6.437 FINAL PROJECT — WRITE UP I

JUAN M ORTIZ

Problem I: A Bayesian Framework.

(a)

$$\begin{aligned} p_{\mathbf{y}|f}(\mathbf{y}|f) &= \mathbb{P}\left(\mathbf{x}_0 = f^{-1}(\mathbf{y}_0)\right) \prod_{j=1} \mathbb{P}\left(\mathbf{x}_j = f^{-1}(y_j) | \mathbf{x}_{j-1} = f^{-1}\right) \\ &= p_{f^{-1}(y)}(f^{-1}(\mathbf{y}_0)) \prod_{j=1}^{n} M_{f^{-1}(\mathbf{y})_j, f^{-1}(\mathbf{y})_{j-1}} \end{aligned}$$

(b)

$$\begin{aligned} p_{f|\mathbf{y}}(f|\mathbf{y}) &= \frac{p_{\mathbf{y}|f}(\mathbf{y}|f) p_f(f)}{p_{\mathbf{y}}(y)} \\ &= \frac{p_{f^{-1}(y)}(f^{-1}(\mathbf{y}_0)) \prod_{j=1}^{n} M_{f^{-1}(\mathbf{y})_j, f^{-1}(\mathbf{y})_{j-1}}}{\sum_{g \in \mathcal{F}} (p_{g^{-1}(y)}(g^{-1}(\mathbf{y}_0)) \prod_{j=1}^{n} M_{g^{-1}(\mathbf{y})_j, g^{-1}(\mathbf{y})_{j-1}})} \end{aligned}$$

Where $\mathcal{F}$ is the set of all permutations of $\mathcal{A}$. Thus, the MAP estimator is one that maximizes the above expression or, equivalently, its numerator. Thus

$$\hat{f}_{MAP} = \arg\max_{f \in \mathcal{F}} p_{f^{-1}(y)}(f^{-1}(\mathbf{y}_0)) \prod_{j=1}^{n} M_{f^{-1}(\mathbf{y})_j, f^{-1}(\mathbf{y})_{j-1}}$$

(c) Direct computation of the $\hat{f}_{MAP}$ is infeasable because it requires an optimization over a large, discrete, non-linear set. Computing expression on (b) would optimize over $\mathcal{F}$ which has a size of $|\mathcal{F}| = |\mathcal{A}|! = 28! \simeq 10^{29}$. Additionally, the constraints necessary to enforce that the $f \in \mathcal{F}$ is permutation will be hard to optimize over.

Problem 2: Markov Chain Monte Carlo method.

(a) After the first fixing one of the two cyphering functions $f_1 \in \mathcal{F}$, there are $\binom{|\mathcal{A}|}{2}$ possible $f_2 \in \mathcal{F}$ such that g differs in exactly two symbol assignments (i.e. those that swap two distinct element assignments in f_1). Thus, the probability that f_1 and f_2 differ in exactly two symbol assignments is given by:

$$\frac{\binom{\mathcal{A}}{2}}{|A|!} = \frac{1}{2(|A|-2)!}$$

(b) part b

(c) part c

1. Experimental results.