Learnability-based Syntactic Annotation Design

Roy Schwartz, Omri Abend
and Ari Rappoport

The Hebrew University
In proceedings of COLING 2012
Overview

• In many cases, there is more than one plausible way to annotate syntactic structures
  – A single annotation must be selected

• We propose learnability as a selection criterion
  – A principled learnability-based methodology
  – Use parsers for annotation design

• Selecting the more learnable annotation may result in up to \textbf{35.3\%} error reduction in parsing performance
Different Syntactic Formalisms

Example

• I want to eat the apple
Different Syntactic Formalisms

Example

- I want to eat the apple

<table>
<thead>
<tr>
<th>PRP</th>
<th>VBP</th>
<th>TO</th>
<th>VB</th>
<th>DT</th>
<th>NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>want</td>
<td>to</td>
<td>eat</td>
<td>the</td>
<td>apple</td>
</tr>
</tbody>
</table>
Different Syntactic Formalisms

Example

• I want to eat the apple

(\(S\) (\(NP\) I) (\(VP\) (\(VP\) want ) (\(VP\) to (\(VP\) eat (\(NP\) the apple) ))))

I   want   to   eat   the   apple

PRP   VBP   TO   VB   DT   NN

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Different Syntactic Formalisms

Example

• I want to eat the apple

\[
(S \ (NP \ I) \ (VP \ (VP \ want) \ (VP \ to) \ (VP \ eat \ (NP \ the \ apple))) ) ) )
\]

Learnability-based Syntactic Annotation Design @
Schwartz, Abend and Rappoport, COLING 2012
Disagreement within the same Formalism

I want to eat the apple
Disagreement within the same Formalism

I want to eat the apple
Disagreement within the same Formalism

I want to eat the apple

I want to eat the apple
Disagreement within the same Formalism

I want to eat the apple

I want to eat the apple
Disagreement within the same Formalism

I want to eat the apple

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Disagreement within the same Formalism

I want to eat the apple

I want to eat the apple

I want to eat the apple

I want to eat the apple

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Disagreement within the same Formalism

I want to eat the apple

Varying Syntactic Structures (VSS)
Disagreement within the same Formalism

I want to eat the apple

Varying Syntactic Structures (VSS)

Annotation scheme

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Varying Syntactic Structures (VSS)

• VSSs are very frequent
  – More than 40% of the tokens in PTB participate in at least one VSS*

* Schwartz et al., ACL 2011
Varying Syntactic Structures (VSS)

• VSSs are very frequent
  – More than 40% of the tokens in PTB participate in at least one VSS*

• Evaluation Problems
  – Different parsers train and evaluate against different annotation schemes

• Selecting one alternative over the other in a VSS can affect the performance of a specific parser**

* Schwartz et al., ACL 2011
** Nilsson et al., ACL 2006
Performance Differences

• Learning the correct annotation for a VSS is easy
  – Usually the direction of a single edge

the$_{DT}$ $\leftrightarrow$ apple$_{NN}$
Performance Differences

• Learning the correct annotation for a VSS is easy
  – Usually the direction of a single edge

• An annotation scheme is learned as a whole
  – Not each VSS alone

• There are second order effects
  – The way in which the VSS attaches to the rest of the sentence
  – These can lead to performance differences

the\textsuperscript{DT} \iff apple\textsuperscript{NN}
Example
Performance Differences

He_{PRP} \quad sure_{JJ} \quad about_{IN} \quad everyone_{NN}

(a)
Example
Performance Differences

(a)
Example

Performance Differences

(a)

(b)
Example
Performance Differences

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Example
Performance Differences

(a)

(b)

Definite Error
Example
Performance Differences

MST parser
(McDonald et al. 2005)
Example
Performance Differences

Predictable?

MST parser
(McDonald et al. 2005)
Example
Performance Differences

MST parser
(McDonald et al. 2005)

Predictable?
Parser specific?
Example
Performance Differences

MST parser
(McDonald et al. 2005)

Predictable?
Parser specific?
Magnitude?
Varying Syntactic Structures (VSS)
Varying Syntactic Structures (VSS)

Selecting one Annotation is Required
Varying Syntactic Structures (VSS)

Selecting one Annotation is Required

Selection Can Affect Parsing Performance
Our Solution: **Learnability**

- A straightforward selection criterion
  - Namely, **how easy it is to learn** a given annotation scheme using statistical tools

- Learnability is justified practically
  - Training on more learnable schemes results in more accurate parsers

- This criterion is only applied on linguistically **plausible** annotations
Learnability

• Learnability is widely used theoretically
  – Learnability using distributional methods has been used as an important consideration in designing the phrase structure formalism*

  – It is also used recurrently in cognitive science**

* Chomsky 2006
** Chater and Vitányi 2003, Perfors et al. 2011
Learnability-based Methodology 1
Learnability-based Methodology 1

Standard Parsing Evaluation

parser_1 \downarrow \quad \text{corpus} \quad \downarrow \quad \text{annotation scheme} \quad \downarrow \quad \text{result}_1

parser_2 \downarrow

parser_3 \downarrow

result_2

result_3
Learnability-based Methodology 1

Standard Parsing Evaluation

parser_1 -> parser_2 -> parser_3

\[ \text{corpus} \]

\[ \downarrow \]

\[ \text{annotation scheme} \]

\[ \downarrow \]

\[ \text{result}_1 \]

\[ \downarrow \]

\[ \text{result}_2 \]

\[ \downarrow \]

\[ \text{result}_3 \]

\[ \downarrow \]

\[ \text{parser}^* \]
Learnability-based Methodology 1

**Standard Parsing Evaluation**

- $\text{parser}_1$
- $\text{parser}_2$
- $\text{parser}_3$
- $\text{corpus}$
- $\text{annotation scheme}$
- $\text{result}_1$
- $\text{result}_2$
- $\text{result}_3$

**Our Approach**

- $\text{parser}$
- $\text{corpus}$
- $\text{scheme}_1$
- $\text{scheme}_2$
- $\text{scheme}_3$
- $\text{result}_1$
- $\text{result}_2$
- $\text{result}_3$
Learnability-based Methodology 1

Standard Parsing Evaluation

\( \text{parser}_1 \downarrow \text{parser}_2 \downarrow \text{parser}_3 \downarrow \text{corpus} \downarrow \text{annotation scheme} \downarrow \text{result}_1 \downarrow \text{result}_2 \downarrow \text{result}_3 \)

Our Approach

\( \text{parser} \downarrow \text{corpus} \downarrow \text{scheme}_1 \downarrow \text{scheme}_2 \downarrow \text{scheme}_3 \downarrow \text{result}_1 \downarrow \text{result}_2 \downarrow \text{result}_3 \)

**Learnability-based Syntactic Annotation Design** @ Schwartz, Abend and Rappoport, COLING 2012
Learnability-based Methodology 2

\[
\begin{align*}
\text{parser}_1 & \rightarrow \text{corpus} \\
\downarrow & \downarrow \\
\text{scheme}_1 & \rightarrow \text{result}_1 \\
\downarrow & \\
\text{scheme}_2 & \rightarrow \text{result}_2 \\
\downarrow & \\
\text{scheme}_3 & \rightarrow \text{result}_3
\end{align*}
\]
Learnability-based Methodology 2

parser₁

→
corpus

→
scheme₁

→
result₁

→
scheme₂

→
result₂

→
scheme₃

→
result₃

→
scheme₁*

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Learnability-based Methodology 2

parser_1

corpus

scheme_1  scheme_2  scheme_3

result_1  result_2  result_3

scheme_1*

parser_2

corpus

scheme_1  scheme_2  scheme_3

result_1  result_2  result_3

scheme_2*
Learnability-based Methodology 2

parser\textsubscript{1} → corpus → scheme\textsubscript{1}, scheme\textsubscript{2}, scheme\textsubscript{3} → result\textsubscript{1}, result\textsubscript{2}, result\textsubscript{3} → scheme\textsubscript{1}\textsuperscript{*}

parser\textsubscript{2} → corpus → scheme\textsubscript{1}, scheme\textsubscript{2}, scheme\textsubscript{3} → result\textsubscript{1}, result\textsubscript{2}, result\textsubscript{3} → scheme\textsubscript{2}\textsuperscript{*}

parser\textsubscript{3} → corpus → scheme\textsubscript{1}, scheme\textsubscript{2}, scheme\textsubscript{3} → result\textsubscript{1}, result\textsubscript{2}, result\textsubscript{3} → scheme\textsubscript{3}\textsuperscript{*}

scheme\textsuperscript{*}

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
Experimental Setup 1

- 6 VSSs, $2^6 = 64$ annotation schemes

(a) Prepositional Phrases

(b) Noun Phrases

(c) Coordinations

(d) Verb Groups

(e) Noun Sequences

(f) Infinitive Verbs
Experimental Setup 2

• 5 parsers of different types
  – Graph based parsers
    • MST parser (McDonald et al. 2005)
    • DMV parser (Klein and Manning 2004)
  – Transition based parsers
    • $S_u$ parser (Nivre 2009)
    • Clear parser (Choi and Nicolov 2009)
  – Other
    • NonDir Parser (Goldberg and Elhadad 2010)
Results

• In 3/6 structures, a **unanimously** more learnable annotation was found:
  
  – **Prepositions** (and not NPs) as heads of **PPs**

  \[ \text{about}_{\text{IN}} \rightarrow \text{everyone}_{\text{NN}} \]

  – **Nouns** (and not their determiners) as heads of **NPs**

  \[ \text{the}_{\text{DT}} \leftarrow \text{apple}_{\text{NN}} \]

  – **Conjunctions** as heads of **coordination** structures

\[ \text{John}_{\text{NNP}} \rightarrow \text{and}_{\text{CC}} \rightarrow \text{Mary}_{\text{NNP}} \]
Results

• In 3/6 structures, a **unanimously** more learnable annotation was found:
  - **Prepositions** (and not NPs) as heads of **PPs**
    
    ![Diagram](about_IN_everyone_NN)

  - **Nouns** (and not their determiners) as heads of **NPs**
    
    ![Diagram](the_DT_apple_NN)

  - **Conjuncts** as heads of **coordination** structures
    
    ![Diagram](John_NNP_and_CC_Mary_NNP)

**Consistently** more learnable, **independently** of the **parser** and the other **VSSs**
Magnitude

• Gains are substantial
  – Up to 19.8% error reduction for a single structure
Magnitude

- Gains are substantial
  - Up to **19.8%** error reduction for a single structure

- Gains are additive
  - Selecting the more learnable annotation in all 3 VSSs results in an **even more learnable** scheme
  - Up to **35.3%** error reduction by selecting the most vs. least learnable annotation
And the Winner is...

(a) Coordination
(b) Infinitive Verbs
(c) Noun Phrases
(d) Noun Sequence
(e) Prepositional Phrases
(f) Verb Groups

Available @ http://www.cs.huji.ac.il/~roys02/
Additional Experiments

• Two learnability measures

• High agreement between different parsers

• *Predictability* – a simple information-theoretic measure that yields similar results to learnability
Additional Experiments

• Two learnability measures

• High agreement between different parsers

• *Predictability* – a simple information-theoretic measure that yields similar results to learnability
Some more Ideas

• Apply our methodology to different tasks
  – POS tagging, Phrase Structure parsing, etc.

• Apply our methodology to different languages
  – Ballesteros and Nivre, CL 2013
What does it all Mean?

- Powerful results
  - Some annotations are clearly more learnable than others
What does it all Mean?

• Powerful results
  – Some annotations are clearly more learnable than others

• Linguistic implications?
What does it all Mean?

• Powerful results
  – Some annotations are clearly more learnable than others

• Linguistic implications?

• Cognitive implications?
Summary

• Varying Syntactic Structures (VSS)
  – Sometimes you have to choose

• Learnability as a selection criterion
  – A principled learnability-based methodology
  – Use parsers as research tools

• Selecting one alternative has a substantial and predictable effect on parsing performance
  – Parser independent
  – up to 35.3% error reduction
Summary

• Varying Syntactic Structures (VSS)
  – Sometimes you have to choose

• Learnability as a selection criterion
  – A principled learnability-based methodology
  – Use parsers as research tools

• Selecting one alternative has a substantial and predictable effect on parsing performance
  – Parser independent
  – up to 35.3% error reduction

Direct implications for designing more learnable annotation schemes

Learnability-based Syntactic Annotation Design @ Schwartz, Abend and Rappoport, COLING 2012
• Richard Johansson
• Roi Reichart
• Shai Shalev-Shwartz
• Valentin I. Spitkovsky
roys02@cs.huji.ac.il
http://www.cs.huji.ac.il/~roys02/