How do we formalize these calculations?

Recall the elimination rule

$$P(A) = P(A, B) + P(A, \overline{B}) \quad \text{or also}$$

$$P(A) = P(A|B)\,P(B) + P(A|\overline{B})\,P(\overline{B})$$

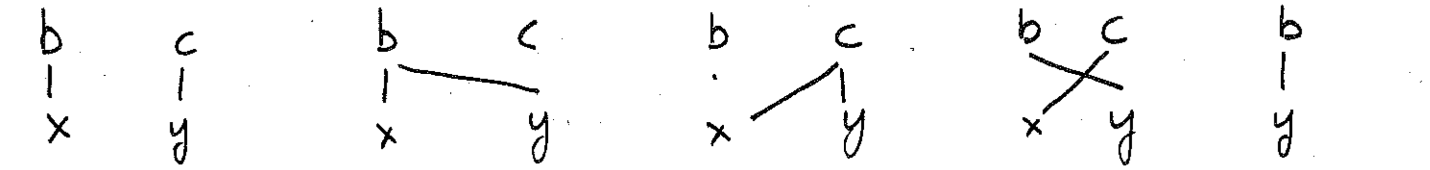
In general,

$$P(A) = \sum_B P(A, B) \quad \text{or} \quad P(A) = \sum_B P(A|B)\,P(B)$$

TRANSLATION PROBABILITY

$$t(y|b) = \frac{t(y, b)}{t(b)} = \frac{\sum_a t(a, y, b)}{\sum_a t(a, b)} \qquad (1)$$

For example, first iteration of step 3, we collect fractional counts (the numerator of equation (1)), and then normalise (the denominator of equation (1)).



$P(a_{11}) = p_{11}$ $\quad$ $P(a_{12}) = p_{12}$ $\quad$ $P(a_{13}) = p_{13}$ $\quad$ $P(a_{14}) = p_{14}$ $\quad$ $P(a_2) = p_2$

$$t(y|b) = \frac{p_{12} + p_{14} + p_5}{p_{11} + p_{12} + p_{13} + p_{14} + p_2} = \frac{\frac{1}{4} + \frac{1}{4} + 1}{\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + 1}$$

PROBABILITY OF ALIGNEMENT

Probability of an alignement given two sentences

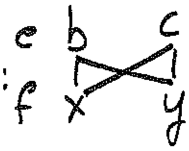
$$P(a|e,f) = \frac{P(a,e,f)}{P(e,f)} = \frac{P(a,f|e)\,\cancel{P(e)}}{P(f|e)\,\cancel{P(e)}}$$

By the elimination rule $P(f|e) = \sum_a P(a, f|e)$

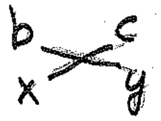
Therefore: $$P(a|e,f) = \frac{P(a,f|e)}{\sum_a P(a,f|e)}$$

where $$P(a,f|e) = \prod_{j=1}^{l_f} t(f_j \mid e_{a_j})$$

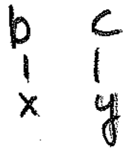
For example, second iteration of step 2 a_{14} :

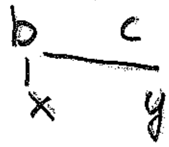


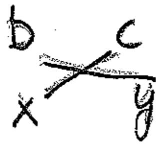
$$P(a_{14}|e,f) = \frac{P(a_{14}, f|e)}{P(a_{11},f|e) + P(a_{12},f|e) + P(a_{13},f|e) + P(a_{14},f|e)}$$



$$= \frac{t(x|c) \times t(y|b)}{\begin{array}{l} t(x|b) \times t(y|c) + \\ t(x|b) \times t(y|b) + \\ t(x|c) \times t(y|c) + \\ t(x|c) \times t(y|b) \end{array}}$$







Simplification of denominator

$$P(f|e) = \sum_{a} P(a, f|e)$$

$$= \sum_{a_1 \ldots a_{l_f} \in \{1, \ldots, l_e\}^{l_f}} \prod_{j=1}^{l_f} t(f_j \mid e_{a_j})$$

For example, the sum of products

$$\left.\begin{array}{l} t(x|b) \times t(y|c) + \\ t(x|b) \times t(y|b) + \end{array}\right\} = t(x|b)\,[t(y|c) + t(y|b)]$$

$$+$$

$$\left.\begin{array}{l} t(x|c) \times t(y|c) + \\ t(x|c) \times t(y|b) \end{array}\right\} = t(x|c)\,[t(y|c) + t(y|b)]$$

$$= [t(x|b) + t(x|c)] \times [t(y|c) + t(y|b)]$$

that is the product of sums expressed by the formula

$$P(f|e) = \prod_{j=1}^{l_f} \sum_{i=1}^{l_e} t(f_j \mid e_i)$$