



A Review of Word Alignment Techniques

— Presenters —

Guo Xuechun 2819****

Ke Huishu 2819****



Southeast University-Monash University Joint
Graduate School(Suzhou)
5/26/2017

Objectives



- There is no review about word alignment(CNKI, ACM Digital Library, Monash Library)
- Introduce the word alignment to you
- Show different models of word alignment techniques

CONTENTS



1

Introduction

Background, Applications, Difficulties

2

Models Explanation

IBM model, Neural model

3

Evaluation Rules

AER, F-measure, CPER, ASAER

4

Conclusions

Conclusion, challenges and future work



Introduction

Ke Huishu



1

Background

- ✓ Word alignment refers to identifying correspondences between the words/phrases in source language and in target language (Nguyen and Dinh, 2012)

□ 文化 是 维系 中华民族 生生不息、蔓延不断 精神纽带。

□ Culture is the spiritual bond that ensures the continuity of our Chinese nation.

What Is
Word Alignment

?

1

Applications

Machine Translation

The translation of text by a computer automatically, without human involvement.

Word Sense Disambiguation

Word is used in different conditions and has multiple meaning.

Dictionary Construction

Construct bilingual dictionary automatically and greatly improve the efficiency.

Text Retrieval

Text retrieval is defined as the matching of some stated user query against a set of free-text records



1

Difficulties

Incompatibility

► Complexity and Diversity

Different languages have much divergence on part of speech, lexicon table, semantics, syntax structure, etc.

Various type

► One-to-many

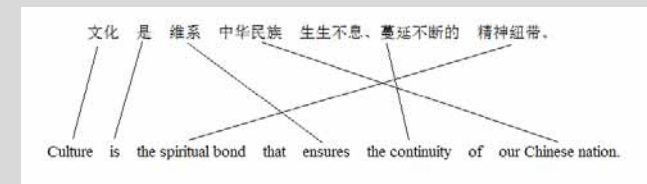
中国 —— China
食品业 —— food industry

► Non-contiguous words

对外开放
Opening...to the outside

Cross Alignment

► Word order and part of speech change





Models Explanation

Guo Xuechun



2

Different Approaches



- Statistical-based method
IBM model
- Syntactical-based method
- Innovative method
Neural model, Monte Carlo model

IBM model

● Nature

A series of five statistical models of the translation process (Brown et al., 1993)

● Prerequisite

A pair of sentences that are translations of one another

● Thought

- Assign a probability to each of the possible word-by-word alignments
- Seek the most probable of these alignments

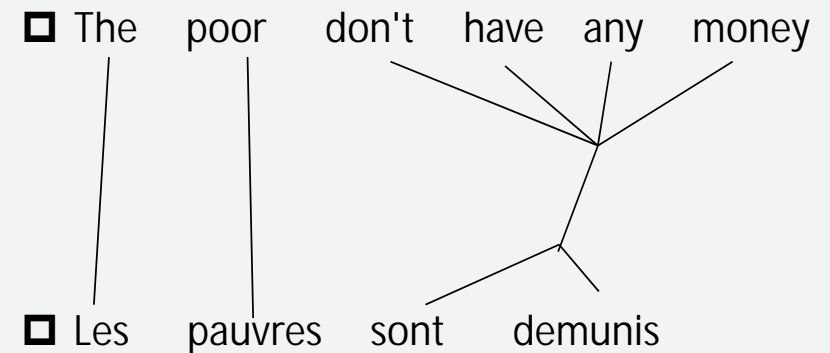


Figure 1. A general alignment (Brown et al., 1993)

2

IBM model



$$\Pr(t | s) = \frac{\Pr(t) \Pr(s|t)}{\Pr(s)}$$

- Denote the sentence pair $\langle s, t \rangle$

Source sentence s containing words s_i for $i \in 1 \dots I$

Target sentence t containing words t_j for $j \in 1 \dots J$

$$\hat{t}_i = \underset{t_i}{\operatorname{argmax}} P(t_i | s_j) = \underset{t_i}{\operatorname{argmax}} (P(t_i) P(s_j | t_i))$$

- **Shortcomings**

- Sensible to the training data
- Hapax logomania

Neural model

● Nature

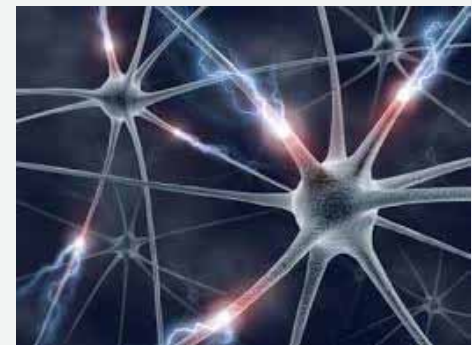
A neural network for word alignment is to compute alignment scores for sentence pairs (Legrand et al., 2016)

● Prerequisite

A pairs of sentences that are translations of one another

● Thought

- Unsupervised training
- A soft-margin objective



2 Neural model

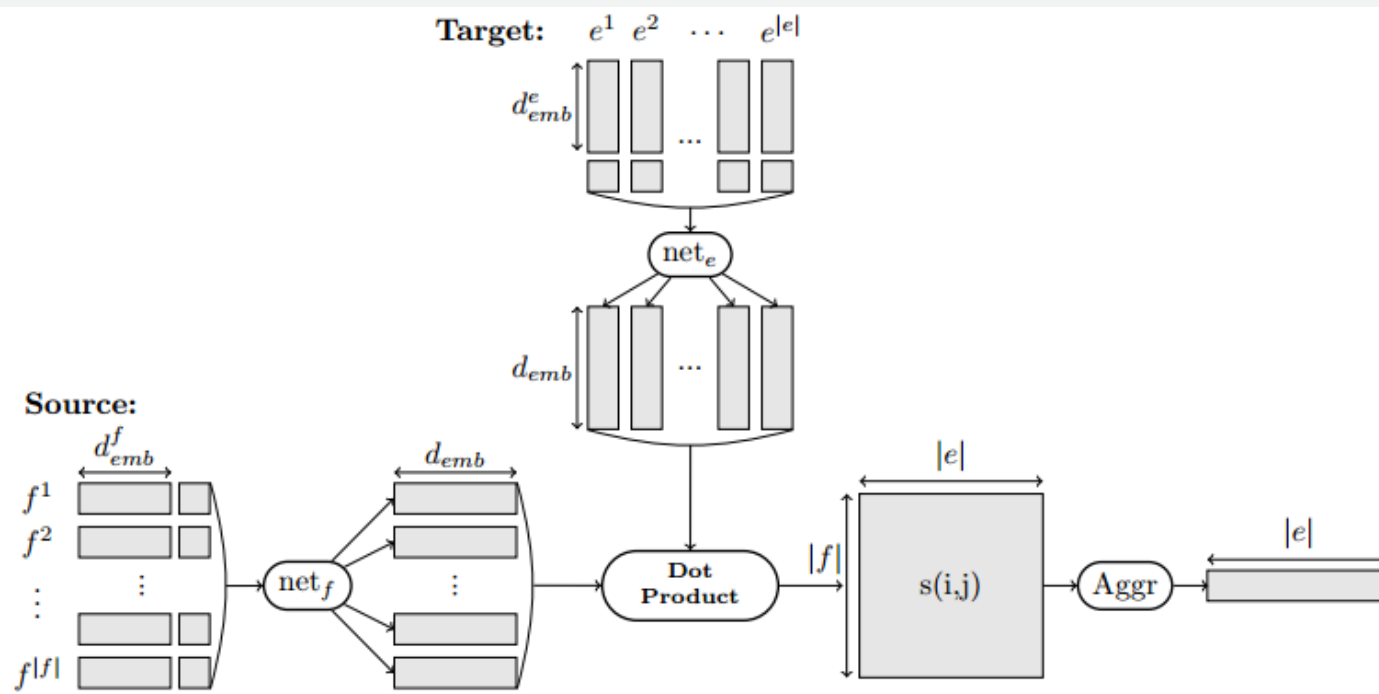


Figure 2. Illustration of the neural model (Legrand et al., 2016)

Denote the sentence pair $\langle e, f \rangle$

Source sentence $\mathbf{f} = (f_1, \dots, f_{|f|})$

Target sentence $\mathbf{e} = (e_1, \dots, e_{|e|})$

$$net_e([e]_i^{d_{win}^e}) \in \mathbb{R}^{d_{emb}}$$

$$net_f([e]_j^{d_{win}^f}) \in \mathbb{R}^{d_{emb}}$$

$$s(i, j) = net_e([e]_i^{d_{win}^e}) \cdot net_f([e]_j^{d_{win}^f})$$



Evaluation rules

Guo Xuechun



3

Recall and Precision



Parameters



Recall



Precision

Sure link—— S

Possible link—— P

Word Alignment Result—— A

$$recall = \frac{|A \cap S|}{|S|}$$

$$precision = \frac{|A \cap P|}{|A|}$$

3

Evaluation Rules

Bilingual Evaluation Understudy ● Papineni et al.,2002

Och,2003 ● Alignment Error Rate

Consistent Phrase Error Rate ● Zhang and Gildea,2004

Huang et al.,2009 ● Error-Sensitive Alignment Error Rate



Conclusions

Ke Huishu



Result of Some Experiments

Alignment		AER	ESAER	ESAER/50	BLEU
Model1	E2F	0.470 5	10.580 9	0.211 6	16.53
	F2E	0.441 7	<u>21.325 5</u>	<u>0.426 5</u>	<u>13.94</u>
	Union	<u>0.471 2</u>	18.010 3	0.360 2	15.10
	GDF	0.344 5	13.275 6	0.265 5	17.63
HMM	E2F	0.430 4	10.717 9	0.214 4	18.66
	F2E	0.391 7	19.797 6	0.396 0	16.71
	Union	0.396 8	16.795 8	0.335 9	17.40
	GDF	0.337 9	13.610 3	0.272 2	18.92
Model4	E2F	0.368 7	<u>8.967 3</u>	<u>0.179 3</u>	<u>20.16</u>
	F2E	0.352 8	17.582 9	0.351 7	16.08
	Union	0.361 7	14.750 9	0.295 0	17.64
	GDF	<u>0.302 3</u>	11.651 3	0.233 0	19.35

Table 1. Using AER, ESAER, BLEU to compare IBM Model-1、HMM、Model-4(Huang et al.,2009)

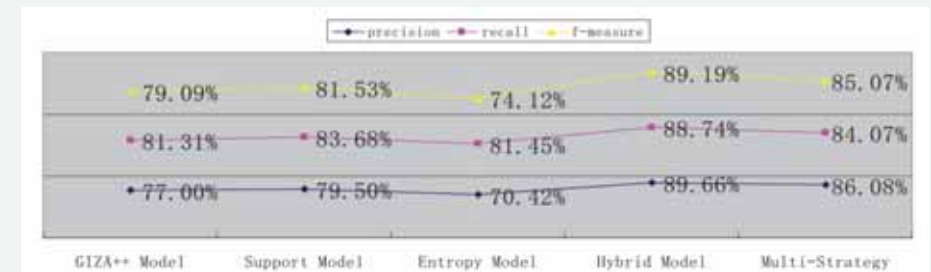


Figure 3. Comparing a hybrid model with other model(Chen et al.,2009)

	English-French			Romanian-English			English-Czech		
	En-Fr	Fr-En	sym	Ro-En	En-Ro	sym	En-Cz	Cz-En	sym
words	22.2	24.2	15.7	47.0	45.5	40.3	36.9	36.3	29.5
+ POS	20.9	23.9	15.3	45.3	42.9	36.9	35.6	33.7	28.2
+ diag	15.1	15.8	12.8	37.6	35.7	32.2	24.8	24.5	21.0
+ POS + diag	13.2	12.1	10.2	33.1	32.2	27.8	24.6	22.9	19.9

Table 2. AER using different input features in each direction in Neural Network-based Model(Legrand et al.,2016)

4

Challenges and Future Work



Three Difficulties

Finding better method to overcome three problems

Evaluation System

To construct a evaluation system to measure the quality of word alignment

Other Techniques

Looking for the probability to combine other techniques with word alignment

4

Take-home Message

Three Difficulties ●

● Two Model for Word Alignment

Evaluation Rules ●

● Future Work



References

- Brown, P. F., Pietra, V. J. D., Pietra, S. A. D., & Mercer, R. L. (1993). The mathematics of statistical machine translation: parameter estimation. *Computational Linguistics*, 19(2), 263-311.
- Chen, L., Xu, J., & Zhang, Y. (2012). A hybrid model for word alignment with bilingual corpus. *IEEE International Conference on Computer Science and Automation Engineering*, 3, 99-103.
- Huang, S., Ning, X. I., Zhao, Y., Dai, X., & Chen, J. (2009). An error-sensitive evaluation metric for word alignment. *Journal of Chinese Information Processing*.
- Legrand, J., Auli, M., & Collobert, R. (2016). Neural network-based word alignment through score aggregation, 66-73.
- Nguyen, G. T., & Dinh, D. (2012). Improving English-Vietnamese Word Alignment Using Translation Model. *IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future*, 1-4.
- Och, F. J. (2003). Minimum error rate training in statistical machine translation. *Meeting on Association for Computational Linguistics*, 32, 160-167.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: a method for automatic evaluation of machine translation. *Meeting on Association for Computational Linguistics*, 4, 311-318.
- Zhang, H., & Gildea, D. (2004). Syntax-based alignment: supervised or unsupervised?. *International Conference on Computational Linguistics*, 418.

