Name: Jonah Biedermann

CS 331 Final Exam
Spring 2023

You have 110 minutes to complete this exam including the 15 minutes for the online quiz in the beginning. You are only allowed to use your textbook, your notes, lecture slides, your assignments and solutions to those assignments and the midterm during this exam. Please read all the questions carefully and follow directions. For all questions, partial credit can sometimes be granted when you show your work, thought process, assumptions, etc. Please submit only pdf files.

Section I: Pre-Midterm questions
This has been turned into "Fnal Quiz." You should submit it by 2:15.

Section II. Logical Agents

2.Represent the following sentences in propositional logic and convert them to clausal form (CNF) (6 points)

(a) John is either at work or at home.
¬At work -> At home : translates to
At Work ∨ At Home

(b) If he is at home his light is on.
At home -> light is on : translates to
¬ At home ∨ Light is on

(c) If his light is on I can see it from my home.

Light is on -> I can see it from my house

¬Light is on ∨ I can see it from my house

(d) I don't see John's light.
¬ See johns light

Show that John is at work using proof by resolution. (4 points)

Negated goal: john is at home

If john is home his light would be on and if the light was on I could see it from my house but since I cannot see the johns light he must not be home.

Section II: Probability

3. An AI spam filter has recently been installed in all computer systems. It has 95% sensitivity (probability of flagging a spam message). However it has 20% false positive rate, i.e., it incorrectly flags 20% of non-spam messages as spam. Past data indicates that 40% of the messages are spam. Define two random variables S for whether the message is spam or not and F for the spam filter flagging the message as spam.

Answer the following questions. It is important to show your work.

a) Represent all the given information in the probabilistic notation. [1 point]

S = Event that the message is spam F = Event that the spam filter flags the message as spam

Given probabilities: P(S) = 0.4 (40% of messages are spam) P(F|S) = 0.95 (sensitivity - probability of flagging a spam message) P(F|¬S) = 0.2 (false positive rate - probability of incorrectly flagging a non-spam message)

We can now express the probabilities using the probabilistic notation:

P(S = true) = 0.4

P(S = false) = 1 - P(S = true) = 0.6

P(F = true | S = true) = 0.95

P(F = false | S = true) = 1 - P(F = true | S = true) = 0.05

P(F = true | S = false) = 0.2

$P(F = false \mid S = false) = 1 - P(F = true \mid S = false) = 0.8$

b) What is the joint probability that the message is spam and the spam filter detects it? [1 point]

$P(S = true, F = true) = P(F = true \mid S = true) * P(S = true)$

$P(S = true, F = true) = 0.95 * 0.4 = 0.38$

c) What is the joint probability that the message is not spam and the spam filter still flags it as spam? [1 point]

$P(S = false, F = true) = P(F = true \mid S = false) * P(S = false)$

$P(S = false, F = true) = 0.2 * 0.6 = 0.12$

d) What is the probability that the spam filter flags a random message as spam? [1 point]

$P(F = true) = P(S = true, F = true) + P(S = false, F = true)$

$P(F = true) = 0.38 + 0.12 = 0.5$

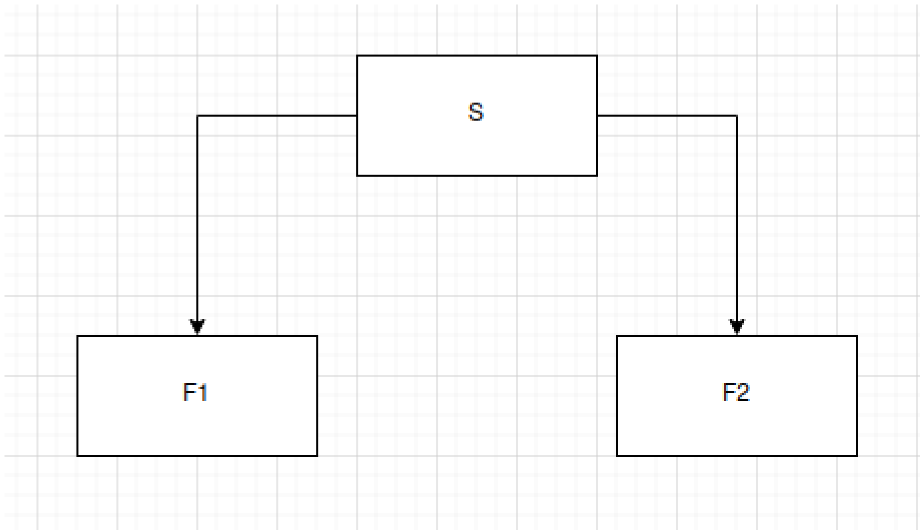
e) What is the probability that a message which is flagged as spam is indeed spam? [1 point]

$P(S = true \mid F = true) = P(S = true, F = true) / P(F = true)$

$P(S = true \mid F = true) = 0.38 / 0.5 = 0.76$

Assume now that there are two spam filters made by two companies. Introduce two random variables F1 and F2 to represent their flags. Assume that the flags of the two filters are independent given the spam status (spam or not) of the message.

f) Draw a Bayesian network that represents this relationship and show what conditional probability distributions are needed to fully specify it. [2 points]



g) Derive an expression for the message being a spam if both the spam filters flag it as a function of the conditional probabilities of your network. [2 points]

We want to calculate the probability of S given that both F1 and F2 occur, i.e., P(S | F1, F2).

h) Assuming that the second filter has the same sensitivity and false positive rate as the first one, calculate the exact probability in (g) [1 point]

Using Bayes' theorem, we have:

P(S | F1, F2) = P(F1, F2 | S) * P(S) / P(F1, F2)

We can expand the numerator using the chain rule:

P(F1, F2 | S) = P(F2 | F1, S) * P(F1 | S) * P(S)

Assuming that the two filters, F1 and F2, are independent given the spam status (S) of the message, we can simplify the expression:

P(F2 | F1, S) = P(F2 | S)

Substituting the simplified expression into the numerator, we have:

P(F1, F2 | S) = P(F2 | S) * P(F1 | S) * P(S)

Now we can rewrite the expression for P(S | F1, F2):

$P(S | F1, F2) = (P(F2 | S) * P(F1 | S) * P(S)) / P(F1, F2)$

This expression represents the probability of the message being a spam given that both spam filters flag it, using the conditional probabilities of the network.

Assuming that the second filter has the same sensitivity and false positive rate as the first one, we can say that $P(F2 | S) = P(F1 | S)$, as they have the same conditional probability.

Section III. Bayesian Networks

3. Consider the following Bayes net, with variables for the season (summer or winter) and binary variables indicating whether it is raining, whether the sprinkler is on, whether the pavement is wet, and whether the pavement is slippery. For the binary variables, only the independent parameters are supplied in the CPTs; the rest of the values can be inferred from these.
   (a) Using these CPTs and the Bayes net structure, compute the joint probability that it is summer ($S = summer$), the sprinkler is off ($P = false$), and the pavement is wet ($W = true$). Show your work. [5 points]

We want to compute P(S=summer, P=false, W=true).

Using the chain rule of probability, we have:

P(S=summer, P=false, W=true) = P(S=summer) * P(P=false | S=summer) * P(W=true | R, P=false) * P(R | S=summer) * P(L=true | W=true)

Let's substitute the values from the CPTs into the expression:

P(S=summer) = 0.5

P(P=false | S=summer) = 1 - P(P=true | S=summer) = 1 - 0.6 = 0.4

P(W=true | R=true, P=false) = 0.95

P(W=true | R=false, P=false) = 0.1

P(R=true | S=summer) = 0.1 P(L=true | W=true) = 0.7

Now we can compute the joint probability:

P(S=summer, P=false, W=true) = 0.5 * 0.4 * 0.95 * 0.1 * 0.7 ≈ 0.0133

   (b) Compute the conditional probability that the pavement is wet in summer if the sprinkler is off. Show your work. [5 points]

The conditional probability can be calculated as follows:

P(W=true | S=Summer, P=false) = P(W=true, S=Summer, P=false) / P(S=Summer, P=false)

we can plug the numerator from part a 0.0133

Next, let's calculate the denominator P(S=Summer, P=false):

P(S=Summer, P=false) = Σ P(W=true, S=Summer, P=false) + Σ P(W=false, S=Summer, P=false)

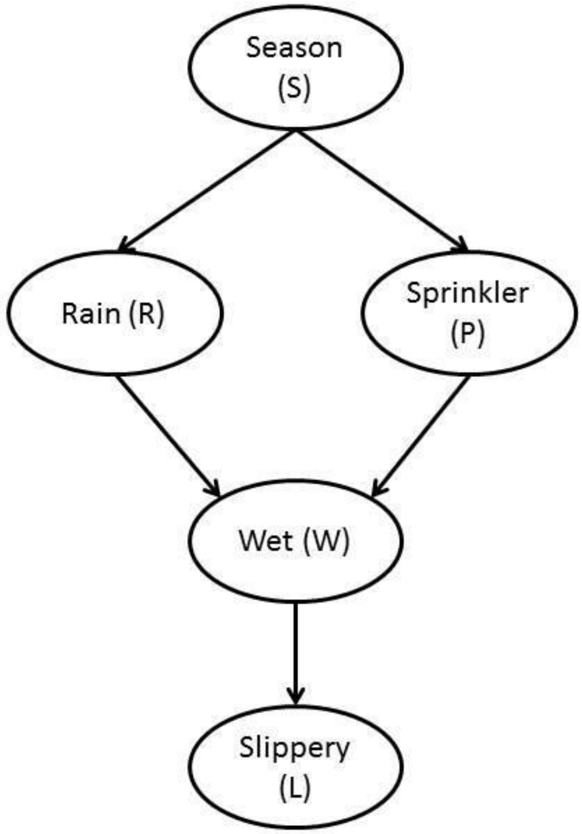
We need to consider all possible values of W (true/false):

P(W=true, S=Summer, P=false) = 0.75 * 0.5 * 1 = 0.0133 (same as the numerator)
P(W=false, S=Summer, P=false) = 0.1 * 0.5 * 1 = 0.05 (from the given CPTs)

Substituting these values, we get: P(S=Summer, P=false) = 0.0133 + 0.05 = 0.633

Finally, we can compute the conditional probability:

P(W=true | S=Summer, P=false) = P(W=true, S=Summer, P=false) / P(S=Summer, P=false) = 0.0133 / 0.633 ≈ 0.210



S	P(S)
summer	0.5
winter	0.5

R	P	W	P(W|R,P)
true	true	true	0.99
true	false	true	0.95
false	true	true	0.75
false	false	true	0.1

S	R	P(R|S)
winter	true	0.8
summer	true	0.1

S	P	P(P|S)
winter	true	0.05
summer	true	0.6

W	L	P(L|W)
true	true	0.7
false	true	0.2