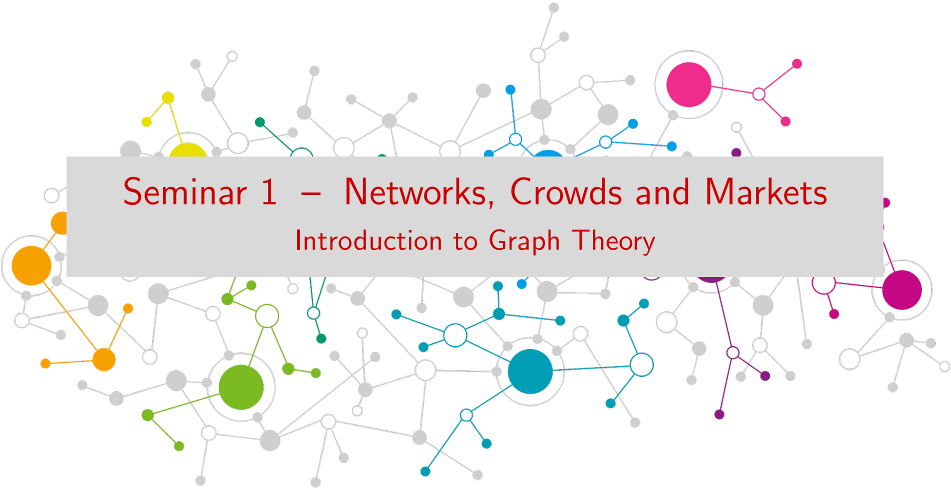
Seminar 1 – Networks, Crowds and Markets

Introduction to Graph Theory



Exercise 1. Adjacency and complements

Let B be the adjacency matrix of a simple *directed* graph G, and A the adjacency matrix of its *underlying undirected* graph (ignore arrow directions; collapse parallel arcs to a single edge). Decide if each statement is true or false and briefly justify.

(a) $a_{ij} = a_{ji}$ for $i \neq j$

(b) $b_{ij} \leq a_{ij}$ for $i \neq j$

(c) $b_{ij} + b_{ji} \leq 2$ for $i \neq j$

(d) $a_{ij} = b_{ij}$ for $i \neq j$

(e) In any simple *undirected* graph: $A(G^c) = J - I - A(G)$.

Exercise 2. Quick degree-sequence checks

Consider the multiset

$$\{0, 1, 1, 1, 2, 2, 2, 3, 6\}.$$

(a) Can this be a degree sequence of a simple undirected graph on $N = 9$ vertices? Justify briefly. (Hint: parity and sum check.)

(b) If yes, how many edges L does such a graph have? What is $\overline{deg}$?

(c) Intend to draw this graph.

Exercise 3: Paths and Connectedness

Show that if a graph (simple undirected) is not connected then its complement is. Provide an example of a four node network that is connected and such that its complement is also connected.

Hint: What happens in $\overline{G}$ is u, v lie in two different components of G? What happens if they lie in the same component?

Exercise 4. Degrees and handshaking

A simple undirected graph G has $N = 1000$ nodes and $L = 800$ edges.

(a) Compute the average degree $\overline{\deg}(G)$.

(b) You observe the degree sequence has N_0 isolated vertices and N_1 vertices of degree 1. Give two independent checks that your N_0, N_1 are consistent with N, L.
(Hint: handshaking and nonnegativity.)
- e.g. why $N_0 = 960$ is impossible?
- why $N_0 = 0$, $N_1 = 399$ is impossible?

Exercise 5: Comparing networks

Imagine two different undirected networks, each with the same number of nodes and links.

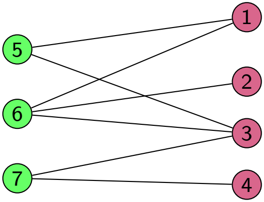
Questions:

1. Must both networks have the same maximum and minimum degree? Explain why or why not.
2. Must they have the same mean degree? Explain why or why not.

Exercise 6: Bipartite Networks

Consider the following bipartite network:

1. Construct its adjacency matrix A. Why the block diagonal blocks are zero?
2. Calculate the average degree of the purple nodes and the average degree of the green nodes in the bipartite network.
3. Without computing A^2 explicitly, what are its $(5, 1)$ and $(5, 6)$ entries?



Bipartite network with 6 purple and 5 green nodes, 10 links.

Exercise 7. Bipartite graphs

(a) Let a bipartite graph have parts of sizes m and n. Show that the maximum number of edges is mn, attained by $K_{m,n}$.

(b) Among all bipartite graphs on N vertices (where the partition is free to choose), show that the maximum number of edges is

$$L_{\max} = \left\lfloor \frac{N^2}{4} \right\rfloor,$$

attained by $K_{\lfloor N/2 \rfloor, \lceil N/2 \rceil}$.

Exercise 8: Netflix bipartite network

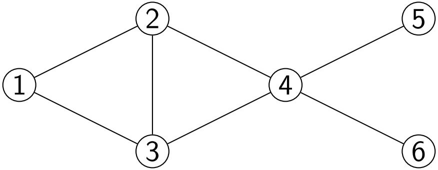
Netflix keeps data on customer preferences using a big bipartite network connecting users to titles they have watched and/or rated. Say, Netflix's movie library contains approximately 20,000 titles.

In the first quarter of 2025, Netflix reported having about 300 million users. Assume the average user's degree in this network is 100.

Questions:

1. Approximately how many links are in this network?
2. Would you consider this network sparse or dense? Explain.

Exercise 9. BFS levels and shortest paths (unweighted)



(a) Run BFS from node 1 by hand: write the level sets $L_0, L_1, L_2, \ldots$.

(b) Give a shortest path from 1 to 6. How many distinct shortest paths from 1 to 6 exist?

(c) Which vertices are at eccentricity equal to the graph diameter?

Exercise 9: Diameter and Degree

Consider a sequence of networks such that each network in the sequence is connected and involves more nodes than the previous network. Show that if the diameter of the networks is bounded, then the maximal degree of the networks is unbounded. That is, show that if there exists a finite number M such that the diameter of every network in the sequence is less than M, then for any integer K there exists a network in the sequence and a node in that network that has more than K neighbors.

Hint: Use the Moore bound.

Exercise 10: Facts about Trees

Recall that a tree is a connected graph with no cycles.

Show the following:

1. The minimal number of edges of a connected graph is $N - 1$.
2. A connected network is a tree if and only if it has $N - 1$ links.
3. A tree has at least two leaves, where leaves are nodes that have exactly one link.
4. In a tree, there is a unique path between any two nodes.

Optional take-home: NetworkX mini-lab

Paste this into Colab:

```
# 1) Setup
!pip -q install networkx matplotlib
import networkx as nx
import matplotlib.pyplot as plt

# 2) Build a small graph by hand
G = nx.Graph()
G.add_edges_from([(1,2),(1,3),(2,4),(3,4),(4,5),(5,6)])

# 3) Basic stats
N, L = G.number_of_nodes(), G.number_of_edges()
print(f"Nodes: {N}, Edges: {L}, Avg degree: {2*L/N:.2f}")

# 4) Degree sequence
deg = [d for _, d in G.degree()]
print("Degree sequence:", sorted(deg))

# 5) Shortest paths from node 1 (BFS)
lengths = nx.single_source_shortest_path_length(G, 1)
print("Distances from 1:", dict(lengths))

# 6) Draw
pos = nx.spring_layout(G, seed=7)
nx.draw(G, pos, with_labels=True, node_color="lightblue",
        node_size=700, edge_color="gray")
plt.show()
```

Stretch goals: change the edge set, recompute degree sequence and distances; add a node with many edges and see how the layout and degree stats change.