The Impact of Source–Side Syntactic Reordering on Hierarchical Phrase-based SMT

Jinhua Du, Andy Way

CNGL, School of Computing, Dublin City University
Outline

Motivation

How does DE construction affect HPB?

Problem of Chinese DE Construction Translation in HPB

Stanford DE Classifier and DE Classifications

Experiments of DE reordering on HPB

Analysis: how DE reordering affects HPB?

Conclusions and Future Work
An Example: DE Construction in Chinese

Original: 澳洲是与北韩有邦交的少数国家之一。Aozhou shi yu Beihan you bangjiao DE shaoshu guojia zhiyi.
Australia is with North Korea have diplomatic relations that few countries one of.

Reference: Australia is [one of the few countries] that have diplomatic relations with North Korea.

Reordered: 澳洲是少数国家之一的与北韩有邦交。Aozhou shi yu Beihan you bangjiao DE shaoshu guojia zhiyi.
Australia is [one of the few countries] [have diplomatic relations with North Korea].

Literal Translation: Australia is [one of the few countries] [have diplomatic relations with North Korea]．

Hierarchical Rule:

<X₁ DE X₂, X₂ that X₁>

This reordering rule generalises the reordering of relative clauses and the DE structure.
Motivation:

- Previous Research on DE Constructions
  - Wang et al. (2007) proposed a set of syntactic rules to reorder NP with DE’s with certain patterns. They experimented using PB-SMT.
  - Chang et al. (2009) proposed a log-linear classifier to classify DE constructions. They also experimented using PB-SMT.

- We are considering that
  - All the previous work were applied using PB-SMT system;
  - The phrases in PB-SMT don’t have a reordering capability themselves;
  - Some phrases/rules in a hierarchical PBSMT (HPB) can perform reordering when decoding.

what would happen if we applied source-side reordering in HPB?
Outline

- Motivation
  - How does DE construction affect HPB?
- Problem of Chinese DE Construction Translation in HPB
- Stanford DE Classifier and DE Classifications
- Experiments of DE reordering on HPB
- Analysis: how DE reordering affects HPB?
- Conclusions and Future Work
Problem of Chinese DE Construction Translation

• DE construction in Chinese is difficult to translate

  ✓ DE can cause a long distance dependency and reordering;

  ✓ DE construction is diverse in Chinese:

  ○ DE translation Errors caused by the HPB decoding

  Source: 当地一所[名声不佳]的[中学]

  Reference: [a] [local] [middle school] [with] [a bad reputation]

  Hyp1: [a bad reputation] [of] [the local] [secondary school]

  Hyp2: [the local] [a bad reputation] [secondary school]

  Hyp3: [a local] [stigma] [secondary schools]

Note: Originated from Chang’s (2009) paper
Our idea

Goal: Perform DE reordering on the Chinese side To facilitate the reordering in HPB decoding

How: Use DE classifier to recognise which should be reordered, and which should not.

Why: Analyse why DE reordering can affect the HPB system
Outline

- Motivation
  - How does DE construction affect HPB?
- Problem of Chinese DE Construction Translation in HPB
- Stanford DE Classifier and DE Classifications
- Experiments of DE reordering on HPB
- Analysis: how DE reordering affects HPB?
- Conclusions and Future Work
Diversity of DE construction in Chinese

A DE B ➔

- A B 他 的 书 ➔ his book
- A’s B 彼得 的 家 ➔ Peter’s home
- B A 桌上 的 电脑 ➔ the laptop on the desk
- B of A 彼得 的 朋友 ➔ friends of Peter
- B that A 有邦交 的 国家 ➔ countries that have diplomatic relations
- B with A 名声不佳 的 中学 ➔ school with a bad reputation
Three non-reordered classes:

- **A B → label: DE_{AB}**
  - A in the Chinese side is translated as a pre-modifier of B. In most cases A is an adjectival form.

- **A ’s B → label: DE_{AsB}**
  - the English translation is an explicits-genitive case.

- **A prep. B → label: DE_{AprepB}**
  - The English translations that fall into this class usually have some number, percentage or level word in the Chinese A.
Two reordered classes

- **B prep. A** → label: **DE_{B{\text{prep}}A}**
  - A and B in English side are reordered around the preposition such as ‘of’.

- **Relative clause** → label: **DE_{\text{relc}}**
  - The relative clause would be introduced by a relative pronoun or be a reduced relative clause.
Stanford DE Classifier

- Training data to label DEs in NPs from:
  - CTB6
  - English-Chinese Translation Treebank
  - the manual word alignment data (LDC2006E93)

- Log-linear Model

- Features:
  - **DEPOS**: the part-of-speech tag of DE;
  - **A-pattern**: indicates the Chinese syntactic patterns appearing before DE;
  - **SemClass**: represents the semantic class of words in a Chinese thesaurus “CILIN”;
  - **Topicality**: denotes the re-occurrence of nouns in the contexts.
  - **Lexical features**: word suffixes rather than word

Note: refer to (Chang et al., 2009) for details
Outline

- Motivation
  - How does DE construction affect HPB?
- Problem of Chinese DE Construction Translation in HPB
- Stanford DE Classifier and DE Classifications
- Experiments of DE Annotation and Reordering for HPB
- Analysis: how DE reordering affects HPB?
- Conclusions and Future Work
Experimental Settings

- **HPB System:** In-house re-implementation of HPB

- **Training data:**
  - Translation direction: Chinese-to-English
  - Amount of pairs: 2,159,232
  - Sources: HK corpus, ISI corpus, part of UN data and other News data from LDC
  - Language model: English part of the parallel data; 5-gram

- **Development set:**
  - NIST 2006 test set: 1,664 sentences
  - References: 4 references for each source sentence

- **Test set:**
  - NIST 2008 test set: 1,357 sentences
  - References: 4 references for each source sentence
DE Annotation and Reordering

Steps:

1. use the Stanford Chinese parser to parse the Chinese side of the MT training data and the devset and test set.
2. use the DE classifier to annotate the DE constructions in NPs in all of the parsed data.
3. pre-process the Chinese data by reordering the sentences only with \texttt{DE}_{B\text{prep}A} and \texttt{DE}_{r\text{elc}} annotations.
4. Continue regular SMT steps….
### Statistics of 5-class DE Annotation

<table>
<thead>
<tr>
<th>DE-class</th>
<th>Training</th>
<th>Devset</th>
<th>Testset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>%</td>
<td>Count</td>
</tr>
<tr>
<td>DE_{AB}</td>
<td>312,590</td>
<td>23.07</td>
<td>544</td>
</tr>
<tr>
<td>DE_{AprepB}</td>
<td>6,953</td>
<td>0.51</td>
<td>9</td>
</tr>
<tr>
<td>DE_{AsB}</td>
<td>13,105</td>
<td>0.97</td>
<td>21</td>
</tr>
<tr>
<td>DE_{BprepA}</td>
<td>658,692</td>
<td>48.62</td>
<td>974</td>
</tr>
<tr>
<td>DE_{relc}</td>
<td>316,675</td>
<td>23.37</td>
<td>392</td>
</tr>
<tr>
<td>DE_{non}</td>
<td>46,752</td>
<td>3.45</td>
<td>87</td>
</tr>
<tr>
<td>Total</td>
<td>1,354,767 100</td>
<td>2027</td>
<td>100</td>
</tr>
<tr>
<td>Average</td>
<td>--</td>
<td>--</td>
<td>1.22</td>
</tr>
</tbody>
</table>

- the reordered DE constructions of $\text{DE}_{BprepA}$ and $\text{DE}_{relc}$ account for 71.99%, 67.39% and 62.83% of the total DE constructions in the training data, devset and test set respectively.
Experimental Results

<table>
<thead>
<tr>
<th>Metrics</th>
<th>HPB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>BLEU</td>
<td>20.53</td>
</tr>
<tr>
<td>METEOR</td>
<td>42.91</td>
</tr>
<tr>
<td>TER</td>
<td>62.81</td>
</tr>
<tr>
<td>Improvements</td>
<td>8.91/2.73/1.45</td>
</tr>
</tbody>
</table>

- Baseline: run HPB on non-reordered training, dev and test data;
- DE reordering: run HPB on DE-reordered training, dev and test data;
Outline

- Motivation
  - How does DE construction affect HPB?
- Problem of Chinese DE Construction Translation in HPB
- Stanford DE Classifier and DE Classifications
- Experiments of DE reordering on HPB

Analysis: how DE reordering affects HPB?

Conclusions and Future Work
Case 1: DE$_{B\text{prep}A}$ Construction

- one of two annotations which would have to be reordered in Chinese sentences.

- This annotation indicates that the Chinese “A” is translated into a prepositional phrase and reordered with “B” in English side.
## An Example of DE\textsubscript{BprepA} Construction

<table>
<thead>
<tr>
<th>Chinese</th>
<th>据 [临汾市 中院 刑庭] 的 [工作 人员] 透露</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. 1</td>
<td>[the staff] [at] [the linfen intermediate criminal court] revealed</td>
</tr>
<tr>
<td>Ref. 2</td>
<td>[a staff member] [at] [the criminal court of linfen intermediate court] disclosed</td>
</tr>
<tr>
<td>Ref. 3</td>
<td>[a staff member] [at] [the criminal court of the linfen intermediate people ’s court] disclosed</td>
</tr>
<tr>
<td>Ref. 4</td>
<td>[the staff] [of] [linfen city intermediate people ’s criminal court] revealed</td>
</tr>
<tr>
<td>BAS-HPB</td>
<td>[linfen intermediate court of the criminal court] [of] [staff] revealed</td>
</tr>
<tr>
<td>DE-HPB</td>
<td>[according to the staff] [of] [the lifen intermediate criminal court] revealed</td>
</tr>
</tbody>
</table>
Process of SCFG derivations for DE_BprepA

(a) without DE annotation

(b) with DE annotation
Case 2: DE_{relc} Construction

- another annotation which requires reordering of the Chinese sentences.

- This annotation indicates that the Chinese “A” is translated into a relative clause.
<table>
<thead>
<tr>
<th>Chinese</th>
<th>我 有 太 多 [现在 无 法 做] A 的 [事情] B 等待 着 我 去 做 。</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. 1</td>
<td>i have too many [things] B waiting for me to do that[i can ’t do at the moment] A .</td>
</tr>
<tr>
<td>Ref. 2</td>
<td>i have too many [things] B which[i cannot do now] A waiting for me to do .</td>
</tr>
<tr>
<td>Ref. 3</td>
<td>i have too many [things] B waiting for me that[i cannot do now] A .</td>
</tr>
<tr>
<td>Ref. 4</td>
<td>i have too many [things] B that[i cannot do now] A waiting for me to do .</td>
</tr>
<tr>
<td>BAS-HPB</td>
<td>i have too much to do things in the waiting for me to do .</td>
</tr>
<tr>
<td>DE-HPB</td>
<td>i have too many [things] B that[couldn ’t make] A waiting for me to do .</td>
</tr>
</tbody>
</table>
Process of SCFG derivations for DE\textsubscript{relc}

(a) without DE annotation

(b) with DE annotation
Influence of DE Annotation on Word Alignment

- Word alignment has a significant impact on the phrase extraction and probability calculation.

- For the HPB system, the initial phrase table and the word alignment links in the phrase pair will impact on the hierarchical phrase generation.
An Example of the Influence of DE Annotation on Word Alignment

with the continuous structural developments in the economy of Hong Kong and the Mainland of China

(a) the word alignment without DE annotation $B_{prepA}$

with the continuous structural developments in the economy of Hong Kong and the Mainland of China

(b) the word alignment with DE annotation $B_{prepA}$
Influence of DE Annotation on Phrase Extraction

<table>
<thead>
<tr>
<th>without DE</th>
<th>with DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>经济 体系 的 \textit{B}_{prepA} 香港</td>
</tr>
<tr>
<td>经济 体系</td>
<td></td>
</tr>
<tr>
<td>香港</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>\textit{X1} 的 \textit{B}_{prepA} X2</td>
</tr>
</tbody>
</table>

Comparison: Initial phrases and SCFG rules extracted from both “non-annotated DE” and “annotated DE” alignment.
Analysis on Altered Sentences by Reordering DE Constructions

Statistics:
- In the NIST2008 testset, there are 839 out of 1357 sentences (61.8%) that have DEs under NPs;
- There are 664 out of 839 sentences (79.1%) that have \(B_{prepA}\) or \(relc\) labels and are reordered.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>HPB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>BLEU</td>
<td>19.77</td>
</tr>
<tr>
<td>METEOR</td>
<td>42.99</td>
</tr>
<tr>
<td>TER</td>
<td>64.44</td>
</tr>
<tr>
<td>Improvements</td>
<td></td>
</tr>
</tbody>
</table>

Experimental results of altered sentences on HPB systems
Summary of the Analysis

- The reordered data has an influence on the word alignment and phrase extraction.

- The DE reordering approach can reduce the influence of the poor generalisation capability of HPB rules.

- The reordered hierarchical rules have a better generalisation capability.
Outline

- Motivation
  - How does DE construction affect HPB?
- Problem of Chinese DE Construction Translation in HPB
- Stanford DE Classifier and DE Classifications
- Experiments of DE reordering on HPB
- Analysis: how DE reordering affects HPB?
- Conclusions and Future Work
Conclusions

- Employ the Stanford DE annotated approach into the hierarchical phrase-based system to verify whether it works or not;

- Experimental results show that
  - Syntactic reordering on the source-side has a significant influence on HPB system;
  - Specifically, the DE construction reordering affected the word alignment, phrase extraction and rule generalisation phrases in building a HPB system.
Future Work

- We plan to carry out a larger scale experiments on the HPB system and verify the consistency of the improvements.

- We plan to apply the DE-annotated approach into a syntax-based MT system and examine the effects.

- We also intend to improve the classification accuracy of the DE classifier as well as investigate other grammatical reordering constructions to further improve the translation quality.

- Our re-implemented HPB doesn’t work better than Moses. In future, we will use Moses Chart decoder to verify the impact.
Acknowledgment

Thanks very much to Dr. Pichuan Chang for providing us the DE classifier, training data so that we can carry out this work.
Thanks for your attention!

Questions?