

CSE 150 Programming Assignment 4

Due November 21, 2016 11:59 PM

1 Overview

In this project, you will be implementing inference algorithms for Bayesian networks. The Bayesian networks we use are the cloudy network shown in Figure 1, and the salmon network shown in Figure 2. For the cloudy network, you have to perform the exact inference algorithm by hand. For both networks, you have to solve the inference problem using the code you write for three approximate inference algorithms: rejection sampling, likelihood weighting, and Gibbs sampling.

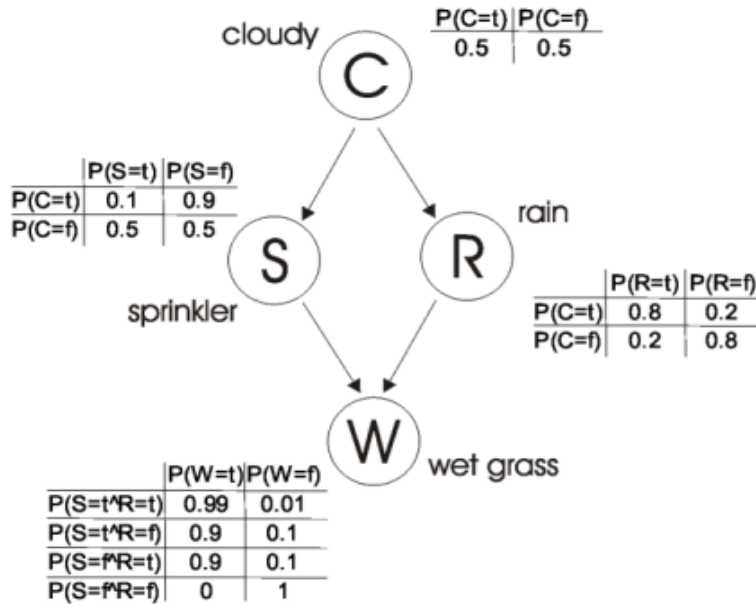


Figure 1: Cloudy Network

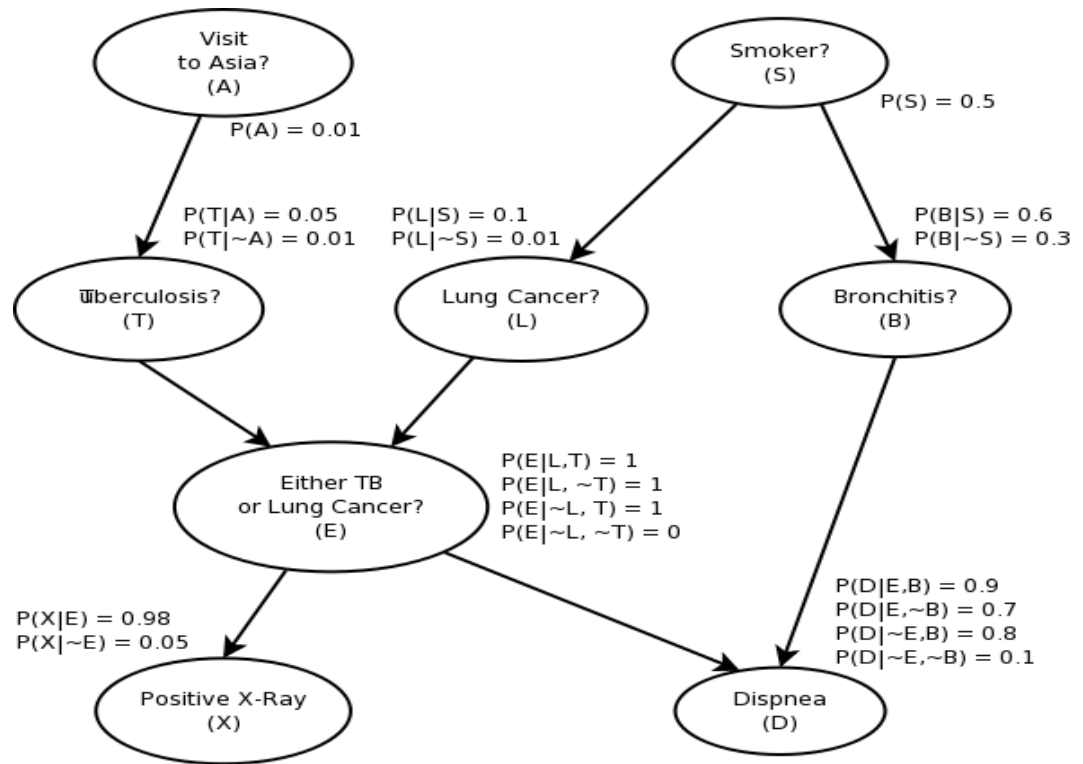


Figure 2: Salmon Network

2 Provided Code

We have provided code to construct a Bayesian network. Some basic elements in a Bayesian network, such as CPT(conditional probability table), Random-Variable, Sample, Edge, and Node classes are provided in Assignment4.py. The BayesianNetwork class in BayesianNetwork.py provides an interface to construct a Bayesian network. All the approximate inference algorithms should also be implemented in this class. You can download the started code from Piazza Resources section. The file would be present as StarterCodePA4.zip

3 Problems

You must implement rejection sampling, weighted sampling, and Gibbs sampling algorithms. The public interfaces of these methods are marked in BayesianNetwork.py with TODO. You must also submit a report, as described below.

3.1 Problem 1(3 Points)

For the cloudy network, manually compute the probability $P(W = T | R = T, C = T)$, show the result and steps you used in the report.

3.2 Problem 2(2 Points)

For the cloudy network, manually compute the probability $P(S = T | R = T)$, show the result and steps you used in the report.

3.3 Problem 2(22 Points)

For the two above queries in the cloudy network, implement rejection sampling, likelihood weighting, and Gibbs sampling. Report the following:

1. Submit code for BayesianNetwork.py which have all the 3 sampling algorithms implemented (14 points)
2. Describe the three sampling algorithms and how you implement them. (3 points)
3. For each algorithm and each query, perform a test with at least 5 different reasonably large values for the number of trials parameter, and graph how the estimated probabilities converge to the correct probability. (3 points)
4. How many trials do you need for each method to give an estimate that was reasonably close to the actual probability? How do these methods compare to each other? Discuss your findings. (2 points)

You can test your code by running - **python CloudyNetwork.py**

3.4 Problem 4(8 Points)

For the salmon network, compute the query of $P(T = \text{True} | A = \text{True}, S = \text{False}, X = \text{True}, D = \text{False})$ using the three sampling methods. Repeat steps for 3.3.3 and 3.3.4

You can test your code by running- **python SalmonNetwork.py**

Hint: You might need iterations on the order of 10^6 for this problem.

3.5 Problem 5

A paragraph from each author stating what their contribution was and what they learned.