

Introduction and housekeeping

Using Networks to Study Microbes

Alberto Santos & Yesid Cuesta-Astroz

Outline

- Who are you?
- Introduction to the course
- Course schedule
- Tools
- Contact

Who are you?

Alberto Santos - Associate Professor and Head of the Data Science Platform at the Danish Technical University (DTU) - Denmark

- Research Group Lead **Multi-omics Network Analytics** (MoNA )

- **Multi-omics**

Using multimodal data to have a comprehensive view on (micro) biology problems

- **Network**

Exploiting graphs to structure, represent, integrate and analyse data

- **Analytics**

Applying Data Science and Machine Learning to answer complex biological questions

Who are you?

Yesid Cuesta-Astroz

Who are you?

Introduction

Introduction to the Course

Networks to Study Microbes

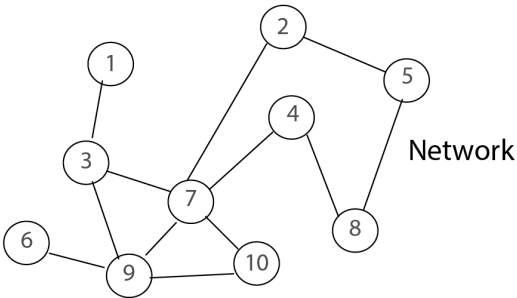
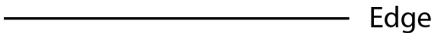
- Technological advances are making **large-scale omics datasets** available to **study microbes and microbial communities**
- We need **computational methods** to **model and understand** these complex data
- This course will provide an overview of **omics data types** and **data resources** available and how to **integrate, analyse** and **interpret** these data with **networks**

Introduction to the Course

What are networks?

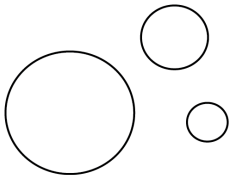
- **Data structures** of components (**nodes**) connected by relationships (**edges**)
- These structures allow:
 - Quick **integration** of **heterogeneous data** based on relationships
 - **Graph theory** methods can be used to **analyse** and **interpret** data, e.g., topological properties can be used to explain:
 - The possible **role** of specific components
 - The **flow** of information
 - The **robustness** of the system
- **Visualize** data

Networks

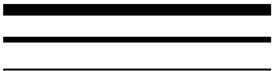


0	0	1	0	0	0	0	0	0	0
0	0	0	0	1	0	1	0	0	0
1	0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	1	1	0	0
0	1	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	1	0
0	1	1	1	0	0	0	0	1	1
0	0	0	1	1	0	0	0	0	0
0	0	1	0	0	1	1	0	0	1
0	0	0	0	0	0	1	0	1	0

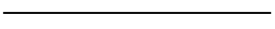
Adjacency matrix



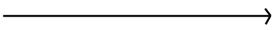
weighted nodes (size)



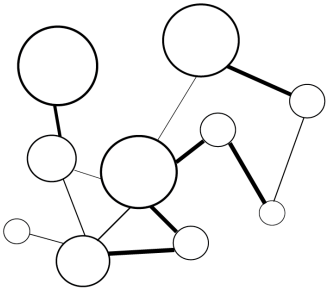
weighted edges (thickness)



undirected edge



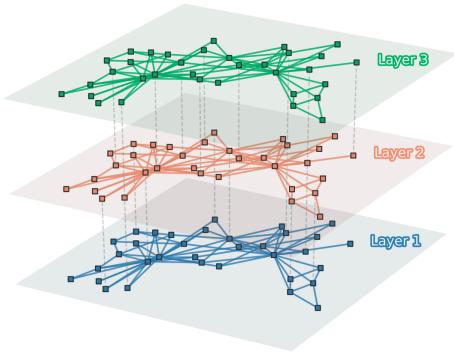
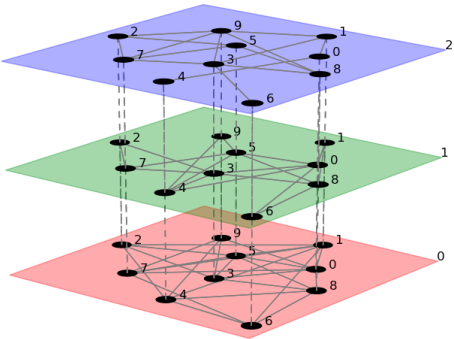
directed edge



weighted undirected network (thickness)

0	0	%	0	0	0	0	0	0	0
0	0	0	0	%	0	%	0	0	0
%	0	0	0	0	0	%	0	%	0
0	0	0	0	0	0	%	%	0	0
0	%	0	0	0	0	0	0	%	0
0	0	0	0	0	0	0	0	0	%
0	%	%	%	0	0	0	0	%	%
0	0	%	%	%	0	0	0	0	0
0	0	0	0	0	0	%	0	0	%
0	0	0	0	0	0	0	0	%	0

Weighted adjacency matrix



Schedule

Time	DAY 1	DAY 2	DAY 3
8:30-9:20	Introduction and Housekeeping	Working with Data in Python I	Analysing Networks I
9:20-10:10	An Omics View on Microbes I	Working with Data in Python II	
10:10-10:30	Coffee break	Coffee break	Coffee break
10:30-11:20	An Omics View on Microbes II	Visualizing Data in Python	Analysing Networks II
11:20-12:10	Data Resources	Visualising Networks I	
12:10-13:30	Lunch	Lunch	Lunch
13:30-14:20	Introduction to Python I	Visualising Networks II	Team Project
14:20-14:40	Coffee break	Coffee break	Coffee break
14:40-16:00	Introduction to Python II	Network Exercises	Team Project
16:00-16:50	Recap and Q & A	Recap and Q & A	Team Project Presentations and Q&A

Tools

Basic Programming

Python



- Python is a **high-level, general-purpose programming language**
- It is a good language for both **beginners** and **advanced programmers**:
 - Its design philosophy **emphasises code readability**
 - Many **learning resources** available
 - **Large community** behind (i.e. stackoverflow for questions)
 - **Excellent Libraries** available, especially for data science

Jupyter Notebooks

<https://jupyter.org/>



- **Web-based development environment** for **creating, running** and **sharing Python** (and other languages) code
- A **notebook** is an interactive document that combines **live code, equations, text or markdown**, and **visualisations** (output of your code)
- Notebooks are divided into **cells** that **run sequentially!** (Need to pay attention)
- It **requires** having **Python installed** on your local machine



Colab Notebooks

<https://research.google.com/colaboratory/faq.html>

- Google Colab is based on **Jupyter Notebook** open source project **hosted on Google's servers**
- Advantages:
 - Requires **no setup** to use (no python installation)
 - Provides **free** access to **computing resources** on Google's servers including GPUs
 - **Notebooks** can be **shared** just as you would with Google Docs or Sheets.
 - You can **import** existing **Jupyter notebooks**
- **Own data and notebooks** need to be accessed through **Google Drive** — Need Google account

Cytoscape

<https://cytoscape.org/>



- An open source **software platform** for **visualising** and **analysing** complex **networks**
- Used for any kind of networks but **specialised** on **biological domains**:
e.g, Molecular interaction networks and biological pathways and integrating these networks with annotations, gene expression profiles and other state data.
- **Additional features** are available as freely available **Apps** (<https://apps.cytoscape.org/>)

Relevant Links

Extra readings and resources

- **Network biology: understanding the cell's functional organisation.** Albert-László Barabási & Zoltán N. Oltvai. Nature Reviews Genetics 2004 (<https://www.nature.com/articles/nrg1272>)
- **Network analysis of protein interaction data** — Online tutorial — EMBL EBI (<https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/>)
- **Network Biology: A short introduction to the core concepts** <https://www.youtube.com/watch?v=H1bGk8PGvf8>
- **Network Visualization: a short introduction to the core concepts of network layout and clustering** <https://www.youtube.com/watch?v=OunX7ISRbgA>
- **A Guide to Conquer the Biological Network Era Using Graph Theory.** Mikaela Koutrouli, Evangelos Karatzas, David Paez-Espino, Georgios A. Pavlopoulos. Frontiers in Bioengineering and Biotechnology 2020 (<https://www.frontiersin.org/articles/10.3389/fbioe.2020.00034/full>)
- **Python for Beginners** <https://www.python.org/about/gettingstarted/>
- **How to Use Jupyter Notebook: A Beginner's Tutorial** <https://www.dataquest.io/blog/jupyter-notebook-tutorial/>
- **Getting started with Python and Jupyter Notebooks** <https://colab.research.google.com/github/jckantor/CBE30338/blob/master/docs/01.01-Getting-Started-with-Python-and-Jupyter-Notebooks.ipynb>