

## Homework 5: Graphical Models, MDPs

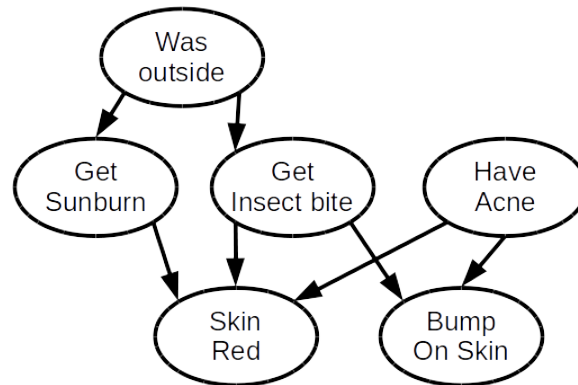
### Introduction

There is a mathematical component and a programming component to this homework. Please submit your **tex**, **PDF**, and **Python files** to Canvas, and push all of your work to your GitHub repository. If a question requires you to make any plots, please include those in the writeup.

## Bayesian Networks [7 pts]

### Problem 1

In this problem we explore the conditional independence properties of a Bayesian Network. Consider the following Bayesian network representing a person's skin condition. Each random variable is binary (true/false).



The random variables are:

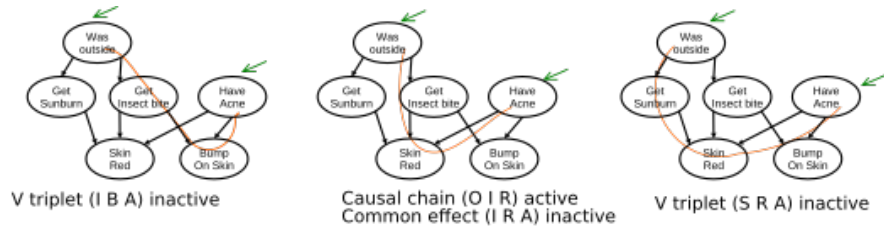
- Was Outside: Was the person outside?
- Get Sunburn: Did the person get sunburn?
- Get Insect Bite: Did the person get an insect bite?
- Have Acne: Does the person have acne?
- Skin Red: Is the skin red?
- Bump on Skin: Is there a bump?

For the following questions,  $I(A, B)$  means that events A and B are independent and  $I(A, B|C)$  means that events A and B are independent conditioned on C. Use the concept of d-separation to answer the questions and show your work.

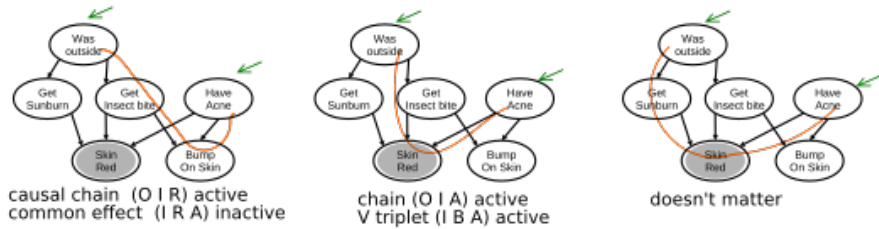
1. Is  $I(\text{Have Acne}, \text{Was Outside})$ ? If NO, give intuition for why.
2. Is  $I(\text{Have Acne}, \text{Was Outside} | \text{Skin Red})$ ? If NO, give intuition for why.
3. Is  $I(\text{Get Sunburn}, \text{Bump on Skin})$ ? If NO, give intuition for why.
4. Is  $I(\text{Get Sunburn}, \text{Bump on Skin} | \text{Get Insect Bite})$ ? If NO, give intuition for why.
5. Suppose the person has taken a medicine to suppress their response to insect bites: they get red skin, but no bumps. Draw the modified network.
6. For this modified network, is  $I(\text{Get Sunburn}, \text{Bump on Skin})$ ? If NO, give an intuition why. If YES, describe what observations (if any) would cause them to no longer be independent.

## Solution Bayesian Networks

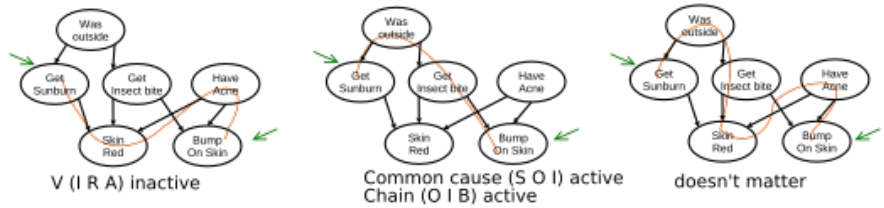
Solution: Please see figure for showing d-separation work.



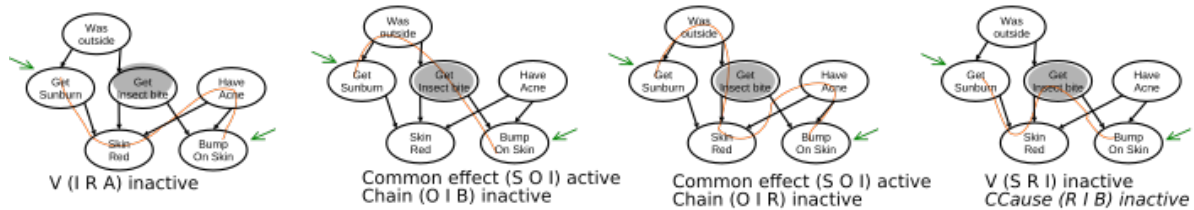
Part 1. YES. All paths between O and A are inactive; and O and A are guaranteed to be independent.



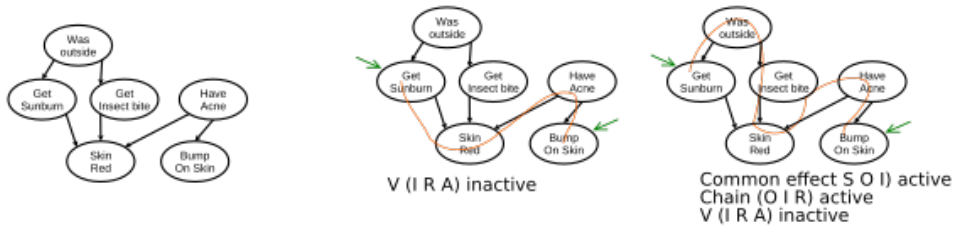
Part 2. NO. Active path between O and A, independence not guaranteed.



Part 3. NO. Active path between S and B, independence not guaranteed.



Part 4. YES. All paths between S and B inactive. Independence is guaranteed.



Part 5. YES. All paths between S and B inactive. Independence is guaranteed.

Part 6. If we observed R, then we would "activate" the V (common effect) structure I R A. There would then be an active path between S and B, and they would no longer be guaranteed to be independent.

Figure 1: Part 1. D-Separation Work

- Is  $I(\text{Have Acne}, \text{Was Outside})$ ? If NO, give intuition for why.

YES, Yes, having acne is independent of going outside. Whether we go outside or not doesn't affect whether we get acne (at least according to this model). Also, having acne doesn't affect whether we go outside or not (according to this model).

- Is  $I(\text{Have Acne}, \text{Was Outside} | \text{Skin Red})$ ? If NO, give intuition for why.

NO, If we know whether or not we have acne, does also knowing whether our skin is red or not impact our estimate of whether we were outside? If we know we have red skin, and knowing we had acne would "explain away" some of our certainty about being outside or not. Thus, having observed R, O and A are no longer independent.

- Is  $I(\text{Get Sunburn}, \text{Bump on Skin})$ ? If NO, give intuition for why.

NO, Not independent. The likelihood of us getting sunburned impacts our estimate of being outside, which in turn impacts our estimate of getting an insect bite and potentially having a bump on our skin.

- Is  $I(\text{Get Sunburn}, \text{Bump on Skin} | \text{Get Insect Bite})$ ? If NO, give intuition for why. YES, If we know whether or not we got an insect bite, do we care about whether we have a bump or not, in order to estimate if we are sunburned or not? No! (thus YES, they are independent). That is, the insect bite directly tells us whether the bump in the skin was caused by being outside, (vs having acne), and whether we got a sunburn or not provides no additional information.
- Suppose the person has taken a medicine to suppress their response to insect bites: they get red skin, but no bumps. Draw the modified network.
- For this modified network, is  $I(\text{Get Sunburn}, \text{Bump on Skin})$ ? If NO, give an intuition why. If YES, describe what observations (if any) would cause them to no longer be independent.

If we observed R, then we would "activate" the V (common effect) structure I R A. There would then be an activate path between S and B, and they would no longer be guaranteed to be independent.

Intuitively, if we know that we have red skin, then knowing that we got sunburned would lower the likelihood that our red skin was caused by acne. Having acne impacts whether we get bumps or not. Therefore, after we know whether or not we were our skin being reddened, then knowing whether we were sunburned or not adds information to our estimate of whether we have bumps or not.

## Kalman Filters [7 pts]

### Problem 2

In this problem, you will implement a one-dimensional Kalman filter. Assume the following dynamical system model:

$$\begin{aligned}z_{t+1} &= z_t + \epsilon_t \\x_t &= z_t + \gamma_t\end{aligned}$$

where  $z$  are the hidden variables and  $x$  are the observed measurements. The random variables  $\epsilon$  and  $\gamma$  are drawn from the following Normal distributions:

$$\begin{aligned}\epsilon_t &\sim N(\mu_\epsilon, \sigma_\epsilon) \\ \gamma_t &\sim N(\mu_\gamma, \sigma_\gamma)\end{aligned}$$

where  $\mu_\epsilon = 0$ ,  $\sigma_\epsilon = 0.05$ ,  $\mu_\gamma = 0$  and  $\sigma_\gamma = 1.0$

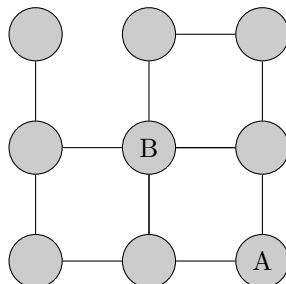
You are provided with the observed data  $x$  and the hidden data  $z$  in `kf-data.csv`, and the prior on the first hidden state is  $p(z_0) = N(\mu_p, \sigma_p)$  where  $\mu_p = 5$  and  $\sigma_p = 1$

- The distribution  $p(z_t | x_0 \dots x_t)$  will be Gaussian  $N(\mu_t, \sigma_t^2)$ . Derive an iterative update for the mean  $\mu_t$  and variance  $\sigma_t^2$  given the mean and variance at the previous time step ( $\mu_{t-1}$  and  $\sigma_{t-1}^2$ ).
- Implement this update and apply it to the observed data above (do not use the hidden data to find these updates). Provide a plot of  $\mu_t$  over time as well as a  $2\sigma_t$  interval around  $\mu_t$  (that is  $\mu_t \pm 2\sigma_t$ ). Does the Kalman filter “catch up” to the true hidden object?
- Repeat the same process but change the observation at time  $t = 10$  to  $x_{t=10} = 10.2$  (an extreme outlier measurement). How does the Kalman filter respond to this outlier?
- Comment on the misspecification of dynamical system model for these data. Based on the previous two parts, how does this misspecification affect the predictions?

## Markov Decision Processes [7 pts]

### Problem 3

In this problem we will explore the calculation of the *MDP value function*  $V$  in a 2D exploration setting, without time discounting and for a finite time horizon. Consider a robot navigating the following grid:



The robot moves in exactly one direction each turn (and must move). The robot's goal is to maximize its score in the game after  $T$  steps. The score is defined in the following way:

- If the robot's movement attempts to move it off the grid, then the robot loses a point (-1) and stays where it is.
  - If the robot ends its motion on node A, it receives 10 points, and if it moves onto node B, it receives -100 points. Otherwise, it receives 0 points.
1. Model this as a Markov decision process: define the states  $S$ , actions  $A$ , reward function  $r : S \times A \mapsto \mathbb{R}$ , and transition model  $p(s' | s, a)$  for  $s', s \in S$  and  $a \in A$ . For now, assume that the robot's actions execute perfectly: if the robot tries to move in a particular direction, it always succeeds in doing so.
  2. Consider a *random policy*  $\pi$ , where in each state the robot moves uniformly at randomly in any of its available directions (including off the board). For every position on the grid calculate the value function,  $V_t^\pi : S \mapsto \mathbb{R}$ , under this policy, for  $t = 2, 1$  steps left to go. You can find LaTeX code for the tables in the solution template. Note that you should have 2 tables, one for each time horizon.
  3. Now assume that the robot plays an *optimal policy*  $\pi_t^*$  (for  $t$  time steps to go). Find the optimal policy in the case of a finite time horizon of  $t = 1, 2$  and give the corresponding MDP value functions  $V_t^* : S \mapsto \mathbb{R}$ , under this optimal policy. You can indicate the optimal policy for each time horizon on the corresponding  $V_t^*$  table via arrows or words in the direction that the robot should move from that state.
  4. Now consider the situation where the robot does not have complete control over its movement. In particular, when it chooses a direction, there is a 80% chance that it will go in that direction, and a 10% chance it will go in the two adjacent (90° left or 90° right) directions. Explain how this changes the elements  $S$ ,  $A$ ,  $r$ , and  $p(s' | s, a)$  of the MDP model.

Assume the robot uses the same policy  $\pi_t^*$  from the previous question (now possibly non-optimal), and write this as  $\pi_t$ , and tie-break in favor of N, then E, then S then W. Give the corresponding MDP value functions  $V_t^\pi : S \mapsto \mathbb{R}$ , for this policy in this partial control world, for  $t = 2, 1$  steps left to go. Is the policy still optimal?

## A solution section

Some solution

Figure 2: some caption

- Is blah blah?

YES, blah blah blah

- test asdfsf



**Problem 4** (Calibration, 1pt)

Approximately how long did this homework take you to complete?

- Name:
- Email:
- Collaborators: