Introduction to Bayesian Networks Part 1

Gennaro De Luca, PhD
Arizona State University

The lecture is based on the slides developed by Prof. Yu Zhang from ASU School of Computing and Augmented Intelligence

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

Objective

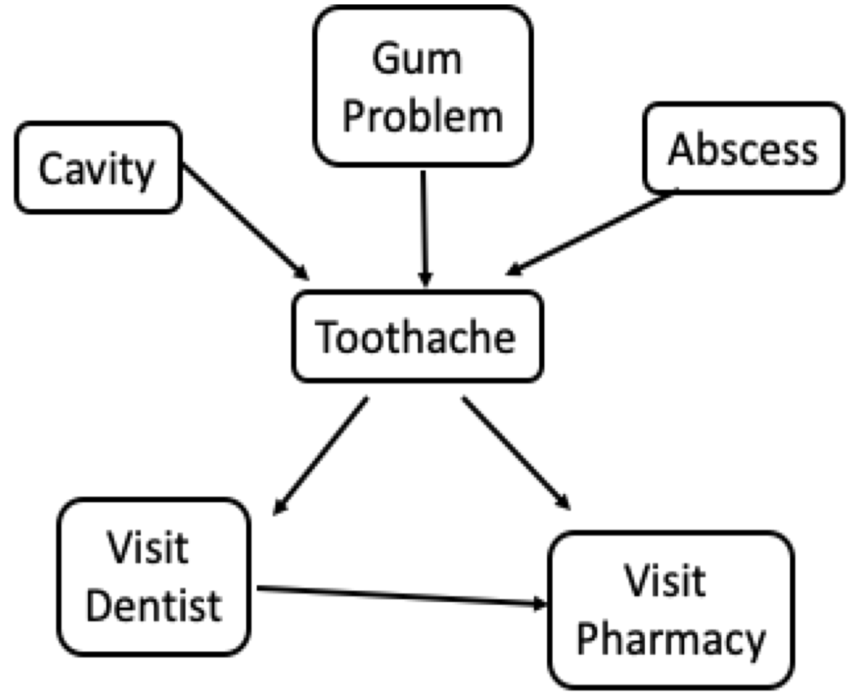
Describe Bayesian Networks

Objective

Illustrate key tasks in implementing Bayesian Networks

Why Do We Use Graphical Models?

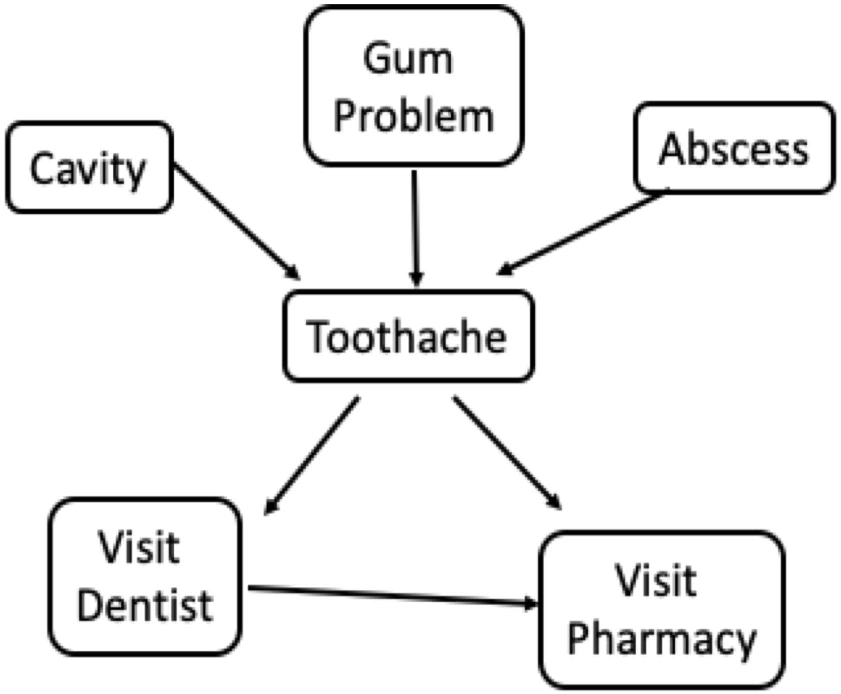
- In machine learning, we are often concerned with joint distributions of many random variables.
- A graph may provide an intuitive way to represent or visualize the relationships of the variables.
  - Much easier for domain experts to build a model.



Graphical Models for Causal Relations

- Graphical models arise naturally from, often causal, independence relations of physical events.

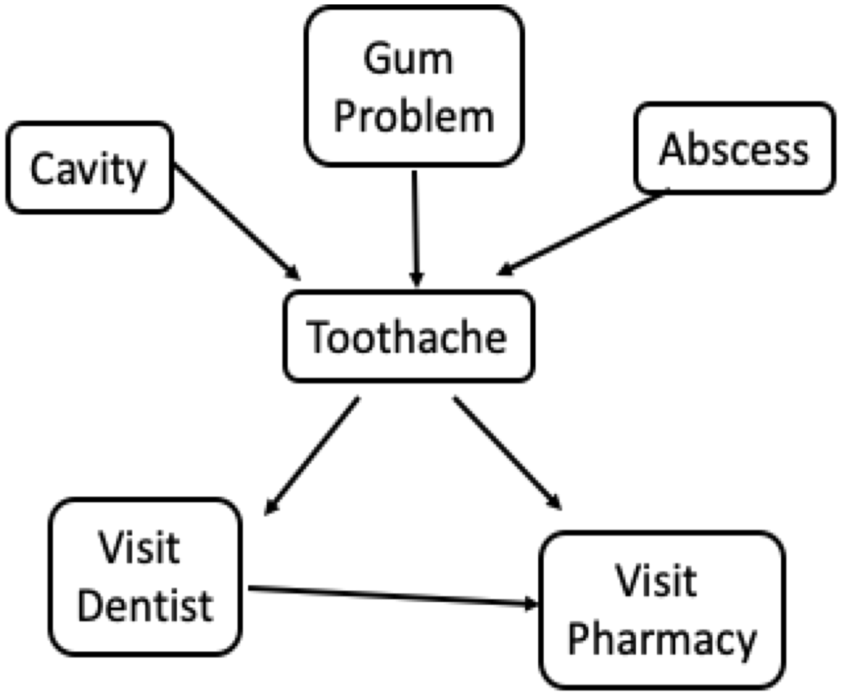
- Caveat: probabilistic relationship does not imply causality.



Representation of Bayesian Network (BN)

A BN is a directed acyclic graph (DAG), where…

- **Nodes** (vertices) represent **random variables**.
    - Can be assigned (observed) or unassigned (unobserved)
- **Directed edges** represent immediate **dependence** of nodes.
    - Indicate "direct influence" between variables.
    - Formally: encode conditional independencies.

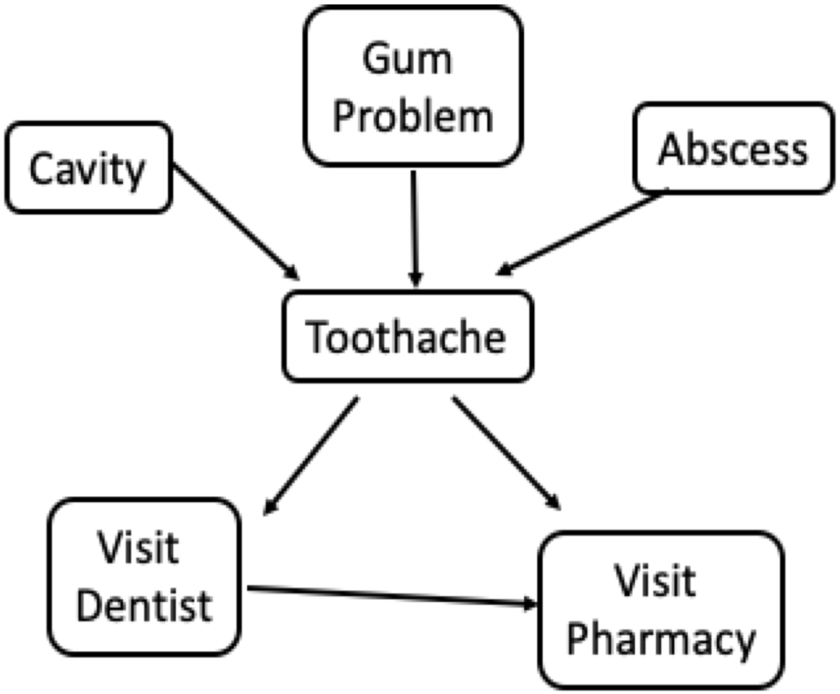


E.g., Toothache (+t) → Cavity (+c)?

Other names: Belief Networks, Bayes Nets, etc.

Size of BN

- **How many values must be stored using BN:**
  - $2 + 2 + 2 + 2^4 + 2^2 + 2^3 = 36$
- **We can save a lot of space!**
  - Space is affected by the largest Conditional Probability Table (CPT) of the BN.



Independence

Bayesian Network:

$$P(x_1, x_2, \ldots x_n) = \Pi_{i=1}^{n} P(x_i | \text{parents}(X_i))$$

The factorization adds additional constraints to $P(X_1, X_2, X_3, \ldots X_n)$.

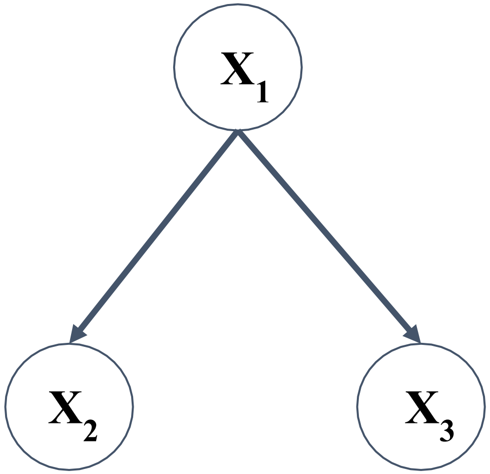
Independence

Bayesian Network:

$$P(x_1,x_2,x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1)$$

Chain rule:

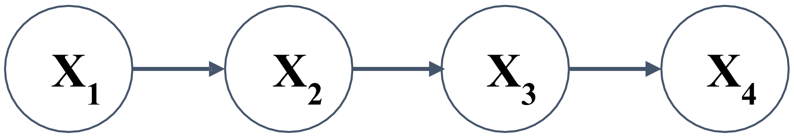
$$P(x_1,x_2,x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1)$$



$$X_3 \perp\!\!\!\perp X_2 \mid X_1$$

In BN, each variable x is independent of all its non-descendants given its parents.

Additional Conditional Independence



- Do we also have $X_1 \perp\!\!\!\perp X_3, X_4 \mid X_2$?
  - Yes! There is a proof for this:

$$P(X_1|X_2, X_3, X_4) = \frac{P(X_1, X_2, X_3, X_4)}{P(X_2, X_3, X_4)}$$

$$= \frac{P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_4|X_3)}{\sum_{x_1} P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_4|X_3)}$$

$$= \frac{P(X_1, X_2)}{P(X_2)} = P(X_1|X_2)$$

Note:

$$\sum_{x_1} P(X_1)P(X_2|X_1) = P(X_2) \; and \; P(X_1)P(X_2|X_1) = P(X_1,X_2)$$