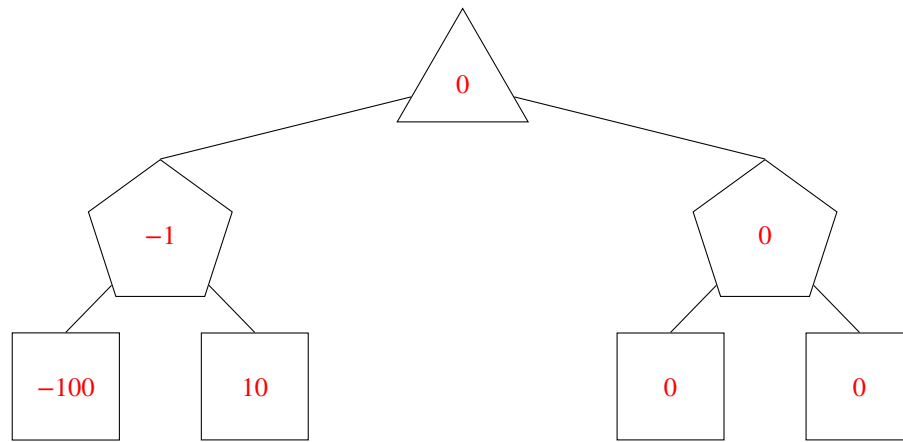


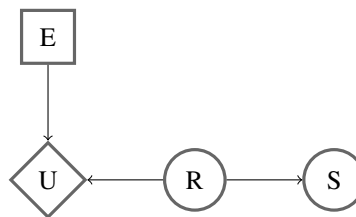
### Q1. Decision Networks and VPI

Valerie has just found a cookie on the ground. She is concerned that the cookie contains raisins, which she really dislikes but she still wants to eat the cookie. If she eats the cookie and it contains raisins she will receive a utility of  $-100$  and if the cookie doesn't contain raisins she will receive a utility of  $10$ . If she doesn't eat the cookie she will get  $0$  utility. The cookie contains raisins with probability  $0.1$ .

- (a) We want to represent this decision network as an expectimax game tree. Fill in the nodes of the tree below, with the top node representing her maximizing choice.



- (b) Should Valerie eat the cookie?  Yes  No
- (c) Valerie can now smell the cookie to judge whether it has raisins before she eats it. However, since she dislikes raisins she does not have much experience with them and cannot recognize their smell well. As a result she will incorrectly identify raisins when there are no raisins with probability  $0.2$  and will incorrectly identify no raisins when there are raisins with probability  $0.3$ . This decision network can be represented by the diagram below where E is her choice to eat, U is her utility earned, R is whether the cookie contains raisins, and S is her attempt at smelling.



Valerie has just smelled the cookie and she thinks it doesn't have raisins. Write the probability,  $X$ , that the cookie has raisins given that she smelled no raisins as a simplest form fraction or decimal.

$X =$  0.04

$$P(+r | -s) = \frac{P(-s|+r)P(+r)}{P(-s)} = \frac{P(-s|+r)P(+r)}{P(-s|+r)P(+r) + P(-s|-r)P(-r)} = \frac{.3*.1}{.3*.1 + .8*.9} = \frac{.03}{.75} = .04$$

- (d) What is her maximum expected utility,  $Y$  given that she smelled no raisins? You can answer in terms of  $X$  or as a simplest form fraction or decimal.

$$Y = \boxed{-100X + 10(1 - X), 5.6}$$

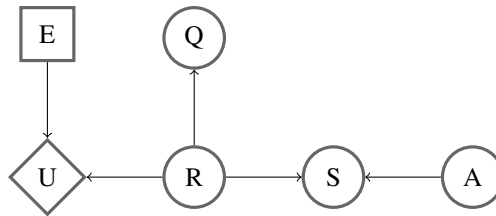
$$\begin{aligned}
 MEU(-s) &= \max(MEU(eating| -s), MEU(noteating| -s)) = \\
 &= \max(P(+r| -s) * EU(eating, +r) + P(-r| -s) * EU(eating, -r), MEU(noteating)) = \\
 &= \max(X * (-100) + (1 - X) * 10, 0) = \\
 &= X * 100 + (1 - X) * 10
 \end{aligned}$$

(e) What is the Value of Perfect Information (VPI) of smelling the cookie? You can answer in terms of X and Y or as a simplest form fraction or decimal.

$$VPI = \boxed{0.75 * Y, 4.2}$$

$$\begin{aligned}
 VPI(S) &= MEU(S) - MEU(\emptyset) \\
 MEU(S) &= P(-s)MEU(-s) + P(+s)MEU(+s) \\
 P(-s) &= .75 \text{ from part (c), } MEU(-s) = Y \\
 MEU(+s) &= 0 \text{ because it was better for her to not eat the raisin without knowing anything, smelling raisins will only} \\
 &\text{make it more likely for the cookie to have raisins and it will still be best for her to not eat and earn a utility of 0. Note this} \\
 &\text{means we do not have to calculate } P(+s). \\
 MEU(\emptyset) &= 0 \\
 VPI(S) &= .75 * Y + 0 - 0 = .75 * Y
 \end{aligned}$$

(f) Valerie is unsatisfied with the previous model and wants to incorporate more variables into her decision network. First, she realizes that the air quality (A) can affect her smelling accuracy. Second, she realizes that she can question (Q) the people around to see if they know where the cookie came from. These additions are reflected in the decision network below.



Choose one for each equation:

	Could Be True	Must Be True	Must Be False
$VPI(A, S) > VPI(A) + VPI(S)$	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
$VPI(A) = 0$	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
$VPI(Q, R) \leq VPI(Q) + VPI(R)$	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
$VPI(S, R) > VPI(R)$	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
$VPI(Q) \geq 0$	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
$VPI(Q, A) > VPI(Q)$	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
$VPI(S A) < VPI(S)$	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
$VPI(A S) > VPI(A)$	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Q2. Naive Bayes: Pacman or Ghost?

You are standing by an exit as either Pacmen or ghosts come out of it. Every time someone comes out, you get two observations: a visual one and an auditory one, denoted by the random variables  $X_v$  and  $X_a$ , respectively. The visual observation informs you that the individual is either a Pacman ( $X_v = 1$ ) or a ghost ( $X_v = 0$ ). The auditory observation  $X_a$  is defined analogously. Your observations are a noisy measurement of the individual's true type, which is denoted by  $Y$ . After the individual comes out, you find out what they really are: either a Pacman ( $Y = 1$ ) or a ghost ( $Y = 0$ ). You have logged your observations and the true types of the first 20 individuals:

individual $i$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
first observation $X_v^{(i)}$	0	0	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	0
second observation $X_a^{(i)}$	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
individual's type $Y^{(i)}$	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0

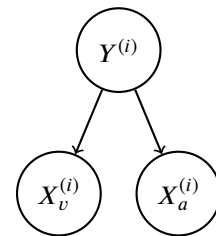
The superscript  $(i)$  denotes that the datum is the  $i$ th one. Now, the individual with  $i = 20$  comes out, and you want to predict the individual's type  $Y^{(20)}$  given that you observed  $X_v^{(20)} = 1$  and  $X_a^{(20)} = 1$ .

- (a) Assume that the types are independent, and that the observations are independent conditioned on the type. You can model this using naïve Bayes, with  $X_v^{(i)}$  and  $X_a^{(i)}$  as the features and  $Y^{(i)}$  as the labels. Assume the probability distributions take on the following form:

$$P(X_v^{(i)} = x_v | Y^{(i)} = y) = \begin{cases} p_v & \text{if } x_v = y \\ 1 - p_v & \text{if } x_v \neq y \end{cases}$$

$$P(X_a^{(i)} = x_a | Y^{(i)} = y) = \begin{cases} p_a & \text{if } x_a = y \\ 1 - p_a & \text{if } x_a \neq y \end{cases}$$

$$P(Y^{(i)} = 1) = q$$



for  $p_v, p_a, q \in [0, 1]$  and  $i \in \mathbb{N}$ .

- (i) What's the maximum likelihood estimate of  $p_v, p_a$  and  $q$ ?

$$p_v = \underline{\frac{4}{5}} \quad p_a = \underline{\frac{3}{5}} \quad q = \underline{\frac{1}{2}}$$

To estimate  $q$ , we count 10  $Y = 1$  and 10  $Y = 0$  in the data. For  $p_v$ , we have  $p_v = 8/10$  cases where  $X_v = 1$  given  $Y = 1$  and  $1 - p_v = 2/10$  cases where  $X_v = 1$  given  $Y = 0$ . So  $p_v = 4/5$ . For  $p_a$ , we have  $p_a = 2/10$  cases where  $X_a = 1$  given  $Y = 1$  and  $1 - p_a = 8/10$  cases where  $X_a = 1$  given  $Y = 0$ . The average of  $2/10$  and  $1$  is  $3/5$ .

- (ii) What is the probability that the next individual is Pacman given your observations? Express your answer in terms of the parameters  $p_v, p_a$  and  $q$  (you might not need all of them).

$$P(Y^{(20)} = 1 | X_v^{(20)} = 1, X_a^{(20)} = 1) = \frac{p_v p_a q}{p_v p_a q + (1 - p_v)(1 - p_a)(1 - q)}$$

The joint distribution  $P(Y = 1, X_v = 1, X_a = 1) = p_v p_a q$ . For the denominator, we need to sum out over  $Y$ , that is, we need  $P(Y = 1, X_v = 1, X_a = 1) + P(Y = 0, X_v = 1, X_a = 1)$ .

Now, assume that you are given additional information: you are told that the individuals are actually coming out of a bus that just arrived, and each bus carries *exactly* 9 individuals. Unlike before, the types of every 9 consecutive individuals are *conditionally* independent given the bus type, which is denoted by  $Z$ . Only after all of the 9 individuals have walked out, you find out the bus type: one that carries mostly Pacmans ( $Z = 1$ ) or one that carries mostly ghosts ( $Z = 0$ ). Thus, you only know the bus type in which the first 18 individuals came in:

individual $i$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
first observation $X_v^{(i)}$	0	0	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	0	
second observation $X_a^{(i)}$	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	
individual's type $Y^{(i)}$	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	
bus $j$										0										1	
bus type $Z^{(j)}$										0										1	

(b) You can model this using a variant of naïve bayes, where now 9 consecutive labels  $Y^{(i)}, \dots, Y^{(i+8)}$  are *conditionally* independent given the bus type  $Z^{(j)}$ , for bus  $j$  and individual  $i = 9j$ . Assume the probability distributions take on the following form:

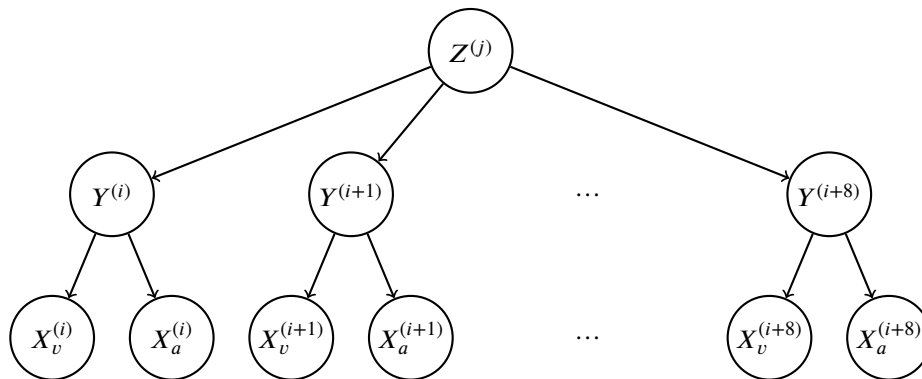
$$P(X_v^{(i)} = x_v | Y^{(i)} = y) = \begin{cases} p_v & \text{if } x_v = y \\ 1 - p_v & \text{if } x_v \neq y \end{cases}$$

$$P(X_a^{(i)} = x_a | Y^{(i)} = y) = \begin{cases} p_a & \text{if } x_a = y \\ 1 - p_a & \text{if } x_a \neq y \end{cases}$$

$$P(Y^{(i)} = 1 | Z^{(j)} = z) = \begin{cases} q_0 & \text{if } z = 0 \\ q_1 & \text{if } z = 1 \end{cases}$$

$$P(Z^{(j)} = 1) = r$$

for  $p, q_0, q_1, r \in [0, 1]$  and  $i, j \in \mathbb{N}$ .



(i) What's the maximum likelihood estimate of  $q_0, q_1$  and  $r$ ?

$$q_0 = \frac{2}{9} \qquad q_1 = \frac{8}{9} \qquad r = \frac{1}{2}$$

For  $r$ , we've seen one ghost bus and one pacman bus, so  $r = 1/2$ . For  $q_0$ , we're finding  $P(Y = 1 | Z = 0)$ , which is  $2/9$ . For  $q_1$ , we're finding  $P(Y = 1 | Z = 1)$ , which is  $8/9$ .

- (ii) Compute the following joint probability. Simplify your answer as much as possible and express it in terms of the parameters  $p_v, p_a, q_0, q_1$  and  $r$  (you might not need all of them).

$$P(Y^{(20)} = 1, X_v^{(20)} = 1, X_a^{(20)} = 1, Y^{(19)} = 1, Y^{(18)} = 1) = \underline{p_a p_v [q_0^3 (1 - r) + q_1^3 r]}$$

$$\begin{aligned} & P(Y^{(20)} = 1, X_v^{(20)} = 1, X_a^{(20)} = 1, Y^{(19)} = 1, Y^{(18)} = 1) \\ &= \sum_z P(Y^{(20)} = 1 | Z^{(2)} = z) P(Z^{(2)} = z) P(X_v^{(20)} = 1 | Y^{(20)} = 1) P(X_a^{(20)} = 1 | Y^{(20)} = 1) \\ &\quad P(Y^{(19)} = 1 | Z^{(2)} = z) P(Y^{(18)} = 1 | Z^{(2)} = z) \\ &= q_0(1 - r)p_a p_v q_0 q_0 + q_1 r p_a p_v q_1 q_1 \\ &= p_a p_v [q_0^3 (1 - r) + q_1^3 r] \end{aligned}$$