

Word Alignment

Input: bitext

Output: Je fais un bureau
I am making a desk

Je vais le faire
I am going to do it

target: English aligned to source: French

\bar{T} = Eng words \bar{S} = French words

\bar{a} = alignments

$a_i = j$ English word i aligns to Fr word j

$\bar{S} =$ Je fais un bureau

$\bar{a} =$ 1 | 2 / 2

$\bar{T} =$ I am making a desk

$a_1 = 1$ $a_4 = 3$

$a_2 = 2$ $a_5 = 4$

$a_3 = 2$

Alignment model: $P_{\theta}(\bar{T}, \bar{a} | \bar{S})$
Unsupervised

generative