

Statistical MT

Given a source (foreign) language sentence $f_1^J = f_1 \dots f_j \dots f_J$,

Produce a target language (English) sentence $e_1^I = e_1 \dots e_j \dots e_I$.

Among all possible target language sentences, choose the sentence with the highest probability:

$$\hat{e}_1^I = \underset{I, e_1^I}{\operatorname{argmax}} p(e_1^I | f_1^J) \quad (1)$$

We stick to the e_1^I, f_1^J notation despite translating from English to Czech.