EE/Ma 127c Error-Correcting Codes - Homework Assignment 3

Ling Li, ling@cs.caltech.edu

April 25, 2001

3.1 *BCJR.* Let

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

be the transition matrix. That is, $T_{ij} = \Pr\{X_t = j | X_{t-1} = i\}$ is the probability that state j follows state i. Define 3×3 matrix $Q^{(k)}$ as (note that $X_{t-1} \to X_t \to Y_t$ is a Markov chain)

$$\begin{aligned} Q_{ij}^{(k)} &\stackrel{\text{def}}{=} \Pr\{X_t = j, Y_t = k | X_{t-1} = i\} \\ &= \Pr\{X_t = j | X_{t-1} = i\} \Pr\{Y_t = k | X_t = j\} \\ &= T_{ij} P_{jk}, \end{aligned}$$

i.e.,

$$Q^{(0)} = \begin{pmatrix} \frac{2}{5} & \frac{1}{20} & 0 \\ \frac{4}{15} & 0 & \frac{1}{15} \\ \frac{8}{15} & \frac{1}{30} & 0 \end{pmatrix}, \; Q^{(1)} = \begin{pmatrix} \frac{1}{20} & \frac{2}{5} & 0 \\ \frac{1}{30} & 0 & \frac{1}{15} \\ \frac{1}{15} & \frac{4}{15} & 0 \end{pmatrix}, \; Q^{(2)} = \begin{pmatrix} \frac{1}{20} & \frac{1}{20} & 0 \\ \frac{1}{30} & 0 & \frac{8}{15} \\ \frac{1}{15} & \frac{1}{30} & 0 \end{pmatrix}.$$

The probability that the Markov process takes a 'path' of states $\mathbf{x} = (x_1, x_2, x_3, x_4)$ and results $\mathbf{y} = (y_1, y_2, y_3, y_4)$, given the start state is x_0, is

$$w(\mathbf{x}, \mathbf{y}) = \prod_{t=1}^{4} \Pr\{X_t = x_t, Y_t = y_t | X_{t-1} = x_{t-1}\} = \prod_{t=1}^{4} Q_{x_{t-1} x_t}^{(y_t)}. \tag{1}$$

What we want to calculate is

$$\begin{aligned} \Pr\{X_i = j | \mathcal{E}\} &= \frac{\Pr\{X_i = j, \mathcal{E}\}}{\Pr\{\mathcal{E}\}} \\ &= K \sum_{\mathbf{x}: x_i = j} w(\mathbf{x}, \mathcal{E}), \end{aligned} \tag{2}$$

where $K^{-1} = \Pr\{\mathcal{E}\} = \sum_{\mathbf{x}} w(\mathbf{x}, \mathcal{E})$.

Now we can use the technique of forward-backward algorithm to get (2). Comparing (1) with the definition of the path weight in a trellis graph, we can define $W_t = Q^{(y_t^e)}$, i.e., $W_1 = Q^{(1)}$, $W_2 = W_3 = Q^{(2)}$, and $W_4 = Q^{(0)}$. Since the start state x_0 sticks to 0, $\alpha_0 = (1, 0, 0)$. Thus

from $\alpha_t = \alpha_{t-1}W_t$, we have $\alpha_1 = \left(\frac{1}{20}, \frac{2}{5}, 0\right)$, $\alpha_2 = \left(\frac{19}{1200}, \frac{1}{400}, \frac{16}{75}\right)$, $\alpha_3 = \left(\frac{1087}{72000}, \frac{569}{72000}, \frac{1}{750}\right)$, and $\alpha_4 = \left(\frac{4783}{540000}, \frac{1151}{1440000}, \frac{569}{1080000}\right)$. From $\beta_4 = (1,1,1)^T$ and $\beta_{t-1} = W_t\beta_t$, we get $\beta_3 = \left(\frac{9}{20}, \frac{1}{3}, \frac{17}{30}\right)^T$, $\beta_2 = \left(\frac{47}{1200}, \frac{571}{1800}, \frac{37}{900}\right)^T$, $\beta_1 = \left(\frac{1283}{72000}, \frac{2509}{108000}, \frac{89}{6750}\right)^T$, and $\beta_0 = \left(\frac{43993}{4320000}, \frac{1909}{1296000}, \frac{23921}{3240000}\right)^T$.
Now we have $K^{-1} = \Pr\{\mathcal{E}\} = \alpha_t\beta_t = \frac{43993}{4320000}$, and $\Pr\{X_i = j|\mathcal{E}\} = K \cdot \alpha_i(j) \cdot \beta_i(j)$:

$\Pr\{X_i = j\|\mathcal{E}\}$	$i = 1$	2	3	4
$j = 0$	$\frac{3849}{43993}$	$\frac{2679}{43993}$	$\frac{29349}{43993}$	$\frac{38264}{43993}$
1	$\frac{40144}{43993}$	$\frac{3426}{43993}$	$\frac{11380}{43993}$	$\frac{3453}{43993}$
2	0	$\frac{1024}{1189}$	$\frac{3264}{43993}$	$\frac{2276}{43993}$

Table above shows that the most probable value for X_i is $X_1 = 1$, $X_2 = 2$, $X_3 = X_4 = 0$.

3.2 *Bayesian network.*

(a) X_2 and X_3 are independent, since

$$\begin{aligned} p(x_2, x_3) &= \sum_{x_1,x_4} p(x_1, x_2, x_3, x_4) \\ &= \sum_{x_1,x_4} p(x_1)p(x_2)p(x_3|x_1)p(x_4|x_1, x_2) \\ &= p(x_2)\sum_{x_1} p(x_1)p(x_3|x_1)\sum_{x_4} p(x_4|x_1, x_2) \\ &= p(x_2)\sum_{x_1} p(x_1)p(x_3|x_1) \\ &= p(x_2)p(x_3). \end{aligned}$$

(b) The pair X_1 and X_2 are also independent, since

$$\begin{aligned} p(x_1, x_2) &= \sum_{x_3,x_4} p(x_1, x_2, x_3, x_4) \\ &= \sum_{x_3,x_4} p(x_1)p(x_2)p(x_3|x_1)p(x_4|x_1, x_2) \\ &= p(x_1)p(x_2)\sum_{x_3} p(x_3|x_1)\sum_{x_4} p(x_4|x_1, x_2) \\ &= p(x_1)p(x_2). \end{aligned}$$

(c) For any variable X in the Bayesian network, define a 'source set' of X, denoted by $S(X)$, as the collection of variables Z that there exits a path from Z to X. That is,

$$S(X) = \{Z|\exists P : Z \mapsto X\}.$$

By definition, $X \in S(X)$.
My guess for the general rule judging whether variables X and Y are independent is: *if $S(X) \cap S(Y) = \emptyset$, they are independent; otherwise they are dependent.*
Take the Bayesian network in the problem as an example. We have

$$S(X_1) = \{X_1\}, \; S(X_2) = \{X_2\}, \; S(X_3) = \{X_1, X_3\}, \; S(X_4) = \{X_1, X_2, X_4\}.$$

Thus by the general rule, there are two pairs, (X_1, X_2) and (X_2, X_3), that are independent.

3.3 *Matrix chain.* $A : p \times q$, $B : q \times r$, $C : r \times s$.

(a) To calculate one entry of AB, we need q scalar multiplications. Thus to computer AB, the number of scalar multiplications is pqr.

(b) When ABC is parenthesized as $(AB)C$, the number of scalar multiplications is $pqr + prs$, since the size of AB is $p \times r$. Similarly, that number of $A(BC)$ is $qrs + pqs$.

(c) When $p = 10$, $q = 100$, $r = 5$, and $s = 50$, computing ABC as $(AB)C$ needs

$$pqr + prs = pr(q + s) = 10 \times 5 \times 150 = 7500$$

scalar multiplications, while computing ABC as $A(BC)$ needs

$$qrs + pqs = qs(p + r) = 100 \times 50 \times 15 = 75000$$

scalar multiplications. Thus to compute ABC via $(AB)C$ is better.