Homework 3

(Due Monday 23:59)

1) Metropolis–Hastings

(10 points)

Continuing the “Categorizing blickets” task in Homework 2 (i.e., use the same task, same PCFG and same likelihood function), for this homework, suppose we have observed only one creature:

Observation	Color	Shape	Texture	Blicket?
o_1	red	round	soft	yes

Now suppose we are running a Metropolis–Hastings sampler over PCFG trees using the subtree-regeneration proposal described in class. The current hypothesis

$$t = \texttt{red},$$

and the proposal selects the root H node and regenerates the proposed hypothesis

$$t' = (\texttt{red} \wedge \texttt{square}).$$

1. Write down the PCFG derivations for t and t'. Then compute $|t|$ and $|t'|$, the number of selectable nonterminal nodes in each tree.

2. Compute $P(D \mid t)$ and $P(D \mid t')$.

3. Using the subtree-regeneration acceptance probability,

   $$a(t \to t') = \min\left\{1, \frac{P(D \mid t')}{P(D \mid t)} \frac{|t|}{|t'|}\right\},$$

   compute the probability of accepting the proposed move from t to t'.

4. Explain how to use the acceptance probability to decide whether to move to t' or stay at t.

   Then say what happens in each of the following two cases:

   - the sampled random number is $u = 0.03$;
   - the sampled random number is $u = 0.07$.

5. In your own words, why is it useful for Metropolis–Hastings to sample hypotheses rather than always choosing the hypothesis with the higher posterior probability? Why might the sampler sometimes accept a worse hypothesis?

2) Hormone and Treatment

(4 points)

Doctors have given a treatment T to a group of patients. Unbeknownst to the doctors, 10% of patients have hormone A in their blood, 90% of patients don't. The Treatment only helps ($D = 0$) the patients with hormone A, and kills the rest ($D = 1$). The SCM is:

$$D := (A \wedge \neg T) \vee (\neg A \wedge T)$$

a) What is the interventional probability $P(\neg D|\text{Do}(T))$ of surviving if given the treatment?

b) Suppose you've observed a patient who did not receive the treatment and died. What is the probability that this patient would have survived if he had received the treatment?

3) Logic and Math Classes

(3 points)

Students can pass the Logic class only if they have high Skill (S). They can pass the Math class if they have high Skill or if they Work hard (W). At the beginning of semester, a student needs to choose a class (if $C = 1$, they choose the math class, otherwise they choose the logic class). The SCM is:

$$M := C \wedge (S \vee W)$$

$$L := \neg C \wedge S$$

John chose the Math class and he passed ($C = 1, M = 1$). If he had chosen the logic class, what is the probability that he would have passed? In other words, compute $P(L = 1_{C=0} \mid M = 1, C = 1)$. Express your answer as a function of $P(S)$ and $P(W)$.

4) The Boulder and the Hiker

(4 points) Consider the following scenario. There is a boulder (i.e. a large rock) rolling toward a hiker. Seeing the boulder, the hiker jumps out of the way. If the hiker had not jumped he would have died.

The SCM for this scenario is:

$$J := B$$

$$D := \neg J \wedge B$$

where B: the boulder rolls down; J: the hiker jumps; D: the hiker dies. In the actual world, $B = 1$, $J = 1$, $D = 0$.

Using the Halpern definition:

a) Compute whether $J = 1$ is an actual cause of the hiker surviving.

b) Compute whether $B = 1$ is an actual cause of the hiker surviving.