Lab week 6: Inference with Uncertainty **Print Last Name:** **Winter**

1. A sales representative wants to sell you an alarm system. They tell you that:
   The alarm system will warn you of a break-in 99% of the time if there is a break-in but the competition will only warn you 85% of the time there is a break-in.
   Suppose that the buyer lives in a relatively safe area and the probability of a break-in on a given day, P(B) is 0.001. In addition, the alarm system is not only set off by a break-in but also by animals such as birds or cats in the yard which gives a total probability of P(A) = 0.1. Do you think the buyer would be happy if they purchase this alarm system? Hint: (Estimate $P(A|\neg B)$)

Sale:

Alarm #1(Selling)	0.99
Alarm #2 (Competition)	0.85

House Alert Situation:

Break-In	0.001
Animal	0.1

$P(A|\neg B) = [P(A) \times P(\neg B|A)] / P(\neg B)$

$\rightarrow P(A) = 0.1$

$\rightarrow P(\neg B) = 1 - 0.001 = 0.999$

$\rightarrow P(\neg B|A) = P(A \cap \neg B) / P(A) = (0.1 - 0.99 \times 0.001) / 0.1 = 0.99$

$= P(A|\neg B) = (0.1 * 0.99) / 0.999 = 0.0991$ -> 9.91% of the time

The buyer would be relatively fine if they bought either system since it would accidentally go off relatively 9% of the time.

Look up the meaning of sensitivity vs specificity?

Sensitivity is : The probability of a positive test result, conditioned on the individual truly being positive (true positive rate); meaning, the fraction of values predicted to be of a positive class out of all the values that truly belong to the positive class

Specificity is : The probability of a negative test result, conditioned on the individual truly being negative (true negative rate); meaning, the fraction of values predicted to be of a negative class out of all the values that truly belong to the negative class

Mammogram posterior probabilities: Approximately 1% of women aged 40-50 have breast cancer. A woman with breast cancer has a 90% chance of a positive test from a mammogram, while a woman without has a 10% chance of a false positive result.
What is the probability a woman has breast cancer given that she just had a positive test?

P(A) = 0.01
P(B | A) = 0.90
P(B | ¬A) = 0.10
P(B) = (0.1 * 0.9) + (0.1 * 0.99) = 0.108

"Given the positive test 'B', what is the probability of a woman having breast cancer 'A'"
P(A | B) = [P(A) × P(B|A)] / P(B)
P(A | B) = [0.01 × 0.90] / 0.108 = 0.083 → 8.3% chance

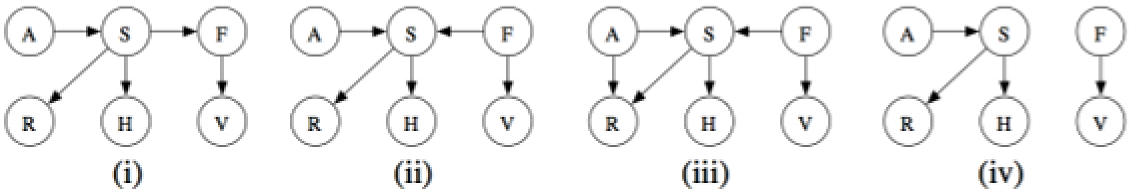
2. **Monty Hall Problem:** Suppose you're on a game show, and you're given the choice of three closed doors: Behind one door is a car; behind the others, a donkey.
   - You decide to pick a door, say #.1,
   - But then without saying anything or doing anything else, the host, who knows what's behind the doors, opens another door, say #.3, which you see has a donkey.
   - He then says to you, "Do you want to pick door No.2?"

   Is it to your advantage to switch your choice? (Whitaker 1990)
   Justify your answer using probability.

   It is not in your advantage to switch your choice. This is because, before the host opened the door, you had a 1/3 chance that the door you selected had the car behind it. After the host opens the door with the donkey, your chances increase to 2/3 since a door has become known. Now, either door you pick, you will have the same exact odds.

3. **Snuffles**: Assume
   - There are two types of conditions: **(S)inus** congestion and**(F)lu.**
   - Sinus congestion is caused by **(A)llerg**y or the **flu**.
   - There are three observed symptoms for these conditions: **(H)eadache, (R)unnynose, and fe(V)er.**
   - **Runnynose** and **headaches** are directly caused by sinus congestion(only),
   - While **fever** comes from having the **flu** (only). For example, **allergies** only cause **runnynoses** indirectly.
   - Assume each variable is boolean.

Consider the four Bayes Nets shown. Circle the one which models the domain as described above best?



CIRCLE: (ii)

Explain why each of the other 3 are not correct:

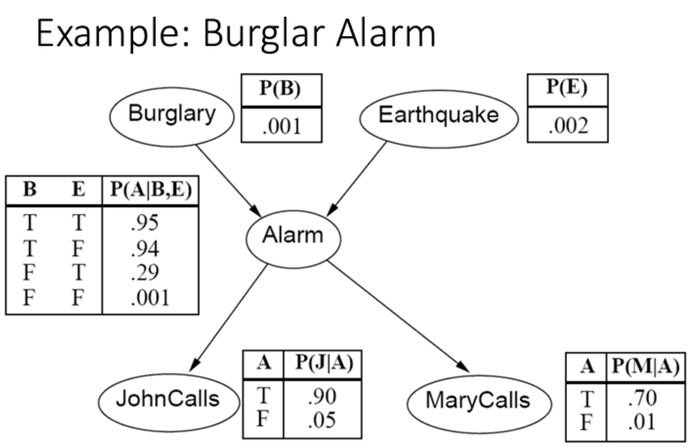
A. (i) is incorrect because the model shows that Sinus congestion *implies* the Flu, when in reality Sinus congestion *is a result of* the Flu.

B. (iii) is incorrect because the model shows that Allergies can *directly* cause a Runny Nose, when in reality allergies only cause runnynoses *indirectly*.

C. (iv) is incorrect because the model shows that Sinus congestion can only occur as a result of *Allergies*, Sinus congestion can occur when from *Allergies or the Flu*.

4. **Los Angeles Burglar Alarm**

- I have a burglar alarm that is sometimes set off by minor earthquakes. My two neighbors, John and Mary, promised to call me at work if they hear the alarm
  - Example inference task: suppose Mary calls and John doesn't call. What is the probability of a burglary?
- What are the random variables?
  - Burglary, Earthquake, Alarm, JohnCalls, MaryCalls
- What are the direct influence relationships?
  - A burglar can set the alarm off
  - An earthquake can set the alarm off
  - The alarm can cause Mary to call
  - The alarm can cause John to call



A. Classification using probabilities. Suppose Mary has called to tell you that you had a burglar alarm. Should you call the police?

Make a decision that maximizes the probability of being correct. This is called a MAP (maximum a posteriori) decision. You decide that you have a burglar in your house if and only if $\boldsymbol{P(Burglary \mid Mary) > P(\neg Burglary \mid Mary)}$

DECISION: You do not call the police because P(B | M) < P(¬B | M)

Need to compute − $P\ (Burglary \mid Mary)$

Step 1: Find the joint probability of B (and $\neg B$), M (and $\neg M$), and any other variables that are necessary in order to link these two together.

$P\ (B, E, A, M) = P\ (B)\ P\ E)\ P(A|B, E)\ P(M|A)$

$P(B,E,A,M)$	$(\neg M,\neg A)$	$(\neg M, A)$	$(M,\neg A)$	(M, A)
(¬B,¬E)	0.9859846427	0.001502	0.0095	0.0000014
(¬B, E)	0.00197	0.000000301	0.0000019926	0.0000000029
(B,¬E)	0.000987	0.0000001567	0.000000956	0.00000000143
(B, E)	0.0000000193	0.000000093	0.000000084	0.00000002

Step 2: marginalize (add) to get rid of the variables you don't care about.

$$P(B, M) = \sum_{E,\neg E} \sum_{A,\neg A} P(B, E, A, M)$$

$P(B,M)$	$\neg M,$	M
¬B	0.98946	0.0009530
B	0.0000996	0.0000009548

Step 3: Ignore (delete) the column that didn't happen.

$P(B,M)$	M
¬B	0.0009530
B	0.0000009548

Step 4: Use the definition of conditional probability to normalize so probabilities sum to 1

$$\mathrm{P(B|M)} = \frac{P(B,M)}{P(B,M) + P(B,M)}$$

$P(B,M)$	M
¬B	0.99999
B	0.00001