Networks theoretical practice problems no.1.

1. During the lectures we have learned that the **adjacency matrix** provides a natural mathematical representation for a network or a graph. Based on A_{ij} we can calculate quite a few quantities describing various properties of the studied network. Here are two simple examples:

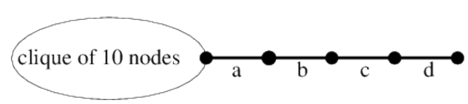
   a) Provide a formula giving the total number of links as a function of the adjacency matrix, (treating the directed and undirected cases separately).

   b) Provide a formula giving the total number of paths of length ℓ between nodes i and j as a function of the adjacency matrix.

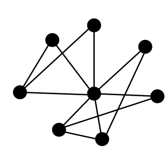
   WARNING!! I am asking for mathematical formulas, **not algorithms**. A formula is expressing a relation between different quantities in a concise way via symbols, and usually takes up no more than a single line. E.g., the formula for the number of links starting from node i, (denoted by $k_i^{\rm out}$), as a function of the adjacency matrix reads $k_i^{\rm out} = \sum_{j=1}^{N} A_{ij}$.

2. We have learned how to characterize nodes (or links) in a network by different measures such as centralities, clustering coefficient, etc. Let us practice the evaluation of these quantities with the help of the following exercises.

   a) Which of the shown links has the highest betweenness centrality? Why?



   b) Calculate the clustering coefficient for the central node:



   c) What is the average clustering coefficient in an undirected Erdős–Rényi graph with $N = 100$ nodes and $M = 990$ links?

3. We have learned that the degree distribution is one of the most important distributions characterising the statistical properties of a network. What is the degree distribution for the following networks?

   a) an undirected chain of 100 nodes,

   b) a balanced undirected binary tree of 127 nodes.
   (In this tree we find always two „children" linked under a given node, except for the leafs at the bottom of the tree, which have no descendants at all.)

4. Draw a small network having the following degree distribution:

$$
\begin{aligned}
p(k=1) &= 0.4\\
p(k=2) &= 0.3\\
p(k=3) &= 0.2\\
p(k=4) &= 0.1
\end{aligned}
$$

5. The closeness centrality C_c we discussed on the lectures is not robust against the network falling to seprate (unconnected) components: The distance between node i and any other node j in a different component becomes $l_{ij} = \infty$, and thus, the closeness of i becomes $C_c(i) = 0$.

   Motivated by this, cook up an alternative definition of the closeness with the following properties:

   - it should give higher values for nodes intuitively closer to the rest of the network, i.e., if for node i the average distance to the rest of the nodes is smaller than for node j, then this implies $C_c(i) > C_c(j)$.
   - it should give a meaningful result for $C_c(i)$ even when there are unreachable nodes from i, (for which $l_{ij} = \infty$).

6. What is the maximum possible number of strongly connected components in a directed network of N nodes, where at least one link exists between every node pairs? (Note that even a single node can be considered as a strongly connected component). Please draw a figure of the network with the max. number of strongly connected nodes in case of $N = 4$ nodes.

7. What is the probability distribution for the **number of links** in an undirected Erdős-Rényi graph with N nodes and p linking probability? (Equivalently: what is the probability for having exactly M links in an Erdos-Rényi graph with N nodes and p linking probability?)