



Improving Word Alignment Using Syntactic Dependencies

Yanjun Ma Sylwia Ozdowska
Yanli Sun Andy Way

National Centre for Language Technology
School of Computing
Dublin City University
Dublin, Ireland



Outline

Context

Syntax-enhanced Model

Experiments

- Experimental Setup

- Intrinsic Evaluation

- Extrinsic Evaluation

Conclusions and Future Work



Outline

Context

Syntax-enhanced Model

Experiments

Experimental Setup

Intrinsic Evaluation

Extrinsic Evaluation

Conclusions and Future Work



Word Alignment

Determining translational correspondences at word level given a parallel corpus of aligned sentences

- Generative word alignment models ([Brown et al., 1993](#); [Vogel et al., 1996](#); [Deng & Byrne, 2005](#); [Liang et al., 2006](#); [Ganchev et al., 2008](#))
- ITG ([Wu, 1997](#))
- Heuristics-based word alignment models ([Melamed, 2000](#); [Tiedemann, 2003](#))
- Discriminative word alignment models ([Liu et al., 2005](#); [Moore, 2005](#); [Taskar et al., 2005](#); [Ittycheriah and Roukos, 2005](#); [Blunsom and Cohn, 2006](#); [Moore et al., 2006](#); [Lacoste-Julien et al., 2006](#); [Cherry and Lin, 2006](#); [Fraser and Marcu, 2007](#); [Niehues and Vogel, 2008](#))



Question 1: Can Syntax Help Word Alignment ?

Syntax	Generative	Heuristics	Discriminative
POS tags	-	1	2
Chunking	-	3	-
Parsing	4	5	6

¹ Tiedemann, 2003;

² Liu et al., 2005; Blunsom and Cohn, 2006;

³ Sun et al., 2000; Ren et al., 2007; [Crego and Habash, 2008](#)

⁴ Wang and Zhou, 2004; Lopez and Resnik, 2005; DeNero and Clein, 2007; May and Knight, 2007

⁵ Ayan et al., 2004; Ozdowska, 2004

⁶ Cherry and Lin, 2006; [Fossum et al., 2008](#); [this work](#)



Question 2: $AER_{\downarrow} = MT_{\uparrow}$? If NOT, Why ?

ACL 2006 T-shirt (cf. Daniel Marcu's invited talk)

- Don't ever ever use AER !

Only Fraser and Marcu (2007) and Ganchev et al. (2008) showed improved word alignment improves MT

The correlation between word alignment quality and MT quality becomes a bottleneck

- $AER_{\downarrow} = MT_{\uparrow}$ iff %\$#%%&...



Question 1: Can syntax help word alignment ?

Question 2: $AER_{\downarrow} = MT_{\uparrow}$? If NOT, why ?



Outline

Context

Syntax-enhanced Model

Experiments

Experimental Setup

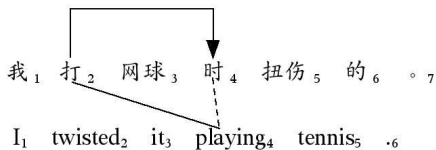
Intrinsic Evaluation

Extrinsic Evaluation

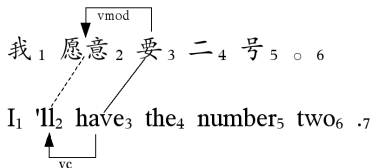
Conclusions and Future Work

Two Cases Where Syntax Helps

1-to- n word alignment



Cross-lingual syntactic correspondence

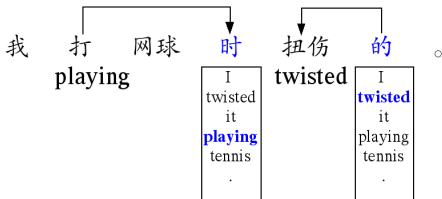


Main Idea

Acquiring a set of fixed links



Aligning the remaining words incorporating dependency information





General model

$$\begin{aligned} p(A|c_1^J, e_1^I) &= \prod_{j=0}^J p(a_j|c_1^J, e_1^I, a_1^{j-1}) \\ &= \frac{1}{Z} \cdot p_\epsilon(A_\Delta|c_1^J, e_1^I) \cdot \\ &\quad \prod_{j \in \bar{\Delta}} p(a_j|c_1^J, e_1^I, a_1^{j-1}, A_\Delta) \end{aligned}$$

- Decomposed into an **anchor alignment model** and a **syntax-enhanced model**
- Assuming the anchor alignment model is available, the syntax-enhanced model can be estimated using discriminative training



Anchor alignment model

- Existing word aligners: GIZA++ intersection

(Local) syntax-enhanced model

$$\begin{aligned}\hat{a}_j &= \operatorname{argmax}_{a_j} \{p_{\lambda_1^M}(a_j | c_1^J, e_1^I, a_1^{j-1}, A_\Delta)\} \\ &= \operatorname{argmax}_{a_j} \left\{ \sum_{m=1}^M \lambda_m h_m(c_1^J, e_1^I, a_1^j, A_\Delta, T_c, T_e) \right\}\end{aligned}$$

where A_Δ is the anchor alignment, T_c and T_e are the dependency trees for Chinese and English sentence pairs



Feature functions

- Statistics-based features: word association measures
- Distortion feature: relative position changes
- Syntactic features: dependency agreement features, features encoding source/target word dependencies, etc.

Discriminative training

- Various machine learning techniques can be applied, e.g. SVM

Search

- From all the candidate links, we select 1) the link with highest score; 2) the link score is greater than threshold t
- t can be optimised on the development set



Outline

Context

Syntax-enhanced Model

Experiments

Experimental Setup

Intrinsic Evaluation

Extrinsic Evaluation

Conclusions and Future Work



Experimental Setting

Evaluation

- AER and MT

Word alignment

- Gold standard: IWSLT development set 3 (502 sentence pairs) annotated by two linguists based on existing word alignment guidelines; 300 training, 50 development and 152 testing
- Dependency parser: Malt dependency parser (Nivre et al., 2007), with 88% labelled attachment scores for English and 84% for Chinese

Machine translation

- IWSLT 2007 data: 40K sentence pairs for training and 489 for testing
- Baseline system: Moses



Intrinsic evaluation: anchor alignment

anchor model	precision	recall	f-measure	AER
Heuristics	0.9774	0.4047	0.5724	0.3947
Model 1	0.9509	0.5011	0.6563	0.3157
HMM	0.9802	0.5327	0.6903	0.2809
Model 4	0.9777	0.5677	0.7179	0.2533

- High precision for all anchor alignment approaches
- High order IBM models can achieve both high precision and recall



Influence of anchor alignment on syntax-enhanced model

- Evaluate the **intersection of the words left unaligned** after anchoring using each of the anchor alignment models
- Heuristics-based and Model 4 anchoring



- Intersection

我₁ 愿意₂ 要₃ 二₄ 号₅ ○₆

I₁ ll₂ have₃ the₄ number₅ two₆ .₇



anchor model	precision	recall	f-score	AER
Heuristics	0.4505	0.3270	0.3790	0.6210
Model 1	0.5538	0.3894	0.4573	0.5427
HMM	0.5932	0.3611	0.4489	0.5511
Model 4	0.5660	0.4216	0.4832	0.5168

- Quality of anchor alignment has a significant influence on syntax-enhanced model



Overall performance

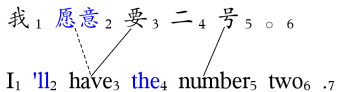
model	precision	recall	f-score	AER
Heuristics				
no syntax	0.8362	0.6751	0.7470	0.2302
w. syntax	0.8376	0.6894	0.7563	0.2240
Model 1				
no syntax	0.8759	0.6902	0.7720	0.2045
w. syntax	0.8542	0.7160	0.7790	0.2011
HMM				
no syntax	0.8655	0.7168	0.7841	0.1952
w. syntax	0.8744	0.7304	0.7959	0.1845
Model 4				
no syntax	0.8697	0.7340	0.7961	0.1832
w. syntax	0.8566	0.7685	0.8102	0.1730

- Syntax plays a positive role in all different anchor alignment configurations
- The influence of syntax grows proportionally to the strength of the anchor alignment model.

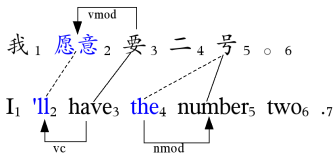


Analysis of Word Alignment

Word alignment without syntax



Syntax-enhanced word alignment





Weights of some informative features

	weight
Model 1 Score	0.1416
POS	0.0540
Log-likelihood Ratio	0.0856
relative distortion	0.0606
DA-1	0.0227
DLA-2	0.0927
tgt-1-PRD	0.0961
tgt-2-AMOD	0.0621

- IBM model 1 score is the most informative feature
- Some dependency types, e.g. adjective/adverb modifier (AMOD) and predicate (PRD) are most informative in alignment



Extrinsic evaluation: Chinese \rightarrow English MT

	dev. set	test set
Baseline	0.5412	0.3510
Syntax-HMM	0.6015	0.3409
Syntax-Model 4	0.5834	0.3585

- Development set: statistically significant improvement with both HMM-based syntax model ($p < 0.002$) and IBM4-based syntax model ($p < 0.008$).
- Test set: no significant improvement



Error Analysis of Translation

Fewer (better) links in word alignment → better coverage in translation ?

	No. word alignment	No. phrases
Baseline	463,491	309,183
Syntax-HMM	390,347	514,273
Syntax-Model 4	413,088	402,239

Chinese: 请再给我一份麦片。

Reference: Can I have another bowl of cereal ?

Baseline: please bring me a medical 麦片 .

Syntax-Model 4: please give me a oat-meal .



Syntax models perform better in fluency

- The baseline outputs tend to be redundant

Chinese: 请 给 我 两 张 二 楼 座 位 的 票 。

Glossary: please give me two second floor seat ticket

Reference: two seats on the upper deck , please .

Baseline: please give me two tickets on the second floor **seats , please .**

Syntax-Model4: please give me two seats on the second floor .



Poor fidelity

- Important information missing due to the noise in phrase table

Chinese: 我赌 **七 号** 马 五 美元 。

Glossary: I bet seven number horse five dollar

Reference: I would like to bet five dollars on horse **number seven**, please .

Baseline: i bet five dollars horse number seven .

Syntax-Model4: I |0-0| bet |1-1| five |5-5| dollars |6-6| **on the** |2-3| horse |4-4| . |7-7|



Outline

Context

Syntax-enhanced Model

Experiments

- Experimental Setup

- Intrinsic Evaluation

- Extrinsic Evaluation

Conclusions and Future Work



Question 1: Can syntax help word alignment ?

- Syntax is useful in word alignment
- Syntax is especially useful in improving recall

Question 2: $AER_{\downarrow} = MT_{\uparrow}$? If NOT, why ?

- To answer this is not trivial
- Some in-depth error analysis is crucial to answer this question



Future work

- Jointly optimise the two submodels
- Better search
- Compare our model with other discriminative word alignment models
- Go on to find answers to Question 2...



Thank you for listening

Our Chinese-English word alignment gold-standard can be downloaded freely from now.

Contact: yma@computing.dcu.ie

<http://www.computing.dcu.ie/~yma>

We would like to thank Prof.Rebecca Hwa (Pittsburgh) and Dr.Yang Liu (ICT) for providing us with the word alignment annotation guidelines.