汉英双语命名实体识别与对齐的联合方法

A Joint Model to Identify and Align Bilingual Named Entities

宗成庆

中国科学院自动化研究所

模式识别国家重点实验室

Outline

1. Introduction
2. Our Motivations
3. Our Joint Model
4. Experiments and Analysis
5. Conclusion

1. Introduction

- **Named entities (NEs) include**: Person Names (PER), Location Names (LOC), Organization Names (ORG), temporal and numerical expressions
- **NEs carry essential meanings in human communication**
- **It is requisite to identify NEs in NLP**
  - To simplify NLP tasks
  - It is a very critical in MT and cross-lingual IR systems

1. Introduction

- **The basic statistical data**
  - **C. Huang(2003) gave the result on '**People's Daily': the number of NEs only takes about 8.7% in total number of words, but the errors in results of Chinese word segmentation take about 59.2% in wrong results.

1. Introduction

➢ Our statistics:

<table>
<tr><th colspan="3">错误类型</th><th>错误数</th><th colspan="3">比例(%)</th><th>例子</th></tr>
<tr><td rowspan="6">集外词</td><td rowspan="4">命名实体</td><td>人名</td><td>31</td><td>25.83</td><td rowspan="4">55.0</td><td rowspan="6">98.33</td><td>约翰 斯坦贝克</td></tr>
<tr><td>地名</td><td>11</td><td>9.17</td><td>米苏拉塔</td></tr>
<tr><td>组织机构名</td><td>10</td><td>8.33</td><td>泰党</td></tr>
<tr><td>时间和数字</td><td>14</td><td>11.67</td><td>37万兆</td></tr>
<tr><td colspan="2">专业术语</td><td>4</td><td colspan="2">3.33</td><td>脱氧核糖核酸</td></tr>
<tr><td colspan="2">普通生词</td><td>48</td><td colspan="2">40.00</td><td>致病原</td></tr>
<tr><td colspan="3">切分歧义</td><td>2</td><td colspan="3">1.67</td><td>歌名为</td></tr>
<tr><td colspan="3">合计</td><td>120</td><td colspan="3">100</td><td></td></tr>
</table>

The statistics are carried out based on 418 sentences, 19,777 Chinese characters and 11,739 words, randomly chose form web pages.

1. Introduction

① We randomly chose 100 sentences from the test set of NIST2005 MT evaluation:

- There are 242 NEs in 87 sentences
  - **PER: 78 LOC: 119 ORG: 45**
- 173 NEs are correctly translated (**P=71.5%**)
- 13 NEs are wrongly translated **$69=50_{PER}+4_{LOC}+15_{ORG}$**
- 56 NEs are not translated (**81.2% of 13+56**)
- The BLEU scores increased from **28.02% to 30.05% (+2.03%)** after we replaced the wrong and NULL translations with correct translations

1. Introduction

② We randomly chose another 100 sentences from the test set of NIST2005 MT evaluation:

- There are 281 NEs in all 100 sentences
- 192 NEs are correctly translated (**P=68.3%**)
- 89 NEs are not translated or wrong translations
  - 84.3% are not translated
  - 15.7% are wrong translations

P* = (71.5% + 68.3%)/2 = 69.9%

1. Introduction

- It is a very challenging work to identify and translate NEs
  - Unknown words occurred frequently
  - Semantic translation interweaved with phonetic transliteration
    - PER: 胡锦涛←→Hu Jintao 金成勋←→Kim Sung-Hoon
      时光←→Shi Guang 成龙←→Jackie Chan
    - LOC: 小斯莫基河←→ Little Smoky River
    - ORG: 北京冶金学院←→ Beijing Institute of Metallurgy

1. Introduction

Why we can’t directly translate NEs ?

Translation modals	Examples	What NEs are fit?
Based on phonetic rules	北京/Beijing;東京/ Tokyo	Some PERs, LOCs, ORGs
Irregular transliteration	吉百利/ Cadbury	Some PERs, trade marks
Full semantic translation	自動化所/ Inst. of Auto.	Some ORGs
Combination of transliteration and translation	北海公園/ Beihai Park; 加勒比共同體/Caribbean Comminity	Some LOCs and ORGs
Partial transliteration and partial translation	星巴克/ Starbucks; 劍橋/ Cambridge	Rare
Liberal translation or paraphrase	蝴蝶夢/ Rebecca; 母女情深/ Terms of Endearment	Film names, book names
Different translations for one word	孫中山/ Sun Zhongshan or Sun Yat-Sen	Forward transliteration

1. Introduction

NE Alignment has been focused

- It is the first step to train (NE) translation model
  - May help to correct word alignment
- NE-pairs can be extracted from internet
  - NEs emerge from time to time, and could be transformed in various ways
  - In many cases we don’t know how to translate a NE

1. Introduction

There are two methods for NE Alignment:

Asymmetrical vs. Symmetrical

- **NE recognition is only done on the source side (Al-Onaizan and Knight, 2002)**
  - Use identified source NEs to locate the corresponding target NEs
- **The strong point is to avoid the target NE recognition errors**
  - However, source NE recognition errors still remain

1. Introduction

Symmetrical NE Alignment

- Recognize both source and target NEs (as initial *anchors*)
- Adopt *expansion strategy* (Huang et al., 2003)
  - Shrinking and enlarging the boundaries of anchors
  - Select the top NE linking-pairs between expanded candidate-sets
- In our experiments it greatly raises the NE-pair including-rate before alignment (from 83.9% to 96.1%)
  - Useless, if NEs are un-recognized initially

1. Introduction

- Example：

据报道 <加拉巴戈斯/PER> 國家公園以及當地漁民...

The report said the [Galapagos/PER] National Park and local fishermen …

加拉巴戈
加拉巴戈斯
加拉巴戈斯国
加拉巴戈斯国家
……
加拉巴戈斯国家公园

Said the Galapagos
the Galapagos
Galapagos
Galapagos National
……
Galapagos National Park

1. Introduction

Chinese text | English text

NEs recognition system

NNT集團銷售第一次下降日本電信電話株式會社記入332億日圓的淨受益

NTT Group sales down for the 1st time … Nippon Telegraph and Telephone Corp. posted a net gain of 33.2 billion yen

Extract candidate translation pairs

Filtering

NNT集團 | NTT Group
日本電信電話株式會社 | Nippon Telegraph and Telephone Corp.

1st step

2nd step

3rd step

1. Introduction

Summary:

- NE recognition errors
- NE recognition in two sides is independent with each other

– **To jointly identify and align NEs**

Outline

1. Introduction
2. **Our Motivations**
3. Our Joint Model
4. Experiments and Analysis
5. Conclusion

2. Our Motivations

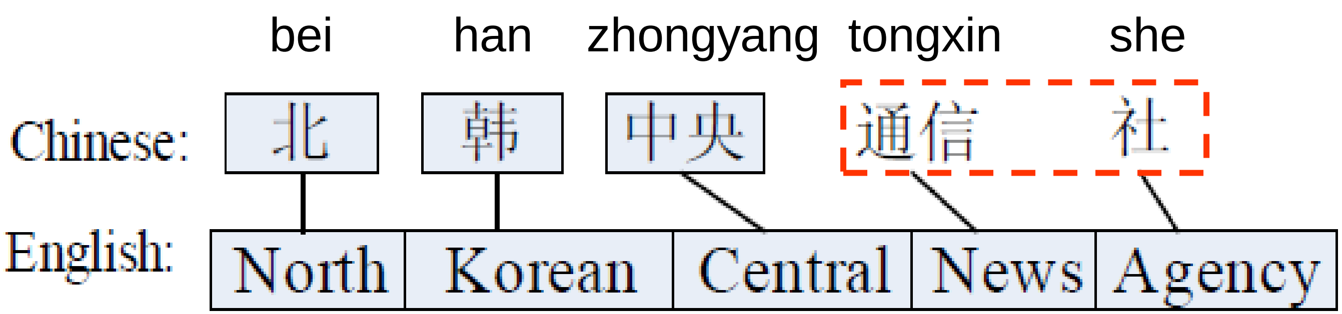
Why we jointly identify and align NEs ?

- Different language has different properties
  - Identifying NE boundaries is difficult for Chinese (un-tokenized language)
  - Classifying NE types is hard for English
- NE alignment may help to recover NER errors
  - To correct the unreliable part from the reliable counterpart in another language

2. Our Motivations

How alignment helps?

- Correcting wrong NE boundaries

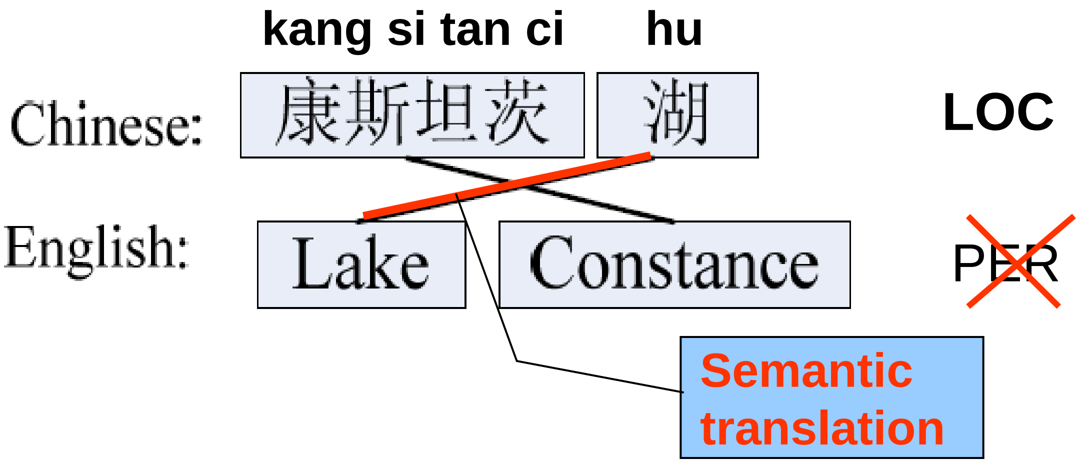


OBSERVATION:

The Chinese has difficulties to identify NE boundaries or even word boundaries

2. Our Motivations

- Correcting wrong NE type

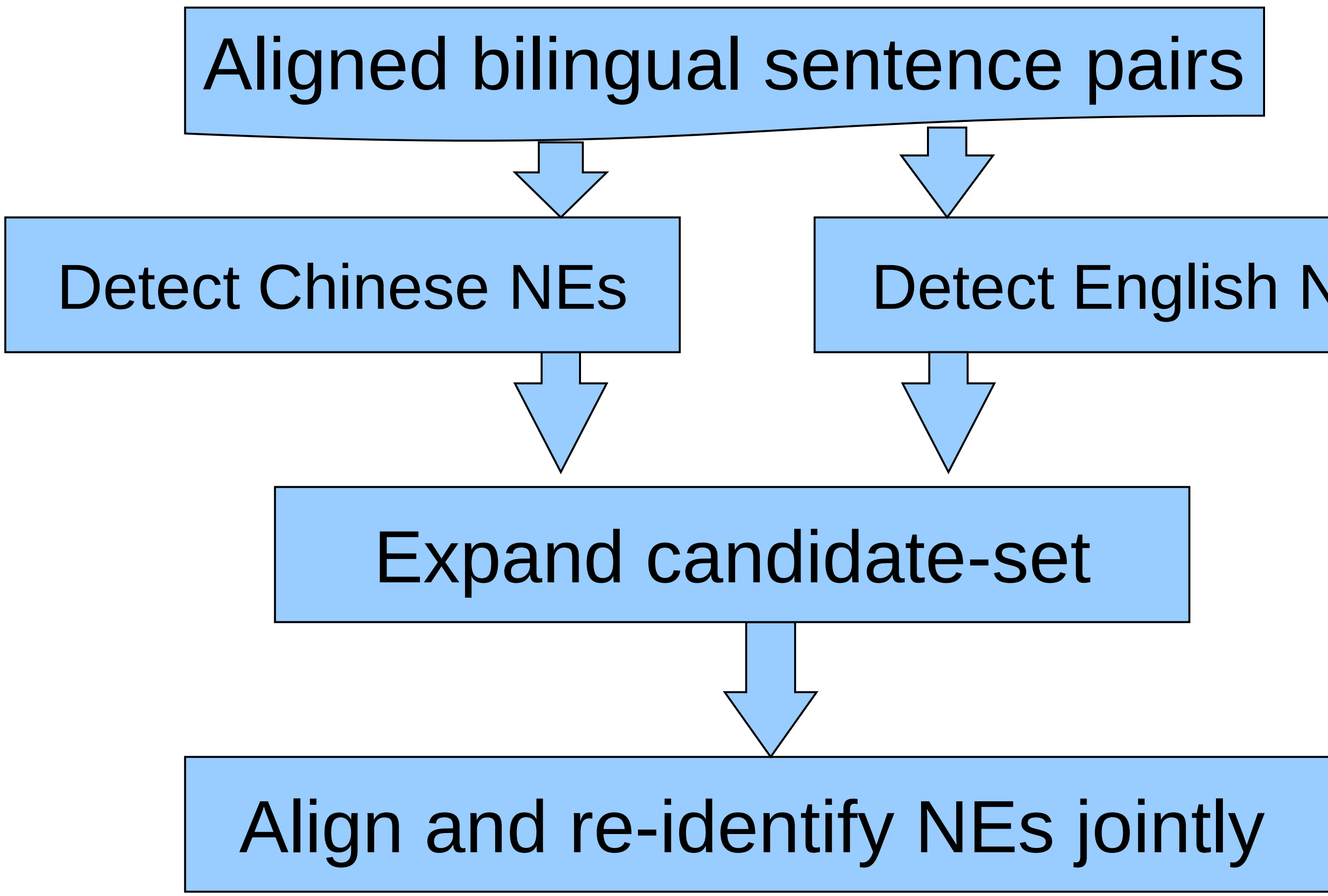


OBSERVATION:

English has difficulty to identify NE types

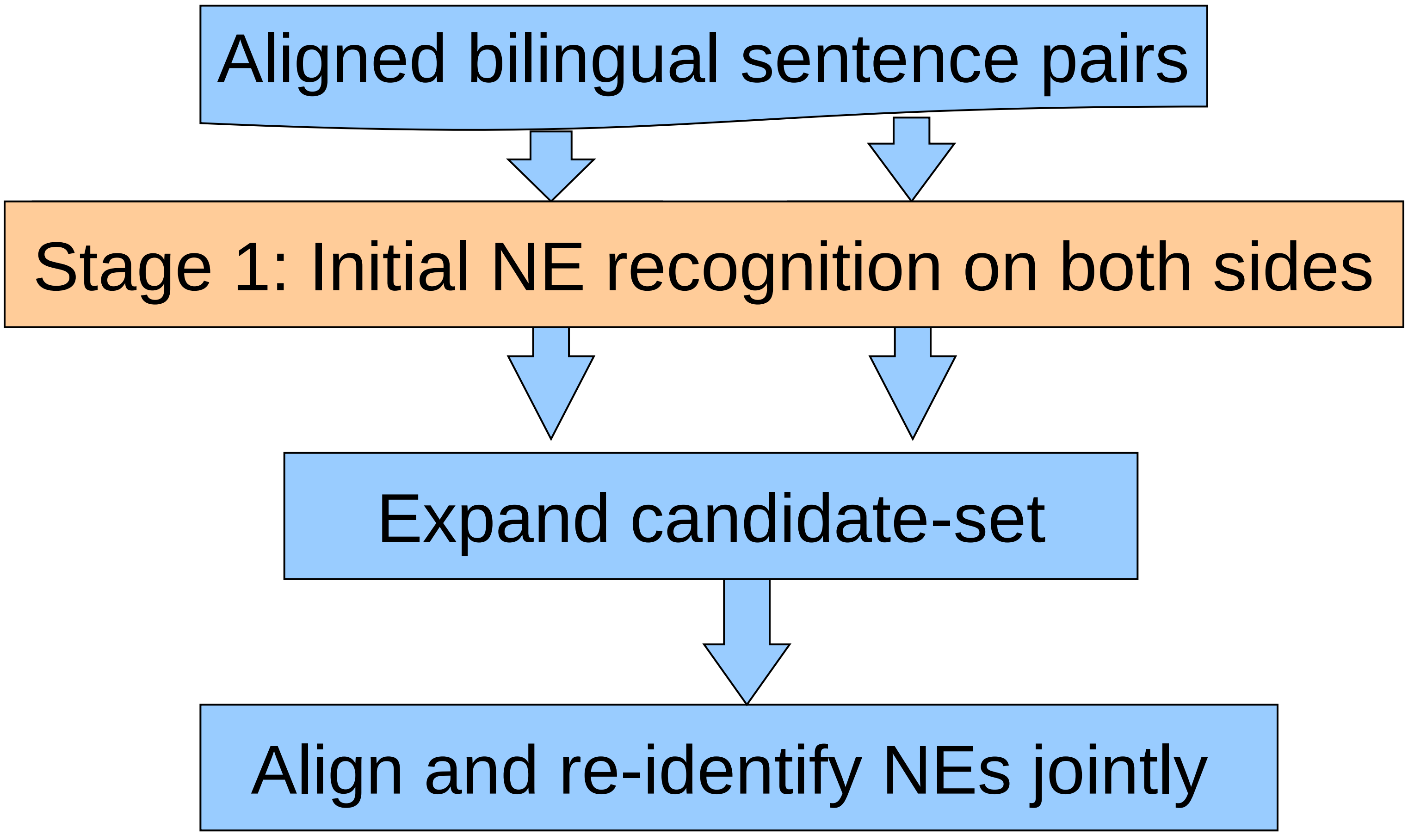
2. Our Motivations

Our joint NE alignment framework:



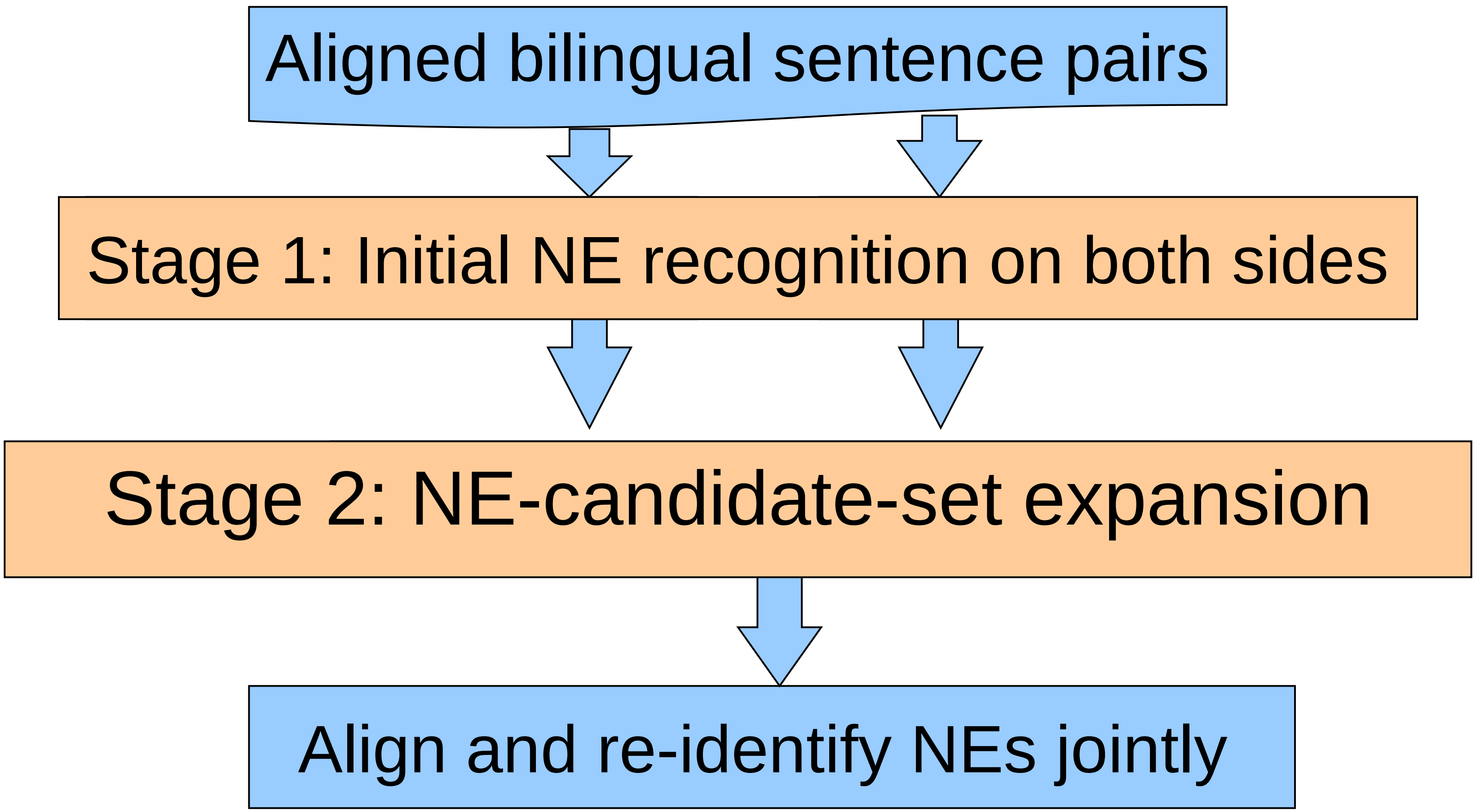
2. Our Motivations

Our joint NE alignment framework:



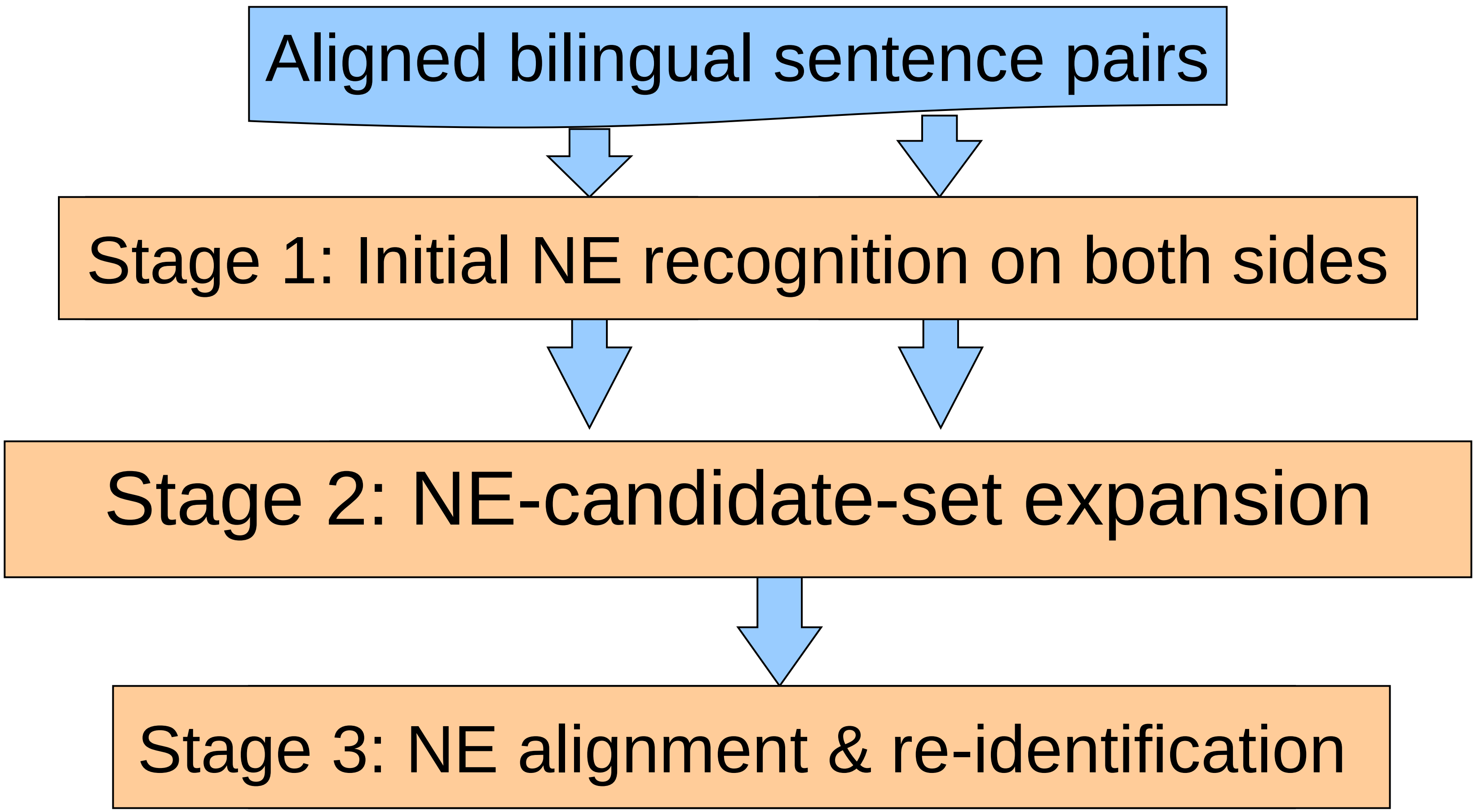
2. Our Motivations

Our joint NE alignment framework:



2. Our Motivations

Our joint NE alignment framework:



2. Our Motivations

Our joint NE alignment framework:

Aligned bilingual sentence pairs

Stage 1: Initial NE recognition on both sides

Four Chinese characters for shrinking and enlarging
Two English words for shrinking and three for enlarging
Including-rate is only 0.8% lower than that without limitation

Stage 3: NE alignment & re-identification

Outline

1. Introduction
2. Our Motivations
3. Our Joint Model
4. Experiments and Analysis
5. Conclusion

3. Our Joint Model

<穆夏拉夫/PER>在<伊斯蘭馬巴德/PER>的記者會上

He said at a press conference in [Islamabad/LOC]:

3. Our Joint Model

<穆夏拉夫 ... 巴德/PER>的記者會上

$$< CNE_i, CType_i >_{i=1}^{S}, S = 2$$

He said at a press conference in [Islamabad/LOC]:

$$RCNE_1^{K_C}$$

穆夏拉
穆夏拉夫
穆夏拉夫在
穆夏拉夫
……
伊斯蘭馬巴
伊斯蘭馬巴德

3. Our Joint Model

<穆夏拉夫 $< CNE_i, CType_i >_{i=1}^{S}, S = 2$ 巴德/PER>的記者會上

He said at a press conference $[ENE_j, EType_j]_{j=1}^{T}, T = 1$ **:**

$RCNE_1^{K_C}$

穆夏拉
穆夏拉夫
穆夏拉夫在
穆夏拉夫
……
伊斯蘭馬巴
伊斯蘭馬巴德

$RENE_1^{K_E}$

Islamabad
In Islamabad
Conference in Islamabad
Islamabad :
In Islamabad :
……

3. Our Joint Model

<穆夏拉夫 $< CNE_i, CType_i >_{i=1}^{S}, S=2$ 巴德/PER>的記者會上

He said at a press conference $[ENE_j, EType_j]_{j=1}^{T}, T=1$ **:**

K=min(S,T)=1

$RCNE_1^{K_C}$

$RENE_1^{K_E}$

穆夏拉

Cartesian product

Islamabad

mabad

:

$$\left\{RCNE^{*}_{<k>}, RENE^{*}_{[k]}, RType^{*}_{k}\right\}_{k=1}^{K} = \underset{\left\{RCNE_{<k>}, RENE_{[k]}\right\}_{k=1}^{K}}{\arg\max} \left[\prod_{k=1}^{K} \boxed{score(RCNE_{<k>}, RENE_{[k]})}\right]$$

3. Our Joint Model

NE pair linking score:

- Given a candidate pair (*RCNE*,*RENE*), *RCNE* (re-generated from *CNE*); *RENE* (re-generated from *ENE*)

$$
\begin{aligned}
&Score(RCNE, RENE) \\
&= \max_{M_{IC}, RType} P\left(\begin{array}{l} \boxed{RCNE, RENE}, \boxed{RType}, \boxed{M_{IC}} \mid \\ <CNE, CType>, CS, [ENE, EType], ES \end{array} \right)
\end{aligned}
$$

Re-generated NE candidates

Re-assigned NE type

Internal component mapping

3. Our Joint Model

$$M_{IC} \equiv < [\boxed{cpn_{<n>}, ew_{[n]}}, \boxed{Mtype}_n]_{n=1}^{N}, \boxed{\delta} >$$

A linked pair of Chinese component and an English word

Mapping type

Mapping type ratio

Mtype: Semantic Translation (TS) and *Phonetic Transliteration* (TL)

Mapping type ratio (δ): percentage of NE internal tokens translated semantically

Ex: <康斯坦茨湖>::[Lake Constance] δ=0.5 (i.e., 1/2):

[cpn1: 康斯坦茨, ew1:*Constance*,TL] and [cpn2:湖, ew2: *Lake*, TS]

3. Our Joint Model

Why we use mapping type ratio δ ?

- Distribution of semantic translation varies among different NE types
  - Semantic translation proportions for PER, LOC, and ORG are 0%, 28.6%, and 74.8% (LDC2005T34 training-set)
  - If δ >0, then it is very unlikely to be PER

3. Our Joint Model

- **Derive the NE linking Model**

$$P\left(M_{IC}, RType, RCNE, RENE \middle| \begin{array}{l} <CNE, CType>, CS, \\ [ENE, EType], ES \end{array} \right)$$

$$\cong \boxed{\begin{array}{l} P\left(M_{IC} \middle| RType, RCNE, RENE\right) \\ \times P\left(RType \mid CNE, ENE, CType, EType\right) \end{array}}$$

$$\boxed{\begin{array}{l} \times P\left(RCNE \mid CNE, CType, CS, RType\right) \\ \times P\left(RENE \mid ENE, EType, ES, RType\right) \end{array}}$$

Bilingual related factors

Monolingual confidence factors

3. Our Joint Model

- **Derive the bilingual related factors**

$$
\begin{aligned}
& P(M_{IC} \mid RType, RENE) \\
& \equiv P([cpn_{<n>}, ew_{[n]}, Mtype_n]_{n=1}^{N}, \delta \mid RType, RENE) \\
& \approx \prod_{n=1}^{N} \begin{bmatrix} P(cpn_{<n>} \mid Mtype_n, ew_{[n]}, RType) \\ \times P(Mtype_n \mid ew_{[n]}, RType) \end{bmatrix} \\
& \quad \times P(\delta \mid RType)
\end{aligned}
$$

3. Our Joint Model

- **Derive the monolingual confidence factors (for Chinese):**

$$
\begin{aligned}
&P(RCNE \mid CNE, CType, CS, RType) \\
&\cong P\left(LeftD, RightD, Str[RCNE] \mid Len_C, CType, RType\right) \\
&\approx \boxed{P(LeftD \mid Len_C, CType, RType) \times P(RightD \mid \boxed{Len_C}, CType, RType)} \\
&\quad \times \boxed{\prod_{m=1}^{M} P(cc_m \mid cc_{m-1}, RType)}
\end{aligned}
$$

make RCNE into a sequence

Left/right distance features

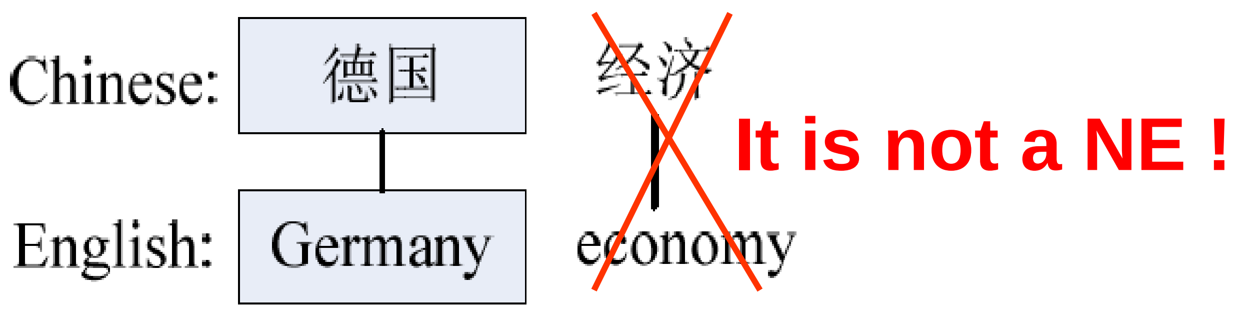
NE bi-gram feature

The length of the originally detected NE

3. Our Joint Model

Why we use LeftD and RightD features?

- Alignment alone is not enough
  - Initially recognized NEs carry NE scope information

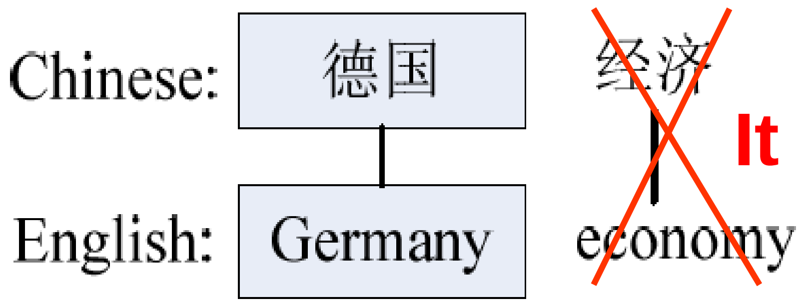


- LeftD and RightD provide anchor information
  - Denote the left and right distances that regenerated NE shrinks/enlarges from the initial anchor.

3. Our Joint Model

Why we use LeftD and RightD features?

- Alignment alone is not enough
  - Initially recognized NEs carry NE scope information



It is not a NE !

- The initial recognition result is: “北韓中央”,
  a candidate is: “韓中央通信社” (enlarging part)
  The LeftD and RightD is “-1” and “+3” respectively

3. Our Joint Model

Weight the alignment model:

- All the bilingual and monolingual factors are weighted differently according to their contributions
- Using the well-known *Minimum Error Rate Training* (MERT) algorithm (Och, 2003) by minimizing the number of associated errors

Outline

1. Introduction
2. Our Motivations
3. Our Joint Model
4. **Experiments and Analysis**
5. Conclusion

4. Experiments and Analysis

◆ Data-Sets

- **Training-Set-I:**
  - 90,412 aligned sentence-pairs, FBIS newswire data
- **Training-Set-II:**
  - LDC2005T34 bilingual NE dictionary (218,772 NE-pairs)
- **Development-Set:** for MERT weight training
  - 200 sentence pairs randomly selected from LDC2005T06
- **Testing-Set:** total 718 gold NE pairs
  - 300 sentence pairs randomly selected from LDC2005T06

4. Experiments and Analysis

◆ Experimental Setting

- **Initial NE recognizers:**
  - Chinese NER (Wu et al., 2005):
    - Overall Performance 84.7% F-score
  - English NER (AK McCallum, 2002, Mallet toolkit):
    - Overall Performance 82.3% F-score

      http://mallet.cs.umass.edu/index.php/Main_Page
- **Evaluation criterion**
  - Precision (P), recall (R) and F-score (F)
  - Type-sensitive and type-insensitive performance

4. Experiments and Analysis

◆ Baseline System (Huang et al., 2003)

- **Adopt symmetric expansion strategy**
- **Utilize the monolingual information**
  - Only a simple bi-gram model
- **Adopt three cost functions:**
  - Transliteration cost $\sum_{(i,j)\in A^*}[\arg\max_{\{cpn_i \in E_{cpn_i}\}} \log P(ew_j \mid cpn_i)]$
  - Translation cost $\log \frac{1}{J^I}\prod_{i=1}^{I}\sum_{j=1}^{J} P(cpn_i \mid ew_j)] + \log \frac{1}{I^J}\prod_{j=1}^{J}\sum_{i=1}^{I} P(ew_j \mid cpn_i)$
  - NE scope tagging cost (bi-gram model)

$$\min_{RType}[-\log(\prod_{m=1}^{M} P(cc_m \mid cc_{m-1}, RType)) - \log(\prod_{n=1}^{N} P(ew_n \mid ew_{n-1}, RType))]$$

4. Experiments and Analysis

◆ Experimental Results

- **The NE-pair alignment performance**

Models	P (%)	R (%)	F (%)
Baseline	77.1 (67.1)	79.7 (69.8)	78.4 (68.4)
***Exp*-1 (*All-N-BiFactors*)**	**77.7 (73.5)**	79.9 (**75.7**)	78.8 (**74.6**)
***Exp*-2 (*Fully-JointModel*)**	**83.7 (78.1)**	**86.2 (80.7)**	**84.9 (79.4)**
***Exp*-3 (*Weighted-Joint Model*)**	**85.9 (80.5)**	**88.4 (83.0)**	**87.1 (81.7)**

4. Experiments and Analysis

◆ Experimental Results

- **The NE-pair alignment performance**

Models	P (%)	R (%)	F (%)
Baseline	77.1 (67.1)	79.7 (69.8)	78.4 (68.4)
Exp-1 **(*All-N-BiFactors*)**	**77.7** (**73.5**)	79.9 (**75.7**)	78.8 (**74.6**)
Exp-2 **(*Full*** [illegible] ***ointModel*)**	**83.7 (78.1)**	**86.2 (80.7)**	**84.9 (79.4)**
Exp-3 **(*Wei*** [illegible] ***oint Model*)**	**85.9 (80.5)**	**88.4 (83.0)**	**87.1 (81.7)**

$$\left[\prod_{n=1}^{N} P(cpn_{<n>} \mid Mtype_n, ew_{[n]}, RType) \times P(Mtype_n \mid ew_{[n]}, RType)\right]^{\boxed{\frac{1}{N}}}$$

4. Experiments and Analysis

◆ Experimental Results

- The NE-pair alignment performance

Models	P (%)	R (%)	F (%)
Baseline	77.1 (67.1)	79.7 (69.8)	78.4 (68.4)
***Exp*-1 (*All-N-BiFactors*)**	**77.7 (73.5)**	79.9 (**75.7**)	78.8 (**74.6**)
***Exp*-2 (*Fully-JointModel*)**	**83.7 (78.1)**	**86.2 (80.7)**	**84.9 (79.4)**
***Exp*-3 (*Weighted-Joint Model*)**	**85.9 (80.5)**	**88.4 (83.0)**	**87.1 (81.7)**

It also improves NE recognition performance F-score (Chinese +3.9%; English +4.6%)

4. Experiments and Analysis

Why we have to use a derived model?

- **Why not ME framework with *primitive features?***
  - No assumptions required
  - Primitive Features: internal component mapping, initial and finial NE types, NE bigram-based string, and left/right distance
- **Tested under various training-sets:**
  - 400, 4,000, 40,000, and 90,412 sentence-pairs of Training-Set-I
  - Adopt YASMET toolkit (http://www.fjoch.com/YASMET.html)

4. Experiments and Analysis

The derived model performs better!

- Model derivation also helps (not just features)
  - Providing additional human knowledge about the dependency between features (implied constraints)

Models	400	4,000	40,000	90,412
ME framework	36.5 (0%)	50.4 (0%)	62.6 (0%)	67.9 (0%)
Un-weighted Joint Model	+4.6 (+12.6%)	+4.5 (+8.9%)	+4.3 (+6.9%)	+4.1 (+6.0%)
Weighted Joint Model	+5.0 (+13.7%)	+4.7 (+9.3%)	+4.6 (+7.3%)	+4.5 (+6.6%)

Gap becomes more noticeable as training-set gets smaller

4. Experiments and Analysis

How about the capability in learning new NEs?

- New NE-pairs are extracted via semi-supervised learning
  - **Alignment imposes constraints**
    - **3.9% and 4.6%** F-score improvements on Chinese and English NER
  - **Split Training-Set-I into two parts:**
    - 50,412 (unlabeled data) and 40,000 (labeled data) sentence pairs
  - **Various seed data-sets:**
    - 100; 400; 4,000 and 40,000 are extracted from labeled data
  - **Iterate until the result converges**

4. Experiments and Analysis

- **Semi-supervised learning for English NER**

Models	100	400	4,000	40,000
Initial-NER	36.7 (0%)	58.6 (0%)	71.4 (0%)	79.1 (0%)
NER-Only	-2.3 (-6.3%)	-0.5 (-0.8%)	-0.3 (-0.4%)	-0.1 (-0.1%)
NER+Baseline (alignment, F. Huang)	+4.9 (+13.4%)	+3.4 (5.8%)	+1.7 (2.4%)	+0.7 (0.9%)
NER+Joint Model (alignment, Our)	+10.7 (+29.2%)	+8.7 (+14.8%)	+4.8 (+6.7%)	+2.3 (+2.9%)

Note: The last three lines show the performance of semi-supervised methods by using different seed data to train 50,000 unlabeled sentences.

4. Experiments and Analysis

◆ Four types of errors:

Error types	Reference NE-Pair	Initial NE anchors	Final output	Percentage
(A) Inconsistency of original NEs or components	{<乌伯林更镇>:: [Berlingen]}	CNE: <乌伯林> ENE: [Berlingen]	{<乌伯林>:: [Berlingen]}	23% (25)
(B) Missing or spurious anchors	{<东协>::[ASEAN]}	CNE: <> ENE: [ASEAN]	No such alignment	24% (27)
(C) Unassumed mapping types	{<葛兰素制药厂>::[GSK]}	CNE: <葛兰素> ENE: [GSK]; [Ziagen]	{<葛兰素>:: [Ziagen]}	27% (30)
	{<明仁>::[Akihito]}	CNE: <明仁> ENE: [Akihito]	{<明仁>:: [Hirohito]}	
(D) Wrong NE scope	{<南北韩>:: [South and North Koreas]}	CNE: <韩> ENE: [North Koreas]	{<北韩>::[North Koreas]}	26% (29)

4. Experiments and Analysis

◆ Four types of errors:

Error types	Reference NE-Pair	Initial NE anchors	Final output	Percentage
(A) Inconsistency of original NE components				
(B) Missing or spurious anchors	{<东 …N]}	CNE: <> ENE: [ASEAN]	No such alignment	24% (27)
(C) Unassumed mapping types	{<葛兰素制药厂>::[GSK]}	CNE: <葛兰素> ENE: [GSK]; [Ziagen]	{<葛兰素>:: [Ziagen]}	27% (30)
	{<明仁>::[Akihito]}	CNE: <明仁> ENE: [Akihito]	{<明仁>:: [Hirohito]}	
(D) Wrong NE scope	{<南北 …	CNE: <韩> ENE: [North …	{<北韩>::[North …	26% (29)

Acronym/Abbreviation: (GSK for "GlaxoSmithKline Factory")

Loanword: Japanese kanji "明仁", pronounced as "Mingren", aligns with "Akihito".

Outline

1. Introduction
2. Our Motivations
3. Our Joint Model
4. Experiments and Analysis
5. **Conclusion**

5. Conclusion

- **We have proposed a joint model derived in a principled way**
  - **Capable to learn new NE-pairs via semi-supervised learning**
  - **Some novel features/factors are introduced:**
    - Bilingual related factors:
      - Mapping type ratio & NE type consistency
    - Monolingual certainty factors:
      - Initial NE anchor information (LeftD and RightD)
      - Normalized NE bi-gram score

5. Conclusion

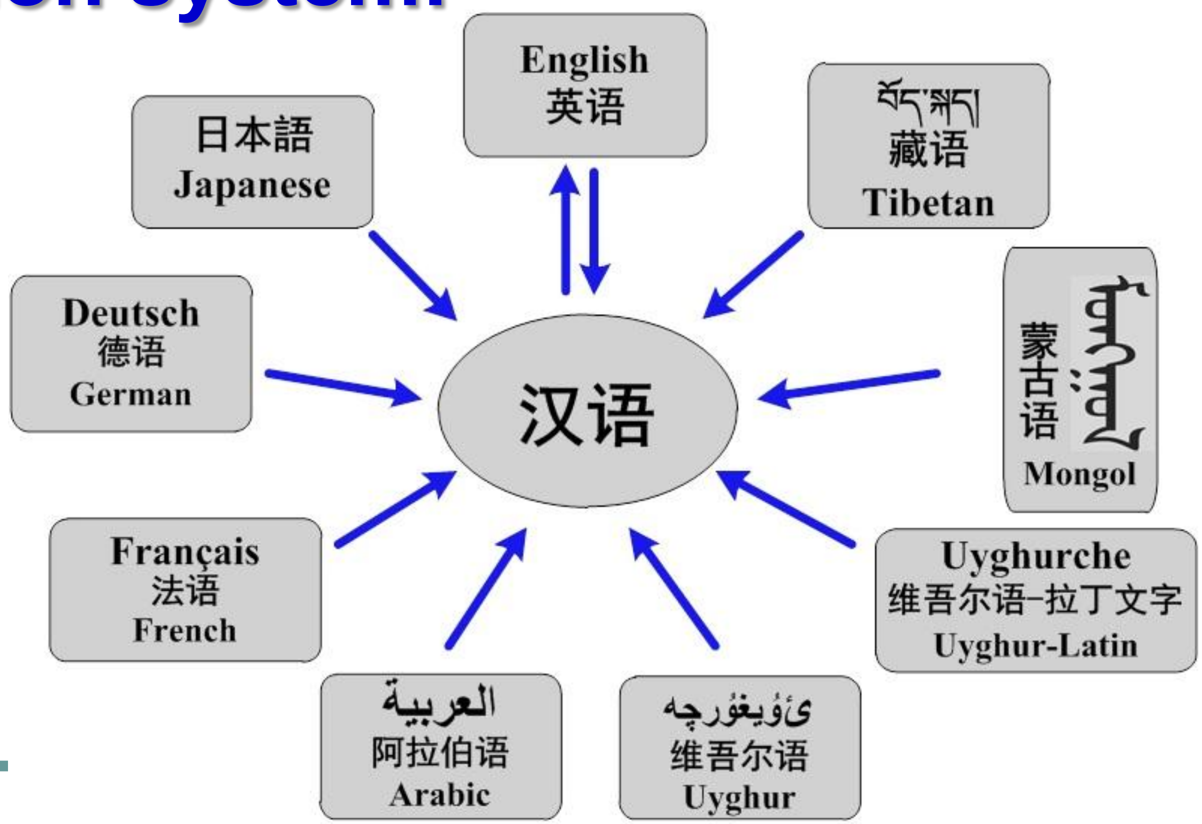
- **Our joint model outperforms the baseline system:**
  - From **68.4% to 81.7%** in type-sensitive F-score, 18.6% relative improvement
  - Outperforms ME model with primitive features: **13.7% relative improvement for 400** sentence-pairs
  - Capable to learn new NE-pairs via semi-supervised learning
  - The experiments have proven that the alignment factors are essential in learning new NE-pairs via semi-supervised learning

5. Conclusion

- **Problems:**
  - The introduced features are not linguistic knowledge dependent, especially no semantic information involved
  - It is not enough only based on parallel corpora
- **Future work**
  - To develop linguistically related features
  - Use internet data

5. Conclusion

- **This work has been partially employed in our English-Chinese machine translation system:**



5. Conclusion

◆ Related publications:

1. Yufeng Chen, Chengqing Zong and Keh-Yih Su. A Joint Model to Simultaneously Identify and Align Bilingual Named Entities. To appear in ***Computational Linguistics***
2. Yufeng Chen, Chengqing Zong, and Keh-Yih Su. On Jointly Recognizing and Aligning Bilingual Named Entities. ***Prof. ACL***, Uppsala, Sweden, July 11–16, 2010. Pages 631-639
3. Yufeng Chen, Chengqing Zong. A Structure-based Model for Chinese Organization Name Translation. ***ACM TALIP***, 7(1): 1-30, 2008
4. Yufeng Chen and Chengqing Zong. A Semantic-Specific Model for Chinese Named Entity Translation. ***Proc. IJCNLP-2011***, pages 138-146
5. 陈钰枫，宗成庆，苏克毅，汉英双语命名实体识别与对齐的交互式方法，**计算机学报**，2011年9月，第34卷第9期，第1688—1696页

Thank you!

谢谢！

感谢陈钰枫博士在本报告的准备过程中给予的大力帮助！