CS 181 Artificial Intelligence (Fall 2023), Final Exam

Instructions

- Time: 10:30 – 12:10 (100 minutes)
- This exam is closed-book, but you may bring one A4-size cheat sheet. Put all the study materials and electronic devices (except a calculator) into your bag and put your bag in the front, back, or sides of the classroom.
- Two blank pieces of paper are attached, which you can use as scratch paper. Raise your hand if you need more paper.
- For multiple choice questions in Question 2–5:
  - □ means you should mark ALL choices that apply;
  - ○ means you should mark exactly ONE choice;
  - When marking a choice, please fill in the bubble or square COMPLETELY (e.g., ● and ■). Ambiguous answers will receive no points.
  - For each question with □ choices, you get half of the points for selecting a non-empty proper subset of the correct answers.

1 Multiple choices (10 pt)

Each question has one or more correct answers. Select all the correct answers. For each question, you get 1 point if your select all the correct answers and nothing else, 0 point if you select one or more wrong answers, and 0.5 point if you select a non-empty proper subset of the correct answers.

1	2	3	4	5
6	7	8	9	10

1. Which of the following statement(s) about Markov models, hidden Markov models (HMM), and dynamic Bayes nets (DBN) is/are correct?

   A. If several states have probability 0 in the initial distribution of a Markov model, then the stationary distribution will never be reached, no matter what the transition model is.

   B. Considering the formula of the Forward algorithm: $\alpha \times P(e_{t+1}|X_{t+1}) \times \sum_{x_t} P(x_t|e_{1:t}) P(X_{t+1}|x_t)$, the three parts separated by multiplication can be interpreted as "Normalize", "Update", and "Predict" from left to right.

   C. The time complexity of the Viterbi algorithm is $O(|X|^2 T)$ and the space complexity is $O(|X|T)$, where $|X|$ is the number of states and T is the total number of time steps.

   D. A discrete DBN with 4 triple-state variables in each time step can be converted to a HMM which has 4^3 possible values for each state variable.

2. Which of the following statement(s) about particle filtering is/are correct?

   A. Likelihood weighting is not proper for HMM because the weight of samples drops quickly.

   B. In the propagate forward step, each particle is moved by sampling its next position from the transition model $P(X_{t+1}|X_t)$.

   C. In the observe step, we need to do "count and normalize" to get the weight of each particle.

   D. In the resample step, each new sample should inherit the weight of the old sample from which it is generated.

3. Suppose we have an MDP with state space of size S and action space of size A. Which of the following statement(s) about value iteration and policy iteration is/are correct?

   A. Value iteration takes $O(SA^2)$ time for each iteration.

   B. In value iteration, the values often converge much faster than the policy extracted from the values.

   C. In policy iteration, we do policy improvement by doing policy extraction from the values computed in policy evaluation.

   D. Both value iteration and policy iteration can find optimal values.

   E. None of the above.

4. Which of the following statement(s) about MDP is/are correct?

   A. MDP could be viewed as randomized search problems, so we can solve them by expectimax search.

   B. In MDP, the successor state depends on not only the current state, but also the previous states.

   C. A policy of MDP is a mapping from each state to a reward value.

   D. A discount factor smaller than one helps the value iteration algorithm converge in general.

   E. None of the above.

5. Which of the following statement(s) about Q-learning is/are correct?

   A. To converge to the optimal policy, you need to act optimally during Q-learning.

   B. In the ϵ-greedy algorithm, agents act randomly with probability $(1 - \epsilon)$.

   C. Minimizing regret is necessary to learn the optimal policy.

   D. To converge to the optimal policy, you have to explore enough and eventually make the learning rate small enough.

   E. None of the above.

6. Which of the following statement(s) about reinforcement learning is/are correct?

   A. In model-based learning, we first learn the empirical MDP model and then solve the learned MDP.

   B. In passive reinforcement learning, we do offline planning without taking actions in the world and compute the optimal policy.

   C. Direct evaluation does not require any knowledge of transition and reward and it utilizes state connections.

   D. In temporal difference learning, the older samples are given exponentially less weights if the discount factor is less than 1.

   E. None of the above.

7. A recent paper finds that the performance of LLM (Large Language Model) degrades on test data when repeating training on the training data for additional epochs. Which **ONE** of the following is the most likely reason for this phenomenon?

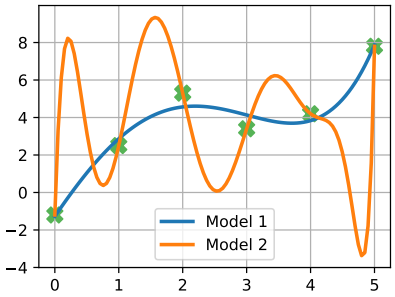
A. The model is overfitting on the test data. Since the model is trained on training data for multiple times, it fits the training data better and better, and it also fits the test data better and better.

B. The model is overfitting on the training data. Since the model is trained on training data for multiple times, it fits the training data better and better, but it does not generalize well to the test data.

C. The model is overfitting on the training data. Since the model is trained on training data for multiple times, the training data is no longer representative of the true data distribution, so the model does not generalize well to the test data.

D. The model is underfitting on the training data. Since the model is trained on training data for multiple times, it does not fit the training data well, and it does not fit the test data well either.

8. Suppose we are doing non-linear regression for a set of data points. The fitting result is shown in the figure below.



Generally, we prefer Model 1 over Model 2. Which **ONE** of the following is the most likely explanation for this preference?

A. According to Ockham's razor, we prefer the simplest hypothesis consistent with the data. Model 2 is obviously more complex than Model 1 and overfits the data.

B. According to Ockham's razor, we prefer the most complex hypothesis consistent with the data. Model 1 is obviously more complex than Model 2 and fits the data well.

C. Model 1 has a lower training error than Model 2, so we prefer Model 1.

D. Model 2 fits all the data points perfectly, which is not a good property for a regression model.

9. Which of the following statement(s) about k-means is/are correct?

A. k-means always converges to global optimal solutions.

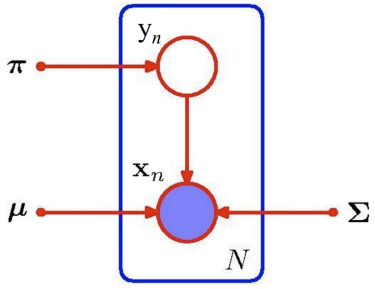
B. The final clustering results of k-means are usually dependent on the initial cluster centers.

C. k-means can only take Euclidean distance as its distance calculation method.

D. The point with the minimum squared Euclidean distance to a set of points is their mean.

E. None of the above.

10. Consider the sampling process of a Gaussian mixture model (GMM) shown in the figure below.



We choose a multivariate Gaussian component i with probability π_i. Then we generate a sample $\mathbf{x}_n$ from $\mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i)$ with its label $y_n = i$. Which of the following statement(s) about GMM is/are correct?

A. Given a data sample x_j, we have $\underset{i}{\operatorname{argmax}}\, P(y_j = i, \mathbf{x}_j) = \underset{i}{\operatorname{argmax}}\, \pi_i$.

B. If we are given label $y_j = i$ for each data sample $\mathbf{x}_j$, we can estimate the parameters of GMM by maximizing the likelihood $\prod_j \pi_i \mathcal{N}(\mathbf{x}_j|\mu_i, \Sigma_i)$.

C. If we do not know label y_j for each data sample $\mathbf{x}_j$, the objective for GMM parameter estimation is to maximize the marginal likelihood $\prod_j \sum_i \pi_i \mathcal{N}(\mathbf{x}_j|\mu_i, \Sigma_i)$.

D. If we do the generation process many times to produce a lot of data samples, it is most likely true that the larger π_i is, the more data samples are labeled as i.

2 Hidden Markov Models (10 pt)

Intrepid scientists and lovers Katia and Maurice Krafft died in a volcanic explosion doing the very thing that brought them together: unraveling the mysteries of volcanoes by capturing the most explosive imagery ever recorded.

Now they have time-traveled to the present! We can use Hidden Markov Models to assist them in safely exploring the volcano, avoiding risks, and rewriting their tragic ending.

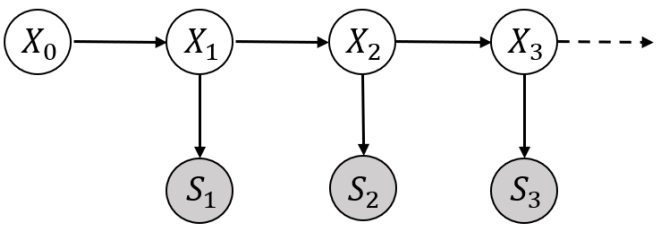
Katia and Maurice Krafft need to go deep into the active volcano to seek for a rare mine. We can only speculate about the existence of rare mines based on the external behavior of the volcano. The behaviour of volcanoes can be divided into three categories: 1. erupting red magma (RED); 2. emitting gray smoke (GREY); 3. being quiet (QUIET).

2.1 Construct Hidden Markov Model (4 pt)

We construct a Hidden Markov Model with the following statistics:

- The prior probability of mine existence (with no observations) is 0.3.
- The probability of mine existence on day t is 0.6 if the mine exists the previous day, and 0.1 if not.
- The probability of the volcano's behaviour being RED/GREY/QUIET is 0.3/0.5/0.2 respectively if the mine exists. On the other hand, the probability of RED/GREY/QUIET is 0.2/0.2/0.6 if the mine doesn't exist.

For state X_t, we have $X_t = 1$ if the mine exists on day t and $X_t = 0$ if not. For evidence S_t on day t, we have $S_t = 0$ if the volcano behaves RED, $S_t = 1$ if the volcano behaves GREY, and $S_t = 2$ if the volcano behaves QUIET.



2.1.1 Complete The Model (2 pt)

Based on the probabilities and model above, fill in the following tables of the initial distribution, transition model and emission model. We have finished part of them for you.

X_0	$P(X_0)$
0	0.7
1	0.3

X_t	X_{t+1}	$P(X_{t+1}\|X_t)$
0	0	
0	1	
1	0	
1	1	

X_t	S_t	$P(S_t\|X_t)$
0	0	0.2
0	1	0.2
0	2	0.6
1	0	
1	1	
1	2	

2.1.2 Stationary Distribution (2 pt)

What is the value of $P(X_\infty = 1)$ in the stationary distribution? Choose the correct answer by filling up the circle like ●.

○ 0.05 ○ 0.10 ○ 0.15 ○ 0.20 ○ 0.25

2.2 Filtering (3 pt)

We can calculate the posterior distribution over the most recent state given all evidence (i.e., $P(X_t|s_{1:t})$) to do filtering. It helps Katia and Maurice determine whether day t is a good day to seek for mine. Given $s_1 = 1$, $s_2 = 0$, $s_3 = 2$, What is the value of $P(X_3 = 1|s_{1:3})$? Choose the correct answer by filling up the circle like ●.

○ 0.10 ○ 0.13 ○ 0.16 ○ 0.19 ○ 0.21

2.3 Most Likely Explanation (3 pt)

Katia and Maurice want to know if they made the right decisions in their past adventures. They sought for the mine in day $t = 2, 3$ and didn't take action in day $t = 0, 1$, given $s_1 = 1$, $s_2 = 0$, $s_3 = 2$. According to most likely explanation (i.e., $\arg\max x_{0:t} P(x_{0:t}|s_{1:t})$), how many right choices have they made in the total four days? Choose the correct answer by filling up the circle like ●.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4

3 MDP and RL: Video Pushing (10 pt)

People nowadays spend more and more time browsing videos on their cellphones. For video platforms such as Bilibili and TikTok, it is important to push proper content to different users to maximize the time they spend.

We now formulate this problem as an MDP (markov decision process). There are three state $\{s_1, s_2, s_3\}$, indicating different emotion of user. The actions on these video platforms are simplified to either "change" or "remain", i.e., whether to change the video content to view next. Different actions may lead to different rewards.

Set the discount factor $\gamma = 0.5$.

3.1 MDP: Value Iteration (5pt)

For this question, suppose we have known the underlying MDP, as shown in the table below.

s	a	s'	$T(s, a, s')$	$R(s, a, s')$
s_1	Remain	s_2	1.0	0.0
s_1	Change	s_3	1.0	-2.0
s_2	Remain	s_1	0.4	-1.0
s_2	Remain	s_3	0.6	2.0
s_2	Change	s_1	0.6	2.0
s_2	Change	s_3	0.4	-1.0
s_3	Remain	s_1	0.6	2.0
s_3	Remain	s_2	0.4	2.0
s_3	Change	s_1	0.4	2.0
s_3	Change	s_2	0.6	0.0

3.1.1 Value iteration (3pt)

Suppose after iteration k of value iteration, we end up with the following values:

$V_k(s_1)$	$V_k(s_2)$	$V_k(s_3)$
0.4	1.4	2.16

Calculate the result after the next iteration:

$V_{k+1}(s_1)$	$V_{k+1}(s_2)$	$V_{k+1}(s_3)$

3.1.2 Policy extraction (2pt)

Suppose value iteration has converged to the following values:

$V^*(s_1)$	$V^*(s_2)$	$V^*(s_3)$
0.8	1.8	2.6

Extract the optimal policy for s_1:

○ Remain ○ Change

Extract the optimal policy for s_2:

○ Remain ○ Change

3.2 Reinforcement learning

In practice, we may not know the underlying MDP, so we need to learn it from data. We now begin with model-based learning.

3.2.1 Model-based learning (2pt)

For the model described above, consider the following samples that the agents have encountered (these samples have been sorted for your convenience):

s	a	s'	r
s_1	Remain	s_2	0.0
s_1	Remain	s_2	0.0
s_1	Remain	s_2	0.0
s_1	Remain	s_3	−10.0
s_1	Remain	s_3	−10.0
s_1	Change	s_3	−8.0
s_1	Change	s_3	−8.0
s_1	Change	s_2	0.0
s_1	Change	s_2	0.0
s_1	Change	s_3	−8.0
s_2	Remain	s_1	−3.0
s_2	Remain	s_1	−3.0
s_2	Remain	s_1	−3.0
s_2	Remain	s_1	−3.0
s_2	Remain	s_3	0.0

s	a	s'	r
s_2	Change	s_1	−10.0
s_2	Change	s_1	−10.0
s_2	Change	s_1	−10.0
s_2	Change	s_1	−10.0
s_2	Change	s_3	0.0
s_3	Remain	s_1	0.0
s_3	Remain	s_2	6.0
s_3	Remain	s_2	6.0
s_3	Remain	s_1	0.0
s_3	Remain	s_1	0.0
s_3	Change	s_2	−8.0
s_3	Change	s_2	−8.0
s_3	Change	s_2	−8.0
s_3	Change	s_1	0.0
s_3	Change	s_2	−8.0

Estimate the following values based on the samples above:

$T(s_1, \text{Remain}, s_2)$ = _____, $R(s_1, \text{Remain}, s_2)$ = _____

$T(s_1, \text{Remain}, s_3)$ = _____, $R(s_1, \text{Remain}, s_3)$ = _____

3.2.2 Model-free learning (2pt)

Now, instead of estimating the transition and reward function, we will directly estimate Q values using Q-learning. Assume the step size for Q-learning $\alpha = 0.5$. Our current Q value is as follows:

	s_1	s_2	s_3
Remain	1.5	−0.5	2.7
Change	3.0	−6.0	2.1

Then the agent encounters the following samples

s	s	s'	r
s_1	Change	s_3	8.0
s_3	Change	s_1	0.0

Calculate the updated Q value after these two samples (round to 3 decimal places):

$Q(s_1, \text{Change})$ = _____, $Q(s_3, \text{Change})$ = _____

3.2.3 Exploration v.s. exploitation (1pt)

If the agent selects its action according to probability:

$$\Pr(a|s) = \frac{e^{Q(s,a)/\tau}}{\sum_{a'} e^{Q(s,a')/\tau}}$$

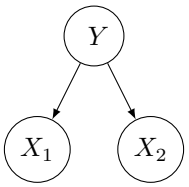
where τ is a parameter that is decreased over time. Then the agent's action is

○ mostly exploration. ○ mostly exploitation. ○ mix of both.

4 Supervised Machine Learning (10 pt)

4.1 Naïve Bayes (5 pt)

Answer questions about the following Naïve Bayes model with label $Y \in \{0, 1\}$ and features $X_1, X_2 \in \{0, 1\}$.



When writing your answers, feel free to use fractions like $\frac{2}{3}$.

4.1.1 Maximum Likelihood Estimation (2 pt)

Consider the training data below:

data	1	2	3	4	5	6	7	8	9	10
X_1	1	0	1	0	0	0	1	0	1	1
X_2	1	0	1	0	1	1	1	0	1	1
Y	0	1	1	1	0	0	1	1	1	1

What are the maximum likelihood estimates for the following probabilities?

(1) $P(Y = 0) =$

(2) $P(X_2 = 1 | Y = 1) =$

4.1.2 Prediction (1 pt)

Now consider a new data point $X_1 = 1, X_2 = 0$. Use the Naïve Bayes classifier learned in the previous question to determine the posterior probability of Y:

$P(Y = 0 | X_1 = 1, X_2 = 0) =$

4.1.3 Laplace Smoothing (1 pt)

Instead of maximum likelihood estimate, we use Laplace Smoothing with strength $k = 1$ to estimate the model with the training data in question 4.1.1. What is the prior $P(Y)$ now?

$P(Y = 0) =$

4.1.4 Laplace Smoothing with Large Strength (1 pt)

In practice, Laplace Smoothing with a very large strength k often performs poorly. Which of the following statement(s) is/are possible reason(s) for this phenomenon?

□ When k is large, the estimate of $P(Y)$ will be close to a uniform distribution. This is very inaccurate.

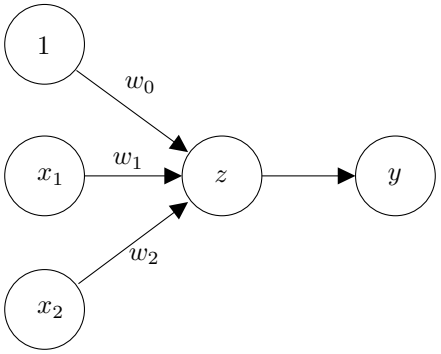
□ When k is large, the estimate of $P(X_i|Y)$ will be close to a uniform distribution, where $i = 1, 2$. This is very inaccurate.

□ When k is large, the estimate of $P(Y)$ will be a sharp distribution with most probabilities close to 0. This is very inaccurate.

□ When k is large, the estimate of $P(X_i|Y)$ will be a sharp distribution with most probabilities close to 0, where $i = 1, 2$. This is very inaccurate.

4.2 Perceptron and Logistic Regression (5 pt)

Consider the following computation graph, for which the inputs are two features $x_1, x_2 \in \mathbb{R}$. The output is a binary label $y \in \{-1, +1\}$.



where $z = w_0 + w_1 x_1 + w_2 x_2$. Suppose we have 2 models: a perceptron and a logistic regression model. The perceptron calculates y as follows:

$$y = \begin{cases} +1 & \text{if } z > 0 \\ -1 & \text{otherwise} \end{cases}$$

The logistic regression model first calculates the probability of $\hat{y} = +1$ as follows:

$$P(\hat{y} = +1|x; w) = \frac{1}{1 + e^{-z}}$$

and then do classification by selecting $y = \arg\max_{\hat{y}} P(\hat{y}|x; w)$.

4.2.1 Classification (2 pt)

What is the output if the weight vector $[w_0\ w_1\ w_2] = [1\ 0\ 0]$?

○ For perceptron, $y = +1$; For logistic regression, $y = +1$.

○ For perceptron, $y = +1$; For logistic regression, $y = -1$.

○ For perceptron, $y = -1$; For logistic regression, $y = +1$.

○ For perceptron, $y = -1$; For logistic regression, $y = -1$.

○ Since we do not know the value of x_1 and x_2, we cannot determine the output.

4.2.2 Update (2 pt)

Suppose we train the perceptron and the logistic regression model respectively on a training data point with input $[x_1\ x_2] = [1\quad -1]$ and label $y = +1$. How will the weight vector $[w_0\ w_1\ w_2]$ be updated if the current weight vector is [1 0 0]? Suppose we are doing gradient ascent (maximizing the log likelihood) for logistic regression with learning rate $\alpha = 1$.

	Perceptron			Logistic Regression		
	w_0	w_1	w_2	w_0	w_1	w_2
Increase	☐	☐	☐	☐	☐	☐
Decrease	☐	☐	☐	☐	☐	☐
No change	☐	☐	☐	☐	☐	☐

4.2.3 Gradient Descent (1 pt)

Suppose we have N samples where N is very large. We use gradient ascent to train the logistic regression model. Which **ONE** of the following is the most likely reason for gradient ascent to be very slow, compared with stochastic gradient ascent?

○ For each step we need to compute the gradient on all the N samples, which is very time-consuming.

○ The learning rate α is too small, so that the weight vector does not change much in each step.

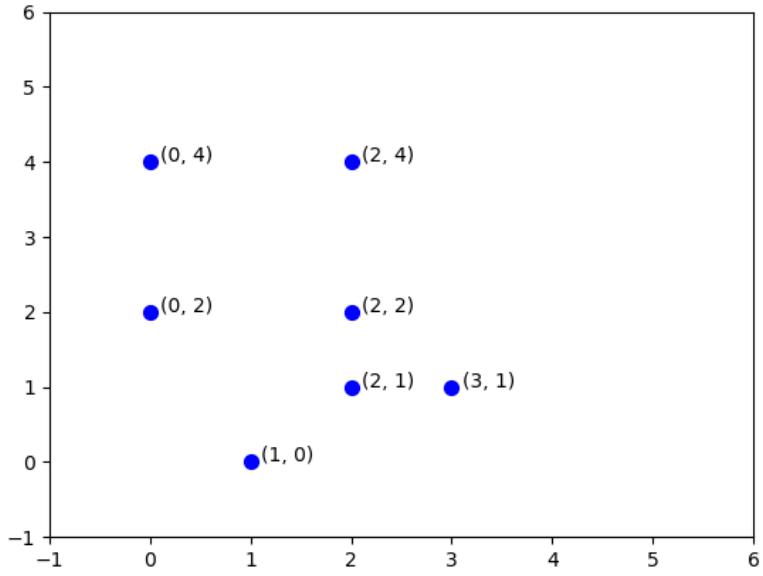
○ The learning rate α is too large, so that the weight vector changes too much in each step and the algorithm cannot converge.

○ The loss function has a very flat region, so that gradient ascent converges very slowly.

5 Unsupervised Machine Learning (10pts)

5.1 K-Means (6pts)

Consider the data points in the figure below.



Each data point $s_i \in \mathbb{R}^2$ is represented by a pair (x_i, y_i). We want to group them into **3** clusters using k-means.

Each cluster center is represented by $c_j^{(t)}$, where $j \in \{0, 1, 2\}$ is the cluster index and $t \in \{0, 1, 2, ...\}$ is the iteration number. We start from three **initial cluster centers** $c_0^{(0)} = (0, 2)$, $c_1^{(0)} = (2, 1)$, $c_2^{(0)} = (3, 5)$.

We use the Euclidean distance for distance calculation between points, i.e.,

$$\textbf{dist}(s_i, s_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

We run k-means for one iteration in the following questions.

5.1.1 Data Point Assignment (2.5pts)

Assign data points to the closest cluster center. Complete the table below by selecting one of the cluster centers for each data point. We have completed selection for $(0, 2)$ and $(2, 1)$ as an example.

	$(0,2)$	$(0,4)$	$(1,0)$	$(2,1)$	$(2,2)$	$(2,4)$	$(3,1)$
$c_0^{(0)} = (0,2)$	■	□	□	□	□	□	□
$c_1^{(0)} = (2,1)$	□	□	□	■	□	□	□
$c_2^{(0)} = (3,5)$	□	□	□	□	□	□	□

5.1.2 Update Cluster Center (1.5pts)

Assign each center to the average of its assigned points. Calculate the new cluster centers and fill in the blanks below.

(1) $c_0^{(1)}$ =

(2) $c_1^{(1)}$ =

(3) $c_2^{(1)}$ =

5.1.3 Convergence (2pts)

Does the k-means algorithm converge with $c_0^{(1)}$, $c_1^{(1)}$, $c_2^{(1)}$ as the final cluster centers? (Hint: run one more iteration and check whether the points' assignments will change or not.)

○ Yes, it converges at $c_0^{(1)}$, $c_1^{(1)}$, $c_2^{(1)}$.

○ No, it doesn't converge at $c_0^{(1)}$, $c_1^{(1)}$, $c_2^{(1)}$.

5.2 Expectation Maximization (4pts)

For this section, we will go through a classic problem of Expectation Maximization (EM): the coin flipping problem.

Each coin has two sides: Head (H) and Tail (T). Here, we observe **three** attempts of coin flipping:

attempt 1	HTHTH
attempt 2	HHHHT
attempt 3	THTTH

Suppose we have **two** different coins with **unknown** probabilities of showing Head or Tail. For each attempt, we randomly choose one coin and flip it for 5 times. The results are listed above in order from left to right. For example, the first attempt shows that we flipped one coin 5 times and got Head, Tail, Head, Tail, Head in order.

We define θ_i as the probability of the i_{th} coin showing Head (H) when flipping, define $y_j = i \in \{1, 2\}$ as the label that attempt j is the result of flipping the i_{th} coin, and define $x_j \in \{\text{H}, \text{T}\}^5$ as the result of attempt j.

Now we use the EM algorithm to estimate θ_1 and θ_2. We start with a guess that $\theta_1^{(0)} = 0.7$ and $\theta_2^{(0)} = 0.4$. ($\theta_i^{(t)}$ means the estimation of θ_i after t iterations.)

5.2.1 E-Step (2pts)

In E-step, we should compute the label distribution of each attempt j, i.e. $P(y_j|x_j)$. Calculate the distribution of the first attempt and select the correct answer below.

Hint: If one attempt x_k only contains one flip and the result is H, then we have:

$$P(y_k = 1|x_k = \text{H}) = \frac{P(x_k = \text{H}|y_k = 1)}{P(x_k = \text{H}|y_k = 1) + P(x_k = \text{H}|y_k = 2)} = \frac{7}{11}$$

The correct value of $P(y_1 = 1|x_1 = (\text{H,T,H,T,H}))$ is:

○ 0.573

○ 0.526

○ 0.451

Please also calculate the distribution for the other two attempts and record the results. We need these results for the next question.

5.2.2 M-Step (2pts)

Now we should update the parameter θ in M-step. For the coin flipping setting, we update the new θ_i by:

$$\frac{\text{Expected total number of Head (H) by coin } i}{\text{Expected total number of flips by coin } i}$$

Calculate $\theta_1^{(1)}$ and select the correct answer below. (Hint: For one attempt x_k, the expected number of H by coin i is $P(y_k = i|x_k) \times \{\text{number of H in } x_k\}$.)

The correct value of $\theta_1^{(1)}$ is:

○ 0.713

○ 0.679

○ 0.665

This is a complete iteration of EM for the coin flipping problem. We can keep doing E-step and M-step in turn until convergence to get a final estimation of θ_1 and θ_2.