**CSC529: Advanced Data Mining**

**Practice Handout: Bayesian Networks**

**Learning objective**: how to build network models to reason under uncertainty according to the laws of probability theory

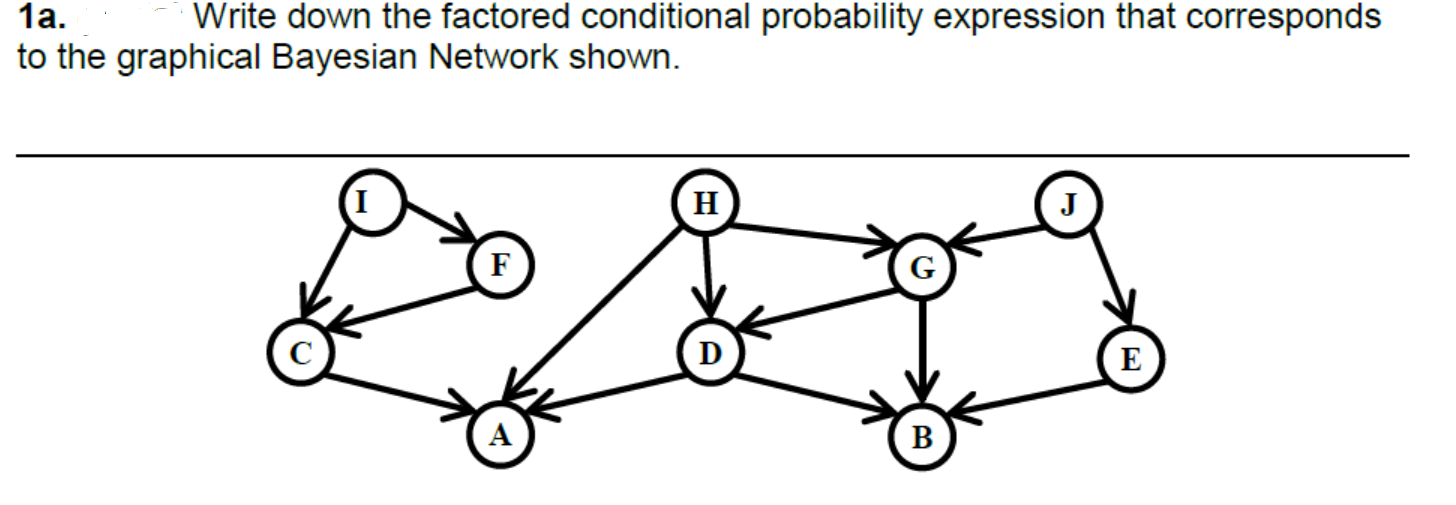
**Definition:** A Bayesian network is a directed graph in which each node is annotated with quantitative probability information. The full specification is as follows:

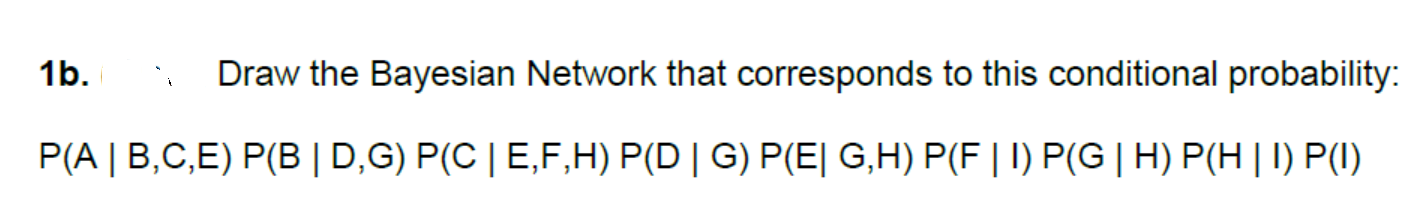
1. Each node corresponds to a random variable, which may be discrete or continuous.

2. A set of directed links or arrows connects pairs of nodes. If there is an arrow from node X to node Y, X is said to be a parent of Y. The graph has no directed cycles (and hence is a directed acyclic graph, or DAG.

3. Each node Xi has a conditional probability distribution P(Xi | Parents(Xi))that quantifies the effect of the parents on node.

**Problem 1:**





**Problem 2**: You have a new burglar alarm installed at home. It is fairly reliable at detecting a burglary, but also responds on occasion to minor earthquakes. (This example is due to Judea Pearl, a resident of Los Angeles—hence the acute interest in earthquakes.) You also have two neighbors, John and Mary, who have promised to call you at work when they hear the alarm. John nearly always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too. Mary, on the other hand, likes rather loud music and often misses the alarm altogether. Given the evidence of who has or has not called, we would like to estimate the probability of a burglary.

