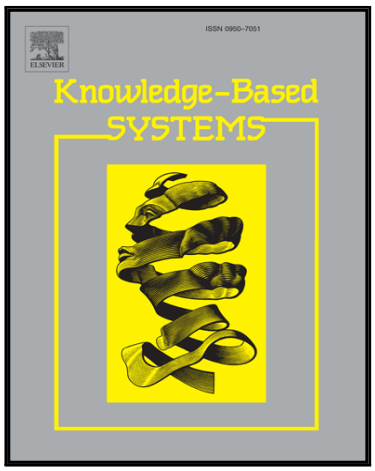

Alignment-consistent Recursive Neural Networks for Bilingual Phrase Embeddings

Jinsong Su, Biao Zhang, Deyi Xiong, Yang Liu, Min Zhang

PII: S0950-7051(18)30211-9
DOI: 10.1016/j.knosys.2018.05.003
Reference: KNOSYS 4322

To appear in: *Knowledge-Based Systems*

Received date: 5 October 2017
Revised date: 28 April 2018
Accepted date: 2 May 2018

Please cite this article as: Jinsong Su, Biao Zhang, Deyi Xiong, Yang Liu, Min Zhang, Alignment-consistent Recursive Neural Networks for Bilingual Phrase Embeddings, *Knowledge-Based Systems* (2018), doi: 10.1016/j.knosys.2018.05.003

Alignment-consistent Recursive Neural Networks for Bilingual Phrase Embeddings

Jinsong Su[a,b], Biao Zhang[a], Deyi Xiong[c,*], Yang Liu[d], Min Zhang[c]

[a] *Xiamen University, Xiamen 361005, China*
[b] *Beijing advanced innovation center for language resources and intelligence, Beijing 100083, China*
[c] *Soochow University, Suzhou 215006, China*
[d] *Tsinghua University, Beijing 100084, China*

Abstract

Learning semantic representations of bilingual phrases is very important for statistical machine translation to overcome data sparsity and exploit semantic information. In this paper, we consider word alignments as a semantic bridge between the source and target phrases, and propose two neural networks based on the conventional recursive autocoder, which exploit word alignments to generate alignment-consistent bilingual phrase structures: One is Alignment Enhanced Recursive Autoencoder that incorporates a word-alignment-related error into the final objective function; The other is Alignment Guided Recursive Neural Network which treats word alignments as direct signals to guide phrase structure constructions. Then, we further establish the semantic correspondences between the source and target nodes of the generated bilingual phrase structures via word alignments. By jointly minimizing recursive autoencoder reconstruction errors, structural alignment consistency errors and cross-lingual reconstruction errors, our model not only generates alignment-consistent phrase structures, but also captures different levels of semantic correspondences within bilingual phrases. Experiments on the NIST Chinese-English translation task show that our model achieves significant improvements over the baseline.

*Corresponding author
Email addresses: jssu@xmu.edu.cn (Jinsong Su), zb@stu.xmu.edu.cn (Biao Zhang), dyxiong@suda.edu.cn (Deyi Xiong), liuyang2011@tsinghua.edu.cn (Yang Liu), minzhang@suda.edu.cn (Min Zhang)