

CS188: Artificial Intelligence, Spring 2009

Written Assignment 3: Probabilistic Models

Due: April 16 (extended!)

You can work on this in groups, but everyone should turn in his/her own work.

Don't forget your **name and login**.

Name:

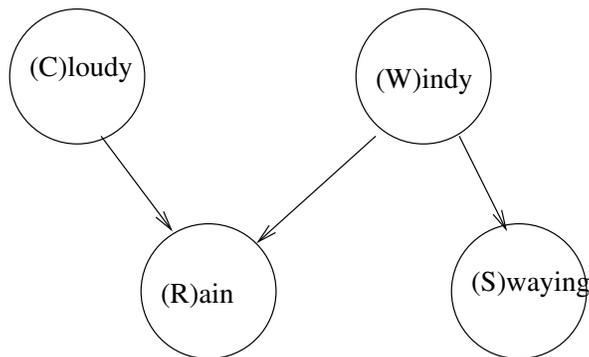
Login:

GSI:

Section Time:

1 Question 1: Bayes Net Inference

Dr. A. Gent, an armchair climatologist, wants to predict whether or not it's going to (R)ain. When it's (C)loudy, it tends to (R)ain unless it's (W)indy. From her armchair, Dr. Gent can observe the clouds (C) but not the wind. However, if it's (W)indy, then the tree outside tends to (S)waying, which she can observe. Sometimes the tree (S)ways even when it's not windy. Squirrels, perhaps?



$P(C)$	
C	$P(C)$
c	0.4

$P(W)$	
W	$P(W)$
w	0.6

$P(R C, W)$			
C	W	R	$P(R = r C, W)$
c	w	r	0.40
c	$\neg w$	r	0.80
$\neg c$	w	r	0.10
$\neg c$	$\neg w$	r	0.20

$P(S W)$		
W	S	$P(S = s W)$
w	s	0.9
$\neg w$	s	0.3

Note: the tables above are abbreviated, but fully specified, e.g., $P(\neg r|c, w) = 1 - P(r|c, w) = 0.6$.

a) What is the probability of perfect weather, $P(\neg c, \neg r, \neg w)$?

$P(C)$	
C	$P(C)$
c	0.4

$P(W)$	
W	$P(W)$
w	0.6

$P(R C, W)$			
C	W	R	$P(R = r C, W)$
c	w	r	0.40
c	$\neg w$	r	0.80
$\neg c$	w	r	0.10
$\neg c$	$\neg w$	r	0.20

$P(S W)$		
W	S	$P(S = s W)$
w	s	0.9
$\neg w$	s	0.3

Conditional probability tables (CPTs) are repeated for convenience.

b) A *dangling* node is a leaf node that is *unobserved* in a query. When computing $P(C, R, W)$, S is dangling. Show that the probabilities in the CPT for $P(S|W)$ cannot affect $P(C, R, W)$ for this network, regardless of the values in the CPTs above.

c) Can the CPT of a dangling leaf node for a query ever affect the outcome of that query? If so, when? If not, what preprocessing step does this fact suggest for variable elimination?

d) Compute $P(R)$ using the variable elimination algorithm, perhaps modified by your answer to (c). Write down the intermediate factors after summing out each hidden variable. *Hint: you can answer this question with one intermediate factor, in addition to $P(R)$.*

e) Now, approximate $P(R|C = \neg c, S = \neg s)$ using each of the two sets of samples below.

Prior samples (rejection sampling):

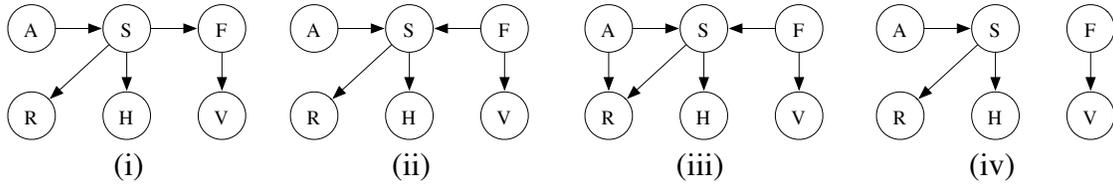
$c, \neg w, s, r$	$\neg c, w, s, \neg r$	$c, w, s, \neg r$	$\neg c, \neg w, \neg s, \neg r$	$P(R = r C = \neg c, S = \neg s) \approx$
$\neg c, w, s, \neg r$	$\neg c, w, \neg s, \neg r$	$c, \neg w, \neg s, \neg r$	c, w, s, r	
$\neg c, w, s, r$	$c, \neg w, \neg s, r$	$\neg c, w, s, \neg r$	$\neg c, w, s, \neg r$	
$\neg c, \neg w, \neg s, \neg r$	$\neg c, w, s, \neg r$	$c, w, s, \neg r$	$\neg c, \neg w, s, r$	

Weighted samples (likelihood weighting):

$\neg c, w, \neg s, \neg r$	$\neg c, w, \neg s, r$	$P(R = r C = \neg c, S = \neg s) \approx$
$\neg c, \neg w, \neg s, \neg r$	$\neg c, w, \neg s, \neg r$	

2 Question 2: Conditional Independence Review

Consider the following Bayes' nets.



For each of the following conditional independence properties, circle the number of each network for which the statement is guaranteed to be true.

$F \perp\!\!\!\perp R$ (i) (ii) (iii) (iv)

$F \perp\!\!\!\perp A$ (i) (ii) (iii) (iv)

$V \perp\!\!\!\perp A \mid F$ (i) (ii) (iii) (iv)

$V \perp\!\!\!\perp A \mid R$ (i) (ii) (iii) (iv)

$F \perp\!\!\!\perp R \mid S$ (i) (ii) (iii) (iv)

$A \perp\!\!\!\perp R \mid V$ (i) (ii) (iii) (iv)

3 Question 3: The Captain's Treasure

You are searching for the treasure of Captain Long Beard, lost for fifty years. You've narrowed down the search to four sites on the corners of Dinghy Island. Here is what you know:

- i. The treasure is in one of four locations: the North-East, North-West, South-East, or South-West. You assign equal probability to each. Let $L \in \{ne, nw, se, sw\}$ be the (unknown) location of the treasure.
- ii. You can only dig in one of the four locations. Let $A \in \{ne, nw, se, sw\}$ indicate where you dig.
- iii. The North of the island is rocky, and the South sandy. If the treasure is in the North and you dig in the correct location, you retrieve the treasure with probability 0.5, whilst if the treasure is in the South and you dig in the correct place, you retrieve it with probability 0.9. Let $T \in \{t, f\}$ be the variable recording whether you retrieve the treasure.
- iv. The town's guard claims to have seen whether Long Beard took the road to the North or to the South of the island when burying his treasure. The guard is willing to (honestly) tell you what he remembers, but his memory is shaky and biases Northward. If the Captain went North, the guard would remember. If the Captain went South, the guard has a $\frac{1}{4}$ chance of remembering North anyway. Let $G \in \{n, s\}$ be the variable indicating whether the guard remembers the Captain burying the treasure in the North or the South of the island.
- v. Your utility U is measured in dollars. The treasure chest is filled with gold, worth \$1000. The East of the island is reachable only by river, so if you dig there you must buy a raft for \$40.

a) Draw the decision network for this problem and fill in the conditional probability tables for each variable and the utility function table.

L	$P(L)$
ne	
nw	
se	
sw	

L	A	T	$P(T L, A)$
ne	ne	t	
nw	nw	t	
se	se	t	
sw	sw	t	
$L \neq A$		t	

L	G	$P(G L)$
ne	n	
nw	n	
se	n	
sw	n	

A	T	$U(A, T)$
ne	t	
ne	f	
nw	t	
nw	f	
se	t	
se	f	
sw	t	
sw	f	

b) What is the best place to dig if you *do not* know what the town guard remembers. What is the expected utility of this action?

c) Suppose the guard tells you that he remembers the Captain going North. Now what is the best place to dig, and what is the expected utility of this action? *Hint: first compute the distribution $P(L | G = n)$.*

d) Now suppose the guard tells you that he remembers the Captain going South instead. What is the best place to dig based on this information, and what is the expected utility of this action?

e) What is the value of the guard telling you what he remembers (assuming you don't already know)?

f) The guard will only tell you the information if you promise to dig in the East (his wife sells rafts). How much money do you expect to gain (or lose) by accepting his offer, versus digging without his information?

4 Question 4: Playing Poker

You are playing poker against a friend. Over the years, you've collected data about her tendency to bluff (as opposed to playing her cards straight up) and decide to model her habits with the following HMM.

State	$P(X_1)$
<i>bluff</i>	0.2
<i>straight up</i>	0.8

Previous State	Current State	$P(X_t X_{t-1})$
<i>bluff</i>	<i>bluff</i>	0.2
<i>bluff</i>	<i>straight up</i>	0.8
<i>straight up</i>	<i>bluff</i>	0.6
<i>straight up</i>	<i>straight up</i>	0.4

State	Observation	$P(O_t X_t)$
<i>bluff</i>	<i>rolls eyes</i>	0.6
<i>bluff</i>	<i>scratches shoulder</i>	0.4
<i>straight up</i>	<i>rolls eyes</i>	0.3
<i>straight up</i>	<i>scratches shoulder</i>	0.7

- a) On the first hand, your friend rolls her eyes. What is the posterior probability that she is bluffing on the first hand?
- b) You observe two consecutive eye rolls on hands 1 and 2. What's the probability that she is bluffing on the second hand?
- c) You confirm that she was in fact bluffing on the first two hands. What is the probability that she will bluff again on the third hand?
- d) Without any observations, what is the probability that she is bluffing on the k^{th} hand as $k \rightarrow \infty$?

You now decide to track your friend's behavior via particle filtering. **You have only two particles, one initially assigned to *bluff* and the other to *straight up*.**

e) You observe an eye roll on the first hand. What is the probability that your particles will still be *bluff* and *straight up* after resampling based on this evidence?

f) After resampling, what is the expected value of the error in your estimate of $P(X_1 = \textit{bluff} \mid O_1 = \textit{rolls eyes})$?

g) After observing $O_1 = \textit{rolls eyes}$, you have the option of calling her bluff. A correct call nets you \$3; an incorrect call nets you -\$1. If you do not call her bluff, you net \$0. In dollars, how much more or less money should you expect to make if you use a particle filter with two particles (starting with one for each outcome) to estimate $P(X_1 \mid O_1 = \textit{rolls eyes})$ versus computing it with Bayes rule, assuming you maximize your expected utility given your (possibly approximate) beliefs?