## A Simple, *Fast* Strategy for Weighted Alignment Hypergraph

Zhaopeng Tu<sup>1,2</sup> Jun Xie<sup>2</sup> Yajuan Lv<sup>2</sup> Qun Liu<sup>2,3</sup>

I. University of California, Davis, USA
2. Institute of Computing Technology, CAS, China
3. Dublin City University, Ireland

## Outlines

- Weighted Alignment Hypergraph (WAH)
- Rule Extraction on WAH
- Experiments
- Conclusion

## Outlines

- Weighted Alignment Hypergraph (WAH)
- Rule Extraction on WAH
- Experiments
- Conclusion

- A **graph-based** compact representation of multiple alignments
  - Alignment as minimal connected sub-graph (MCS)



Usually not connected

- A **graph-based** compact representation of multiple alignments
  - Alignment as minimal connected sub-graph (MCS)



Usually not connected

- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



# 在 桌子 上—> is on the desk

- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help





- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



- A **graph-based** compact representation of multiple alignments
  - Multiple alignments help



## Outlines

- Weighted Alignment Hypergraph (WAH)
- Rule Extraction on WAH
- Experiments
- Conclusion

- Challenge
  - compute the count of the phrases  $\boldsymbol{p}$  in a hypergraph  $\mathbf{H}$

 $count(p|H) = \frac{p(all \ alignments \ with \ which \ p \ is \ consistent)}{p(all \ alignments \ encoded \ in \ H)}$ 

- Challenge
  - compute the count of the phrases  $\boldsymbol{p}$  in a hypergraph  $\mathbf{H}$

 $count(p|H) = \frac{p(all \ alignments \ with \ which \ p \ is \ consistent)}{p(all \ alignments \ encoded \ in \ H)}$ 

#### To enumerate all alignments in H is *NP-complete*

- Solution 1: Bull's alogrithm
  - Key observation: only consider overlapped subhypergraph (independence)



- Solution 1: Bull's alogrithm
  - Key observation: only consider overlapped subhypergraph (independence)



- Solution 1: Bull's alogrithm
  - Key observation: only consider overlapped subhypergraph (independence)



- Solution 1: Bull's alogrithm
  - Key observation: only consider overlapped subhypergraph (independence)



- Solution 1: Bull's alogrithm
  - Key observation: only consider overlapped subhypergraph (independence)



- Solution 2: inside-outside algorithm
  - In-depth independence

- Solution 2: inside-outside algorithm
  - In-depth independence (



在

- Solution 2: inside-outside algorithm
  - In-depth independence (



在

e₄

Solution 2: inside-outside algorithm

$$p(A|H) = \prod_{h_i \in H} p(A_i|h_i)$$

$$p(A|H,P) = \prod_{h_i \in OS} p(A_i|h_i,P) \prod_{h_j \in NS} p(A_j|h_j)$$

$$c(P|H) = \frac{p(A|H, P)}{p(A|H)} = \frac{\prod_{h_i \in OS} p(A|h_i, P)}{\prod_{h_i \in OS} p(A|h_i)}$$

## Outlines

- Weighted Alignment Hypergraph (WAH)
- Rule Extraction on WAH
- Experiments
- Conclusion

### Setup

- HPB
- FBIS corpus
- Same settings as (*Liu, Tu, and Lin. ACL 2013*)
  - Build hypergraphs from 100-best lists
  - A subhypergraph has at most 10 hyperedges

#### **Translation Results**

Outperforms both 1-best and 10-best alignments



Multiple alignments helps!

#### Bull's vs Inside-outside



## Outlines

- Weighted Alignment Hypergraph (WAH)
- Rule Extraction on WAH
- Experiments
- Conclusion

### Conclusion

- A graph-based weighted alignment hyper graph
  - Easier to exploit independence
  - Inside-outside algorithm is twice faster than Bull's algorithm

## Thanks!

*Q&A*