COMPSCI 370	**Instructor:** Ronald Parr
Homework 5	**Due:** Thursday April 21, 2022

Note: We will be updating this PDF with specific boxes in which to write your answers. We will be requesting that, unlike previous assignments, you write your answers in specific parts of a PDF we give you and that you upload this modified PDF to gradescope. This is to allow us to grade your assignments and get them back to you before the final exam.

Also, unlike previous assigments, **we will not be able to accept late submissions**. Disallowing late submissions is necessary to allow us to release the solutions and grades promptly.

1 Filling CPTs (10 points)

Consider the sinus inflamation (AFSHN) network from the slides, and compute the CPTs for this network from the joint distribution shown below. Note that you will not get full credit for less than 20 numbers.

Assignment	Probability	Assignment	Probability
$\overline{afshn}$	$\frac{1323}{4096}$	$a\overline{fshn}$	$\frac{126}{4096}$
$\overline{afsh}n$	$\frac{189}{4096}$	$a\overline{fsh}n$	$\frac{18}{4096}$
$\overline{afs}h\overline{n}$	$\frac{441}{4096}$	$a\overline{fs}h\overline{n}$	$\frac{42}{4096}$
$\overline{afs}hn$	$\frac{63}{4096}$	$a\overline{fs}hn$	$\frac{6}{4096}$
$\overline{af}s\overline{hn}$	$\frac{18}{4096}$	$a\overline{f}s\overline{hn}$	$\frac{36}{4096}$
$\overline{af}s\overline{h}n$	$\frac{54}{4096}$	$a\overline{f}s\overline{h}n$	$\frac{108}{4096}$
$\overline{af}sh\overline{n}$	$\frac{54}{4096}$	$a\overline{f}sh\overline{n}$	$\frac{108}{4096}$
$\overline{af}shn$	$\frac{162}{4096}$	$a\overline{f}shn$	$\frac{324}{4096}$
$\overline{a}f\overline{shn}$	$\frac{126}{4096}$	$af\overline{shn}$	$\frac{21}{4096}$
$\overline{a}f\overline{sh}n$	$\frac{18}{4096}$	$af\overline{sh}n$	$\frac{3}{4096}$
$\overline{a}f\overline{s}h\overline{n}$	$\frac{42}{4096}$	$af\overline{s}h\overline{n}$	$\frac{7}{4096}$
$\overline{a}f\overline{s}hn$	$\frac{6}{4096}$	$af\overline{s}hn$	$\frac{1}{4096}$
$\overline{a}fs\overline{hn}$	$\frac{36}{4096}$	$afs\overline{hn}$	$\frac{14}{4096}$
$\overline{a}fs\overline{h}n$	$\frac{108}{4096}$	$afs\overline{h}n$	$\frac{42}{4096}$
$\overline{a}fsh\overline{n}$	$\frac{108}{4096}$	$afsh\overline{n}$	$\frac{42}{4096}$
$\overline{a}fshn$	$\frac{324}{4096}$	$afshn$	$\frac{126}{4096}$

Hint: These may look like nasty fractions, but they've been chosen carefully so that your CPT entries will all simplify to fractions with a single digit in the numerator and the denominator.

$P(a)$:

$P(\overline{a})$:

$P(f)$:

$P(\overline{f})$:

$P(s|af)$:

$P(s|a\overline{f})$:

$P(s|\overline{a}f)$:

$P(s|\overline{af})$:

$P(\overline{s}|af)$:

$P(\overline{s}|a\overline{f})$:

$P(\overline{s}|\overline{a}f)$:

$P(\overline{s}|\overline{af})$:

$P(h|s)$:

$P(h|\overline{s})$:

$P(\overline{h}|s)$:

$P(\overline{h}|\overline{s})$:

$P(n|s)$:

$P(n|\overline{s})$:

$P(\overline{n}|s)$:

$P(\overline{n}|\overline{s})$:

2 Variable Elimination (10 points)

Use the variable elimination algorithm to compute $P(S|H)$. Your answer must include 4 numbers to get full credit. Remember that you can compute $P(S|H)$ from $P(SH)$, so it's fine to compute $P(SH)$, then derive the conditional probabilities from there.

Show your work, but write your final answers in these boxes:

$P(s|h)$:

$P(\overline{s}|h)$:

$P(s|\overline{h})$:

$P(\overline{s}|\overline{h})$:

3 Hiden Markov Models (25 points)

For this up problem, we will *initially* consider a simple Markov model (nothing hidden). Algorithmically, this is the same as ignoring the term for the probability of the evidence, or just assuming that all observations have probability 1.

Suppose we are tracking two states h and u for whether a patient is healthy or unhealthy, and that $P(S^{t+1} = h|S^t = h) = \frac{3}{4}$, and $P(S^{t+1} = h|S^t = u) = \frac{1}{2}$. Suppose further that the patient is believed to be healthy with probability $\frac{1}{2}$ at time step 0, and that we are considering the patient through time step 2. One motivation for this type of model could be figuring out when a patient initially became ill for the purposes of contact tracing.

1. Write down the probability of each path through the state space.

2. Use the Viterbi algorithm to compute the highest probability path, and verify that it matches the path with the highest probability from your previous answer. We will be asking you to write down the table stored by the Viterbi algorithm for middle time step (time 1), and we will focus on that as a way of verifying that you have executed the algorithm correctly..

3. Suppose you are told that the state at time step 2 is u, but you have no observations of the state at time steps 0 or 1. Compute and write down the Viterbi path. **You should show your work, but write your final answer in the provided box.**

4. Given the same assumptions as the previous question, what is the smoothed (hindsight) distribution over the state at time 1, i.e., $P(S^1|S^2 = u)$? (Note that this requires a slight modification to the smoothing algorithm to work with a known future state, rather than future observations.) **You should show your work, but write your final answer in the provided boxes.**

5. Suppose that instead of being told the state, there is a test variable $T \in \{t, \bar{t}\}$ that can be observed. Define $P(t|u) = 1.0$, and $P(t|h) = x$ for some $0 \leq x \leq 1$. Observe that if $x = 1$, it is a worthless test, since it always returns the same result regardless of the state. On the other hand, if $x = 0$, then a result of t tells you exactly what state the patient is in at the time the test variable is observed. Assume that you observe a positive test at time step 2, i.e., $T^2 = t$, but make no other observations. Derive the largest value of x that can cause the Viterbi path to change from your answer in part 2? **We will ask for a single number in a box, but you should also show your work.**

6. Observe that for the forward (monitoring) equations, the highest probability state at step 1 is h. Using the same assumptions as the previous question, we can ask about the largest value of x is that could cause h to no longer be the highest probability state at time step 1 after we have computed the smoothed distribution. Note that, in general, it's possible for the highest probability state at a particular time step to *not* be on the Viterbi path, so don't assume that your answer will be the same as what you got for the previous question. **We will ask for a single number in a box, but you should also show your work.**

3.1 (8 points)

$P(hhh)$: ☐

$P(hhu)$: ☐

$P(huh)$: ☐

$P(huu)$: ☐

$P(uhh)$: ☐

$P(uhu)$: ☐

$P(uuh)$: ☐

$P(uuu)$: ☐

3.2 (6 points)

Highest probability path:

Entry for h state at time step 1:

Entry for u state at time step 1:

3.3 (3 points)

Highest probability path:

Don't forget to show your work as well.

3.4 (4 points)

$P(S^1 = h|S^2 = u)$: ☐

$P(S^1 = u|S^2 = u)$: ☐

Don't forget to show your work as well.

3.5 (2 points)

$x =$ ☐

Don't forget to show your work as well.

3.6 (2 points)

$x = \boxed{}$

Don't forget to show your work as well.