Homework 5

Mehul Yesminkumar - 3824981440

1 Hidden Markov Model

1.1 Probability of an observed sequence

Used Forward algorithm.

$$\begin{aligned}
&\alpha_1(1) = \pi_1 * b_{1A} = 0.7 * 0.4 = 0.28 \\
&\alpha_1(2) = \pi_2 * b_{2A} = 0.3 * 0.2 = 0.06 \\
&a_{11} = 0.8 \qquad a_{21} = 0.4 \qquad a_{12} = 0.2 \qquad a_{22} = 0.6 \\
&b_{1A} = 0.4 \qquad b_{1C} = 0.1 \qquad b_{1G} = 0.4 \qquad b_{1T} = 0.1 \\
&b_{2A} = 0.2 \qquad b_{2C} = 0.3 \qquad b_{2G} = 0.2 \qquad b_{2T} = 0.3
\end{aligned}$$

$$\begin{aligned}
&\alpha_2(1) = b_{1G} * (\alpha_1(1) * a_{11} + \alpha_1(2) * a_{21}) = 0.0992 \\
&\alpha_2(2) = b_{2G} * (\alpha_1(1) * a_{12} + \alpha_1(2) * a_{22}) = 0.0184 \\
&\alpha_3(1) = b_{1C} * (\alpha_2(1) * a_{11} + \alpha_2(2) * a_{21}) = 0.008672 \\
&\alpha_3(2) = b_{2C} * (\alpha_2(1) * a_{12} + \alpha_2(2) * a_{22}) = 0.009264 \\
&\alpha_4(1) = b_{1G} * (\alpha_3(1) * a_{11} + \alpha_3(2) * a_{21}) = 0.00425728 \\
&\alpha_4(2) = b_{2G} * (\alpha_3(1) * a_{12} + \alpha_3(2) * a_{22}) = 0.00145856 \\
&\alpha_5(1) = b_{1T} * (\alpha_4(1) * a_{11} + \alpha_4(2) * a_{21}) = 0.0003989248 \\
&\alpha_5(2) = b_{2T} * (\alpha_4(1) * a_{12} + \alpha_4(2) * a_{22}) = 0.0005179776 \\
&\alpha_6(1) = b_{1A} * (\alpha_5(1) * a_{11} + \alpha_5(2) * a_{21}) = 0.00021053252 \\
&\alpha_6(2) = b_{2A} * (\alpha_5(1) * a_{12} + \alpha_5(2) * a_{21}) = 0.000078114304
\end{aligned}$$

Hence, $P(X_{1:6} = O_{1:6}; \theta) = \alpha_6(1) + \alpha_6(2) = \mathbf{0.000288646824}$

1.2 Explanation

Using Viterbi algorithm.
$\delta_1(1) = \pi_1 * P(x_1 = A|z_1 = 1) = 0.4 * 0.7 = 0.28$
$\delta_1(2) = \pi_2 * P(x_1 = A|z_1 = 2) = 0.3 * 0.2 = 0.06$

$$\begin{aligned}
\delta_2(1) &= \max(P(x_2 = G|z_2 = 1) * P(z_2 = 1|z_1 = 1) * \delta_1(1), P(x_2 = G|z_2 = 1) * P(z_2 = 1|z_1 = 2) * \delta_1(2)) \\
&= \max(0.4 * 0.8 * 0.28, 0.4 * 0.4 * 0.06) = \mathbf{0.0896} \\
\delta_2(2) &= \max(P(x_2 = G|z_2 = 2) * P(z_2 = 2|z_1 = 1) * \delta_1(1), P(x_2 = G|z_2 = 2) * P(z_2 = 2|z_1 = 2) * \delta_1(2)) \\
&= \max(0.2 * 0.2 * 0.28, 0.2 * 0.6 * 0.06) = 0.0112
\end{aligned}$$

$$\begin{aligned}
\delta_3(1) &= \max(P(x_3 = C|z_3 = 1) * P(z_3 = 1|z_2 = 1) * \delta_2(1), P(x_3 = C|z_3 = 1) * P(z_3 = 1|z_2 = 2) * \delta_2(2)) \\
&= \max(0.1 * 0.8 * 0.0896, 0.1 * 0.4 * 0.0112) = \mathbf{0.007168} \\
\delta_3(2) &= \max(P(x_3 = C|z_3 = 2) * P(z_3 = 2|z_2 = 1) * \delta_2(1), P(x_3 = C|z_3 = 2) * P(z_3 = 2|z_2 = 1) * \delta_2(2)) \\
&= \max(0.3 * 0.2 * 0.0896, 0.3 * 0.6 * 0.0112) = 0.005376
\end{aligned}$$

$$
\begin{aligned}
\delta_4(1) &= \max(P(x_4 = G|z_4 = 1) * P(z_4 = 1|z_3 = 1) * \delta_3(1), P(x_4 = G|z_4 = 1) * P(z_4 = 1|z_3 = 2) * \delta_3(2)) \\
&= \max(0.4 * 0.8 * 0.007168, 0.4 * 0.4 * 0.005376) = \mathbf{0.00229376} \\
\delta_4(2) &= \max(P(x_4 = G|z_4 = 2) * P(z_4 = 2|z_3 = 1) * \delta_3(1), P(x_4 = G|z_4 = 2) * P(z_4 = 2|z_3 = 1) * \delta_3(2)) \\
&= \max(0.2 * 0.2 * 0.007168, 0.2 * 0.6 * 0.005376) = 0.000645
\end{aligned}
$$

$$
\begin{aligned}
\delta_5(1) &= \max(P(x_5 = T|z_5 = 1) * P(z_5 = 1|z_4 = 1) * \delta_4(1), P(x_5 = T|z_5 = 1) * P(z_5 = 1|z_4 = 2) * \delta_4(2)) \\
&= \max(0.1 * 0.8 * 0.00229376, 0.1 * 0.4 * 0.000645) = \mathbf{0.0001835} \\
\delta_5(2) &= \max(P(x_5 = T|z_5 = 2) * P(z_5 = 2|z_4 = 1) * \delta_4(1), P(x_5 = T|z_5 = 2) * P(z_5 = 2|z_4 = 2) * \delta_4(2)) \\
&= \max(0.3 * 0.2 * 0.00229736, 0.3 * 0.6 * 0.000645) = 0.000137625
\end{aligned}
$$

$$
\begin{aligned}
\delta_6(1) &= \max(P(x_6 = A|z_6 = 1) * P(z_6 = 1|z_5 = 1) * \delta_5(1), P(x_6 = A|z_6 = 1) * P(z_6 = 1|z_5 = 2) * \delta_5(2)) \\
&= \max(0.4 * 0.8 * 0.0001835, 0.4 * 0.4 * 0.0001376256) = \mathbf{0.00005872} \\
\delta_6(2) &= \max(P(x_6 = A|z_6 = 2) * P(z_6 = 2|z_5 = 1) * \delta_5(1), P(x_6 = A|z_6 = 2) * P(z_6 = 2|z_5 = 2) * \delta_5(2)) \\
&= \max(0.2 * 0.2 * 0.0001835, 0.2 * 0.6 * 0.0001376256) = 0.000016515
\end{aligned}
$$

Hence, the most likely sequence is: $s_1 s_1 s_1 s_1 s_1 s_1$

1.3 Prediction

$$
\begin{aligned}
P(X_7 = x|X_{1:6} = O_{1:6}; \theta) &= \frac{P(X_7 = x, X_{1:6} = O_{1:6}; \theta)}{P(X_{1:6} = O_{1:6}; \theta)} \\
&= \frac{P(X_{1:7} = O_{1:7}; \theta)}{P(X_{1:6} = O_{1:6}; \theta)}
\end{aligned}
$$

Prediction depends on the numerator, as denominator is constant

Numerator $= \alpha_7(1) + \alpha_7(2)$

$\alpha_6(1) = 0.00021053252 \quad \alpha_6(2) = 0.000078114304$

$a_{11} = 0.8 \quad a_{21} = 0.4 \quad a_{12} = 0.2 \quad a_{22} = 0.6$

$b_{1A} = 0.4 \quad b_{1C} = 0.1 \quad b_{1G} = 0.4 \quad b_{1T} = 0.1$

$b_{2A} = 0.2 \quad b_{2C} = 0.3 \quad b_{2G} = 0.2 \quad b_{2T} = 0.3$

When $X_7 = A$ **:**

$$
\begin{aligned}
\alpha_7(1) &= b_{1A} * (a_{11} * \alpha_6(1) + a_{21} * \alpha_6(2)) \\
&= 0.4(0.8 * 0.00021053252 + 0.4 * 0.000078114304) = 0.00007986869504 \\
\alpha_7(2) &= b_{2A} * (a_{12} * \alpha_6(1) + a_{22} * \alpha_6(2)) \\
&= 0.2(0.2 * 0.00021053252 + 0.6 * 0.000078114304) = 0.00001779501728
\end{aligned}
$$

$$\implies X_7 = A \qquad P(X_{1:7} = O_{1:7}; \theta) = \alpha_7(1) + \alpha_7(2) = \mathbf{0.00009766371229}$$

When $X_7 = C$ **:**

$$
\begin{aligned}
\alpha_7(1) &= b_{1C} * (a_{11} * \alpha_6(1) + a_{21} * \alpha_6(2)) \\
&= 0.1(0.8 * 0.00021053252 + 0.4 * 0.000078114304) = 0.00001996717376 \\
\alpha_7(2) &= b_{2C} * (a_{12} * \alpha_6(1) + a_{22} * \alpha_6(2)) \\
&= 0.3(0.2 * 0.00021053252 + 0.6 * 0.000078114304) = 0.00002669252592
\end{aligned}
$$

$$\implies X_7 = C \qquad P(X_{1:7} = O_{1:7}; \theta) = \alpha_7(1) + \alpha_7(2) = \mathbf{0.00004665969968}$$

When $X_7 = G$ **:**

$$\begin{aligned}
\alpha_7(1) &= b_{1G} * (a_{11} * \alpha_6(1) + a_{21} * \alpha_6(2)) \\
&= 0.4(0.8 * 0.00021053252 + 0.4 * 0.000078114304) = 0.00007986869504 \\
\alpha_7(2) &= b_{2G} * (a_{12} * \alpha_6(1) + a_{22} * \alpha_6(2)) \\
&= 0.2(0.2 * 0.00021053252 + 0.6 * 0.000078114304) = 0.00001779501728
\end{aligned}$$

$$\implies \quad X_7 = G \qquad P(X_{1:7} = O_{1:7}; \theta) = \alpha_7(1) + \alpha_7(2) = \mathbf{0.00009766371229}$$

When $X_7 = T$ **:**

$$\begin{aligned}
\alpha_7(1) &= b_{1T} * (a_{11} * \alpha_6(1) + a_{21} * \alpha_6(2)) \\
&= 0.1(0.8 * 0.00021053252 + 0.4 * 0.000078114304) = 0.00001996717376 \\
\alpha_7(2) &= b_{2T} * (a_{12} * \alpha_6(1) + a_{22} * \alpha_6(2)) \\
&= 0.3(0.2 * 0.00021053252 + 0.6 * 0.000078114304) = 0.00002669252592
\end{aligned}$$

$$\implies \quad X_7 = T \qquad P(X_{1:7} = O_{1:7}; \theta) = \alpha_7(1) + \alpha_7(2) = \mathbf{0.00004665969968}$$

Thus we get, $P(X_{1:7} = O_{1:7}; \theta)$ are maximum when $X_7 = A$ or $X_7 = G$
Thus, the prediction for the next observed values are either A or G.