



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 X_1	有	邦交]A	的	[少数	国家 X_2	之一]B	。
	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]	X_2	that	[have diplomatic relations	X_1	with	North Korea]	.	
Reordered:	澳洲	是	[少数国家之一]B	的	[与北韩有邦交]A	。					
Literal Translation:	Australia	is	[one of the few countries]	[have diplomatic relations with North Korea]	.						

An example of DE construction reordering (extended from the original figure in (Chiang, 2005))

Hierarchical Rule:

$\langle X_1 \text{ DE } X_2, X_2 \text{ that } X_1 \rangle$

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]_A 的 [中学]_B

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]

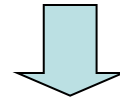
Hyp3: [a local] [stigma] [secondary schools]

Note: Originated from Chang's (2009) paper

Our idea

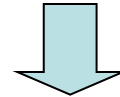
Goal:

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

same order

● 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

reordered

● B with A 名声不佳的中学 school with a bad reputation

5 Classes of DE Construction (Chang et al., 2009)

- 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 explicit-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.

5 Classes of DE Construction (Cont.)

● Two reordered classes

● *B prep. A* → label: **DE**_{-BprepA}

- A and B in English side are reordered around the preposition such as 'of'.

● *Relative clause* → label: **DE**_{-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 $DE_{-BprepA}$ and DE_{-relc} annotations.
4. Continue regular SMT steps.....

Statistics of 5-class DE Annotation

DE-class	Training		Devset		Testset	
	Count	%	Count	%	Count	%
<i>DE_AB</i>	312,590	23.07	544	26.84	464	29.50
<i>DE_AprepB</i>	6,953	0.51	9	0.44	7	0.44
<i>DE_AsB</i>	13,105	0.97	21	1.04	11	0.70
<i>DE_BprepA</i>	658,692	48.62	974	48.05	663	42.12
<i>DE_relc</i>	316,675	23.37	392	19.34	326	20.71
<i>DE_non</i>	46,752	3.45	87	4.29	103	6.54
Total	1,354,767	100	2027	100	1574	100
Average	--	--	1.22	--	1.16	--

- the reordered DE constructions of *DE_BprepA* and *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

Metrics	HPB	
	Baseline	DE reordered
BLEU	20.53	22.36
METEOR	42.91	44.08
TER	62.81	61.90
Improvements	8.91/2.73/1.45	

- 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

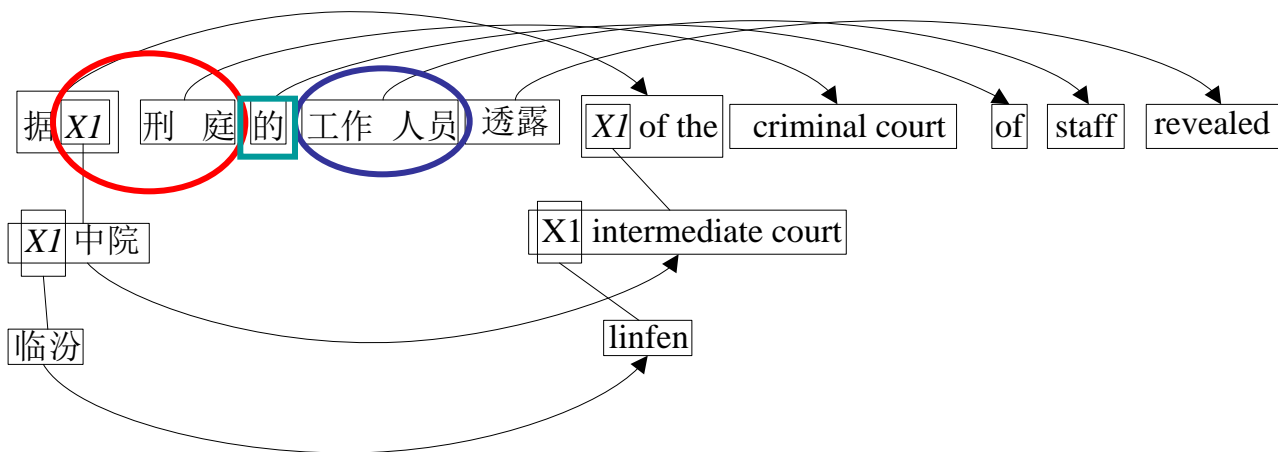
Case 1: DE_{—BprepA} Construction

- 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.

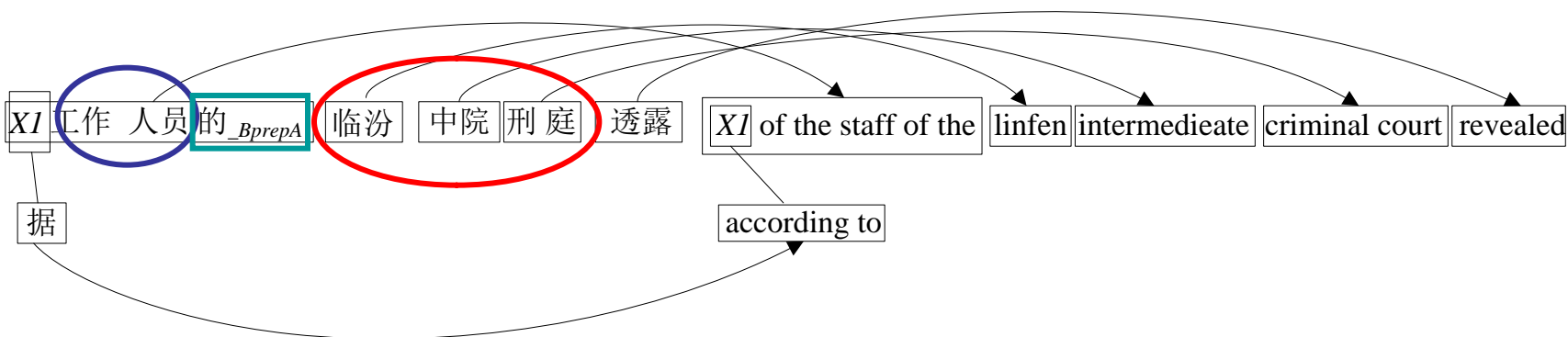
An Example of DE_{-BprepA} Construction

Chinese	据 [临汾市 中院 刑庭] _A 的 [工作 人员] _B 透露
Ref. 1	[the staff] _B at [the linfen intermediate criminal court] _A revealed
Ref. 2	[a staff member] _B at [the criminal court of linfen intermediate court] _A disclosed
Ref. 3	[a staff member] _B at [the criminal court of the linfen intermediate people 's court] _A disclosed
Ref. 4	[the staff] _B of [linfen city intermediate people 's criminal court] _A revealed
BAS-HPB	[linfen intermediate court of the criminal court] _A of [staff] _B revealed
DE-HPB	[<u>according to the staff</u>] _B of [<u>the lifen intermediate criminal court</u>] _A revealed

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.

An Example of DE_{-relc} Construction

Chinese 我有太多 [现在无法 **做**_A 的 [事情]_B 等待着我去做。

Ref. 1 i have too many [things]_B waiting for me to do **that** [i can't do at the moment]_A .

Ref. 2 i have too many [things]_B **which** [i cannot do now]_A waiting for me to do .

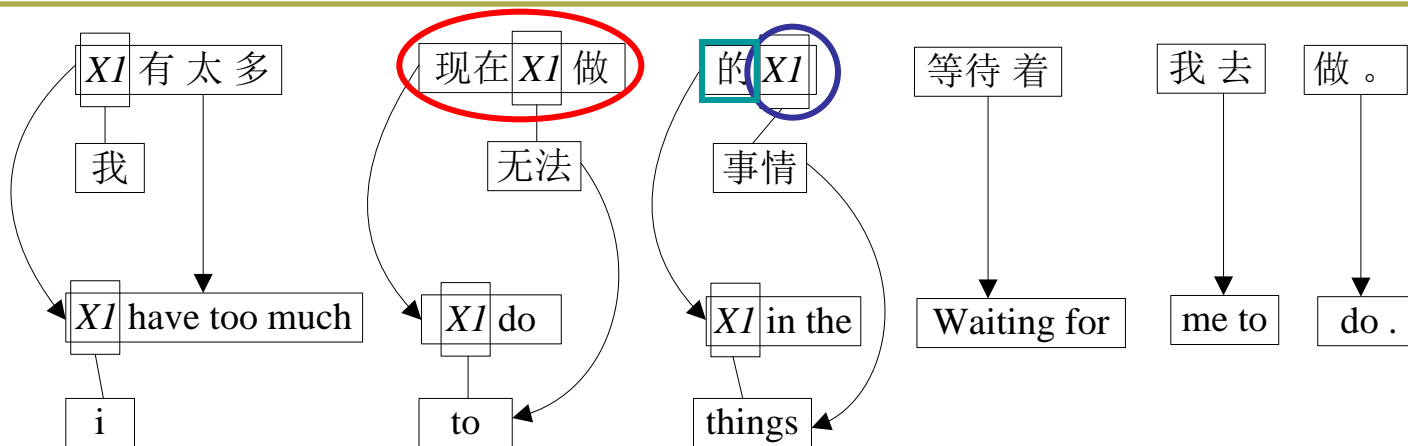
Ref. 3 i have too many [things]_B waiting for me **that** [i cannot do now]_A .

Ref. 4 i have too many [things]_B **that** [i cannot do now]_A waiting for me to do .

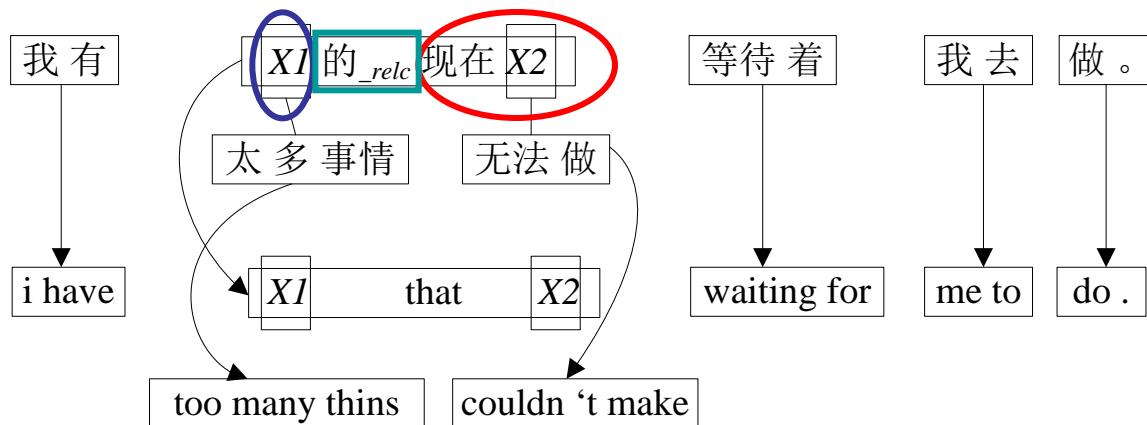
BAS-HPB i have too much to do things in the waiting for me to do .

DE-HPB i have too many [things]_B **that** [couldn't make]_A waiting for me to do .

Process of SCFG derivations for DE_{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
基于香港与中国内地~~的~~经济体系在结构上持续发展
economy of hong kong and the mainland of china
(a) the word alignment without DE annotation *BprepA*

with the continuous structural developments in the
基于经济体系的~~的~~*BprepA*香港与中国内地在结构上持续发展
economy of hong kong and the mainland of china
(b) the word alignment with DE annotation *BprepA*

Influence of DE Annotation on Phrase Extraction

without DE	with DE
–	<u>经济体系的</u> _{BprepA} <u>香港</u> <u>economy of hong kong</u>
经济体系 economy	<u>经济体系</u> <u>economy</u>
香港 hong kong	<u>香港</u> <u>hong kong</u>
–	<u>X1 的</u> _{BprepA} <u>X2</u> <u>X1 of X2</u>

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 *BprepA* or *relc* labels and are reordered.

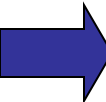
Metrics	HPB	
	Baseline	DE reordering
BLEU	19.77	22.47
METEOR	42.99	44.48
TER	64.44	62.86
Improvements	13.66/3.47/2.45	

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?

