**CS 63018 & CS 73018 Probabilistic Data Management**

**Homework 2**

**Instructor:** Xiang Lian

**Due Date:** Please refer to the course website

1. Please search related topics on Bayes' formula on the Web, give the formula, and explain its rationale. **[10 points]**

2. Given *P*(*X*1) and *P*(*X*2 | *X*1) for two nodes in the Markov sequence as shown in Figure 1, please provide the joint probability *P*(*x*1, *x*2). **[10 points]**



**Figure 1.** Markov Sequence

3. What are the joint probabilities of the Bayesian network below (**Hint:** using Bayes' formula)? **[30 points]**



**Figure 2.** Bayesian Network

3(a). Joint probability *P*(A, B) [15points]

3(b). Joint probability *P*(A, C) [15 points]

4. Read the abstract and Sections 1-2 of the following SIGMOD 2003 paper. Write a description of the problem studied in this paper, including the motivation, problems/queries, and challenges of this problem. **[50 points]**

* Reynold Cheng, Dmitri V. Kalashnikov, Sunil Prabhakar. Evaluating Probabilistic Queries over Imprecise Data. In *SIGMOD*, 2003.

<https://www.ics.uci.edu/~dvk/pub/SIGMOD03_dvk.pdf>

**Bonus Question** What is the joint probability of *P*(A, B, C, D) in Figure 2 (**Hint:** using Bayes' formula for multiple times) [extra 20 points]

**Submission**

Submit an electronic copy of your homework solution to the [Blackboard](https://learn.kent.edu/).

18