

Word Association Models and Search Strategies for Discriminative Word Alignment



UNIVERSITAT POLITÈCNICA DE CATALUNYA



Barcelona Media

Patrik Lambert¹, Rafael Banchs²

¹ Universitat Politècnica de Catalunya

² Barcelona Media Innovation Center

ABSTRACT

This paper deals with core aspects of discriminative word alignment systems. We compare various low-computational-cost word association models: χ^2 score, log-likelihood ratio and IBM model 1. We also compare three beam-search strategies.

We show that it is more flexible and accurate to let links to the same word compete together, than introducing them sequentially in the alignment hypotheses, which is the strategy followed in several systems.

SEARCH STRATEGIES

Introduction

Search in discriminative alignment system: find the alignment which minimises the sum of weighted feature costs

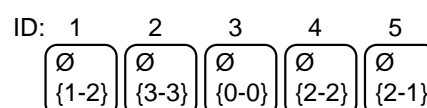
Promising (or possible) link list: in order to limit the search space, a set of promising links is first selected

(0)the (1)member (2)state (3).
(0)los (1)pais (2)miembr (3).

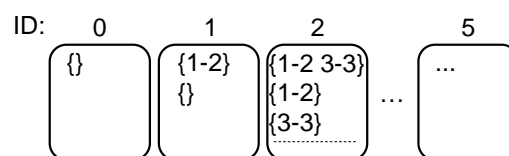
Link	Cost	Corresponding words
1-2	0.1736	member-miembr
3-3	0.6758	-.
0-0	1.3865	the-los
2-2	1.8285	state-miembr
2-1	2.4027	state-pais

Baseline Search

Possible-links stacks:



Alignment hypothesis stacks:



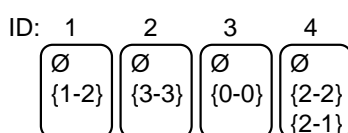
main issue: the final alignment depends on the order in which links are introduced

Proposed Improvements

- successive iterations of the alignment algorithm. Start from final alignment of previous iteration instead of empty alignment

- consider links to same source (or target) word at the same time

Possible-links stacks:



Example

the member state .
| |
| |
los pais miembr .

Next possible links: 1.8 state-miembr
2.4 state-pais

Model weights:
1 word association (wa)
0.5 distortion (d)
2 unlinked word (um)

EXPERIMENTAL RESULTS

Basic Word Association Models

Model	Recall	Precision	AER
χ^2	62.4	86.7	27.1
LLR	59.4	75.7	33.2
IBM1	65.9	90.3	23.5
IBM1+UM	67.1	92.5	21.9

Search Strategies

Strategy	Recall	Precision	AER
Baseline	67.1	92.5	21.9
SWS	67.1	93.5	21.6
SWP	66.3	91.5	22.8
SWP 2lt	66.7	93.2	21.9
SWP 3lt	67.3	93.2	21.5

(SWP: source word position ; SWS: source word score)

CONCLUSIONS

- Results suggest that the log-likelihood ratio is not an adequate word association measure for discriminative word alignment.
- Even the simplest IBM model probabilities allowed a significant alignment quality improvement with respect to other word association measures.

- We compared three beam-search strategies. We showed that starting from the empty alignment is not the best choice, and that it is more flexible and accurate to let links to the same word compete together, than to introduce them sequentially in the alignment hypotheses.