From Bayes to Log-Linear Model

Och (2002) discusses some problems of Equation 3:

- Models estimated unreliably $\Rightarrow$ maybe LM more important:

$$\hat{e}_1^{\hat{I}} = \operatorname*{argmax}_{I, e_1^I} p(f_1^J | e_1^I) (p(e_1^I))^2 \tag{6}$$

- In practice, "direct" translation model equally good:

$$\hat{e}_1^{\hat{I}} = \operatorname*{argmax}_{I, e_1^I} p(e_1^I | f_1^J) p(e_1^I) \tag{7}$$

- Complicated to *correctly* introduce other dependencies.
  $\Rightarrow$ Use log-linear model instead.