**ECE/CS 498 DSU/DSG Spring 2020**
**In-Class Activity 4**

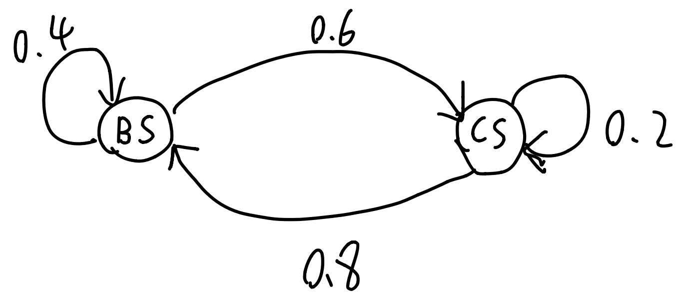
NetID: chuhaof2

The purpose of the in-class activity is for you to:
(i) Understand how to model a time series prediction problem as an HMM.
(ii) Go through the forward-backward algorithm for predicting the most likely hidden state given the time series observations.

**Problem 1**

The security state of a computer can be either in a **benign state (BS)** or in a **compromised state (CS)**. The computer is constantly being attacked by hackers. The probability that an attack is successful, and the computer moves from benign to compromised is **0.6**. The probability that the computer, given that is in a compromised state, detects the attacker and transitions from the compromised state to the benign state is **0.8**. In all other situations, the state remains unchanged. **The transition probabilities are independent of the past states given the current state of the system.** At any point in time, it is believed, that the computer is in the **benign state** with probability **0.9**.

a) Draw the states with the state transition probabilities that describe this system:




There is no way of directly observing the state of the computer. On the other hand, there are **system events** like **port scanning (PS)** and **web browsing (WB)** that can be observed. The probability of observing an event depends only on the state of the computer. The probability that a benign user does a port scan is **0.4** and does web browsing is **0.6**. An attacker will perform a port scan with a probability of **0.7** and web browsing with a probability of **0.3**.

During an observation period, the following sequence of events was observed: [WB, PS, WB] corresponding to t=1, 2 and 3 respectively. Answer the following questions.

b) Is it possible to identify the exact state of the computer at time instant two (t=2)? If not, state a condition when you can fully determine (with 100% probability) the system's state after observing an event.

No, not possible

The only condition is: benign user does WB with probability 1, the attacker does PS with probability 1. (At this condition, WB must be BS, PS must be CS)

c) Is it possible to identify the most likely state of the computer at t=2?

Yes

d) Mathematically express the property of the transition probability of states mentioned above. What is the property known as?

$$Q = \begin{array}{cc} \begin{bmatrix} P(S_{t+1}=BS \mid S_t=BS) & P(S_{t+1}=CS \mid S_t=BS) \\ P(S_t=BS \mid S_t=CS) & P(S_{t+1}=CS \mid S_t=CS) \end{bmatrix} & \begin{array}{c} BS \\ CS \end{array} \\ \begin{array}{cc} \quad BS \quad\quad\quad\quad & \quad\quad CS \end{array} & \end{array}$$

Markov assumption (state variable only depends on the previous state)

e) Which of the following models can be used to answer the question in part (c)? Explain your answer.

Linear Regression:

We can only use the Hidden Markov Models:
reason: ① ②

Markov Models:

① The current state probability only depends on the previous state, Which is the Markov assumption. we must use the Markov based models

Hidden Markov Models:

②
Because we can not observe the state directly (Hidden state)
We must use the HMM to represent that.

Now that is clear that we need to use HMM to solve answer the question in part (c), let us set up the model.

f) Write down the state transition probability matrix **A**, observation matrix **B** and the initial distribution of hidden states $\boldsymbol{\pi}$.

$1 := BS \quad 2 := CS$

$q_{ij}$: from $j$ state to $i$ state

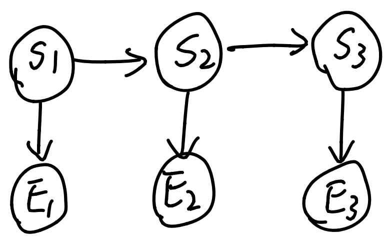
$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \begin{bmatrix} 0.4 & 0.8 \\ 0.6 & 0.2 \end{bmatrix}$$

$$B \quad \begin{array}{cc} & \begin{array}{cc} PS & WB \end{array} \\ \begin{array}{c} BS \\ CS \end{array} & \begin{bmatrix} 0.4 & 0.6 \\ 0.7 & 0.3 \end{bmatrix} \end{array}$$

$$\pi \quad \begin{array}{c} \begin{array}{cc} BS & CS \end{array} \\ \begin{bmatrix} 0.9 & 0.1 \end{bmatrix} \end{array}$$

g) Draw the HMM model. Denote the hidden states as $S_t$ for $t \in \{1,2,3\}$.

$S_t \in [BS, CS]$ $\qquad$ $E_t \in [PS, WB]$




To predict the most likely state for $S_2$ , we need to compute

$$S_2^* = \underset{\sigma_j \in \{BS,CS\}}{\operatorname{argmax}} \gamma_2(j)$$

where

$$\gamma_2(j) = P(S_2 = \sigma_j | E_1, E_2, E_3)$$

Recall from the lecture slides, that to calculate $\gamma_2$, we need to perform the forward algorithm which gives us $\alpha_2$, and the backward algorithm that produces $\beta_2$.

h) Compute $\alpha_2$ recursively using the forward algorithm.

| <WB> (t=1) | | | |
|---|---|---|---|
| States | $\boldsymbol{\alpha_1}$ | | Normalize $\alpha_1$ |
| BS | $\alpha_1(BS)$ | $P(S_1 = BS) \times P(E_1 = WB \mid S_1 = BS)$ $= 0.9 \times 0.6 = 0.54$ | $\frac{\alpha_1(BS)}{\alpha_1(BS) + \alpha_1(CS)}$ $= 0.947$ |
| CS | $\alpha_1(CS)$ | $P(S_1 = CS) \times P(E_1 = WB \mid S_1 = CS)$ $= 0.1 \times 0.3 = 0.03$ | $\frac{\alpha_1(CS)}{\alpha_1(BS) + \alpha_1(CS)}$ $= 0.053$ |

| <WB, PS> (t=2) | | | |
|---|---|---|---|
| States | $\boldsymbol{\alpha_2}$ | | Normalize $\alpha_2$ |
| BS | $\alpha_2(BS)$ | $[\alpha_1(BS) \times P(S_2 = BS \mid S_1 = BS)$ $+ \alpha_1(CS) \times P(S_2 = BS \mid S_1 = CS)] \times P(E_2 = PS \mid S_2 = BS)$ $= (0.947 \times 0.4 + 0.053 \times 0.8) \times 0.4 = 0.16848$ | $\frac{\alpha_2(BS)}{\alpha_2(BS) + \alpha_2(CS)}$ $= 0.2937$ |
| CS | $\alpha_2(CS)$ | $[\alpha_1(BS) \times P(S_2 = CS \mid S_1 = BS)$ $+ \alpha_1(CS) \times P(S_2 = CS \mid S_1 = CS)]$ $\times P(E_2 = PS \mid S_2 = CS)$ $= (0.947 \times 0.6 + 0.053 \times 0.2) \times 0.7 = 0.40516$ | $\frac{\alpha_2(CS)}{\alpha_2(BS) + \alpha_2(CS)}$ $= 0.7063$ |

i) Compute $\beta_2$ recursively using the backward algorithm.
Note: We initialize $\beta_3(BS) = 1,\ \beta_3(CS) = 1$

| <WB, PS> (t=2) (WB observed at t=3) | | |
|---|---|---|
| States | $\boldsymbol{\beta_2}$ | |
| BS | $\beta_2(BS)$ | $P(S_3 = BS \mid S_2 = BS) \times P(E_3 = WB \mid S_3 = BS) \times \beta_3(BS)$ $+ P(S_3 = CS \mid S_2 = BS) \times P(E_3 = WB \mid S_3 = CS) \times \beta_3(CS)$ $= 0.4 \times 0.6 \times 1 + 0.6 \times 0.3 \times 1$ $= 0.24 + 0.18 = 0.42$ |
| CS | $\beta_2(CS)$ | $P(S_3 = BS \mid S_2 = CS) \times P(E_3 = WB \mid S_3 = BS) \times \beta_3(BS)$ $+ P(S_3 = CS \mid S_2 = CS) \times P(E_3 = WB \mid S_3 = CS) \times \beta_3(CS)$ $= 0.8 \times 0.6 \times 1 + 0.2 \times 0.3 \times 1$ $= 0.48 + 0.06 = 0.54$ |

j) Compute $\gamma_2$ and find $S_2^*$

| <WB, PS> (t=2) | | | |
|---|---|---|---|
| States | $\boldsymbol{\gamma_2}$ | | Normalize $\gamma_2$ |
| BS | $\gamma_2(BS)$ | $\beta_2(BS) \times \alpha_2(BS) =$ 0.123354 | $\frac{\gamma_2(BS)}{\gamma_2(BS)+\gamma_2(CS)}$ = 0.2444 |
| CS | $\gamma_2(CS)$ | $\beta_2(CS) \times \alpha_2(CS) =$ 0.381402 | $\frac{\gamma_2(CS)}{\gamma_2(CS)+\gamma_2(BS)}$ = 0.7556 |

$S_2^* =$ CS

k) What if the observation matrix B was modified to be the following: $B = \begin{bmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{bmatrix}$.
What is the most likely state in this case? (Hint: The observation matrix does not give us any additional information of which hidden state is more likely given an observation)

$$\alpha_1(BS) = \frac{0.9\times0.1}{0.9\times0.1+0.1\times0.1} = 0.9$$
$$\alpha_1(CS) = \frac{0.1\times0.1}{0.9\times0.1+0.1\times0.1} = 0.1$$

$$\alpha_2(BS) = \frac{(0.9\times0.4+0.1\times0.8)\times0.4}{(0.9\times0.4+0.1\times0.8)\times0.4+(0.9\times0.6+0.1\times0.2)\times0.7} = 0.310$$
$$\alpha_2(CS) = \frac{(0.9\times0.6+0.1\times0.2)\times0.7}{(0.9\times0.4+0.1\times0.8)\times0.4+(0.9\times0.6+0.1\times0.2)\times0.7} = 0.690$$

$$\beta_2(BS) = 0.4 \times 0.1 \times 1 + 0.6 \times 0.1 \times 1 = 0.1$$
$$\beta_2(CS) = 0.8 \times 0.1 \times 1 + 0.2 \times 0.1 \times 1 = 0.1$$

$$\gamma_2(BS) = \frac{\beta_2(BS)\times\alpha_2(BS)}{\beta_2(BS)\times\alpha_2(BS)+\beta_2(CS)\times\alpha_2(CS)} = 0.31$$
$$\gamma_2(CS) = \frac{\beta_2(CS)\times\alpha_2(CS)}{\beta_2(BS)\times\alpha_2(BS)+\beta_2(CS)\times\alpha_2(CS)} = 0.69$$

$$\therefore S_2^* = CS$$