Why is it easier to deal with?

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

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*