Homework 3: Answer sheet

1 Metropolis-Hastings

1.1 Tree sizes **2 points**

The current hypothesis is

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

Its derivation is

$$H \to A,$$

followed by

$$A \to \texttt{red}.$$

Thus the nonterminal nodes in the tree are one H node and one A node. Therefore,

$$|t| = 2.$$

The proposed hypothesis is

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

Its derivation is

$$H \to (H \wedge H).$$

The left branch is

$$H \to A \to \texttt{red},$$

and the right branch is

$$H \to A \to \texttt{square}.$$

Thus the nonterminal nodes are:

- the root H node,
- the left child H node,
- the left child A node,
- the right child H node,

- the right child A node.

Therefore,

$$|t'| = 5.$$

1.2 Likelihoods **2 points**

The observed creature is red, round, and soft, and it is a blicket.

The hypothesis

$$t = \texttt{red}$$

is true of this creature, because the creature is red. Therefore t predicts "yes," which matches the observation. Hence,

$$P(D \mid t) = 0.9.$$

The hypothesis

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

is false of this creature, because although the creature is red, it is not square. It is round. Therefore t' predicts "no," which does not match the observation. Hence,

$$P(D \mid t') = 0.1.$$

1.3 Metropolis–Hastings acceptance probability **3 points**

The acceptance probability is

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

Substituting the likelihoods and tree sizes:

$$a(t \to t') = \min\left\{1, \frac{0.1}{0.9} \cdot \frac{2}{5}\right\}.$$

Since

$$\frac{0.1}{0.9} = \frac{1}{9},$$

we get

$$a(t \to t') = \min\left\{1, \frac{1}{9} \cdot \frac{2}{5}\right\} = \min\left\{1, \frac{2}{45}\right\}.$$

Therefore,

$$a(t \to t') = \frac{2}{45} \approx 0.044.$$

So the proposed move is accepted with probability approximately 4.4%.

1.4 Accept or reject? **1 point**

To decide whether to accept the proposed move, sample a random number

$$u \sim \text{Uniform}(0, 1).$$

If

$$u \leq a(t \to t'),$$

then accept the proposal and move to t'. Otherwise, reject the proposal and stay at t. Here,

$$a(t \to t') = \frac{2}{45} \approx 0.044.$$

If $u = 0.03$, then

$$0.03 < 0.044,$$

so we accept the proposal and move to

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

If $u = 0.07$, then

$$0.07 > 0.044,$$

so we reject the proposal and stay at

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

1.5 Conceptual question **2 points**

Metropolis–Hastings is designed to sample from the posterior distribution, not simply to find the single best hypothesis. Sampling is useful because the posterior distribution represents uncertainty over hypotheses. Instead of committing to only one hypothesis, we can approximate how plausible many different hypotheses are.

The sampler sometimes accepts worse hypotheses because this helps it explore the hypothesis space. If it always chose the better hypothesis, it could get stuck in a local optimum and fail to discover other important regions of the posterior. Accepting worse moves with some probability helps maintain the correct stationary distribution and allows the sampler to approximate the full posterior distribution over hypotheses.

2 Hormone and Treatment

1. From the structural equation we can see that the interventional probability is $P(\neg D|\text{Do}(T)) = P(A) = 0.1$.

2. Formally we need to compute $P(\neg D_{T=1} \mid D, \neg T)$. This is equal to:

$$\sum_A P(\neg D \mid \mathrm{Do}(T), A)\, P(A \mid D, \neg T)$$

Since $P(\neg D \mid \mathrm{Do}(T), A) = 1$ and $P(\neg D \mid \mathrm{Do}(T), \neg A) = 0$, this reduces to $P(A \mid D, \neg T)$. This can be expanded as:

$$\frac{P(A, D, \neg T)}{P(D, \neg T)} = \frac{P(A, D, \neg T)}{\sum_A P(A, D, \neg T)} = 1$$

So the patient would have definitely survived.

3 Logic and math classes

We have:

$$P(L = 1_{C=0} \mid M = 1, C = 1) = \sum_{C,W,S} P(L \mid \mathrm{Do}(C = 0), C, W, S)\, P(C, W, S \mid M, C = 1)$$

Note that L does not depend on W, and we intervened on C. Also, $P(S|M, C) = P(S|M)$ because $M = 1$ implies $C = 1$. So we can simplify the above as:

$$\sum_S P(L \mid \mathrm{Do}(C = 0), S)\, P(S \mid M)$$

$$= P(L \mid \mathrm{Do}(C = 0), S)\, P(S \mid M) + P(L \mid \mathrm{Do}(C = 0), \neg S)\, P(\neg S \mid M)$$
$$= P(S \mid M)$$

We can compute $P(S \mid M)$ using Bayes' rule:

$$P(S \mid M) = \frac{P(M \mid S)\, P(S)}{P(M)} = \frac{P(C)P(S)}{P(M)} = \frac{P(S)P(C)}{P(C)[P(W) + P(S) - P(S)P(W)]}$$
$$= \frac{P(S)}{P(W) + P(S) - P(S)P(W)}$$

4 The Boulder and the Hiker

If J had been 0, then we have $D := 1 \wedge 1$, so the hiker would have died. Therefore, J is a cause.

For B: if B had not rolled down, J would have been 0. With these values D is still 0. Freezing J to its actual value doesn't change this. Is (B, J) a cause? No, because of the minimality condition: $\{J\}$ is a proper subset of $\{B, J\}$, and $\{J\}$ is already a cause. Therefore B is not an actual cause of the hiker surviving.

Marking Sheet

Part	Criteria	Points
1.1	Correct derivation for $t = \texttt{red}$ and correct count $\lvert t\rvert = 2$	1
	Correct derivation for $t' = (\texttt{red} \wedge \texttt{square})$ and correct count $\lvert t'\rvert = 5$	1
1.2	Correctly states that t predicts "yes" and gives $P(D \mid t) = 0.9$	1
	Correctly states that t' predicts "no" and gives $P(D \mid t') = 0.1$	1
1.3	Correctly writes or applies the MH acceptance formula	1
	Correctly substitutes likelihoods and tree sizes	1
	Correct final acceptance probability $a(t \to t') = \frac{2}{45} \approx 0.044$	1
1.4	Correct accept/reject rule using $u \sim \text{Uniform}(0, 1)$	0.5
	Correct decisions: accept for $u = 0.03$, reject for $u = 0.07$	0.5
1.5	Explains that MCMC samples from the posterior rather than merely optimizing	1
	Explains why accepting worse hypotheses can help exploration or avoid getting stuck in local optima	1
2.1	At marker's discretion	2
2.2	At marker's discretion	2
3	At marker's discretion	3
3.1	At marker's discretion	2
3.2	At marker's discretion	2
	Total	21