

# Probabilistic Graphical Models

## HWK#2 Part A

Assigned Sunday, Feb. 9, 2020

Due : Thursday, March 5, 2020

### Problem 1 (Variable Elimination)

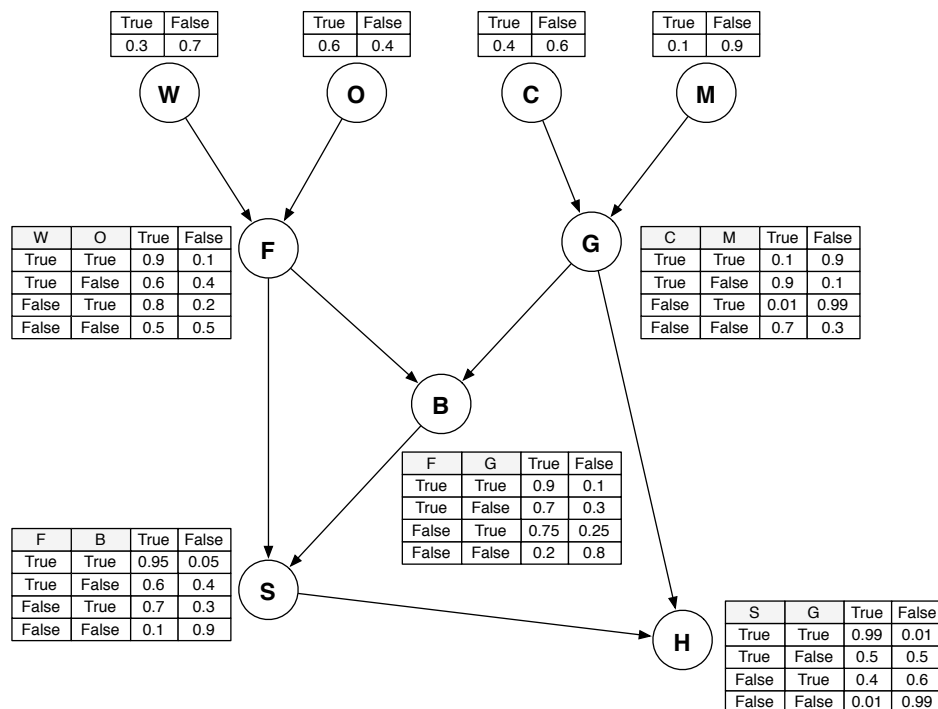


Figure 1: Football network

After two losses in three games to start the season, the world champion Pittsburgh Steelers seem to be in a world of trouble. Not knowing what else to do, Coach Cowher has enlisted your help in trying to turn things around during the bye week. In Figure 1, you will find a Bayes net that hopefully contains the key to a successful season. The variables of interest are: Weather (W), Opponent (O), Fans (F), Big Ben (B), Steelers Win (S), Good Health (G), Chunky Soup (C), Motorcycle (M), and Happy (H).

Coach Cowher would like to know the answers to the following questions, but keep in mind, he does not like to be kept waiting. There's only one thing he hates more than losing, and that is exponential computational complexity.

1. What is the chance that the Steelers win given that Big Ben eats his Chunky Soup? [  $P(S = T|C = T) = ?$  ]
2. How fired up are the fans given that Big Ben is in good health? [  $P(F = T|G = T) = ?$  ]
3.  $P(M = T|G = T) = ?$
4.  $P(M = T|G = T, S = T) = ?$
5.  $P(W = T|G = T, B = F, S = T) = ?$

Additionally, report the ordering used and the factors produced after eliminating each variable for the first query  $[P(S = T|C = T)]$ .

## Problem 2 (Inference in HMM)

There are a set of firstnames: david, anton, fred, jim, barry which are numbered 1 (david) to 5 (barry) and a set of surnames: barber, ilsung, fox, chain, fitzwilliam, quinceadams, grafvonunterhosen, which are numbered 1 (barber) to 7 (grafvonunterhosen). A string is generated by first randomly sampling a character from a to z. Then we either continue this random character sampling with probability 0.8, or we start to generate a firstname. A firstname is chosen uniformly at random. After generating a firstname, we generate a character at random. We continue to generate another letter at random with probability 0.8, or start to generate a surname (chosen uniformly from the set of surnames). We continue until the last letter of the surname is generated. The process then goes back to start (unless we are at the end timepoint  $T = 10000$ ). For example, we might generate then dtyjimdfilsungffdavidmjfox.... The catch is that the process of character generation is very noisy. The character we want to generate is only generated correctly with probability 0.3, with probability 0.7 that another character (uniformly at random) will be generated. Answer the following problem:

Given 10000 character sequence in the file ``noisystring.mat'', you must decode the ``noisystring' to find the most likely intended 'clean' sequence. Once you have this sequence, you will be able to find a set of (firstname,surname) pairs as you traverse along the clean sequence. Construct a matrix  $m(i,j)$  with the counts of the number of occurrences of the pair (firstname(i),surname(j)) in your clean sequence and [display/print](#) this matrix as your final answer in your solution.

You must provide your code as well.

### Problem 3 (Maximum Probable Explanation)

Consider a distribution defined over binary variables:

$$P(a, b, c) \equiv P(a|b)P(b|c)P(c)$$

with

$$P(a = tr \mid b = tr) = 0.3 \ , \ P(a = tr \mid b = fa) = 0.2 \ , \ P(b = tr \mid c = tr) = 0.75$$

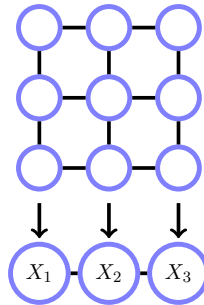
$$P(b = tr \mid c = fa) = 0.1 \ , \quad P(c = tr) = 0.4,$$

What is the most likely joint configuration? That is  $\operatorname{argmax}_{a,b,c} P(a, b, c)$  ?

Hint: It is not acceptable to just naively work out the 8 possible states to solve the problem. Naïve approach is only feasible for this example but not for a general problem with many variables. Instead, you must use the max-product algorithm to solve this problem.

### Problem 4. (Inference in Ising Model- Markov Networks)

Consider a binary variable Markov Random Field  $p(x) = Z^{-1} \prod_{i>j} \phi(x_i, x_j)$ , defined on the  $n \times n$  lattice with  $\phi(x_i, x_j) = e^{\mathbb{I}[x_i=x_j]}$  for  $i$  a neighbour of  $j$  on the lattice and  $i > j$ . A naïve way to perform inference is to first stack all the variables in the  $t^{\text{th}}$  column and call this cluster variable  $X_t$ , as shown. The resulting graph is then singly connected. What is the complexity of computing the normalisation constant based on this cluster representation? Compute  $\log Z$  for  $n = 10$ .



You need to provide your software code for answering  $\log(Z)$  for  $n=10$ . Give also your answer for  $\log(Z)$  so that we do not need to run your code.

### Problem 5 (Inference in BN Networks)

The file ‘diseaseNet.mat’ contains the potentials for a disease bi-partite belief network, with 20 diseases  $d_1, \dots, d_{20}$  and 40 symptoms,  $s_1, \dots, s_{40}$ . The disease variables are numbered from 1 to 20 and the symptoms from 21 to 60. Each disease and symptom is a binary variable, and each symptom connects to 3 parent diseases.

1. Using the BRMLtoolbox, construct a junction tree for this distribution and use it to compute all the marginals of the symptoms, i.e.,  $p(s_i = 1)$ . In addition to providing your code, you must display/print the results in your solution paper work.
2. Explain how to compute the marginals  $p(s_i = 1)$  (in the previous part) in a way more efficient than using the junction tree formalism. By implementing this method, compare it with the results from the junction tree algorithm from the previous part.
3. Symptoms 1 to 5 are present (state 1), symptoms 6 to 10 not present (state 2), and the rest not known. Compute the marginal  $p(d_i = 1 | s_{1:10})$  for all diseases.

**In all above questions, in addition to providing your codes, you must display/print the results in your solution paper work.**

## Problem 6   Clique Trees

*Consider the distribution*

$$p(x_1, x_2, x_3, x_4) = \phi(x_1, x_2)\phi(x_2, x_3)\phi(x_3, x_4)\phi(x_4, x_1)$$

1. *Write down a junction tree for the above distribution.*
2. *Carry out the absorption procedure and demonstrate that this gives the correct result for the marginal  $p(x_1)$ .*

**Three more problems in Part B**