

Algorithm: Forward algorithm for profile HMMsInitialisation: $f_{M_0}(0) = 1$.

$$\text{Recursion: } f_{M_k}(i) = e_{M_k}(i)[f_{M_{k-1}}(i-1)a_{M_{k-1}M_k} + f_{I_{k-1}}(i-1)a_{I_{k-1}M_k} \\ + f_{D_{k-1}}(i-1)a_{D_{k-1}M_k}];$$

$$f_{I_k}(i) = e_{I_k}(i)[f_{M_k}(i-1)a_{M_kI_k} + f_{I_k}(i-1)a_{I_kI_k} \\ + f_{D_k}(i-1)a_{D_kI_k}];$$

$$f_{D_k}(i) = f_{M_{k-1}}(i)a_{M_{k-1}D_k} + f_{I_{k-1}}(i)a_{I_{k-1}D_k} + f_{D_{k-1}}(i)a_{D_{k-1}D_k}.$$

$$\text{Termination: } f_{M_{M+1}}(L+1) = f_{M_M}(L)a_{M_M M_{M+1}} + f_{I_M}(L)a_{I_M M_{M+1}} \\ + f_{D_M}(L)a_{D_M M_{M+1}}. \quad \triangleleft$$

Algorithm: Backward algorithm for profile HMMsInitialisation: $b_{M_{M+1}}(L+1) = 1$;

$$b_{M_M}(L) = a_{M_M M_{M+1}};$$

$$b_{I_M}(L) = a_{I_M M_{M+1}};$$

$$b_{D_M}(L) = a_{D_M M_{M+1}}.$$

$$\text{Recursion: } b_{M_k}(i) = b_{M_{k+1}}(i+1)a_{M_k M_{k+1}}e_{M_{k+1}}(x_{i+1}) \\ + b_{I_k}(i+1)a_{M_k I_k}e_{I_k}(x_{i+1}) + b_{D_{k+1}}(i)a_{M_k D_{k+1}};$$

$$b_{I_k}(i) = b_{M_{k+1}}(i+1)a_{I_k M_{k+1}}e_{M_{k+1}}(x_{i+1}) \\ + b_{I_k}(i+1)a_{I_k I_k}e_{I_k}(x_{i+1}) + b_{D_{k+1}}(i)a_{I_k D_{k+1}};$$

$$b_{D_k}(i) = b_{M_{k+1}}(i+1)a_{D_k M_{k+1}}e_{M_{k+1}}(x_{i+1}) \\ + b_{I_k}(i+1)a_{D_k I_k}e_{I_k}(x_{i+1}) + b_{D_{k+1}}(i)a_{D_k D_{k+1}}. \quad \triangleleft$$

The forward and backward variables can then be combined to re-estimate emission and transition probability parameters as follows:

Algorithm: Baum-Welch re-estimation equations for profile HMMsExpected emission counts from sequence x :

$$E_{M_k}(a) = \frac{1}{P(x)} \sum_{i|x_i=a} f_{M_k}(i)b_{M_k}(i);$$

$$E_{I_k}(a) = \frac{1}{P(x)} \sum_{i|x_i=a} f_{I_k}(i)b_{I_k}(i).$$

Expected transition counts from sequence x :

$$A_{X_k M_{k+1}} = \frac{1}{P(x)} \sum_i f_{X_k}(i)a_{X_k M_{k+1}}e_{M_{k+1}}(x_{i+1})b_{M_{k+1}}(i+1);$$

$$A_{X_k I_k} = \frac{1}{P(x)} \sum_i f_{X_k}(i)a_{X_k I_k}e_{I_k}(x_{i+1})b_{I_k}(i+1);$$

$$A_{X_k D_{k+1}} = \frac{1}{P(x)} \sum_i f_{X_k}(i)a_{X_k D_{k+1}}b_{D_{k+1}}(i).$$