COMPSCI 370	**Instructor:** Ronald Parr
Homework 5	**Due:** Wednesday, April 21, 2021

1 Bayes Nets I (5 points)

Consider a Bayesian network with a simple, chain-like strucure and binary variables, $A \ldots E$, where A has no parents, B has A as its only parent, C has B as its only parent, and so on. (We'll use the letters A through E for conciseness, but if you'd like a simple example of how such a network might arise, A could be *discipline*, B could be *studying*, C could be *knowledge*, D could be *grades* and E could be *Internship quality*).

Suppose we wish to compute $P(E = e)$, by eliminating A, then B, then, C and so on. How many multiplications and additions are required to do this? You should explain your reasoning and give an exact number, **not** a $O()$ bound in this case.

2 Bayes Nets II (10 points)

Consider the AFSHN network from class, and the following conditional probabilities:

- $P(a) = P(f) = \frac{1}{4}$
- $P(s|af) = 1$, $P(s|a\overline{f}) = \frac{3}{4}$, $P(s|\overline{a}f) = \frac{3}{4}$, $P(s|\overline{af}) = \frac{1}{4}$
- $P(n|s) = \frac{3}{4}, P(h|s) = \frac{1}{2}$, $P(n|\overline{s})\frac{1}{8}$, $P(h|\overline{s}) = \frac{1}{4}$

Compute the joint probability distribution $P(HN)$ by variable elimination. You should write out your answer in a way that makes the order in which you are doing summations explicit. You should also show all entries for every table that you compute as a result of eliminating a variable, but you don't need to show every calculation to compute these entries, i.e., you can use a calculator and just plug the numbers into your table if you want.

3 HMMs (15 points)

Consider a hidden Markov model with two states, x and y, two observations a and b, and the following somewhat unusual parameters. We use the random variable S_t for the state at time t:

- $P(S_{t+1} = x|S_t = x) = P(S_{t+1} = y|S_t = y) = 1$
- $P(a|S_t = x) = P(b|S_t = x) = 0.5$
- $P(b|S_t = y) = 1.0$

These parameters were chosen to help you realize some things about HMMs as you work through the basic algorithms. Assume that at time 0, the distribution over states is: $P(S_0 = x) = 0.6$, and that there are no observations at time 0. At time 1, the observation is b. At time 2, the observation is a.

a) Show the work to run the forward algorithm (AKA monitoring or tracking) and compute the distribution over states at time steps 1 and 2, i.e., $P(S_1|O_1 = b)$ and $P(S_2|O_1 = b, O_2 = a)$.

b) Show the work to compute the smoothed distribution over states at time step 1, i.e., $P(S_1|O_1 = b, O_2 = a)$, where O_t is the observation at time t. (Hint: You should see a big difference between the forward probabilities at time 1 and the smoothed probabilities.)

c) Show the work to compute the Viterbi path. Note that your answer should be a sequence of 3 states.