S	$\pi(S)$	T_{SA}	T_{SB}	E_{S0}	E_{S1}
A	0.99	0.99	0.01	0.8	0.2
B	0.01	0.01	0.99	0.1	0.9

d) For the sequence $O_1 = 0, O_2 = 1, O_3 = 0$:
Using $\alpha_1(s) = E_{s o_1} \pi(s)$,
$\forall t \geq 2$, $\alpha_t(s) = E_{s o_t} \sum_{s' \in S} T_{s's} \alpha_{t-1}(s')$

Time t	1	2	3
Obs O_t	0	1	0
$\alpha_t(A)$	0.792	0.156818	0.124264
$\alpha_t(B)$	0.001	0.008019	0.000950699

$\alpha_1(A) : 0.8 \times 0.99 = 0.792$
$\alpha_1(B) : 0.1 \times 0.01 = 0.001$
$\alpha_2(A) : 0.2(0.792 \times 0.99 + 0.001 \times 0.01)$
$= 0.156818$
$\alpha_2(B) = 0.9(0.792 \times 0.01 + 0.001 \times 0.99)$
$= 0.008019$
$\alpha_3(A) = 0.8(0.156818 \times 0.99 + 0.008019 \times 0.01)$
$= 0.124264$
$\alpha_3(B) = 0.1(0.156818 \times 0.01 + 0.008019 \times 0.99)$
$= 0.000950699$

$p(O_1 = 0, O_2 = 1, O_3 = 0)$
$= \sum_{s \in S} \alpha_T(s)$
$= \alpha_3(A) + \alpha_3(B)$
$= 0.124264 + 0.000950699$
$= 0.125214699$

e) For the same sequence, using the backward algo:
Using $\beta_T(s) = 1$
$\beta_t(s) = \sum_{s' \in S} T_{ss'} E_{s' o_{t+1}} \beta_{t+1}(s')$

Time t	1	2	3
Obs O_t	0	1	0
$\beta_t(A)$	0.157977	0.793	1
$\beta_t(B)$	0.096923	0.107	1

$\beta_3(A) = 1$
$\beta_3(B) = 1$
$\beta_2(A) = 0.99 \times 0.8 \times 1 + 0.01 \times 0.1 \times 1$
$= 0.793$
$\beta_2(B) = 0.01 \times 0.08 \times 1 + 0.99 \times 0.01 \times 1$
$= 0.107$
$\beta_1(A) = 0.99 \times 0.02 \times 0.793 + 0.01 \times 0.9 \times 0.107$
$= 0.157977$
$\beta_1(B) = 0.01 \times 0.2 \times 0.793 + 0.99 \times 0.9 \times 0.107$
$= 0.096923$

$p(O_1 = 0, O_2 = 1, O_3 = 0)$
$= 0.792 \times 0.157977 + 0.001 \times 0.096923$
$= 0.125214707$

f) Yes, the two results from the forward & backward algorithms agree.
To find the most likely sequence of values of these states, we compare $\alpha_t(A) \times \beta_t(A)$ with $\alpha_t(B) \times \beta_t(B)$. If $\alpha_t(A) \cdot \beta_t(A)$ is greater, most likely state at time t will be A & vice versa.
The comparisons are:

$\alpha_1(A)\beta_1(A) = 0.792 \times 0.159977 = 0.1257178$

$\alpha_1(B)\beta_1(B) = 0.001 \times 0.096923 = 0.000096923$

For $t=1$, state is A

$\alpha_2(A)\beta_2(A) = 0.156818 \times 0.793 = 0.1243567$

$\alpha_2(B)\beta_2(B) = 0.008019 \times 0.107 = 0.000858033$

For $t=2$, state is A

$\alpha_3(A)\beta_3(A) = 0.124264 \times 1 = 0.124264$

$\alpha_3(B)\beta_3(B) = 0.000950699 \times 1 = 0.000950699$

For $t=3$, state is A

Therefore the most likely sequence is A, A, A.

g) The Viterbi sequence is computed as follows:

$V_1(A) = 0.99 \times 0.8 = 0.792$

$V_1(B) = 0.01 \times 0.1 = 0.001$

For $t=1$, state is A

$V_2(A) = 0.792 \times 0.99 \times 0.2 = 0.156816$

$V_2(B) = 0.792 \times 0.01 \times 0.9 = 0.007128$

For $t=2$, state is A

$V_3(A) = 0.156816 \times 0.99 \times 0.8 = 0.1241963$

$V_3(B) = 0.156816 \times 0.01 \times 0.1 = 0.000156816$

For $t=3$, state is A

Therefore, the most likely sequence using Viterbi Algorithm is A, A, A

h) To find the most likely sequence of values for each state separately, we would just have to compare products of $\pi(S) \cdot E_{So_t}$.

For sequence 0, 1, 0:

$t=1$, $p(A): 0.8 \times 0.99 = 0.792 \rightarrow A$
$p(B): 0.1 \times 0.01 = 0.001$

$t=2$, $p(A): 0.2 \times 0.99 = 0.198 \rightarrow A$
$p(B): 0.9 \times 0.01 = 0.009$

$t=3$, $p(A): 0.8 \times 0.99 = 0.792 \rightarrow A$
$p(B): 0.1 \times 0.01 = 0.001$

The resulting sequence A, A, A is equivalent to what we observed in part (f) & (g). This is because $\pi(B)$ is really low in this case. However this does not hold in general. In any case where $\pi(B)$ is moderately higher, the sequence would possibly change.