

Recap: Bayes' rule

$$e^* = \operatorname{argmax}_{e \in E} \underbrace{p(e)}_{\text{Language model}} \underbrace{p(f|e)}_{\text{Translation model}}$$

Language model

Translation model

- $p(e)$ models the *fluency* of the translation
- $p(f|e)$ models the *adequacy* of the translation
- argmax is the search problem implemented by a *decoder*