

IBM-1 Pseudocode

initialize $P(f|e)$ Slide 8

for a number of iterations:

 set $tcount(f, e)$ to 0 for all f, e

 set $total(e)$ to 0 for all e

 for each sentence pair (F, E) in training corpus:

 for each **unique** word f in F :

$denom_c = 0$

 for each unique word e in E :

$denom_c += P(f|e) * F.count(f)$

 for each **unique** word e in E :

$tcount(f, e) += P(f|e) * F.count(f) * E.count(e) / denom_c$

$total(e) += P(f|e) * F.count(f) * E.count(e) / denom_c$

 for each e in $domain(total(:))$:

 for each f in $domain(tcount(:, e))$:

$P(f|e) = tcount(f, e) / total(e)$

Slide 13

Slide 12