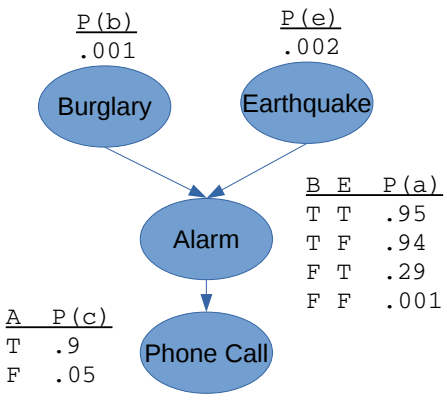
CS444 Naive Bayes Classifier and Inference Exercises

Reminders:

Naive Bayes' Classifier. The c_i are the categories, the X_i are the evidence variables and the u_i are the observed evidence.	$C^{predict} = \underset{c_i \in C}{\text{argmax}} P(c_i) \prod_{j=1}^{M} P(X_j = u_j \mid c_i)$
Bayes' Optimal Classifier	$C^{predict} = \underset{c_i \in C}{\text{argmax}} P(c_i) P(X_1 = u_1, ..., X_M = u_m \mid c_i)$
Marginalization or "summing out". Note that Y and Z can represent sets of variables. Given a table with the full joint PD we can calculate the probability of Y by adding up all of the rows that match.	$P(Y) = \sum_{z \in Z} P(Y, z)$
Definition of conditional probability	$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$
Calculating joint probabilities from a belief network.	$P(X_1, X_2, ..., X_n) = \prod_{i=1}^{N} P(X_i \mid parents(X_i))$

1. Consider the following belief network:



Use "brute-force" marginalization to compute the following probabilities:

(a) $P(c)$

(b) $P(e \mid c)$

2. In a medical study, 100 patients all fell into one of three classes: Pneumonia, Flu, or Healthy. The following database indicates how many patients in each class had fever and headache.

   Consider a patient with a fever but no headache.

   (a) What values would a Bayes optimal classifier assign to the three diagnoses? (A Bayes optimal classifier *doesn't* make any independence assumptions about the evidence variables.) For this question, the three values need not sum to 1. Recall that the Bayes' classifier drops the denominator because it is the same for all three classes.

   (b) What values would a naive Bayes classifier assign to the three possible diagnoses? Show your work. Again, your answers for this question need not sum to 1.

   (c) What *probability* would a Bayes optimal classifier assign to the proposition that a patient has Pneumonia. Show your work. (For this question, the three values should sum to 1.)

   (d) What *probability* would a naive Bayes classifier assign to the proposition that a patient has Pneumonia. Show your work. (For this question, the three values should sum to 1.)

Pneumonia		
Fever	Headache	count
T	T	5
T	F	0
F	T	4
F	F	1
total:		10

Flu		
Fever	Headache	count
T	T	9
T	F	6
F	T	3
F	F	2
total:		20

Healthy		
Fever	Headache	count
T	T	2
T	F	3
F	T	7
F	F	58
total:		70

3. Use the AIspace Belief Network Tool to create a belief network corresponding to the naive Bayes classifier above. Confirm that the network gives the same probability distribution as you found in question 2d above. Submit a screen capture or print-out of your network with the correct observations set and monitoring enabled for the diagnosis node.