Network Analysis -
Quiz 1

Q1. What is the difference between *in-degree* and *out-degree* in a directed graph?

A) Number of Edges pointing to the Node | into the node is Indegree
Number of edges going out or node points to is out-degree

Q2. Define *betweenness centrality* and explain what it reveals about a node's role in a network.

A) The number of shortest paths between pairs of other nodes that passes through this node reflects the role of node, bridge between the flow of information flow or control called betweeness Centrality.

Q3. What is meant by the *diameter* of a graph? How is it different from a *shortest path*?

A) the shortest path is the minimum distance between a specific pair of nodes but the diameter of a graph is the longest shortest path between any nodes in the graph.
we can say it as Global maximum.

Q4. What is meant by community detection in network analysis? Briefly describe.

A) It is the process of finding groups in a network where nodes are densely connected within and sparcely connected between. there are fwe methods to do this like modularity optimization and spectral clustering

Q5. **True or False:** If two directed acyclic graphs (DAGs) represent the same set of conditional independencies, then their Bayesian Networks define the same set of joint probability distributions.

False as Two DAG's may encode same conditional independences but still define different parametrizations which implies different joint distributions if the conditional probabilities

Q6. What are the two forms of "learning" required for a Bayesian Network? Briefly describe.

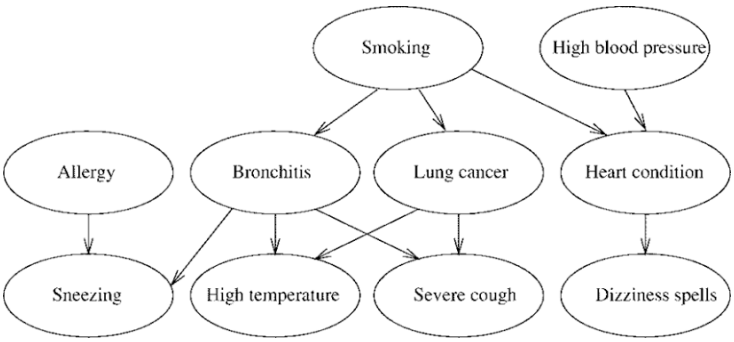
A) The two forms are

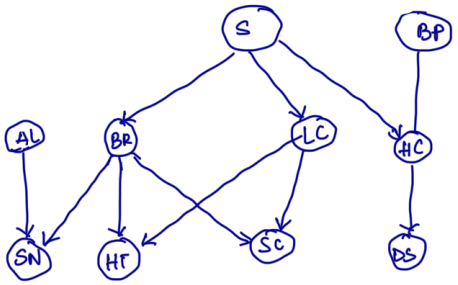
Struture learning
- learning graph structure from data
- discovering independicies and dependencies

Parmeter learning
- conditional probability tables
- maximum likelihood / Bayesian estimation are utilized

For Q7-Q10

The following network is a Bayesian Network inspired by Santos et al. (1998). **Use abbreviated variables as needed**, and define all symbols used in your responses to Questions 7-10 in a key (e.g., HT = High Temperature, BP = Blood Pressure, etc.).





Q7. Express the joint probability distribution of the network in a product of conditional probabilities (factored form). Ensure all parent-child relationships are captured.

$P(S, LC, BR, HC, BP, DS, AL, SN, SC, HT)$

$= P(S) \cdot P(LC|S) \cdot P(BR|S) \cdot P(HC|S) \cdot P(BP|HC) \cdot P(DS|HC) \cdot P(AL) \cdot P(SN|AL) \cdot P(SC|LC, BR) \cdot P(HT|LC, BR)$

Q8. List all v-structures in the Bayesian Network. Recall, a v-structure is a configuration of the form `X → Z ← Y` where `X` and `Y` are not directly connected.

LC → SC ← BR , LC → HT ← BR

lung cancer → severe cough ← Bronchitis , lung cancer → High temperature ← Bronchitis

Q9. Assume each variable in the Medical Diagnosis BN is binary. Specify the size of the conditional probability table (CPT) for each of the following variables:

Q8.1) High Temperature parents lung cancer, Bronchitis CPT size $= 2^2 = 4$ rows

Q8.2) Severe Cough parents lung cancer, Bronchitis cpt size $2^2 = 4$ rows

Q8.3) High Blood Pressure no parents only 1 child Heart condition
cpt size $= 2^0 = 1$ row

Q10. Based on the structure of the Medical Diagnosis BN:

a) Are *Smoking* and *Dizziness Spells* d-separated given *Lung Cancer*?

Smoking → Heart condition → Dizziness Spells No collider with lung cancer so not not D-separated

b) Are *Smoking* and *Dizziness Spells* d-separated given *Heart Condition*?

Smoking → Heart condition → Dizziness yes they are D-Separated

c) Are *Allergy* and *Severe Cough* d-separated given *High Temperature*?

Allergy → Bronchitis → Severe Cough , AL → BR → LC → SC so no not D separated with High temperature.

d) Is there an active path between *Heart Condition* and *Sneezing*?
If yes, provide one such path.

Heart condition ← Smoking → Bronchitis → Sneezing
Yes colliders in between Heart condition and Sneezing it is blocked at Smoking
V Structure and no other active paths