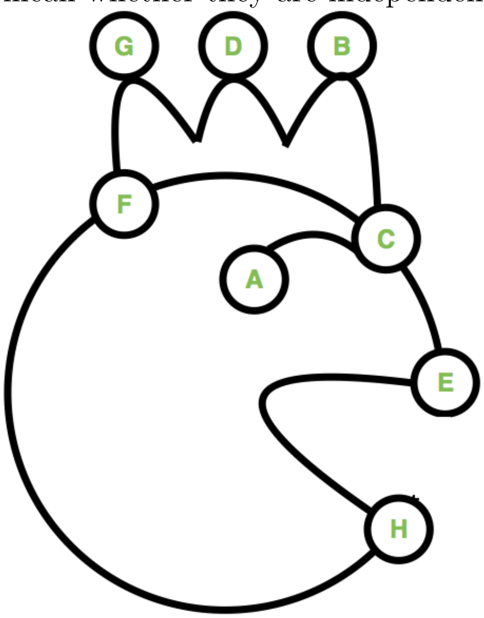
Introduction to Artificial Intelligence — Dieter Fox
CSE 473, Spring 2017

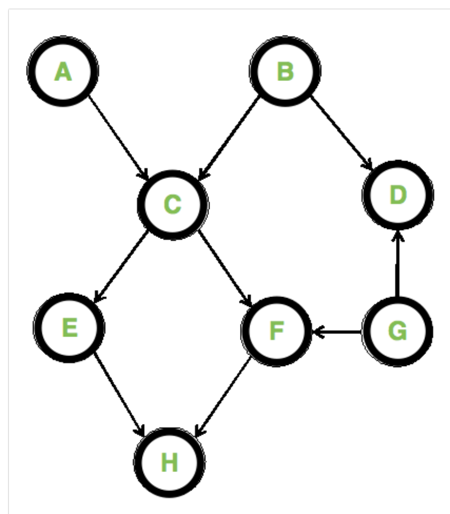
Homework 5

Due on Jun 2, 2017

1. **Bayes Net: Independence**

   Consider the Bayes Net shown below (Please use the right one for clarity. The left one is just an equivalent but cuter version). The following questions are worth 1 point each with a negative point for incorrect answers (don't guess randomly). By independent we mean whether they are independent for any setting of the CPTs.





   (a) Are A and B independent given C? F

   (b) Are A and H independent? F

   (c) Are A and H independent given E? F

   (d) Are E and F independent given H? F

   (e) Are E and F independent given C? T

   (f) Are E and F independent given C and D? T

   (g) Are A and F independent given C and H? T

   (h) Are A and F independent given C and D? F

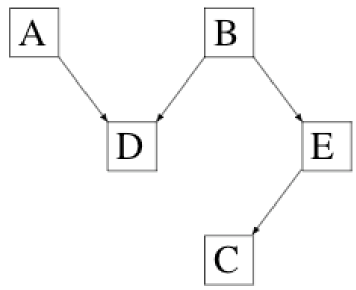
   (i) Are A and F independent given C and G? T

   (j) Are A and F independent given C? T

   (k) Are C and G independent given H? F

2. **Bayes Net: Inference**

   Below you see the structure of a Bayesian network.



   (a) What are the probability distributions that have to be specified in order to completely define the network?

   $$P(A)P(B)P(D \mid A, B)P(E \mid B)P(C \mid E)$$

   (b) How can you compute $P(D \mid A = a, E = e)$? Describe the individual steps of your reasoning. You can assume that all variables are discrete, please use the following notation if you want to sum over the values of a variable, for example, X: $\sum_x P(X = x)$. You should define new factors such as $f_2(X, y) = \sum_z f_1(X, y, Z = z)P(X|Z = z)$. When you eliminate variables, please do so in alphabetical order.

   $P(D \mid A = a, E = e)$
   Initial factors: $P(a)P(B)P(e|B)P(C|e)P(D|a, B)$
   Eliminate B: $f_1(D, a, e) = \sum_b P(B = b)P(e|B = b)P(D|a, B = b)$
   New factors: $f_1(D, a, e)P(a)P(C|e)$
   Eliminate C: $f_2(e) = \sum_c P(C = c|e)$
   New factors: $f_1(D, a, e)P(a)f_2(e)$
   Join to get: $f_3(D, a, e) = f_1(D, a, e)P(a)f_2(e)$
   Normalize over D to get: $P(D|a, e)$

3. **Probabilities**

Consider the *joint probability distribution* below.

A	**B**	**C**	**P(A,B,C)**
false	false	false	0.2
false	false	true	0.05
false	true	false	0.2
false	true	true	0.05
true	false	false	0.1
true	false	true	0.15
true	true	false	0.1
true	true	true	0.15

(a) What is $P(A = true)$? Provide the individual terms involved in this probability.

0.1 + 0.15 + 0.1 + 0.15 = 0.5

(b) What is $P(A = false \mid B = true)$? Provide the individual terms involved in this probability.

$$\frac{P(A = false, B = true)}{p(B = true)} = \frac{0.2 + 0.5}{0.2 + 0.05 + 0.1 + 0.15} = \frac{0.25}{0.5} = \frac{1}{2}$$

(c) Are A and B independent, that is, $A \perp\!\!\!\perp B$? Justify your answer.

True

P(A) = 0.5,0.5

P(B) = 0.5,0.5

P(A,B) = 0.25,0.25, 0.25, 0.25

P(A)P(B) = 0.25,0.25, 0.25, 0.25

Alternate answer: see that $P(A) = P(A \mid B)$ from previous 2 answers

4. **(Optional, not graded) Hidden Markov Models**

Consider a Hidden Markov Model where the hidden state X_t can be one of three values $\{A, B, C\}$. The transition probabilities are provided in the following table, where the row corresponds to X_{t-1} and the column to X_t.

	A	B	C
A	0.7	0.3	0
B	0.1	0.7	0.2
C	0	0.4	0.6

For example, $P(X_t = A \mid X_{t-1} = B) = 0.1$.

The noisy sensor model for evidence E_t corresponding to X_t gives the true hidden state with probability 0.8, and one of the other two states each with probability 0.1. For example, $P(E_t = B \mid X_t = A) = 0.1$.

(a) Assume our belief about the hidden state X_t is

X_t	$P(X_t)$
A	0.5
B	0.5
C	0

Compute the belief about the hidden state X_{t+1} before considering noisy evidence (no need to normalize):

X_{t+1}	$P(X_{t+1})$
A	$0.35 + 0.05 = 0.4$
B	$0.35 + 0.15 = 0.5$
C	0.1

(b) Given your answer from the previous question, now assume we have the noisy sensor reading $E_{t+1} = C$. Compute our posterior belief taking this evidence into account (no need to normalize):

X_{t+1}	$P(X_{t+1})$
A	$0.1 * 0.4 = 0.04$
B	$0.1 * 0.5 = 0.05$
C	$0.8 * 0.1 = 0.08$

continued on next page

(c) Assume now we are using a particle filter with 3 particles to approximate our belief instead of using exact inference. Imagine we have just applied transition model sampling (elapse-time) from state X_t to X_{t+1}, and now have the set of particles $\{A, A, B\}$. What is our belief about X_{t+1} before considering noisy evidence?

X_{t+1}	$P(X_{t+1})$
A	2/3
B	1/3
C	0

(d) Now assume we receive sensor evidence $E_{t+1} = B$. What is the weight for each particle, and what is our belief now about X_{t+1} (before weighted resampling)?

Particle	Weight
A	0.1
A	0.1
B	0.8

X_{t+1}	$P(X_{t+1})$
A	0.2
B	0.8
C	0

(e) Will performing weighted resampling on these weighted particles to obtain our final three particle representation for X_{t+1} cause our belief to change? **Briefly** explain why or why not.

Yes, because there will be three unweighted particles which can't represent this belief

5. **(Optional, not graded) Create Bayes Net**

   Create a Bayes net with exactly four states {A,B,C,D}, that follows all of the independence constraints below.

   (a) $A \perp\!\!\!\perp B$

   (b) $A \not\perp\!\!\!\perp D|B$

   (c) $A \perp\!\!\!\perp D|C$

   (d) $A \not\perp\!\!\!\perp C$

   (e) $B \not\perp\!\!\!\perp C$

   (f) $A \not\perp\!\!\!\perp B|D$

   (g) $B \perp\!\!\!\perp D|A, C$

   Answer: