How to Train a Theorem Prover for Natural Language Inference

Lasha Abzianidze

Universiteit Utrecht

LAComplIng, 17.12.2021
There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians.

The artificial languages of logicians usually come with proof systems. **What do proof systems of natural languages look like?**

I will present the **Natural Tableau** that is a semantic tableau-based proof system for English.
Outline

Natural Tableau

Natural language theorem prover

Learning as abduction

Experiments & results

Conclusion & references
Natural Tableau
A lark is not moving
contradicts
Every bird is flying

Syntactic parsing, \[ \lambda \]

\[ \exists x (\text{lark}(x) \land \neg \text{move}(x)) \]
\[ \forall x (\text{bird}(x) \rightarrow \text{fly}(x)) \]

contradicts

Syntactic parsing, rewriting

a lark (not (be move))
every bird (be move)

Natural theorem proving

Blackburn & Bos (2005)
## Lambda Logical Forms

<table>
<thead>
<tr>
<th>English</th>
<th>Lambda Logical Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>John sleeps</td>
<td>sleep john</td>
</tr>
<tr>
<td>Every bird flies</td>
<td>every bird fly</td>
</tr>
<tr>
<td>No bird flies</td>
<td>no bird fly</td>
</tr>
<tr>
<td>The lark does not fly</td>
<td>the lark (not fly)</td>
</tr>
<tr>
<td>John sleeps and Mary sings</td>
<td>and (sleep john) (sing mary)</td>
</tr>
<tr>
<td>Every man loves a woman</td>
<td>every man (λs. a woman (λo. love o s))</td>
</tr>
<tr>
<td></td>
<td>a woman (λo. every man (love o))</td>
</tr>
</tbody>
</table>

No \(\land, \lor, \rightarrow, \exists, \neg, \forall\) are used.

R. Muskens (2010): An analytic tableau system for natural logic

\[
A B C \equiv (A B) C
\]
Natural tableau entries

\[ (\text{Term } a_1, a_2, \ldots, a_n) \text{ is true} \]

\[ (\text{Term } a_1, a_2, \ldots, a_n) \text{ is false} \]

John sleeps

It is false that the lark doesn't fly

\[ \text{sleep john} \]

\[ \text{the lark (not fly)} \]

\[ \text{sleep john} \]

\[ \text{the lark not fly} \]

\[ \text{the lark, not fly} \]

R. Muskens (2010): An analytic tableau system for natural logic
A lark is not moving
contradicts

Natural tableau proof

A lark is not moving
covers contradictions
Every bird is flying

\[
\begin{align*}
\text{a lark (not (be move))} & \text{ every bird (be fly)} \\
\text{lark} & \text{ c} \\
\text{not (be move)} & \text{ c} \\
\text{be move} & \text{ c} \\
\text{move} & \text{ c} \\
\text{bird} & \text{ c} \\
\text{be fly} & \text{ c} \\
\text{fly} & \text{ c} \\
\end{align*}
\]


\[
\begin{align*}
\text{every S V} & \\
\text{S c} & \text{ V c} \\
\text{be X a} & \text{ X a} \\
\text{be X a} & \text{ X a} \\
\text{X a} & \text{ Y a} \\
\text{X a} & \text{ Y a} \\
\end{align*}
\]

\[
\begin{align*}
\text{a S V} & \\
\text{S c} & \text{ V c} \\
\text{not X a} & \text{ X a} \\
\end{align*}
\]

Proved

R. Muskens (2010): An analytic tableau system for natural logic
**Natural logic** is the study of inference in natural language, done as close as possible to the "surface forms" (Moss, 2010)

LLFs are terms of **higher-order logic**, namely, simple type theory:

\[
\begin{align*}
\text{a (et)(et)t} & \quad \text{lark et} \\
\underline{\text{a lark (et)t}} & \quad \text{not (et)et} \\
\underline{\text{a lark (not (be move)) t}}
\end{align*}
\]

**Tableau method** is very intuitive: theorem-proving as a model construction
Natural Language Theorem Prover
Scaling up the natural tableau

Fine-grained entries: \( m_1, \ldots, m_k \) \( \text{Term} \) \( a_1, \ldots, a_n \)

Syntactic types: \( \{ e, t \} + \{ n, np, s, pp \} \)

\( s <: t \quad e <: np \quad n <: et \quad pp <: et \)
\( \alpha\beta <: \gamma\delta \quad \text{if} \quad \gamma <: \alpha \quad \text{and} \quad \beta <: \delta \)

More tableau rules:
- Adjectives
- Prepositions/particles
- Passive constrictions
- Subcategorization
- Definite NPs
- Expletives
- Open compound nouns
- Light verb constructions
- Attitude verbs
- PP attachments

L. Abzianidze (2016): A natural proof system for natural language
Natural Language Theorem Prover

LangPro

Premises
Hypothesis

CCG parser
Trees
Logical Form generator
Terms
Natural Tableau prover
Prediction

C&C-based LangPro
EasyCCG-based LangPro

KB
WordNet

Relations

ONLINE DEMO
https://naturallogic.pro/LangPro

L. Abzianidze (2017): LangPro: Natural language theorem prover
Natural Language Inference (NLI)

aka Recognizing Textual Entailment (RTE)

Gold inference label = Entailment | Contradiction | Neutral

A man in a hat is playing a **harp**

E | A man is playing an **instrument**

The woman is **not** wearing glasses **or** a headdress

C | A woman is wearing an **Egyptian** headdress

A woman is dancing and singing with other woman

N | A woman is dancing and singing **in the rain**

From SICK (Sentences Involving Compositional Knowledge) (Marelli et al., 2014)
Findings on SICK

Someone is playing with a toad
N/E Someone is playing with a frog

A couple is not looking at a map
N/C A couple is looking at a map

A soccer ball is not rolling into a goal net
C/C A soccer ball is rolling into a goal net

A man in a hat is playing a harp
E A man is playing an instrument

The woman is not wearing glasses or a headdress
C A woman is wearing an Egyptian headdress

Wrong gold
Inconsistent gold
Solved hard problems

L. Abzianidze (2015): A tableau prover for natural logic and language
Failed proofs

A small boy in a yellow shirt is laughing on the beach

\[ C/N \] There is no small boy [in a yellow shirt [laughing on the beach]]

A man and a woman are shaking hands

\[ E/N \] Two persons are shaking hands

Missing inference rule

It is raining on a walking man

\[ E/N \] A man is walking in the rain

25% of errors

A man and woman are talking

\[ C/N \] A man and a woman are silent

Missing lexical knowledge

Someone is holding a hedgehog

\[ E/N \] Someone is holding a small animal

50% of errors

L. Abzianidze (2015): A tableau prover for natural logic and language
Pros & Cons

- Higher-order logic
- Semantic parsing ≈ syntactic parsing
- Highly reliable proofs
- Heavily hinges on syntactic parsing
- Knowledge is restricted to KBs
Learning as Abduction
Approaches to NLI

Reasoning with Logic

- puppy $\equiv$ young dog
- slice $\equiv$ cut
- guitar | person
- put on | take off

Learn from Data
- robust
- scalable
- easy to deploy
- no semanticists
A clown is putting eyeshadow on
The person is quickly removing make-up

Contradiction ≠ put on | remove

Neutral

A clown is putting eyeshadow on
The person is quickly removing make-up
Questions

1. What is a feasible way to learn from data via abduction?

2. How to prevent learning pseudo-knowledge?

3. Is the learned knowledge enough alone?

4. How does the learned knowledge boosts performance?
Abductive reasoning in logic

Inference to the best explanation

$E \lor \text{Premise} \land \text{Conclusion}$

**In the NLI context**

- Lovely little puppy is sleeping:
  - lovely $|$ little
  - lovely $\sqsubseteq$ little

- puppy $\sqsubseteq$ dog; lovely $\sqsubseteq$ little
  - puppy $\sqsubseteq$ dog

- lovely little black puppy $\sqsubseteq$ dog
  - puppy $\sqsubseteq$ dog

$E$ is consistent with Premise

$E$ is minimal

$E$ has a pre-defined forms

Mayer & Pirri (1993)
Learning as abduction

A nestling is not moving

Every bird is flying

A lark is not moving

Every bird is flying

nestling ⊑ bird

nestling ⊑ move

bird ⊑ move

be fly ⊑ move

be move ⊑ move

move ⊑ move

Reverse app

nestling ⊑ bird

Lasha Abzianidze
Abduction in Natural Tableau

General picture

\[ X \vdash \bar{a} \]

\[ Y \vdash \bar{a} \]

and other 16 closure rules

If \( X \subseteq Y \)

5,184 Possible \( E \)

Shrink Explanation Space
Abductive learning workflow

Calculate impact based on relevant problems

Incorporate learned relations and restart the learning phase if new relations were learned

LangPro

Initial KB

Learned KB

All problems

Unsolved problems

Learned relations

Abductive inference

puppy ⊑ dog
puppy ⊑ little
puppy ⊑ young dog

The toy car is rolling down a hill

The toy car ⊑ car

The toy ⊑ car

The car ⊑ plastic vehicle

Note ⊑ paper

+3 vs -0

+1 vs -2

len 3

len 2

Learned
Experiments & Results
Experiment setup

Sentences Involving Compositional Knowledge (SICK, Marelli et al., 2014)

Train 5,000 3-fold cross validation
Test 4,927 unseen data

CCG parsers:

- C&C (Clark and Curran, 2007)
- EasyCCG (Lewis and Steedman, 2014)
- DepCCG (Yoshikawa et al., 2017)

LangPro:

Rule application limit: 50 & 800

https://github.com/kovvalsky/LangPro
Ablation & efficiency

3-fold CV on the training data, using only C&C parser

<table>
<thead>
<tr>
<th>LangPro + Abduction</th>
<th>Av. accuracy on train fold %</th>
<th>Av. accuracy on test fold %</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 800 rule app.</td>
<td>89.02</td>
<td>82.90</td>
</tr>
<tr>
<td>≤ 50 rule app.</td>
<td>88.57</td>
<td>82.64</td>
</tr>
<tr>
<td>- WordNet</td>
<td>-0.98</td>
<td>-2.44</td>
</tr>
<tr>
<td>+ KB</td>
<td>+0.04</td>
<td>+0.36</td>
</tr>
</tbody>
</table>

~10× faster
## Results & comparison

<table>
<thead>
<tr>
<th>LangPro + WN (accuracy % on test)</th>
<th>C&amp;C</th>
<th>EasyCCG</th>
<th>DepCCG</th>
<th>DEC &amp; C&amp;C</th>
<th>+ Abduction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>81.3</td>
<td>81.1</td>
<td>81.8</td>
<td>83.0</td>
<td>+1.4</td>
</tr>
</tbody>
</table>

### Prec % | Rec% | Acc %
---|---|---
94.3 | 67.9 | 84.4
84.2 | 77.3 | 84.3

Abzianidze & Kogkalidis (2021): 77.2

A Logic-Based Framework for Natural Language Inference in Dutch

Yanaka et al. (2018)

Some meat is being cut into pieces by a woman

A lady is cutting up some meat precisely

\[
\text{preci} \quad \text{cut up} \quad w \\
\text{s piece } (\lambda x. \text{cut into } x) m \quad w
\]
Learned knowledge

- A barefoot man in pajamas is looking toward the **stars** and ...  
- A barefoot man in pajamas is looking toward the **sky** and ...  
- A cyclist is biking in a snowy forest **in the dark**  
- A cyclist is biking in a snowy forest **at night**  
- **Someone** is drilling a hole in a strip of wood with a power drill  
- **A man** is drilling a hole in a piece of wood  
- ... and performing **acrobatics** on a rail  
- ... and performing a **trick** on a rail  
- A squirrel is **lying down**  
- A squirrel is **running around** in circles

<table>
<thead>
<tr>
<th>Correct</th>
<th>Wrong</th>
<th>Contextual</th>
<th>Reversed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>26,60%</td>
<td>28,80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28,90%</td>
<td>15,70%</td>
</tr>
</tbody>
</table>

In total 312
1. What is a feasible way to learn from data via abduction?
Reduce the explanation space & apply the filters.

2. How to prevent learning pseudo-knowledge?
The filters partially prevent pseudo-knowledge. However, it doesn’t harm the performance.

3. Is the learned knowledge enough alone?
No. WordNet is crucial for reaching the state-of-the-art results.

4. How does the learned knowledge boosts performance?
It consistently increases Acc. while Prec. is still high
Conclusion

Natural proof system for natural logic (and language)

Learning as abduction: reverse rule applications to abduce knowledge from labeled inferences (currently severely restricted due to efficiency)

The tableau proofs are rarely wrong:

- helps to detect errors in gold labels of NLI datasets
- can be used to validate labels in NLI datasets
- provides explanatory reasoning
• Lasha Abzianidze. 2015. A tableau prover for natural logic and language. EMNLP.


• Lasha Abzianidze. 2017. LangPro: Natural language theorem prover. EMNLP.


• Mike Lewis and Mark Steedman. 2014. A* CCG parsing with a supertag-factored model. EMNLP.


• Hitomi Yanaka, Koji Mineshima, Pascual MartinezGomez, and Daisuke Bekki. 2018. Acquisition of phrase correspondences using natural deduction proofs. NAACL.
References (2)

- Hitomi Yanaka, Koji Mineshima, Pascual MartínezGomez, and Daisuke Bekki. 2018. Acquisition of phrase correspondences using natural deduction proofs. NAACL.
Attribution

• Free Icons from the Streamline Icons Pack: https://streamlineicons.com

• Most of the cliparts are take from: https://pixabay.com

https://giphy.com/gifs/research-madebydot-gifathon-26FPC5oAdfeFPkQQE

https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcT0oYysClFwZIg9E8sbT-0QoKnro58nShr4_IMOVr24dwgsD_GN&s


https://freesvg.org/img/hand-pencil.png