SOLUTIONS to Assignment 6

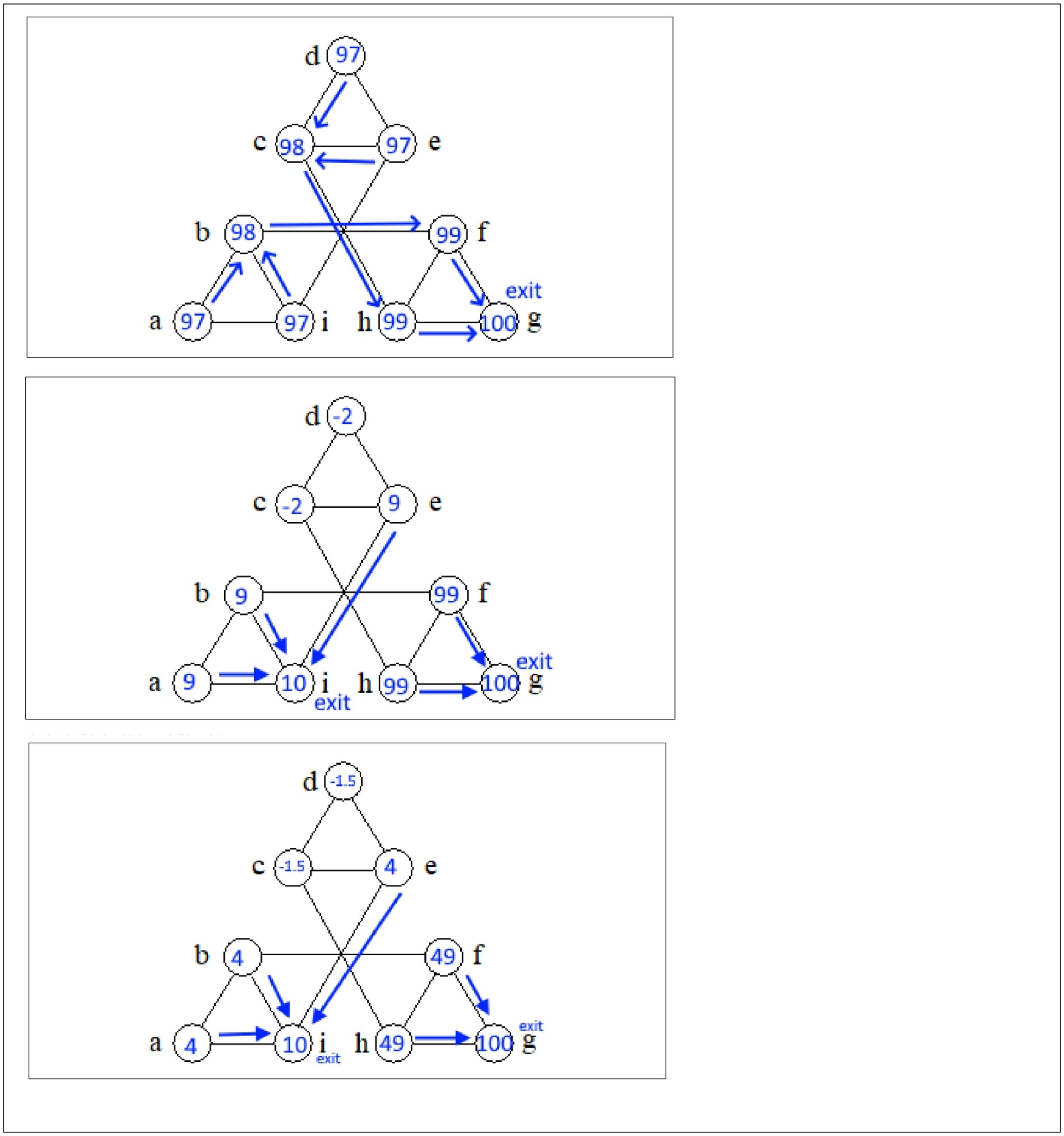
V.1.0

Prepared by the teaching staff.

CSE 415, Introduction to Artificial Intelligence

Paul G. Allen School of Computer Science and Engineering
University of Washington
Winter 2025.

1 MDP Basics (Xinming)



Important Note: The transition from state g (or i) to the terminal state consumes one move. Although no additional -1 reward is incurred in this move, there is instead a reward of 100 or 10 for making the transition. This means that for scenarios with a two-move time horizon (as in questions (c), (d), (e), and (f)), this extra move must be included in our calculations.

For example, in question (c), the values for nodes c and d are computed as:

$$\text{Value}(c) = \text{Value}(d) = (-1) + (-1) = -2.$$

This is because, within two moves, it is not possible to reach the terminal state and receive its reward.
In questions (e) and (f), when accounting for the discount factor, the calculations become:

$$\text{Value}(f) = \frac{100}{2} - 1 = 49, \quad \text{and similarly, Value}(h) = 49,$$

$$\text{Value}(a) = \frac{10}{2} - 1 = 4, \quad \text{and similarly, Value}(b) = \text{Value}(e) = 4,$$

Additionally, the value at node c is:

$$\text{Value}(c) = \frac{-1}{2} + (-1) = -1.5.$$

2 Reinforcement Learning (Sahil)

Use this MDP diagram to help you as you answer the following questions about Q Learning.

A	B (-10)	C	D	E	F	G (+10)

Use Q Learning to estimate the true q values of this MDP given the observed state transitions below. For each state transition, indicate which q value is changing by giving the state and action associated with that value and giving its new value. You may wish to keep your own diagram of the current values while you work through the given transitions.

You may assume that at states $\{A, C, D, E, F\}$ there are two allowed actions, {Left, Right}, and at states $\{B, G\}$, the only allowable action is the Exit action. Make sure to process the transitions in the order provided. All Q values are initialized to 0. Use a discount factor of $\gamma = 0.8$ and a learning rate of $\alpha = 0.5$.

Note that although a description of the MDP parameters are given in the next part of this question, Q Learning does not rely on knowing these parameters and this problem may not use the same rules and parameters.

State Transitions				New Q Values		
State (s)	Action (a)	New State (s')	Reward (r)	State	Action	Q(S,A)
A	Right	B	-1	A	Right	-0.5
B	Exit	Terminal	-10	B	Exit	-5
E	Right	D	-1	E	Right	-0.5
D	Right	E	-1	D	Right	-0.5
E	Right	F	-1	E	Right	-0.75
F	Right	G	-1	F	Right	-0.5
G	Exit	Terminal	10	G	Exit	5

Q Learning Equation: $Q_{k+1}(s,a) = (1-\alpha)Q_k(s,a) + \alpha\left(R(s,a,s') + \gamma \max_{a'} Q_k(s',a')\right)$

In the 5th transition, we have already observed a transition for the same Q value earlier in the process. Because of this, we update the Q Value using the learning rate to "average" between the two values, giving a different value in this transition than in other steps.

3 Joint Distributions and Inference (Emilia)

Let C represent the proposition that it is cloudy in Seattle. Let R represent the proposition that it is raining in Seattle.
Consider the table given below.

C	R	$P(C, R)$
cloudy	*rain*	0.53
cloudy	*sun*	0.13
clear	*rain*	0.02
clear	*sun*	0.32

(a) (2 point) Compute the marginal distribution $P(C)$ and express it as a table.

C	$P(C)$
cloudy	0.66
clear	0.34

(b) (2 point) Similarly, compute the marginal distribution $P(R)$ and express it as a table.

R	$P(R)$
rain	0.55
sun	0.45

(c) (2 point) Compute the conditional distribution $P(R|C = cloudy)$ and express it as a table. Show your work/calculations.

R	$P(R\|C = cloudy)$
rain	$0.53/0.66 = 0.80\overline{3}$
sun	$0.13/0.66 = 0.19\overline{6}$

(d) (2 point) Compute the conditional distribution $P(C|R = sun)$ and express it as a table. Show your work/calculations.

C	$P(C\|R = sun)$
cloudy	$0.13/0.45 = 0.2\overline{8}$
clear	$0.32/0.45 = 0.7\overline{1}$

(e) (3 points) Is it true that $C \perp\!\!\!\perp R$? (i.e., are they statistically independent?) Explain your reasoning.
No, because C and R are independent if and only if $P(C)P(R) = P(C, R)$.

(f) (4 points) Suppose you decide to track additional weather patterns of Seattle such as temperature (hot/cold), humidity (humid/dry), and wind (windy/calm) denoted as the random variables T, W, H respectively. Is it possible to compute $P(C, R, T, W, H)$ as a product of five terms? If so, show your work. What assumptions need to be made, if any? Otherwise, explain why it is not possible.

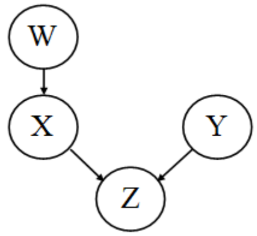
Yes, using the chain rule we get

$$P(C, R, T, W, H) = P(C|R, T, W, H)P(R|T, W, H)P(T|W, H)P(W|H)P(H).$$

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

4 Bayes Net Structure and Meaning (Krish)

Consider a Bayes net whose graph is shown below.



Random variable W has a domain with two values $\{w_1, w_2\}$; 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 two values: $\{z_1, z_2\}$.

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

(b) (1 point) How many probability values belong in the (full) joint distribution table for this set of random variables? $2 \cdot 3 \cdot 3 \cdot 2 = 36$

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

W: 2

X: 6

Y: 3

Z: 18

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

W: 1

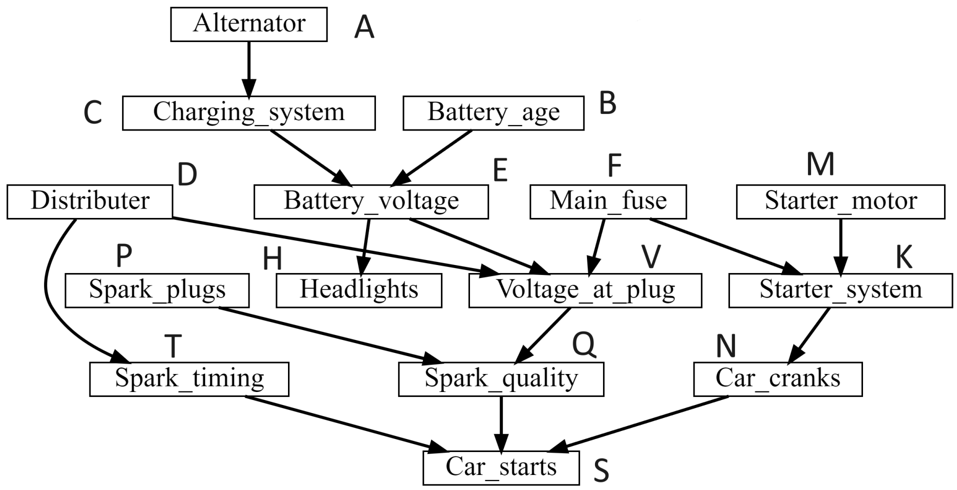
X: 4

Y: 2

Z: 9

5 D-Separation (Krish)

(20 points)
Consider the automobile engine-diagnosis Bayes Net below.



(a) (1 point) Assuming no observations have been made, are A and B independent?

Yes. They are both roots of the Bayes Net.

(b) (1 point) Assuming no observations have been made, are A and C necessarily (i.e., no matter what values the conditional probability table for C given A might contain) independent?

No, C is generally dependent on A, because of the arrow from A to C, although its CPT could be set to make it independent of A.

(c) (5 points) List all "undirected" paths from C to S.

CEVQS, CEVDTS, CEVFKNS

(d) (3 points) Which of the above are active paths?

those active: CEVQS

(e) (2 points) Is it guaranteed that $A \perp\!\!\!\perp S | Q, C$? If not, give an active path from A to S.

Yes, the triple ACE is inactive because C is observed.

(f) (2 points) Is it guaranteed that $B \perp\!\!\!\perp S | V, K, T$? If not, give an active path from B to S.

Yes, each undirected path is inactive: BEVFKNS is blocked at K BEVQS is blocked at V BEVDTS is blocked at T.

(g) (2 points) Is it guaranteed that $P \perp\!\!\!\perp B | S$? If not, give an active path from P to B.

No, an active path is PQVEB. The triple PQV is active because S is an observed descendant of Q.

(h) (2 points) Is it guaranteed that $A \perp\!\!\!\perp M | F, V, S$? If not, give an active path from A to M.

No. ACEVDTSNKM is active.

(i) (2 points) What is the longest loop-free undirected path you can find in this graph? What nodes would need to be observed to make it an active path?

> Note that the answer to (h) provides a valid part of an answer here. More generally, though, ACEVDTSNKM, ACEVDTSNKF and ACEVFKNSTD or the reversal of any of these three are all paths of length 9 edges (or 10 nodes).
> Observe S to make all of these active. Observing V in addition to S does not change the activity of the paths, but slightly changes the justification for triples where V is a common effect. When V is not observed, S is an observed descendant of V. If V is observed, then it becomes a common effect observed.

6 Choosing Hidden Sequences (Xinming)

(40 points) Consider the Jones family – Mom, Dad, Jim, Sally. Jim and Sally are twin siblings, and both 16 years old. Once a month, the family goes out for ethnic food – it's either Thai or Indian. Jim has his list of favorite restaurants, and Sally has her own list of her favorites. Each month, Dad uses one of the lists to choose what type of food the family will eat during their outing. He tends to stick with the same list more often than switch, and he tends to slightly favor Jim's list, maybe because of his own preferences.
With the current list, Dad selects a restaurant at random, assuming the restaurants on the list are equally likely to be picked. But they don't necessarily go to that restaurant. If the restaurant is Indian, Jim gets to pick any Indian restaurant from his list, and they go there. If the restaurant is Thai, then Sally gets to pick any Thai restaurant from her list, and they go there.

```
Jim's list:
  My-Pad-Thai (Thai), Punjabi Kitchen (Indian), Delhi Curry (Indian).
Sally's list:
  Bangkok Bites (Thai), Mango Mansion (Thai),
  Siam Shack (Thai), Delhi Curry (Indian).
```

Then after Dad makes his random selection from the current list, he uses some kind of spinning device that we don't understand, and he determines whether to switch lists for next time according to the following transition probabilities.
If the current list is Jim's, the probably of staying with Jim's is 0.7, and switching to Sally's is 0.3. If the current list is Sally's the probably of staying with Sally's is 0.6, and switching to Jim's is 0.4.

The way to solve this problem is to specify a Hidden Markov Model that represents the processes of this family activity. Let $t = 0$ represent the first month (January in the starting year) of the family eating out. The next month, February, corresponds to $t = 1$, March to $t = 2$, etc. The state variable for month t is written X_t and its two possible values are $X_t =$ Sally or $X_t =$ Jim, meaning Sally's list or Jim's list. The possible observations, "Indian" and "Thai" will be values of emission variables E_t. For example, $E_0 =$ Indian. The transition model comes from the story.

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
Jim	Jim	0.7
Jim	Sally	0.3
Sally	Jim	0.4
Sally	Sally	0.6

(a) (4 points) Assume Dad starts the year using Sally's list for January's outing. He's public about that, but in subsequent months, he does not tell anyone which list he is using.

One possible 3-month sequence is list uses is: Sally's, Sally's, Jim's (for January, February, March). What is the probability of that sequence of list usages, without considering the ethnicities of the restaurants they visited?

The probability of of the sequence Sally, Sally, Jim can be obtained from the transition model, without regard to the emission model.
$P(X_0 = \text{Sally}) = 1$, because that is given in the problem.
$P(X_1 = \text{Sally } | X_0 = \text{Sally}) = 0.6$.
$P(X_2 = \text{Jim } | X_1 = \text{Sally}) = 0.4$.
The probability of the sequence is the product of these probabilities: $1 \cdot 0.6 \cdot 0.4 = 0.24$.

(b) (3 points) Suppose that quarter of the year they end up going to an Indian restaurant in January, a Thai restaurant in February, and a Thai restaurant in March. Taking that additional information into consideration. what is the probability of that same list-usage sequence and seeing that sequence of cuisine choices?
(i.e., compute $\text{P}(X_1 = \text{Sally}, X_2 = \text{Sally}, X_3 = \text{Jim}, E_1 = \text{Indian}, E_2 = \text{Thai}, E_3 = \text{Thai})$.

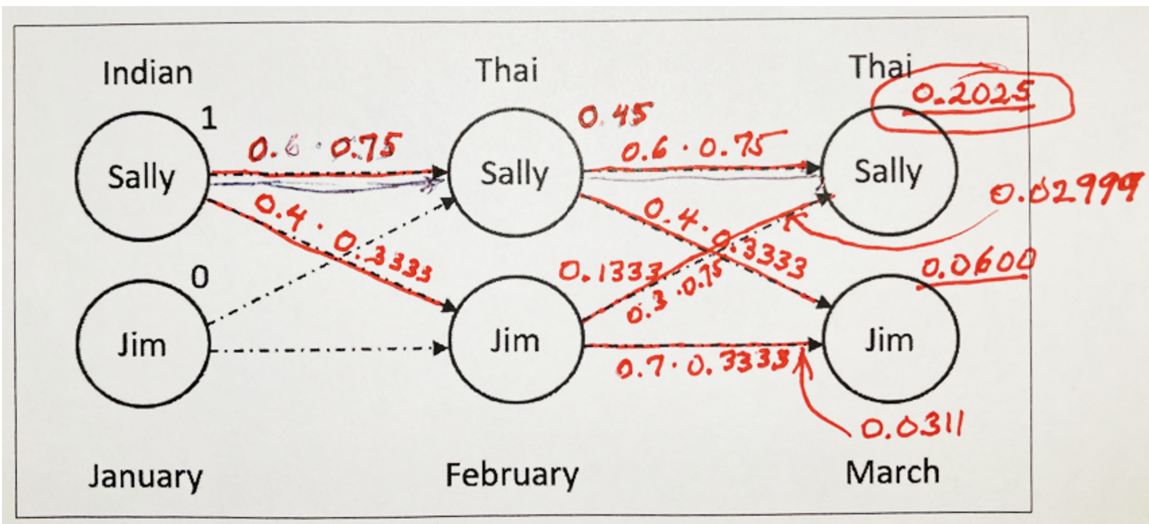
Using the chain rule, and conditional independence in Bayes nets, we have
$\text{P}(X_1 = \text{Sally}, X_2 = \text{Sally}, X_3 = \text{Jim}, E_1 = \text{Indian}, E_2 = \text{Thai}, E_3 = \text{Thai}) =$

$$\text{P}(X_1 = \text{Sally})\cdot\text{P}(E_1 = \text{Indian}|X_1 = \text{Sally})\cdot\text{P}(X_2 = \text{Sally}|X_1 = \text{Sally})\cdot$$
$$\text{P}(E_2 = \text{Thai}|X_2 = \text{Sally})\cdot\text{P}(X_3 = \text{Jim}|X_2 = \text{Sally})\cdot\text{P}(E_3 = \text{Thai}|X_3 = \text{Jim})$$

This is $1.0 \cdot (1/4) \cdot 0.6 \cdot (3/4) \cdot 0.4 \cdot (1/3) = 0.015$

(c) (16 points) What is the most likely sequence of list usages by Dad given that same restaurant ethnicity sequence? (Indian, Thai, Thai). Use the diagram template below to create a trellis diagram for this problem. If the probability coming into a dashed arrow is 0, leave it dashed. Otherwise, write over the dashed line to make it a more solid arrow. Before identifying the most likely list-usage sequence, compute, at each node, the probability of reaching that node along a most likely path from the Sally-in-January node. The probability value 1 is provided on that starting node, since it's a given in this exercise. Naturally, the probability of getting to the Jim-in-January node is 0, because we're given that Sally's list is the one being used in the first month (January). Use the Viterbi algorithm to get these probabilities at the other 4 nodes of the diagram. Finally, highlight the most probable path that starts at Sally-in-January and ends at one of the nodes in the March column. (Hint: the HMM worksheet given out on Nov. 6 involves a similar computation.) Show the factors you are multiplying on the appropriate edges of the trellis. For the March column, at least, use a calculator to show the two probabilities to 4 decimal places.

To compute the most probable sequence of list usages, given the sequence of food ethnicities, we use the Viterbi algorithm. Proceeding left to right, it computes, for each node of the trellis diagram, the probability of arriving at that node along a most probable path. The image below shows the computations needed in red. Generally each edge going out from a non-zero probability node has a product of two values on it: the transition probability for making that state transition, multiplied by the likelihood of observing the specific emission given the destination state of the edge. That product is multiplied with the probability computed at the tail endpoint of the edge to get a provisional value for the probability at the head (destination) of the edge. It's provisional because there can be multiple edges coming into that destination, and we'll use the maximum probability among these provisional values. Once we have these probability values at each of the March nodes, we'll identify the larger one and trace the path back to the left by using the argmax (rather than the max used to compute the value) at each node reached.



Here we can see that the probability of reaching node Sally-in-February by a most probable path is $1 \cdot 0.6 \cdot 0.75 = 0.45$. This path starts at Sally-in-January and goes to Sally-in-February. The probability of coming to Sally-in-February from Jim-in-January is 0, since the probability of Jim-in-January itself is 0.

There are two ways to arrive at Sally-in-March. One is coming from Sally-in-February which has probability $0.45 \cdot 0.6 \cdot 0.3333 = 0.2025$. The other is coming from Jim-in-February which has probability $0.1333 \cdot 0.3 \cdot 0.75 = 0.02999$. The max of these is 0.2025, meaning that arriving from Sally-in-February is more likely than arriving from Jim-in-February. We also have to consider the possibility of ending up in Jim-in-March. So we compute the probabilities of arriving there from each of Sally-in-February and Jim-in-February. We get 0.0600 and 0.0311 respectively, of which 0.0600 is the max. Comparing this 0.0600 to Sally-in-March's max (0.2025) indicates that the most likely path ends at Sally-in-March, not Jim-in-March. Backtracing gives us the sequence Sally, Sally, Sally as the most likely, given what we know about the models (transition model and emission model) and the observation sequence.

(d) (2 points) Suppose Mom somehow gets suspicious that Sally's preferences are not being considered fairly, in comparison with Jim's. What is a possible basis for that?

Perhaps she knows Dad has some preference for Indian food himself, or simply suspects that he is using Jim's list more often than Sally's. This might be, correctly or not, the random result that it seems to Mon that the family seems to go for Indian more often than Thai.

(e) (8 points) Compute the stationary probability of Jim's list vs Sally's list being used. (Assume that when the sequence extends beyond 12 months, Dad does NOT automatically go back to Sally's list each January. In other words, the Markov Model represented by the transition CPT is not limited to 12 time steps.)

Show the needed equations before you solve them, and show the steps you use in solving them.

The equations:
Let $x = \mathrm{P}_\infty(\text{Sally}) = 0.6{\cdot}\mathrm{P}_\infty(\text{Sally}) + 0.3{\cdot}\mathrm{P}_\infty(\text{Jim})$
$\mathrm{P}_\infty(\text{Jim}) = 0.4{\cdot}\mathrm{P}_\infty(\text{Sally}) + 0.7{\cdot}\mathrm{P}_\infty(\text{Jim})$
$\mathrm{P}_\infty(\text{Jim}) = 1 - x$
Solving for x:
$0.7x = 0.3$
$x = 0.3/0.7 = 0.429 = \mathrm{P}_\infty(\text{Sally})$
$\mathrm{P}_\infty(\text{Jim}) = 0.571$

(f) (4 points) Compute the marginal probabilities of the family's having Indian and Thai food on their outings, corresponding to the stationary probabilities of list usages.

$\mathrm{P}(\text{Indian}) = \mathrm{P}(\text{Sally}) \cdot \mathrm{P}(\text{Indian} \mid \text{Sally}) + \mathrm{P}(\text{Jim}) \cdot \mathrm{P}(\text{Indian} \mid \text{Jim})$
$= 0.429 \cdot (1/4) + 0.571 \cdot (2/3)$
$= 0.488$
P(Thai) = 1 - P(Indian) $= 1 - 0.488 = 0.512$.

(g) (3 points) To what extent are these distributions (that you computed in parts e and f) biased against or for each child's preferences? Explain. There are three pertinent points you can describe here.

The stationary distribution shows that in the long term, after some months have passed, there is a bias towards greater use of Jim's list than Sally's. However, a mitigating factor is that Sally's is used first, no matter what, and Dad tends to stick with the current list more often than switch. Thus there is some bias not only at the first month of the process, but that bias lingers for a while. The next point is that whenever Sally's list is being used, there is a 0.75 probability that the family will go to a Thai restaurant (presumably Sally's favorite food ethnicity), but when Jim's list is being used, there is only a 0.667 probability of going for Indian (presumably his favorite food ethnicity). This is an argument for a bias in Sally's favor. Finally, one can make the point that Both Sally's list and Jim's list contain a mix of Indian and Thai restaurants, and from this data alone, we can't conclude that, say, Sally actually prefers Thai food to Indian food in general. Perhaps the bias might be that whenever they eat Indian, Jim gets to choose the restaurant and it might never be the sole Indian restaurant on Sally's list. That's also true the other way around. Sally might never choose the Thai restaurant on Jim's list. And there is that slight bias in the marginal probability of Indian, that this happens more often, in the long run, with Jim choosing than with Sally choosing. There are more possible justifications for either an argument for bias or for "random enough" but these cover the main points.

7 Perceptrons (Sahil)

(20 points) You are planning an exciting road trip across California and need to decide which cities to include in your itinerary. You classify cities into two categories: Must-Visit Cities and Skip Cities. A city is labeled as Must-Visit (+1) if it is highly attractive for tourists, while a city is labeled as Skip (−1) if it does not meet your travel preferences. You decide to use a perceptron model to classify cities based on their features:
Dataset Features and Specification:
Natural Attractions: Represented as low $= -2$, moderate $= 0$, high $= 2$.
Accommodation Cost: Represented as cheap $= 1$, moderate $= 2$, expensive $= 3$.
Travel Distance: Represented as far $= -2$, near $= 2$.

Perceptron Parameters

- The perceptron has a weight vector, w, with four components: bias term, natural attractions, accommodation cost, and travel distance.
- Initial weights: $w = [1, 0, 0, 0]$.
- Threshold: 0.
- Learning rate: 1.

Example Number	Natural Attractions	Accommodation Cost	Travel Distance	Label
1	high	expensive	near	Must-Visit
2	moderate	moderate	far	Skip
3	high	moderate	near	Skip
4	moderate	expensive	near	Must-Visit
5	low	cheap	fast	???

Table 1: Road Trip Dataset

(a) (4 points) What would be the updated weights $\mathbf{w}$ after processing example number 1?

- **Input:** $x = [1, 2, 3, 2]$ (bias = 1, high = 2, expensive = 3, near = 2)
- **Label:** $y = +1$ (Must-Visit)
- **Initial Weights:** $w = [1, 0, 0, 0]$
- **Prediction:**

$$\hat{y} = \text{sign}(w \cdot x) = \text{sign}(1 \cdot 1 + 0 \cdot 2 + 0 \cdot 3 + 0 \cdot 2) = \text{sign}(1) = +1$$

Since the prediction matches the label ($\hat{y} = y$), the weights remain unchanged:

$$w = [1, 0, 0, 0]$$

(b) (3 points) What would be the updated weights $\mathbf{w}$ after processing example number 2?

- **Input:** $x = [1, 0, 2, -2]$ (bias = 1, moderate = 0, moderate = 2, far = -2)
- **Label:** $y = -1$ (Skip)
- **Weights:** $w = [1, 0, 0, 0]$
- **Prediction:**

$$\hat{y} = \text{sign}(w \cdot x) = \text{sign}(1 \cdot 1 + 0 \cdot 0 + 0 \cdot 2 + 0 \cdot -2) = \text{sign}(1) = +1$$

The prediction ($\hat{y} = +1$) is incorrect ($y = -1$), so the weights are updated:

$$w \leftarrow w + \eta y x = [1, 0, 0, 0] + 1 \cdot (-1) \cdot [1, 0, 2, -2] = [0, 0, -2, 2]$$

Updated Weights:

$$w = [0, 0, -2, 2]$$

(c) (3 points) What would be the updated weights $\mathbf{w}$ after processing example number 3?

- **Input:** $x = [1, 2, 2, 2]$ (bias = 1, high = 2, moderate = 2, near = 2)
- **Label:** $y = -1$ (Skip)
- **Weights:** $w = [0, 0, -2, 2]$
- **Prediction:**

  $$\hat{y} = \text{sign}(w \cdot x) = \text{sign}(0 \cdot 1 + 0 \cdot 2 - 2 \cdot 2 + 2 \cdot 2) = \text{sign}(-4 + 4) = \text{sign}(0) = 0$$

  Since $\hat{y} = 0$, the prediction does not match $y = -1$. The weights are updated:

  $$w \leftarrow w + \eta y x = [0, 0, -2, 2] + 1 \cdot (-1) \cdot [1, 2, 2, 2] = [-1, -2, -4, 0]$$

  Updated Weights:

  $$w = [-1, -2, -4, 0]$$

(d) (3 points) What would be the updated weights **w** after processing example number 4?

- **Input:** $x = [1, 0, 3, 2]$ (bias = 1, moderate = 0, expensive = 3, near = 2)
- **Label:** $y = +1$ (Must-Visit)
- **Weights:** $w = [-1, -2, -4, 0]$
- **Prediction:**

  $$\hat{y} = \text{sign}(w \cdot x) = \text{sign}(-1 \cdot 1 + -2 \cdot 0 + -4 \cdot 3 + 0 \cdot 2) = \text{sign}(-1 - 12) = \text{sign}(-13) = -1$$

  The prediction ($\hat{y} = -1$) is incorrect ($y = +1$), so the weights are updated:

  $$w \leftarrow w + \eta y x = [-1, -2, -4, 0] + 1 \cdot (+1) \cdot [1, 0, 3, 2] = [0, -2, -1, 2]$$

  Updated Weights:

  $$w = [0, -2, -1, 2]$$

(e) (3 points) What is your prediction for the label of example number 5 based on the final weights?

- **Input:** $x = [1, -2, 1, -2]$ (bias = 1, low = -2, cheap = 1, far = -2)
- **Weights:** $w = [0, -2, -1, 2]$
- **Prediction:**

  $$\hat{y} = \text{sign}(w \cdot x) = \text{sign}(0 \cdot 1 + -2 \cdot -2 + -1 \cdot 1 + 2 \cdot -2) = \text{sign}(4 - 1 - 4) = \text{sign}(-1) = -1$$

  Prediction: Skip ($y = -1$).

(f) (4 points) Is convergence guaranteed for this perceptron and data set? Why or why not?

(g) **Answer:** Convergence is not guaranteed if the dataset is not linearly separable. If a linear boundary cannot separate the data points x, the perceptron will oscillate and fail to converge.

8 AI and the Potential for Harm (Emilia)

(20 points)
Answers will vary.