

Ranking Function for JM Smoothing

$$f(q, d) = \sum_{\substack{w_i \in d \\ w_i \in q}} c(w, q) \left[\log \frac{p_{\text{Seen}}(w_i | d)}{\alpha_d p(w_i | C)} \right] + n \log \alpha_d$$

$$p(w | d) = (1 - \lambda) \frac{c(w, d)}{|d|} + \lambda p(w | C) \quad \lambda \in [0, 1]$$

$$\frac{p_{\text{seen}}(w | d)}{\alpha_d p(w | C)} = \frac{(1 - \lambda) p_{\text{ML}}(w | d) + \lambda p(w | C)}{\lambda p(w | C)} = 1 + \frac{1 - \lambda}{\lambda} \frac{c(w, d)}{|d| p(w | C)}$$

$$f_{\text{JM}}(q, d) = \sum_{\substack{w \in d \\ w \in q}} c(w, q) \log \left[1 + \frac{1 - \lambda}{\lambda} \frac{c(w, d)}{|d| p(w | C)} \right]$$