Do the following exercises. These are intended to take 10-30 minutes each if you know how to do them. Each is worth 10 to 20 points. Names of responsible staff members are given for each question.

---

# 1 Value Iteration (Vivek)

Consider an MDP with two states $s_1$ and $s_2$ and transition function $T(s, a, s')$ and reward function $R(s, a, s')$. Let's also assume that we have an agent whose discount factor is $\gamma = 1$. From each state, the agent can take three possible actions $a \in \{x, y, z\}$ . The transition probabilities for taking each action and the rewards for transitions are shown below.

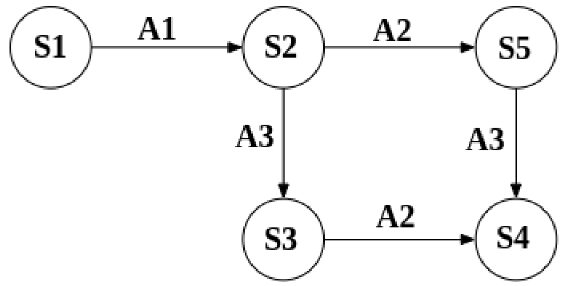
| $s$ | $a$ | $s'$ | $T(s,a,s')$ | $R(s,a,s')$ |
|---|---|---|---|---|
| $s_1$ | $x$ | $s_1$ | 0.4 | 0 |
| $s_1$ | $x$ | $s_2$ | 0.6 | 0 |
| $s_1$ | $y$ | $s_1$ | 0 | 999 |
| $s_1$ | $y$ | $s_2$ | 1 | 7 |
| $s_1$ | $z$ | $s_1$ | 0.5 | 0 |
| $s_1$ | $z$ | $s_2$ | 0.5 | 0 |

| $s$ | $a$ | $s'$ | $T(s,a,s')$ | $R(s,a,s')$ |
|---|---|---|---|---|
| $s_2$ | $x$ | $s_1$ | 0.3 | 20 |
| $s_2$ | $x$ | $s_2$ | 0.7 | 25 |
| $s_2$ | $y$ | $s_1$ | 1 | 0 |
| $s_2$ | $y$ | $s_2$ | 0 | 0 |
| $s_2$ | $z$ | $s_1$ | 0.7 | 5 |
| $s_2$ | $z$ | $s_2$ | 0.3 | 10 |

Compute $V_0$ , $V_1$ and $V_2$ for states $s_1$ and $s_2$:
(a). $V_0(s_1) = 0$
(b). $V_0(s_2) = 0$
(c). $V_1(s_1) = 7$ (taking action $y$).
(d). $V_1(s_2) = 23.5$ (taking action $x$).
(e). $V_2(s_1) = 30.5$ (taking action $y$).
(f). $V_2(s_2) = 42.05$ (taking action $x$).

## 2 Q-Learning updates (Vivek)

(10 points) Consider an agent traveling on the graph below. The states are represented by the nodes and actions are represented by the edges in the following graph.




(a) (6 points) Consider the following episodes performed in this state space. The experience tuples are of the form [s, a, s', r], where the agent starts in state s, performs action a, ends up in state s', and receives immediate reward r, which is determined by the state entered. Let $\gamma = 1.0$ for this MDP. Fill in the values computed by the Q-learning algorithm with a learning rate of = 0.2. All Q values are initially 0, and you should fill out each row using values you have computed in previous rows.

| | |
|---|---|
| [S2, A2, S5, 5] | Q(S2, A2) = $(1-0.2)*0+0.2*(5+0) = 1.0$ |
| [S5, A3, S4, 2] | Q(S5, A3) = $(1-0.2)*0+0.2*(2+0) = 0.4$ |
| [S2, A3, S3, -7] | Q(S2, A3) = $(1-0.2)*0+0.2*(-7+0) = -1.4$ |
| [S3, A2, S4, 7] | Q(S3, A2) = $(1-0.2)*0+0.2*(7+0) = 1.4$ |
| [S2, A3, S5, 2] | Q(S2, A3) = $(1-0.2)*-1.4+0.2*(2+0.4) = -0.64$ |
| [S2, A2, S3, -2] | Q(S2, A2) = $(1-0.2)*1.0+0.2*(-2+1.4) = 0.68$ |
| [S1, A1, S1, 8] | Q(S1, A1) = $(1-0.2)*0+0.2*(8+0) = 1.6$ |

(b) (3 points) Now, based on the record table in the previous problem, we want to approximate the transition function:

$T(S2, A2, S3) = 0.5$

$T(S2, A2, S5) = 0.5$

$T(S2, A3, S5) = 0.5$

$T(S2, A3, S3) = 0.5$

$T(S1, A1, S2) = 0$

$T(S1, A1, S1) = 1$

(c) (1 points) What's the key difference between Q-learning and Value Iteration? What's one advantage of each of the methods in general?
Q-learning is Model−Free learning and Value Iteration is Model−Based learning. Model−Free learning is more computationally efficient, and tends to focus its sampling more on rele-

vant parts of the state space, whereas, Model−Based learning tends to arrive at a more complete understanding of the state space, but at the cost of much more sampling.

## 3 Joint Distributions and Inference (Phuong)

(10 points) Let $N$ be the random variable that represent whether you are on the nice or naughty list. Let $C$ be the random variable that represent whether you will receive chocolate or coal from Santa.

Consider the table given below.

| $N$ | $C$ | $\mathbb{P}(N, C)$ |
|---|---|---|
| *nice* | *chocolate* | 0.45 |
| *nice* | *coal* | 0.15 |
| *naughty* | *chocolate* | 0.05 |
| *naughty* | *coal* | 0.35 |

(a) (1 point) Compute the marginal distribution $\mathbb{P}(N)$ and express it as a table.

| $N$ | $\mathbb{P}(N)$ |
|---|---|
| *nice* | 0.60 |
| *naughty* | 0.40 |

(b) (1 point) Similarly, compute the marginal distribution $\mathbb{P}(C)$ and express it as a table.

| $C$ | $P(C)$ |
|---|---|
| *chocolate* | 0.50 |
| *coal* | 0.50 |

(c) (2 points) Compute the conditional distribution $\mathbb{P}(C|N = naughty)$ and express it as a table. Show your work/calculations.

| $C$ | $\mathbb{P}(C\|N = naughty)$ |
|---|---|
| *chocolate* | $0.05/0.40 = 0.125$ |
| *coal* | $0.35/0.40 = 0.875$ |

(d) (2 points) Compute the conditional distribution $\mathbb{P}(N|C = chocolate)$ and express it as a table. Show your work/calculations.

| $N$ | $\mathbb{P}(N\|C = chocolate)$ |
|---|---|
| *nice* | $0.45/0.50 = 0.90$ |
| *naughty* | $0.05/0.50 = 0.10$ |

(e) (1 point) Is it true that $N \perp\!\!\!\perp C$? (i.e., are they statistically independent?) Explain your reasoning.

No, because $N$ and $C$ are independent if and only if $\mathbb{P}(N)P(C) = P(N, C)$.

(f) (3 points) Suppose Santa has other secret lists that track your other behaviors and qualities throughout the year such as manners (polite/rude), house-chores completion (many/few), school grades (high/low), denoted as the random variables $M$, $H$, $G$ respectively. Is it possible to compute $\mathbb{P}(N, C, M, H, G)$ as a product of five terms? We require that each term be in the form $P(X_i|X_j, \ldots, X_k)$ where there is one random variable to the left of the vertical bar "|" and there may be 0 or more random variables to the right of the vertical bar. (If there are none to the right, then the vertical bar does not appear. Thus $P(X_i)$ as a term is fine.) If it is possible to express the distribution as such a product, then show the details of such a product. What assumptions need to be made, if any? Otherwise, explain why it is not possible.

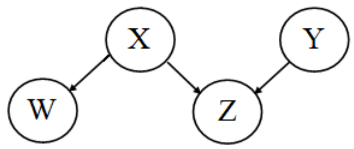
Yes, using the chain rule we get

$$\mathbb{P}(N, C, M, H, G) = \mathbb{P}(N)\mathbb{P}(C|N)\mathbb{P}(M|N, C)\mathbb{P}(H|N, C, M)\mathbb{P}(G|N, C, M, H).$$

No other assumptions need to be made in order to use the chain rule.

## 4 Bayes Net Structure and Meaning (Phuong)

(10 points) Consider the Bayes net with its graph is shown below:




Random variable $W$ has a domain with four values $\{w_1, w_2, w_3, w_4\}$; the domain for $X$ has three values: $\{x_1, x_2, x_3$; $Y$'s domain has three values: $\{y_1, y_2, y_3\}$; and $Z$'s domain has four values: $\{z_1, z_2, z_3, z_4\}$.

(a) (3 points) Give a formula for the joint distribution of all four random variables, in terms of the marginals (e.g., $\mathbb{P}(X = x_i)$), and conditionals that must be part of the Bayes net (e.g., $\mathbb{P}(Z = z_m | X = x_j, Y = y_k)$).

$$\boxed{\mathbb{P}(X)\mathbb{P}(Y)\mathbb{P}(W|X)\mathbb{P}(Z|X,Y)}$$

(b) (1 point) How many probability values belong in the (full) joint distribution table for this set of random variables?

$$\boxed{4 \cdot 3 \cdot 3 \cdot 4 = 144}$$

(c) (2 points) For each random variable: give the number of probability values in its marginal (for $X$ and $Y$) or conditional distribution table (for the others).

$W$: $4 \cdot 3 = 12$

$X$: $3$

$Y$: $3$

$Z$: $3 \cdot 3 \cdot 4 = 36$

(d) (4 points) For each random variable, give the number of *non-redundant* probability values in its table from (c).

$W$: $(4 - 1) \cdot 3 = 9$

$X$: $(3 - 1) = 2$

$Y$**:** $(3-1)=2$

$Z$**:** $(4-1)\cdot 3\cdot 3=27$

The (-1) comes from the fact that we know that the conditioned variable must sum to 1. Once we have computed the probability of one value of the conditioned variable, we can determine the other one (degrees of freedom).

# 5 D-Separation (Steve)

(20 points) Consider the Bayes Net graph $\beta$ below, which represents the topology of a web-server security model. Here the random variables have the following interpretations:

**V** = Vulnerability exists in web-server code or configs.

**C** = Complexity to access the server is high. (Passwords, 2-factor auth., etc.)

**S** = Server accessibility is high. (Firewall settings, and configs on blocked IPs are permissive).

**A** = Attacker is active.

**L** = Logging infrastructure is state-of-the-art.

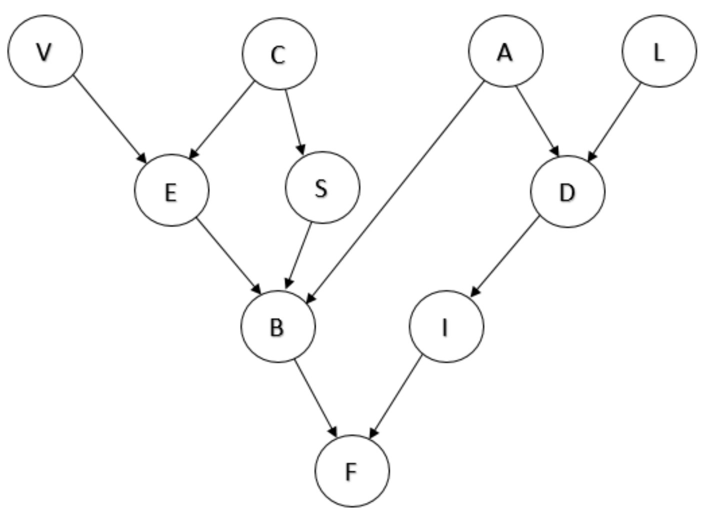
**E** = Exposure to vulnerability is high.

**D** = Detection of intrusion attempt.

**B** = Break-in; the web server is compromised.

**I** = Incident response is effective.

**F** = Financial losses are high (due to data loss, customer dissatisfaction, etc).



Let $\beta'$ be the undirected graph obtained from $\beta$ by removing the arrowheads from the edges of $\beta$. By an "undirected path" in $\beta$ we mean any path in $\beta'$. A "loop-free" path is any path in which no vertex is repeated.

(a) (5 points) List all loop-free undirected paths from A to C in the graph $\beta$. (ABEC), (ABSC), (ADIFBEC), (ADIFBSC)

(b) (5 points) Suppose random variable B is observed, and no others are observed. Then which (if any) of those paths would be active paths? Justify your answer. ABEC and ABSC are active paths. ABEC is active because its triples are ABE (common effect observed) and BEA (causal chain). Similarly ABSC is active because its triples are ABS (common effect observed) and BSC (causal chain). The path ADIFBEC and the path ADIFBSC are both inactive. Both contain the triple IFB (common effect unobserved and with no descendant observed) which blocks activity. That is enough to make the paths inactive. But they also have either the triple FBE or FBS which are blocked causal chains.

For each of the following statements, indicate whether (True) or not (False) the topology of the net guarantees that that the statement is true. If False, identify a path ("undirected") through which influence propagates between the two random variables being considered. (Be sure that the path follows the D-Separation rules covered in lecture.) The first one is done for you.

(c) $E \perp\!\!\!\perp S$: False (ECS)

(d) (1 point) $L \perp\!\!\!\perp B \mid A, I$ True

(e) (1 point) $I \perp\!\!\!\perp A \mid D, F$ False (IFBA)

(f) (1 point) $L \perp\!\!\!\perp V \mid D, E, F$ False (LDABSCEV)

(g) (1 point) $L \perp\!\!\!\perp V \mid A, D$ True

(h) (1 point) $D \perp\!\!\!\perp E \mid F, I$ False (DABE)

(i) (5 points) Suppose that the company hired an outside expert to examine the system and she determines that D is true: She has detected an intrusion event. Given this information, your job is to explain to management why getting additional information about L (Logging infrastructure being up-to-date) could have an impact on the probability of F (rfinancial losses being high). Give your explanation, for the manager of the company, using about between 2 and 10 lines of text, which should be based on what you know about D-separation, applied to this situation. However, your explanation should not use the terminology of D-separation but be in plain English. (You can certainly use words like “influence”, “probability”, “given”, but not “active path”, “triple”, or even “conditionally independent”).

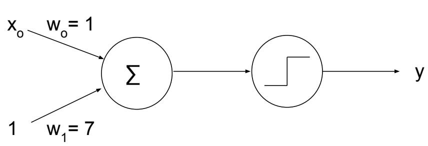
The question is whether learning anything new about L (logging infrastructure being up-to-date) could help us update our belief in F (financial losses high). We are given that D is true (there has been an intrusion event). Based on the diagram, we can see two immediate influences on D: one is A (an Attacker is active), and one is L. If we were to learn that L is likely true, then that suggests that the cause of the intrusion is indeed the attacker being active and not some deficiency in the logging infrastructure. This new information about the attacker being active will lead to a change in belief about B (Break-in to the web server), and that, in turn, impacts the probability of high financial losses.

# 6 Perceptrons (Anna)

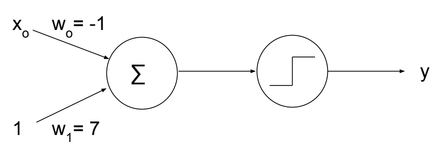
For all parts of this question perceptrons should output 1 if $\sum w_i x_i \geq \theta$ and 0 otherwise.

(a) (5 points) Assuming there will be two inputs $x_0$ and $x_1$, each with possible values in $\{0, 1\}$, give values for a triple of weights $\langle w_0, w_1, w_2 \rangle$ such that the corresponding perceptron would act as a NAND gate for the two inputs. (Weight $w_2$ is the bias weight.) Note that a NAND gate outputs 1 when at least one of the inputs is 0; it outputs 0 otherwise.
$\langle w_0, w_1, w_2 \rangle = \langle -1, -1, 1 \rangle$

(b) (5 points) Draw a perceptron, with weights, that accepts a single integer $x$ and outputs 1 if and only if the input is greater than or equal to $-7$. Be sure to include the bias input of value 1 and its weight in your diagram. Draw another perceptron that outputs 1 if and only if the input is less than or equal to 7.

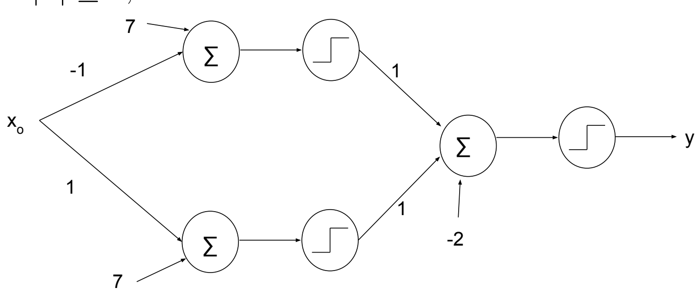



a.




a.

(c) (2 points) Using the previous perceptrons, create a two-layer perceptron that outputs 1 if $|x| \leq 7$, and 0 otherwise.




(d) (5 points) Suppose we want to train a perceptron to compare two numbers $x_0$ and $x_1$ and produce output $y = 1$ provided that $x_1$ exceeds $x_0$ by at least 5. Assume that the initial weight vector is: $\langle w_0, w_1, w_2 \rangle = \langle 0, 0, 1 \rangle$. Consider a first training example: $(\langle x_0, x_1 \rangle, y) = (\langle 1, 2 \rangle, 0)$. This says that with inputs 1, and 2, the output $y$ should be 0, since 2 exceeds 1 by only 1. What will be the new values of the weights after this training example has been processed one time? Assume the learning rate is 2.

The example is misclassified as a false positive. To update the weights, we subtract each the training example vector times learning rate from the weight vector to get (-2, -4). To update the bias, we subtract learning rate from the bias term to get -1. All together, the new weight vector is $\langle w_0, w_1, w_2 \rangle = \langle -2, -4, -1 \rangle$

(e) (3 points) Continuing with the last example, now suppose that the next step of training involves a different training example: $(\langle 2, 8 \rangle, 1)$. The output for this example should be 1, since 8 does exceed 2 by at least 5. Starting with the weights already learned in the first step, determine what the adjusted weights should be after this new example has also been processed once.

This example also gets incorrectly classified, though this time as a false negative. This time, we add the training example vector times the learning rate to the weight vector, and add the learning rate to the bias term to get $\langle w_0, w_1, w_2 \rangle = \langle 2, 12, 1 \rangle$

# 7 The Laws of Robotics (Emilia)

In the 1940's, Isaac Asimov introduced a set of three laws to govern robot behavior:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

(NOTE: You might also want to take a look at this cartoon `https://xkcd.com/1613/`)

Imagine you live in a world where personal robotic assistants are a bit more advanced than they are at present. They are still not commonplace, but in another 10 years (in our scenario), they could very well be. Further, imagine that you are a medical student who has had to move to virtual schooling due to a viral pandemic. Your well-endowed medical school, in partnership with faculty and industry collaborators, has presented you and all your classmates with "Standardized Patient Robots" – beta version humanoid personal assistant robots that they had quickly modified with knowledge of medical conditions, symptoms, and treatments (along with other enhancements) to enable you to continue your clinical training while in quarantine.

You were a computer science student before attending medical school, and you're also a science fiction aficionado. Your favorite author – Isaac Asimov. You can't resist hacking your new robot so that it now is programmed to obey Asimov's three laws before considering any other part of its programming. You feel quite pleased with yourself. Now, not only do you have a personal robot, you also have a personal bodyguard.

In other news, you recently volunteered to take part in a vaccine study being run out of your medical school and you've just been notified that you've been accepted. Furthermore, since the robot has programming that qualifies it as a medical assistant, you won't even need to leave your house to take part – the vaccine trials will be delivered to your front door and administered by the robot, which will then monitor you for any side effects.

Considering the above information, please answer the following questions. For each multiple choice question (MCQ), select what you think would be the most likely outcome. In the free response question that follows each MCQ explain what has happened and provide a justification for your selection, referring to Asimov's Laws by number as appropriate (e.g. "According to Asimov's 1st Law, ...").

For the multiple choice questions, any selection COULD be correct, depending on how the choice is justified. There is **no ONE correct answer**, so if students are having trouble with this question, please make sure they understand that. Consequently, there will be many, varied solutions. Some examples are provided for each question below.

(a) (1 point) It is time for your first vaccine dose. Which of the following takes place?

  (a) The robot calmly gives you the vaccine dose on schedule.
  (b) The robot runs an analysis on the vaccine contents before giving the dose on schedule.
  (c) The robot runs an analysis on the vaccine contents, after which it refuses to administer the dose.
  (d) The robot refuses to give you the scheduled vaccine dose.

(b) (4 points) Please justify the scenario you selected.

  If item a is chosen, it could be justified with the explanation that vaccines protect people, so the robot wants to protect the person from harm due to the disease, according to Asimov's First Law. If item d is chosen, it could be justified using the First Law as well, because 1) given the vaccine may cause pain (harm) or because vaccines can have bad side effects (even though it is violating the 2nd Law about obeying orders).

(c) (1 point) As a single student with an insane work schedule (normally), you tended to eat most of your meals out. Now, you are faced with the problem of how to feed yourself. You look around your kitchen and realize that all your frozen pizzas and instant noodles have run out. You get ready to make a quick run to the store down the street. What happens next?

  (a) The robot calmly reminds you to wear your mask and bring along some hand sanitizer.
  (b) The robot refuses to let you leave your house, and informs you it will do the shopping instead. You tell it the items you want and it returns an hour later with everything on your list.
  (c) The robot refuses to let you leave your house, and informs you it will do the shopping instead. You tell it the items you want and it returns an hour later with several bags of shopping, but few (if any) of the items that were on your list.

(d) The robot heads off to the store. Several hours later, it still has not returned and you go out to find it. You get to the store and find a long line of people, spaced 6 feet apart. You ask one of them what's going on, and one of them tells you some crazy robot is getting people groceries one at a time.

(d) (4 points) Please justify the scenario you selected.

For this question, I would hope people didn't choose a – leaving the house is never without risks, even when not in a pandemic, but if they did, they might still come up with a reasonable justification. If they selected c, the justification could be that frozen pizzas and ramen are not healthy food items and the robot has purchased much more nutritious items, in accordance with the First Law against allowing harm (from poor diet) to come to the human (even though it is violating the 2nd Law about obeying orders).

(e) (1 point) As your quarantine extends from week to week, you start feeling more and more stressed and depressed. You try to think of the robot as a kind of housemate, but its social interaction modules seem to be lacking. You make plans for a friend to visit and hang out with you in your backyard. What happens after he arrives?

(a) The robot accompanies you out to the yard and monitors your self distancing etiquette.

(b) The robot refuses to let your friend enter the backyard, telling him instead to go home and respect the quarantine.

(c) The robot welcomes your friend and monitors your visit, but refuses to let your friend leave several hours later.

(d) The robot lets your friend enter the backyard, but refuses to let you join him there, instead insisting that you communicate with your friend over the phone while looking at each other through a window.

(f) (4 points) Please justify the scenario you selected.

For this question, if c was selected, it could be justified by the robot because, according to the first law, the robot can't let a human being come to harm. All the precautions the robot would take with you, it is now going to take with your friend.

(g) (5 points) Now consider the following, slightly different, scenario. At some later point in time, Asimov added a 4th law to his Laws of Robotics – this law is known as the zeroth law and is meant to take priority over the original three laws. This new law is provided below:

0. A robot may not injure humanity, or, by inaction, allow humanity to come to harm.

Consider how the robot's responses to the situations described above would differ if you had programmed it with all four of these laws. Which selections might change, and how and why would they change?

Here, again, there will be different responses. One might be that in the first question, the student's answer didn't have the robot giving the vaccine. With the new rule, now it does, because even if the vaccine harms this individual, the knowledge gained from being in a vaccine trial can benefit humanity.

## 8 Markov Models (Shane)

(20 points) According to an unnamed source, the stock market can be modeled using a Markov model, where there are two states "bull" and "bear." The dynamics of the model are given:

| $S_{t-1}$ | $S_t$ | $P(S_t \vert S_{t-1})$ |
|---|---|---|
| bull | bull | 0.7 |
| bull | bear | 0.3 |
| bear | bull | 0.2 |
| bear | bear | 0.8 |

(a) (2 points) Suppose it's given that $S_0 = \text{bull}$. Compute the probability that $S_2 = \text{bull}$.

First, $P(S_1 = \text{bull}) = 0.7$, and $P(S_1 = \text{bear}) = 0.3$. Then, $P(S_2 = \text{bull}) = P(S_1 = \text{bull}) \cdot 0.7 + P(S_1 = \text{bear}) \cdot 0.3 = 0.7 \cdot 0.7 + 0.3 \cdot 0.2 = 0.55$.

(b) (4 points) Compute the stationary probabilities for bull and bear.

$P_\infty(\text{ bull }) = P_\infty(\text{ bull }) \cdot 0.7 + P_\infty(\text{ bear }) \cdot 0.2$.
$P_\infty(\text{ bear }) = 1 - P_\infty(\text{ bull })$.
$P_\infty(\text{ bull }) = 0.4$. $P_\infty(\text{ bear }) = 0.6$.

(c) (2 points)Now suppose that whenever it's a bull market, a certain company's (Acme, Inc) stock stays the same or rises in value with probability 0.8 and falls in value with probably 0.2.

When it's a bear market Acme's stock value stays the same or rises with probability 0.4 and falls with probability 0.6.

Suppose an observer cannot directly tell whether the state of the stock market is bull or bear, but can only see whether Acme's stock is "rising" or "falling."

| State $S$ | Observation $Q$ | $P(Q \vert S)$ |
|---|---|---|
| bull | rising | 0.8 |
| bull | falling | 0.2 |
| bear | rising | 0.4 |
| bear | falling | 0.6 |

Suppose $P(S_0 = \text{bull}) = 0.5$. If the observation at time 1 is "rising," what is the belief in $S$=bull right after the observation?

First compute $P(S_1 = \text{bull })$ without the observation using the forward transitions from $S_0 = \text{bull}$ and $S_0 = \text{bear}$. $P(S_1 = \text{bull }) = 0.5 * 0.7 + 0.5 * 0.2 = 0.45$.
Now take the observation into account using Bayes' rule:
$P(S_1 = \text{bull } | Q_1 = \text{rising }) = P(S_1 = \text{bull}, Q_1 = \text{rising })/P(Q_1 = \text{rising })$
$= P(Q_1 = \text{rising } | S_1 = \text{bull })P(S_1 = \text{bull })/P(Q_1 = \text{rising })$
Here $P(Q_1 = \text{rising }) = P(S_1 = \text{bull}, Q_1 = \text{rising }) + P(S_1 = \text{bear}, Q_1 = \text{rising }) = P(Q_1 = \text{rising } | S_1 = \text{bull})P(S_1 = \text{bull}) + P(Q_1 = \text{rising } | S_1 = \text{bear})P(S_1 = \text{bear}) = 0.8 \cdot 0.45 + 0.4 \cdot 0.55 = 0.36 + 0.22 = 0.58$.
So, $P(S_1 = \text{bull } | Q_1 = \text{rising }) = 0.36/0.58 = 0.621$.

(d) (2 points) Suppose at time 2, the observation is "falling". what is the belief in $S = \text{bull}$ right after that observation? (This belief will take into consideration the previous belief you computed above.)

First compute $P(S_2 = \text{bull })$ without the observation using the forward transitions from $S_1 = \text{bull}$ and $S_1 = \text{bear}$. $P(S_2 = \text{bull }) = 0.621 * 0.7 + 0.379 * 0.2 = 0.511$.
Now take the observation into account using Bayes' rule:
$P(S_2 = \text{bull } | Q_2 = \text{falling }) = P(S_2 = \text{bull}, Q_2 = \text{falling })/P(Q_2 = \text{falling })$
$= P(Q_2 = \text{falling } | S_2 = \text{bull })P(S_2 = \text{bull })/P(Q_2 = \text{falling })$
Here $P(Q_2 = \text{falling }) = P(S_2 = \text{bull}, Q_2 = \text{falling }) + P(S_2 = \text{bear}, Q_2 = \text{falling }) = P(Q_2 = \text{falling } | S_2 = \text{bull})P(S_2 = \text{bull}) + P(Q_2 = \text{falling } | S_1 = \text{bear})P(S_1 = \text{bear}) = 0.2 \cdot 0.511 + 0.6 \cdot 0.489 = 0.102 + 0.294 = 0.396$.
So, $P(S_2 = \text{bull } | Q_2 = \text{falling }) = 0.102/0.396 = 0.258$.

(e) (2 points) Suppose that the actual state sequence for the first four time steps is bear, bear, bear, bull

What is the probably of observing the sequence (starting at $t = 1$) rising, falling, rising?

Note that the first state $S_0$ has no influence, since no observation takes place at time 0, and state $S_1$ is given. We simply multiply the likelihoods of the emissions for each time step to get the probability of the sequence of observations:
$P(\text{ rising falling rising } | \text{ bear bear bear bull }) = P(\text{ rising } | \text{ bear }) \cdot P(\text{ falling } | \text{ bear }) \cdot P(\text{ rising } | \text{ bull }) = 0.4 \cdot 0.6 \cdot 0.8 = 0.192$.

(f) (3 points) Now, what if the state sequence was bull, bear, bear, bull. What is the probability of observing the same sequence of stock changes as above?

Same, since the state 0 has no effect.

(g) (2 points) Which of these two state sequences is more likely, given that sequence of observations?

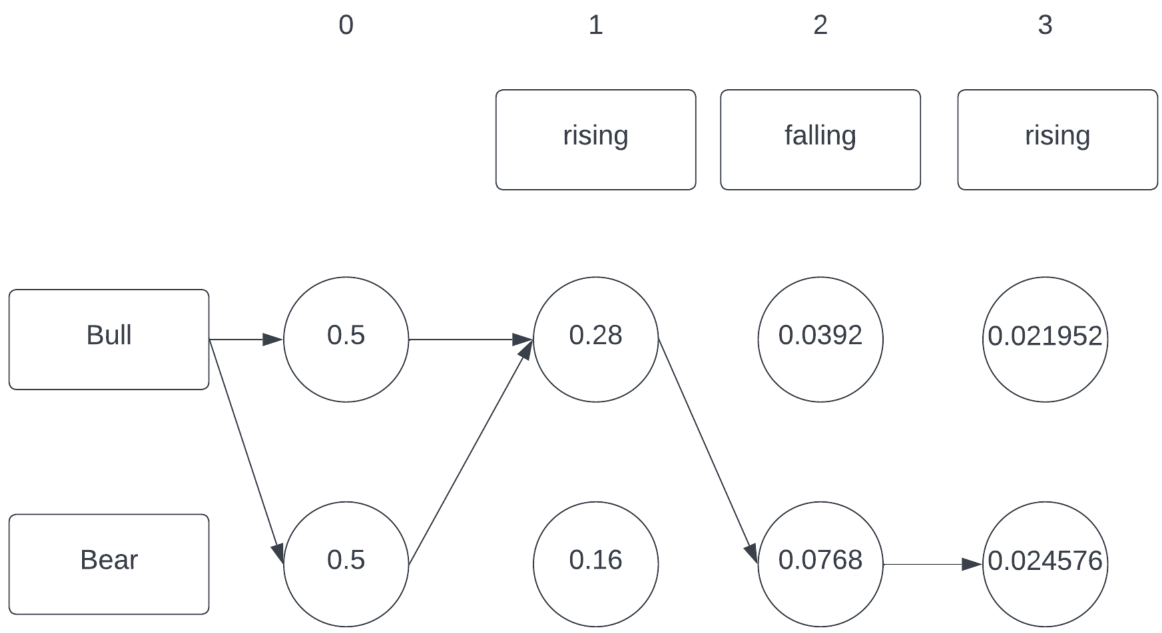
Equally likely.

(h) (3 points) Is there another state sequence that is even more likely? Explain.

The sequence bull bull bear bull, or the sequence bear bull bear bull, would have a likelihood of $0.8 \cdot 0.6 \cdot 0.8 = 0.384$, which is much more likely than either of the other sequences.

(i) (optional, not for credit) Use the Viterbi algorithm to compute the most likely state sequence for this observation sequence. You should draw a trellis diagram. Assume that the initial probability distribution is $P(S_0 = \text{bull}) = 0.5$; $P(S_0 = \text{bear}) = 0.5$. A very concise and relatively clear video presentation of a simple Viterbi algorithm example is one by Luis Serrano at:

`https://www.youtube.com/watch?v=mHEKZ8jv2SY`




According to the diagram, the most likely state sequence for this observation sequence is bull/bear bull bear bear.

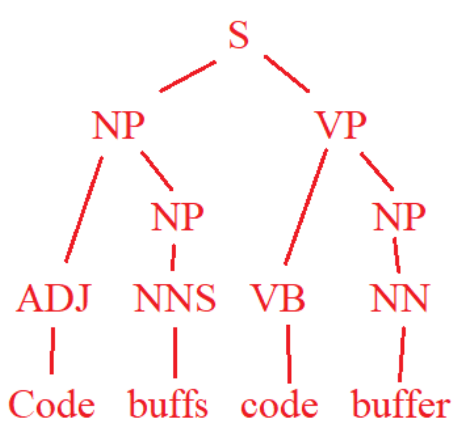
# 9 Probabilistic Context-Free Grammars (Steve)

(15 points) Consider the sentence, "Code buffs code buffer." Here there is ambiguity at multiple levels: lexical (word parts of speech and meanings), and syntactic (phrase structure). The semantics also vary. The sentence could mean that programming enthusiasts ("code buffs") write a program that manages a storage area (a buffer). It could also mean that a computer program smooths (buffs) the operation of a buffer that holds source code (a code buffer).

With the probabilistic context-free grammar given below, compare two parses, and compute a score for each one. Then identify the most probable parse using the scores. Assume the number at the right of a production is its conditional probability of being applied, given that the symbol to be expanded is that production's left-hand side.
(a) (5 points) Convert each probability into a score by taking score $= -\log_{10}(p)$. Round scores to 2 decimal places of accuracy. Write the production scores in the "__.__" blanks.

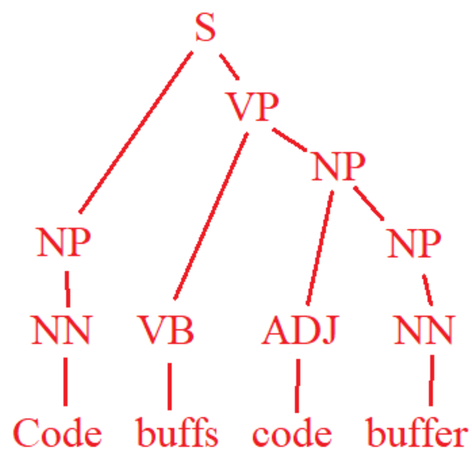
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| S | ::= VP | 0.1 | 1.00 | | VP | ::= VB NP | 0.4 | 0.40 |
| S | ::= NP VP | 0.8 | 0.10 | | VB | ::= code | 0.05 | 1.30 |
| NP | ::= ADJ NP | 0.4 | 0.40 | | VB | ::= buffs | 0.03 | 1.52 |
| NP | ::= NP PRP | 0.2 | 0.70 | | NN | ::= code | 0.06 | 1.22 |
| NP | ::= NN | 0.2 | 0.70 | | NN | ::= buffer | 0.02 | 1.70 |
| NP | ::= NNS | 0.2 | 0.70 | | NNS | ::= buffs | 0.01 | 2.00 |
| VP | ::= VB S | 0.2 | 0.70 | | ADJ | ::= code | 0.06 | 1.22 |
| VP | ::= VBG NP | 0.2 | 0.70 | | | | | |

(b) (3 points) Here is a first parse for the sentence. Compute the (total) score for this parse.



Using individual scores left-to-right, top-to-bottom, Score is: $0.10 + 0.40 + 0.40 + 0.70 + 0.70 + 1.22 + 2.00 + 1.30 + 1.70 = 8.52$.

(c) (5 points) Here is the second parse. Compute its score.




Using individual scores left-to-right, top-to-bottom, Score is: $0.10 + 0.40 + 0.40 + 0.70 + 0.70 + 1.22 + 1.52 + 1.22 + 1.70 = 7.96$.

(d) (2 points) Convert each score back to a probability and write them here as P1 and P2. Then tell which parse is more probable.

The probability of the first parse is $10^{-8.52} = 3.02\text{E}{-9}$.

The probability of the second parse is $10^{-7.96} = 1.10\text{E}{-8}$.

The first parse has a higher probability than that of the second parse, and so the first parse is the preferred parse, according to this PCFG.