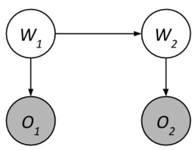
CS 188
Summer 2023

Discussion 4B Solutions

1 Particle Filtering

Let's use Particle Filtering to estimate the distribution of $P(W_2|O_1 = a, O_2 = b)$. Here's the HMM again.



W_1	$P(W_1)$
0	0.3
1	0.7

W_t	W_{t+1}	$P(W_{t+1}\|W_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

W_t	O_t	$P(O_t\|W_t)$
0	a	0.9
0	b	0.1
1	a	0.5
1	b	0.5

We start with two particles representing our distribution for W_1.
$P_1 : W_1 = 0$
$P_2 : W_1 = 1$
Use the following random numbers to run particle filtering:

[0.22, 0.05, 0.33, 0.20, 0.84, 0.54, 0.79, 0.66, 0.14, 0.96]

(a) Observe: Compute the weight of the two particles after evidence $O_1 = a$.
$w(P_1) = P(O_t = a|W_t = 0) = 0.9$
$w(P_2) = P(O_t = a|W_t = 1) = 0.5$

(b) Resample: Using the random numbers, resample P_1 and P_2 based on the weights.
We now sample from the weighted distribution we found above. Using the first two random samples, we find:
$P_1 = sample(weights, 0.22) = 0$
$P_2 = sample(weights, 0.05) = 0$

(c) Predict: Sample P_1 and P_2 from applying the time update.
$P_1 = sample(P(W_{t+1}|W_t = 0), 0.33) = 0$
$P_2 = sample(P(W_{t+1}|W_t = 0), 0.20) = 0$

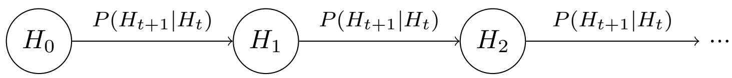
(d) Update: Compute the weight of the two particles after evidence $O_2 = b$.
$w(P_1) = P(O_t = b|W_t = 0) = 0.1$
$w(P_2) = P(O_t = b|W_t = 0) = 0.1$

(e) Resample: Using the random numbers, resample P_1 and P_2 based on the weights.
Because both of our particles have $X = 0$, resampling will still leave us with two particles with $X = 0$.
$P_1 = 0$
$P_2 = 0$

(f) What is our estimated distribution for $P(W_2|O_1 = a, O_2 = b)$?
$P(W_2 = 0|O_1 = a, O_2 = b) = 2/2 = 1$
$P(W_2 = 1|O_1 = a, O_2 = b) = 0/2 = 0$

2 MangoBot Human Detector

Your startup company MangoBot wants to build robots that delivers packages on the road. One core module of the robot's software is to detect whether a human is standing in front of it. We model the presence of humans with a Markov model:



where $H_t \in \{0, 1\}$ corresponds to a human being absent or present respectively. The initial distribution and the transition probabilities are given as follows:

H_0	$P(H_0)$
0	p
1	$1-p$

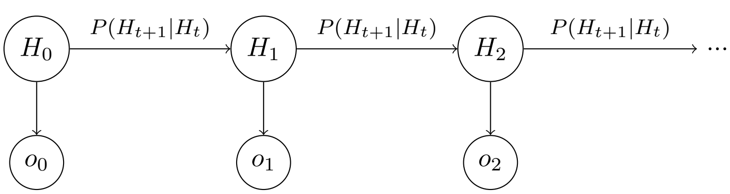
H_t	H_{t+1}	$P(H_{t+1}\|H_t)$
0	0	0.9
0	1	0.1
1	0	0.8
1	1	0.2

(a) Express the following quantities in terms of p:

(i) $P(H_1 = 1) = -0.1p + 0.2$
$P(H_1 = 1) = P(H_1 = 1, H_0 = 0) + P(H_1 = 1, H_0 = 1) = P(H_0 = 0)P(H_1 = 1|H_0 = 0) + P(H_0 = 1)P(H_1 = 1|H_0 = 1) = 0.2(1-p) + 0.1p = 0.2 - 0.1p$

(ii) $\lim_{t\to\infty} P(H_t = 0) = 8/9$
As $t \to \infty$, the system converges to the stationary distribution π, which satisfies $\pi = T^\top \pi$, where T is the transition probability matrix $\begin{bmatrix} 0.9 & 0.1 \\ 0.8 & 0.2 \end{bmatrix}$. Assume $\pi = [q \quad (1-q)]^\top$ and solve for q in $\pi = T^\top \pi$ gives $q = 8/9$.

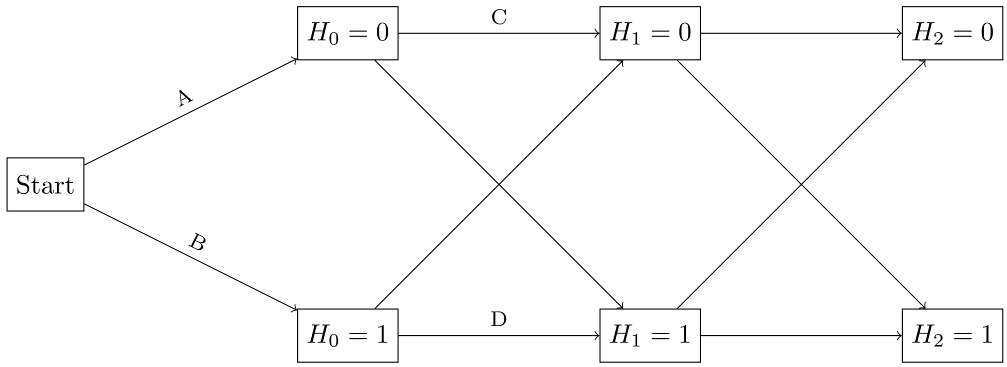
To make things simple, we stick to the original first-order Markov chain formulation. To make the detection more accurate, the company built a sensor that returns an observation O_t each time step as a noisy measurement of the unknown H_t. The new model is illustrated in the figure, and the relationship between H_t and O_t is provided in the table below.



H_t	O_t	$P(O_t\|H_t)$
0	0	0.8
0	1	0.2
1	0	0.3
1	1	0.7

(b) Based on the observed sensor values $o_0, o_1, \cdots, o_t$, we now want the robot to find the most likely sequence $H_0, H_1, \cdots, H_t$ indicating the presence/absence of a human up to the current time.

(i) Suppose that $[o_0, o_1, o_2] = [0, 1, 1]$ are observed. The following "trellis diagram" shows the possible state transitions. Fill in the values for the arcs labeled A, B, C, and D with the product of the transition probability and the observation likelihood for the destination state. The values may depend on p.



A: $0.8p$; B: $0.3(1-p)$; C: $0.9 * 0.2 = 0.18$; D: $0.2 * 0.7 = 0.14$.

(ii) There are two possible most likely state sequences, depending on the value of p. Complete the following (Write the sequence as "x,y,z" (without quotes), where x, y, z are either 0 or 1):
Hint: it might be helpful to complete the labelling of the trellis diagram above.

- When $p <$ [0.25], the most likely sequence H_0, H_1, H_2 is [1,0,0].
- Otherwise, the most likely sequence H_0, H_1, H_2 is [0, 0, 0].

After filling out the full trellis diagram, we can easily observe that the sequence with largest probability given $H_0 = 0$ is (0, 0, 0) and the sequence with largest probability given $H_0 = 1$ is (1, 0, 0). (To see what is the sequence with largest probability, we run search from $H_0 = 0$ to either $H_2 = 0$ or $H_0 = 1$, but instead of adding the costs we multiply the probabilities.) Therefore the two possible most likely state sequences are (0, 0, 0), with probability $0.8p * 0.18 * 0.18$, and (1, 0, 0), with probability $0.3(1-p) * 0.16 * 0.18$. Setting up the equation $0.8p * 0.18 * 0.18 = 0.3(1-p) * 0.16 * 0.18$ gives $p = 0.25$ to be the threshold.
Note that this is a bit counter-intuitive since the observations suggest the exact opposite thing. However, in this problem the transition probabilities for 0 to 0 and 1 to 0 are so large that they dominates the computation.

(c) True or False: For a fixed p value and observations $\{o_0, o_1, o_2\}$ in general, H_1^*, the most likely value for H_1, is always the same as the value of H_1 in the most likely sequence H_0, H_1, H_2. ○ True ● False
The maximum likelihood estimation (MLE) for a single variable is in general not the same as the value of that variable in the most likely sequence estimation (MLSE). For example, in this problem, when $p = 0.3$, the MLE for H_0 is 1, but the value of H_0 in the MLSE is 1. There exists similar examples for H_1.