**CSci 384: Artificial Intelligence Spring, 2017**

**Instructor: Dr. M. E. Kim** **Date: April 20, 2017**

**Due: 5:00 PM, April 28th (Fri.), 2017**

**Home Assignment 4: Uncertainty & Probabilistic Reasoning (124/ 150 )**

**Instruction:**

For each question, show the computational steps precisely.

1. Define the variables: e.g.) A: Alarm, F: Fire, etc.
2. Define what needs to be computed: e.g.) P(A|F)
3. Derive the formulas for computation, step by step:

e.g.) P(A|F) = P(F| A)⋅P(A)

1. Assign the values to the derived formula to complete the computation.

e.g.) P(A|F) = P(F|A)⋅P(A) = 0.2 ⋅ 0.6 = 0.12

Save your assignment file as ‘HW4-YourLastNameOnly’, then upload it to submission section.: e.g.) HW4-Kim.docx

**Q1. [20]** There are nine marbles in an urn, three are blue, three pink, and three green. Two marbles are drawn randomly from the urn at the same time in your eyes closed.

1. [10] What is the probability that both marbles are pink?

3 choose 2 \* 3 choose 0 \* 3 choose 0 / 9 choose 2 = 0.0833333333

or 1/3 \* ¼ = 1/12 = 0.0833

divide by 9 choose 2 because there are 9 total and your selecting 2 from the 9.

The first time getting a marble there is a 1/3 change that you get a pink one. With both marbles the probability drop because there are now fewer pinks. 0.0833333333

1. [10] Suppose you know that both marbles drawn are the same color. What is the probability that both marbles are pink?

0.33 or 1/3 because there is only three colors blue, pink and green in the urn.

Correct, but you have to formulate it.

If you formulate it, what probability do you have to compute and how to compute them?

i.e. P(Pink | Same Color) = P(Pink & Same) / P(Same)

where P(Same) = P(Pink ∧ Pink) + P(Blue ∧ Blue) + P(Green ∧ Green) = 1/12 \* 3 = ¼,

P(Pink ∧ Same) = P(Pink ∧ Pink) = 1/12

So, P(Pink | SameColor) = (1/12) / (1/4) = 1/3

**Q2. [10]** A physical exam is administered to all new inmates at a prison. Suppose that 80% of all healthy individuals pass this exam, 60% of all individuals with minor ailments pass, and 30% of all prisoners with serious ailments also pass. Suppose that 25% of these new prisoners are actually in good health (event **E1**), 50% have minor ailments (**E2**), and 25% have major health issues (**E3**). Given that an inmate passes this physical (event **B**), what is the *posterior probability that the inmate is in good health*?

Compute it using Bayes’ Theorem, **not** by the enumeration of Full Joint Probability Distribution.

80% healthy pass the exam

60% minor ailment pass

30% serious ailment also pass

h = healthy, m – minor ailments, s – serious ailments, p – physical exam , y – passed

p(y|h ) - 0.8, p(y|m) - 0.6, p(y|s) - 0.3, p(g)- 0.25, p(m) - 0.5, p(S) = 0.25

p(h|y) - (p(y|h) \*( p\*h)) / p(y) = (0.8\*0.25)/(0.575 = 0.348

**Q3. [25]** Consider two medical tests, A and B, for a virus.

Test A (**TA**) is 95% effective at recognizing the virus (**V**) when it is present, but has a 10% false positive rate (indicating that the virus is present, when it is not). Test B (**TB**) is 90% effective at recognizing the virus, but has a 5% false positive rate. The two tests use independent methods of identifying the virus.

The virus is carried by 1 % of all people. Say that a person is tested for the virus using only one of the tests, and that test comes back positive for carrying the virus. Query: Which test returning positive is more indicative of someone really carrying the virus?

1. [5] Define the variables and their domain that are required in this scenario.
   * 1. V = virus ∈ {present, absent}
     2. TA = Test A ∈ {positive, negative}
     3. TB = Test B ∈ {positive, negative}
2. [5] Using the defined variables in 1), define the formula(s) of probability that you need to compute to answer the query. e.g.) P(TA|TB)
   * 1. P(V|TA), P(V|TB)

- P(TA=+ | V=true) = 95% - .95

- P(TA=+ | V=false) = 10% - .1

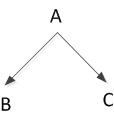
- P(TB=+ | V=true) = 90% - .9

- P(TB=+ | V=false) = 5% - .05

- P(V = true) = 1% - .01

1. [10] Compute the probability in 2) using ***Bayes’ Theorem***, **not** by the enumeration of Full Joint Probability Distribution.
   * 1. P(V|TA) = P(TA|V)P(V)/P(TA) where P(TA) = P(TA & V) +P(TA & ~V) = .087
     2. P(V|TB) = P(TB|V)P(V)/P(TB) where P(TB) = P(TB & V) +P(TB & ~V) = .154
2. [5] What is the answer to query? i.e. which test returning positive is more indicative of someone really carrying the virus? Justify your answer based on your computational result in 3).
   1. Test B(TB) is indicative – in question three shows that the rate is higher

**Q4.[0/ 10]** In the given Bayes net where P(*a*) = .5, P(*b*|*a*) = .2, P(*b*|*¬a*) = .2, P(*c*|*a*) = .8, P(*c*|*¬a*) =.4, compute the distribution of P(B | *c*).



P(a) = 0.5

P(b|a) = 0.2

P(b| øa) = 0.2

P(c|a) = 0.8

P( c| ø a) = 0.4

P(A,B,C) = P(b|a) P(c|a) P(a)

P(A, B, C) = ~~0.28~~(0.8)(0.5)

P(A, B, C) = 0.08

P(B|c) = P(b|øa) ~~P(c | øa)~~ P(~a | c) + P(b|a)P(a|c)

= P(b|~a)P(c|~a)P(~a) + P(b|a)P(c|a)P(a)

= .2 \* .4\* .5 + .2 \* .8 \*.5 = .04 + .08 = .12 when B=true

Similarly, compute P(~b|c) = .48.

Then, normalize them. <.12/(.12+.48), .48/(.12+.48) > = <.2, .8>

~~P(B|c) = (0.2)(0.4)~~

~~P(B|c) = 0.08~~

When you write it in the capital letter, e.g. B, it’s a variable.

For a Boolean variable B whose value is true, i.e. the proposition, you have to use a small letter, b; For B=false, you should write it as ¬b.

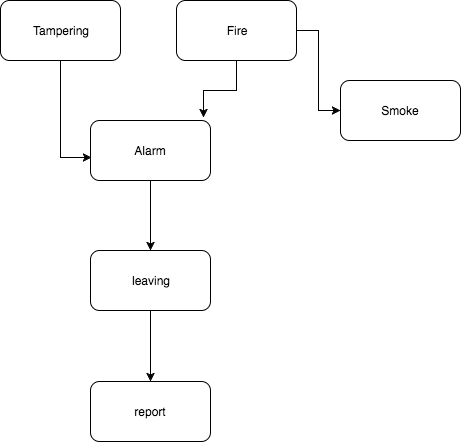
In your notation, P(B|c), you should compute the conditional distribution of B,

i.e. both <P(b|c), P(~b|c) > = <0.2, 0.8> .

**Q5. [10]** Suppose we want to use the diagnostic assistant to diagnose whether there is a fire in a building based on noisy sensor information and possibly conflicting explanations of what could be going on. The agent receives a report about whether everyone is leaving the building. Suppose the report sensor is noisy: It sometimes reports leaving when there is no exodus (a false positive), and it sometimes does not report when everyone is leaving (a false negative). Suppose the fire alarm going off can cause the leaving, but this is not a deterministic relationship. Either tempering or fire could affect the alarm. Fire also causes smoke to rise from the building.

***Construct the Bayesian network*** using the following Boolean variables in the following order:

* *Tampering* is true when there is tampering with the alarm.
* *Fire* is true when there is a fire.
* *Alarm* is true when the alarm sounds.
* *Smoke* is true when there is smoke.
* *Leaving* is true if there are many people leaving the building at once.
* *Report* is true if there is a report given by someone of people leaving. *Repor*t is false if there is no report of leaving.



**Q6.[30]** Consider the simple Bayes net below with Boolean variables *B = BrokeElectionLaw,*

*I = Indicted, M = PoliticallyMotivatedProsecutor, G = FoundGuilty, J = Jailed.*

1. [10] Which of the following are asserted by the network structure? Explain your answer.
2. P(B, I, M) = P(B)P(I)P(M).

* No this is because its need to be P(B,I,M) = P(B)P(I|B,M)P(M) instead of P(B)P(I)P(M).

1. P(J|G) = P(J|G,I)

* Yes this is because J is independent of its no descendant because of its parent G.

1. P(M|G,B,I) = P(M|G,B,I,J)

* Yes this is because M is independent of J – Children ( I, G) – children’s parent (B)

1. [10] Compute P(*b, I ¬ m, g, j*). Show the steps of your computation.

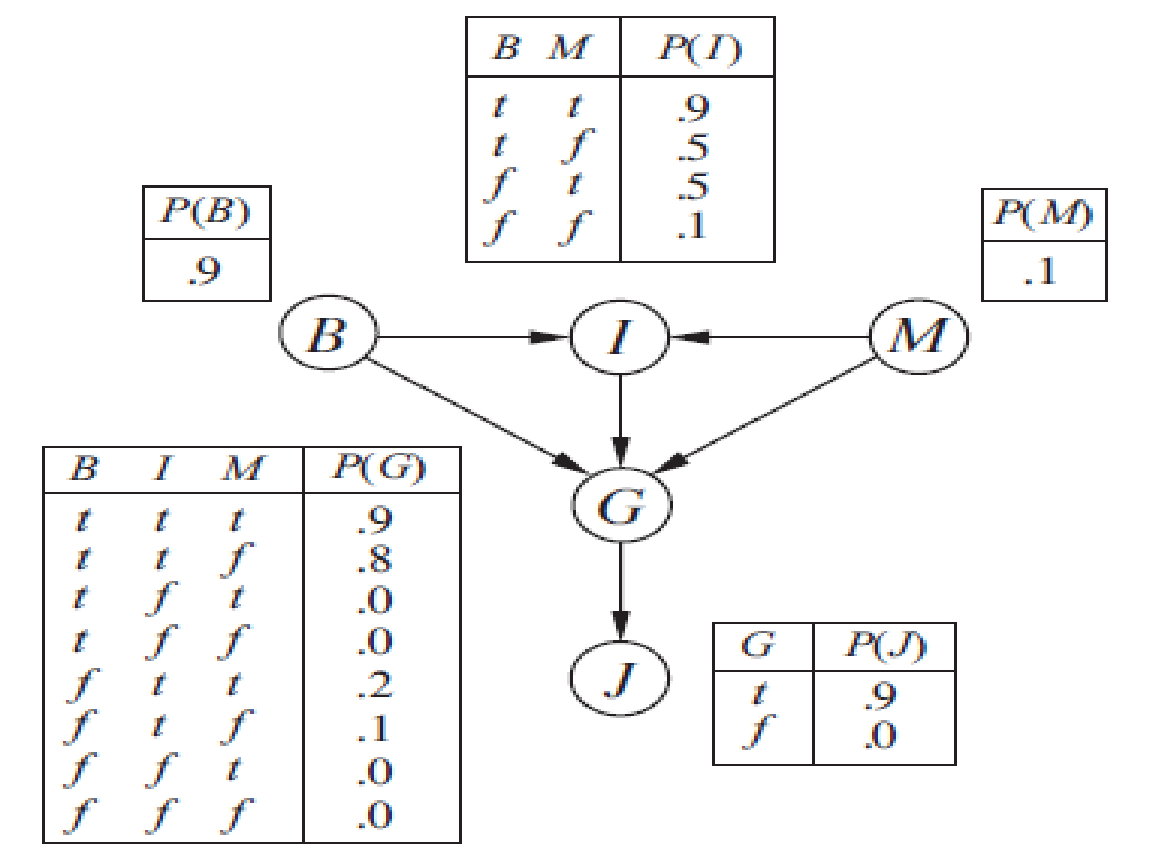
i.e. *P(B=true, I, M=false, G=true, J=true)*.

* 1. < P (b,I, ¬m,j). P(b, ¬I, ¬m, g, j) > = < 2916, 0> - > P(b, I, ¬m, g, j) = P(b) \* P(i|b,¬m) \*P (¬m) \*P (g|I,b, ¬m) \*P (j|g) =

0.9 \* 0.9 \* 0.5 \* 0.8 \* 0.9 = 0.2916

P(b, ¬I, ¬m,g,j) = P9B ) P (¬i|b, ¬m) (P(¬m) P(g|b, ¬I, ¬m) P(j|g) =

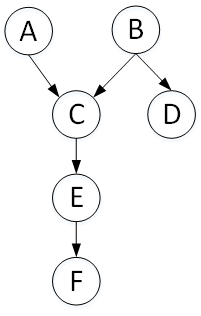
(0.9)(0.5)(0.9)(0.0)(0.9) = 0



1. [10] Compute the probability that someone goes to jail given that they broke the law, have been indicated, and face a politically motivated prosecutor. Define the probability that needs to be computed and compute it, showing its computational steps.

* P(J|B,I,M)=P(J|G)\*P(G|B,I,M) = .9 \* .9 = 0.81

**Q7. [0/ 30]** Consider the following belief network with Boolean variables and the following conditional probabilities:



.

P(*a*) = 0.02 P(*d* | *b* ) = 0.9

P(*b*) = 0.01 P(*d* | *¬ b*) = 0.01

P(*c* | *a, b*) = 0.5 P(e | *c)* = 0.88

P(*c* | *a*, *¬ b*) = 0.99 P(*e* | *¬ c)* = 0.001

P(*c* | *¬ a, b*) = 0.85 P(*f* | *e*) = 0.75

P(*c* | *¬ a, ¬ b*) = 0.0001 P(*f* | *¬ e*) = 0.01

1. [**0/** 15] Compute P(A | *d*, *f* ) using Variable Elimination (VE). Which variables are irrelevant to inference? First, prune irrelevant variables. Show the complete computation steps with the factors that are created for a given elimination ordering.

P(A) =  Σ P(C|A) = .9

P(B) = Σ P(C|B) = .2

P(C) = Σ P(A,B,C) = Σ P(C|A,B))\*P(A)\*P(B) + Σ P(C|A,¬B)\*P(A)\*P(¬B)+ ΣP(C|¬A,B)\*P(¬A)\*P(B)+P(C|¬A, ¬B)\*P(¬A)\*P(¬B)

P(E) = Σ P(C,E) = Σ P(E|C)P(C) + P(E|¬C)\*P(¬C)

d) P(B') = 1-0.10 = 0.90

a) P(D) = 1-P(A)-P(B)-P(C)+P(AD)

= 0.57

b) P(A or D) = 0.57+0.20-0.03

= 0.74

c) P(A/D) = P(AD)/P(D)

= 0.03/0.57

=0.0526

1. [0/ 15] c) Compute P( *e* | *b*) using VE. a) Which variables are irrelevant to inference? b) Which factors can be reused from (1)? d) Show the factors that are different from those in (1).

**Q8. [29/ 15 + Optional]** In the given Bayesian network, compute the following probabilities.

1. [15] What is the probability that an elderly male who is over 60 gets malignant cancer?

P(C=malignant | old, male) = α ∑e ∑s P(C=malignant, old, male, e, s)

= α ∑e ∑s P(C=malignant | e, s) P( e | old) P(s | old, male) P(old) P(male)

= α P(old) P(male) ∑e ∑s P(e | old) P(C = malignant | e, s) P(s | old, male)

= α P(old) P(male) [ P(e | old) P(malignant | e, s) P( s | old, male) + P(e| old) P(malignant | e,~s) P(~s | old, male) + P(~e | old) P(malignant | ~e, s) P( s | old, male) + P(~e | old) P(malignant | ~e, ~s) P(~s | old, male) ]

= α P(old) P(male) [0.7 \* 0.9 \* 0.3 + 0.7 \* 0.6 \* 0.7 + 0.3 \* 0.6 \* 0.3 + 0.3 \* 0.1 \* 0.7 ]

= α P(old) P(male) \* 0.558

=P(C=benign| old, male ) = = α P(old) P(male) \* P( C= benign) = α P(old) P(male) \* 0.442

P(malignant| old, male), P(benign | old, male) = 0.558 / ( 0.558 + 0.442), .442/(0.558 + 0.442) = 0.558, 0.442

1. [14/15, Optional] What is the probability that an elderly male patient with high Serum Calcium to have malignant cancer?

P(cancer | SC = high, Age > 60, G = Male)

P(SC = high | cancer) \* P(Cancer | Age > 60, G = Male) / P(SC =High | Age > 60, Gender = Male)

P(SC = high | cancer) = ~~0.3~~ 0.6

P(cancer | Age > 60, G = Male) = 0.558

P(SC = High | Age > 60, G = Male) = ∑Cancer P(SC=high | cancer, Age > 60, G = Male) \* P(cancer | Age > 60, G= Male) = ~~0.3~~  0.6\* 0.558 + 0.1 \*(1 - 0.558) = ~~0.1674~~ .3348 + 0.0442 = ~~0.2116~~ .379

P(cancer | Serum Calcium = high, Age > 60, Gender = Male) = (~~0.3~~ 0.6 \* 0.558) / ~~0.2116~~ .379 = ~~0.791~~ .8834 -1

