COMP307

Assignment 3

Uncertainty and Probability

William Kilty

300473541

Part 1:

1. Question 1:
    a. Product rule:

$P(A, B) = P(B|A) * P(A)$

$\therefore P(X, Y, Z) = P(Z|X, Y) * P(X, Y)$

$P(X, Y)$ is as shown in Q1.b.

As $X \perp Z|Y$, we use rule of conditional independence:

$P(A|B, C) = P(A|C) \therefore P(Z|X, Y) = P(Z|Y)$

$P(Z|Y)$ is given to us. So now we combine;

X	Y	Z	$P(X, Y, Z)$
0	0	0	$P(Z = 0\|Y = 0) * P(X = 0, Y = 0)$ $= 0.70 * 0.035$ $= 0.0245$
0	0	1	$P(Z = 1\|Y = 0) * P(X = 0, Y = 0)$ $= 0.30 * 0.035$ $= 0.0105$
0	1	0	$P(Z = 0\|Y = 1) * P(X = 0, Y = 1)$ $= 0.20 * 0.315$ $= 0.063$
0	1	1	$P(Z = 1\|Y = 1) * P(X = 0, Y = 1)$ $= 0.80 * 0.315$ $= 0.252$
1	0	0	$P(Z = 0\|Y = 0) * P(X = 1, Y = 0)$ $= 0.70 * 0.39$ $= 0.273$
1	0	1	$P(Z = 1\|Y = 0) * P(X = 1, Y = 0)$ $= 0.30 * 0.39$ $= 0.117$
1	1	0	$P(Z = 0\|Y = 1) * P(X = 1, Y = 1)$ $= 0.20 * 0.26$ $= 0.052$
1	1	1	$P(Z = 1\|Y = 1) * P(X = 1, Y = 1)$ $= 0.80 * 0.26$ $= 0.208$

b. Product rule:

$P(X,Y) = P(X) * P(Y|X)$, so;

X	Y	$P(X,Y)$
0	0	$P(X=0) * P(Y=0\|X=0)$ $= 0.35 * 0.10$ $= 0.035$
0	1	$P(X=0) * P(Y=1\|X=0)$ $= 0.35 * 0.90$ $= 0.315$
1	0	$P(X=1) * P(Y=0\|X=1)$ $= 0.65 * 0.60$ $= 0.39$
1	1	$P(X=1) * P(Y=1\|X=1)$ $= 0.65 * 0.40$ $= 0.26$

c.

i. Sum rule:

$$P(Z = 0) = \sum_{x\in\{1,0\},y\in\{1,0\}} P(X = x, Y = y, Z = 0)$$
$$= 0.0245 + 0.063 + 0.273 + 0.052$$
$$= 0.4125$$

ii. Sum rule:

$$P(X = 0, Z = 0) = \sum_{y\in\{1,0\}} P(X = 0, Y = y, Z = 0)$$
$$= 0.0245 + 0.063$$
$$= 0.0875$$

iii. reverse product rule:

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

Normalisation rule:

$$P(Z = 1) = 1 - P(Z = 0) = 1 - 0.4125$$
$$= 0.5875$$

$$= \frac{0.117}{0.5875}$$
$$= 0.1991(4dp)$$

iv. reverse product rule:

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

Sum rule:

$$P(Y = 0, Z = 0) = \sum_{x\in\{1,0\}} P(X = x, Y = 0, Z = 0)$$
$$= 0.0245 + 0.273$$
$$= 0.2975$$

$$= \frac{0.0245}{0.2975}$$
$$= 0.0824(4dp)$$

2. Question 2:
    a. Product rule:

$$P(B = t, C = t) = P(C = t) * P(B = t|C = t)$$

$$= 0.4 * 0.2$$

$$= 0.08$$

    b. Normalisation rule:

$$P(A = f|B = t) = 1 - P(A = t|B = t)$$

$$= 1 - 0.3$$

$$= 0.7$$

    c. Conditional independence rule:

$$P(A = t, B = t|C = t) = P(A = t|C = t) * P(B = t|C = t)$$

$$= 0.5 * 0.2$$

$$= 0.1$$

    d. Conditional independence rule:

$$P(A = t|B = t, C = t) = P(A = t|C = t)$$

$$= 0.5$$

    e. Product rule:

$$P(A = t, B = t, C = t) = P(A = t|B = t, C = t) * P(B = t, C = t)$$

$$= Q2.d * Q2.a$$

$$= 0.5 * 0.08$$

$$= 0.04$$

Part 2:

1.	no-recurrence-events	recurrence-events
P(Class)	0.7063197026	0.2936802974
P(age = 10-19 | Class)	0.005050505051	0.01149425287
P(age = 20-29 | Class)	0.0101010101	0.01149425287
P(age = 30-39 | Class)	0.1111111111	0.183908046
P(age = 40-49 | Class)	0.3131313131	0.3103448276
P(age = 50-59 | Class)	0.3282828283	0.2528735632
P(age = 60-69 | Class)	0.1919191919	0.1954022989
P(age = 70-79 | Class)	0.0303030303	0.01149425287
P(age = 80-89 | Class)	0.005050505051	0.01149425287
P(age = 90-99 | Class)	0.005050505051	0.01149425287
P(menopause = lt40 | Class)	0.03125	0.01234567901
P(menopause = ge40 | Class)	0.4583333333	0.3827160494
P(menopause = premeno | Class)	0.5104166667	0.6049382716
P(tumor-size = 0-4 | Class)	0.03980099502	0.02222222222
P(tumor-size = 5-9 | Class)	0.02487562189	0.01111111111
P(tumor-size = 10-14 | Class)	0.1293532338	0.02222222222
P(tumor-size = 15-19 | Class)	0.1144278607	0.07777777778
P(tumor-size = 20-24 | Class)	0.1741293532	0.1555555556
P(tumor-size = 25-29 | Class)	0.1592039801	0.2111111111
P(tumor-size = 30-34 | Class)	0.1691542289	0.2555555556
P(tumor-size = 35-39 | Class)	0.05970149254	0.08888888889
P(tumor-size = 40-44 | Class)	0.08457711443	0.07777777778
P(tumor-size = 45-49 | Class)	0.01492537313	0.02222222222
P(tumor-size = 50-54 | Class)	0.02487562189	0.04444444444
P(tumor-size = 55-59 | Class)	0.004975124378	0.01111111111
P(inv-nodes = 0-2 | Class)	0.797029703	0.4725274725
P(inv-nodes = 3-5 | Class)	0.08415841584	0.1758241758
P(inv-nodes = 6-8 | Class)	0.0396039604	0.1208791209
P(inv-nodes = 9-11 | Class)	0.01485148515	0.06593406593
P(inv-nodes = 12-14 | Class)	0.009900990099	0.03296703297
P(inv-nodes = 15-17 | Class)	0.0198019802	0.04395604396
P(inv-nodes = 18-20 | Class)	0.00495049505	0.01098901099
P(inv-nodes = 21-23 | Class)	0.00495049505	0.01098901099
P(inv-nodes = 24-26 | Class)	0.00495049505	0.02197802198
P(inv-nodes = 27-29 | Class)	0.00495049505	0.01098901099
P(inv-nodes = 30-32 | Class)	0.00495049505	0.01098901099
P(inv-nodes = 33-35 | Class)	0.00495049505	0.01098901099
P(inv-nodes = 36-39 | Class)	0.00495049505	0.01098901099
P(node-caps = yes | Class)	0.1256544503	0.4
P(node-caps = no | Class)	0.8743455497	0.6
P(deg-malig = 1 | Class)	0.2916666667	0.1111111111
P(deg-malig = 2 | Class)	0.5104166667	0.3580246914
P(deg-malig = 3 | Class)	0.1979166667	0.5308641975
P(breast = left | Class)	0.5078534031	0.55
P(breast = right | Class)	0.4921465969	0.45
P(breast-quad = left_up | Class)	0.3453608247	0.3012048193
P(breast-quad = left_low | Class)	0.3659793814	0.3855421687
P(breast-quad = right_up | Class)	0.1082474227	0.1686746988
P(breast-quad = right_low | Class)	0.09278350515	0.0843373494
P(breast-quad = central | Class)	0.08762886598	0.06024096386
P(irradiat = yes | Class)	0.1570680628	0.3875
P(irradiat = no | Class)	0.8429319372	0.6125

2.

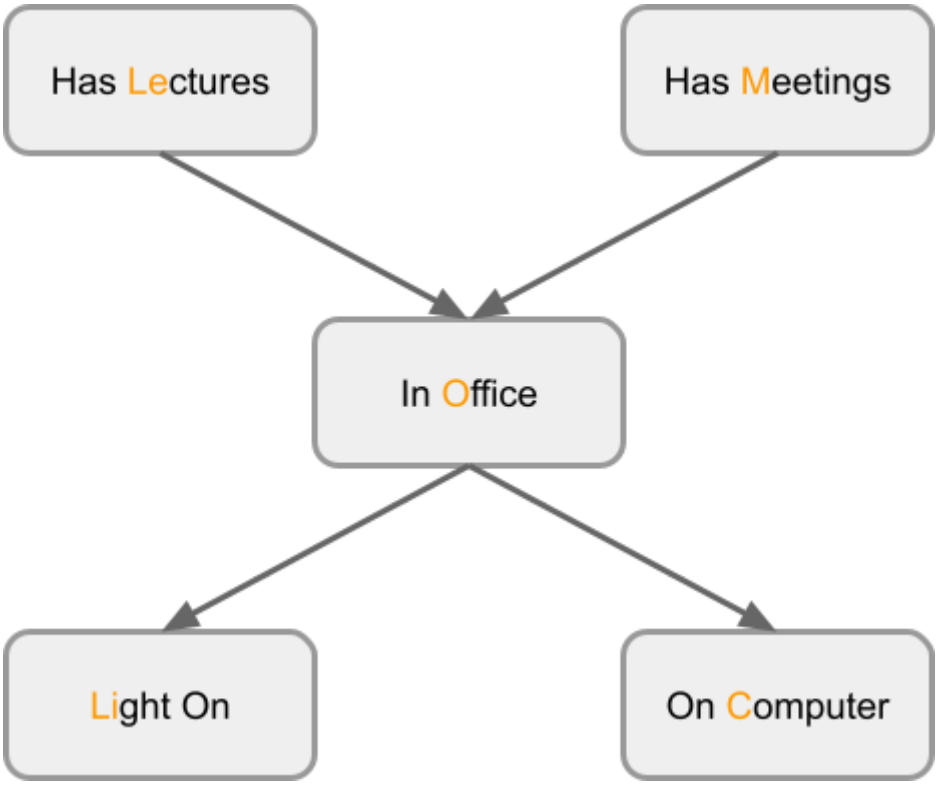
	no-recurrence-events	recurrence-events
P(Class)	0.7063197026	0.2936802974

3.

Instance 0:
Score for no-recurrence-events: 4.017731924138001E-6
Score for recurrence-events: 7.096642912782161E-6
Chosen class: recurrence-events. Incorrect...
Instance 1:
Score for no-recurrence-events: 3.3365875538186835E-4
Score for recurrence-events: 2.6021024013534593E-5
Chosen class: no-recurrence-events. Correct!
Instance 2:
Score for no-recurrence-events: 4.707378797592419E-5
Score for recurrence-events: 9.715041077797865E-7
Chosen class: no-recurrence-events. Correct!
Instance 3:
Score for no-recurrence-events: 1.5163354504655905E-4
Score for recurrence-events: 1.0219165794406317E-5
Chosen class: no-recurrence-events. Correct!
Instance 4:
Score for no-recurrence-events: 4.34523614378287E-6
Score for recurrence-events: 1.9205993914751727E-6
Chosen class: no-recurrence-events. Correct!
Instance 5:
Score for no-recurrence-events: 6.000123072810078E-4
Score for recurrence-events: 3.880403296274196E-5
Chosen class: no-recurrence-events. Correct!
Instance 6:
Score for no-recurrence-events: 2.071279390811688E-4
Score for recurrence-events: 7.415569159796725E-5
Chosen class: no-recurrence-events. Correct!
Instance 7:
Score for no-recurrence-events: 3.1504865826699285E-4
Score for recurrence-events: 8.934762413310772E-6
Chosen class: no-recurrence-events. Incorrect...
Instance 8:
Score for no-recurrence-events: 3.9123945169671666E-5
Score for recurrence-events: 6.483884504037111E-5
Chosen class: recurrence-events. Correct!
Instance 9:
Score for no-recurrence-events: 4.1017039290784815E-5
Score for recurrence-events: 5.283165151437645E-5
Chosen class: recurrence-events. Correct!
Accuracy: 80.00%

Part 3:

1. Domain Variables: Meetings, Lectures, Office, Light, Computer



L	P(L)
T	0.6

M	P(M)
T	0.7

L	M	O	P(O|L,M)
T	T	T	0.95
T	F	T	0.80
F	T	T	0.75
F	F	T	0.06

O	Li	P(Li|O)
T	T	0.50
F	T	0.02

O	C	P(C|O)
T	T	0.80
F	T	0.20

2. The number of free parameters is:
   $1 + 1 + 4 + 2 + 2 = 10$

3. Product rule:
   $P(L = t, M = f, O = t, C = t, Li = f)$
   $= P(L = t, M = f, O = t, C = t) * P(Li = f | L = t, M = f, O = t, C = t)$
   Conditional Independence rule:
   $P(Li = f | L = t, M = f, O = t, C = t) = P(Li = f | O = t)$
   $= P(L = t, M = f, O = t) * P(C = t | L = t, M = f, O = t) * P(Li = f | O = t)$
   Conditional Independence rule:
   $P(C = t | L = t, M = f, O = t) = P(C = t | O = t)$
   $= P(L = t, M = f) * P(O = t | L = t, M = f) * P(C = t | O = t) * P(Li = f | O = t)$
   $= P(L = t) * P(M = f | L = t) * P(O = t | L = t, M = f) * P(C = t | O = t) * P(Li = f | O = t)$
   Independence rule:
   $P(M = f | L = t) = P(M = f)$
   $= P(L = t) * P(M = f) * P(O = t | L = t, M = f) * P(C = t | O = t) * P(Li = f | O = t)$
   $= 0.6 * P(M = f) * 0.8 * 0.8 * P(Li = f | O = t)$
   Normalisation rule:
   $P(M = f) = 1 - P(M = t)$
   $= 1 - 0.7 = 0.3$
   Normalisation rule:
   $P(Li = f | O = t) = 1 - P(Li = t | O = t)$
   $= 1 - 0.8 = 0.2$
   $= 0.6 * 0.3 * 0.8 * 0.8 * 0.2$
   $= 0.02304$

4. Sum rule:
   $P(O = t) = P(O = t, L = t, M = t) + P(O = t, L = t, M = f)$
   $+ P(O = t, L = f, M = t) + P(O = t, L = f, M = f)$
   Product rule:
   $P(O, L, M) = P(L, M) * P(O | L, M)$
   Independence rule:
   $P(L, M) = P(L) * P(M)$
   $= P(L = t) * P(M = t) * P(O = t | L = t, M = t) + P(L = t) * P(M = f) * P(O = t | L = t, M = f)$
   $+ P(L = f) * P(M = t) * P(O = t | L = f, M = t) + P(L = f) * P(M = f) * P(O = t | L = f, M = f)$
   $= 0.6 * 0.7 * 0.95 + 0.6 * 0.3 * 0.8 + 0.4 * 0.7 * 0.75 + 0.4 * 0.3 * 0.06$
   $= 0.399 + 0.144 + 0.21 + 0.0072$
   $= 0.7602$

5. Conditional Independence rule:
   $P(C = t, Li = f | O = t) = P(C = t | O = t) * P(Li = f | O = t)$
   $= 0.8 * 0.5$
   $= 0.4$

Part 4:

1. Evidence: $\{X = t\ (x)\}$
   Hidden: $\{S,\ C,\ D\}$
   Query: $\{P\}$

2. Product rule:
   $P(P|x) = \alpha * P(P, x)$ (Product rule)
   $= \alpha * \Sigma_{C,S,D} P(P, C, S, D, x)$ (Sum rule)
   $= \alpha * \Sigma_{C,S,D} P(P) * P(S) * P(C|P, S) * P(x|C) * P(D|C)$ (Independence rules)
   Variable elimination:
   $f_1(P) =$

P	$P(P)$
t	0.9
f	0.1

   $f_2(S) =$

S	$P(S)$
t	0.3
f	0.7

   $f_3(C, P, S) =$

S	**P**	**C**	$P(C\|P, S)$
t	t	t	0.05
f	t	t	0.02
t	f	t	0.03
f	f	t	0.001
t	t	f	0.95
f	t	f	0.98
t	f	f	0.97
f	f	f	0.999

$f_4(C) =$

C	$P(x\|C)$
t	0.9
f	0.2

$f_5(D, C) =$

C	**D**	$P(D\|C)$
t	t	0.65
f	t	0.30
t	f	0.35
f	f	0.70

$\Sigma_{C,S,D} f_1(P) \otimes f_2(S) \otimes f_3(C, P, S) \otimes f_4(C) \otimes f_5(D, C)$

Order of hidden variables, D, C, S

Join factors containing D, and eliminate D:

$f_6(C) = \Sigma_D f_5(D, C)$

C	$\Sigma_D P(D\|C)$
t	1
f	1

$\Sigma_{C,S} f_1(P) \otimes f_2(S) \otimes f_3(C, P, S) \otimes f_4(C) \otimes f_6(C)$

Join factors containing C, and eliminate C:

$f_7(P, S) = \Sigma_C f_3(C, P, S) \otimes f_4(C) \otimes f_6(C)$

S	**P**	$\Sigma_C P(C\|P, S) * P(x\|C) * \Sigma_D P(D\|C)$
t	t	$(0.05 * 0.9 * 1) + (0.95 * 0.2 * 1) = 0.235$
f	t	$(0.02 * 0.9 * 1) + (0.98 * 0.2 * 1) = 0.214$
t	f	$(0.03 * 0.9 * 1) + (0.97 * 0.2 * 1) = 0.221$
f	f	$(0.001 * 0.9 * 1) + (0.999 * 0.2 * 1) = 0.2007$

$\Sigma_S f_1(P) \otimes f_2(S) \otimes f_7(P, S)$

Join factors containing S, and eliminate S:

$f_8(P) = \Sigma_S f_2(S) \otimes f_7(P, S)$

P	$\Sigma_S P(S) * \Sigma_C P(C\|P,S) * P(x\|C) * \Sigma_D P(D\|C)$
t	$(0.3 * 0.235) + (0.7 * 0.214) = 0.2203$
f	$(0.3 * 0.221) + (0.7 * 0.2007) = 0.20679$

$f_1(P) \otimes f_8(P)$

Now, we can create the final table

$f_9(P) = f_1(P) \otimes f_8(P)$

P	$P(P) * \Sigma_S P(S) * \Sigma_C P(C\|P,S) * P(x\|C) * \Sigma_D P(D\|C)$
t	$0.9 * 0.2203 = 0.19827$
f	$0.1 * 0.20679 = 0.020679$

3. $P(P|x) = \alpha * f_9(P)$

   $P(P = t|x) = \alpha * 0.19827$

   $= 0.19827/(0.19827 + 0.020679)$

   $= 0.905553348$

   $= 0.9056\ (4dp)$