Beijing Institute of Technology

Prosodic annotation enriched statistical machine translation

Peidong Guo, Heyan Huang, Ping Jian, Yuhang Guo

Department of Computer Science and Technology, Beijing Institute of technology

Beijing Engineering Research Center of High Volume Language Information Processing and Cloud Computing Applications

OUTLINE

- Introduction
- Methods
- Experiments
- Conclusion

INTRODUCTION

- Prosodic features
  - Continuous prosodic features

    Duration time, speaking rate, fundamental frequencies ,energy
  - Discrete prosodic annotations

    Boundaries, emphasis, pronunciation, tone
- Statistical machine translation (SMT) with prosodic features
  - Use energy, duration and F0 for speed-up parsing and translation[1]
  - Prosody enriched factored translation model[2]
  - Rule extraction according to prosodic boundaries[3]

INTRODUCTION

- Challenge
  - Conventional approaches ignore the key information beyond the text
- Prosodic features in our machine translation system
  - Improve source language analysis in Chinese-English translation
    - Factored models, word lattices with prosody
  - Help target language generation in English-Chinese translation
    - Re-ranking model

OUTLINE

- Introduction
- Methods
- Experiments
- Conclusion

METHODS

- Basic Assumption
  I. The translations of some words rely on their prosodic features
  II. Normal Chinese sentences have reasonable prosodic structures
- Prosodic features we used:
  - Pronunciation

    The phonetic alphabet and the tone
  - Prosody boundaries

    Boundaries of sentences, phrases and prosodic words
  - Emphasis

    Emphasis of sentences and of phrases

METHODS

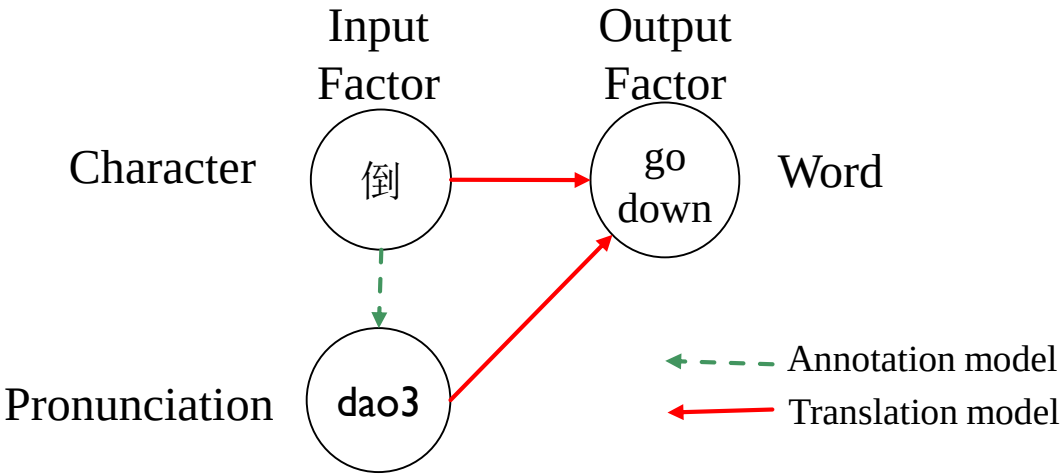
- Features used in our CRF prosody tagger

Type	Features
Character Feature	character(i) (i=-2,-1,0,+1,+2)
Pronunciation Feature	pronunciation(0)
Composite Feature	character(-2,-1,0)
	character(0,1,2)
	pronunciation(-2,-1,0)
	pronunciation(0,1,2)

METHODS

- Factored translation model with prosody

[SRC]上|shang4 周|zhou1 日|ri4 台|tai2 风|feng1 , **倒|dao3** 了|le5 一|yi4 棵|ke1 木|mu4 瓜|gua1 树|shu4
[REF]A papaya tree **went down** in the typhoon last Sunday
[RES]There is a typhoon on Sunday, **pour** a papaya tree



METHODS

- Factored translation models with prosody
  - Chinese-English (C2E) translation

$$h_{T,C2E}(e,f) = \lambda_0 \log P(f|e) + \sum_{i=1}^{3} \lambda_i \log P(f|f_i) P(f_i|e)$$

$$\approx \lambda_0 \log P(f|e) + \sum_{i=1}^{3} \lambda_i \log P(f_i'|e)$$

Annotation model for prosodic annotations

Translation model for prosodic factors

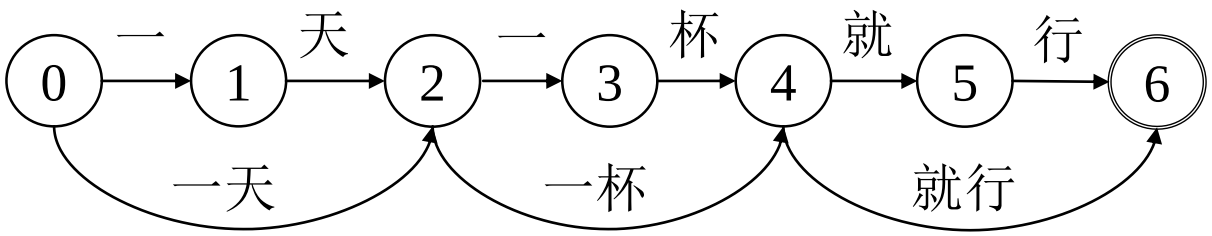
  - English-Chinese (E2C) translation

$$h_{T,E2C}(e,f) = \lambda_0 \log P(f|e) + \sum_{i=1}^{3} \lambda_i \log P(f|e_i) P(e_i|e)$$

Generation step for prosodic factors

METHODS

- Word lattice with prosodic boundaries
  - Take full advantages of the prosodic structures
  - Expand the space of translation candidates



METHODS

- Re-ranking with language model of prosodic tags
  - Target language probabilities interpolated with prosodic features:

$$P'(e) = \frac{1}{Z}\exp\left(\alpha_0 \log P(e_0) + \sum_{i=1}^{3} \alpha_i \log P(e_i)\right)$$

  - Re-ranking score for n-best lists:

$$Score(e|f) = Score'(e|f) + \frac{1}{Z}\exp\left(\sum_{i=1}^{3} \alpha_i \log P(e_i)\right)$$

OUTLINE

- Introduction
- Methods
- Experiments
- Conclusion

EXPERIMENTS

- Corpus
  - DISCOURSE-CASS (prosodic corpus) :
    - 244 audio files (12 hours, 8 kHz sampling rate)
    - Domains: insurance, bank, China Telecom, China Unicom and QQ chat
    - Provided by Phonetics and Speech lab, Institute of Linguistics, CASS
  - BOLT (parallel corpus):
    - Chinese-English parallel corpus of chat messages
    - Training(150,123 sentence pairs), development (4,932 sentence pairs) ,testing (4,977 sentence pairs)
    - Provided by LDC

EXPERIMENTS

- Baseline System
  - Moses for training and decoding
  - GIZA++ with sL0 norm For Alignment
  - SRILM for language model
  - MERT for tuning parameters
  - BLEU for evaluation (Character-level computation for Chinese)
- CRF prosody tagger
  - Accuracy for prosodic boundaries prediction: 69.66%
  - Accuracy for emphasis annotations prediction: 62.02%

EXPERIMENTS

- Results of the factored models with prosody
  - Character(Char)
  - Pronunciation(Pron)
  - Boundaries(Boun)
  - Emphasis(Emph)

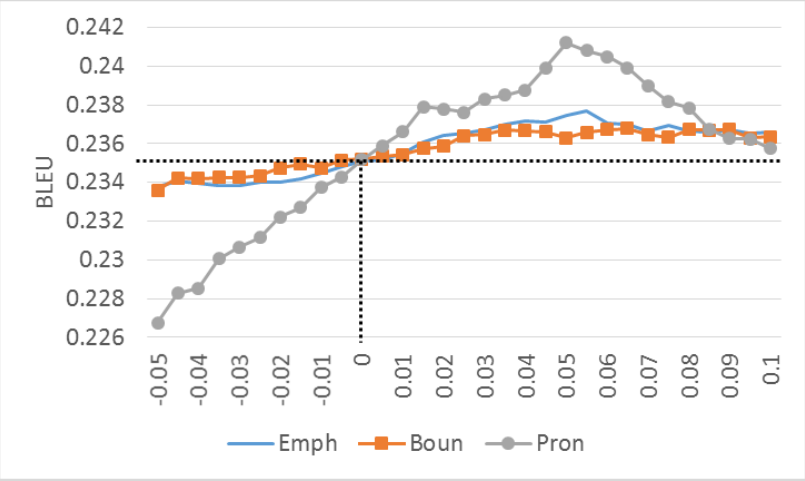
Language pair	Factor combination	Dev BLEU	Test BLEU
Chinese-English	Char→Word(Baseline)	19.23	13.60
	Char+Pron→Word	19.35	13.71
	Char+Boun→Word	19.33	**13.78***
	Char+Emph→Word	19.10	13.44
	Char+Pron+Boun→Word	19.30	13.70
English-Chinese	Word→Char(Baseline)	23.51	20.78
	Word→Char+Pron	23.60	20.92
	Word→Char+Boun	23.17	20.38
	Word→Char+Emph	23.56	20.91
	Word→Char+Boun+Emph	23.62	**20.99***

EXPERIMENTS

- Results of word lattice decoding with prosodic boundaries

Experiment	Test BLEU
Baseline	13.60
Word Lattice	**13.75**

- Results of re-ranking model with prosody



Rescore features	Dev BLEU	Test BLEU
Baseline	23.51	20.78
Pron	24.12	21.12*
Boun	23.67	20.79
Emph	23.77	21.02*
Pron+Boun+Emph	24.15	**21.29***

OUTLINE

- Introduction
- Methods
- Experiments
- Conclusion

CONCLUSION

- In this work, we exploited ways to enrich statistical machine translation with prosodic annotations:
  - Factored models with prosodic annotation
  - Word lattice decoding with prosodic boundaries
  - Re-ranking models with prosodic feature
- Character-level prosodic features can improve the performance of translation
- Further studies will focus on the effectiveness of high-level prosodic features for statistical machine translation.

REFERENCES

- [1] E. Noth, A. Batliner, A. Kiessling, and R. Kompe, "VERBMOBIL: the use of prosody in the linguistic components of a speech understanding system," *Speech & Audio Processing IEEE Transactions on*, vol. 8, pp. 519-532, 2000.
- [2] V. K. R. Sridhar, S. Bangalore and S. S. Narayanan, "Factored translation models for enriching spoken language translation with prosody," in *INTERSPEECH*, 2008, pp. 2723-2726.
- [3] X. Gan and Y. Chen, "Improving of Hierarchical Phrase-based SMT Model Based on Prosodic Structure Information," *Computer Knowledge and Technology*, pp. 2860-2863, 2013.

THANKS!

Q&A