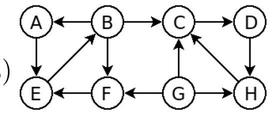
Essential Concepts

1 - *Directed Networks.* Consider the right figure, representing a simple directed network:



a) Identify all **strongly connected components** (SCCs) (and their correspondent nodes)

b) Redraw the network as a **DAG** (directed acyclic graph) of SCCs.

Graph Models

2 - Erdös-Rényi $G_{(n,p)}$ model (ER). Imagine you generate an ER graph with $n = 1000$ and $p = 0.01$.

a) What is the **average degree** of its nodes?

b) Draw a plot showing what you expect to be its **degree distribution**.

c) How many **connected components** do you expect the graph to have? Why?

d) Do you expect to find many big **cliques** on the network? Why?

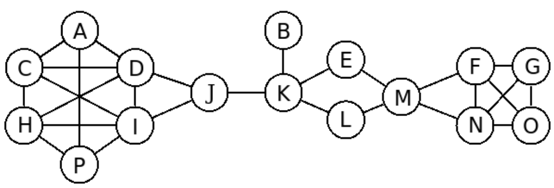
3 - Small-World Networks. Imagine you generate a small world network using the Watts-Strogatz (WS) model.

a) Which model generates a network with higher **clustering coefficient**: ER or WS? Why?

b) A purely regular network can have an high **average path length**. Explain how the WS model lowers this value.

Node Centrality

4 - Centrality Metrics. Consider the following network, and indicate:



a) a node with high degree but low betweenness centrality

b) a node with high closeness centrality but low betweenness centrality

c) a node with high betweenness centrality but low closeness centrality

d) two nodes with the same betweenness centrality

5 - The PageRank Algorithm.

a) Problems. Explain what a **spider trap** is, why it might be a problem for the PageRank algorithm and how we can avoid that problem.

Community Structure

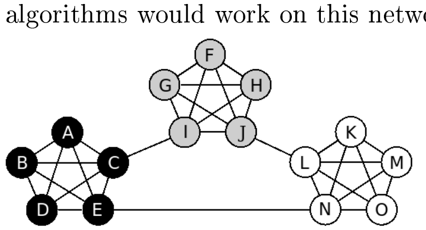
6 - Girvan-Newman and Louvain algorithms. Consider these two algorithms for community discovery.

a) Classify both algorithms in terms of being **divisive or agglomerative** .

b) Why can we say that both algorithms are **hierarchical**?

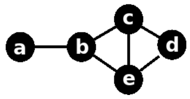
c) Which one is faster in terms of **execution time**?

d) Consider the following network with three very natural and intuitive communities. Explain how each of the two algorithms would work on this network, by explaining a couple of their initial greedy iterations.



Subgraph Patterns

For the following questions consider the graph G depicted in the following figure:



7 - *Subgraph census.* Calculate the **frequency of all possible induced subgraphs of size 3** on graph G (draw the topology of each possible subgraph and indicate all occurrences as sets of nodes).

8 - *Network Motifs.*

a) Explain the concept of **network motifs**. Indicate the necessary conditions for a subgraph to be considered a motif.

b) Explain how you could use network motifs to compare the **similarity of two networks**.

9 - *Graphlets and Orbits.*

a) Give a brief explanation of what a **orbit** is and indicate all possible orbits of graph G.

b) Calculate the **graphlet degree vector** of node c in graph G (consider graphlets up to size 3)

Network Construction

10 - *Multimode Network Transformation.* Imagine you have a bipartite network of genes and diseases.

a) If you construct a one mode projection on the diseases you will loose some information. Exemplify by **drawing two different initial bipartite networks that will give origin to the same projection** (suppose that a single common neighbor will give origin to an edge in the projection).

b) Explain the concept of the **Jaccard Index** and how you could use it to build network projections.

11 - Time series and Network Science:

a) Suppose you are using sensors to measure the air pollution of different locations (creating a time series for each location). **How you could build a network from this data?** (what are the nodes? what do the edges represent?)

b) Draw the natural **visibility graph** for the following time series with 8 equally spaced points in time: $6, 3, 8, 5, 4, 6, 7, 10$