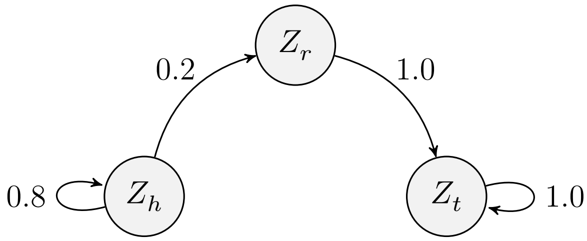
Midterm exam, Genome Informatics 20221025 *Write your name on each sheet.*



	Z_h	Z_r	Z_t
H	0.9	0.5	0.1
T	0.1	0.5	0.9

Emission probabilities of {H,T} for the 3 states.

The model is a "left-to-right" model, always starting in state Z_h.
The observed output sequence is $X =$ HHHTTT.

Notation

Here I suggest some notation to use when showing your work.
Let $\mathrm{P}[X_{1-i}]$ be a shorthand for the first i letters of the output, e.g. $\mathrm{P}[X_{1-4}] =$ HHHT.
Use Q to denote a hidden state sequence, Q_i is the hidden state at time step i. For brevity let Q_{ik} denote $Q_i = Z_k$, e.g. Q_{4r} means the model was in state Z_r in time step 4.
To further simplify notation you may use $\mathrm{P}[h \to r]$ as a shorthand for the transition probability $\mathrm{P}[Q_{i+1} = Z_r | Q_i = Z_h]$.

You will work with numerical values.
We will consider any answer correct in the first 3 digits to be correct.
But please write as many digits as possible, as this can help us trace your work.
(爲了幫助我們確認你的計算，請盡量不要四捨五入)

Problem 1.
What is the probability of the output X given the model, $\mathrm{P}[X|\lambda, \texttt{len}(X) = 6]$?

Solution: We use the "forward" algorithm.
1st step. Always starts in state Z_h.

$$\mathrm{P}[X_1 = \texttt{H}] = \mathrm{P}[X_1 = \texttt{H}, Q_1 = Z_h] \;\; = \;\; 0.9.$$

2nd step. State Z_t is still unreachable, so two possibilities for Q_{1-2}, $Z_t Z_t$ and $Z_t Z_r$

$$\begin{aligned}
\mathrm{P}[X_{1-2}, Q_{2h}] &= \mathrm{P}[X_1]\,\mathrm{P}[h \to h]\,\mathrm{P}[X_2|Q_{2h}] = (0.9)(0.8)(0.9) = \mathbf{0.648} \\
\mathrm{P}[X_{1-2}, Q_{2r}] &= \mathrm{P}[X_1]\,\mathrm{P}[h \to r]\,\mathrm{P}[X_2|Q_{2r}] = (0.9)(0.2)(0.5) = \mathbf{0.09}
\end{aligned}$$

Thus
3rd step. In this step the possiblities are:

$$\begin{aligned}
\mathrm{P}[X_{1-3}, Q_{3h}] &= \mathrm{P}[X_{1-2}, Q_{2h}]\,\mathrm{P}[h \to h]\,\mathrm{P}[X_3|Q_{3h}] = (0.648)(0.8)(0.9) &= \mathbf{0.46656} \\
\mathrm{P}[X_{1-3}, Q_{3r}] &= \mathrm{P}[X_{1-2}, Q_{2h}]\,\mathrm{P}[h \to r]\,\mathrm{P}[X_3|Q_{3r}] = (0.648)(0.2)(0.5) &= \mathbf{0.0648} \\
\mathrm{P}[X_{1-3}, Q_{3t}] &= \mathrm{P}[X_{1-2}, Q_{2r}]\,\mathrm{P}[r \to t]\,\mathrm{P}[X_3|Q_{3t}] = (0.09)(1.0)(0.1) &= \mathbf{0.009}
\end{aligned}$$

4th step.

$$\begin{aligned}
\mathrm{P}[X_{1-4}, Q_{4h}] &= \mathrm{P}[X_{1-3}, Q_{3h}]\,\mathrm{P}[h \to h]\,\mathrm{P}[X_4|Q_{4h}] &= (0.46656)(0.8)(0.1) = &\quad \mathbf{0.0373248} \\
\mathrm{P}[X_{1-4}, Q_{4r}] &= \mathrm{P}[X_{1-3}, Q_{3h}]\,\mathrm{P}[h \to r]\,\mathrm{P}[X_4|Q_{4r}] &= (0.46656)(0.2)(0.5) = &\quad \mathbf{0.046656} \\
\mathrm{P}[X_{1-4}, Q_{4t}] &= \mathrm{P}[X_{1-3}, Q_{3r}]\,\mathrm{P}[r \to t]\,\mathrm{P}[X_4|Q_{4t}] &= (0.0648)(1.0)(0.9) = &\quad 0.05832 \\
&\quad + \mathrm{P}[X_{1-3}, Q_{3t}]\,\mathrm{P}[t \to t]\,\mathrm{P}[X_4|Q_{4t}] &= (0.009)(1.0)(0.9) \quad &\quad +0.00810 \\
&& = &\quad \mathbf{0.06642}
\end{aligned}$$

5th step.

$$\begin{aligned}
\mathrm{P}[X_{1-5}, Q_{5h}] &= \mathrm{P}[X_{1-4}, Q_{4h}]\,\mathrm{P}[h \to h]\,\mathrm{P}[X_5|Q_{5h}] &= (0.0373248)(0.8)(0.1) = &\quad \mathbf{0.00298598} \\
\mathrm{P}[X_{1-5}, Q_{5r}] &= \mathrm{P}[X_{1-4}, Q_{4h}]\,\mathrm{P}[h \to r]\,\mathrm{P}[X_5|Q_{5r}] &= (0.0373248)(0.2)(0.5) = &\quad \mathbf{0.00373248} \\
\mathrm{P}[X_{1-5}, Q_{5t}] &= \mathrm{P}[X_{1-4}, Q_{4r}]\,\mathrm{P}[r \to t]\,\mathrm{P}[X_5|Q_{5t}] &= (0.046656)(1.0)(0.9) = &\quad 0.0419904 \\
&\quad + \mathrm{P}[X_{1-4}, Q_{4t}]\,\mathrm{P}[t \to t]\,\mathrm{P}[X_5|Q_{5t}] &= (0.06642)(1.0)(0.9) \quad &\quad +0.0597780 \\
&& = &\quad \mathbf{0.1016784}
\end{aligned}$$

6th step.

$$
\begin{aligned}
\mathrm{P}[X_{1-6}, Q_{6h}] &= \mathrm{P}[X_{1-5}, Q_{5h}]\,\mathrm{P}[h \to h]\,\mathrm{P}[X_6|Q_{6h}] &&= (0.00298598)(0.8)(0.1) = &&\mathbf{0.000238878} \\
\mathrm{P}[X_{1-6}, Q_{6r}] &= \mathrm{P}[X_{1-5}, Q_{5h}]\,\mathrm{P}[h \to r]\,\mathrm{P}[X_6|Q_{6r}] &&= (0.00298598)(0.2)(0.5) = &&\mathbf{0.000298598} \\
\mathrm{P}[X_{1-6}, Q_{6t}] &= \mathrm{P}[X_{1-5}, Q_{5r}]\,\mathrm{P}[r \to t]\,\mathrm{P}[X_6|Q_{6t}] &&= (0.00373248)(1.0)(0.9) = &&0.00335923 \\
&\quad + \mathrm{P}[X_{1-5}, Q_{5t}]\,\mathrm{P}[t \to t]\,\mathrm{P}[X_6|Q_{6t}] &&= (0.1016784)(1.0)(0.9) &&+0.09151060 \\
& && = &&\mathbf{0.09486983}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P}[X_{1-6}] &= \mathrm{P}[X_{1-6}, Q_{6h}] + \mathrm{P}[X_{1-6}, Q_{6r}] + \mathrm{P}[X_{1-6}, Q_{6t}] \\
&= 0.000238878 + 0.000298598 + 0.09486983 = 0.0954073 \approx 0.0954
\end{aligned}
$$

Problem 2.
What is the maximum likelihood state sequence (Viterbi decoding)?
What is the likelihood of that sequence?

In other words, compute $\mathrm{P}[Q^*|X]$,
where Q^* denotes the maximum likelihood path: $\arg\max_{Q \in \{\mathbf{Q_{1-6}}\}} \mathrm{P}[Q|X_{1-6}]$.
and $\mathbf{Q_{1-6}}$ denotes the set of all state sequences of length 6.

For intermediate calculations, use δ_{ik} to denote: $\max_{Q \in \{\mathbf{Q_{1-ik}}\}} \mathrm{P}[Q|X_{1-i}]$,
where $\mathbf{Q_{1-ik}}$ denotes the set of all state sequences of length i, ending in state Z_k.

Solution: Following Rabiner, we will find it convenient to compute $\mathrm{P}[Q, X] \propto \mathrm{P}[Q|X]$.
1st step. Always starts in state Z_h, so $\delta_{1h} = P[X_1] = \mathbf{0.9}$, $\delta_{1r} = \delta_{1t} = 0$.

2nd step. HMM either stays in state Z_h or advances to Z_r.

$$\delta_{2h} = \delta_{1h}\,\mathrm{P}[X_2|Q_{2h}]\,\mathrm{P}[h \to h] = (0.9)(0.9)(0.8) = \qquad \mathbf{0.648}$$
$$\delta_{2r} = \delta_{1h}\,\mathrm{P}[\mathbf{X}_2|\mathbf{Q}_{2r}]\,\mathrm{P}[\mathbf{h} \to \mathbf{r}] = (0.9)(0.5)(0.2) = \qquad 0.090$$

3rd step.

$$\delta_{3h} = \delta_{2h}\,\mathrm{P}[\mathbf{X}_3|\mathbf{Q}_{3h}]\,\mathrm{P}[\mathbf{h} \to \mathbf{h}] = (0.648)(0.9)(0.8) = \qquad 0.46656$$
$$\delta_{3r} = \delta_{2h}\,\mathrm{P}[\mathbf{X}_3|\mathbf{Q}_{3r}]\,\mathrm{P}[\mathbf{h} \to \mathbf{r}] = (0.648)(0.5)(0.2) = \qquad \mathbf{0.0648}$$
$$\delta_{3t} = \delta_{2r}\,\mathrm{P}[X_3|Q_{3t}]\,\mathrm{P}[r \to t] = (0.09)(0.1)(1.0) = \qquad 0.0090$$

4th step.

$$\delta_{4h} = \delta_{3h}\,\mathrm{P}[X_4|Q_{4h}]\,\mathrm{P}[h \to h] = (0.46656)(0.1)(0.8) = \qquad 0.0373248$$
$$\delta_{4r} = \delta_{3h}\,\mathrm{P}[X_4|Q_{4r}]\,\mathrm{P}[h \to r] = (0.46656)(0.5)(0.2) = \qquad 0.046656$$
$$\delta_{4t}|Q_{3r} = \delta_{3r}\,\mathrm{P}[X_4|Q_{4t}]\,\mathrm{P}[r \to t] = (0.0648)(0.9)(1.0) = \qquad \mathbf{0.05832}$$
$$\delta_{4t}|Q_{3t} = \delta_{3t}\,\mathrm{P}[X_4|Q_{4t}]\,\mathrm{P}[t \to t] = (0.009)(0.9)(1.0) = \qquad 0.00810$$

Where $\delta_{4t}|Q_{3r}$ denotes the maximum likelihood path under the constraint that it includes $Q_{3r}Q_{4t}$.

5th step.

$$\delta_{5h} = \delta_{4h}\,\mathrm{P}[X_5|Q_{5h}]\,\mathrm{P}[h \to h] = (0.0373248)(0.1)(0.8) = \qquad 0.00298598$$
$$\delta_{5r} = \delta_{4h}\,\mathrm{P}[X_5|Q_{5r}]\,\mathrm{P}[h \to r] = (0.0373248)(0.5)(0.2) = \qquad 0.00373248$$
$$\delta_{5t}|Q_{4r} = \delta_{4r}\,\mathrm{P}[X_5|Q_{5t}]\,\mathrm{P}[r \to t] = (0.046656)(0.9)(1.0) = \qquad 0.0419904$$
$$\delta_{5t}|Q_{4t} = \delta_{4t}\,\mathrm{P}[X_5|Q_{5t}]\,\mathrm{P}[t \to t] = (0.05832)(0.9)(1.0) = \qquad \mathbf{0.0524880}$$

6th step.

$$\begin{aligned}
\delta_{6h} =& \delta_{5h}\,\mathrm{P}[X_6|Q_{6h}]\,\mathrm{P}[h\to h] &&= (0.00298598)(0.1)(0.8) = 0.000238878\\
\delta_{6r} =& \delta_{5h}\,\mathrm{P}[X_6|Q_{6r}]\,\mathrm{P}[h\to r] &&= (0.00298598)(0.5)(0.2) = 0.000298598\\
\delta_{6t}|Q_{5r} =& \delta_{5r}\,\mathrm{P}[X_6|Q_{6t}]\,\mathrm{P}[r\to t] &&= (0.00373248)(0.9)(1.0) = 0.00335923\\
\delta_{6t}|Q_{5t} =& \delta_{5t}\,\mathrm{P}[X_6|Q_{6t}]\,\mathrm{P}[t\to t] &&= (0.0524880)(0.9)(1.0) = \mathbf{0.04723920}
\end{aligned}$$

The Viterbi path is: $Q^* = Z_h Z_h Z_r Z_t Z_t Z_t$
$\mathrm{P}[Q^*, X] = 0.0472392$ so,

$$\mathrm{P}[Q^*|X] = \mathrm{P}[Q^*, X]/\mathrm{P}[X] = 0.0472392/0.0954073 = 0.4951319238674609 \approx 0.495$$

Problem 3.
What is the posterior decoding?
In other words, what is the state sequence: $Q^M = Q_1^M Q_2^M \cdots Q_6^M$
where $Q_i^M \stackrel{\text{def}}{=} \max_{k \in \{Z_h, Z_r, Z_t\}} \mathrm{P}[Q_i = Z_k | X]$.

For each position i and state $k \in \{Z_h, Z_r, Z_t\}$, give the probability $\mathrm{P}[Q_i = Z_k | X]$.
Hint: Note that the fact that some transitions have probability one can be used to simplify the backward algorithm computation.

Solution: To solve this, we can combine probabilities computed from the "forward" and "backward" algorithms. We already did the forward algorithm in a previous problem. Here we use the backward algorithm to compute: $\mathrm{P}[X_{i+1}, \cdots X_6 | Q_{ik}]$. Using β_{ik} to denote $\mathrm{P}[X_{i+1}, \cdots, X_6 | Q_{ik}]$
As the hint in the question states, the deterministic transition from Z_r to Z_t can help us. In particular, note that if we are in *either* state Z_r to Z_t in step i, we will definitely be in state Z_t in step. Since β_{ik} is affected only by the output *after* step i, we have $\beta_{ir} \equiv \beta_{it}$. Therefore we omit β_{it} in the calculations below.

5th step (counting backwards)

$$
\begin{aligned}
\mathrm{P}[X_6 | Q_{5r}] &= \mathrm{P}[r \to t]\, \mathrm{P}[X_6 | Q_{6t}] &&= (1.0)(0.9) = \mathbf{0.90} \\
\mathrm{P}[X_6 | Q_{5h}, Q_{6h}] &= \mathrm{P}[h \to h]\, \mathrm{P}[X_6 | Q_{6h}] &&= (0.8)(0.1) = 0.08 \\
\mathrm{P}[X_6 | Q_{5h}, Q_{6r}] &= \mathrm{P}[h \to r]\, \mathrm{P}[X_6 | Q_{6h}] &&= (0.2)(0.5) = 0.10 \\
\mathrm{P}[X_6 | Q_{5h}] &= \mathrm{P}[X_6 | Q_{5h}, Q_{6h}] + \mathrm{P}[X_6 | Q_{5h}, Q_{6r}] &&= 0.08 + 0.10 = \mathbf{0.18}
\end{aligned}
$$

4th step

$$
\begin{aligned}
\mathrm{P}[X_{5-6}|Q_{4r}] &= \mathrm{P}[r \to t]\,\mathrm{P}[X_5|Q_{5t}]\,\mathrm{P}[X_6|Q_{5t}] &&= (1.0)(0.9)(0.9) = \mathbf{0.81} \\
\mathrm{P}[X_{5-6}|Q_{4h}, Q_{5r}] &= \mathrm{P}[h \to r]\,\mathrm{P}[X_5|Q_{5r}]\,\mathrm{P}[X_6|Q_{5r}] &&= (0.2)(0.5)(0.9) = 0.09 \\
\mathrm{P}[X_{5-6}|Q_{4h}, Q_{5h}] &= \mathrm{P}[h \to h]\,\mathrm{P}[X_5|Q_{5h}]\,\mathrm{P}[X_6|Q_{5h}] &&= (0.8)(0.1)(0.18) = 0.0144 \\
\mathrm{P}[X_{5-6}|Q_{4h}] &= \mathrm{P}[X_{5-6}|Q_{4h}, Q_{5r}] + \mathrm{P}[X_{5-6}|Q_{4h}, Q_{5h}] &&= 0.09 + 0.0144 = \mathbf{0.1044}
\end{aligned}
$$

3th step

$$
\begin{aligned}
\mathrm{P}[X_{4-6}|Q_{3r}] &= \mathrm{P}[r \to t]\,\mathrm{P}[X_4|Q_{4t}]\,\mathrm{P}[X_{5-6}|Q_{4t}] &&= (1.0)(0.9)(0.81) = \mathbf{0.729} \\
\mathrm{P}[X_{4-6}|Q_{3h}, Q_{4r}] &= \mathrm{P}[h \to r]\,\mathrm{P}[X_4|Q_{4r}]\,\mathrm{P}[X_{5-6}|Q_{4r}] &&= (0.2)(0.5)(0.81) = 0.081 \\
\mathrm{P}[X_{4-6}|Q_{3h}, Q_{4h}] &= \mathrm{P}[h \to h]\,\mathrm{P}[X_4|Q_{4h}]\,\mathrm{P}[X_{5-6}|Q_{4h}] &&= (0.8)(0.1)(0.1044) = 0.008352 \\
\mathrm{P}[X_{4-6}|Q_{3h}] &= \mathrm{P}[X_{4-6}|Q_{3h}, Q_{4t}] + \mathrm{P}[X_{4-6}|Q_{3h}, Q_{4h}] &&= 0.081 + 0.008352 = \mathbf{0.089352}
\end{aligned}
$$

2nd step

$$
\begin{aligned}
\mathrm{P}[X_{3-6}|Q_{2r}] &= \mathrm{P}[r \to t]\,\mathrm{P}[X_3|Q_{3t}]\,\mathrm{P}[X_{4-6}|Q_{3r}] &&= (1.0)(0.1)(0.729) = \mathbf{0.0729} \\
\mathrm{P}[X_{3-6}|Q_{2h}, Q_{3r}] &= \mathrm{P}[h \to r]\,\mathrm{P}[X_3|Q_{3r}]\,\mathrm{P}[X_{4-6}|Q_{3r}] &&= (0.2)(0.5)(0.729) = 0.0729 \\
\mathrm{P}[X_{3-6}|Q_{2h}, Q_{3h}] &= \mathrm{P}[h \to h]\,\mathrm{P}[X_3|Q_{3h}]\,\mathrm{P}[X_{4-6}|Q_{3h}] &&= (0.8)(0.9)(0.089352) = 0.0643334 \\
\mathrm{P}[X_{3-6}|Q_{2h}] &= \mathrm{P}[X_{3-6}|Q_{2h}, Q_{3t}] + \mathrm{P}[X_{3-6}|Q_{2h}, Q_{3h}] &&= 0.0729 + 0.0643334 = \mathbf{0.137233}
\end{aligned}
$$

1st step

$$
\begin{aligned}
\mathrm{P}[X_{2-6}|Q_{1r}] &= \mathrm{P}[r \to t]\,\mathrm{P}[X_2|Q_{2t}]\,\mathrm{P}[X_{3-6}|Q_{2r}] &&= (1.0)(0.1)(0.0729) = \mathbf{0.00729} \\
\mathrm{P}[X_{2-6}|Q_{1h}, Q_{2r}] &= \mathrm{P}[h \to r]\,\mathrm{P}[X_2|Q_{2r}]\,\mathrm{P}[X_{3-6}|Q_{2r}] &&= (0.2)(0.5)(0.0729) = 0.00729 \\
\mathrm{P}[X_{2-6}|Q_{1h}, Q_{2h}] &= \mathrm{P}[h \to h]\,\mathrm{P}[X_2|Q_{2h}]\,\mathrm{P}[X_{3-6}|Q_{2h}] &&= (0.8)(0.9)(0.137233) = 0.0988078 \\
\mathrm{P}[X_{2-6}|Q_{1h}] &= \mathrm{P}[X_{2-6}|Q_{1h}, Q_{2t}] + \mathrm{P}[X_{2-6}|Q_{1h}, Q_{2h}] &&= 0.00729 + 0.0988078 = \mathbf{0.106098}
\end{aligned}
$$

Now we multiply the forward and backward probabilities to obtain $\mathrm{P}[Q_i = Z_k|X]$ for all combinations of step i and state k.

step i	state k	α_{ik}	β_{ik}	$\alpha_{ik}\ \beta_{ik}$
1	Z_h	0.9	0.106098	**0.0954882**
	Z_r	0	0.00729	0
	Z_t	0	0.00729	0
2	Z_h	0.648	0.137233	**0.088927**
	Z_r	0.09	0.0729	0.006561
	Z_t	0	0.0729	0
3	Z_h	0.46656	0.089352	0.0416881
	Z_r	0.0648	0.729	**0.0472392**
	Z_t	0.009	0.729	0.006561
4	Z_h	0.0373248	0.1044	0.00389671
	Z_r	0.046656	0.81	0.0377914
	Z_t	0.06642	0.81	**0.0538002**
5	Z_h	0.00298598	0.18	0.000537476
	Z_r	0.00373248	0.9	0.00335923
	Z_t	0.1016784	0.9	**0.0915106**
6	Z_h	0.000238878	1	0.000238878
	Z_r	0.000298598	1	0.000298598
	Z_t	0.09486983	1	**0.09486983**

The posterior decoding is obtained by concatenating the maximum probability state in each time step (shown in bold in the preceding table).

We obtain $Z_h Z_h Z_r Z_t Z_t Z_t$. Rather boring result in that in this case it happens to be the same as the Viterbi path.