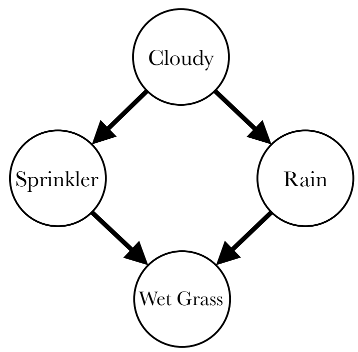
Computational Cognitive Science
Problem Set: Bayes Nets

January 24, 2014

1 A simple Bayes net

In this problem you will derive predctions about the Sprinkler Bayes net (see Figure 1). This problem is intended to provide experience specifying and performing inference in Bayes nets. This problem should take around one hour.

P(C=F)	P(C=T)
0.5	0.5



C	P(S=F)	P(S=T)
F	0.5	0.5
T	0.9	0.1

C	P(R=F)	P(R=T)
F	0.8	0.2
T	0.2	0.8

S	R	P(W=F)	P(W=T)
F	F	1.0	0.0
T	F	0.1	0.9
F	T	0.1	0.9
T	T	0.01	0.99

Figure 1: A simple Bayes net.

1. Write down the full joint distribution for the sprinkler Bayes net, first expanding the joint distribution using the chain rule for probabilities, then simplifying this expression by exploiting the dependency structure expressed in the Bayes net. Justify your work.

2. Manually calculate the probability that the sprinkler is on given that the grass is wet and it is cloudy. Show your work.

3. Call the previous Bayes net the Human-controlled model. Set up an alternative Bayes net the automatic sprinkler model. Under this model, the sprinkler has a rain sensor and turns itself on when it is not raining. (This means your Bayes net will have a different structure.) The conditional probability table (CPT) for the sprinkler is now:

Raining	P(Sprinkler is on)
true	0.05
false	0.5

All other conditional probability tables are the same as before. Calculate the probability that the sprinkler is on given that the grass is wet.

2 Monte Carlo: baby steps

Monte Carlo methods compute results through the use of repeated random sampling. They're used when exact enumeration is difficult or intractable. When we wish to calculate a conditional probability of an event, x'; $x' \in X$, the set of all events, given evidence, e, we must calculate the probability of the event and evidence occurring together, $P(x', e)$, then divide that by the probability of the evidence[1]. Formally,

$$\begin{aligned} P(x'|e) &= \frac{\sum_{\{x_i \in X : x', e\}} P(x_i)}{\sum_{\{x_j \in X : e\}} P(x_j)} \\ &= \frac{P(x', e)}{P(e)} \end{aligned} \tag{1}$$

Unfortunately, depending on the size of X, these sums (especially the denominator) can be difficult to calculate. This is where Monte Carlo comes in.

2.1 Rejection sampling

One of the simplest (albeit dumbest) flavors of Monte Carlo simulation is *rejection sampling.* Rejection sampling works just as you might imagine, it samples from a distribution specified by the Bayes net (our tables in Figure 1) and rejects the samples that don't match the evidence. The desired probability is then computed by counting how many samples contain both x' and e out of the total number of samples. Thus, Equation 1 becomes,

$$P(X|e) \approx \frac{Count(x', e)}{Count(e)} \tag{2}$$

It is important to note that our equals symbol has become an approximation. The accuracy of the simulation is dependent on the number of samples drawn.

Suppose that we wish to find the probability that the grass is wet given it's cloudy. We run the rejection sampler for 100 samples. Imagine that of those 100 samples, 46 of those contain e, that it's cloudy, and 33 of those samples contain x', the grass is wet, we have[2]:

$$P(X|e) \approx \frac{Count(x', e)}{Count(e)} = \frac{33}{46} = 0.717 \tag{3}$$

2.1.1 Walkthrough

We'll stick with the sprinkler example in Figure 1. Suppose we want to calculate the probability that the grass is wet given that it's cloudy. For each sample we:

1. Set the value of *cloudy* by randomly sampling based on it's CPT
2. - Sample the value of *rain* conditional on the value of *cloudy*

[1] For an introduction to the baffling world of set-builder notation, see below. This will definitely help you out, as predicate logic tends to come up a lot in the literature because it's a compact way to describe relationships and theories.

[2] This is an actual answer taken from the solution code. It might be a good way to check your own code.

- Sample the value of *sprinkler* conditional on the value of *cloudy*

3. Sample the value of *grass is wet* conditional on the values of *rain* and *sprinkler*

4. - Increment a count, $Count(e)$, if we see the evidence: *cloudy*
   - Increment a count, $Count(x', e)$, if we see the evidence and the event: it is cloudy and the grass is wet

The answer is simply the counts plugged into Equation 2.

2.2 Problems

5. Implement your own rejection sampler and use it to verify your answers to questions 2 and 3.

6. Compute the answer for 5 using 10, 100, 1000, and 10000 samples. Do this multiple times, and compare the error between answers.

7. Discuss, in no more than three paragraphs total, your stance on the following issues. You can provide a mixed treatment (i.e. address a few of them simultaneously) but be sure to include at least a few sentences on each.

   (a) Some psychologists have criticized Bayesian models of cognition on the grounds that it seems implausible to expect to find Bayes rule in the head. Do you find that criticism compelling? If so, why? If not, why not?

   (b) Based on lecture, this problem set, and your readings, do you think there are other, more psychologically (or computationally) compelling strategies for approximately implementing Bayesian inference on computing machines? Very good answers to this question could lead to cognitive science or computer science papers.

Appendix: notation

symbol	**meaning**
$\{\}$	denotes a set
$\forall$	For all, for each
$\exists$	There exists, there is at least one
$\exists!$	There exists exactly one
$:$	such that
$\in$	in, belongs to
$\notin$	not in, does not belong to
$\wedge$	and
$\vee$	or

Examples

- $\{x \in X : x', e\}$
  The set of x in [some other set] X, such that x' and e [are true].

- $\{x \in \mathbb{R} : x^2 = 1\}$
  The set of x in the real numbers, such that x^2 is equal to 1, i.e., 1 and -1.
  Note that this is equivalent to $\{x \in \mathbb{R} : |x| = 1\}$. The set of x in the reals, such that the absolute value of x is equal to 1.[3]

- $\{k : (\exists n \in \mathbb{N})(k = 2n)\}$
  The set of k such that for some natural number n, k is equal to twice n, i.e., all even natural numbers.

- $\{k : \exists p, q \in \mathbb{Z}(q \neq 0) \wedge (qk = p)\}$
  The set of k such there there exists two integers p and q, where q is not 0, *and* k multiplied by q is equal to p, i.e., the set of rational numbers, or $\mathbb{Q}$.

[3] Examples taken from wikipedia set-builder notation page. (January 24, 2014)