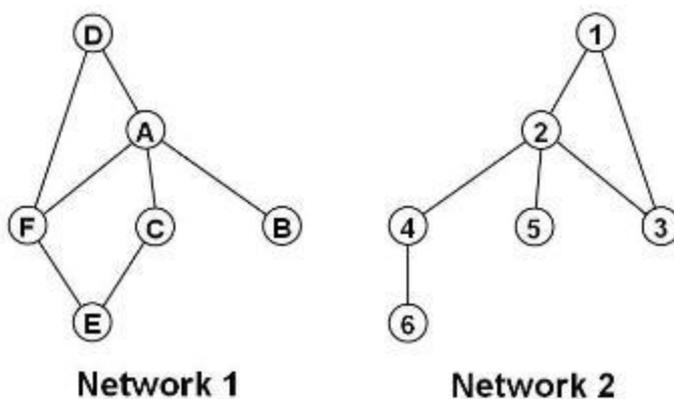


Hands-on practical: Performing network alignment via the Stable Marriage Algorithm (SMA)

The figure below represents two protein-protein interaction (PPI) networks from two different organisms. In such networks, nodes and edges respectively represent proteins and the interactions between them. Your task here is to align these networks using the stable marriage (Gale-Shapley) algorithm.

Network alignment is about finding correspondence between the topology of two (or more) networks. Alignment of biological networks is only meaningful if they are homologous, implying that corresponding proteins have similar functions and interactions. Alignments of biological networks can be useful if we know a lot about some of the nodes in one network and almost nothing about topologically similar nodes in the other network. In that case, specialized knowledge about one may tell us something new about the other.



Step A. Calculate the clustering coefficient C for each node in both networks, using this formula:

$$C = \left\{ \begin{array}{ll} 2E / (N(N-1)) & \text{if } N > 1 \\ -1 & \text{if } N = 1 \end{array} \right\}$$

Here, N is the number of neighbours and E is de number of edges (links) between those neighbours. To answer this question, you may copy the tables below and fill in the calculated clustering coefficients.

Network 1	
A	
B	
C	
D	
E	
F	
Network 2	
1	
2	
3	
4	
5	
6	

Step B. Make a preference list for each node, in which the preference to the nodes in the other network is ranked. This ranking must be done based on the absolute difference in clustering coefficients (calculated in step A). So if a node in the other network has an equal clustering coefficient, it should be ranked first. In case of ties, when a node prefers two other nodes equally well, you should order those nodes by their names (*i.e.* 'A' before 'B' and '1' before '2'). To answer this question, you may copy the tables below and fill in the ranked node names.

Preferences Network 1						
A						
B						
C						
D						
E						
F						
Preferences Network 2						
1						
2						
3						
4						
5						
6						

Step C. Perform the Gale-Shapley algorithm to align Network 1 and 2. Use the preferences you obtained in step B, and assume that the nodes in Network 1 and 2 are respectively male and female. Briefly write down the decision for each step of the algorithm.

Step D. Provide the final network alignment. To answer this question, you may copy the table below and fill in the mapping of the proteins.

A	B	C	D	E	F

Step E. The final step is to infer biological knowledge from the network alignment. Given that Network 1 has been fully experimentally verified, and that Network 2 is still provisional, which extra protein interaction could you predict in Network 2? Also explain why!

Some further questions about the stable marriage algorithm (SMA):

1. Is the SMA order dependent? How can you test this?
2. It is claimed that the Gale-Shapley (SMA) pairing, in the form used here, is male-optimal and female-pessimal (it would be the reverse, of course, if the roles of "male" and "female" participants in the algorithm were interchanged). In other words, the algorithm is sexist. How can you test this?