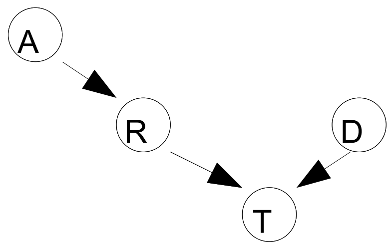
Question 1 (Homework) - More Bayes Nets

In this problem, you will build a small model of the performance of a CS 188 student on the final. A student can either pass or fail (T). This is influenced by his exam readiness (R) and the difficulty of the final (D). His readiness in turn depends on whether he attends lecture (A). Assume all variables in are binary (e.g. D = {d,¬d}).

a) Draw the Bayes net for this model.



What conditional probability tables (CPTs) describe this model?

P(T|R,D), P(R|A), P(A), P(D)

b) Consider the following probability tables.

Table 1

T	A	D	
pass	a	d	0.15
pass	a	¬d	0.01
pass	¬a	d	0.05
pass	¬a	¬d	0.09
fail	a	d	0.1
fail	a	¬d	0.3
fail	¬a	d	0.27
fail	¬a	¬d	0.03

Table 2

T	A	D	
pass	a	d	0.2
pass	a	¬d	0.15
pass	¬a	d	0.6
pass	¬a	¬d	0.5
fail	a	d	0.8
fail	a	¬d	0.85
fail	¬a	d	0.4
fail	¬a	¬d	0.5

Table 3

T	A	D	
pass	a	d	0.018
pass	a	¬d	0.008
pass	¬a	d	0.27
pass	¬a	¬d	0.216
fail	a	d	0.042
fail	a	¬d	0.032
fail	¬a	d	0.27
fail	¬a	¬d	0.144

Which of these tables can represent the joint distribution on test performance, lecture attendance, and exam difficutly P(T,A,D)? Which can model test performance **given** attendance and exam difficulty P(T|A,D)? For now, do not consider whether they respect any particular Bayes net.

Table 1 and 3 are valid joint distributions since the sum of all entries is 1. Table 2 is a valid conditional probability table for P(T|A,D) because the sum over T for each value of (A,D) is 1.

c) Of the tables you identified as modeling the joint distribution P(T,A,D), which can represent the joint distribution for the model represented by the Bayes net in (a)?

Table 3, since in Table1, we do not have P(A,D) = P(A)*P(D) as asserted by the Bayes net.

d) Suppose the following table represents the joint probability over all variables P(T,DA,R).

T	D	A	R	P(T,D,A,R)
pass	*d*	*a*	*r*	0.3024
pass	*d*	*a*	*¬r*	0.0324
pass	*d*	*¬a*	*r*	0.0024
pass	*d*	*¬a*	*¬r*	0.0114
pass	*¬d*	*a*	*r*	0.2394
pass	*¬d*	*a*	*¬r*	0.054
pass	*¬d*	*¬a*	*r*	0.0019
pass	*¬d*	*¬a*	*¬r*	0.019
fail	*d*	*a*	*r*	0.0756
fail	*d*	*a*	*¬r*	0.1296
fail	*d*	*¬a*	*r*	0.0006
fail	*d*	*¬a*	*¬r*	0.0456
fail	*¬d*	*a*	*r*	0.0126
fail	*¬d*	*a*	*¬r*	0.054
fail	*¬d*	*¬a*	*r*	0.0001
fail	*¬d*	*¬a*	*¬r*	0.019

For each of conditional probability tables (CPTs) in theBayes net from part (a), fill in the values of the probabilities such that the joint probability over all variables is given by the table. (You can use the back of the sheet to do this).

P(A)

a	*0.9*
¬a	*0.1*

P(D)

d	*0.6*
¬d	*0.4*

P(R|A)

r	*a*	*0.7*
r	*¬a*	*0.05*
¬r	*a*	*0.3*
¬r	*¬a*	*0.95*

P(T|R,D)

pass	*r*	*d*	*0.8*
pass	*r*	*¬d*	*0.95*
pass	*¬r*	*d*	*0.2*
pass	*¬r*	*¬d*	*0.5*
fail	*r*	*d*	*0.2*
fail	*r*	*¬d*	*0.05*
fail	*¬r*	*d*	*0.8*
fail	*¬r*	*¬d*	*0.5*

e) Suppose we wish to do sampling on this Bayes net to calculate P(T=pass|A=a). Suppose further that you are given 3 sets of numbers generated randomly on [0,1], and that each set contains one random number for each variable in the Bayes net. Describe how you would you use these numbers and the CPTs from part (d) to generate 3 samples according to from the joint distribution of the Bayes net. (Hint: samples produced by rejection sampling **before** rejection are sampled according to the joint distribution of a Bayes net).

We can treat the random number as a way of giving us a "weighted coin" to flip. For example, if we are deciding on a binary variable, if we take the variable to be true when a random number is < 0.9, this event will happen %90 of the time. We can use weighted coin flips for each variable to generate a value. For example, to generate a value for A, would generate A=a if a random number is < 0.9. Note that you could also do this if the number is > 0.1, but you should be consistent about which convention you use. We use the first convention for our solution to (f). To generate a value for T, we must condition on the values that we have generated for its parents R and D. For example, if we have already generated R=r, and D=d, then we would generate T=t if random number < 0.8.

f) Now suppose you are given the following sets of random numbers.

Set 1: {T: 0.45, A: 0.6, D:0.7, R:0.02}

Set 2: {T: 0.1, A:0.95, D:0.5, R:0.5}

Set 3: {T: 0.9, A: 0.8, D: 0.1, R: 0.8}

What samples would you produce under rejection sampling? What about likelihood weighting?

Rejection sampling (before rejection):

pass,a,¬d,r

fail,¬a,d,¬r

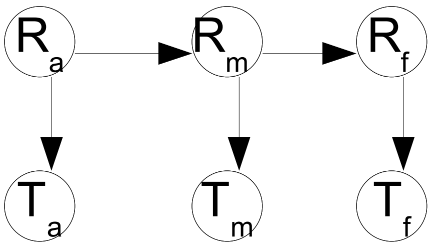
fail,a, d,¬r

After rejection, only [pass,a,¬d,r] and [fail,a, d,¬r] would remain.

For likelihood weighting, we do not need to generate a value for A, so we do not the random number assigned to A. The samples we get are:

pass,a, ¬d, r
fail, a, d, r
fail,a, d,¬r

g) Now suppose that we want to model the performance of a studenton both the the first project, the midterm, and the final. This time, we model a student's performance (T={pass,fail}) as depending only on their test readiness (R={r,¬r}). However, this time, we believe that the amount of preparation a student does for each task depends on the amount of preparation done for the previous task. Draw a Bayes net for this particular model of student performance. (The order of events here is project (p), midterm (m), and final and (f). You may denote, for example, the student's readiness for the midterm as R_m.



h) Viewing this model as an HMM, what are the hidden states? What are the observations?

The readiness variables (R_a, R_m, R_f) are the hidden variables, while the test variables (T_a, T_m, T_f).

i) Suppose we know that the student passed all three tasks. Write an expression for the forward probability that the student was ready for the final. Your answer should have the form P([some variables = some values] | [other variables = other values])

$P(R_f=r|T_p=pass,T_m=pass,T_f=pass)$

j) Write an expression for the same forward probability as in (i), but this time, in terms of the forward probability at the midterm, the probability $P(T_f|R_f)$, and the probability $P(R_f|R_m)$

$$P(R_f=r|T_p=pass,T_m=pass,T_f=pass) \propto P(T_f=pass|R_f=r)(P(R_f=r|R_m=r)\ P(R_m=r|T_m=pass,T_f=pass) + P(R_f=r|R_m=\neg r)\ P(R_m=\neg r|T_m=pass,T_f=pass))$$

This number must be normalized to get the value of

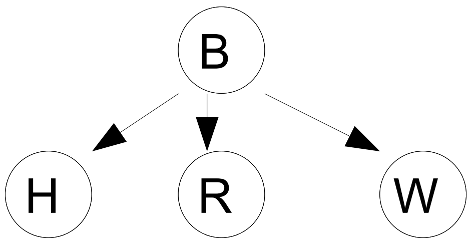
$P(R_f=r|T_p=pass,T_m=pass,T_f=pass)$.

Question 2 (Homework) - Even more Bayes nets

You've been hired to build a quality control system to decide whether car engines coming off an assembly line are *bad* or *ok*. However, this decision must be based on three noisy boolean observations: the engine may be *wobbly* (motion sensor), *rumbly* (sound sensor), or *hot* (heat sensor). Each sensor gives a Boolean observation: *true* or *false.*

You formalize the problem using the following variables and domains:

(a) Draw a Bayes net for a model of these variables in which the sensor readings are independent of each other for a given engine quality.



(b) We would like to use this model to determine the probability that an engine is *ok* given the sensor observations. Write down an expression for the probability P(B=*ok*| *W=true,*R=*false*,*H=true)* in terms of the conditional probabilities that describe the Bayes net you drew in part (a).

$$P(B=ok|W=true, R=false, H=true)=\frac{P(W=true|B=ok)P(R=false|B=ok)P(H=true|B=ok)}{\Sigma_{b'\in\{ok, bad\}}P(W=true|B=b')P(R=false|B=b')P(H=true|B=b')}$$

(c) Consider the following conditional probability tables for this Bayes net.

P(B)

bad	0.375
ok	0.625

P(R|B)

true	*bad*	0.25
false	*bad*	0.125
true	*ok*	0.25
false	*ok*	0.375

P(W|B)

true	*bad*	0.25
false	*bad*	0.125
true	*ok*	0.125
false	*ok*	0.5

P(H|B)

true	*bad*	0.125
false	*bad*	0.25
true	*ok*	0.125
false	*ok*	0.5

Suppose that you are given the sensor readings for a new engine with W=false, R=true, H=false. Using the (estimated) CPTs from (c), what valueof B does your model predict?

Predict B=bad.

$$P(b|\neg w, r, \neg h) \propto \frac{3}{8} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{1}{18}$$
$$P(\neg b|\neg w, r, \neg h) \propto \frac{5}{8} \cdot \frac{4}{5} \cdot \frac{2}{5} \cdot \frac{4}{5} = \frac{4}{25}$$

(d) What is the posterior probability of that prediction (i.e. what is P(BIW=false,R=true,H=false) for whatever value of B your model predicts)?

Normalize above probability, get $\dfrac{\frac{4}{25}}{\frac{1}{18}+\frac{4}{25}} = \dfrac{72}{97}$

(e) The company informs you that, according to their utility function, the utility of rejecting an engine is –1, regardless of whether or not it is actually bad. However, theutility of accepting an engine is 2 for ok engines and –20 for bad engines. Given these utilities, what is the minimum posterior probability your agent needs to have for B= ok before accepting an engine becomes the rational action?

Let $p = P(B = ok|W, R, B)$, then $EU(reject) = -1$ and $EU(accept) = p \cdot 2 + (1 - p) \cdot -20$. Accepting would be a rational decision if $EU(accept) \geq EU(reject)$. These expected utilities reach equality when:

$$-1 = p \cdot 2 + (1 - p) \cdot -20$$

Or when $p = \frac{19}{22}$.