1 Derivation Approach 1

$$\begin{split} f_k^* &= \arg\max_{f_k} E[P(y^d | \vec{x}^d, F_{k-1}^*, f_k)] \\ f_k^* &= \arg\max_{f_k} \frac{1}{|D|} \sum_{d \in D} P(y^d | \vec{x}^d, F_{k-1}^*, f_k) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} \frac{P(y^d, \vec{x}^d, F_{k-1}^*, f_k)}{P(\vec{x}^d, F_{k-1}^*, f_k)} \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} \sum_{\{y_i^d\}_{1 \leq i \leq k}} \frac{P(y^d, \vec{x}^d, F_{k-1}^*, f_k, \{y_i^d\}_{1 \leq i \leq k})}{P(\vec{x}^d, F_{k-1}^*, f_k)} \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} \sum_{\{y_i^d\}_{1 \leq i \leq k}} \frac{P(y^d | \{y^d\}_{1 \leq i \leq k}) \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k) P(\vec{x}^d) \prod_{i=1}^{k-1} (P(f_i)) P(f_k)}{P(\vec{x}^d, F_{k-1}^*, f_k)} \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} \sum_{\{y_i^d\}_{1 \leq i \leq k}} P(y^d | \{y^d\}_{1 \leq i \leq k}) \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k) \end{split}$$

 $P(y^d | \{y^d\}_{1 \le i \le k}) \hspace{1cm} = \hspace{1cm} I[y^d \hspace{1cm} = \hspace{1cm} \wedge_{i=1}^k] \hspace{1cm} = \hspace{1cm} \begin{cases} y^d = 1 \Rightarrow \{y_i^d = 1\}_{1 \le i \le k} \\ y^d = 0 \Rightarrow \{y_i^d\}_{1 \le i \le k, \exists y_i^d | \|y_i^d = 0\} \end{cases}$

1.1 Precision case

$$\begin{split} f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[\overrightarrow{y} = \wedge_{i=1}^k] \prod_{i=1}^{k-1} (P(y_i^d | \overrightarrow{x}^d, f_i)) P(y_k^d | \overrightarrow{x}^d, f_k) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = 1] \prod_{i=1}^{k-1} (P(y_i^d = 1 | \overrightarrow{x}^d, f_i)) P(y_k^d = 1 | \overrightarrow{x}^d, f_k) \\ &+ I[y^d = 0] \sum_{\substack{\{y_i^d\}_{1 \leq i \leq k}, \\ \exists y_i^d | | y_i^d = 0}} \prod_{i=1}^{k-1} (P(y_i^d | \overrightarrow{x}^d, f_i)) P(y_k^d | \overrightarrow{x}^d, f_k) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = 1] \prod_{i=1}^{k-1} (P(y_i^d = 1 | \overrightarrow{x}^d, f_i)) P(y_k^d = 1 | \overrightarrow{x}^d, f_k) \\ &+ I[y^d = 0] (1 - \prod_{i=1}^{k-1} (P(y_i^d = 1 | \overrightarrow{x}^d, f_i)) P(y_k^d = 1 | \overrightarrow{x}^d, f_k)) \end{split}$$

1.2 Recall case

$$P(y^{d}|\{y^{d}\}_{1 \le i \le k}) = I[y^{d} = \bigvee_{i=1}^{k}] = \begin{cases} y^{d} = 1 \Rightarrow \{y_{i}^{d}\}_{1 \le i \le k, \exists y_{i}^{d} || y_{i}^{d} = 1 \\ y^{d} = 0 \Rightarrow \{y_{i}^{d} = 0\}_{1 \le i \le k} \end{cases}$$

$$\begin{split} f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = \vee_{i=1}^k] \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = 0] \prod_{i=1}^{k-1} (P(y_i^d = 0 | \vec{x}^d, f_i)) P(y_k^d = 0 | \vec{x}^d, f_k) \\ &+ I[y^d = 1] \sum_{\substack{\{y_i^d\}_{1 \leq i \leq k}, \\ \exists y_i^d | | y_i^d = 1}} \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = 0] \prod_{i=1}^{k-1} (P(y_i^d = 0 | \vec{x}^d, f_i)) P(y_k^d = 0 | \vec{x}^d, f_k) \\ &+ I[y^d = 1] (1 - \prod_{i=1}^{k-1} (P(y_i^d = 0 | \vec{x}^d, f_i)) P(y_k^d = 0 | \vec{x}^d, f_k)) \end{split}$$

2 Derivation Approach 2

$$\begin{split} f_k^* &= \arg\max_{f_k} log(P(D|F_{k-1}^*, f_k)) \\ f_k^* &= \arg\max_{f_k} log(\prod_{d \in D} P(y^d | \vec{x}^d, F_{k-1}^*, f_k)) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} log\left(\frac{P(y^d, \vec{x}^d, F_{k-1}^*, f_k)}{P(\vec{x}^d, F_{k-1}^*, f_k)}\right) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} log\left(\sum_{\{y_i^d\}_{1 \le i \le k}} \frac{P(y^d, \vec{x}^d, F_{k-1}^*, f_k, \{y_i^d\}_{1 \le i \le k})}{P(\vec{x}^d, F_{k-1}^*, f_k)}\right) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} log\left(\sum_{\{y_i^d\}_{1 \le i \le k}} \frac{P(y^d | \{y^d\}_{1 \le i \le k}) \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k) P(\vec{x}^d) \prod_{i=1}^{k-1} (P(f_i)) P(f_k)}{P(\vec{x}^d, F_{k-1}^*, f_k)}\right) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} log\left(\sum_{\{y_i^d\}_{1 \le i \le k}} P(y^d | \{y^d\}_{1 \le i \le k}) \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k)\right) \end{split}$$

2.1 Precision case

$$\begin{split} f_k^* &= \arg\max_{f_k} \sum_{d \in D} log \left(I[y^d = \wedge_{i=1}^k] \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d | \vec{x}^d, f_k) \right) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} log \left(I[y^d = 1] \prod_{i=1}^{k-1} (P(y_i^d = 1 | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k) + I[y^d = 0] \sum_{\substack{\{y_i^d\}_{1 \leq i \leq k, \\ \exists y_i^d | \|y_i^d = 0}}} \prod_{i=1}^{k-1} (P(y_i^d | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k) + I[y^d = 0] (1 - \prod_{i=1}^{k-1} (P(y_i^d = 1 | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k) + I[y^d = 0] (1 - \prod_{i=1}^{k-1} (P(y_i^d = 1 | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k)) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = 1] log \left(\prod_{i=1}^{k-1} (P(y_i^d = 1 | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k) \right) \\ &+ I[y^d = 0] log \left(1 - \prod_{i=1}^{k-1} (P(y_i^d = 1 | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k) \right) \\ f_k^* &= \arg\max_{f_k} \sum_{d \in D} I[y^d = 1] \sum_{i=1}^{k-1} log \left(P(y_i^d = 1 | \vec{x}^d, f_i) \right) + log \left(P(y_k^d = 1 | \vec{x}^d, f_k) \right) \\ &+ I[y^d = 0] log \left(1 - \prod_{i=1}^{k-1} (P(y_i^d = 1 | \vec{x}^d, f_i)) P(y_k^d = 1 | \vec{x}^d, f_k) \right) \end{split}$$

2.2 Recall case

$$\begin{split} P(y^d|\{y^d\}_{1 \leq i \leq k}) &= I[y^d = \bigvee_{i=1}^k] = \begin{cases} y^d = 1 \Rightarrow \{y^d_i\}_{1 \leq i \leq k} : y^d_i | |y^d_i = 1 \\ y^d = 0 \Rightarrow \{y^d_i = 0\}_{1 \leq i \leq k} \end{cases} \\ f^*_k &= \underset{f_k}{\operatorname{arg}} \max \sum_{d \in D} \log \left(I[y^d = \bigvee_{i=1}^k] \prod_{i=1}^{k-1} (P(y^d_i | \vec{x}^d, f_i)) P(y^d_k | \vec{x}^d, f_k) \right) \\ f^*_k &= \underset{f_k}{\operatorname{arg}} \max \sum_{d \in D} \log \left(I[y^d = 0] \prod_{i=1}^{k-1} (P(y^d_i = 0 | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k) + I[y^d = 1] \sum_{\substack{\{y^d_i\}_{1 \leq i \leq k, k} \\ y^d_i | | y^d_i = 1}} \prod_{i=1}^{k-1} (P(y^d_i | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k) + I[y^d = 1] (1 - \prod_{i=1}^{k-1} (P(y^d_i = 0 | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k)) \\ f^*_k &= \underset{f_k}{\operatorname{arg}} \max \sum_{d \in D} I[y^d = 0] \log \left(\prod_{i=1}^{k-1} (P(y^d_i = 0 | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k) \right) \\ &+ I[y^d = 1] log \left(1 - \prod_{i=1}^{k-1} (P(y^d_i = 0 | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k) \right) \\ f^*_k &= \underset{f_k}{\operatorname{arg}} \max \sum_{d \in D} I[y^d = 0] \sum_{i=1}^{k-1} log \left(P(y^d_i = 0 | \vec{x}^d, f_i) \right) P(y^d_k = 0 | \vec{x}^d, f_k) \right) \\ &+ I[y^d = 1] log \left(1 - \prod_{i=1}^{k-1} (P(y^d_i = 0 | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k) \right) \\ &+ I[y^d = 1] log \left(1 - \prod_{i=1}^{k-1} (P(y^d_i = 0 | \vec{x}^d, f_i)) P(y^d_k = 0 | \vec{x}^d, f_k) \right) \end{aligned}$$