Probability and Bayesian Networks Exercises

Prof. Jim Rehg
CS 3600 Introduction to Artificial Intelligence
School of Interactive Computing
Georgia Institute of Technology

March 25, 2018

These practice exercises are for your benefit in preparing for the exams. They will not be collected or graded. Solutions will be provided. Note that not all questions are representative of questions you will find on the exam, but the material covered by these questions will also be covered by the exams.

Question 1. **Events and Combinatorics**

For each problem below, *state* the sample space Ω, the event space A, and the probability of the event $P(A)$. Assume the standard probability models hold for dice, cards, etc.

(a) (1 pts) You roll two 6-sided dice, what is the probability of getting a 7 or an 11?

(b) (1 pts) You roll two 6-sided dice, what is the probability that the total is greater than 9 or even? What is the probability that it is greater than 9 *and* even?

(c) (1 pts) You roll *three* 6-sided dice, what is the probability that no two of the numbers are the same (i.e. no combinations such as $(1, 3, 1)$ or $(2, 2, 2)$ occur).

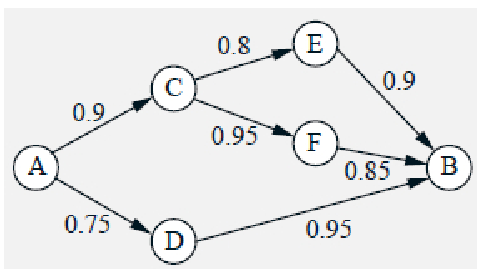
(d) (1 pts) What is the probability of being dealt a flush (all cards of the same suit) in 5 card poker?

(e) (1 pts) In the game of blackjack (sometimes called 21), the goal is to beat the dealer by having the highest count of cards without going bust (exceeding 21). The dealer has one card face-down (the hole card) and all other cards are dealt face up. The dealer must hit (take a card) if his total is 16 or less, and stand (end his turn) otherwise. The face cards (jack, queen, and king) are worth 10, assume that the ace is worth 11. Suppose your total is 13, what is the probability of a bust if you hit? Note: Assume that the probabilities for all cards are equal, in other words that the deck is infinite (or equivalently that cards are dealt with replacement).

(f) (1 pts) Same setting as in (e), and in addition the dealer is showing a 6. Suppose the dealer ignores the rules and is going to stand no matter what her hole card is. If you hit once and then stand, what is the probability that you will win? Note: In the event of a tie, the dealer wins.

(g) (1 pts) An agent starts at the point $(0, 0)$ in the plane and moves to the goal point $(100, 101)$ by taking a sequence of steps, where each step *either* goes to the right one unit (R) or goes up one unit (U). Clearly, there are many possible valid paths that reach the goal. If the agent selects a valid path at random, what is the probability that it consists of 100 R-steps followed by 101 U-steps?

(h) (1 pts) Consider the following communication network, where each arrow represents a link between two nodes. The number for each link is the probability that it is working. What is the probability that a path exists between A and B such that all of the links along the path are working? Note that the link outcomes are independent events.



Hint: Let p_i denote the probability that the ith link is working. In order to follow a series of links (e.g. $C \to E \to B$), each link in the series must be working (i.e. an *AND* connection). In that case, the probability that a path across m links in series is working is $p_1 p_2 \cdots p_m$. In contrast, when there are multiple links leaving the same node in parallel (e.g. $C \to E$ or $C \to F$), then if any one of the links works, there will be a path out of that node (i.e. an *OR* connection). In that case, the probability of success for m links in parallel is given by: $P(\text{connected}) = 1 - P(\text{not-connected}) = 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n)$. Note that it's much easier to compute the failure probability than it is to sum all of the successful cases.

Question 2. **Conditional Probabilities**

(a) (1 pts) Suppose two 6-sided dice are rolled. Given that the roll resulted in a sum of 4 or less, what is the probability that the roll was a double?

(b) (1 pts) Suppose two 6-sided dice are rolled. Given that the two dice land on different numbers, what is the probability that at least one die roll is a 6.

(c) (1 pts) Suppose that three 6-sided dice are rolled. What is the probability that no two numbers are the same *and* there is a 1 or a 6 among the numbers? *Hint:* Use the result from problem 1(c).

(d) (1 pts) We are given three coins: one has heads on both faces, one has tails on both faces, and one has a head and a tail and is fair. We choose a coin at random and then toss it, and it comes up heads. What is the probability that the opposite face is a tail (i.e. that it is the coin with a head and a tail)?

(e) (1 pts) In a game show (Monty Hall), you have to choose one of three doors. Behind one door is a new car, behind the other two are old goats. You choose, but your chosen door is not opened immediately. Instead, the presenter opens a different door, showing you what's behind it. He then gives you the opportunity to choose the third, unopened door *instead* of the one you originally selected. The question is whether you should do so. Answer the question by first assuming that the presenter *always* shows you a goat (in other words, when your door conceals a goat, he always chooses the door with the other goat, and when your door conceals the car, he will choose one of the other two doors at random.) Then answer the question a second time, assuming that the presenter chooses one of the other two doors at random, and opens it for you. If there is a goat behind that door, should you switch?

Question 3. **Conditional Independence** You are going to play two games of go against an opponent that you never played before. Your opponent is equally likely to be a beginner, intermediate, or expert player. Depending on which they are, your probability of winning is 90%, 50%, or 30% respectively.

(a) (1 pts) What is the probability that you win your first game?

(b) (1 pts) Suppose that you win the first game. What is the probability that you will also win the second game? *Note:* you should assume that the outcomes of the games are independent, *given* the skill level of your opponent.

(c) (1 pts) In part (b) you assumed that the outcomes of the two games are conditionally-indpendent given the opponent's skill level. Comment on the validity of this assumption relative to the alternative assumption that the game outcomes are marginally independent.

Question 4. **Probability Tables** Consider the following joint probability distribution $P(A, B, C)$ in three discrete random variables A, B, C:

$p(A = 1, B, C) =$

	$B = 1$	$B = 2$	$B = 3$
$C = 1$	0	0.05	0.05
$C = 2$	0.05	0.05	0.05
$C = 3$	0.05	0	0.05

$p(A = 2, B, C) =$

	$B = 1$	$B = 2$	$B = 3$
$C = 1$	0.1	0.1	0.2
$C = 2$	0.1	0	0
$C = 3$	0	0.1	0.05

(a) (2 pts) Compute $P(B|A = 2, C = 1)$.

(b) (2 pts) What is the *a priori* probability distribution over C?

Question 5. **Professor Bayes** A professor is heading back into the lab late at night and is trying to guess whether his student is working on their paper. He has noticed that whenever this student is working, their car is in the garage and there is music playing in the lab. Let W, C, M be three binary random variables corresponding to the events "student Working on paper", "Car in parking lot", and "Music playing." Through careful observation, the professor has determined the following conditional probabilities:

(1) $P(C|W, M) = P(C|W) =$

	$W = 1$	$W = 0$
$C = 1$	0.8	0.1
$C = 0$	0.2	0.9

(2) $P(M|W) = P(C|W)$

(3) $P(W = 1) = 0.8$

Note: Equation (2) says that C and M have the same conditional distribution (e.g. $P(C = 1|W = 1) = P(M = 1|W = 1) = 0.8$.)

(a) (1 pts) Draw a Bayesian network (i.e. a directed graphical model) for this problem.

(b) (2 pts) Compute the marginal distributions $P(C)$ and $P(M)$, before any evidence is availble

(c) (2 pts) Upon entering the garage, the professor notices that the student's car is parked there. Calculate $P(W|C = 1)$.

(d) (2 pts) Calculate the updated probability that music is playing in the lab, taking the evidence $C = 1$ into account.

Question 6. Safety Monitor

You are worried about the safety of your home, and constantly monitor whether your front door is open (D) using an iphone app. When your door is open there are two possible causes: your spouse is at home (S), and may have left the door open by accident, or your house was robbed (R). D, S, R are binary random variables. R and S are marginally independent (i.e. $R \perp S$) with prior probabilities $P(S = 1) = 0.5$ and $P(R = 1) = 0.1$. The following conditional probability table gives the probability of the door being open:

S	R	$P(D = 1 \mid S, R)$	$P(D = 0 \mid S, R)$
0	0	0.01	0.99
0	1	0.25	0.75
1	0	0.05	0.95
1	1	0.75	0.25

(a) (1 pts) Why don't each of the columns in the probability table above sum to 1?

(b) (1 pts) Draw a Bayes net model (i.e. a directed graphical model) for this problem.

(c) (1 pts) What is the probability that the door is open conditioned on being told that your spouse is at home and your house was not robbed?

(d) (1 pts) Calculate the joint probability $P(D = 1, S = 1, R = 0)$. *Hint:* Use the definition of conditional probability. How does this probability differ from your answer in part (c)? Explain the difference.

(e) (2 pts) Suppose that your only evidence is that your spouse is at home. Compute the conditional probability that the door is open (e.g. $P(D = 1|S = 1)$).

(f) (2 pts) Now suppose that your only evidence is that the door is open. Compute the posterior distribution of S and R.

Question 7. **Recursive Updating**

Derive the standard recursive updating rule for inference in a Hidden Markov Model that is used to obtain $P(x_t|e_{1:t})$, where x_t is the hidden state at time t and $e_{1:t}$ is the sequence of observations up to time t. The rule should be recursive (or incremental) in the sense that $P(x_t|e_{1:t})$ can be obtained given *only* $P(x_{t-1}|e_{1:t-1})$ and e_t (in other words, there is no need to store the previous measurements $e_{1:t-1}$ or reprocess them.)

Question 8. Dynamic Dean

The AI faculty decide to build an agent to guess the mood of Dean Zvi Galil from hour to hour (so visits to the Dean's office can be carefully timed). After detailed observation, they determine that when the Dean is in a good mood, there is a 75% chance he will be in a good mood an hour later. But when he is in a bad mood, there is a 50% chance he will be in a bad mood an hour later. Unfortunately the Dean's mood can't be observed directly. However, through careful study the faculty have determined that when the Dean is in a good mode, the likelihood that he will send a reminder email about CIOS is 80%, but there is only a 60% chance of sending a CIOS email when he is in a bad mood.

(a) (1 pts) Draw a dynamic Bayesian network model for this problem. Be sure to show at least two time slices. Label the nodes in your graph. Use the random variable M to indicate the Dean's mood. M can take on the values "good" or "bad." Use the random variable C to indicate whether a CIOS email was sent. C can take on the values "true" or "false." Use subscripts on M and C to indicate time slices.

(b) (2 pts) Fill in the conditional probability tables for the transition model and sensor model elements of the DBN. Be sure the columns and rows are labeled properly in the spaces provided.

Transition Model:

Measurement Model:

(c) (1 pts) The faculty are interested in tracking the Dean's mood carefully during the Final Exam period. The day before Finals start, the faculty have no knowledge of the Dean's mood. In other words, $P(M_0 = \text{good}) = 0.5$. Predict (through forward simulation) the probability of the Dean's mood on the first day of Finals:

$P(M_1 = \text{good}) =$

(d) (2 pts) Now suppose that the Dean sends a CIOS reminder email on the first day of Finals, in other words $C_1 = T$. Compute an updated posterior estimate for the Dean's mood M_1.

Question 9. **Particle Filter**

This question explores the particle filter algorithm for inference in an HMM using sampling methods. The hidden state is the location of an agent in a 5x5 grid world, which we parameterize as (u, v) for discrete u and v random variables. The transition model for a state (u, v) has 40% probability of staying in the same state and 15% probability of moving in each of the four compass directions N, S, E, and W. The measurement model assigns 60% probability of measuring the discrete state correctly, and 10% probability of generating an erroneous measurement corresponding to a displacement in one of the four compass directions.

(a) (2 pts) Write down the transition model relating (u_{t-1}, v_{t-1}) and (u_t, v_t) and the measurement model relating (u_t, v_t) and the corresponding measurements (x_t, y_t).

Transition Model:

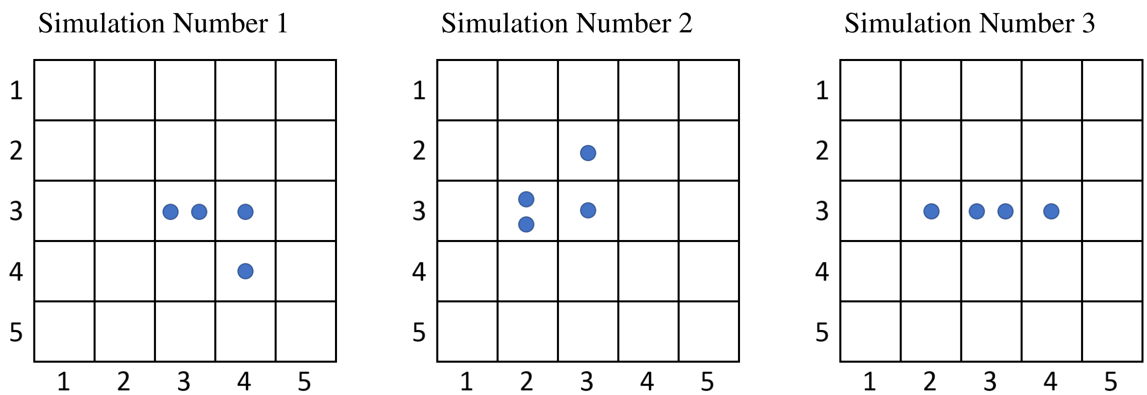
$$P(u_t = i, v_t = j | u_{t-1} = i, v_{t-1} = j) =$$
$$P(u_t = i + 1, v_t = j | u_{t-1} = i, v_{t-1} = j) =$$
$$P(u_t = i - 1, v_t = j | u_{t-1} = i, v_{t-1} = j) =$$
$$P(u_t = i, v_t = j + 1 | u_{t-1} = i, v_{t-1} = j) =$$
$$P(u_t = i, v_t = j - 1 | u_{t-1} = i, v_{t-1} = j) =$$

Measurement Model:

$$P(x_t = i, y_t = j | u_t = i, v_t = j) =$$
$$P(x_t = i + 1, y_t = j | u_t = i, v_t = j) =$$
$$P(x_t = i - 1, y_t = j | u_t = i, v_t = j) =$$
$$P(x_t = i, y_t = j + 1 | u_t = i, v_t = j) =$$
$$P(x_t = i, y_t = j - 1 | u_t = i, v_t = j) =$$

(b) (2 pts) Suppose that at time $t-1$ we have a set of four particles that are all located at $(3,3)$. Let $s^k_{t-1} = (u^k_{t-1}, v^k_{t-1})$ denote the kth particle, and let $S_{t-1} = \{s^1_{t-1}, s^2_{t-1}, s^3_{t-1}, s^4_{t-1}\}$ denote the particle set at time $t-1$. Then we have $s^k_{t-1} = (3,3), \quad k = 1, \ldots, 4$. The three figures below show three *different* possible sets of particles resulting from *one step* of forward simulation (i.e. the result of sampling once from $P(u_t, v_t | u_{t-1}, v_{t-1})$ for each particle). Which of these three particle sets is the *most likely* to result from the transition model in part (a)? Justify your answer numerically. *Hint:* Compute the likelihood of each particle set, which is given by:

$$P(S_t) = \prod_{k=1}^{4} P(u^k_t, v^k_t | u^k_{t-1}, v^k_{t-1}).$$



(c) (2 pts) Now suppose that the result of forward simulation is the distribution of particles in Simulation Number 2 from part (b). We obtain a measurement of the target position at time t, which is $m_t = (2, 3)$. Write the list of particles from Simulation Number 2 with their *likelihood weights*, following the measurement update step.

(d) (2 pts) Suppose that we perform *resampling* using the list of weighted particles given below. Calculate the probability of drawing each of the particles during resampling.

$(1, 2)$ $w = 0.1$
$(2, 1)$ $w = 0.5$
$(2, 2)$ $w = 0.6$
$(2, 2)$ $w = 0.6$
$(2, 4)$ $w = 0.3$
$(3, 2)$ $w = 0.2$
$(3, 3)$ $w = 0.4$
$(3, 3)$ $w = 0.4$
$(3, 3)$ $w = 0.4$
$(3, 5)$ $w = 0.2$
$(4, 3)$ $w = 0.8$
$(4, 4)$ $w = 0.7$