Machine Learning: Data to Models
Assignment 1b: Bayesian Network

Qun Gao, JHED ID qgao6:
Li-Yi Lin, JHED ID: llin34

2.3 Parameter Estimation [20 points]

a. [5 points]

If X is binary, given its parents $Pa(X)$, we can get the parameters $\theta_{X^0|Pa(X)}$ and $\theta_{X^1|Pa(X)}$. Since $\theta_{X^0|Pa(X)} + \theta_{X^1|Pa(X)} = 1$, so we only need to estimate one of them. In this case, the parameters of this network are: $\theta_{a^0}, \theta_{b^0|a^0}, \theta_{b^0|a^1}, \theta_{c^0|a^0,b^0}, \theta_{c^0|a^0,b^1}, \theta_{c^0|a^1,b^0}, \theta_{c^0|a^1,b^1}, \theta_{d^0|b^0}, \theta_{d^0|b^1}$, $\theta_{e^0|b^0,c^0}, \theta_{e^0|b^0,c^1}, \theta_{e^0|b^1,c^0}, \theta_{e^0|b^1,c^1}$, 13 parameters in total.

b. [15 points]

Algorithm 1 Initialization for EM algorithm

```
1: procedure INITIALIZE
2:     //G: Bayesian network with nodes X_A, X_B, X_C, X_D, X_E
3:     //θ^0: randomly set parameters in (a)
4:     //D: Partial observed data in the Table 1 of assignment 1b handout.
5:     //ε: convergence threshold
6:     for i ∈ {A, B, C, D, E} do
7:         for each x_i, u_i ∈ Val(X_i, Pa^G_{X_i}) do
8:             M̄_{θ^0}[x_i, u_i] = 0;
```

Algorithm 2 EM Algorithm for learning parameters in Bayes Network

procedure EM-ALGORITHM
    INITIALIZE;
    $isConverge = False$;
    $\epsilon =$ one small number;
    t=0;
    while $isConverge == False$ **do**
        //E-Step (Compute the expected sufficient statistics)
        for $m = 1...n$ **do** //n is sample size
            Run inference on $\langle \mathcal{G}, \theta \rangle$ using evidence $o[m]$;
            for $i \in \{A, B, C, D, E\}$ **do**
                for $x_i, u_i \in Val(X_i, Pa_{X_i}^{\mathcal{G}})$ **do**
                    $\bar{M}_{\theta^t}[x_i, u_i] = \bar{M}_{\theta^t}[x_i, u_i] + P(x_i, u_i \mid o[m], \theta^t)$;
        //M-Step
        for $i \in \{A, B, C, D, E\}$ **do**
            for $x_i, u_i \in Val(X_i, Pa_{X_i}^{\mathcal{G}})$ **do**
                // $Val(X_i, Pa_{X_i}^{\mathcal{G}})$ finds all the parents of X_i and
                // return a set of paired values (X_i, parent of X_i)
                $\theta_{x_i|u_i}^{t+1} = \frac{\bar{M}_{\theta^t}[x_i, u_i]}{\bar{M}_{\theta^t}[u_i]}$
        //Check convergence
        if the difference between every parameter at t and $t+1 < \epsilon$ **then**
            $isConvergence = True$;
            return θ^{t+1}
        $t = t + 1$;