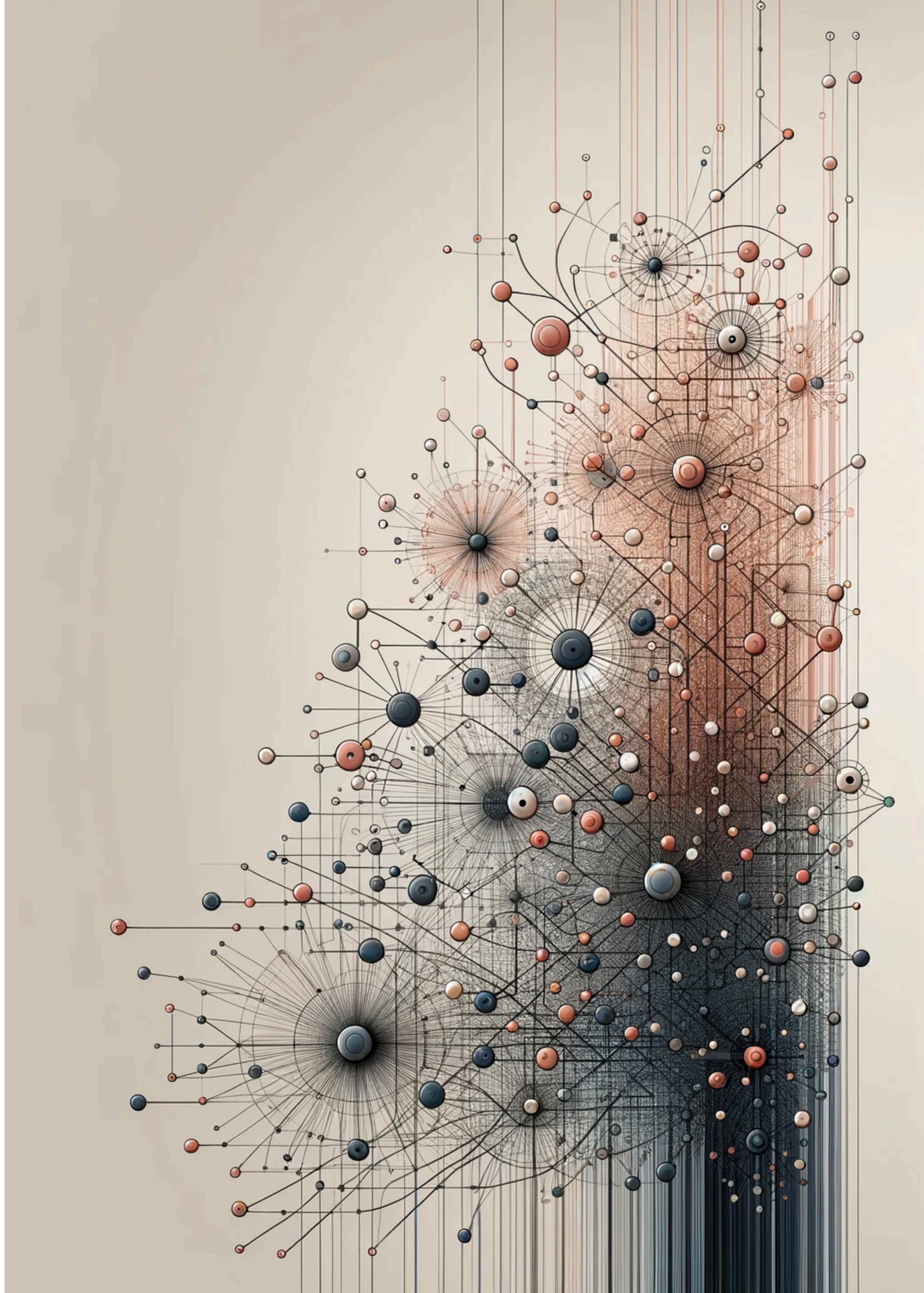
UNIVERSITÀ DI TORINO

Analisi e Visualizzazione delle Reti Complesse

NS05 - Complex Network Analysis

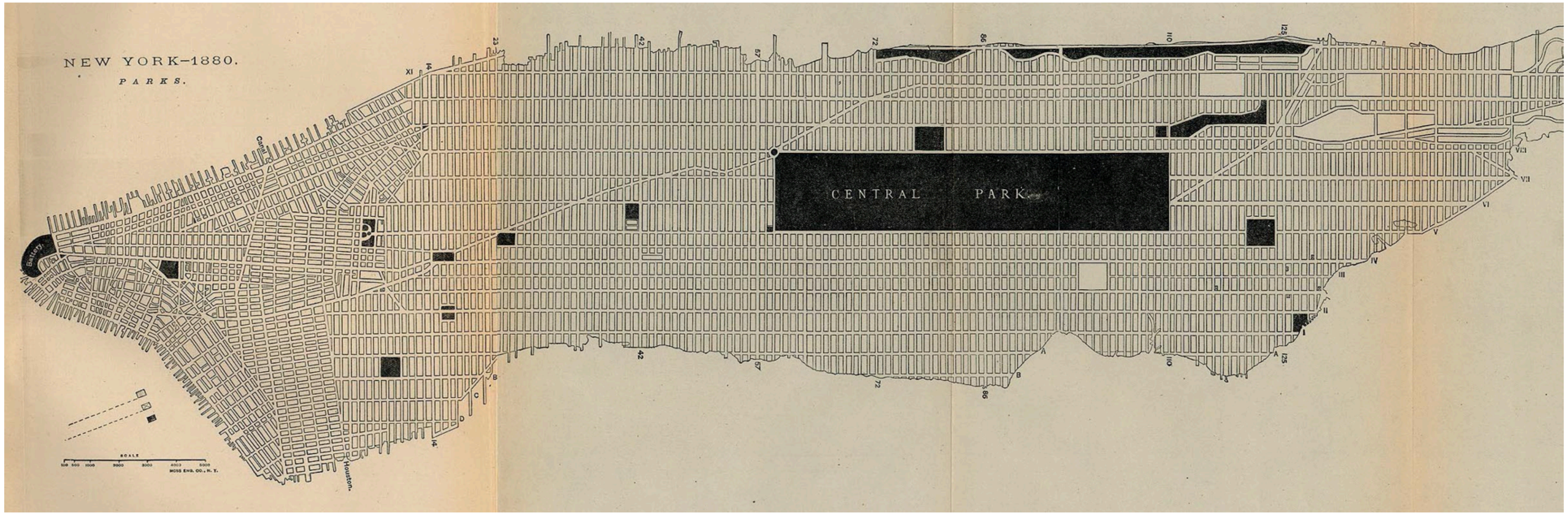
Prof. Rossano Schifanella



Exercises

Exercise 1

Consider the road map figure



Exercise 1

- If one were creating a network representation of traffic patterns, which of the following would be the best choice to make up the links of the network?
  - Hint: your answer to the next question may inform your answer to this question, and vice-versa.

1. Road segments, e.g., 5th Ave. between 12th and 13th streets
2. Pedestrians traveling along the streets
3. Vehicles traveling on the roads
4. Entire roads, e.g., 5th Ave.

UNIVERSITÀ DI TORINO

Exercise 2

Consider again the road map shown in previous question.

- In a network representation of traffic patterns, which of the following would be the best choice to make up the nodes of the network?
- Hint: your answer to the previous question may inform your answer to this question, and vice-versa.

a. Pedestrians moving along the streets

b. Street intersections, e.g., 5th Ave. and 12th St.

c. City blocks, e.g., the block between 5th-6th avenues and 12th-13th streets

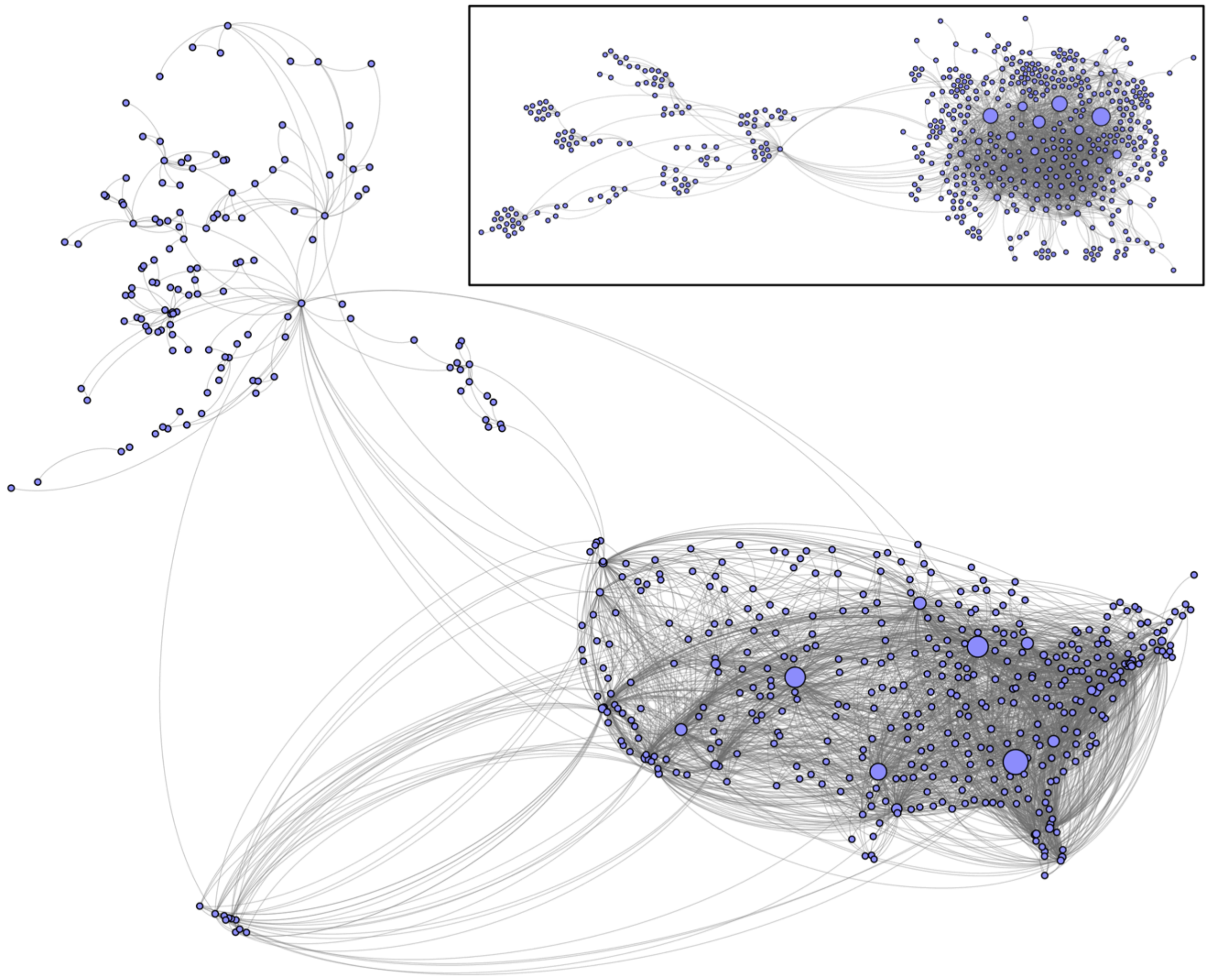
d. Vehicles traveling on the roads

UNIVERSITÀ DI TORINO

Exercise 3

Consider the US air transportation network shown on the right

- Nodes in this network represent airports. What could a link between two airports represent?



UNIVERSITÀ DI TORINO

Exercise 4

Compare the US air transportation network in Exercise 3 with the Manhattan road map in Exercise 1 and 2.

The air transportation network displays a distinguishing feature that the Manhattan road network lacks.

- What is this key characteristic?
  i. Hub nodes with many links
  ii. Singleton nodes with no links
  iii. Nodes with more than one connected link
  iv. Multiple routes between nodes

Exercise 5

In a social graph from Facebook, which type of link best represents the **friend** relation? Directed or undirected?

Exercise 6

In a social graph from Twitter, which type of link best represents the **follower** relation? Directed or undirected?

Exercise 7

Consider a network with N nodes.

Given a single link, what is the maximum number of nodes that link can connect?

Given a single node, what is the maximum number of links that can connect to that node?

1. A link connects two nodes. A node can be connected to at most N other nodes.
2. A link connects two nodes. A node can be connected to at most two other nodes.
3. A link connects one nodes. A node can be connected to at most $N-1$ other nodes.
4. A link connects two nodes. A node can be connected to at most $N-1$ other nodes.

Exercise 8

Consider a **directed** network of N nodes.

Now consider the total **in-degree**, i.e., the sum of the in-degree over all nodes in the network. Compare this to the analogous total **out-degree**.

Which of the following must hold true for any such network?

1. Total in-degree must be greater than total out-degree
2. Total in-degree must be less than total out-degree
3. Total in-degree must be equal to total out-degree
4. None of these hold true in all instances

Exercise 9

Consider a **Twitter retweet network**, where users are nodes, and we want to show how many times a given user has retweeted another user.

What link type best captures this relation?

1. Directed, weighted
2. Undirected, weighted
3. Undirected, unweighted
4. Directed, unweighted

Exercise 10

Given a complete network A with N nodes, and a bipartite network B also with N nodes, which of the following holds true for any $N > 2$

1. None of these hold true for all such $N > 2$
2. Network A has fewer links than network B
3. Network A has more links than network B
4. Network A has the same number of links as network B

Exercise 11

Consider this adjacency matrix:

An entry in the i^{th} row and j^{th} column indicates the weight of the link from node i to node j.

For instance, the entry in the 2^{nd} row and 3^{rd} column is 2, meaning the weight of the link from node B to node C is 2.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & 1 & 3 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

What kind of network does this matrix represent?

1. Directed, unweighted
2. Directed, weighted
3. Undirected, weighted
4. Undirected, unweighted

Exercise 12

Consider the network defined by the adjacency matrix shown in the previous Exercise. A sink is defined as a node with in-links but no out-links.

Which nodes in the network, if any, have this property?

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

$$
\begin{array}{c|c}
 & \begin{matrix} A & B & C & D & E & F \end{matrix} \\
\begin{matrix} A \\ B \\ C \\ D \\ E \\ F \end{matrix} &
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
2 & 1 & 3 & 1 & 1 & 0
\end{pmatrix}
\end{array}
$$

Exercise 13

Convert the network defined by the adjacency matrix used in the previous exercises to an undirected, unweighted graph.

- When converting a directed graph to an undirected one, nodes i and j are connected in the undirected graph if there is a directed link from i to j, or from j to i, or both.
- Print out the resulting matrix and/or draw a network diagram for reference.

How many nodes are in this converted network? How many links?

1. N = 6, L = 9
2. N = 6, L = 10
3. N = 7, L = 10
4. N = 5, L = 9
5. N = 5, L = 10

Exercise 14

Any directed network of N nodes with at least N links must contain a cycle.

1. True
2. False

Exercise 15

Consider an arbitrary non-complete undirected network.

Now add a single link.

How has the number of nodes in this network's giant component changed as a result of this addition?

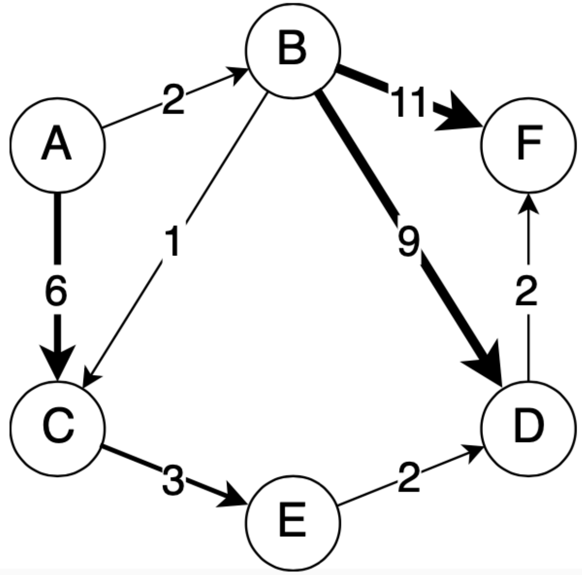
1. It has strictly increased
2. It has strictly decreased
3. It has decreased or stayed the same
4. It has increased or stayed the same

Exercise 16

Consider the weighted directed network in the figure.

Which of the following most accurately describes the connectedness of this network?

1. Disconnected
2. None of the others
3. Weakly connected
4. Strongly connected

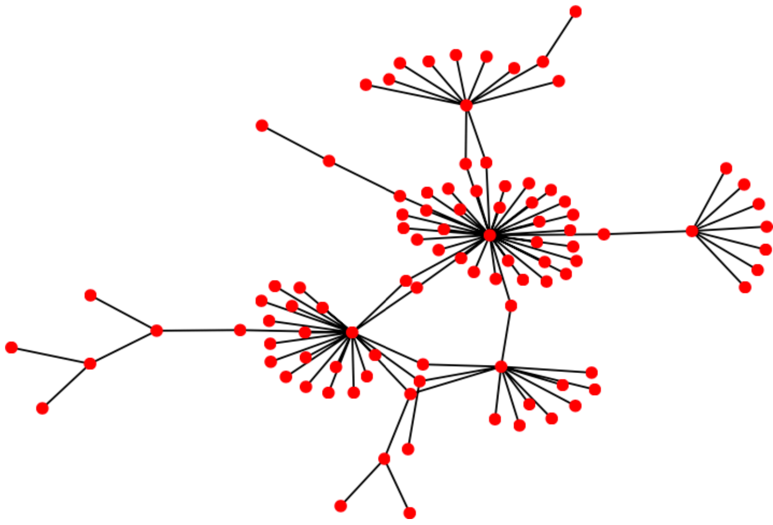


Exercise 17

Consider the network in the figure.

Which of the following is the best estimate of the diameter?

- Note: the figure shows a small sub-network of the Drosophila melanogaster (a.k.a. fruit fly), protein interaction network.
- Each node represents a protein that interacts with other proteins to perform the essential work of the cell.
- Experimental evidence has demonstrated that linked proteins form a molecular bond to accomplish some biological function.



1. 2
2. 4
3. 20
4. 10

Exercise 18

What is the central idea behind the notion of “six degrees of separation” (or also "the small world phenomenon")?

1. Social networks are sparse.
2. Social networks have small average path length.
3. Social networks have many high-degree nodes.
4. Social networks have high clustering coefficients.

Exercise 19

Which of the following seemingly conflicting properties are true of social networks?

1. Social networks have short paths, yet large diameter.
2. Social networks have many high-degree nodes, yet are disconnected.
3. Social networks are highly clustered, yet are not dense.
4. Social networks have small diameter, yet large average path length.

Exercise 20

We said that Kevin Bacon is the Center of the Universe, as probably any other popular actor in the network: it is very easy to find a very short path from anyone else to Kevin Bacon as well as any other famous actor. Conversely, you are asked to check yourself how much is difficult to find the opposite situation!

- Can you find two stars separated by more that four links? Play the game and try!
- Go to the Oracle of Bacon , try your pairs of actors, and when you find a path longer than four links, just complete your assignment!

Exercise 21

Download the file *lesmiserables_edges.csv*.

It contains the edges list representing the famous "Les Miserables" graph, where each node identifies one of the characters of the novel, and a link is established between two nodes if corresponding characters are mentioned in the same page at least once. Weights take into account the number of mentions' occurrences

- Can you draw a graph representing this dataset?
- You can use scripts if you are familiar with some languages or, considered the limited size of the graph, you can also try to draw it with your pencil.
- Try to calculate the degree correlation function knn(k) and to plot the set of points (k, knn(k)). Is it possible to say if the graph is assortative, disassortative, or neutral?

Exercise 22

Consider the social network represented in Figure.

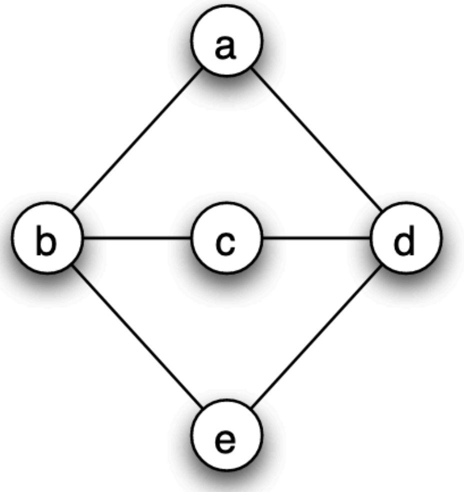
Suppose that this social network was obtained by observing a group of people at a particular point in time and recording all their friendship relations.

Now suppose that we come back at some point in the future and observe it again.

According to the theories based on empirical studies of triadic closure in networks, which new edge is most likely to be present?

- i.e., which pair of unconnected nodes are most likely to be linked by an edge when we return to take the second observation?

Also, give a brief explanation for your answer.

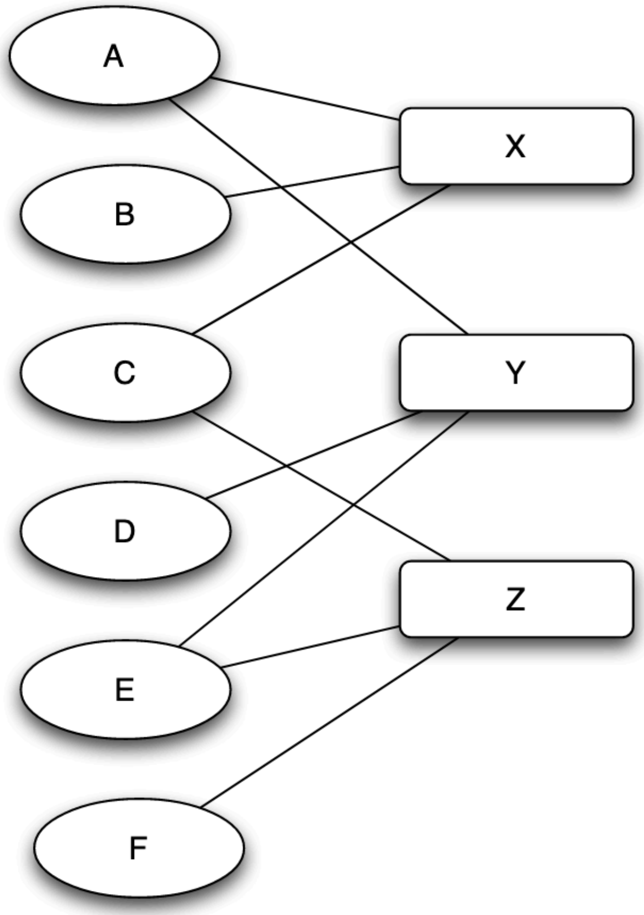


UNIVERSITÀ DI TORINO

Exercise 23

Consider the affiliation network in figure, with six people labeled A–F, and three foci labeled X,Y, and Z.

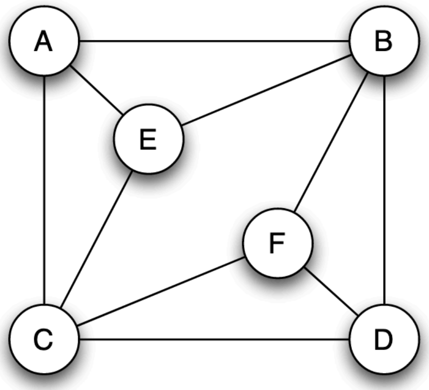
1. Draw the derived network (or projection) on just the six people, joining two people when they share a focus.
2. In the resulting network of people, can you identify a sense in which the triangle on the nodes A, C, and E has a qualitatively different meaning than the other triangles that appear in the network? Explain.



Exercise 24

Given a network showing pairs of people sharing activities, we can reconstruct an affiliation network consistent with this data.

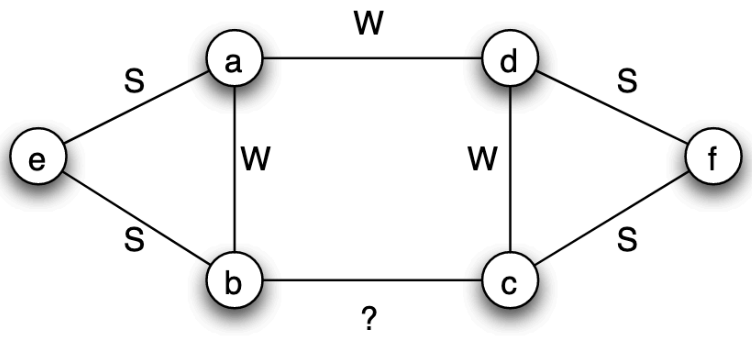
- For example, suppose that you are trying to infer the structure of a bipartite affiliation network. By indirect observation, you've obtained the projected network on just the set of people: an edge is joining each pair of people who share a focus. This projected network is shown in the Figure.



1. Draw an affiliation network involving these six people, together with four foci you should define, whose projected network is the graph shown in the figure.
2. Explain why any affiliation network capable of producing the projected network in the figure must have at least four foci.

Exercise 25

Consider the graph in the Figure below, where each edge — except the edge connecting b and c — is labeled as a strong tie (S) or a weak tie (W).



According to the theory of strong and weak ties, with the strong triadic closure assumption, how would you expect the edge connecting b and c to be labeled?

Give a brief (1-3 sentence) explanation for your answer.

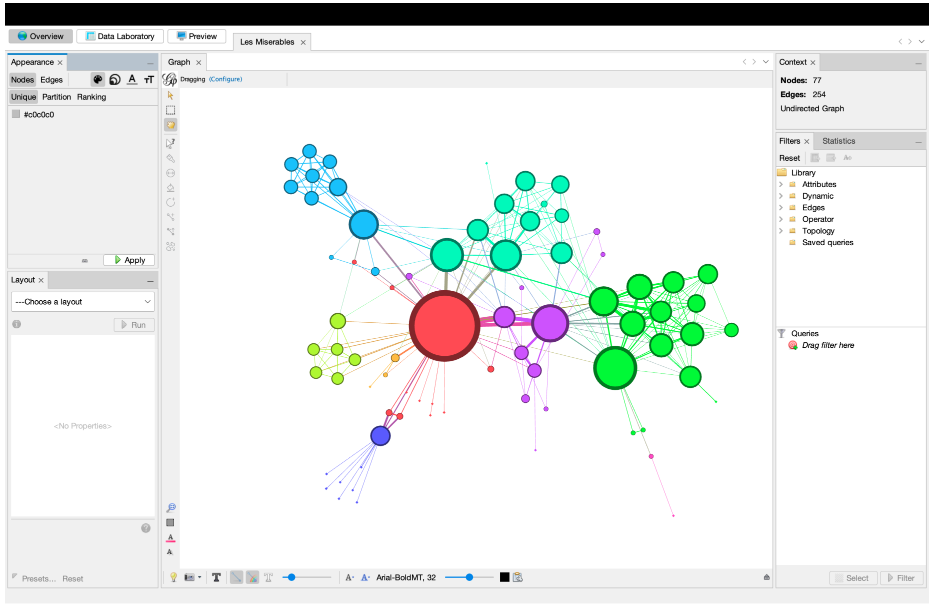
CNA: Complex Network Analysis

Network Analysis software

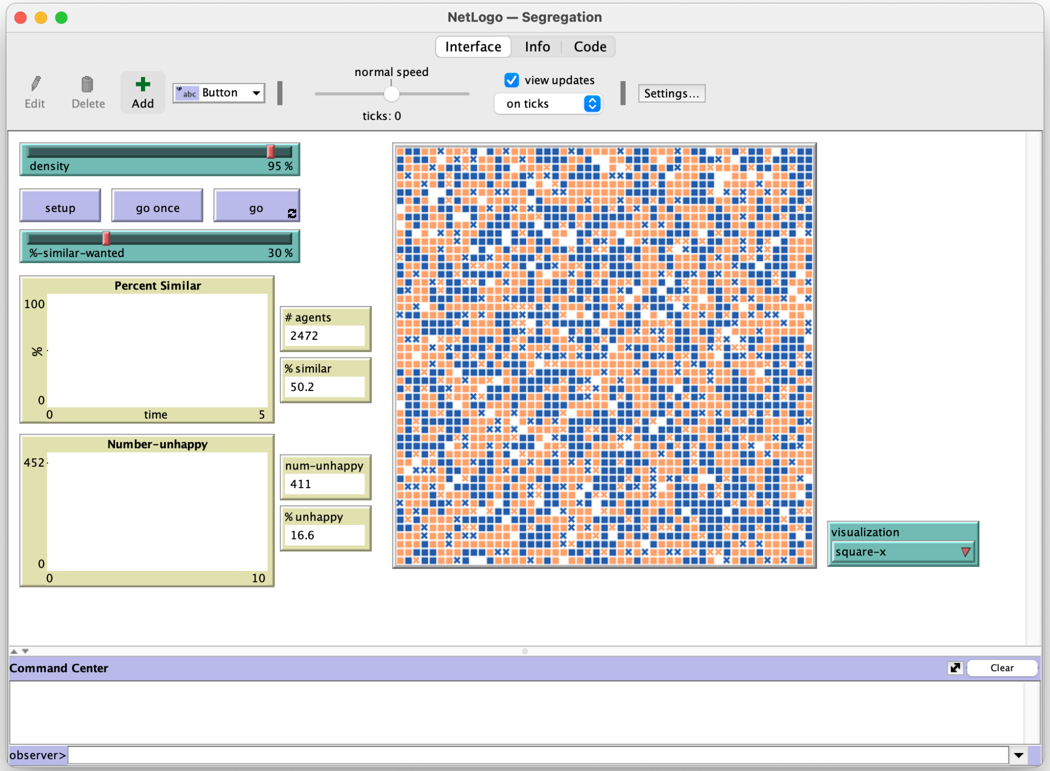
- **Python 3 + Project Jupyter/JupyterLab**
  - Alternatively: R
  - The use of the R ecosystem is out of the scope of this course; however, you can replicate all the practical sessions with alternative computational frameworks
- Students can pick the platform of their preference:
  - local distribution: Anaconda or base Python + pip
  - cloud distributions (many free): e.g., Google Colab, Binder Kaggle Kernels, Azure Notebooks,Datalore
- **NetworkX**
  - Even if there are other very valid alternative packages, e.g., igraph, graph-tool, this course will provide solutions to the practical sessions using networkX.

Other software and tools

- Gephi
  - Tutorials and learning resources



- NetLogo
  - Beginner's guide



Your first iPython notebooks

Official NetworkX Tutorial

Official NetwrokX Guides

Tutorials are taken or adapted from A First Course in Network Science - Tutorials

- Go to the **tutorials** folder:
    - 01-intro-graphs.ipynb
    - 02-paths-components.ipynb

If you need a refresher on Python basics, the book provides one:

- Appendix - Python Tutorial.ipynb

Reading material

References

[ns1] **Chapter 1 & 2 - References to Python scripts and exercises**

[ns1] **Appendix A: Python Tutorial**

Q&A