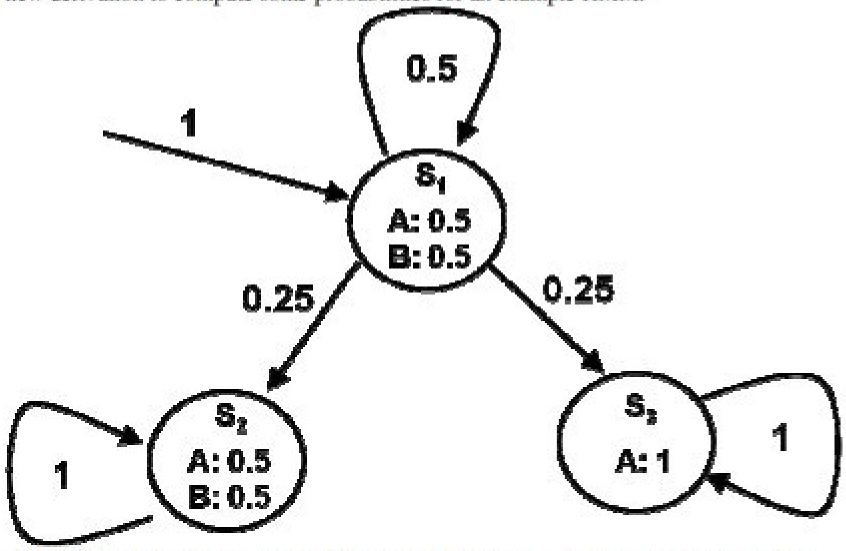
Question 5 – HMM

In class we used the forward term (which we defined as α) to compute the following probability for a set of observed outputs:

$P(O_1 \ldots O_T)$

In this problem we will use a different term for deriving this probability and will use this new derivation to compute some probabilities for an example HMM.



(1). Let $v_i^t = p(O_1 \ldots O_T \mid q_t = s_i)$. Write a formula for $P(O_1 \ldots O_T)$ using *only* v_i^t and $p_t(i)$ which we defined in class (in class we defined $p_t(i) = p(q_t = s_i)$).

Answer:

$$\begin{aligned} p(O_1 \ldots O_T) &= \sum_i p(O_1 \ldots O_T, q_t = s_i) \\ &= \sum_i p(O_1 \ldots O_T \mid q_t = s_i) p(q_t = s_i) \\ &= \sum_i v_i^t p_t(i) \end{aligned}$$

For the next two questions consider the HMM in figure 1. Initial and transition probabilities are listed next to the corresponding edges. Emission probabilities and the states' names are listed inside each node. For example, for state S_2 the emission probabilities are: 0.5 for A and 0.5 for B.

(2). Use only v_i^t and $p_t(i)$ as in (1) to compute $p(O_1 = B, \ldots, O_{200} = B)$ (the probability of observing 200 B's in a row). You need to write an appropriate t for this computation and then explicitly derive the values of v_i^t and $p_t(i)$ for the t that you have chosen and show how you can use these values to compute the probability of this output.
Hint: for computing $p_t(i)$ note that the transitions to and from S_2 and S_3 are symmetric and so for any t, $p_t(S_2) = p_t(S_3)$.

Answer: We will use t=200. For this t we have $p(O_1 = B, \ldots, O_{200} = B \mid S_3) = 0$ and $p(O_1 = B, \ldots, O_{200} = B \mid S_1) = p(O_1 = B, \ldots, O_{200} = B \mid S_2) = (1/2)^{200}$
Also, $p_{200}(1) = (1/2)^{199}$ and based on the hint $p_{200}(2) = (1 - (1/2)^{199})/2$

Putting this together we get: $(1/2)^{399} + (1/2)^{201} (1 - (1/2)^{199})$

(3). Use only v_i^t and $p_t(i)$ as in (1) to compute $p(O_1 = A, \ldots, O_{200} = A)$ (the probability of observing 200 A's in a row). Again, you would need to find an appropriate t for this computation. However, for this part you can use v_1^t for the t that you have chosen in your solution (that is, you do not need to derive the value of v_1^t). Note that this applies only to v_1^t. You would still need to derive the actual values of v_2^t and v_3^t and $p_t(i)$ (for all i) for the t that you have chosen and show how you can use these values to compute the probability of this output

Hint – the t for (3) may be different from the t you selected for (2).

Answer: We will select t=2 for this part. For this t we have
$p(O_1 = A, \ldots, O_{200} = A \mid q_2 = S_2) = (1/2)^{200}$
$p(O_1 = A, \ldots, O_{200} = A \mid q_2 = S_3) = (1/2)$

$p(O_1 = B, \ldots, O_{200} = B \mid q_2 = S_1) = v_1^2$

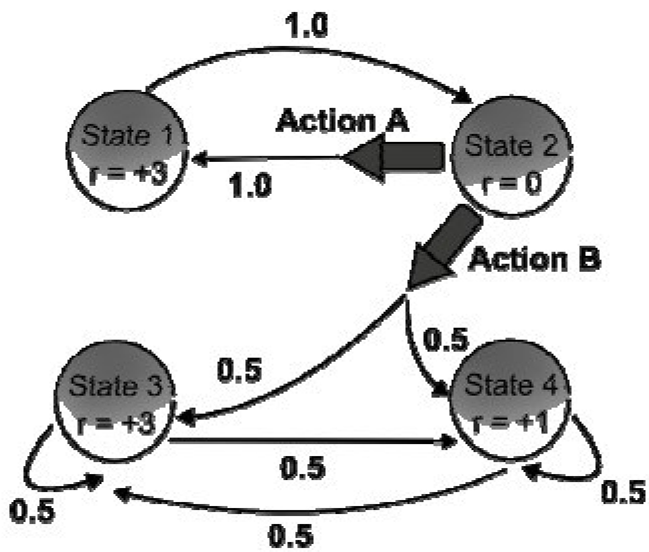
and
$p_2(1) = 0.5$, $p_2(2) = p_2(3) = 0.25$

Putting it together we get:

$0.5\, v_1^2 + (1/2)^{202} + (1/2)^3$

Question 6 – Markov Decision Process

You are given the following Markov decision process, where r denotes the reward at each state :



1. Which action, A or B, maximizes our expected reward on the following turn, for the starting state State 2 ?

Action A

2. Which action from State 2 maximizes the total expected discounted future reward, with a discount factor γ of 0.9? What is the expected discounted future reward for each action?

Action B:

R(A) = sum_i=0^infty $3*\gamma^{2i+1} + 0*\gamma^{2i} = 3\gamma$ sum_k=0^infty $(\gamma^2)^k = 3\,\gamma/(1-\gamma^2) = 14.21$

R(B) = sum_i=1^infty $.5[3\,\gamma^i + 1\,\gamma^i] = 2$ sum_i=1^infty $\gamma^i = 2/(1-\gamma) - 2 = 18$

3. For what value of γ does the expected discounted future reward for each action from State 2 become equal ?

$3\,\gamma/(1-\gamma^2) = 2/(1-\gamma) - 2$ iff $3\,\gamma = 2\,(1+\gamma) - 2\,(1+\gamma)\,(1-\gamma)$ iff $0 = -\gamma + 2\,\gamma^2$ iff $\gamma=1/2$
(since $0 < \gamma < 1$)

Problem 1

1. **(T/F – 2 points)** The sequence of output symbols sampled from a hidden Markov model satisfies the first order Markov property

$\boxed{\text{F}}$

2. **(T/F – 2 points)** Increasing the number of values for the the hidden states in an HMM has much greater effect on the computational cost of forward-backward algorithm than increasing the length of the observation sequence.

$\boxed{\text{T}}$

3. **(T/F – 2 points)** In HMMs, if there are at least two distinct most likely hidden state sequences and the two state sequences cross in the middle (share a single state at an intermediate time point), then there are at least four most likely state sequences.

$\boxed{\text{T}}$

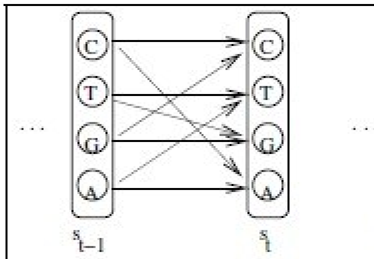
Problem 4

Consider the following pair of observed sequences:

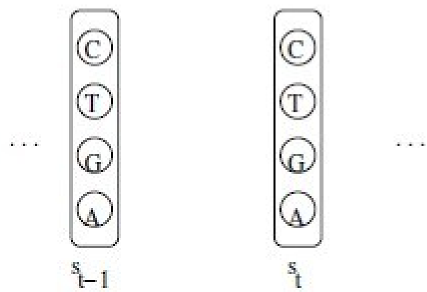
Sequence 1 (s_t):	`A A T T G G C C  A A T T G G C C`	...
Sequence 2 (x_t):	`1 1 2 2 1 1 2 2  1 1 2 2 1 1 2 2`	...
Position t:	`0 1 2 3 4 ...`	

where we assume that the pattern (highlighted with the spaces) will continue forever. Let $s_t \in \{A, G, T, C\}$, $t = 0, 1, 2, \ldots$ denote the variables associated with the first sequence, and $x_t \in \{1, 2\}$, $t = 0, 1, 2, \ldots$ the variables characterizing the second sequence. So, for example, given the sequences above, the observed values for these variables are $s_0 = A$, $s_1 = A$, $s_2 = C, \ldots$, and, similarly, $x_0 = 1, x_1 = 1, x_2 = 2, \ldots$.

1. **(4 points)** If we use a simple first order homogeneous markov model to predict the first sequence (values for s_t only), what is the maximum likelihood solution that we would find? In the *transition diagram* below, please draw the relevant transitions and the associated probabilities (this should not require much calculation)

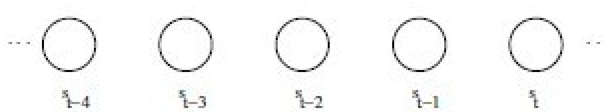


All the transition probabilities are 1/2.

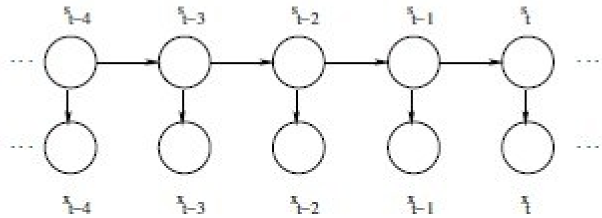


2. **(T/F – 2 points)** The resulting first order Markov model is *ergodic* T

3. **(4 points)** To improve the Markov model a bit we would like to define a graphical model that predicts the value of s_t on the basis of the previous observed values $s_{t-1}, s_{t-2}, \ldots$ (looking as far back as needed). The model parameters/structure are assumed to remain the same if we shift the model one step. In other words, it is the same graphical model that predicts s_t on the basis of $s_{t-1}, s_{t-2}, \ldots$ as the model that predicts s_{t-1} on the basis of $s_{t-2}, s_{t-3}, \ldots$. In the graph below, draw the *minimum number of arrows* that are needed to predict the first observed sequence perfectly (disregarding the first few symbols in the sequence). Since we slide the model along the sequence, you can draw the arrows only for s_t.



4. Now, to incorporate the second observation sequence, we will use a standard hidden Markov model:



where again $s_t \in \{A, G, T, C\}$ and $x_t \in \{1, 2\}$. We will estimate the parameters of this HMM in two different ways.

(I) Treat the pair of observed sequences (s_t, x_t) (given above) as complete observations of the variables in the model and estimate the parameters in the maximum likelihood sense. The initial state distribution $P_0(s_0)$ is set according to the overall frequency of symbols in the first observed sequence (uniform).

(II) Use only the second observed sequence (x_t) in estimating the parameters, again in the maximum likelihood sense. The initial state distribution is again uniform across the four symbols.

We assume that both estimation processes will be successfull relative to their criteria.

a) **(3 points)** What are the observation probabilities $P(x|s)$ ($x \in \{1, 2\}$, $s \in \{A, G, T, C\}$) resulting from the first estimation approach? (should not require much calculation)

> $P(x = 1|s = A) = 1$, $P(x = 1|s = G) = 1$, $P(x = 2|s = T) = 1$, $P(x = 2|s = C) = 1$, all other probabilities are zero.

b) **(3 points)** Which estimation approach is likely to yield a more accurate model over the second observed sequence (x_t)? Briefly explain why.

> The second one (II) since we can use the available four states to exactly capture the variability in the x_t sequence. Using the observation probabilities above, we'd get a model which assigns probability 1/4 to each observed sequence of the above type (the only thing to predict is the starting state).

5. Consider now the two HMMs resulting from using each of the estimation approaches (approaches I and II above). These HMMs are estimated on the basis of the pair of observed sequences given above. We'd like to evaluate the probability that these two models assign to a new (different) observation sequence `1 2 1 2`, i.e., $x_0 = 1, x_1 = 2, x_2 = 1, x_3 = 2$. For the first model, for which we have some idea about what the s_t variables will capture, we also want to know the the associated most likely hidden state sequence. (these should not require much calculation)

   a) **(2 points)** What is the probability that the first model (approach I) assigns to this new sequence of observations?

1/16

   b) **(2 points)** What is the probability that the second model (approach II) gives to the new sequence of observations?

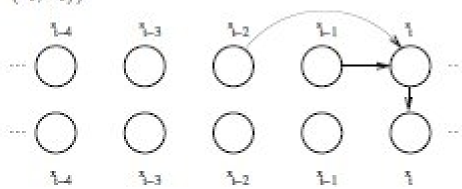
zero (generates only repeated symbol sequences of the type `1 1 2 2 ...`)

   c) **(2 points)** What is the most likely hidden state sequence in the first model (from approach I) associated with the new observed sequence?

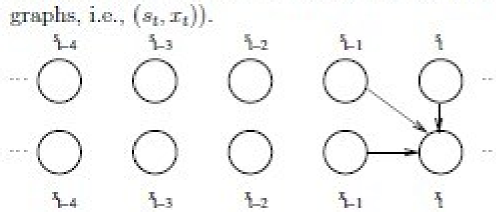
$A\ T\ G\ C$

6. **(4 points)** Finally, let's assume that we observe only the second sequence (x_t) (the same sequence as given above). In building a graphical model over this sequence we are no longer limiting ourselves to HMMs. However, we only consider models whose structure/parameters remain the same as we slide along the sequence. The variables s_t are included as before as they might come handy as hidden variables in predicting the observed sequence.

   a) In the figure below, draw the arrows that any reasonable model selection criterion would find given an unlimited supply of the observed sequence $x_t, x_{t+1}, \ldots$. You only need to draw the arrows for the last pair of variables in the graphs, i.e., (s_t, x_t)).



b) Given only a small number of obervations, the model selection criterion might select a different model. In the figure below, indicate a possible alternate model that any reasonable model selection criterion would find given only a few examples. You only need to draw the arrows for the last pair of variables in the graphs, i.e., (s_t, x_t)).



s	1	2
P(x=1)	0	0.1
P(x=2)	0.199	0
P(x=3)	0.8	0.7
P(x=4)	0.001	0.2

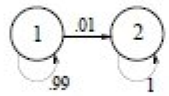


Figure 2: A two-state HMM for Problem 2

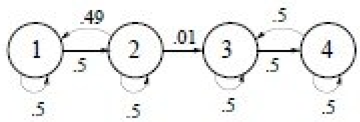


Figure 3: An alternative, four-state HMM for Problem 2

Problem 2

Figure 2 shows a two-state HMM. The transition probabilities of the Markov chain are given in the transition diagram. The output distribution corresponding to each state is defined over $\{1, 2, 3, 4\}$ and is given in the table next to the diagram. The HMM is equally likely to start from either of the two states.

1. **(3 points)** Give an example of an output sequence of length 2 which can not be generated by the HMM in Figure 2.

   1,2

2. **(2 points)** We generated a sequence of $6,867^{2002}$ observations from the HMM, and found that the last observation in the sequence was 3. What is the most likely hidden state corresponding to that last observation?

   2

3. **(2 points)** Consider an output sequence 3 3. What is the most likely sequence of hidden states corresponding to these observations?

   1,1

4. **(2 points)** Now, consider an output sequence 3 3 4. What are *the first two states* of the most likely hidden state sequence?

   2,2

5. **(4 points)** We can try to increase the modeling capacity of the HMM a bit by breaking each state into two states. Following this idea, we created the diagram in Figure 3. Can we set the initial state distribution and the output distributions so that this 4-state model, with the transition probabilities indicated in the diagram, would be equivalent to the original 2-state model? If yes, how? If no, why not?

   No we cannot. First note that we have to associate the first two states in the 4-state model with state 1 of the 2-state model. The probability of leaving the first two states in the 4-state model, however, depends on time (whether the chain happens to be in state 1 or 2). In contrast, in the 2-state model the probability of transitioning to 2 is always 0.01.

6. **(T/F – 2 points)** The Markov chain in Figure 3 is ergodic

   F

Problem 4

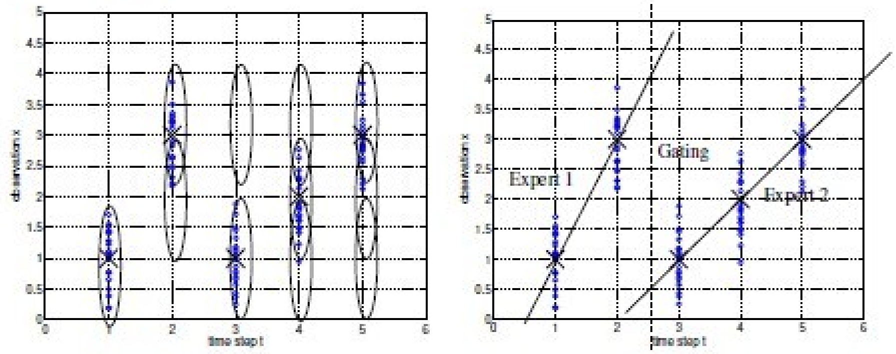
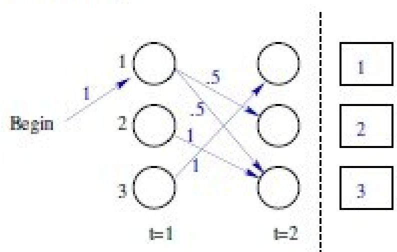


Figure 1a) Figure 1b)

Figure 1: Time dependent observations. The data points in the figure are generated as sets of five consecutive time dependent observations, $x_1, \ldots, x_5$. The clusters come from repeatedly generating five consecutive samples. Each visible cluster consists of 20 points, and has approximately the same variance. The mean of each cluster is shown with a large X.

Consider the data in Figure 1 (see the caption for details). We begin by modeling this data with a three state HMM, where each state has a Gaussian output distribution with some mean and variance (means and variances can be set independently for each state).

4.1. (4 points) Draw the state transition diagram and the initial state distribution for a three state HMM that models the data in Figure 1 in the maximum likelihood sense. Indicate the possible transitions and their probabilities in the figure below (whether or not the state is reachable after the first two steps). In order words, your drawing should characterize the 1st order homogeneous Markov chain govering the evolution of the states. Also indicate the means of the corresponding Gaussian output distributions (please use the boxes).



4.2. (4 points) In Figure 1a draw as ovals the clusters of outputs that would form if we repeatedly generated samples from your HMM over time steps $t = 1, \ldots, 5$. The height of the ovals should reflect the variance of the clusters.

4.3. (4 points) Suppose at time $t = 2$ we observe $x_2 = 1.5$ but don't see the observations for other time points. What is the most likely state at $t = 2$ according to the marginal posterior probability $\gamma_2(s)$ defined as $P(s_2 = s|x_2 = 1.5)$.

$\boxed{2}$

We have only two possible paths for the first three states, 1,2,3, or 1,3,1. The marginal posterior probability comes from averaging the state occupancies across these possible paths, weighted by the corresponding probabilities. Given the observation at $t = 2$ (mean of the output distribution from state 2), the first path is more likely.

4.4. (2 points) What would be the most likely state at $t = 2$ if we also saw $x_3 = 0$ at $t = 3$? In this case $\gamma_2(s) = P(s_2 = s|x_2 = 1.5, x_3 = 0)$.

$\boxed{3}$

The new observation at $t = 3$ is very unlikely to have come from state 3, thus we switch to state sequence 1,3,1.

4.5. (4 points) We can also try to model the data with conditional mixtures (mixtures of experts), where the conditioning is based on the time step. Suppose we only use two experts which are linear regression models with additive Gaussian noise, i.e.,

$$P(x|t, \theta_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{ -\frac{1}{2\sigma_i^2}(x - \theta_{i0} - \theta_{i1}t)^2 \right\}$$

for $i = 1, 2$. The gating network is a logistic regression model from t to binary selection of the experts. Assuming your estimation of the conditional mixture model is successfully in the maximum likelihood sense, draw the resulting mean predictions of the two linear regression models as a function of time t in Figure 1b). Also, with a vertical line, indicate where the gating network would change it's preference from one expert to the other.

4.6. (T/F – 2 points) Claim: by repeatedly sampling from your conditional mixture model at successive time points $t = 1, 2, 3, 4, 5$, the resulting samples would resemble the data in Figure 1

$\boxed{\text{T}}$

.

See the figure.

4.7. (4 points) Having two competing models for the same data, the HMM and the mixture of experts model, we'd like to select the better one. We think that any reasonable model selection criterion would be able to select the better model in this case. Which model would we choose? Provide a brief justification.

The mixture of experts model assigns a higher probability to the available data since the HMM puts some of the probability mass where there are no points. The HMM also has more parameters so any reasonable model selection criterion should select the mixture of experts model.

Problem 5

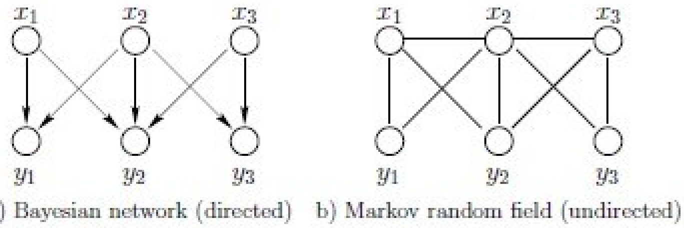


Figure 2: Graphical models

5.1. (2 points) List two different types of independence properties satisfied by the Bayesian network model in Figure 2a.

1) x_1 and x_2 are marginally independent.
2) y_1 and y_2 are conditionally independent given x_1 and x_2.
Lots of other possibilities.

5.2. (2 points) Write the factorization of the joint distribution implied by the directed graph in Figure 2a.

$P(x_1)P(x_2)P(x_3)P(y_1|x_1, x_2)P(y_2|x_1, x_2, x_3)P(y_3|x_2, x_3)$.

5.3. (2 points) Provide an alternative factorization of the joint distribution, different from the previous one. Your factorization should be consistent with all the properties of the directed graph in Figure 2a. Consistency here means: whatever is implied by the graph should hold for the associated distribution.

Any factorization that incorporates all the independencies from the graph and a few more would be possible. For example, $P(x_1)P(x_2)P(x_3)P(y_1)P(y_2)P(y_3)$. In this case all the variables are independent, so any independence statement that we can derive from the graph clearly holds for this distribution as well.

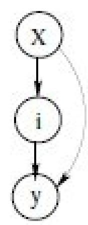
5.4. (4 points) Provide an independence statement that holds for the undirected model in Figure 2b but does NOT hold for the Bayesian network. Which edge(s) should we add to the undirected model so that it would be consistent with (wouldn't imply anything that is not true for) the Bayesian network?

x_1 is independent of x_3 given x_2 and y_2. In the bayesian network any knowledge of y_2 would make x_1 and x_3 dependent. Adding an edge between x_1 and x_3 would suffice (cf. moralization).

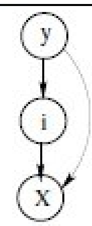
5.5. (2 points) Is your resulting undirected graph triangulated (Y/N)?

Y

5.6. (4 points) Provide two directed graphs representing 1) a mixture of two experts model for classification, and 2) a mixture of Gaussians classifiers with two mixture components per class. Please use the following notation: $\mathbf{x}$ for the input observation, y for the class, and i for any selection of components.



A mixture of experts classifier, where $i = 1, 2$ selects the expert.



A mixture of Gaussians model, where $i = 1, 2$ selects the Gaussian component within each class.

Problem 5

Assume that the following sequences are very long and the pattern highlighted with spaces is repeated:

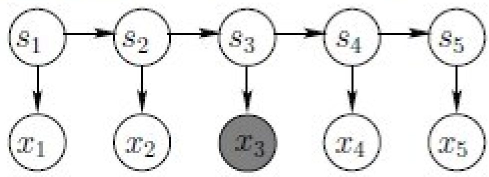
Sequence 1: `1 0 0   1 0 0   1 0 0   1 0 0 ... 1 0 0`

Sequence 2: `1   1 0 0   1 0 0   1 0 0 ... 1 0 0`

1. **(4 points)** If we model each sequence with a different first-order HMM, what is the number of hidden states that a reasonable model selection method would report?

	HMM for Sequence 1	HMM for Sequence 2
No. of hidden states	3	4

2. **(2 points)** The following Bayesian network depicts a sequence of 5 observations from an HMM, where s_1, s_2, s_3, s_4, s_5 is the hidden state sequence.



Are x_1 and x_5 independent given x_3? Briefly justify your answer.

They are not independent. The moralized ancestral graph corresponding to x_1, x_3, and x_5 is the same graph with arrows replaced with undirected edges. x_1 and x_5 are not separated given x_3, and thus not independent.

3. **(3 points)** Does the order of Markov dependencies in the observed sequence always determine the number of hidden states of the HMM that generated the sequence? Provide a brief justification.

No. The answer to the previous question implies that observations corresponding to (typical) HMMs have no Markov properties (of any order). This holds, for example, when there are only two possible hidden states. Thus Markov properties of the observation sequence cannot in general determine the number of hidden states.

(i) Give one similarity and one difference between HMM and MDP. produces a smaller, new set

similarity: Marov assumptions

difference: The Markov chain in HMM is hidden; in MDP, the states are fully observed

(j) For each of the following datasets, is it appropriate to use HMM? Provide a brief reasoning for your answer.

- Gene sequence dataset.
- A database of movie reviews (eg., the IMDB database).
- Stock market price dataset.
- Daily precipitation data from the Northwest of the US.

Time-series data ; Markov assumption may be reasonable

6 Hidden Markov Models (12pts)

Consider an HMM with states $Y_t \in \{S_1, S_2, S_3\}$, observations $X_t \in \{A, B, C\}$, and parameters

$\pi_1 = 1$	$a_{11} = 1/2$	$a_{12} = 1/4$	$a_{13} = 1/4$	$b_1(A) = 1/2$	$b_1(B) = 1/2$	$b_1(C) = 0$
$\pi_2 = 0$	$a_{21} = 0$	$a_{22} = 1/2$	$a_{23} = 1/2$	$b_2(A) = 1/2$	$b_2(B) = 0$	$b_2(C) = 1/2$
$\pi_3 = 0$	$a_{31} = 0$	$a_{32} = 0$	$a_{33} = 1$	$b_3(A) = 0$	$b_3(B) = 1/2$	$b_3(C) = 1/2$

(a) (3pts) What is $P(Y_5 = S_3)$?

$$1 - P(Y_5 = S_1) - P(Y_5 = S_2)$$
$$= 1 - \frac{1}{16} - 4 \times \frac{1}{32}$$
$$= \frac{13}{16}$$

For 6(b)-(d), suppose we observe $AABCABC$, starting at time point 1.

(b) (2pts) What is $P(Y_5 = S_3 | X_{1:7} = AABCABC)$?

0

(c) (4pts) Fill in the following table assuming the observation $AABCABC$. The α's are values obtained during the forward algorithm: $\alpha_t(i) = P(X_1, \ldots, X_t, Y_t = i)$.

t	$\alpha_t(1)$	$\alpha_t(2)$	$\alpha_t(3)$
1	$\frac{1}{2}$	0	0
2	$\frac{1}{8}$	$\frac{1}{16}$	0
3	$\frac{1}{32}$	0	$\frac{1}{32}$
4	0	$\frac{1}{2^8}$	$\frac{5}{2^8}$
5	0	$\frac{1}{2^{10}}$	0
6	0	0	$\frac{1}{2^{12}}$
7	0	0	$\frac{1}{2^{13}}$

(d) (3pts) Write down the sequence of $Y_{1:7}$ with the maximal posterior probability assuming the observation $AABCABC$. What is that posterior probability?

$S_1 S_1 S_1 S_2 S_2 S_3 S_3$

posterior probability : 1

5. (**True or False**, 2 pts) The Markov Blanket of a node x in a graph with vertex set X is the smallest set Z such that $x \perp X/\{Z \cup x\} | Z$

Solutions: T

8 Hidden Markov Models with continuous emissions (10 points)

In this question, we will study hidden markov models with continuous emissions. We will use the notation used in class, with x^i denoting the output at time i, and y_i denoting the corresponding hidden state. The HMM has K states $\{1 \dots K\}$. The output for state k is obtained by sampling a Gaussian distribution parameterized by mean μ_k and standard deviation σ_k. Thus, we can write the emission probabilitye as $p(x_i|y_i = k, \theta) = \mathcal{N}(x_i|\mu_k, \sigma_k)$. θ is the set of parameters of the HMM, which includes the initial probabilities π, transition probability matrix A and the means and standard deviations $\{\mu_1, \dots, \mu_K, \sigma_1, \dots, \sigma_k\}$.

8.1 Log-likelihood (1 point)

Write down the log-likelihood for a sequence of observations of the emissions $\{x_1, \dots, x_n\}$ when the states (also observed) are $\{y_1, \dots, y_n\}$.

Solution:

$$\begin{aligned} \log p(x_1, \dots, x_n|y_1, \dots, y_n) &= \log \prod_i p(x_i|y_i) && (1) \\ &= \sum_i \log(\mathcal{N}(x_i|\mu_{y_i}, \sigma_{y_i})) && (2) \end{aligned}$$

8.2 Forward and backward updates (2 points)

Write the forward and backward update equations for this HMM. Explain in a single line how they are different from the updates we studied in class.

Solution:

$$\begin{aligned} \alpha_i^k &= \mathcal{N}(x_i|\mu_k, \sigma_k) \sum_l \alpha_{i-1}^l a_{l,k} && (3) \\ \beta_i^k &= \sum_l a_{k,l} \beta_{i+1}^l \mathcal{N}(x_i|\mu_l, \sigma_l) && (4) \end{aligned}$$

The equations are similar in form. But in this case, the output probabilities are gaussian rather than multinomial. The outputs are also continuous rather than discrete.

8.3 Supervised parameter learning

We are given a sequence of observations $X = \{x_1, \dots, x_n\}$ and the corresponding hidden states $Y = \{y_1, \dots, y_n\}$. We want to find the parameters θ for the HMM.

1. Are the update equations for A_{ij} and π_i different from the ones obtained for the HMM we studied in class? Explain why or why not (2 points).

   Solution: The update equations for A_{ij} and π_i are the same. They involve only the state transition counts and so are independent of the form chosen for emission probabilities.

2. What are the update equations for the Gaussian parameters μ_k and σ_k ? (Hint: You do not need to derive them. Given the hidden states, the outputs are all independent of each other, and each is sampled from one out of K gaussians.) (2 points)

   Solution:

$$\mu_k = \frac{\sum_t \mathcal{I}[y_t = k] x_t}{\sum_t \mathcal{I}[y_t = k]} \tag{5}$$

$$\sigma_k^2 = \frac{\sum_t \mathcal{I}[y_t = k](x_t - \mu_k)^2}{\sum_t \mathcal{I}[y_t = k]} \tag{6}$$

8.4 Unsupervised parameter learning

Now, we are only given a sequence of observations $X = \{x_1, \ldots, x_n\}$. We want to find the parameters θ for the HMM. (Slide 47 and 48 for the HMM lecture describe the unsupervised learning algorithm for the HMM discussed in class)

8.5 Objective function

The unsupervised learning algorithm optimizes the expected complete log-likelihood. Why is that a reasonable choice for the objective function? (1 point)

Solution: The expected complete log-likelihood is a lower bound to the complete log-likelihood. It is guaranteed to converge to a local optimum of the complete likelihood. Hence it is a reasonable choice for the objective function.

8.5.1 Expected complete LL

What is the expected complete log-likelihood ($\langle l_c(\theta; x, y)\rangle$) for the HMM with continuous gaussian emissions? Just write the expression, a derivation is not necessary.(1 point)

Solution:

$$\langle l_c(\theta; x, y)\rangle = \sum_n \left(\langle y_{n,1}^i\rangle \log \pi_i\right) + \sum_n \sum_{t=2}^{T} \left(\left\langle y_{n,t-1}^i y_{n,t}^j \right\rangle \log a_{i,j}\right) + \sum_n \sum_{t=1}^{T} \left(\langle y_{n,t}^i\rangle \log \mathcal{N}(x_n, t|\mu_i, \sigma_i)\right) \tag{7}$$

8.5.2 Gaussian Parameter estimation

Suppose you want to find ML estimates $\hat{\mu}_k$ and $\hat{\sigma}_k$ for parameters μ_k and σ_k . Will the ML expressions have the same form as those obtained for the means and variances in a mixture of gaussians? Explain in one line. (Hint: Write down the terms in $\langle l_c(\theta; x, y)\rangle$ that are relevant to the optimization (i.e, contain μ_k and σ_k))(1 point)

Solution: Yes, the ML expressions will have same form. The relevant term in $\langle l_c(\theta; x, y)\rangle$ is only the last term, which closely resembles the expected complete log-likelihood for a gaussian mixture model. In this case the $p(y = 1|x)$ term is computed using the forward backward algorithm rather than by simply using Bayes rule (as is done for a mixture of gaussians).

2 Bayes Nets and HMMs

(a) Let nbs(m) = the number of possible Bayes Network graph structures using m attributes. (Note that two networks with the same structure but different probabilities in their tables do not count as different structures). Which of the following statements is true?

- (i) $\text{nbs}(m) < m$
- (ii) $m \leq \text{nbs}(m) < \frac{m(m-1)}{2}$
- (iii) $\frac{m(m-1)}{2} \leq \text{nbs}(m) < 2^m$
- (iv) $2^m \leq \text{nbs}(m) < 2^{\frac{m(m-1)}{2}}$
- (v) $2^{\frac{m(m-1)}{2}} \leq \text{nbs}(m)$

```
Answer is (v) because the number of undirected graphs with n vertices
is 2^[m choose 2], and there are even more acyclic directed graphs
```

(b) Remember that $I < X, Y, Z >$ means

X is conditionally independent of Z given Y

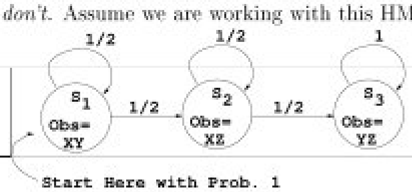
Assuming the conventional assumptions and notation of Hidden Markov Models, in which q_t denotes the hidden state at time t and O_t denotes the observation at time t, which of the following are true of all HMMs? Write "True" or "False" next to each statement.

- (i) $I < q_{t+1}, q_t, q_{t-1} >$
- (ii) $I < q_{t+2}, q_t, q_{t-1} >$
- (iii) $I < q_{t+1}, q_t, q_{t-2} >$
- (iv) $I < O_{t+1}, O_t, O_{t-1} >$
- (v) $I < O_{t+2}, O_t, O_{t-1} >$
- (vi) $I < O_{t+1}, O_t, O_{t-2} >$

```
(i) (ii) (iii) all TRUE
(iv) (v) (vi) all FALSE
```

7 Hidden Markov Models

Warning: this is a question that will take a few minutes if you really understand HMMs, but could take hours if you don't. Assume we are working with this HMM



$a_{11} = 1/2$	$a_{12} = 1/2$	$a_{13} = 0$	$b_1(X) = 1/2$	$b_1(Y) = 1/2$	$b_1(Z) = 0$	$\pi_1 = 1$
$a_{21} = 0$	$a_{22} = 1/2$	$a_{23} = 1/2$	$b_2(X) = 1/2$	$b_2(Y) = 0$	$b_2(Z) = 1/2$	$\pi_2 = 0$
$a_{31} = 0$	$a_{32} = 0$	$a_{33} = 1$	$b_3(X) = 0$	$b_3(Y) = 1/2$	$b_3(Z) = 1/2$	$\pi_3 = 0$

Where

$$\begin{aligned} a_{ij} &= P(q_{t+1} = S_j | q_t = S_i) \\ b_i(k) &= P(O_t = k | q_t = S_i) \end{aligned}$$

Suppose we have observed this sequence

XZXYYZYZZ

(in long-hand: $O_1 = X, O_2 = Z, O_3 = X, O_4 = Y, O_5 = Y, O_6 = Z, O_7 = Y, O_8 = Z, O_9 = Z$). Fill in this table with $\alpha_t(i)$ values, remembering the definition:

$$\alpha_i(t) = P(O_1 \wedge O_2 \wedge ... O_t \wedge q_t = s_i)$$

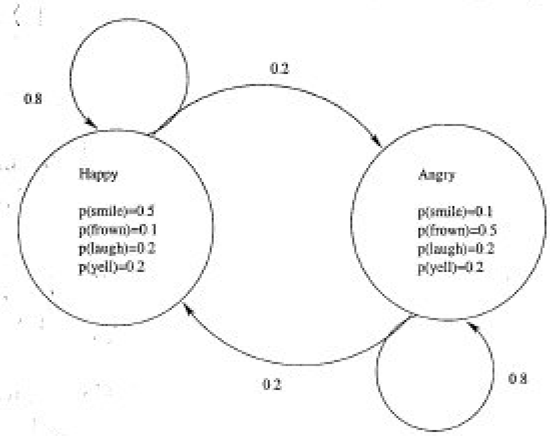
So for example,

$$\alpha_3(2) = P(O_1 = X \wedge O_2 = Z \wedge O_3 = X \wedge q_3 = S_2)$$

t	$\alpha_t(1)$	$\alpha_t(2)$	$\alpha_t(3)$
1	1/2		
2		1/8	
3		1/32	
4			1/128
5			1/256
6			1/512
7			1/1024
8			1/2048
9			1/4096

3 HMMs

Andrew lives a simple life. Some days he's Angry and some days he's Happy. But he hides his emotional state, and so all you can observe is whether he smiles, frowns, laughs, or yells. We start on day 1 in the Happy state, and there's one transition per day.



Definitions:
q_t = state on day t.
O_t = observation on day t.

(a) What is $P(q_2 = Happy)$? 0.8

(b) What is $P(O_2 = frown)$? $\frac{8}{10} \times \frac{1}{10} + \frac{2}{10} \times \frac{1}{2} = \frac{8}{100} + \frac{10}{100} = \frac{18}{100}$

(c) What is $P(q_2 = Happy | O_2 = frown)$? $\frac{P(O_2 = frown \mid q_2 = H)\, P(q_2 = H)}{P(O_2 = frown)}$

$= \frac{\frac{1}{10} \times \frac{8}{10}}{\left(\frac{18}{100}\right)} = 4/9$

(d) What is $P(O_{100} = yell)$?

$P(O_{100} = yell) = P(O_{100} = yell \mid q_{100} = H) P(q_{100} = H) + P(O_{100} = yell \mid q_{100} = A) P(q_{100} = A)$
$= 2/10 \times (P(q_{100} = H) + P(q_{100} = A)) = 2/10 \times 1 = 2/10$

(e) Assume that $O_1 = frown$, $O_2 = frown$, $O_3 = frown$, $O_4 = frown$, and $O_5 = frown$. What is the most likely sequence of states?

HAAAA

9 Markov Decision Processes

Consider the following MDP, assuming a discount factor of $\gamma = 0.5$. Note that the action "Party" carries an immediate reward of +10. The action "Study" unfortunately carries no immediate reward, except during the senior year, when a reward of +100 is provided upon transition to the terminal state "Employed".

(a) What is the probability that a freshman will fail to graduate to the "Employed" state within four years, even if they study at every opportunity?

(b) Draw the diagram for the Markov Process (not the MDP, the MP) that corresponds to the policy "study whenever possible."

(c) What is the value associated with the state "Junior" under the "study whenever possible" policy?

(d) Exactly how rewarding would parties have to be during junior year in order to make it advisable for a junior to party rather than study (assuming, of course, that they wish to optimize their cumulative discounted reward)?

(e) Answer the following true or false. If true, give a one-sentence argument. If false, give a counterexample.

- **(True or False?)** If partying during junior year an optimal action when it is assigned reward r, then it will also be an optimal action for a freshman when assigned reward r.

- **(True or False?)** If partying during junior year is an optimal action when it is assigned reward r, then it will also be an optimal action for a freshman when assigned reward r.

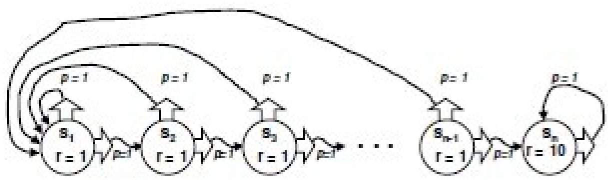
- **(True or False?)** HMM's are a special case of MDP's.

2 Markov Decision Processes (13 points)

For this question it might be helpful to recall the following geometric identities, which assume $0 \leq \alpha < 1$.

$$\sum_{i=0}^{k} \alpha^i = \frac{1-\alpha^{k+1}}{1-\alpha} \qquad \sum_{i=0}^{\infty} \alpha^i = \frac{1}{1-\alpha}$$

The following figure shows an MDP with N states. All states have two actions (North and Right) except S_n, which can only self-loop. Unlike most MDPs, all state transitions are deterministic. Assume discount factor γ.



For questions (a)–(e), express your answer as a finite expression (no summation signs or . . . 's) in terms of n and/or γ.

(a) What is $J^*(S_n)$?

$$J^*(S_n) = 10 + \gamma \cdot J^*(S_n) \Longrightarrow J^*(S_n) = \frac{10}{1-\gamma}$$

(b) There is a unique optimal policy. What is it?

$A_i =$ Right $(i = 1, \ldots, n)$

(c) What is $J^*(S_1)$?

$$J^*(S_1) = 1 + \gamma + \cdots + \gamma^{n-2} + J^*(S_n) \cdot \gamma^{n-1} = \frac{1 + 9\gamma^{n-1}}{1-\gamma}$$

(d) Suppose you try to solve this MDP using value iteration. What is $J^1(S_1)$?

$J^1(S_1) = 1$

(e) Suppose you try to solve this MDP using value iteration. What is $J^2(S_1)$?

$J^2(S_1) = 1 + \gamma$

(f) Suppose your computer has exact arithmetic (no rounding errors). How many iterations of value iteration will be needed before all states record their exact (correct to infinite decimal places) J^* value? Pick one:

(i) Less than $2n$

(ii) Between $2n$ and n^2

(iii) Between $n^2 + 1$ and 2^n

(iv) It will never happen

It's a limiting process.

(g) Suppose you run policy iteration. During one step of policy iteration you compute the value of the current policy by computing the exact solution to the appropriate system of n equations in n unknowns. Suppose too that when choosing the action during the policy improvement step, ties are broken by choosing North.

Suppose policy iteration begins with all states choosing North.

How many steps of policy iteration will be needed before all states record their exact (correct to infinite decimal places) J^* value? Pick one:

(i) Less than $2n$

(ii) Between $2n$ and n^2

(iii) Between $n^2 + 1$ and 2^n

(iv) It will never happen

After i policy iterations, we have

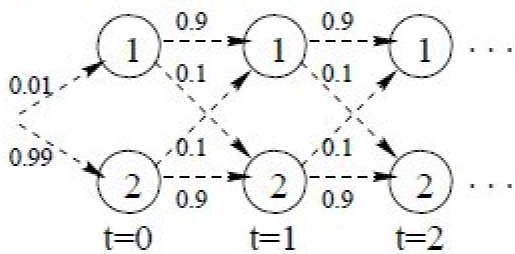
$$Action(S_j) = \begin{cases} Right & \text{if } n - i < j < n \\ North & \text{otherwise.} \end{cases}$$

10 Hidden Markov Models (8 points)

Consider a hidden Markov model illustrated as the figure shown below, which shows the hidden state transitions and the associated probabilities along with the initial state distribution. We assume that the state dependent outputs (coin flips) are governed by the following distributions

$P(x = heads|s = 1) = 0.51$
$P(x = heads|s = 2) = 0.49$
$P(x = tails|s = 1) = 0.49$
$P(x = tails|s = 2) = 0.51$

In other words, our coin is slightly biased towards *heads* in state 1 whereas in state 2 *tails* is a somewhat more probable outcome.



(a) Now, suppose we observe three coin flips all resulting in *heads*. The sequence of observations is therefore *heads*; *heads*; *heads*. What is the most likely state sequence given these three observations? (It is not necessary to use the Viterbi algorithm to deduce this, nor any subsequent questions).

2,2,2

The probabilities of outputting head are nearly identical in two states and it is very likely that the system starts from state 2 and stay there. It loses a factor of 9 in probability if it ever switchs to state 1.

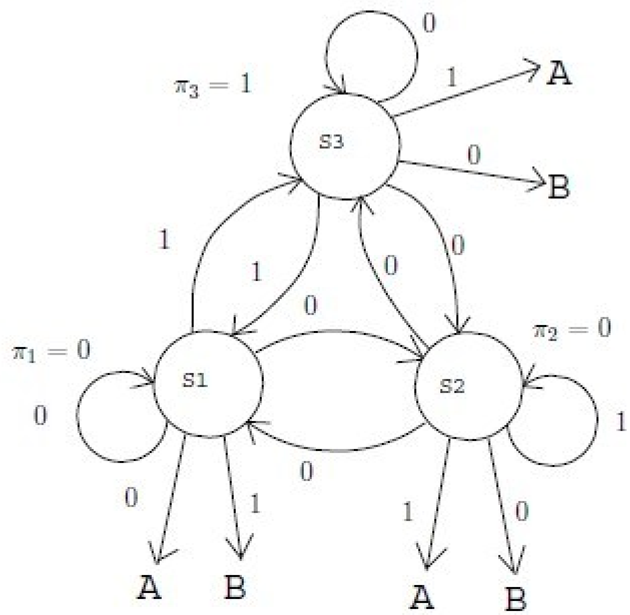
(b) What happens to the most likely state sequence if we observe a long sequence of all heads (e.g., 10^6 heads in a row)?

2,1,1,1,...

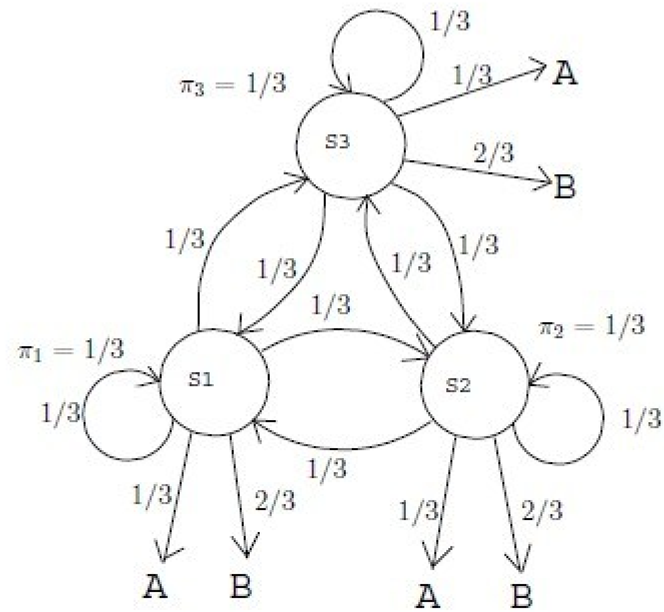
When the number of continuous observations of heads increases, the pressure for the system to switch to state 1 also increases, as state 1 has a slight advantage per observation. Eventually the switch will take place and then there's no benefit from ever switching back to state 2. The cost of the transition switching from state 2 to state 1 is the same regardless of when it takes place. But switching earlier is better than later, since the likelihood of observing the long sequence of all heads is greater. However, it is somewhat better to go via state 2 initially and switch right after (0.99*0.49*0.1 ...) rather than start from state 1 to begin with (0.01*0.51*0.9 ...).

(c) Consider the following 3-state HMM, π_1, π_2 and π_3 are the probabilities of starting from each state $S1$, $S2$ and $S3$. Give a set of values so that the resulting HMM maximizes the likelihood of the output sequence ABA.

There are many possible solutions, and they are all correct as long as they output ABA with probability 1, and the parameter settings of the models are sound. Here is one possible solution:



(d) We're going to use EM to learn the parameters for the following HMM. Before the first iteration of EM we have initialized the parameters as shown in the following figure. (**True** or **False**) For these initial values, EM will successfully converge to the model that maximizes the likelihood of the training sequence ABA.



Note the symmetry of the initial set of values over S_1, S_2 and S_3. After each EM iteration, the transition matrix will keep the same ($a_{ij} = 1/3$). The observation matrix may change, but the symmetry still holds ($b_i(A) = b_j(A)$).

(e) (**True** or **False**) In general when are trying to learn an HMM with a small number of states from a large number of observations, we can almost always increase the training data likelihood by permitting more hidden states.

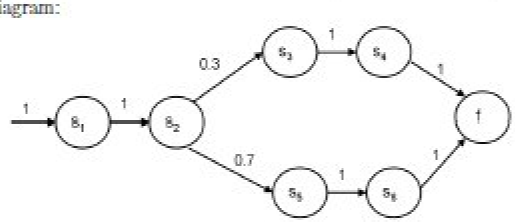
To model any finite length sequence, we can increase the number of hidden states in an HMM to be the number of observations in the sequence and therefore (with appropriate parameter choices) generate the observed sequence with probability 1. Given a fixed number of finite sequences (say n), we would still be able to assign probability $1/n$ for generating each sequence. This is not useful, of course, but highlights the fact that the complexity of HMMs is not limited.

Problem 5. HMM (12 points)

Consider the HMM defined by the transition and emission probabilities in the table below. This HMM has six states (plus a start and end states) and an alphabet with four symbols (A,C, G and T). Thus, the probability of transitioning from state S_1 to state S_2 is 1, and the probability of emitting A while in state S_1 is 0.3.

	0	S_1	S_2	S_3	S_4	S_5	S_6	f	A	C	G	T
0	0	1	0	0	0	0	0	0				
S_1	0	0	1	0	0	0	0	0	0.5	0.3	0	0.2
S_2	0	0	0	0.3	0	0.7	0	0	0.1	0.1	0.2	0.6
S_3	0	0	0	0	1	0	0	0	0.2	0	0.1	0.7
S_4	0	0	0	0	0	0	0	1	0.1	0.3	0.4	0.2
S_5	0	0	0	0	0	0	1	0	0.1	0.3	0.3	0.3
S_6	0	0	0	0	0	0	0	1	0.2	0.3	0	0.5

Here is the state diagram:



For each of the pairs belows, place $<$, $>$ or $=$ between the right and left components of each pair. (*2 pts each*):

(a) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_1 = S_1, q_2 = S_2)$
$P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_1 = S_1, q_2 = S_2)$

Below we will use a shortened notation. Specifically we will us $P(A, C, T, A, S_1, S_2)$ instead of $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_1 = S_1, q_2 = S_2)$, $P(A, C, T, A)$ instead of $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A)$ and so forth.

Answer: $=$
$P(A, C, T, A, S_1, S_2) = P(A, C, T, A|S_1, S_2)P(S_1, S_2) = P(A, C, T, A|S_1, S_2)$, since $P(S_1, S_2) = 1$

(b) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4)$
$P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_3 = S_3, q_4 = S_4)$

Answer: $<$
As in (b), $P(A, C, T, A, S_3, S_4) = P(A, C, T, A|S_3, S_4)P(S_3, S_4)$ however, since $P(S_3, S_4) = 0.3$, then the right hand side is bigger.

(c) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4)$
$P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_5, q_4 = S_6)$

Answer: $<$
The first two emissions (A and C) do not matter since they are the same. Thus, the right hand side translates to $P(O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4) = P(O_3 = T, O_4 = A|q_3 = S_3, q_4 = S_4)P(S_3, S_4) = 0.7 * 0.1 * 0.3 = 0.021$ while the right hand side is $0.3 * 0.2 * 0.7 = 0.042$.

(d) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A)$ $\qquad$ $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4)$

Answer: $>$

Here the left hand side is: $P(A, C, T, A, S_3, S_4) + P(A, C, T, A, S_5, S_6)$. The right side of the summation is the right hand side above. Since the left side of the summation is greater than 0, the left hand side is greater.

(e) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A)$ $\qquad$ $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_3 = S_3, q_4 = S_4)$

Answer: $<$

As mentioned for (e) the left hand side is: $P(A, C, T, A, S_3, S_4) + P(A, C, T, A, S_5, S_6) = P(A, C, T, A | S_3, S_4) P(S_3, S_4) + P(A, C, T, A | S_5, S_6) P(S_5, S_6)$. Since $P(A, C, T, A | S_3, S_4) > P(A, C, T, A | S_5, S_6)$ the left hand side is lower from the right hand side.

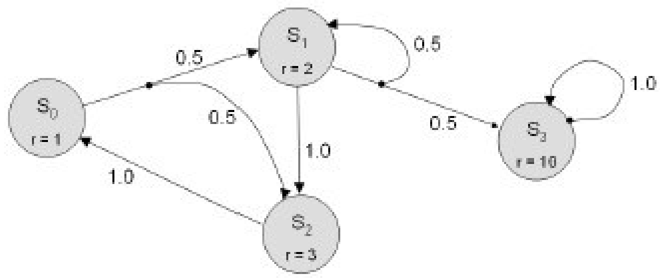
(f) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A)$ $\qquad$ $P(O_1 = A, O_2 = T, O_3 = T, O_4 = G)$

Answer: $<$

Since the first and third letters are the same, we only need to worry about the second and fourth. The left hand side is: $0.1*(0.3*0.1 + 0.7*0.2) = 0.017$ while the right hand side is: $0.6*(0.7*0 + 0.3*0.4) = 0.072$.

Problem 9. Markov Decision Processes (11pts)

(a) *(8 points)* Consider the MDP given in the figure below. Assume the discount factor $\gamma = 0.9$. The r-values are rewards, while the numbers next to arrows are probabilities of outcomes. Note that only state S_1 has two actions. The other states have only one action for each state.



(a.1) *(4 points)* Write down the numerical value of $J(S_1)$ after the first and the second iterations of Value Iteration.

Initial value function: $J^0(S_0) = 0$; $J^0(S_1) = 0$; $J^0(S_2) = 0$; $J^0(S_3) = 0$;

$J^1(S_1) =$

$J^2(S_1) =$

Answer:

$J^1(S_1) = 2$

$$\begin{aligned} J^2(S_1) &= \max(2 + 0.9(0.5 * J^1(S_1) + 0.5 * J^1(S_3)), 2 + 0.9 * J^1(S_2)) \\ &= \max(2 + 0.9(0.5 * 2 + 0.5 * 10), 2 + 0.9 * 3) \\ &= 7.4 \end{aligned}$$

(a.2) *(4 points)* Write down the optimal value of state S_1. There are few ways to solve it, and for one of them you may find useful the following equality: $\sum_{i=0}^{\infty} \alpha^i = \frac{1}{1-\alpha}$ for any $0 \leq \alpha < 1$.

$J^*(S_1) =$

Answer:

It is pretty clear from the given MDP that the optimal policy from S_1 will involve trying to move from S_1 to S_3 as this is the only state that has a large reward. First, we compute optimal value for S_3:

$J^*(S_3) = 10 + 0.9 * J^*(S_3)$

$J^*(S_3) = \frac{10}{0.1} = 100$

We can now compute optimal value for S_1:

$J^*(S_1) = 2 + 0.9(0.5 * J^*(S_1) + 0.5 * J^*(S_3)) = 2 + 0.9(0.5 * J^*(S_1) + 50)$;

Solving for $J^*(S_1)$ we get:

$J^*(S_1) = \frac{47}{0.55} = 87.\overline{45}$

(b) *(3 points)* A general MDP with N states is guaranteed to converge in the limit for Value Iteration as long as $\gamma < 1$. In practice one cannot perform infinitely many value iterations to guarantee convergence. Circle **all** the statements below that are **true**.

- (1) Any MDP with N states converges after N value iterations for $\gamma = 0.5$;

  Answer: False

- (2) Any MDP converges after the 1st value iteration for $\gamma = 1$;
  Answer: False

- (3) Any MDP converges after the 1st value iteration for a discount factor $\gamma = 0$;

  Answer: True, since all the converged values will be just immediate rewards.

- (4) An acyclic MDP with N states converges after N value iterations for any $0 \le \gamma \le 1$.

  Answer: True, since there are no cycles and therefore after each iteration at least one state whose value was not optimal before is guaranteed to have its value set to an optimal value (even when $\gamma = 1$), unless all state values are already converged.

- (5) An MDP with N states and no stochastic actions (that is, each action has only one outcome) converges after N value iterations for any $0 \le \gamma < 1$.

  Answer: False. Consider a situation where there are no absorbing goal states.

- (6) One usually stops value iterations after iteration $k+1$ if: $\max_{0 \le i \le N-1} |J^{k+1}(S_i) - J^k(S_i)| < \xi$, for some small constant $\xi > 0$.

  Answer: True.

6 Hidden Markov Models [11 points]

6.1 [3 points]

Assume we have temporal data from two classes (for example, 10 days closing prices for stocks that increased / decreased on the following day). How can we use HMMs to classify this data?

Solution: We would need to learn two HMMs, one for each class. For a new test data vector, we will run the Viterby algorithm for this vector in both HMMs and chose the HMM with higher likelihood as the class for this vector.

6.2 [3 points]

Derive the probablity of:

$$P(o_1, \ldots, o_t, q_{t-1} = s_v, q_t = s_j) \tag{15}$$

You may use any of the model parameters (starting, emission and transition probabilities) and the following constructs (defined and derived in class) as part of your derivation:

$$p_t(i) = p(q_t = s_i) \tag{16}$$
$$\alpha_t(i) = p(o_1, \ldots, o_t, q_t = s_i) \tag{17}$$
$$\delta_t(i) = \max_{q_1,\ldots,q_{t-1}} p(q_1, \ldots, q_{t-1}, q_t = s_i, O_1, \ldots, O_t) \tag{18}$$

Note that you may not need to use all of these constructs to fully define the function listed above.

Solution:

$$P(o_1, \ldots, o_t, q_{t-1} = s_v, q_t = s_j) \tag{19}$$
$$=P(o_t, q_t = s_j \mid o_1, \ldots, o_{t-1}, q_{t-1} = s_v)P(o_1, \ldots, o_{t-1}, q_{t-1} = s_v) \tag{20}$$
$$=P(o_t, q_t = s_j \mid; q_{t-1} = s_v)\alpha_{t-1}(v) \tag{21}$$
$$=P(o_t \mid q_t = s_j, q_{t-1} = s_v)P(q_t = s_j \mid q_{t-1} = s_v)\alpha_{t-1}(v) \tag{22}$$
$$=b_j(o_t)a_{v,j}\alpha_{t-1}(v) \tag{23}$$

6.3 [5 points]

The following questions refer to figure 5. In that figure we present a HMM and specify both the transition and emission probabilities. Let p_A^t be the probability of being in state A at time t. Similarly define p_B^t and p_C^t.

1. What is p_C^3?

   Solution: $p_C^3 = \text{A1A2}$.

2. Define p_2^t as the probability of observing 2 at time point t. Express p_2^t as a function of p_B^t and p_C^t (you can also use any of the model parameters defined in the figure if you need).

   Solution: $0.3 * (1 - p_B^t - p_C^t)$.

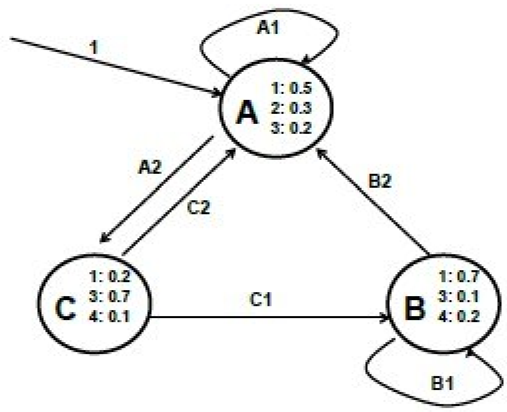
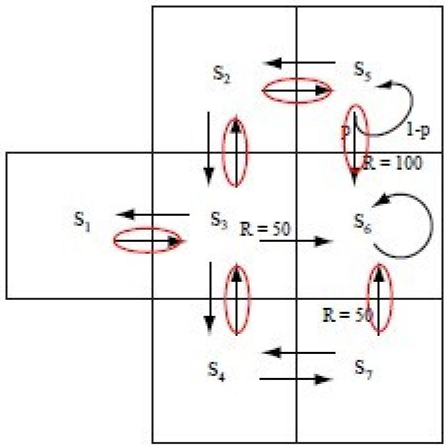


Figure 5: The HMM for 6.3.

8 Markov Decision Processes [12 points]

Consider the following Markov Decision Process (MDP), describing a simple robot grid world. The values of the *immediate rewards* R are written next to the transitions. Transitions with no value have an immediate reward of 0. Note that the result of the action "go south" from state S_5 results in one of two outcomes. With probability p the robot succeeds in transitioning to state S_6 and receives immediate reward 100. However, with probability $(1-p)$ it gets stuck in sand, and remains in state S_5 with zero immediate reward. **Assume the discount factor** $\gamma = 0.8$. **Assume the probability** $p = 0.9$.



1. Mark the state-action transition arrows that correspond to one *optimal* policy. If there is a tie, always choose the state with the smallest index.

   Solution: See figure.

2. Is it possible to change the value for γ so that the optimal policy is changed? If yes, give a new value for γ and describe the change in policy that it causes. Otherwise **briefly** explain why this is impossible.

   Solution: One possible solution: $\gamma = .7$. The optimal policy now takes action $S_3 \to S_6$.

3. Is it possible to change the immediate reward function so that V^* changes but the optimal policy π^* remains unchanged? If yes, give such a change and describe the resulting changes to V^*. Otherwise **briefly** explain why this is impossible.

   Solution: Yes. One possible answer is to double each reward. V^* is also doubled but the policy remains unchanged.

4. How sticky does the sand have to get before the robot will prefer to completely avoid it? Answer this question by giving a probability for p below which the optimal policy chooses actions that completely avoid the sand, even choosing the action "go west" over "go south" when the robot is in state S_5.

   Solution:

$$50\gamma^2 = \frac{100p}{1-\gamma(1-p)} \implies p = \frac{\gamma^2 - \gamma^3}{2 - \gamma^3} = \frac{8}{93} \approx 0.086$$

1 Hidden Markov Model [60 points]

Hidden Markov Model is an instance of the state space model in which the latent variables are discrete. Let K be the number of hidden states. We use the following notations: $\mathbf{x}$ are the observed variables, $\mathbf{z}$ are the hidden state variables (we use 1-of-K representation: $z_k = 1$, $z_{j \neq k} = 0$ means the hidden state is k). The transition probabilities are given by a $K \times K$ matrix $\mathbf{A}$, where $A_{jk} = p(z_{n,k} = 1 | z_{n-1,j} = 1)$ and the initial state variable $\mathbf{z}_1$ are given by a vector of probabilities π: $p(\mathbf{z}_1|\pi) = \prod_{k=1}^{K} \pi_k^{z_{1k}}$. Finally, the emission distribution for a hidden state k is parametrized by ϕ_k: $p(\mathbf{x}_n|\phi_k)$. Let $\mathbf{\Theta} = \{\mathbf{A}, \pi, \phi\}$.

1.1 The full likelihood of a data set

If we have a data set $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$:

1. **[3 points]** What is the full likelihood of observed and latent variables: $p(\mathbf{X}, \mathbf{Z}|\mathbf{\Theta})$? Note $\mathbf{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$ are the hidden states of the corresponding observations.

$$p(\mathbf{X}, \mathbf{Z}|\mathbf{\Theta}) = p(\mathbf{z}_1) \prod_{n=2}^{N} p(\mathbf{z}_n|\mathbf{z}_{n-1}) \prod_{n=1}^{N} p(x_n|z_n) \tag{1}$$

$$= \prod_{k=1}^{K} \pi_k^{z_{1k}} \prod_{n=2}^{N} \prod_{k=1}^{K} \prod_{j=1}^{K} A_{jk}^{z_{nk} z_{n-1,j}} \prod_{n=1}^{N} \prod_{k=1}^{K} p(x_n|\phi_k)^{z_{nk}} \tag{2}$$

2. **[2 points]** What is the likelihood of the data set? (e.g. $p(\mathbf{X}|\mathbf{\Theta})$).

$$p(\mathbf{X}|\mathbf{\Theta}) = \sum_{\mathbf{Z}} p(\mathbf{X}, \mathbf{Z}|\mathbf{\Theta}) \tag{3}$$

1.2 Expectation-Maximization (EM) for Maximum Likelihood Learning

We'd like to derive formulas for estimating $\mathbf{A}$ and ϕ to maximize the likelihood of the data set $p(\mathbf{X}|\mathbf{\Theta})$.

1. **[5 points]** Assume we can compute $p(\mathbf{X}, \mathbf{Z}|\mathbf{\Theta})$ in $O(1)$ time complexity, what is the time complexity of computing $p(\mathbf{X}|\mathbf{\Theta})$?

$O(K^N)$

We use EM algorithm for this task:

- In the E step, we take the current parameter values and compute the posterior distribution of the latent variables $p(\mathbf{Z}|\mathbf{X}, \mathbf{\Theta}^{\text{old}})$.
- In the M step, we find the new parameter values by solving an optimization problem:

$$\Theta^{\text{new}} = \text{argmax}_{\mathbf{\Theta}} Q(\mathbf{\Theta}, \mathbf{\Theta}^{\text{old}}) \tag{4}$$

where

$$Q(\Theta, \Theta^{\text{old}}) = \sum_{\mathbf{Z}} p(\mathbf{Z}|\mathbf{X}, \mathbf{\Theta}^{\text{old}}) \ln p(\mathbf{X}, \mathbf{Z}|\mathbf{\Theta}) \tag{5}$$

2. **[10 points]** Show that

$$Q(\Theta, \Theta^{\text{old}}) = \sum_{k=1}^{K} \gamma(z_{1k}) \ln \pi_k + \sum_{n=2}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln A_{jk} \tag{6}$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln p(\mathbf{x}_n | \phi_k) \tag{7}$$

where

$$\gamma(z_{nk}) = \mathbb{E}_{p(\mathbf{z}_n | \mathbf{X}, \Theta^{\text{old}})}[z_{nk}] \tag{8}$$

$$\xi(z_{n-1,j}, z_{nk}) = \mathbb{E}_{p(\mathbf{z}_{n-1}, \mathbf{z}_n | \mathbf{X}, \Theta^{\text{old}})}[z_{n-1,j} z_{nk}] \tag{9}$$

Show your derivations.

$$Q(\Theta, \Theta^{\text{old}}) = \sum_{\mathbf{Z}} p(\mathbf{Z}|\mathbf{X}, \Theta^{\text{old}}) \ln p(\mathbf{X}, \mathbf{Z}|\Theta) \tag{10}$$

$$= \sum_{\mathbf{Z}} \Big\{ \sum_{k=1}^{K} z_{1k} \log \pi_k + \sum_{n=2}^{N} \sum_{k=1}^{K} \sum_{j=1}^{K} z_{nk} z_{n-1,j} \log A_{jk} \tag{11}$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} z_{nk} \log p(x_n | \phi_k) \Big\} \tag{12}$$

$$= \sum_{k=1}^{K} \gamma(z_{1k}) \ln \pi_k + \sum_{n=2}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} \xi(z_{n-1,j}, z_{nk}) \ln A_{jk} \tag{13}$$

$$+ \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma(z_{nk}) \ln p(\mathbf{x}_n | \phi_k) \tag{14}$$

3. **[5 points]** Show that

$$p(\mathbf{X}|\mathbf{z}_{n-1}, \mathbf{z}_n) = p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_{n-1}) p(\mathbf{x}_n|\mathbf{z}_n) p(\mathbf{x}_{n+1}, \ldots \mathbf{x}_N|\mathbf{z}_n) \tag{15}$$

This follows from the following D-separations:

$$p(\mathbf{X}|\mathbf{z}_{n-1}, \mathbf{z}_n) = p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_{n-1}, \mathbf{z}_n) p(\mathbf{x}_n|\mathbf{z}_{n-1}, \mathbf{z}_n) p(\mathbf{x}_{n+1}, \ldots \mathbf{x}_N|\mathbf{z}_{n-1}, \mathbf{z}_n) \tag{16}$$

$$= p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_{n-1}) p(\mathbf{x}_n|\mathbf{z}_n) p(\mathbf{x}_{n+1}, \ldots \mathbf{x}_N|\mathbf{z}_n) \tag{17}$$

4. **[10 points]** In class, we discuss how to compute:

$$\alpha(\mathbf{z}_n) = p(\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{z}_n) \tag{18}$$
$$\beta(\mathbf{z}_n) = p(\mathbf{x}_{n+1}, \ldots, \mathbf{x}_N | \mathbf{z}_n) \tag{19}$$

Show that

$$\xi(\mathbf{z}_{n-1}, \mathbf{z}_n) = p(\mathbf{z}_{n-1}, \mathbf{z}_n | \mathbf{X}) \tag{20}$$
$$= \frac{\alpha(\mathbf{z}_{n-1}) p(\mathbf{x}_n|\mathbf{z}_n) p(\mathbf{z}_n|\mathbf{z}_{n-1}) \beta(\mathbf{z}_n)}{p(\mathbf{X})} \tag{21}$$

$$\xi(\mathbf{z}_{n-1}, \mathbf{z}_n) = p(\mathbf{z}_{n-1}, \mathbf{z}_n | \mathbf{X}) \tag{22}$$
$$= \frac{p(\mathbf{X}|\mathbf{z}_{n-1}, \mathbf{z}_n) p(\mathbf{z}_{n-1}, \mathbf{z}_n)}{p(\mathbf{X})} \tag{23}$$
$$= \frac{p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_{n-1}) p(\mathbf{x}_n|\mathbf{z}_n) p(\mathbf{x}_{n+1}, \ldots, \mathbf{x}_N|\mathbf{z}_n) p(\mathbf{z}_n|\mathbf{z}_{n-1}) p(\mathbf{z}_{n-1})}{p(\mathbf{X})} \tag{24}$$
$$= \frac{\alpha(\mathbf{z}_{n-1}) p(\mathbf{x}_n|\mathbf{z}_n) p(\mathbf{z}_n|\mathbf{z}_{n-1}) \beta(\mathbf{z}_n)}{p(\mathbf{X})} \tag{25}$$

How would you compute $p(\mathbf{X})$?

$$p(\mathbf{X}) = \sum_{\mathbf{z}_N} \alpha(\mathbf{z}_N) \tag{26}$$

5. **[5 points]** Show how to compute $\gamma(z_{nk})$ and $\xi(z_{n-1,j}, z_{nk})$ using $\alpha(\mathbf{z}_n)$, $\beta(\mathbf{z}_n)$ and $\xi(\mathbf{z}_{n-1}, \mathbf{z}_n)$.

$$\gamma(z_{nk}) = \mathrm{E}_{p(\mathbf{z}_n|\mathbf{X}, \mathbf{\Theta}^{\text{old}})}[z_{nk}] \tag{27}$$
$$= p(z_{nk} = 1 | \mathbf{X}, \mathbf{\Theta}^{\text{old}}) \tag{28}$$
$$= \frac{\alpha(\mathbf{z} : z_{nk} = 1) \beta(\mathbf{z} : z_{nk} = 1)}{p(\mathbf{X})} \tag{29}$$
$$\xi(z_{n-1,j}, z_{nk}) = \mathrm{E}_{p(\mathbf{z}_{n-1}, \mathbf{z}_n|\mathbf{X}, \mathbf{\Theta}^{\text{old}})}[z_{n-1,j} z_{nk}] \tag{30}$$
$$= p(z_{n-1,j} = 1, z_{nk} = 1 | \mathbf{X}, \mathbf{\Theta}^{\text{old}}) \tag{31}$$
$$= \xi(\mathbf{z}_{n-1}, \mathbf{z}_n : z_{n-1,j} = 1, z_{nk} = 1) \tag{32}$$

6. **[5 points]** Show that if any elements of the parameters π or $\mathbf{A}$ for a hidden Markov model are initially set to 0, then those elements will remain zero in all subsequent updates of the EM algorithm.

> The update rules for π and $\mathbf{A}$:
>
> $$\pi_k = \frac{\gamma(z_{1k})}{\sum_{j=1}^{K} \gamma(z_{1j})} \tag{33}$$
>
> $$A_{jk} = \frac{\sum_{n=1}^{N} \xi(z_{n-1,j} z_{nk})}{\sum_{l=1}^{K} \sum_{n=2}^{N} \xi(z_{n-1,j} z_{nl})} \tag{34}$$
>
> Considering the equation (25), if $A_{jk} = 0$ in the current iteration, $\xi(z_{n-1,j} z_{nk})$ is also zero (the term $p(\mathbf{z}_n|\mathbf{z}_{n-1})$ is zero). Therefore, all the future updates always set this to zero.

1.3 A coin game [15 points]

Two students X and Y from Cranberry Lemon University play a stochastic game with a fair coin. X and Y take turn with X going first. All the coin flips are recorded and the game finishes when a sequence of THT first appears. The player who last flips the coin is the winner. Two players can flip the coin many times as follows. At his turn, each time X flips the original coin, he also flips an extra biased coin ($p(H) = 0.3$.) He stops only if the extra coin lands head, otherwise he repeats flipping the original and extra coins, ... (The flips of this extra coin are not recorded.) On the other hand, at his turn, Y flips the coin until T appears (All of his flips are recorded).

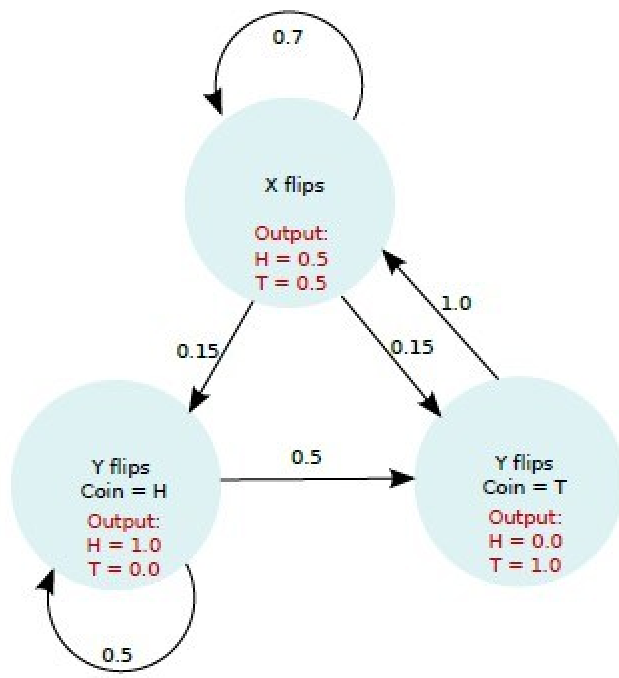
You are given a sequence of recorded coin flips, you would like to infer the winner and as well as the flips of each player.

1. **[10 points]** Describe a HMM to model this game.

> Figure 1. Note that for a valid HMM, the output of the current state is independent of the next state, given the current state ($\mathbf{x_n} \perp \mathbf{z}_{n+1}|\mathbf{z}_n$). Many answers violate this constraint.

2. **[5 points]**How would you use this HMM model to infer the (most probable) winner and the (most probable) flips of each player?

> It is straightforward to use Viterbi's algorithm to infer the most probably values of hidden variables. From this sequence of values, we can infer the the (most probable) winner and the (most probable) flips of each player.


0.7
X flips
Output:
H = 0.5
T = 0.5
0.15
0.15
1.0
Y flips
Coin = H
Output:
H = 1.0
T = 0.0
0.5
Y flips
Coin = T
Output:
H = 0.0
T = 1.0
0.5