Problem 1:

$$P(c \mid d, x, \neg s) = \alpha P(c, d, x, \neg s)$$
$$= \alpha \sum_{p} P(c, d, x, \neg s, p)$$
$$=$$
$$\alpha P(p) P(\neg s) P(c \mid p, \neg s) P(x \mid c) P(d \mid c) +$$
$$\alpha P(\neg p) P(\neg s) P(c \mid \neg p, \neg s) P(x \mid c) P(d \mid c)$$
$$=$$
$$\alpha \begin{pmatrix} 0.1 \times 0.7 \times 0.02 \times 0.9 \times 0.65 + \\ 0.9 \times 0.7 \times 0.001 \times 0.9 \times 0.65 \end{pmatrix} = \alpha \times 0.000855855$$

$$P(\neg c \mid d, x, \neg s) =$$
$$\alpha P(p) P(\neg s) P(\neg c \mid p, \neg s) P(x \mid \neg c) P(d \mid \neg c) +$$
$$\alpha P(\neg p) P(\neg s) P(\neg c \mid \neg p, \neg s) P(x \mid \neg c) P(d \mid \neg c) =$$
$$\alpha \begin{bmatrix} 0.1 \times 0.7 \times 0.98 \times 0.2 \times 0.3 + \\ 0.9 \times 0.7 \times 0.999 \times 0.2 \times 0.3 \end{bmatrix} = \alpha \times 0.0418782$$

Thus, $\alpha = 23.4$

$$\boxed{P(c \mid d, x, \neg s) = 0.02}$$

Problem 2:

1) Model

Transition model:

X_t	$P(X_{t+1}=00 \mid X_t)$	$P(X_{t+1}=01 \mid X_t)$	$P(X_{t+1}=10 \mid X_t)$	$P(X_{t+1}=11 \mid X_t)$
0 0	0.75	0.25	0.00	0.00
0 1	0.25	0.50	0.25	0.00
1 0	0.00	0.25	0.50	0.25
1 1	0.00	0.00	0.25	0.75

Observation model:

X_t	$P(e_t=1 \mid X_t)$	$P(e_t=0 \mid X_t)$
0 0	0	1
0 1	0	1
1 0	1	0
1 1	1	0

2) Initial belief: $P(X_0) = [0.25, 0.25, 0.25]$

Belief at time $t=1$: $P(X_1)$

$$P(x_1 = 00 \mid e_1) = \alpha P(e_1 \mid x_1) \sum_{x_0} P(x_0) P(x_1 \mid x_0)$$
$$= \alpha\, 1 \times [0.25 \times 0.75 + 0.25 \times 0.25 + 0 + 0]$$
$$= \alpha\, 0.25$$

$$P(x_1 = 01 \mid e_1) = \alpha P(e_1 \mid x_1 = 01) \sum_{x_0} P(x_0) P(x_1 = 01 \mid x_0)$$
$$= \alpha \times 1 \times [0.25 \times 0.25 + 0.25 \times 0.5 + 0.25 \times 0.25]$$
$$= \alpha \times 0.25$$

$$P(x_1 = 10 \mid e_1) = \alpha P(e_1 \mid x_1 = 10) \times \sum_{x_0} P(x_0) P(x_1 = 10 \mid x_0)$$
$$= \alpha \times 0$$
$$= 0$$

$$P(x_1 = 11 \mid e_1) = \alpha P(e_1 \mid x_1 = 11) \times \sum_{x_0} P(x_0) P(x_1 = 11 \mid x_0)$$
$$= \alpha \times 0$$
$$= 0$$

Thus $\alpha = \frac{1}{2}$

$$P(X_1) = [0.50, 0.50, 0.0, 0.0]$$

$$P(X_2 \mid e_1, e_2) = \alpha P(e_2 \mid X_2) \times \sum_{x_1} P(x_1 \mid e_1) P(X_2 \mid x_1)$$

$$P(x_2 = 00 \mid e_1, e_2) = \alpha \times 1 \times [0.5 \times 0.75 + 0.5 \times 0.25]$$
$$= \alpha \times 0.5$$

$$P(x_2 = 01 \mid e_1, e_2) = \alpha \times 1 \times [0.5 \times 0.25 + 0.50 \times 0.50]$$
$$= \alpha \times 0.375$$

$$P(x_2 = 10 \mid e_1, e_2) = \alpha \times 0 = 0$$

$$P(x_2 = 11 \mid e_1, e_2) = \alpha \times 0 = 0$$

$$\alpha = \frac{1}{0.875}$$

$$P(X_2 \mid e_1, e_2) = [0.57, 0.43, 0.0, 0.0]$$

3) <u>Smoothing:</u>

$$P(X_1 | e_1, e_2) = \alpha \cdot P(e_2 | X_1) P(X_1 | e_1)$$

$$= \alpha \cdot P(X_1 | e_1) \sum_{X_2} P(X_2 | X_1) P(e_2 | X_1)$$

$$P(e_2 | X_1 = 00) = 0.75 \times 1 + 0.25 \times 1 + 0 + 0$$

$$= 1$$

$$P(e_2 | X_1 = 01) = 0.25 \times 1 + 0.50 \times 1 + 0.25 \times 0$$

$$= 0.75$$

$$P(X_1 = 00 | e_1, e_2) = \alpha \times 1 \times 0.5$$

$$P(X_2 = 01 | e_1, e_2) = \alpha \times 0.75 \times 0.5$$

$$\alpha = \frac{1}{0.875}$$

Thus $P(X_1 | e_1, e_2) = [0.57, 0.43, 0.0, 0.0]$

4) Viterbi Algorithm:

State	t=0	t=1	t=2
0 0	0.25	$0.1875 \times 1 = 0.1875$	0.140625 ✓
0 1	0.25	$0.125 \times 1 = 0.125$	0.0625
1 0	0.25	$\times 0 = 0$	
1 1	0.25	$\times 0 = 0$	

Most likely sequence is: $(0,0) \rightarrow (0,0) \rightarrow (0,0)$ (t=0, t=1, t=2)

5) Prediction:

$$P(e_3 \mid e_1, e_2) = \sum_{X_2} \overbrace{P(X_2 \mid e_1, e_2)}^{\text{Filtering}} P(e_3 \mid X_2)$$

$$P(e_3 \mid X_2) = \sum_{x_3} P(x_3 \mid X_2) P(e_3 \mid x_3)$$

$$= \begin{bmatrix} 0.75 \times 0 + 0.25 \times 0 + 0 \times 1 + 0 \times 1, & 0.25 \times 0 + 0.5 \times 0 + 0.25 \times 1 + 0 \times 1 \\ 0 \times 0 + 0.25 \times 0 + 0.5 \times 1 + 0.25 \times 1, & 0 \times 0 + 0 \times 0 + 0.25 \times 1 + 0.75 \times 1 \end{bmatrix}$$

$$= [0, 0.25, 0.75, 1]$$

$$P(e_3 \mid e_1, e_2) = 0 \times 0.57 + 0.25 \times 0.43 + 0.75 \times 0 + 1 \times 0 = 0.1075$$

Filtering

Problem 3:

1) $V^{\pi}(s) = R(s,a) + \gamma \sum_{s'} T(s,a,s') V^{\pi}(s')$

2)

State	V_0	V_1	V_2
A	0	0	$0+1\times 0=0$
B	0	5	$5+1\times 5=10$

3) Policy Improvement:

State	Action 1	Action 2
A	$0+1\times 0 = 0$	$-1+1\times(0.5\times 0+0.5\times 1$
B	$5+1\times 10 = \underline{15}$	$0+1\times 10 = 10$

New policy: $\pi_{new}(A) = 2$

$\pi_{new}(B) = 1$

Problem 4:

a) False
b) True
c) False
d) True
e) Flase
f) True
g) True
h) False
i) True
j) False

Problem 3:

4) Bellman Equation for a stochastic policy

$$V^{\pi}(s) = \sum_{a} \pi(s,a) \times \left(R(s,a) + \gamma \sum_{s'} T(s,a,s') V^{\pi}(s')\right)$$