COMP 307 Assignment 2

Yan Zichu

300476924

Part 1: Reasoning Under Uncertainty Basics

Question 1

1. Create the full joint probability table of X and Y , i.e. the table containing the following four joint probabilities P(X=0,Y =0),P(X=0,Y =1),P(X=1,Y =0),P(X=1,Y =1). Also explain which probability rules you used.

X	Y	P(X,Y)
0	0	0.0900
1	0	0.5600
0	1	0.2100
1	1	0.1400

I used product rule ($P(A,B) = P(B) \times P(A|B) = P(A)\times P(B|A)$) for calculating values in the table above:

$$P(X=0,Y=0) = P(X=0) \times P(Y=0|X=0) = 0.300 \times 0.300 = 0.0900$$
$$P(X=0,Y=1) = P(X=0) \times P(Y=1|X=0) = 0.300 \times 0.700 = 0.2100$$
$$P(X=1,Y=0) = P(X=1) \times P(Y=0|X=1) = 0.700 \times 0.800 = 0.5600$$
$$P(X=1,Y=1) = P(X=1) \times P(Y=1|X=1) = 0.700 \times 0.200 = 0.1400$$

2. If given $P(X = 1,Y = 0,Z = 0) = 0.336$, $P(X = 0,Y = 1,Z = 0) = 0.168$, $P(X = 0,Y = 0,Z = 1) = 0.036$, and $P(X = 0,Y = 1,Z = 1) = 0.042$, create the full joint probability table of the three variables X, Y , and Z. Also explain which probability rules you used.

X	Y	Z	P(X, Y, Z)
0	0	0	0.054
1	0	0	0.336
0	1	0	0.168
1	1	0	0.112
0	1	1	0.042
1	1	1	0.028
1	0	1	0.224
0	0	1	0.036

I used product rule ($P(A,B,C) = P(A) \times P(B|A) \times P(C|A,B)$ for calculating values in the table above:

$P(X=0,Y=0,Z=0) = P(X=0) \times P(Y=0 \mid X=0) \times P(Z=0|X=0,Y=0)$
$= P(X=0) \times P(Y=0 \mid X=0) \times P(Z=0|Y=0)$ (Because Z is independent from X given Y)
$= 0.300 \times 0.300 \times 0.600$
$= 0.054$

$P(X=1,Y=1,Z=0) = P(X=1) \times P(Y=1 \mid X=1) \times P(Z=0|X=1,Y=1)$
$= P(X=1) \times P(Y=1 \mid X=1) \times P(Z=0|Y=1)$ (Because Z is independent from X given Y)
$= 0.700 \times 0.200 \times 0.800$
$= 0.112$

P(X=1,Y=1,Z=1) = P(X=1) × P(Y=1 | X=1) × P(Z=1|X=1,Y=1)
= P(X=1) × P(Y=1 | X=1) × P(Z=1|Y=1)   (Because Z is independent from X given Y)
= 0.700 × 0.200 × 0.200
= 0.028

P(X=1,Y=0,Z=1) = P(X=1) × P(Y=0 | X=1) × P(Z=1|X=1,Y=0)
= P(X=1) × P(Y=0 | X=1) × P(Z=1|Y=0)   (Because Z is independent from X given Y)
= 0.700 × 0.800 × 0.400
= 0.225

3. From the above joint probability table of X, Y , and Z:
(i) calculate the probability of P(Z = 0) and P(X = 0,Z = 0),
(ii) judge whether X and Z are independent to each other and explain why.

(i) P(Z=0) = P(X=0,Y=0,Z=0) + P(X=1,Y=1,Z=0) + P(X=0,Y=1,Z=0) + P(X=1,Y=0,Z=0)
= 0.054 + 0.336 + 0.168 + 0.112
= 0.67

P(X = 0,Z = 0) = P(X=0,Y=0,Z=0) + P(X=0,Y=1,Z=0)
= 0.054 + 0.168
= 0.222

(ii) P(X = 0,Z = 0) = 0.222
P(X=0) × P(Z=0) = 0.3 × 0.67 = 0.201
P(X = 0,Z = 0) ≠ P(X=0) × P(Z=0)
So X and Z are not independent to each other.

4.From the above joint probability table of X, Y , and Z:
(i) calculate the probability of P (X = 1, Y = 0|Z = 1),
(ii) calculate the probability of P (X = 0|Y = 0, Z = 0).

(i) P(X = 1, Y = 0| Z = 1) = $\dfrac{P(X = 1,Y = 0,Z = 1)}{P(Z = 1)}$

$= \dfrac{0.224}{1 - 0.67}$

= 0.679 (3dp)

(ii) P(X = 0|Y = 0, Z = 0) = $\dfrac{P(X = 0,Y = 0 \,|\, Z = 0)}{P(Y = 0,Z = 0)}$

$= \dfrac{P(X = 0,Y = 0 \,|\, Z = 0)}{P(X = 0,Y = 0,Z = 0) + P(X = 1,Y = 0,Z = 0)}$

$= \dfrac{0.054}{0.39}$

= 0.138 (3dp)

Question 2

Consider three Boolean variables A, B, and C, $A \perp B|C$, P(B) = 0.7, P(C) = 0.4, P(A|B) = 0.3, P(A|C) = 0.5, and P(B|C) = 0.2, calculate the following probabilities. Show your working.

(i) $P(B,C) = P(C) \times P(B|C) = 0.4 \times 0.2 = 0.08$

(ii) $P(\neg A|B) = 1 - P(A|B) = 1 - 0.3 = 0.7$

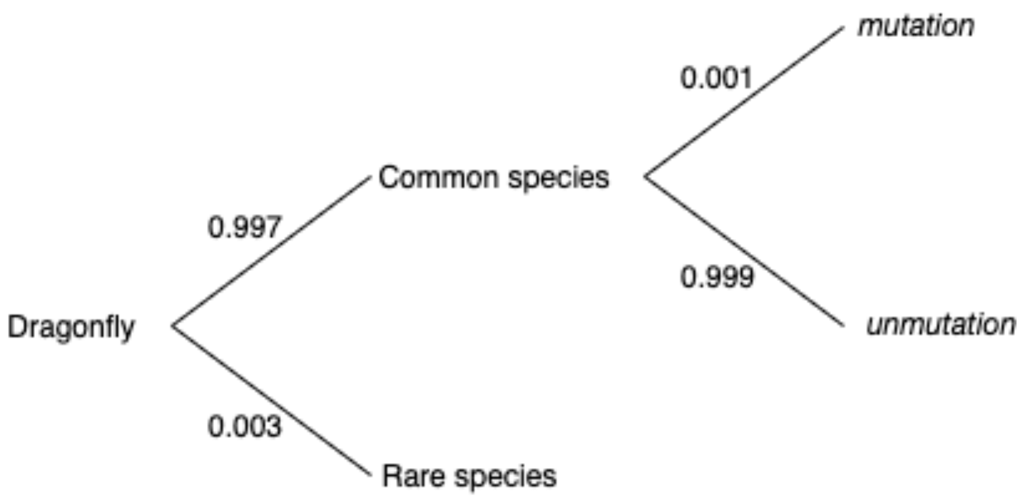
(iii) $P(A,B|C) = P(A|C) \times P(B|C)$ (Because $A \perp B|C$)
$= 0.5 \times 0.2$
$= 0.1$

(iv) $P(A|B,C) = P(A|C)$ (Because $A \perp B|C$) $= 0.5$

(v) $P(A,B,C) = P(A,B) \times P(C|B)$
$= P(B) \times P(A|B) \times (P(C) \times P(B|C) \div P(B))$
$= 0.7 \times 0.3 \times (0.4 \times 0.2 \div 0.7)$
$= 0.024$

Question 3

Dragonfly has a rare species, which always has an extra set of wings. However, common dragonflies can sometimes mutate and get an extra set of wings. A dragonfly either belongs to the common species or the rare species with the extra wings. There are 0.3% dragonflies belonging to the rare species with the extra set of wings. For the common dragonflies, the probability of the extra-wing mutation is 0.1%. Now you see a dragonfly with an extra pair of wings. What is the probability that it belongs to the rare species? Show your working.



Extra Wing = E

$P(E) = P(R) + P(M,C) = 0.003 + 0.997 \times 0.001 = 0.003997$

$$P(R|E) = \frac{P(E|R) \times P(R)}{P(E)} = \frac{1 \times 0.003}{0.003997} = 0.75056292$$

So the probability of a dragonfly with an extra pair of wings and belongs to the rare species is 0.75056292.

Part 2: Naive Bayes Method

i.the probabilities P(Fi|c) for each feature

1. 1|0 0.46979865771812085
   0|0 0.6644295302013422
   1|1 0.34
   0|1 0.20000000000000004
   ----
2. 0|0 0.4161073825503356
   1|0 0.3355704697986577
   0|1 0.34
   1|1 0.8000000000000002
   ----
3. 1|0 0.3355704697986577
   0|0 0.7114093959731543
   1|1 0.8000000000000002
   0|1 0.34
   ----
4. 0|0 0.6644295302013422
   1|0 0.24161073825503357
   0|1 0.20000000000000004
   1|1 0.34
   ----
5. 1|0 0.3355704697986577
   0|0 0.7114093959731543
   1|1 0.8000000000000002
   0|1 0.34
   ----
6. 1|0 0.3355704697986577
   0|0 0.7583892617449665
   1|1 0.8000000000000002
   0|1 0.66
   ----
7. 0|0 0.6644295302013422
   1|0 0.348993288590604
   0|1 0.20000000000000004
   1|1 0.78
   ----
8. 0|0 0.7114093959731543
   1|0 0.3355704697986577
   0|1 0.34
   1|1 0.8000000000000002
   ----
9. 1|0 0.3355704697986577
   0|0 0.4161073825503356
   1|1 0.8000000000000002
   0|1 0.34
   ----
10. 1|0 0.5838926174496645
    0|0 0.6644295302013422
    1|1 0.66
    0|1 0.20000000000000004

ii.For each instance in the unlabelled set, given the input vector F, the probability P(S|D), the probability P(S|D), and the predicted class of the input vector. Here D is an email represented by F , S refers to class spam and S refers to class non-spam.

1. {0: 0.0013853888496881347, 1: 8.59552804020102e-09}
   [1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0] belong to 0
   ----
2. {0: 5.5496824542951195e-06, 1: 0.00010174904744419714}
   [0 0 1 1 0 0 1 1 1 0 0 1] belong to 1
   ----
3. {0: 3.0837663619820645e-05, 1: 0.0005633165256426141}
   [1 1 1 1 1 0 1 0 0 0 1 1] belong to 1
   ----
4. {0: 0.00026651534839374 89, 1: 5.05619296482413e-09}
   [0 1 0 0 0 0 1 0 1 0 0 0] belong to 0
   ----
5. {0: 3.0837663619820645e-05, 1: 0.0005633165256426141}
   [1 1 1 0 1 1 0 1 0 0 1 1] belong to 1
   ----
6. {0: 1.7622330440013875e-05, 1: 0.009695241188213083}
   [1 1 1 1 1 1 0 0 0 1 1 1] belong to 1
   ----
7. {0: 0.0008032019484513938, 1: 6.104789547738702e-08}
   [0 0 0 0 1 1 0 1 0 0 0 0] belong to 0
   ----
8. {0: 6.53758468740198e-05, 1: 0.0002394095233981109 6}
   [0 1 0 1 1 1 1 0 0 0 1 1] belong to 1
   ----
9. {0: 3.6093148115863145e-06, 1: 0.000563316525642614}
   [1 1 1 1 1 0 1 0 0 1 0 1] belong to 1
   ----
10. {0: 0.002904962940479658, 1: 4.0276022544723685e-07}
    [1 1 0 0 0 1 0 1 0 0 1 0] belong to 0

iii.The derivation of the Naive Bayes algorithm assumes that the attributes are conditionally independent. Why is this like to be an invalid assumption for the spam data? Discuss the possible effect of two attributes not being independent.

We assume that the attributes are conditionally independent to each other in Naive Bayes algorithm, but in real life, attributes normally affect each other. Take this case as an example, when "MILLION DOLLARS" and some phrases related to fraud appear in the email, usually from an invalid reply-to address. This shows that the features are not independent of each other

Part 3: Bayesian Network : Questions

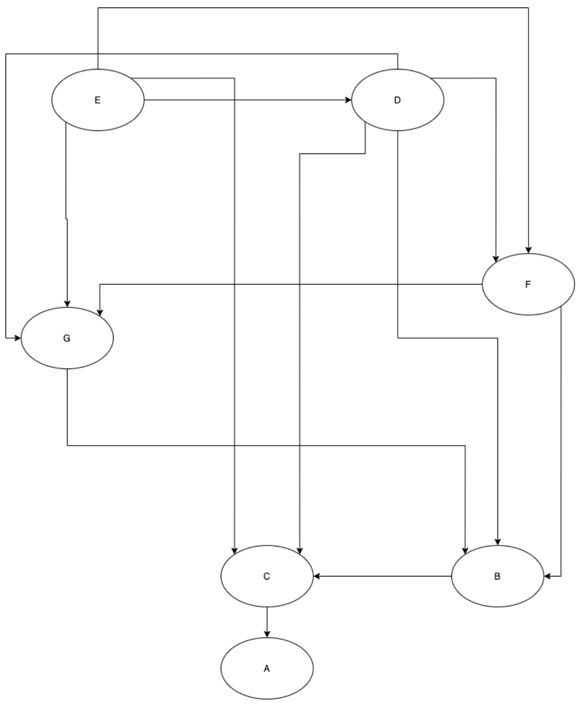
1.Write the factorisation of this Bayesian network.

P(A,B,C,D,F,E,G) = P(A) × P(B) × P(C|A) × P(F|B) × P(D|B,C) × P(E|C) × P(G|E,D,F)

2.Describe under which condition the following pair of nodes will be (conditionally) independent. For example, A and E are conditionally independent given C.

i. A and B are independent with each other.
ii. A and D are conditionally independent given C.
iii. B and G are conditionally independent given D and F.
iv. D and F are conditionally independent given B.
v. C and G are conditionally independent given D and E.

3.Draw the Bayesian network if the node ordering is E → D → F → G → B → C → A.



Step1: add node E
Step2: add node D
P(F|D,E) = P(F)? No
P(F|D,E) = P(F|D)? No
P(F|D,E) = P(F|E)? No,E →F,D →F
Step3: add node G
P(G|F,D,E) = P(G)? Yes, F→G,D→G,E→G
Step4: add node B
P(B|G,D,E,F) = P(B)? No
P(B|G,D,E,F) = P(B|G)? No
…
P(B|G,D,E,F) = P(B|G,D,E)? No
P(B|G,D,E,F) = P(B|G,D,F)?
Yes,G→B,D→B,F→B
Step5: add node C
P(C|B,G,D,E,F) = P(C)? No
P(C|B,G,D,E,F) = P(C|B)? No
…
P(C|B,G,D,E,F) = P(C|B,D,F)? No
P(C|B,G,D,E,F) = P(C|B,D,E)?
Yes,B→C,D→C,E→C
Step6: add node A
P(A|B,G,D,E,F,C) = P(A)? No
P(A|B,G,D,E,F,C) = P(A|C)? Yes,C→A

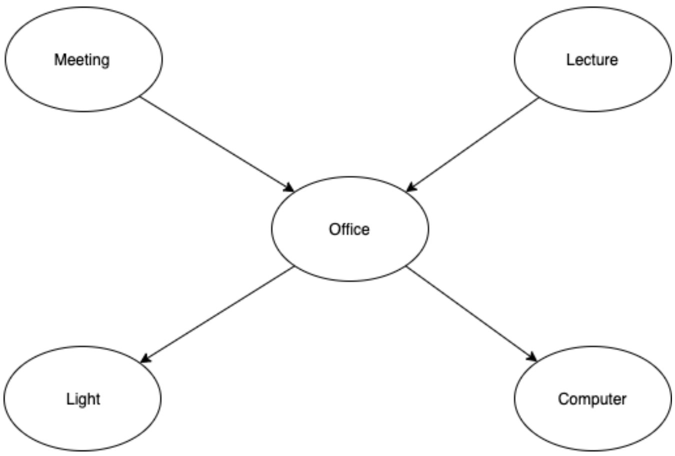
4.Suppose P(A) = 0.7, P(B) = 0.2, P(C|A) = 0.6, P(C|¬A) = 0.3, P (D|B, C) = 0.7, P (D|B, ¬C) = 0.6, P (D|¬B, C) = 0.5, P (D|¬B, ¬C) = 0.2. Calculate the probability P(D).

$P(D) = P(D,B,C) + P(D,\neg B,C) + P(D,B,\neg C) + P(D,\neg B,\neg C)$
$= P(B) \times P(C) \times P(D|B,C) + P(\neg B) \times P(C) \times P(D|\neg B,C) + P(B) \times P(C\neg) \times P(D|B,\neg C) +$
$P(\neg B) \times P(\neg C) \times P(D|\neg B,\neg C)$ — (formula A)
In formula A, P(C) is unknown so:
$P(C) = P(C,A) + P(C,\neg A)$
$= P(A) \times P(C|A) + P(\neg A) \times P(C|\neg A)$
$= 0.7 \times 0.6 + (1 - 0.7) \times 0.3$
$= 0.51$
Then we can continue formula A:
$P(D) =$
$P(B) \times P(C) \times P(D|B,C) + P(\neg B) \times P(C) \times P(D|\neg B,C) + P(B) \times P(C\neg) \times P(D|B,\neg C) +$
$P(\neg B) \times P(\neg C) \times P(D|\neg B,\neg C)$ — (formula A)
$= 0.2 \times 0.51 \times 0.7 + (1 - 0.2) \times 0.51 \times 0.5 + 0.2 \times (1 - 0.51) \times 0.6 + (1 - 0.2) + (1 - 0.51) \times 0.2$
$= 0.0714 + 0.204 + 0.0588 + 0.0784$
$= 0.4126$

Part 4: Building Bayesian Network

1. Construct a Bayesian network to represent the above scenario.

M	P(M)
+M	0.7
-M	0.3



L	P(L)
+L	0.6
-L	0.4

LT	O	P(LT|O)
+LT	-O	0.98
-LT	+O	0.5
+LT	-O	0.02
-LT	+O	0.5

C	O	P(C|O)
+C	-O	0.8
-C	+O	0.2
+C	-O	0.2
-C	+O	0.8

O	M	L	P(O|M,L)
-O	-M	-L	0.94
-O	-M	+L	0.2
-O	+M	-L	0.25
-O	+M	+L	0.05
+O	-M	-L	0.06
+O	-M	+L	0.8
+O	+M	-L	0.75
+O	+M	+L	0.95

2. Calculate how many free parameters in your Bayesian network ?

$$2 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 = 10 \textit{ free parameters}$$

3. What is the joint probability that Rachel has lectures, has no meetings, she is in her office and logged on her computer but with lights off.

P(L=+L,M=-M,O=+O,C=+C,LT=-LT)
=P(O=+O|M=-M,L=+L) × P(M=-M) × P(L=+L) × P(LT=-LT|O=+O) × P(C=+C|O=+O)
=0.6 × 0.3 × 0.8 × 0.8 × 0.5 = 0.0576

4. Calculate the probability that Rachel is in the office.

Sum rule :
P(O= +O) =P(O,M=+M,L=+L) + P(O,M=+M,L=-L)+ P(O,M=-M,L=+L)+ P(O,M=-M,L=-L)

Product rule :
P(O=+O)= P(O|M=+M,L=+L) × P(M=+M,L=+L)+ P(O,M=+M,L=-L) × P(M=+M,L=-L)
+P(O,M=-M,L=+L) × P(M==M,L=+L)+ P(O,M=-M,L=-L) × P(M= -M, L = -L)
= 0.95 × 0.7 × 0.6 + 0.75 × 0.7 × 0.4 + 0.8 × 0.3 × 0.6 + 0.06 × 0.3 × 0.4
= 0.7602

5*. If Rachel is in the office, what is the probability that she is logged on, but her light is off.*

P(L=-F,C=-C|O=+O) = P(L|O) ×P(C|O) = 0.8 × 0.5 = 0.4

6*. Suppose a student checks Rachel's login status and sees that she is logged on. What effect does this have on the student's belief that Rachel's light is on ?*

It is a common cause effect, so there is no effect. If a common cause is known, the effect becomes independent. Since we know the probability of Rachel in the office, the probability of turning on the lights and the probability of the computer become independent.