



LAB 5: WORKING WITH GRAPHS & ADVANCED DEBUGGING

University of Washington ECE 241

Author: Jimin Kim (jk55@uw.edu)

Version: v1.7.0

OUTLINE

Part 1: Introduction to graphs

- Fundamental graph elements
- Graph degrees
- Paths and connectedness
- Removing or Adding vertices/edges

Part 2: Application of graph data structure

- Brain connectome
- Social network
- Power grid
- Knowledge graph

Part 3: Working with Graph data

- Graphs as a data structure
- Visualizing graphs
- Computing graph properties
- Removing or adding edges

Part 4: Advanced debugging

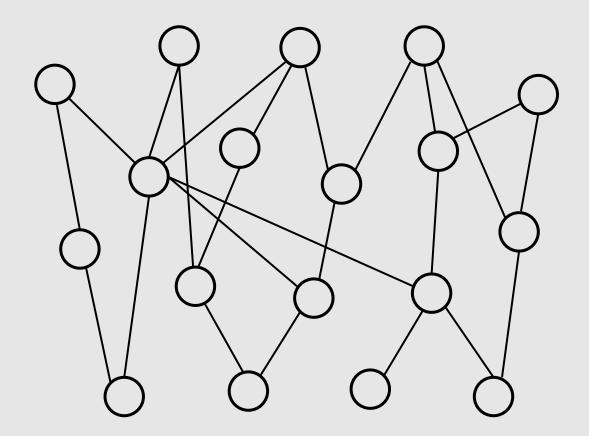
- Introduction to Python Debugger
- Python Debugger Basics
- Useful Commands

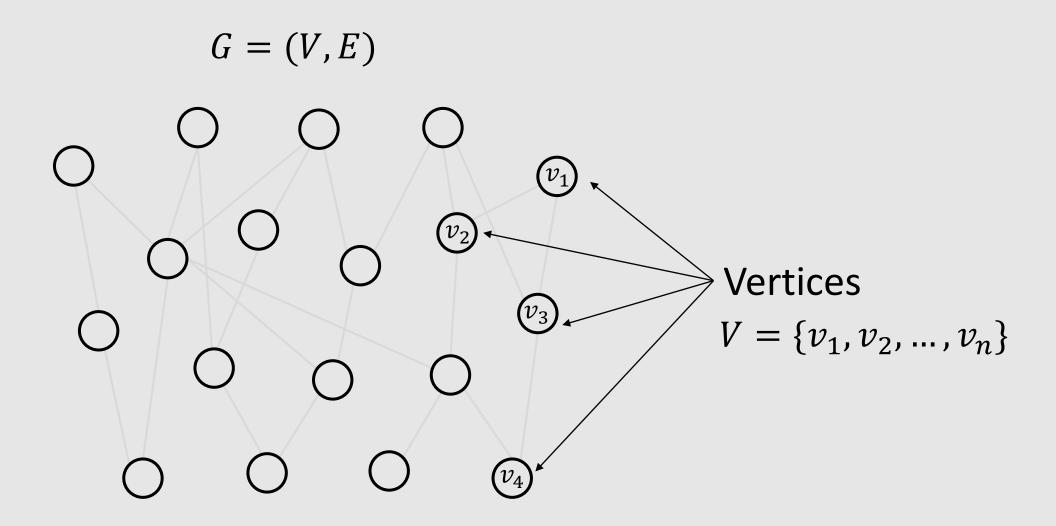
Part 5: Lab Assignments

- Exercise 1 – 5

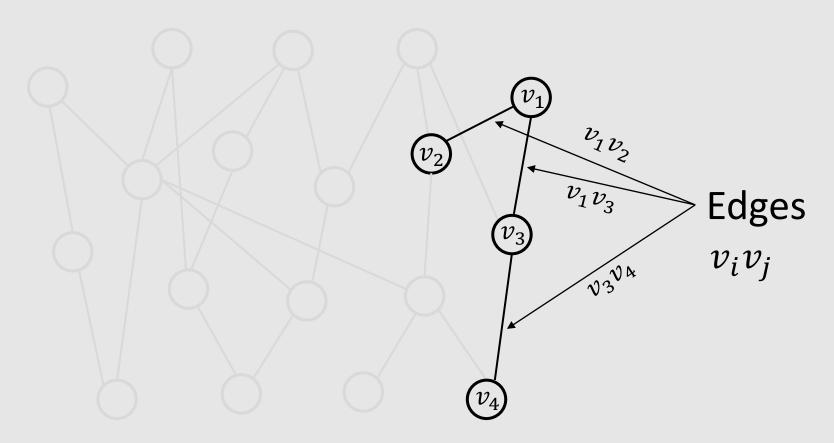
PART 1: INTRODUCTION TO GRAPHS

$$G = (V, E)$$

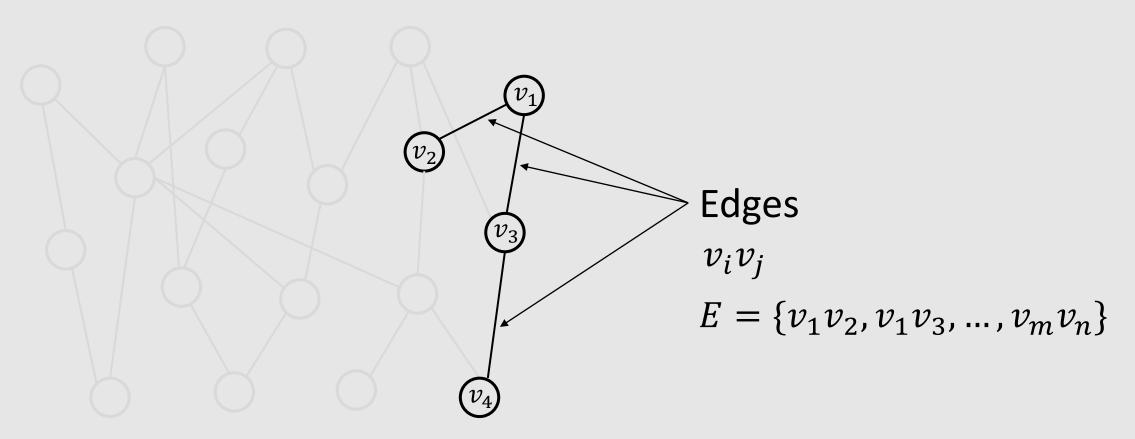




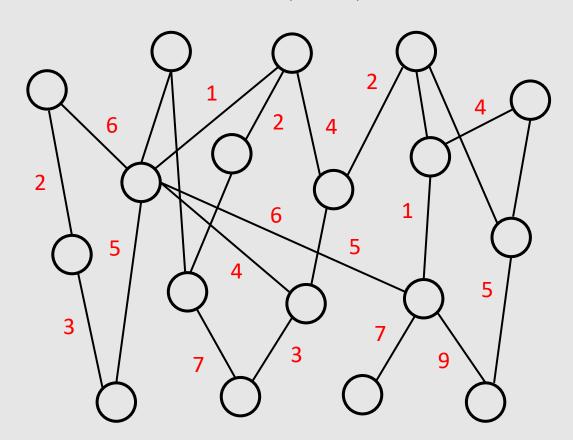
$$G = (V, E)$$



$$G = (V, E)$$



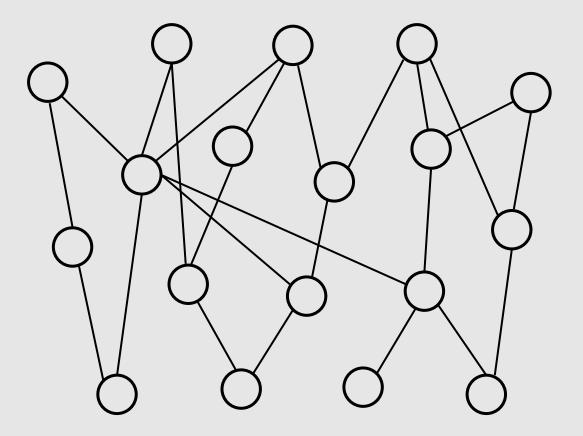
$$G = (V, E)$$



Edges can have values attached to them to represent:

- Number or weight of connections between vertices
- Probability of an event from another event (vertex = event)
- Relationship between vertices (i.e. predicate)
- etc...

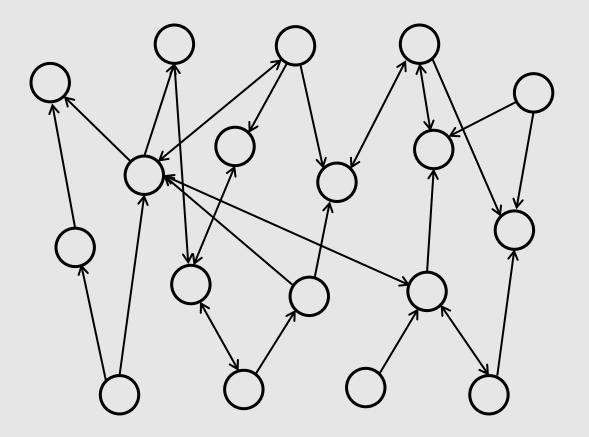
$$G = (V, E)$$



Edges can be:

- Undirected (i.e. bidirectional)

$$G = (V, E)$$



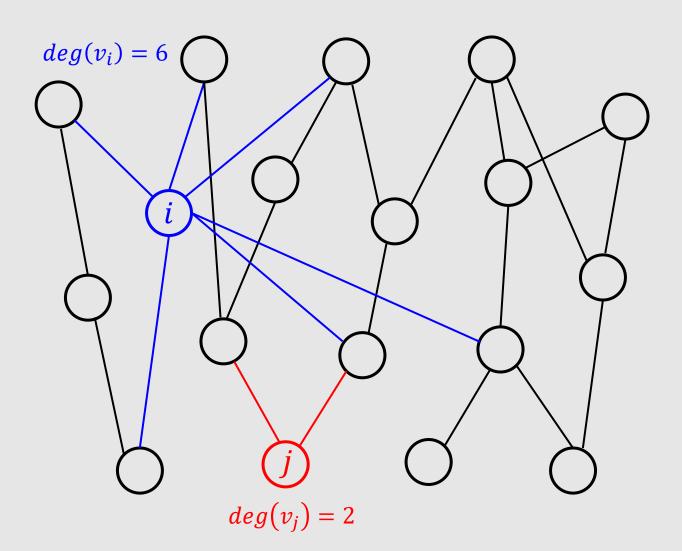
Edges can be:

- Undirected (i.e. bidirectional)
- Directed (i.e. unidirectional)

Note:

Directed graph can have bidirectional connection between two vertices via having two edges $A \rightarrow B$ and $B \rightarrow A$

GRAPH DEGREES

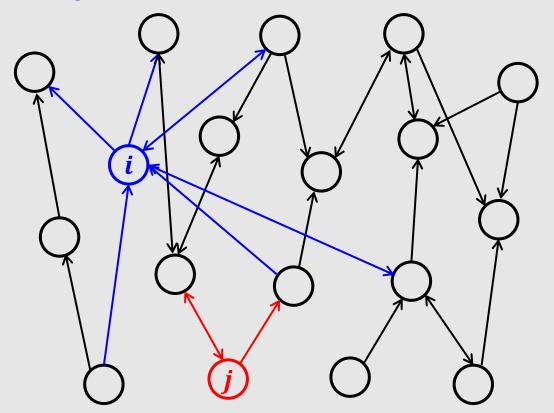


Undirected graphs

Degree of a vertex v_i = number of edges or sum of edge weights **incident** to the vertex

GRAPH DEGREES

```
outdeg = 4
indeg = 4
```

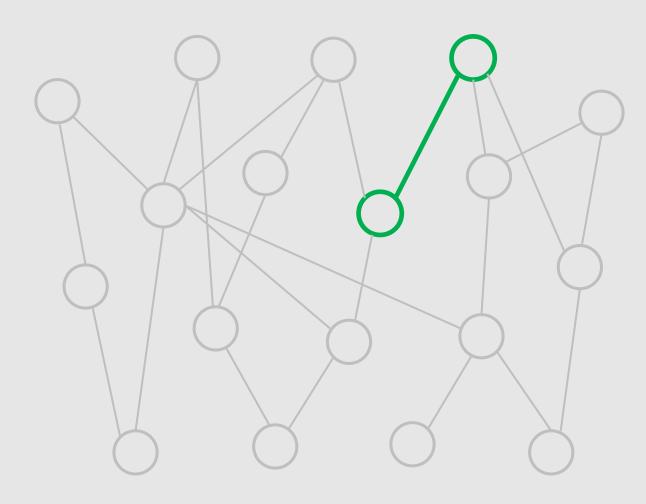


Directed graphs

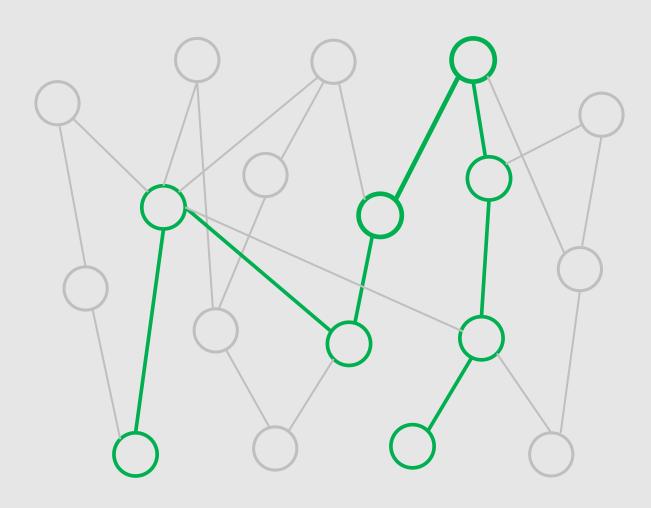
Out-degree: number of incident tail ends or outgoing edge weights

In-degree: number of incident head ends or incoming edge weights

outdeg = 2indeg = 1

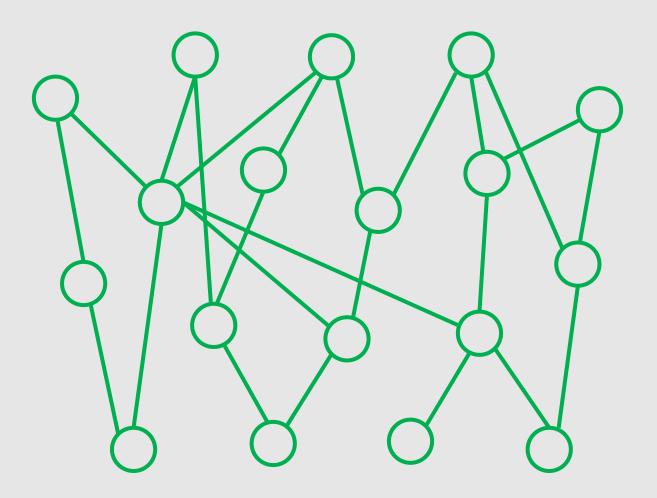


Two vertices connected by an edge are said to be **adjacent**



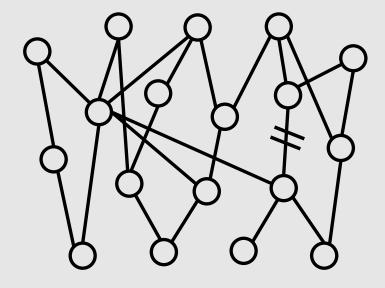
A path of length *m* in *G*:

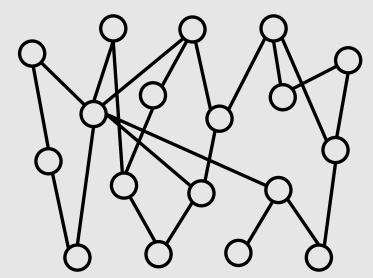
A sequence of distinct adjacent vertices



A graph is **connected** if **all pairs of vertices** are connected by a **path**

REMOVING VERTICES/EDGES

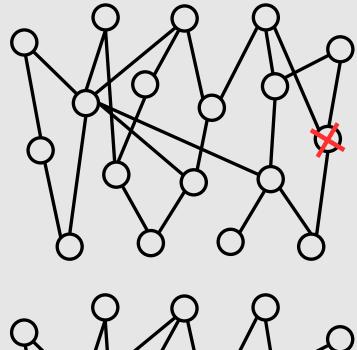


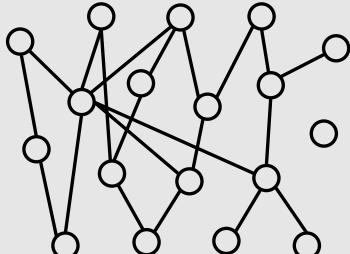


Both vertices and edges can be removed from a graph

Removing an edge **disconnects a connection** between 2 vertices

REMOVING VERTICES/EDGES

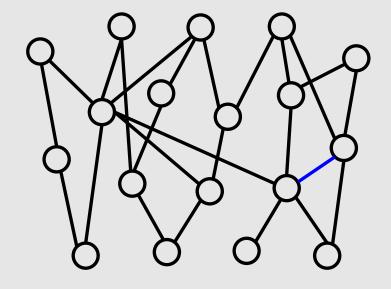


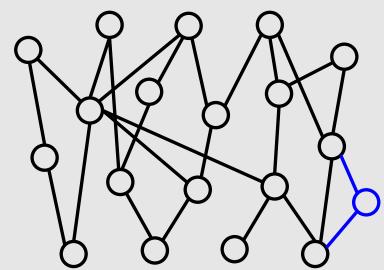


Both vertices and edges can be removed from a graph

Removing a vertex removes the **all the edges** (both **incoming and outgoing** if directed) that connects to it

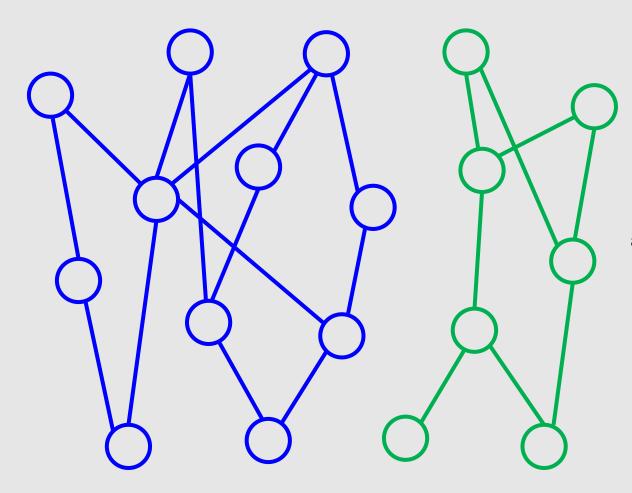
ADDING VERTICES/EDGES





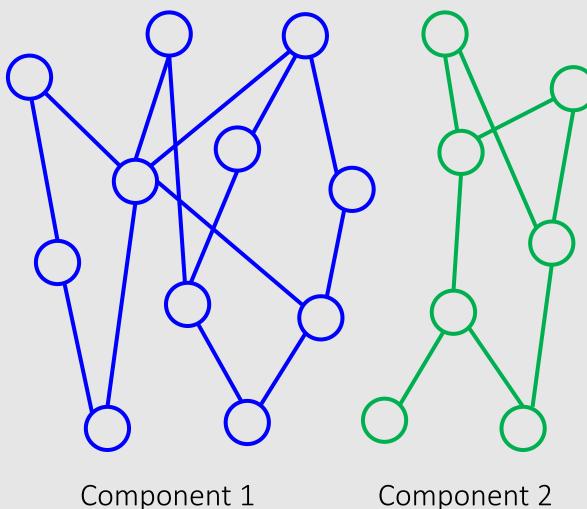
One can also ADD new edges or vertices into a graph

Adding a vertex **also adds edges** that connects the new vertex to the graph



Removing edges or vertices may result in two separate graphs that are **not connected by any path**

In such a case, a graph is said to be **Disconnected**



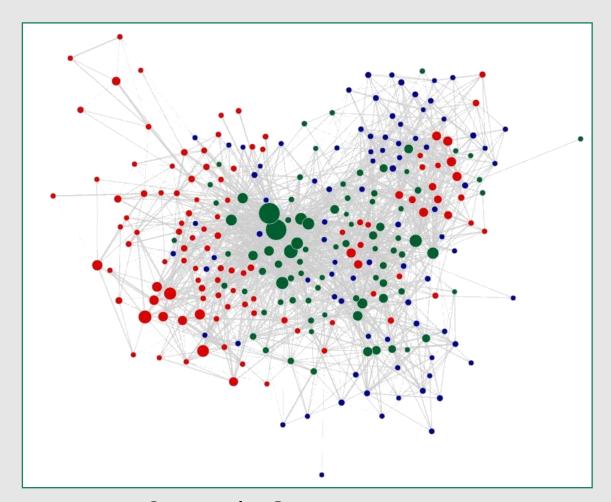
Removing edges or vertices may result in two separate graphs that are not connected by any path

In such a case, a graph is said to be **Disconnected**

Each disconnected graph is called a 'Component'

PART 2: APPLICATIONS OF GRAPH STRUCTURES

BRAIN CONNECTOME



Nervous System of C. elegans Nematode

Vertices:

Neurons

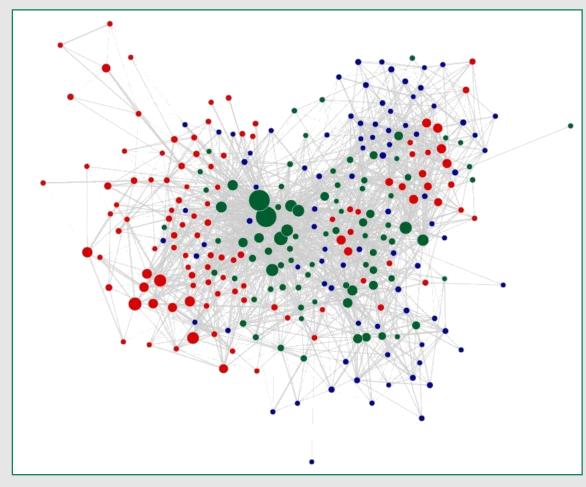
Edges:

Synaptic connection

Synaptic Connectome (Weighted Directed Network)

Image credit: Jimin Kim, Will Leahy, Eli Shlizerman, Neural Interactome: Interactive Simulations of Neuronal Networks Frontiers in Computational Neuroscience, 2019

BRAIN CONNECTOME



Synaptic Connectome (Weighted Directed Network)

In-degree of a vertex:

Number of synaptic connections a neuron receives from others

Out-degree of a vertex:

Number of synaptic connections a neuron makes to others

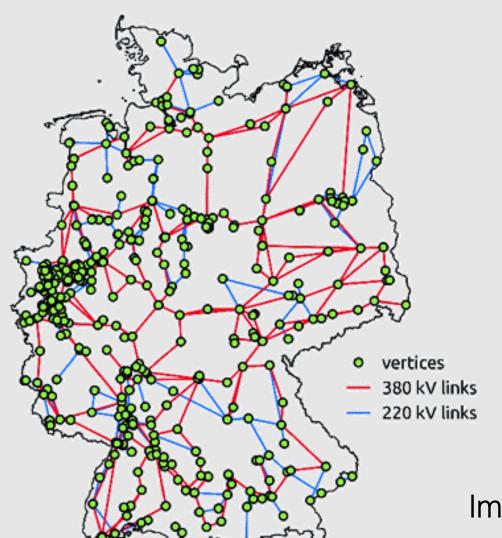
Removing an edge:

Disconnecting a synaptic connection between two neurons

Removing a vertex:

Removing a neuron and its incoming + outgoing synaptic connections from the brain

POWER GRID



Vertices:

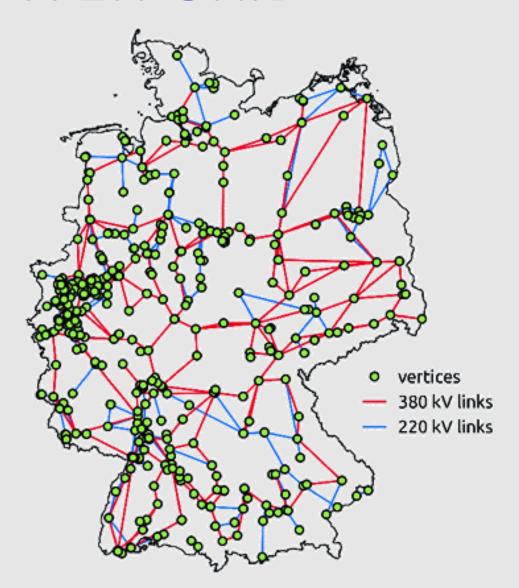
Power stations

Edges:

Power connections

Image credit: Germany's power grid

POWER GRID



Degree of a vertex:

Number of power links a station is connected

Removing an edge:

Disconnecting a power link between two stations

Removing a vertex:

Removing a station and its power links from the grid

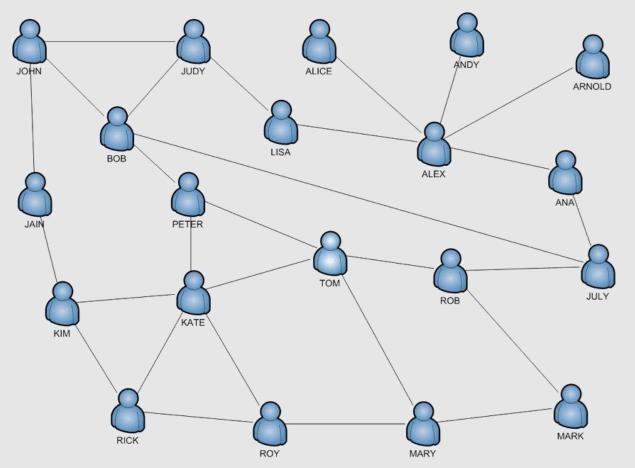
Adding an edge

Adding a new power link between two stations

Adding a vertex

Adding a new power station and power links that connect to grid

SOCIAL NETWORK



Vertices:

People

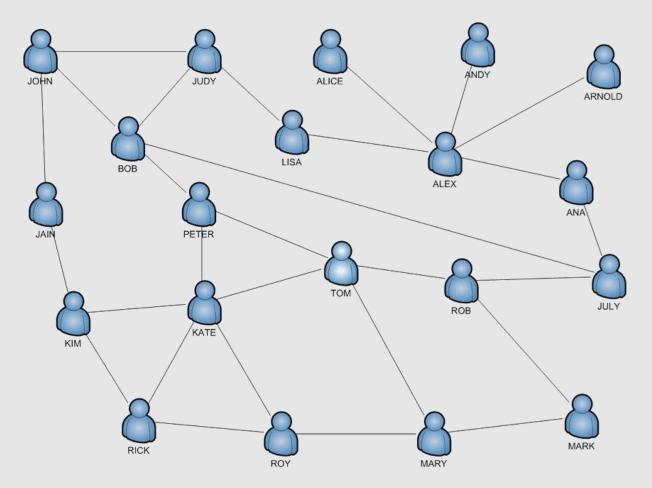
Edges:

Friend, follow status, etc

Image credit: Tomasz Filipowski

Web-based knowledge exchange through social links in the workplace Behavior and Information Technology, 2012

SOCIAL NETWORK



Degree of a vertex:

Number of social connections (e.g. friends) a person (vertex) has

Removing an edge:

Disconnecting a social connection between two people (e.g. un-follow)

Removing a vertex:

Removing a person and his/her social connections from the network (e.g. deleting account)

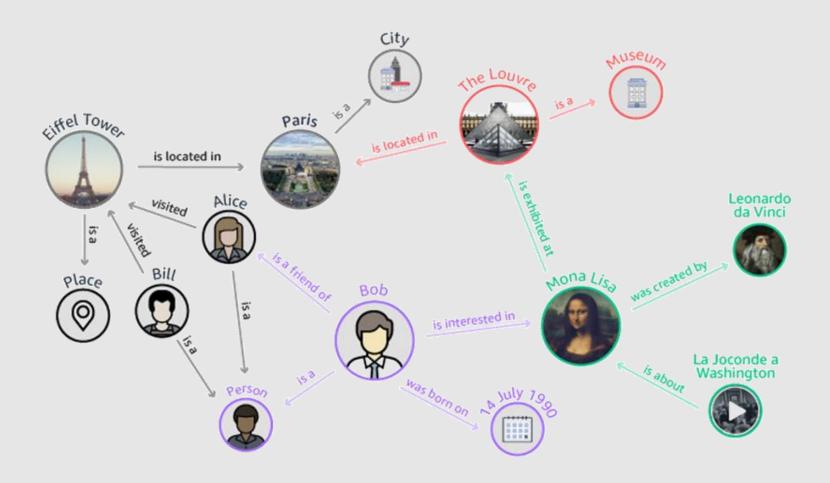
Adding an edge

Adding a friend, following an account

Adding a vertex

Newly registering to the social network

KNOWLEDGE GRAPH



Vertices:

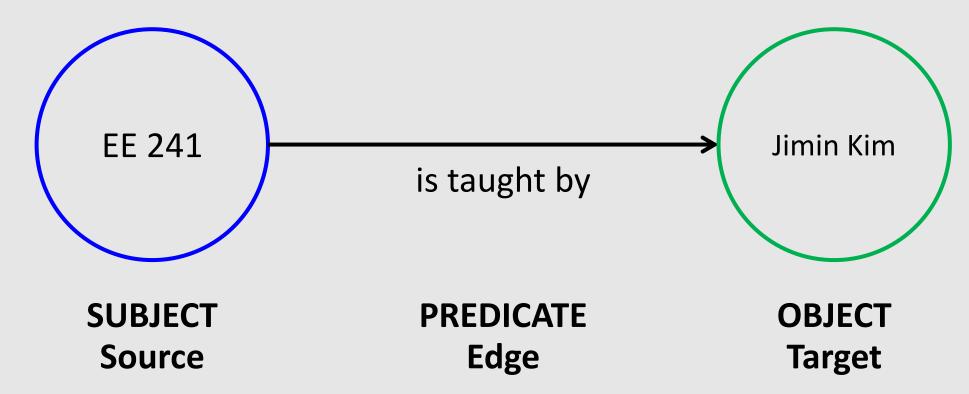
Subjects/Objects

Edges:

Predicates

Image credit: Amazon AWS

KNOWLEDGE GRAPH: TRIPLES

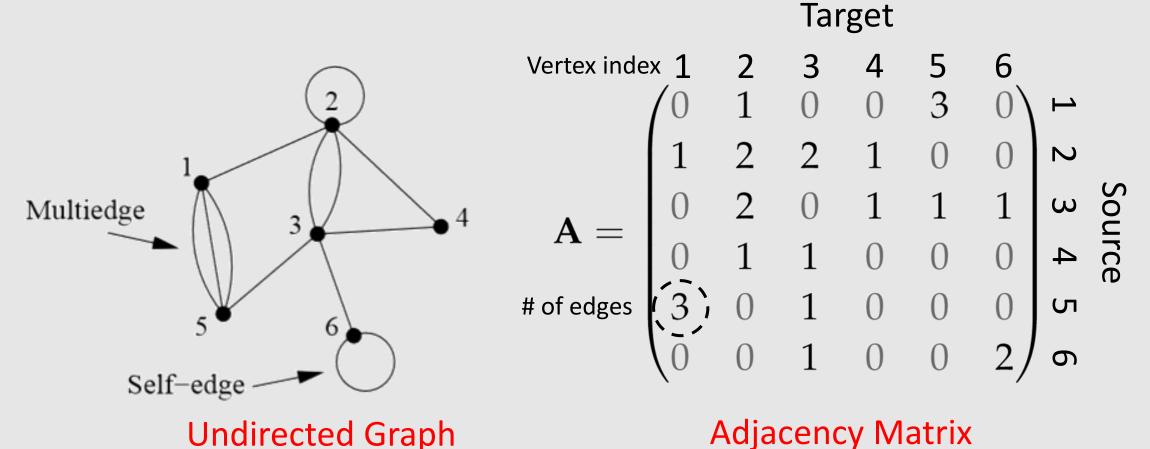


Q: What would these mean in the context of knowledge graph?

- In-degree
- Out-degree
- Removing/adding an edge
- Removing/adding a vertex

PART 3: WORKING WITH GRAPH DATA

GRAPH AS A DATA STRUCTURE: ADJACENCY MATRIX

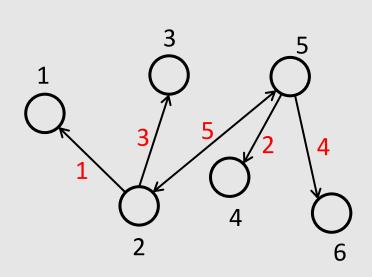


Adjacency Matrix

NOTE: Undirected graph has *Symmetric* Adjacency Matrix (i.e. $A = A^T$)

Image credit: Newman, Mark. Network: An Introduction. Oxford University Press, 2018.

GRAPH AS A DATA STRUCTURE: ADJACENCY MATRIX





Directed Graph

Adjacency Matrix

NOTE: Directed graph is **NOT** necessarily *Symmetric*

NOTE: Weight of the edge is often equivalent to # of edges

GRAPH AS A DATA STRUCTURE: TRIPLES ARRAY

Loading a graph as adjacency matrix could be problematic if the graph is:

- Sparse (i.e. many zeros)
- Very large (e.g. social network)

One way to address this is to represent graphs as lists of triples arrays

 Saves memory via only recording nonzero edges

$$\begin{array}{c} v_0 \ v_1 \ \text{Target} \\ v_0 \\ v_1 \\ A = \begin{pmatrix} 0 & 1 & 0 & 0 & 3 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix} \\ \begin{array}{c} \text{SOUTS} \\ \text{Ce} \end{array}$$

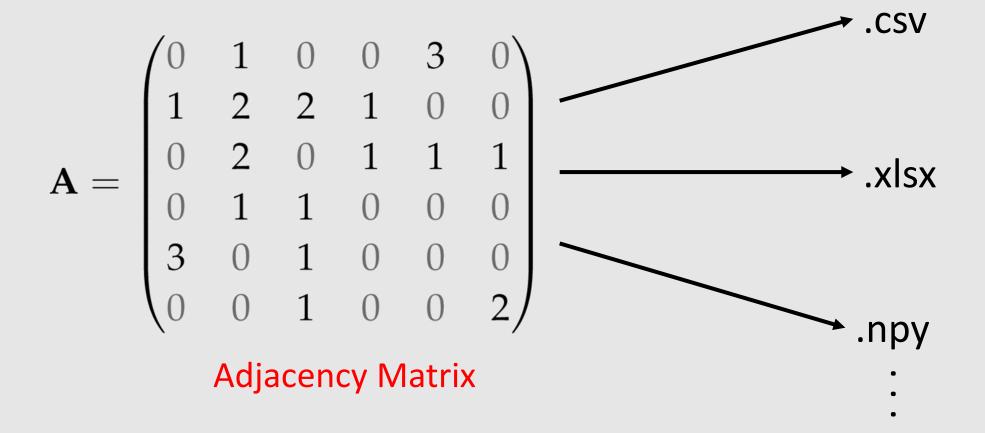
Example of a sparse adjacency matrix

Source	Target	Weight	
1	0	1	
1	1	2	
1	2	2	
1	3	1	

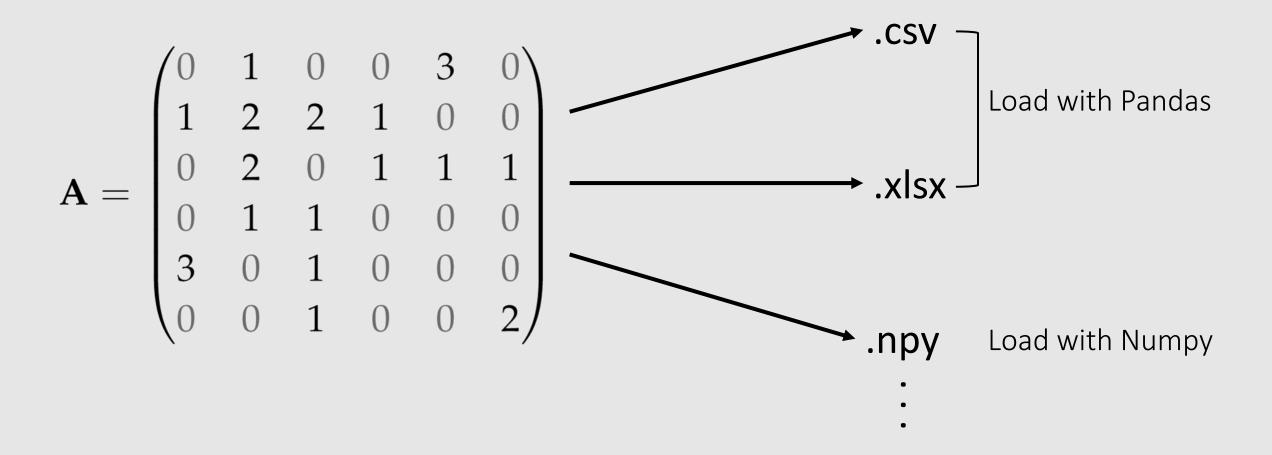
Vertex v_0

Vertex $v_{\scriptscriptstyle 1}$

GRAPH AS A DATA STRUCTURE



GRAPH AS A DATA STRUCTURE



LOADING GRAPH DATA

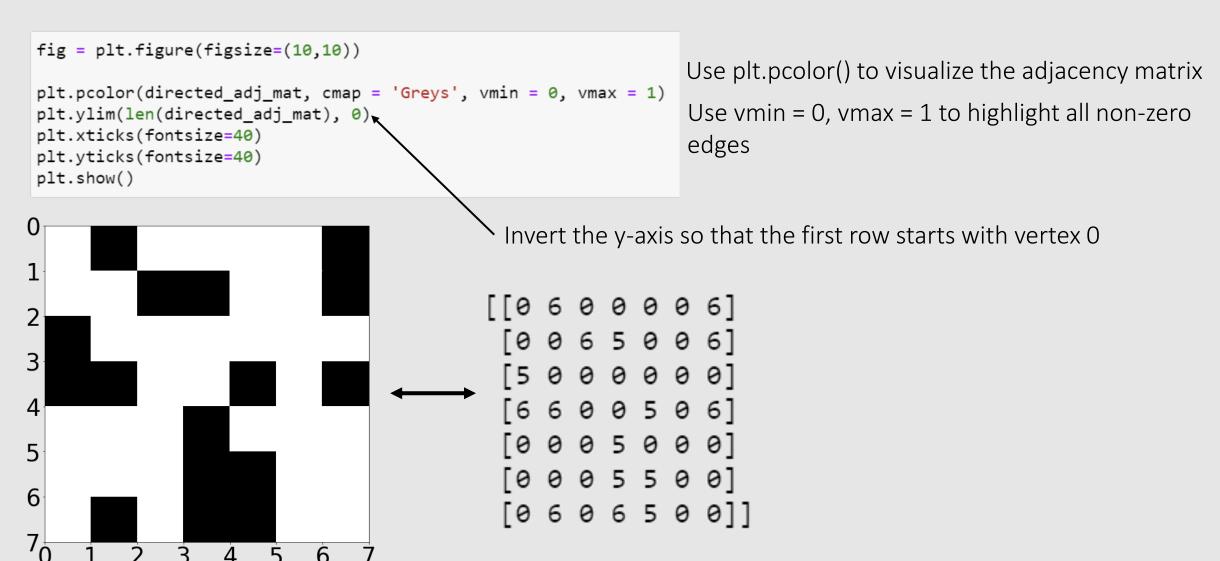
```
directed_adj_mat = pd.read_excel('directed_sample.xlsx')
directed_adj_mat = np.array(directed_adj_mat)
print(directed adj mat)
[[0600006]
[0 0 6 5 0 0 6]
[5 0 0 0 0 0 0]
[6600506]
 [0 0 0 5 0 0 0]
[0 0 0 5 5 0 0]
 [0606500]]
directed_adj_mat = pd.read_csv('directed_sample.csv')
directed adj mat = np.array(directed adj mat)
print(directed adj mat)
[[0600006]
[0 0 6 5 0 0 6]
[5000000]
 [6600506]
 [0 0 0 5 0 0 0]
 [0 0 0 5 5 0 0]
 [0 6 0 6 5 0 0]]
directed_adj_mat = np.load('directed_sample.npy')
print(directed adj mat)
[[0600006]
[0 0 6 5 0 0 6]
[5 0 0 0 0 0 0]
[6 6 0 0 5 0 6]
 [0 0 0 5 0 0 0]
 [0 0 0 5 5 0 0]
[0606500]]
```

Loading an adjacency matrix in .xlsx form

Loading .csv form

Loading .npy form

VISUALIZING AN ADJACENCY MATRIX



VISUALIZING A GRAPH USING NetworkX

What is NetworkX?

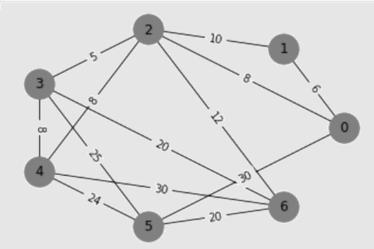
- A Python package for creating, manipulating and analyzing networks
- Provides tools for the study of the network structure
- Graph visualization tools built with matplotlib
- Works seamlessly with Numpy
- Included in Anaconda3



VISUALIZING A GRAPH (Undirected)

```
import pandas as pd
import networkx as nx
import matplotlib.pyplot as plt
```

Import pandas, networkx, matplotlib



Use nx.get_edge_attributes() to obtain the edge weights according to adjacency matrix

Use nx.draw_networkx_edge_labels() to label the edges with the weights

VISUALIZING A GRAPH (Directed)

```
directed_adj_mat_panda = pd.read_excel('directed_sample.xlsx')

directed_adj_mat_np = np.array(directed_adj_mat_panda)

directed_adj_mat_nx = nx.from_numpy_array(directed_adj_mat_np, create_using=nx.DiGraph())

Add create_using=nx.DiGraph() to specify that the graph is directed pos=nx.circular_layout(directed_adj_mat_nx)

Obtain positions of vertex according to circular graph layout

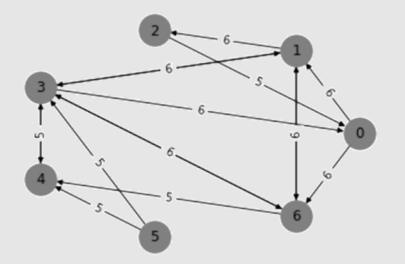
nx.draw_networkx(directed_adj_mat_nx, pos, with_labels = True, node_size = 750, node_color='grey')

labels = nx.get_edge_attributes(directed_adj_mat_nx, 'weight')

nx.draw_networkx_edge_labels(directed_adj_mat_nx, pos,edge_labels=labels)

plt.axis('off')

Use nx.draw_networkx(), nx.get_edge_attributes(), nx.draw_networkx_edge_labels() to plt.show()
```



NOTE: By default, NetworkX uses **rows = source, columns = target** orientation for adjacency matrix

COMPUTING DEGREES (UNDIRECTED)

```
undirected_adj_mat = pd.read_excel('undirected_sample.xlsx')
undirected_adj_mat_np = np.array(undirected_adj_mat)
```

```
Load the included undirected adjacency matrix sample
print(undirected_adj_mat_np)
                                                We will compute the degree of the vertex 3
    0 0 25 24 0 20]
  0 0 12 20 30 20 0]]
np.sum(undirected adj mat np[3, :])
                                                Summing over the row (or column) gives the degree
58
```

np.sum(undirected_adj_mat_np, axis = 1)

Summing over all the columns gives a list of degrees for all the neurons

NOTE: Due to adjacency matrix being *symmetric*, the operation will yield identical degrees when computed in respect to *columns* instead of *rows*

COMPUTING DEGREES (DIRECTED)

```
directed_adj_mat = pd.read_excel('directed_sample.xlsx')
directed_adj_mat_np = np.array(directed_adj_mat)

print(directed_adj_mat_np)

[[0 6 0 0 0 0 6]
  [0 0 6 5 0 0 6]
  [5 0 0 0 0 0 0]
```

Load the included undirected adjacency matrix sample

We will compute both **in-degree and out-degree** of the vertex 3

We will compute both **in-degree and out-degree** of the vertex 3

in-degrees of vertex 3

[0 6 0 6 5 0 0]]

COMPUTING DEGREES (DIRECTED)

<pre>np.sum(directed_adj_mat_np[3, :]) 23</pre>	Summing over all numbers along the columns with respect to a single vertex gives out-degree of a vertex
<pre>np.sum(directed_adj_mat_np, axis = 1) array([12, 17, 5, 23, 5, 10, 17], dtype=int64)</pre>	Summing over all the columns gives out-degrees of all the vertices
<pre>np.sum(directed_adj_mat_np[:, 3]) 21</pre>	Summing over all numbers along the rows with respect to a single vertex gives in-degree of a vertex
<pre>np.sum(directed_adj_mat_np, axis = 0) array([11, 18, 6, 21, 15, 0, 18], dtype=int64)</pre>	Summing over all the rows gives in-degrees of all the vertices

NOTE: We are using **rows = source**, **columns = target** orientation for adjacency matrix

REMOVING EDGES

```
directed_adj_mat = pd.read_excel('directed_sample.xlsx')
                                                          Load the included directed adjacency matrix sample
directed_adj_mat = np.array(directed_adj_mat)
directed_adj_mat[0, 1] = 0
                                                                                                    Set following edges to zero
directed_adj_mat[1, 2] = 0
                                                                                                        vertex 0 \rightarrow \text{vertex } 1
directed_graph_sample = nx.from_numpy_array(directed_adj_mat, create_using=nx.DiGraph())
                                                                                                        vertex 1 \rightarrow \text{vertex } 2
pos=nx.circular_layout(directed_graph_sample)
nx.draw_networkx(directed_graph_sample,pos,with_labels = True, node_size = 750, node_color='grey')
labels = nx.get_edge_attributes(directed_graph_sample, 'weight')
nx.draw_networkx_edge_labels(directed_graph_sample,pos,edge_labels=labels)
plt.axis('off')
                                                          Plot the graph with removed edges
plt.show()
```

0 6 5 0 0]]

Q: How would you expand the operation to remove a vertex?

ADDING EDGES

```
directed adj mat = pd.read excel('directed sample.xlsx')
                                                          Load the included directed adjacency matrix sample
directed adj mat = np.array(directed adj mat)
directed_adj_mat[2, 3] = 4
                                                                                                  Add a new edge with weight = 4
                                                                                                      vertex 2 \rightarrow \text{vertex } 3
directed_graph_sample = nx.from_numpy_array(directed_adj_mat, create_using=nx.DiGraph())
pos=nx.circular layout(directed graph sample)
nx.draw_networkx(directed_graph_sample,pos,with_labels = True, node_size = 750, node_color='grey')
labels = nx.get_edge_attributes(directed_graph_sample, 'weight')
nx.draw_networkx_edge_labels(directed_graph_sample,pos,edge_labels=labels)
plt.axis('off')
plt.show()
                                                                   New connection = 4
```

0 6 5 0 0]]

Q: How would you expand the operation to add a vertex?

PART 4: ADVANCED DEBUGGING

INTRODUCTION TO PYTHON DEBUGGER

What is Python Debugger (PDB)?

- Debugging tool built in Python
- No installation required
- Provides Interactive environment
- Useful when debugging using logging is problematic



BASICS

```
def divide(divisor):
                                  Upon execution of the code, activates Python
   breakpoint() =
                                  debugger console at this line
   val = 1/divisor
   return val
divide(0)
> <ipython-input-24-03e5b3dd8265>(5)divide()
           breakpoint()
      3
          val = 1/divisor
           return val
                                                         Python debugger console
ipdb>
```

USEFUL COMMANDS

USEFUL COMMANDS: help

ipdb>

```
divide(0)
                                                                              h: help
> <ipython-input-24-03e5b3dd8265>(5)divide()
                                                                              w: where
          breakpoint()
                                                                              n: next
          val = 1/divisor
                                                                              c: continue
          return val
     7
ipdb> h
                                                                              p: print
Documented commands (type help <topic>):
                                                                              l: list
      commands
                enable
                                                          until
                                                                              q: quit
                         longlist psource skip_hidden
      condition exit
                                                          uр
alias cont
                                          skip predicates
      context
                help
                                  auit
                                          source
                                                          whatis
args
                         next
      continue
                ignore
                                          step
                                                          where
                interact
                         pdef
                                  restart tbreak
break d
      debug
                         pdoc
bt
                                  return
      disable
                                          unalias
                jump
                         pfile
                                  retval
      display
                         pinfo
                                          undisplay
                                  run
clear down
                         pinfo2
                                          unt
Miscellaneous help topics:
______
exec pdb
                                                            Provide documentation for the command
ipdb> h l
Print lines of code from the current stack frame
```

USEFUL COMMANDS: where

```
divide(2)
                                                                         h: help
> <ipython-input-1-f4a9c3842530>(7)divide()
                                                                         w: where
          val = 1/divisor
          breakpoint()
                                                                         n: next
                                                                         c: continue
---> 7
          return val
                                                                         p: print
ipdb> w
    [... skipping 27 hidden frame(s)]
                                                                         l: list
  <ipython-input-2-cfdb0f794c84>(1)<module>()
  --> 1 divide(2)
                                                                         q: quit
> <ipython-input-1-f4a9c3842530>(7)divide()
          val = 1/divisor
          breakpoint()
          return val
                                                     Print a stack trace (i.e. the order of code being
                                                     executed), with the most recent frame at the
ipdb>
                                                     bottom.
```

USEFUL COMMANDS: next

```
h: help
divide(2)
                                                                 w: where
> <ipython-input-1-03e5b3dd8265>(5)divide()
           breakpoint()
                                                                 n: next
                                                                 c: continue
         val = 1/divisor
                                                                 p: print
           return val
                                                                 I: list
> <ipython-input-1-03e5b3dd8265>(7)divide()
                                                                 q: quit
           breakpoint()
           val = 1/divisor
           return val
                                                         Executes the next line of code
ipdb>
```

USEFUL COMMANDS: continue

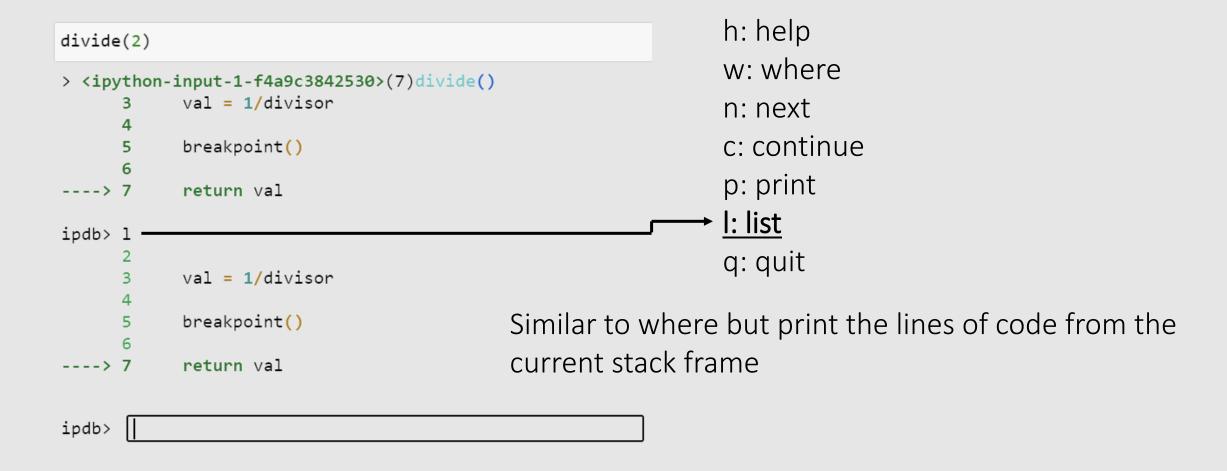
```
h: help
divide(2)
                                                                   w: where
> <ipython-input-1-a669feb31165>(5)divide()
           breakpoint()
                                                                   n: next
                                                                   c: continue
           val = 1/divisor
                                                                   p: print
           breakpoint()
                                                                   l: list
ipdb> c -
> <ipython-input-1-a669feb31165>(9)divide()
                                                                   q: quit
           val = 1/divisor
           breakpoint()
           return val
----> 9
                                                          Run the code until the next breakpoint()
ipdb>
```

USEFUL COMMANDS: print

```
h: help
divide(2)
                                                                  w: where
> <ipython-input-1-f4a9c3842530>(7)divide()
           val = 1/divisor
                                                                  n: next
                                                                  c: continue
           breakpoint()
                                                                  p: print
           return val
---> 7
                                                                  l: list
ipdb> p val —
0.5
                                                                  q: quit
ipdb>
```

Print the value of the desired variable

USEFUL COMMANDS: list



USEFUL COMMANDS: quit

if self.stop here(frame) or self.break here(frame):

self.user line(frame)

return self.trace dispatch

if self.quitting: raise BdbQuit

```
divide(2)
> <ipython-input-1-f4a9c3842530>(7)divide()
           val = 1/divisor
                                                                                               h: help
           breakpoint()
                                                                                               w: where
---> 7
           return val
                                                                                               n: next
ipdb> a
                                                                                               c: continue
BdbOuit
                                       Traceback (most recent call last)
<ipython-input-2-cfdb0f794c84> in <module>
                                                                                               p: print
----> 1 divide(2)
<ipython-input-1-f4a9c3842530> in divide(divisor)
                                                                                               l: list
           breakpoint()
---> 7
           return val
                                                                                              q: quit
<ipython-input-1-f4a9c3842530> in divide(divisor)
           breakpoint()
---> 7
           return val
~\anaconda3\lib\bdb.py in trace_dispatch(self, frame, event, arg)
    86
                  return # None
    87
              if event == 'line':
                  return self.dispatch line(frame)
---> 88
                                                                           Exit the debugger console
              if event == 'call':
                  return self.dispatch_call(frame, arg)
~\anaconda3\lib\bdb.py in dispatch line(self, frame)
```

Useful video tutorial of basic usage of Python debugger:

https://www.youtube.com/watch?v=aZJnGOwzHtU

BdbQuit:

111

112

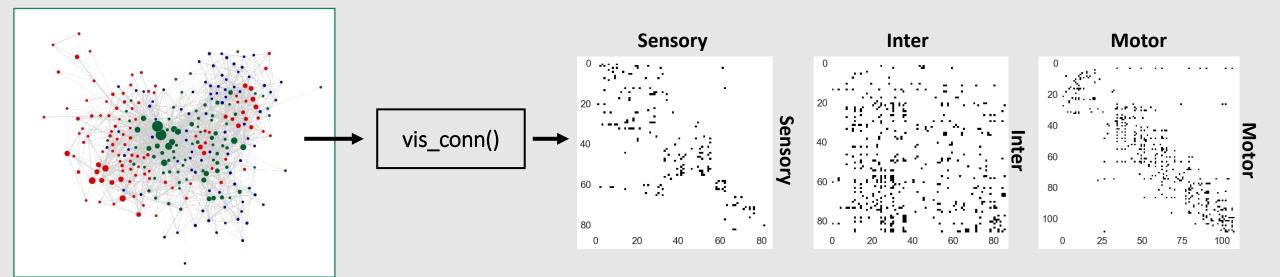
114 115

--> 113

LAB ASSIGNMENTS

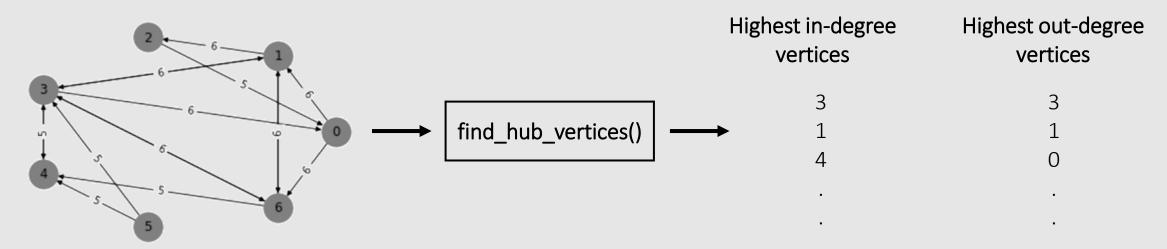
Download ipynb template in Canvas page:
Assignments/Lab 5 report -> click "Lab 5 Report Templates"

EXERCISE 1: Visualize brain connectomes



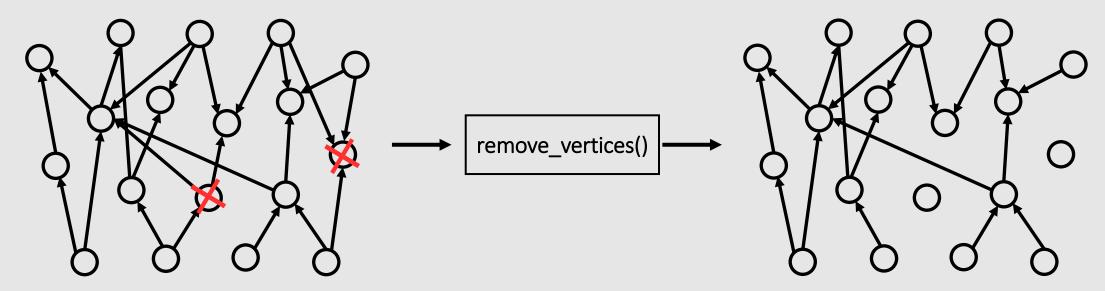
- Recalling from slides 22-23, create a function vis_conn() which takes adjacency matrix of the real *C. elegans* synaptic connectome, list of neuron classes (i.e. labels) for each neuron (sensory, inter, motor) and outputs 1 x 3 subplots of three adjacency matrices corresponding to sub-networks of Sensory, Inter and Motorneurons (e.g. sub-network of sensory neurons is a wiring map between sensory neurons only).
- The function should accept following parameters
 - syn_conn the adjacency matrix of the *C. elegans* synaptic connectome (279 * 279)
 - neuron_classes A list of neuron classes for each neuron (length = 279). Follows the same order of neurons as syn_conn.
- Use plt.pcolor() with cmap = 'Greys' and vmin = 0, vmax = 1.
- Don't forget to invert the y-axis so that the first row starts with vertex 0 (See slide 37)
- Include appropriate x and y labels for each sub-network plot.

EXERCISE 2: Locating the most connected vertices



- Create a function find_hub_vertices() which takes adjacency matrix of a directed graph and outputs two lists where the first list: indices of vertices with the highest in-degree and the second list: indices of vertices with the highest out-degree.
- The function should accept following parameters
 - adj_mat the input adjacency matrix
 - num_vertices the number of the highest degree vertices to find for each list
- The output lists should be ordered such that the vertices with highest degree comes at the top, second highest comes at the second etc.
- Test you function with *C. elegans* synaptic connectome and a sample social network graph.

EXERCISE 3: Removing vertices from a graph



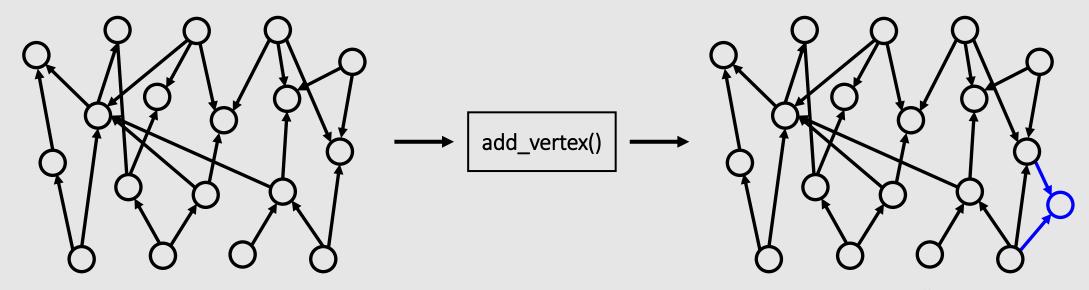
Expanding the concept of removing edges from a graph, write a function named **remove_vertices()** which takes **adjacency matrix** of an existing graph, **list of vertices to be removed** and returns a new adjacency matrix with removed vertices.

The function should take following parameters as inputs:

- adj_mat the input adjacency matrix
- vertices_2b_removed list of vertices to be removed from the graph

Assume that the adjacency matrix follows **row = source, column = to orientation.**Removing a vertex is equivalent to removing all the incoming/outgoing edges from that vertex.
Test your function with provided adjacency matrix and lists of vertices to be removed.

EXERCISE 4: Adding a new vertex to a graph



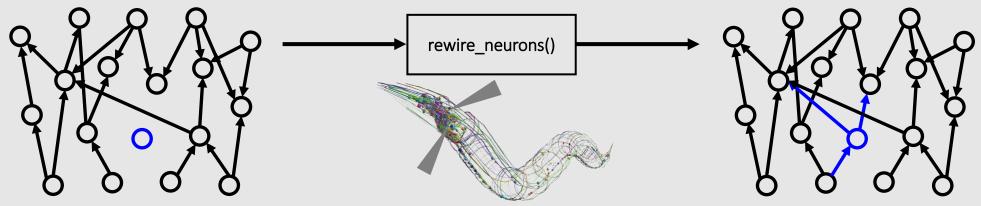
Expanding the concept of adding edges to the graph, write a function named **add_vertex()** which takes adjacency matrix of an existing graph and returns a new adjacency matrix with an added vertex.

The function should take following parameters as inputs:

- adj_mat the input adjacency matrix.
- outgoing_edges list of vertices in which the new vertex connects to.
- incoming_edges list of vertices in which the new vertex receives connections from.

Assume that the adjacency matrix follows **row = source**, **column = target orientation**. Test your function with provided adjacency matrix and lists of outgoing and incoming edges.

EXERCISE 5: Re-wire neurons to restore behavior of *C. elegans*



In this problem we will work with the mathematical model of *C. elegans'* nervous system and body which can simulate the organism's brain, muscles and the body according to its real neural wiring data and its interactions with muscles [1].

Two neurons - 'AVAL' and 'AVAR' play essential roles during the worm's escape maneuver via forward crawling upon a gentle touch on its tail. Unfortunately, the worm has damaged its brain and now lost all the synaptic connections (both incoming and outgoing) for these two neurons. Your task is to re-wire these two neurons to the rest of the nervous system to recover the worm's forward crawling function.

Write a function called **rewire_neurons()** which take following parameters as inputs:

- damaged_brain_adj_matrix Adjacency matrix of the damaged synaptic connectome
- rewiring_instructions_AVAL Triples array describing the re-wiring instructions for AVAL
- rewiring_instructions_AVAR Triples array describing the re-wiring instructions for AVAR

[1] Whole integration of neural connectomics, dynamics and bio-mechanics for identification of behavioral sensorimotor pathways in *Caenorhabditis elegans*, *Jimin Kim, Julia Santos, Mark Alkema, Eli Shlizerman, BioArxiv, 2019*

Use rewiring instructions for AVAL, AVAR to correctly re-wire each neuron to the rest of the neurons in the system. Note that instructions contain both outgoing and incoming connections for both neurons. Your output should be an adjacency matrix of the same dimension as the input adjacency matrix which now contains new connections for AVAL and AVAR. See slide 33 to review the concept of triples array.