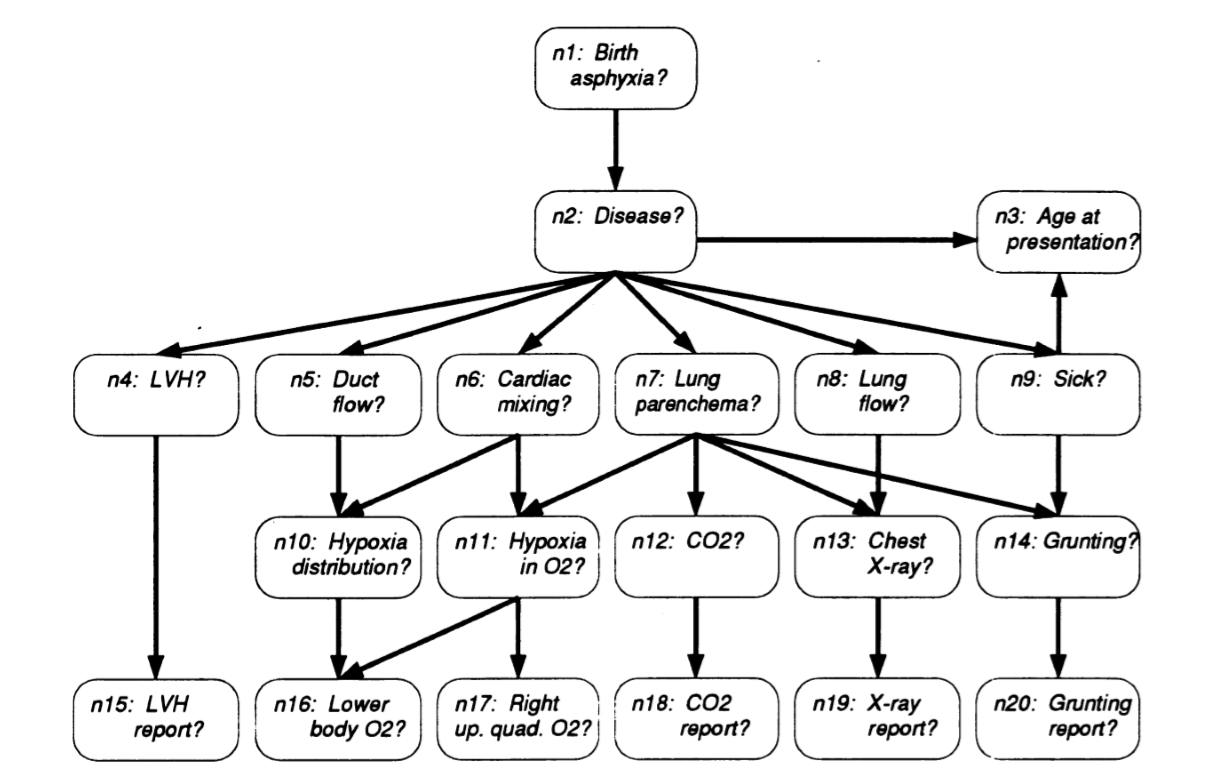
ECSE 4810/6810 Fall, 2020

Project 2

**Approximate Inference**

Due 11:59 pm, Oct 30, 2020

In this project, you will implement approximate inference methods to perform inference on the CHILD Bayesian Network below.



The CHILD network is for diagnosing congenital heart disease in a new born “blue” baby. It provides a mechanism so that both clinical expertise and available data can be exploited to generate diagnostic aid. Further information on the network can be found in [1]. Figure 1 shows the CHILD Bayesian Network. The basic information of the network is as follows:

Number of nodes: 20

Number of arcs: 25

Number of parameters: 230

Size of node: 2,3,4,5 or 6

Each node is discrete, with different number of states. Possible states for each variable are:

Birth Asphyxia: {“yes”, “no”}

Disease: {“PFC”, “TGA”, “Fallot”, “PAIVS”, “TAPVD”, “Lung”}

Age: {“0-3 days”, “4-10 days”, “11-30 days”}

LVH: {“yes”, “no”}

Duct flow: {“Lt to Rt”, “None”, “Rt to Lt”}

Cardiac mixing: {“None”, “Mild”, “Complete”, “Transparent”}

Lung parenchema: {“Normal”, “Oedema”, “Abnormal”}

Lung flow: {“Normal”, “Low”, “High”}

Sick: {“yes”, “no”}

Hypoxia distribution: {“equal”, “unequal”}

Hypoxia in O2: {“None”, “Moderate”, “Severe”}

CO2: {“Normal”, “Low”, “High”}

Chest X-ray: {“Normal”, “Oligaemic”, “Plethoric”, “Grd.\_Glass”, “Asy/Patchy”}

Grunting: {“yes”, “no”}

LVH report: {“yes”, “no”}

Lower body O2: {“<5”, “5-12”, “12+”}

RUQ O2: {“<5”, “5-12”, “12+”}

CO2 report: {“<7.5”, “>=7.5”}

X-ray report: {“Normal”, “Oligaemic”, “Plethoric”, “Grd.\_Glass”, “Asy/Patchy”}

Grunting Report:{“yes”, “no”}

For each node, a CPT is specified. For example, the CPTs for LVH and LVH-report nodes are given below.

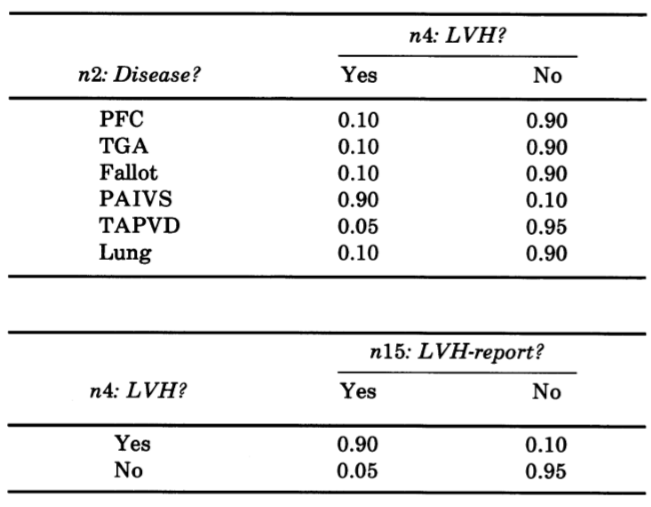


Table 1: CPT for links nd2--nd4 and nd4--nd15

The CPTs for all nodes can be acquired from the link below. Follow the readme file to extract the CPT table for each node.

<https://rpi.box.com/s/10ff90vrec3zjjo973pe9v8vdpj1nign>

Given the BN and its parameterization, perform the following inference tasks

1. Posterior probability inference: Compute p(Birth Asphyxia=yes | CO2report =”<7.5”, LVHReport=yes, and X-rayReport=Plethoric)
2. MAP inference: Disease\*=argmax Disease p(Disease | CO2report =”<7.5”, LVHReport=yes, and X-rayReport=Plethoric)

using the following approximated inference methods

1. The likelihood weighted sampling
2. The Gibbs sampling
3. The mean field method (6000 level students only)

Submit the following to your RPI box:

1. a report that summarizes the inference methods and the results. The report should include the following
2. An introduction that discusses the aims of this project and the inference methods you use
3. Discuss each inference method, its theory, and provide a pseudo code
4. For weighted likelihood sampling method, explain if you observe any bias in its estimation and if the sampling order matters.
5. For Gibbs sampling method, use tables and figures to show the effect of initialization, burn-in time, and skip time on the inference results. Also, discuss how you empirically decide if the sampling converges, i.e., the end of the burn-in period.
6. For the mean field method, show the derivations and explain why the results are different from other two methods.
7. Discussion of the results, the differences between different inference methods, strengths and weaknesses of each method, issues encountered, what you have learned, and conclusions
8. the source code of your inference methods. Document how to compile and run your code to perform both posterior and MAP inference.

Reference:

[1] Spiegelhalter, David J., et al. *Bayesian Analysis in Expert Systems*. Statistical Science, vol. 8, no. 3, 1993, pp. 219–247.

Your project will be evaluated based on the following criteria:

1. Report (40%)-the report should be professionally written, with introduction,

problem statement, theory, experiments, conclusion, and related references if applicable. All figures and tables should be clearly numbered and cited, with proper captions.

1. Results (40%) – correctness of inference results for the required inferences as well as additional new inferences we will perform on your code.
2. Code (20%) – document your code adequately. For each function, clearly define its variables, write down its purpose, and identify its input and output arguments. Also explain how to compile and run your code.