1. In *Viterbi training*, the most probable path is used, as opposed to using the entire forward and backward tables as in Baum-Welch EM. Suppose we have an HMM with hidden state space $S = \{1, 2\}$ representing two weighted coins, and emission state space $\Sigma = \{H, T\}$ representing the observed outcomes of coin tosses. At a certain iteration, suppose you have obtained the following transition and emission probabilities

$$\begin{pmatrix} a_{11} = \frac{1}{2} & a_{12} = \frac{1}{2} \\ a_{21} = \frac{1}{5} & a_{22} = \frac{4}{5} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} e_1(H) = \frac{2}{3} & e_1(T) = \frac{1}{3} \\ e_2(H) = \frac{1}{4} & e_2(T) = \frac{3}{4} \end{pmatrix}$$

   (a) Now we want to find the most likely path (Viterbi path) of hidden states for a given dataset using dynamic programming. Let $v_t(k)$ be the probability of the most probable path that ends in hidden state k at position t in the data. What is the base case ($t = 1$) and recursive call?

   (b) Given the observed sequence $x = (H, T, H)$ and the probabilities above, fill in the table for v below, then use backpointers to find the most likely sequence of hidden states.

	H	T	H
1			
2			

   (c) Now suppose you have a longer observed sequence, and have completed the Viterbi decoding to obtain the most probable hidden state sequence below. What are the updated transition and emission probabilities a_{kl} and $e_k(\sigma)$?

hidden state	2	1	2	1	1	1	2	2	2	2	2	2	2	2	1	2	2
emitted state	T	H	H	H	T	H	T	T	H	T	T	T	T	H	H	T	T

2. The Viterbi algorithm finds the most probable path ending in state k at position t in the sequence x, i.e.

$$v_t(k) = \max_{q_{1:t-1}} \left\{ \mathbb{P}(x_{1:t}, Q_{1:t-1} = q_{1:t-1}, Q_t = k) \right\}.$$

   Prove the Viterbi recursion from question 1.