## CS188 Spring 2014 Section 6: Hidden Markov Models

## 1 Basic computations

Consider the following Hidden Markov Model.


| $X_{1}$ | $\operatorname{Pr}\left(X_{1}\right)$ |
| :---: | :---: |
| 0 | 0.3 |
| 1 | 0.7 |


| $X_{t}$ | $X_{t+1}$ | $\operatorname{Pr}\left(X_{t+1} \mid X_{t}\right)$ |
| :---: | :---: | :---: |
| 0 | 0 | 0.4 |
| 0 | 1 | 0.6 |
| 1 | 0 | 0.8 |
| 1 | 1 | 0.2 |


| $X_{t}$ | $O_{t}$ | $\operatorname{Pr}\left(O_{t} \mid X_{t}\right)$ |
| :---: | :---: | :---: |
| 0 | $A$ | 0.9 |
| 0 | $B$ | 0.1 |
| 1 | $A$ | 0.5 |
| 1 | $B$ | 0.5 |

Suppose that $O_{1}=A$ and $O_{2}=B$ is observed.
Use the Forward algorithm to compute the probability distribution $\operatorname{Pr}\left(X_{2}, O_{1}=A, O_{2}=B\right)$. Show your work. You do not need to evaluate arithmetic expressions involving only numbers.

| $X_{1}$ | $\operatorname{Pr}\left(X_{1}, O_{1}=A\right)$ |
| :---: | :---: |
| 0 | $0.3 \cdot 0.9$ |
| 1 | $0.7 \cdot 0.5$ |


| $X_{2}$ | $\operatorname{Pr}\left(X_{2}, O_{1}=A, O_{2}=B\right)$ |
| :---: | :---: |
| 0 | $0.1 \cdot[0.4 \cdot(0.3 \cdot 0.9)+0.8 \cdot(0.7 \cdot 0.5)]=0.0388$ |
| 1 | $0.5 \cdot[0.6 \cdot(0.3 \cdot 0.9)+0.2 \cdot(0.7 \cdot 0.5)]=0.1160$ |

Use the Viterbi algorithm to compute the maximum probability sequence $X_{1}, X_{2}$. Show your work.

| $X_{1}$ | $\operatorname{Pr}\left(X_{1}, O_{1}=A\right)$ |
| :---: | :---: |
| 0 | $0.3 \cdot 0.9$ |
| 1 | $0.7 \cdot 0.5$ |


| $X_{2}$ | $\max _{x_{1}} \operatorname{Pr}\left(X_{1}=x_{1}, X_{2}, O_{1}=A, O_{2}=B\right)$ | arg max |
| :---: | :---: | :---: |
| 0 | $0.1 \cdot \max (0.4 \cdot(0.3 \cdot 0.9), 0.8 \cdot(0.7 \cdot 0.5))=0.1 \cdot \max (0.108,0.28)=0.028$ | $X_{1}=1$ |
| 1 | $0.5 \cdot \max (0.6 \cdot(0.3 \cdot 0.9), 0.2 \cdot(0.7 \cdot 0.5))=0.5 \cdot \max (0.162,0.07)=0.081$ | $X_{1}=0$ |

Thus, in the maximum probability sequence, $X_{2}=1$ and $X_{1}=0$.

True or false: Variable elimination is generally more accurate than the Forward algorithm. Explain your answer. They both perform exact inference.

## 2 Tracking a Jabberwock

You have been put in charge of a Jabberwock for your friend Lewis. The Jabberwock is kept in a large tugley wood which is conveniently divided into an $N \times N$ grid. It wanders freely around the $N^{2}$ possible cells. At each time step $t=1,2,3, \ldots$, the Jabberwock is in some cell $X_{t} \in\{1, \ldots, N\}^{2}$, and it moves to cell $X_{t+1}$ randomly as follows: with probability $1-\epsilon$, it chooses one of the (up to 4) valid neighboring cells uniformly at random; with probability $\epsilon$, it uses its magical powers to teleport to a random cell uniformly at random among the $N^{2}$ possibilities (it might teleport to the same cell). Suppose $\epsilon=\frac{1}{2}, N=10$ and that the Jabberwock always starts in $X_{1}=(1,1)$.

1. Compute the probability that the Jabberwock will be in $X_{2}=(2,1)$ at time step 2 . What about $\operatorname{Pr}\left(X_{2}=\right.$ $(4,4))$ ?
$\operatorname{Pr}\left(X_{2}=(2,1)\right)=1 / 2 \cdot 1 / 2+1 / 2 \cdot 1 / 100=0.255$
$\operatorname{Pr}\left(X_{2}=(4,4)\right)=1 / 2 \cdot 1 / 100=0.005$

At each time step $t$, you dont see $X_{t}$ but see $E_{t}$, which is the row that the Jabberwock is in; that is, if $X_{t}=(r, c)$, then $E_{t}=r$. You still know that $X_{1}=(1,1)$.
2. Suppose we see that $E_{1}=1, E_{2}=2, E_{3}=10$. Fill in the following table with the distribution over $X_{t}$ after each time step, taking into consideration the evidence. Your answer should be concise. Hint: you should not need to do any heavy calculations.

| $t$ | $\operatorname{Pr}\left(X_{t}, e_{1: t-1}\right)$ | $\operatorname{Pr}\left(X_{t}, e_{1: t}\right)$ |
| :---: | :---: | :---: |
| 1 | $(1,1): 1.0$, (others):0.0 | $(1,1): 1.0$, (others):0.0 |
|  |  |  |
| 2 | $(1,2),(2,1): 51 / 200$, (others):1/200 | $(2,1): 51 / 200,(2,2+): 1 / 200$ |
|  |  |  |

