**DSC190: Network Science and Graph Theory** **Winter 2025**

# Homework Assignment 1

**Topics:** (Matrix Formalism, Graph Representation, Network Metrics)

Instructor: **Prof. Gal Mishne**
Deadline: **Friday Jan. 17 (1/17/2025)**
Submission: **Gradescope**

Note: GenAI tools are permitted solely for problems: 4.

1. *Matrix Formalism*

   Let $\mathbf{A}$ be the $N$x$N$ adjacency matrix of an undirected unweighted network, without self-loops. Let $\mathbf{1}$ be the column vector of $N$ elements such that $\mathbf{1} = (\mathbf{1}, \mathbf{1}, ...., \mathbf{1})^T$ where the superscript $T$ indicates the transpose operation. Use the matrix formalism (multiplicative constants, multiplication row by column, matrix operations like transpose and trace, etc, but avoid the sum symbol $\Sigma$) to write expressions for:

   (a) The vector $k$ whose elements are the degrees $k_i$ of all nodes $i = 1, 2, ..., N$.

   (b) The total number of links, $L$, in the network.

   (c) The number of triangles $T$ present in the network, where a triangle means three nodes, each connected by links to the other two (Hint: you can use the trace of a matrix).

   (d) The vector $k_{nn}$ whose element $i$ is the sum of the degrees of node $i$'s neighbors.

   (e) The vector $k_{nnn}$ whose element $i$ is the sum of the degrees of node $i$'s second neighbors. Note: do not include a node among its own second neighbors.

2. *Graph Representation of Networks*

   The Adjacency Matrix ($\mathbf{A_{ij}}$) is a useful graph representation for many analytical calculations. However, when we need to store a network in a computer, we can save computer memory by offering the list of links in a $L$x2 matrix, whose rows contain the starting and end point $i$ and $j$ of each link.

   Calculate the following for the Networks $(a)$ and $(b)$ given below:

   (a) The corresponding adjacency matrices.

   (b) The corresponding link lists.

   (c) Determine the average clustering coefficient of Network $(a)$.

   (d) If you switch the labels of nodes 1 and 2 in Network $(a)$, how does that move change the adjacency matrix?

   (e) And the link list?

   (f) What kind of information can you not infer from the link list representation of the network that you can infer from the adjacency matrix?

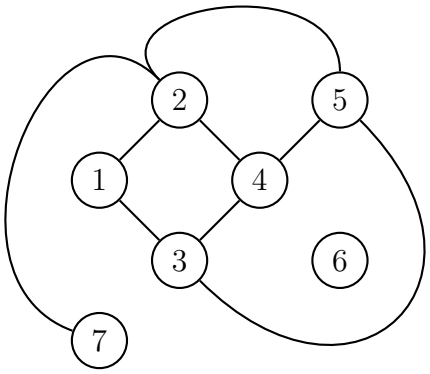


Figure 1: Network ($a$)

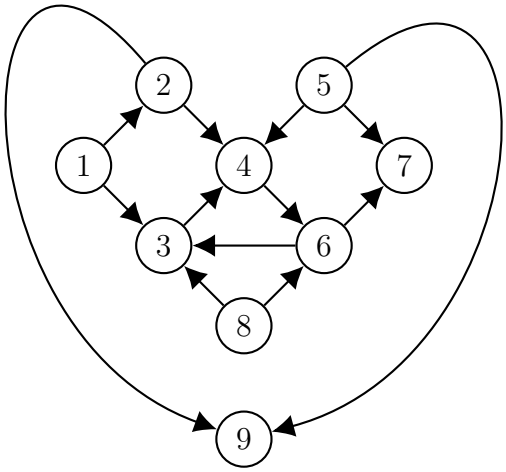


Figure 2: Network ($b$)

(g) In the ($a$) network, how many paths (with possible repetition of nodes and links) of length 4 exist starting from node 1 and ending at node 3? And in ($b$)?

3. *Degree, Clustering Coefficient and Components*

   (a) Consider an undirected network of size $N$ in which each node has degree $k = 1$. Which condition does $N$ have to satisfy? What is the degree distribution of this network? How many components does the network have?

   (b) Consider now a network in which each node has degree $k = 2$ and clustering coefficient $C = 1$. How does the network look like? What condition does $N$ satisfy in this case?

4. *Calculating Network Metrics (coding question)*

   Using *networkx* library in *python* and the graph network data from **Star Wars Episode IV Social Network**.
   Create the network based on the instructions in `star_wars_network_code.txt`,

   (a) Find the following network metrics.

      i. Degree Distribution P(k) v.s. Degree

      ii. Clustering Coefficient of all nodes

   (b) Plot the degree distribution.

   (c) Visualize the network using two different graph layouts and explain which one is more informative to understand the relationships between characters. Note: Include the character names in your plot.