{n}})
\]

\[
\text{large}^{\text{np},\text{np}} \left( \text{brown}^{\text{np},\text{np}} (\text{a}^{\text{n-np}} \text{dog}^{\text{n}}) \right) \quad \leadsto \quad \text{a}^{\text{n-np}} \left( \text{large}^{\text{n-n}} (\text{brown}^{\text{n-n}} \text{dog}^{\text{n}}) \right)
\]

and \( (\lambda x. \text{brown}(x \text{ dog})) \( (\lambda y. \text{black}(y \text{ dog})) \) no

\[
\leadsto \quad \text{and}^{\text{np},\text{np},\text{np}} \left( \text{no} (\text{brown dog}) \right) \left( \text{no} (\text{black dog}) \right)
\]

\[
\text{cut}^{\text{pp},\text{n-np},s} (\text{in}^{\text{n-pp}} \text{slice}_{\text{n}}) \text{meat}_{\text{n}} \quad \leadsto \quad \text{cut}^{\text{pp},\text{np},\text{np},s} (\text{in}^{\text{np},\text{pp}} \{\text{slice}^{\text{n}}\text{np}\}) \{\text{meat}^{\text{n}}\text{np}\}
\]

- Map POS tags and shift to slightly *Generalized* POS tags: UPOS & Penn
- Use only these syntactic categories: \( \text{n}, \text{np}, \text{x}, \text{s}, \text{pp}, \text{pr} \)
- Function words $\mapsto$ canonical terms (excl. prepositions)
Natural Tableau for Dutch

1771 Entailment
Een jongen speelt vrolijk piano
en piano wordt bespeeld door een persoon

1771 Entailment
A boy is happily playing the piano
A piano is being played by a person
Experiments & Results

<table>
<thead>
<tr>
<th>Parser / POS</th>
<th>Te</th>
<th>Tα:Te</th>
<th>Ee</th>
<th>Tα:Ee</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpino / Alpino</td>
<td>72.7</td>
<td>(+9.3)</td>
<td>74.1</td>
<td>(+1.8)</td>
</tr>
<tr>
<td>Alpino / spaCy</td>
<td>74.8</td>
<td>(+9.5)</td>
<td>75.9</td>
<td>(+1.7)</td>
</tr>
<tr>
<td>NPN / Alpino</td>
<td>72.0</td>
<td>(+8.6)</td>
<td>72.8</td>
<td>(+1.5)</td>
</tr>
<tr>
<td>NPN / spaCy</td>
<td>74.3</td>
<td>(+9.1)</td>
<td>75.0</td>
<td>(+1.4)</td>
</tr>
<tr>
<td><strong>LangPro 2×2</strong></td>
<td>76.0</td>
<td>(+9.8)</td>
<td>77.1</td>
<td>(+1.6)</td>
</tr>
</tbody>
</table>

- POS tagging: spaCy is 1.7\(\leq\) better than Alpino;
- Parsers: Alpino \(>\) NPN? Sentences parsed 98.6\% vs 94.9\%;
- The ensemble is 1.2\(\leq\) better than the best monoLP;
- The number of false proofs increases by 4\% after αductive learning.
Comparison to the transformer-based NLI models

<table>
<thead>
<tr>
<th>Models</th>
<th>All ±Δ</th>
<th>Ent ±Δ</th>
<th>Cont ±Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>LangPro 2×2</td>
<td>78.7</td>
<td>50.6</td>
<td>66.3</td>
</tr>
<tr>
<td>BERTje</td>
<td>82.0</td>
<td>86.2</td>
<td>86.7</td>
</tr>
<tr>
<td>mBERT</td>
<td>79.9</td>
<td>79.0</td>
<td>81.9</td>
</tr>
<tr>
<td>RobBert</td>
<td>81.7</td>
<td>76.9</td>
<td>85.3</td>
</tr>
</tbody>
</table>

Problems failed by all three DL models but solved by LangPro:

- **Entailment**: A man is carrying a *tree*  A man is carrying a *plant*
- **Entailment**: A family is watching a little boy who is hitting a baseball  A boy is hitting a baseball
- **Contradiction**: The person is *not* drawing  A man is drawing a picture
- **Contradiction**: A woman in a red dress is *putting away* an instrument  A woman in a red dress is *playing* an instrument
Findings: SICK NL & Open Dutch WN

SICK NL is more challenging that the original dataset due to MT:

- Transferred gold labels are not gold:
  
<table>
<thead>
<tr>
<th>3181</th>
<th>Neutral?</th>
</tr>
</thead>
<tbody>
<tr>
<td>A man is trekking in the woods</td>
<td>→</td>
</tr>
<tr>
<td>The man is not hiking in the woods</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3181</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Een man is aan het wandelen in het bos</td>
<td>→</td>
</tr>
<tr>
<td>De man is niet aan het wandelen in het bos</td>
<td></td>
</tr>
</tbody>
</table>

- Extra reasoning due to translation shifts:
  
  drawing a picture → een tekening maakt | tekent een foto
dirt bike race → crossmotorwedstrijd | crossmotorrace.

Lexical relations learned from the training set:

- lopen ≡ rennen
- pizza ⊑ voedsel/food
- halter/dumbbell ⊑ gewicht/weight
- leeg/empty | vol/full
Conclusion & Future work

- First Dutch NLI system based on logic
- YES! Logic-based Reasoning Works for Dutch: promising results
- Automatic translation makes the NLI data more challenging

- Employ the Dutch CCG parser
- Qualitative comparison of the Dutch syntactic parsers
- Multilingual LangPro: SICK + UD parsers for λLogical Forms