Uptraining for Accurate Deterministic Question Parsing

Slav Petrov, Pi-Chuan Chang, Michael Ringgaard, Hiyan Alshawi
Let’s build a Question Answering System
Let’s build a Question Answering System

- Need a parser
Let’s build a Question Answering System

- **fast**
- Need a parser
Let’s build a Question Answering System

- Need a parser
- Model: Deterministic Shift-Reduce Parser (Nivre et al.)
Let’s build a Question Answering System

- Need a parser
- Model: Deterministic Shift-Reduce Parser (Nivre et al.)
- Train on Wall Street Journal section of Penn Treebank
Let’s build a Question Answering System

- Need a parser
- Model: Deterministic Shift-Reduce Parser (Nivre et al.)
- Train on Wall Street Journal section of Penn Treebank

```
What does the Peugeot company manufacture?
```

```
Stanford Dependencies

ROOT What does the Peugeot company manufacture ?
```

```
SBARQ

WHNP
WP VBZ NP VP

What does the Peugeot company manufacture ?
```

```
Stanford Dependencies

root dobj aux det nn nsubj p

What does the Peugeot company manufacture ?
```
Let’s build a Question Answering System

- Need a parser
- Model: Deterministic Shift-Reduce Parser (Nivre et al.)
- Train on Wall Street Journal section of Penn Treebank
- Test on QuestionBank (Judge et al. ’06)
Unlabeled Attachment Score

WSJ: 88.24
Questions: 72.23
Unlabeled Attachment Score

-WSJ: 88.24
-Questions: 72.23

-16%
What happened?

**INPUT:**

<table>
<thead>
<tr>
<th>He</th>
<th>ate</th>
<th>the</th>
<th>green</th>
<th>apple</th>
</tr>
</thead>
</table>

**STACK:**

<table>
<thead>
<tr>
<th>green</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
</tr>
</tbody>
</table>

**OUTPUT:**

- \( nsubj \) 

<table>
<thead>
<tr>
<th>He</th>
<th>ate</th>
<th>...</th>
</tr>
</thead>
</table>

What happened?

- Single left-to-right pass
What happened?

INPUT:

What did John eat yesterday?

OUTPUT:

• Single left-to-right pass
What happened?

- Single left-to-right pass
- Only 334 questions in the training set (out of 39,832)

**INPUT:**

| What did John eat yesterday? |

**OUTPUT:**
Unlabeled Attachment Score

- Test on WSJ
- Test on Questions
Unlabeled Attachment Score

Our shift-reduce parser

Nivre et al. '07
McDonald et al. '06
Charniak '00
Charniak & Johnson '05
Petrov et al. '06
Petrov '10

Test on WSJ
Test on Questions
Our shift-reduce parser
Nivre et al. '07
McDonald et al. '06
Charniak '00
Charniak & Johnson '05
Petrov et al. '06
Petrov '10

Test on WSJ
Test on Questions
Unlabeled Attachment Score

- Test on WSJ
- Test on Questions

- Our shift-reduce parser
- Nivre et al. '07
- McDonald et al. '06
- Charniak '00
- Charniak & Johnson '05
- Petrov et al. '06
- Petrov '10
Our shift-reduce parser
Nivre et al. '07
McDonald et al. '06
Charniak '00
Charniak & Johnson '05
Petrov et al. '06
Petrov '10

Test on WSJ
Test on Questions
Unlabeled Attachment Score

- Test on WSJ
- Test on Questions

Chart showing the Unlabeled Attachment Score for various parser models:
- Our shift-reduce parser
- Nivre et al. '07
- McDonald et al. '06
- Charniak '00
- Charniak & Johnson '05
- Petrov et al. '06
- Petrov '10
Unlabeled Attachment Score

Test on WSJ
Test on Questions
Unlabeled Attachment Score

- Test on WSJ
- Test on Questions

<table>
<thead>
<tr>
<th>Method</th>
<th>Test on WSJ</th>
<th>Test on Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our shift-reduce parser</td>
<td>85</td>
<td>65</td>
</tr>
<tr>
<td>Nivre et al. '07</td>
<td>80</td>
<td>55</td>
</tr>
<tr>
<td>McDonald et al. '06</td>
<td>90</td>
<td>70</td>
</tr>
<tr>
<td>Charniak '00</td>
<td>95</td>
<td>80</td>
</tr>
<tr>
<td>Charniak &amp; Johnson '05</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>Petrov et al. '06</td>
<td>95</td>
<td>75</td>
</tr>
<tr>
<td>Petrov '10</td>
<td>100</td>
<td>90</td>
</tr>
</tbody>
</table>
Uptraining
Uptraining

• Latent Variable Parser
  • generalizes best
  • too slow for our needs
Uptraining

- Latent Variable Parser
  - generalizes best
  - too slow for our needs

- Idea: Parse lots of unlabeled questions with Latent Variable Parser and use as (noisy) training data for shift-reduce parser.
Uptraining Results on Questions

- WSJ
- WSJ+QB
Uptunning Results on Questions

WSJ

WSJ+QB
Uptraining Results on Questions

Baseline (Questions)
Uptaining Results on Questions

Baseline (Questions)

- WSJ
- WSJ+QB
Uptaining Results on Questions

Baseline (Questions)

WSJ
WSJ+QB

2K gold questions
Uptraining Results on Questions

![Graph showing uptraining results on questions with lines for WSJ, WSJ+QB, and baselines.](image-url)
Thank you!

Come to the poster for more details!