



Statistical Machine Transliteration with Multi-to-Multi Joint Source Channel Model

Rui Wang
with Yu Chen and Yi Zhang
DFKI GmbH, Germany



Outline

- Motivation
- System Overview
- Evaluation Results
- Conclusion



Motivation

- Previous Work
 - Joint Source-Channel
 - [Li et al., 2004], [Li et al., 2007]
 - Pinyin
 - Statistical Machine Translation (SMT)
- Our Setup
 - Multi-to-Multi Joint Source-Channel
 - System integration

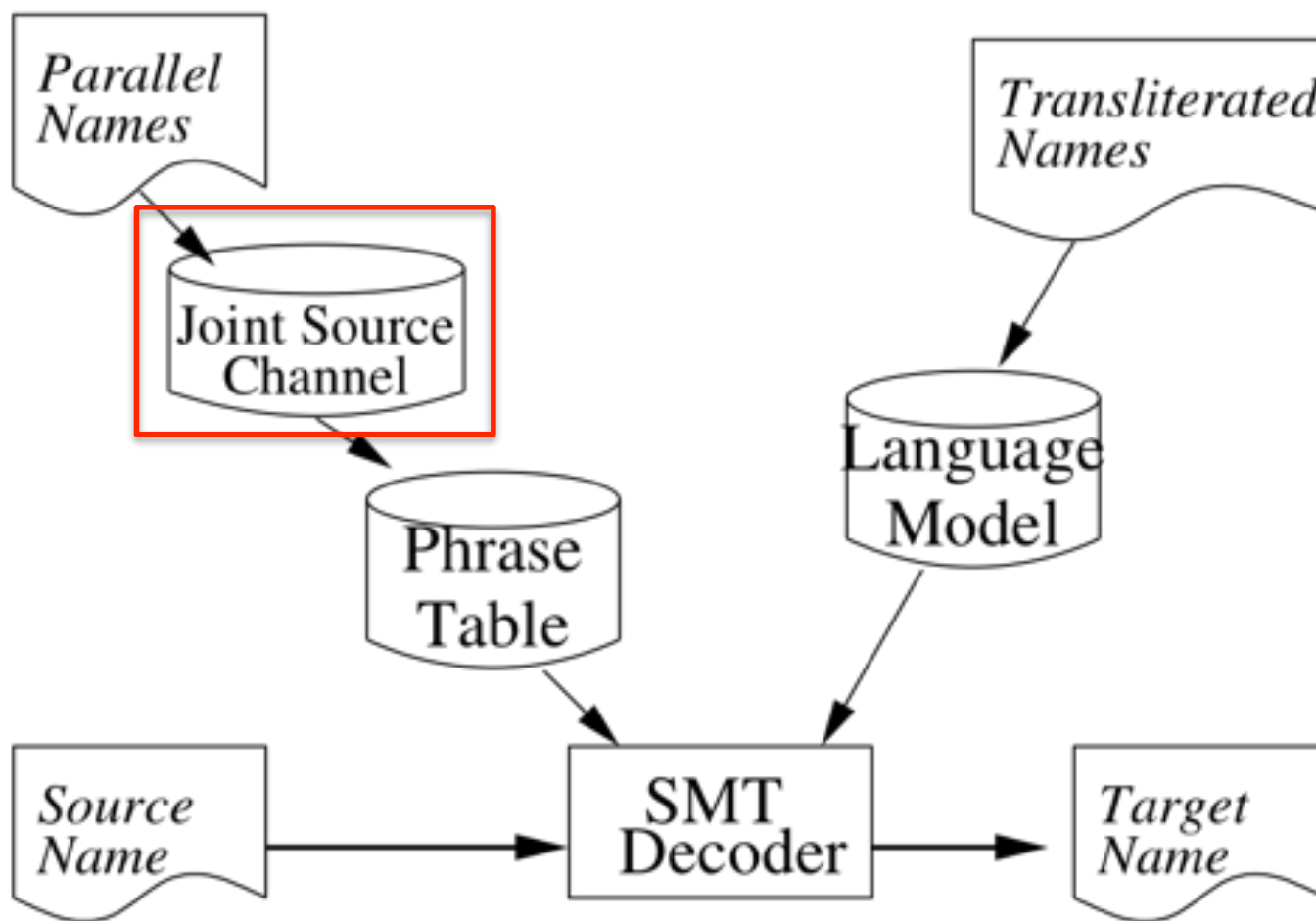


Transliteration in SMT

- EuroMatrixPlus (EU FP7)
 - English, German, French, Spanish \leftrightarrow Chinese
 - Hybrid MT
- Named-Entity translation
 - Statistical
 - Rule-based
 - Hybrid



System Overview



Phrase-based SMT

- Translation model
- Language model

$$\operatorname{argmax}_t P(t | s) = \operatorname{argmax}_t (P(t)P(s | t))$$


Joint Source-Channel Model

- [Li et al., 2004]

$$P(s, t, \alpha) = \sum_{k=1}^K P(\langle e, c \rangle_k | \langle e, c \rangle_{k-n+1}^{k-1})$$

$$\bar{t} = \arg \max_{s, \alpha} P(s, t, \alpha)$$

$$\bar{s} = \arg \max_{t, \alpha} P(s, t, \alpha)$$



Multi-to-Multi JSC

English	Chinese
A/ <u>JA</u> / <u>X</u>	埃/ <u>甲</u> / <u>克斯</u>
A/BA/STE/NIA	阿/巴/斯蒂/尼亚
<u>AHL</u> /BERG	<u>阿尔</u> /伯格

- N-gram → 1-gram
- Random start → Normalize all

Experiments

- Preprocessing
- SMT settings
 - Parameter tuning
- Results



Preprocessing

- Tokenization
- No-phonetic mapping
- Separate training sets



SMT Setting

- Models
 - Translation model
 - “Language model” SRILM [Stolcke, 2002]
 - Length penalty
- Moses Decoder
 - [Koehn et al., 2007]
 - No distortion



Parameter Tuning

- MERT
 - [Och, 2003]
 - Z-MERT [Zaidan, 2009]
- Development Set
 - Randomly 500
- F-score-based tuning
 - Instead of BLEU



(Official) Results

Tasks	System	ACC	Mean F	MRR	Map_ref
English-to-Chinese	M2MJSC +PBSMT	0.320	0.674	0.397	0.308
English-to-Chinese	M2MJSC	0.260	0.638	0.340	0.251
Chinese-to-English	M2MJSC +PBSMT	0.133	0.746	0.210	0.133
Chinese-to-English	M2MJSC	0.117	0.731	0.177	0.117

Additional Results

Tasks	LM	ACC	Mean F	MRR	Map_ref
English-to-Chinese	3-gram	30.62	67.30	38.63	29.71
English-to-Chinese	5-gram	31.71	67.71	39.43	30.70
English-to-Chinese	7-gram	31.62	67.61	39.46	30.66
Chinese-to-English	3-gram	11.67	73.06	18.29	11.68
Chinese-to-English	5-gram	12.48	73.29	18.82	12.47
Chinese-to-English	7-gram	12.97	73.89	19.48	12.97

Conclusion

- First participation
- M2M-JSC
- System integration (20%)
- LM effectiveness



Future Work

- Richer feature set for M2MJSC
- Different models for SMT

- Additional knowledge
 - Pinyin
 - Dictionaries
- System combination
 - [Chen et al., 2009]





Xie Xie!

Questions?

