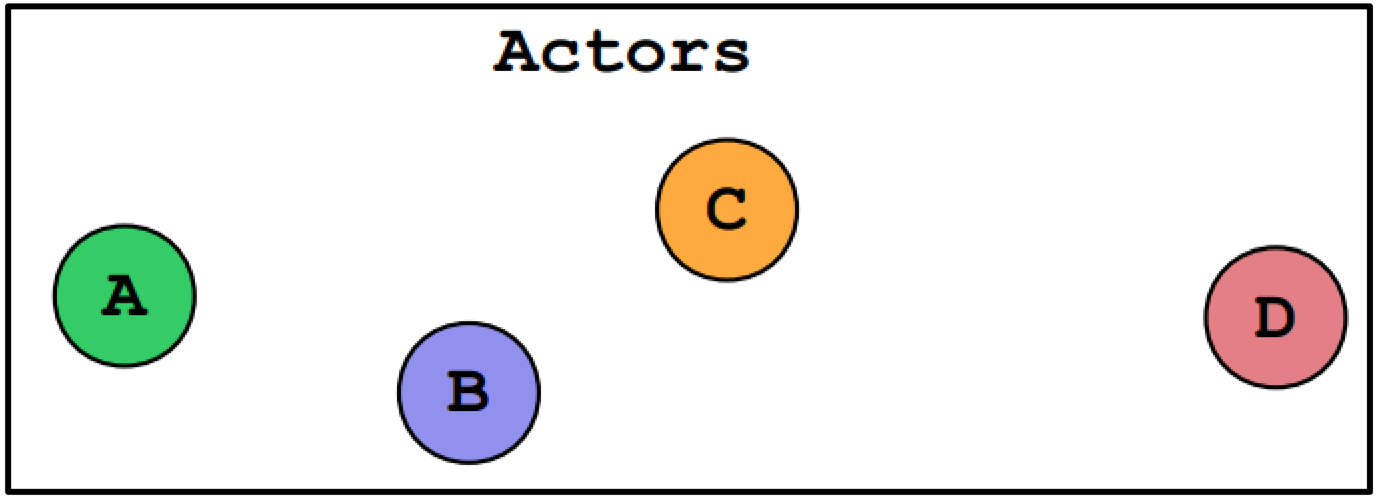
GRAPH THEORY AND SOCIAL NETWORK

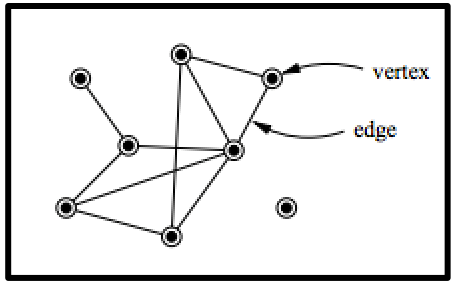
NETWORK ANALYSIS

- ✓ Network analysis is still a growing field with a great deal of opportunity for new and transformative contributions, but its history goes back at least 80 years.
- ✓ While social network theory can be readily applied in theoretical research and qualitative empirical studies, there is a general emphasis on the use of software to analyze and visualize network data once they have been collected.
- ✓ There are a number of different software packages available for this purpose, but two R packages (Statnet and iGraph) have become perhaps the most flexible and powerful tools for performing network analysis.
- ✓ R is a free and open source statistical computing language with a vibrant community of contributors who are constantly updating its functionality through the creation of user defined add-on packages.
- ✓ To get started, check out this website: http://www.r-project.org/.

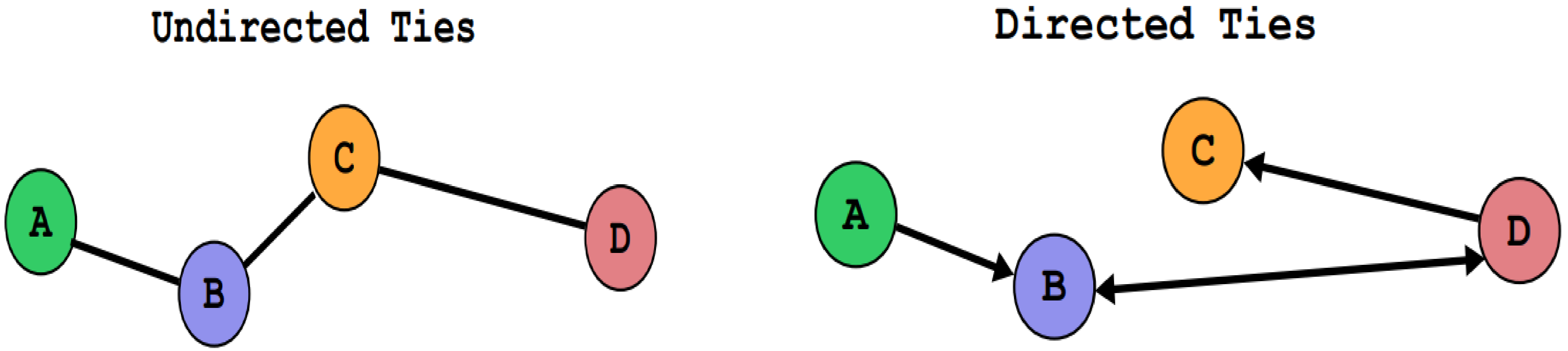
SOCIAL NETWORKING: SOME TERMONOLOGIES

Actor: also called a node or a vertex, referrers to an individual hat can have relationships with other individuals and in this case, an individual or group of individuals we are choosing to study.



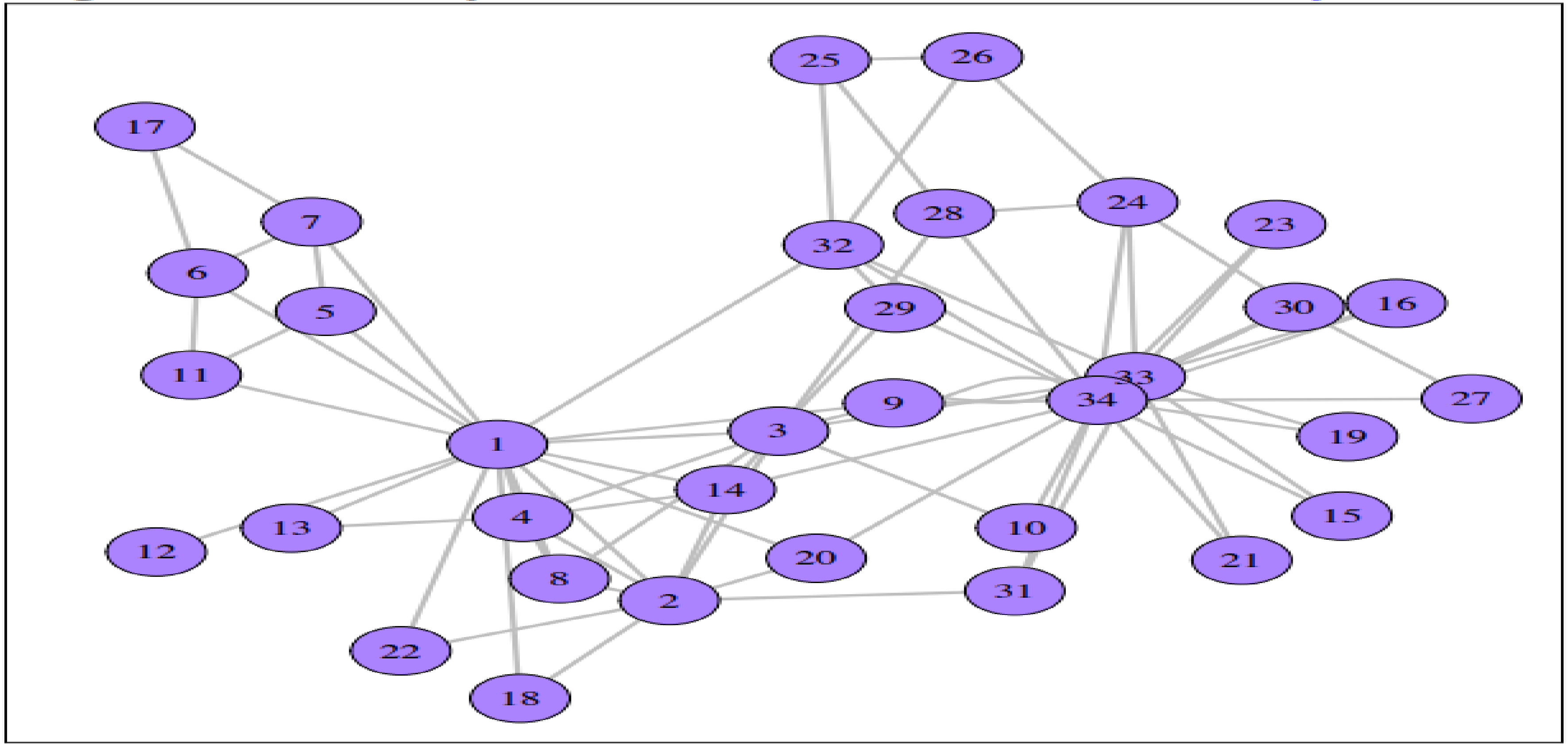


- **Tie**: also called a relation or edge, describes a particular, well specified, relationship between two Actors.
- This could refer to a relationship like "**went to the same school**" or "**likes potato chips**" or something like "**likes**" or "**trades with**".
- Ties can be **un-directed** (like went to the same school), when the **relationship means the same thing to both actors**.



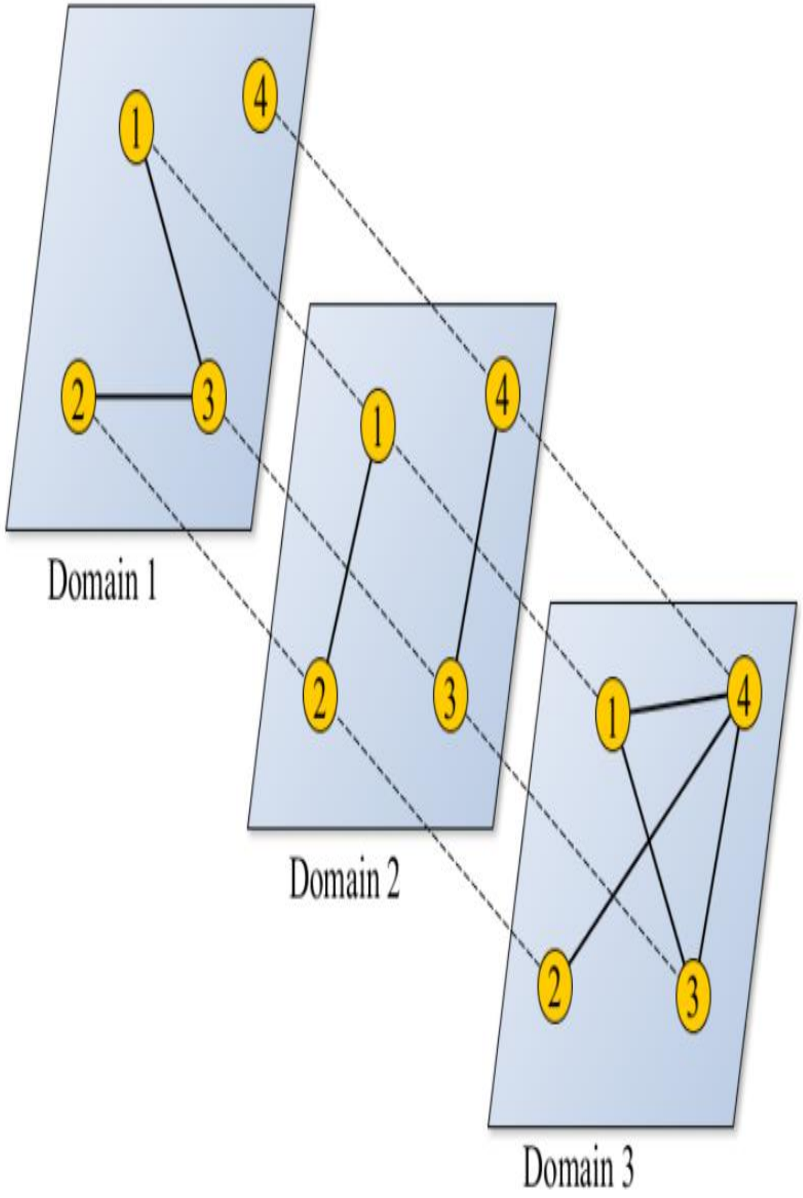
- Ties can also be **directed** (such as "looks up to") and **either one directional or bidirectional**.

- ✓ **Network**: also called a Graph, particularly in the physics and CS literature.
- ✓ Referrers to a collection of Actors and the Ties between them.
- ✓ Figure 2depicts a set of **undirected** friendship relationships between members of a Karate club.

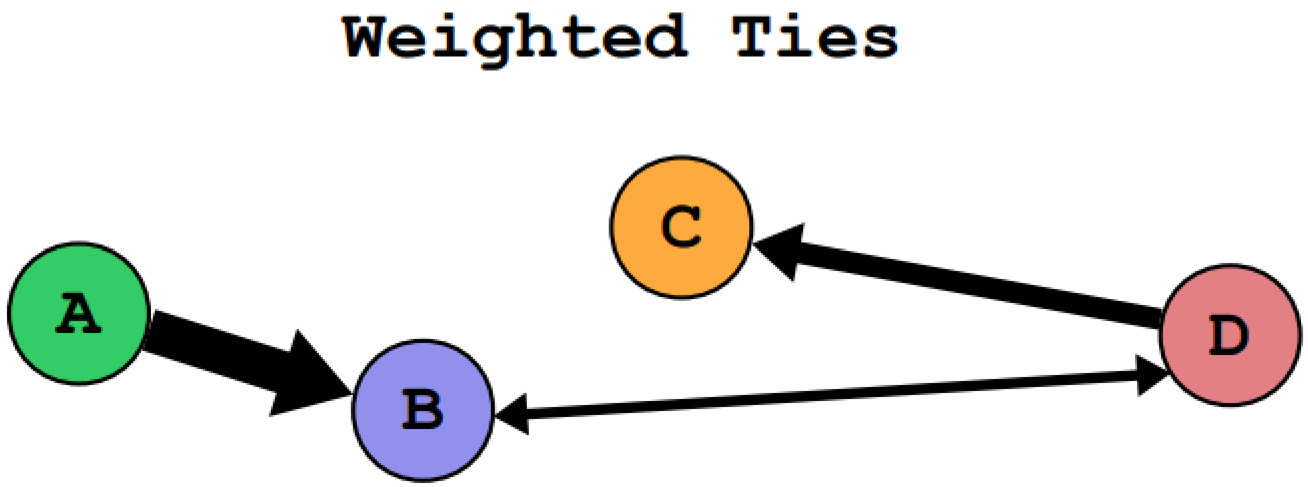


MULTIPLEX NETWORKS

- ✓ **Multiplex networks**.
  - ✓ The set of individuals is the same across all layers. Individuals are not connected to others across layers but can be connected to different sets of people on different layers.
  - ✓ Social networks are representations of relationships that allow us to use methods from graph theory.
  - ✓ Networks consist of nodes, which may be represented as individuals, connected to each other by ties.
  - ✓ The category of multilayer networks encompasses all networks consisting of more than one set of nodes and/or ties, where each layer is defined as a unique set of nodes and ties.
  - ✓ Multiplex networks are the subset of multilayer networks with two basic properties:
    - ✓ (1) all layers share the same set of nodes (i.e., each node replicated in each layer) and ;
    - ✓ (2) all nodes are connected only to themselves across layers (see Fig).
    - ✓ One example of a multiplex network is a social network with layers formed by different domains of interactions, such as hunting, farming, and drinking.
    - ✓ In such a domain-specific multiplex network, all individuals could do all those things (i.e., the same set of nodes is shared across domains), but they may do different things with different sets of people.

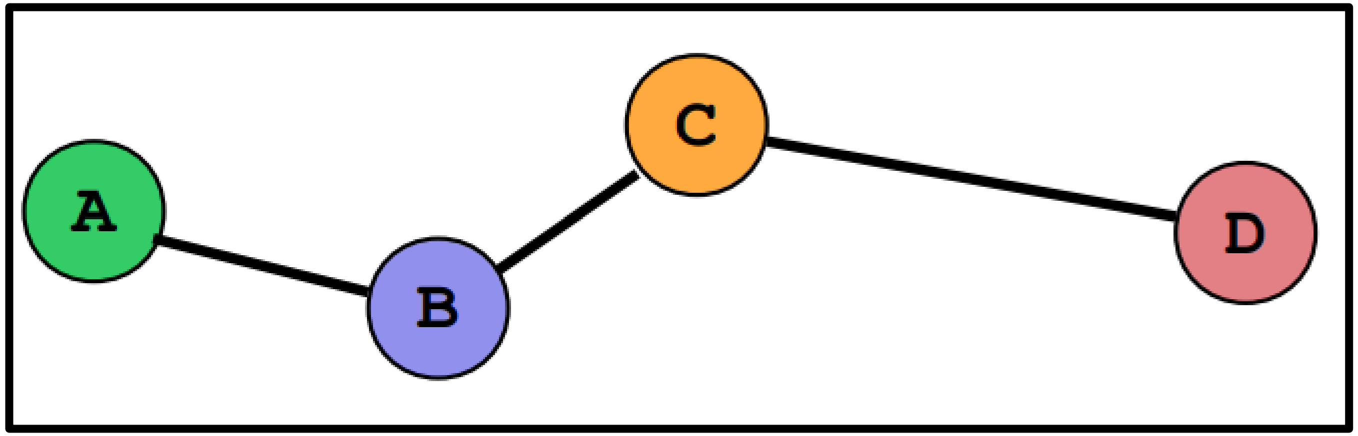


- ✓ **Weighted Ties**. just as networks can contain multiple different kinds of edges between actors, they can also contain relationships of varying strength.
- ✓ For example **A might like B a whole lot**, but **B and C only like each other moderately**.



- ✓ **Group**
  - ✓ A group in a network is just a **subset of the actors which share some characteristic in common**.
  - ✓ If we were to look at an organizational network, **one group could be made up of all actors that work in the human resources department**.
  - ✓ The definition of groups as **commonality on some salient trait allows us to examine a number of network hypotheses** and defined useful measures that are conditional on knowing the group membership of actors.
  - ✓ For example we might **want to test a hypothesis about the number of friendship ties between workers at a company who are part of different departments versus those in the same departments**.

- **Geodesic Distance**
  - It is defined as the least number of connections (ties) that must be traversed to get between any two nodes.
  - For example, in the network depicted below, the **geodesic distance between actor A and actor D is 3**, while the **distance between actor B and C is only 1**.



SOCIAL NETWORK DATA

- ✓There are two main kinds of social network data:-
  - ✓ Edge lists and socio-matricies.
  - ✓Each of these data formats has its own advantages and weaknesses,
    - ✓ mainly having to do with a trade off between ease of entering and storing the data and;
    - ✓ ease of using the data for analysis.

SOCIOMATRIX

- ✓ A Socio-matrix (also known as an Adjacency Matrix): is a way of representing directed or undirected ties between actors using a numerical matrix.
- ✓ There is one column for each actor and one 4 row for each actor.
- ✓ In general, the diagonal elements of this matrix (eg. second row, second column) are always equal, signalling that actors do not tie to themselves.
- ✓ To specify which entry in the matrix we are talking about we always use the same convention: [row i, column j] so that if we were to say the [3,5] entry in the socio-matrix we would be talking about the third row and fifth column.
- ✓ Each row in the socio-matrix represents the ties that Actor i sends to all other actors (j's).
- ✓ As shown in figure, manager one sends a directed friendship tie to manager two, as indicated by the value 1 in the [1,2] entry of the socio-matrix.
- ✓ The upside of taking this approach to storing data about a network is that it naturally encodes the fact that some actors may not send or receive any ties (something we call being a network isolate) and the format is very ready for many statistical analyses.
- ✓ The downside to using this data format is that it can take up a lot of space and be difficult to enter data into by hand.

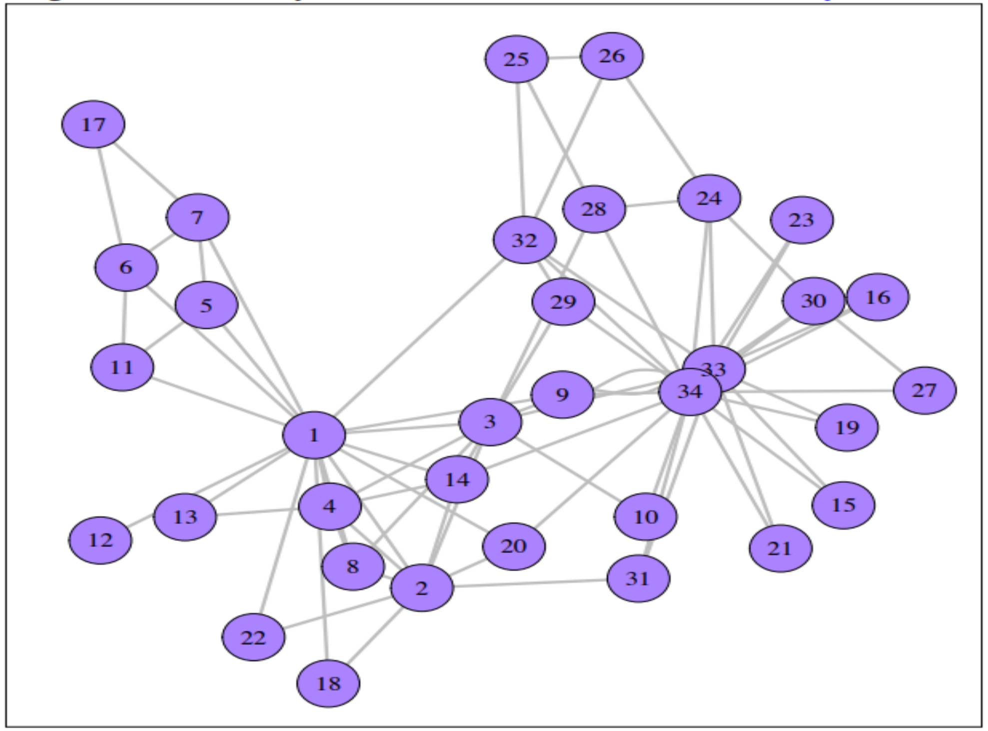
Manager	
1	0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0
2	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1
3	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0
4	1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 0 0
5	0 1 0 0 0 0 0 0 1 0 1 0 0 1 0 0 1 0 1 0 1
6	0 1 0 0 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 0 1
7	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8	0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
9	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
10	0 0 1 0 1 0 0 1 1 0 0 1 0 0 0 1 0 0 0 1 0
11	1 1 1 1 1 0 0 1 1 0 0 1 1 0 1 0 1 1 1 0 0
12	1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1
13	0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
14	0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0
15	1 0 1 0 1 1 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0
16	1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
17	1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 0 0 1 1 1
18	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
19	1 1 1 0 1 0 0 0 0 0 1 1 0 1 1 0 0 0 0 1 0
20	0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0
21	0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0

Socio-matrix of Directed Network of Friendship Ties Between Managers

EDGELIST

1 32	1 4	3 10	6 7	20 34	26 32
1 22	1 3	3 33	7 17	21 34	27 34
1 20	1 2	3 29	9 34	21 33	27 30
1 18	2 31	3 28	9 33	23 34	28 34
1 14	2 22	3 8	9 33	23 33	29 34
1 13	2 20	3 4	10 34	24 30	29 32
1 12	2 18	4 14	14 34	24 34	30 34
1 11	2 14	4 13	15 34	24 33	30 33
1 9	2 8	4 8	15 33	24 28	31 34
1 8	2 4	5 11	16 34	24 26	31 33
1 7	2 3	5 7	16 33	25 32	32 34
1 6	3 14	6 17	19 34	25 28	32 33
1 5	3 9	6 11	19 33	25 26	33 34

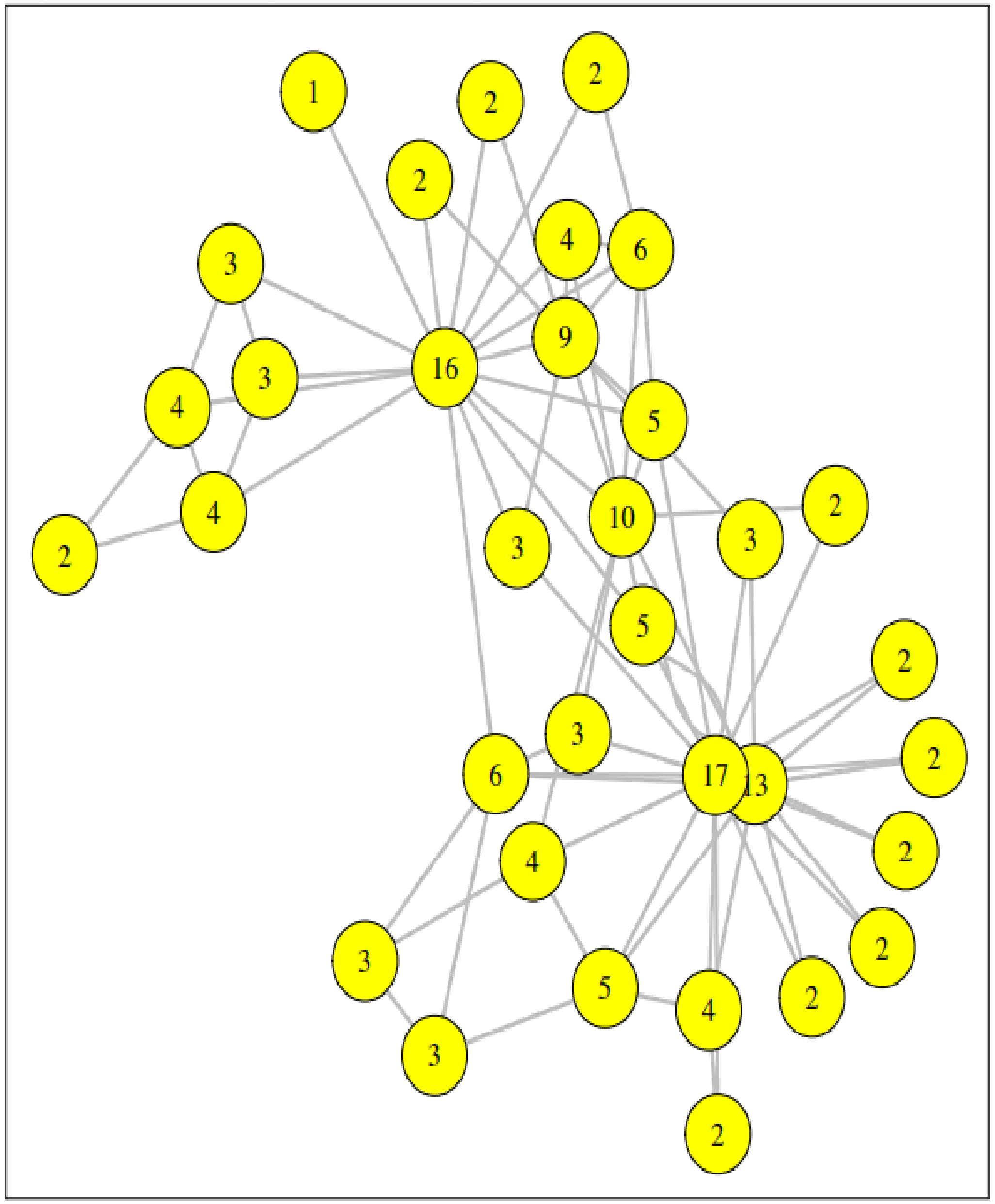
- ✓ An Edgelist is the other primary form of data storage for social network analysis.
- ✓ This only captures information about existing ties so it needs to be supplemented with knowledge of the total number of actors in the network (even if they do not have any ties).
- ✓ In the example edgelist in the Figure above , directed friendship ties for the network shown in the Figure below are presented in edgelist form where the first number on each line denotes the actors sending a tie to the second actor in the row.
- ✓ The figure below is a figure of Karate Club network.
- ✓ This form of data entry is best for storing information about data that are collected by hand as it is very efficient to store and relatively easy to enter, but one must be careful to use a common naming system and keep track of any nodes that do not have any ties to them.



PROPERTIES OF NODE

Properties of Nodes

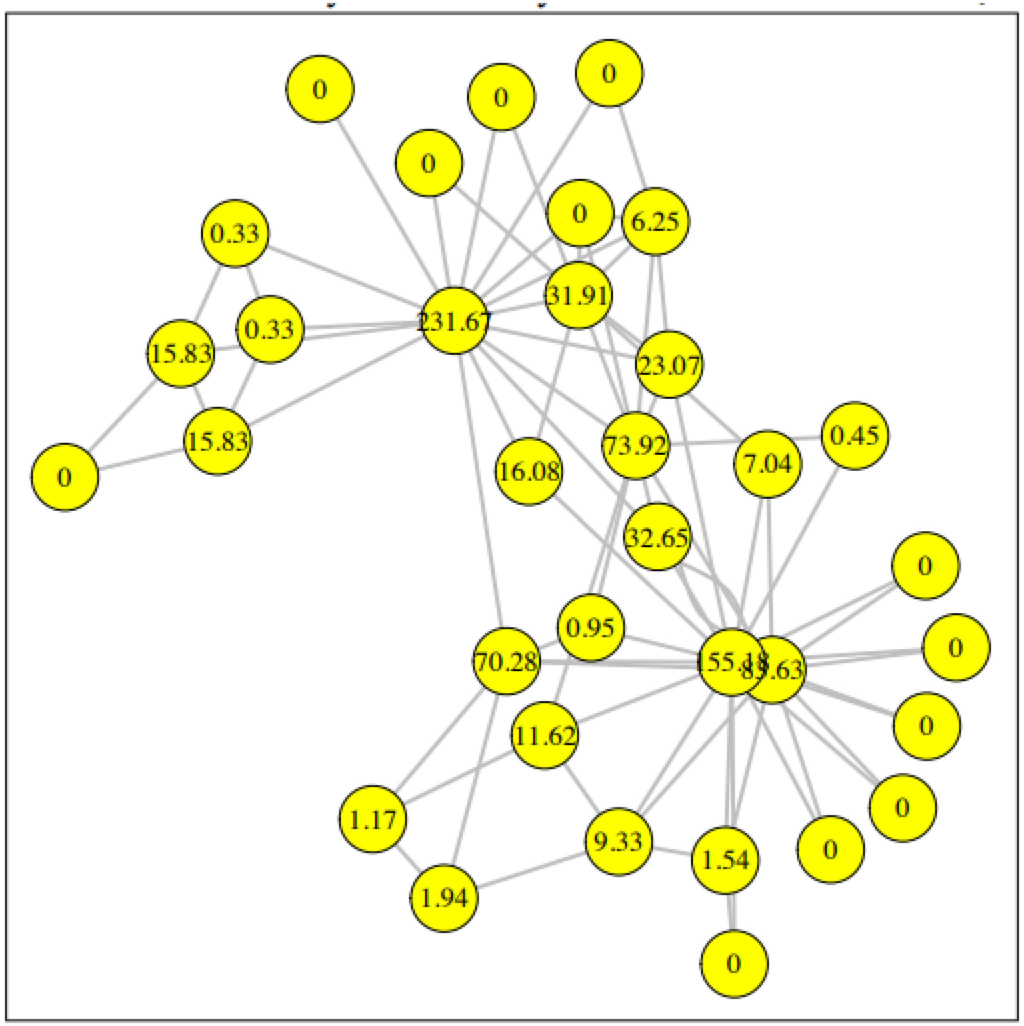
- ✓ Actor level properties that serve as the language for social network analysis.
- ✓ It helps in explaining **how to conceptualize social phenomena and hypotheses in a networks framework .**
- ✓ The goal is to interface and understand theories posed in the literature using a social networks/relational framework.
  - ✓ **Degree Centrality**: is the most basic network measure and captures the number of ties to a given actor.
  - ✓ For **un-directed ties this is simply a count of the number of ties for every actor**. For **directed networks, actors can have both in-degree and out-degree centrality scores**.
  - ✓ As the name implies, **centrality measures how central or well connected an actor is in a network**.
  - ✓ This theoretically signals **importance or power and increased access to information or just general activity level** and **high degree centrality is generally considered to be an asset to an actor**.
  - ✓ Degree centrality is depicted for the Karate club network in Figure where each actor is now labelled with their undirected degree centrality score.



Betweenness Centrality

Betweenness Centrality

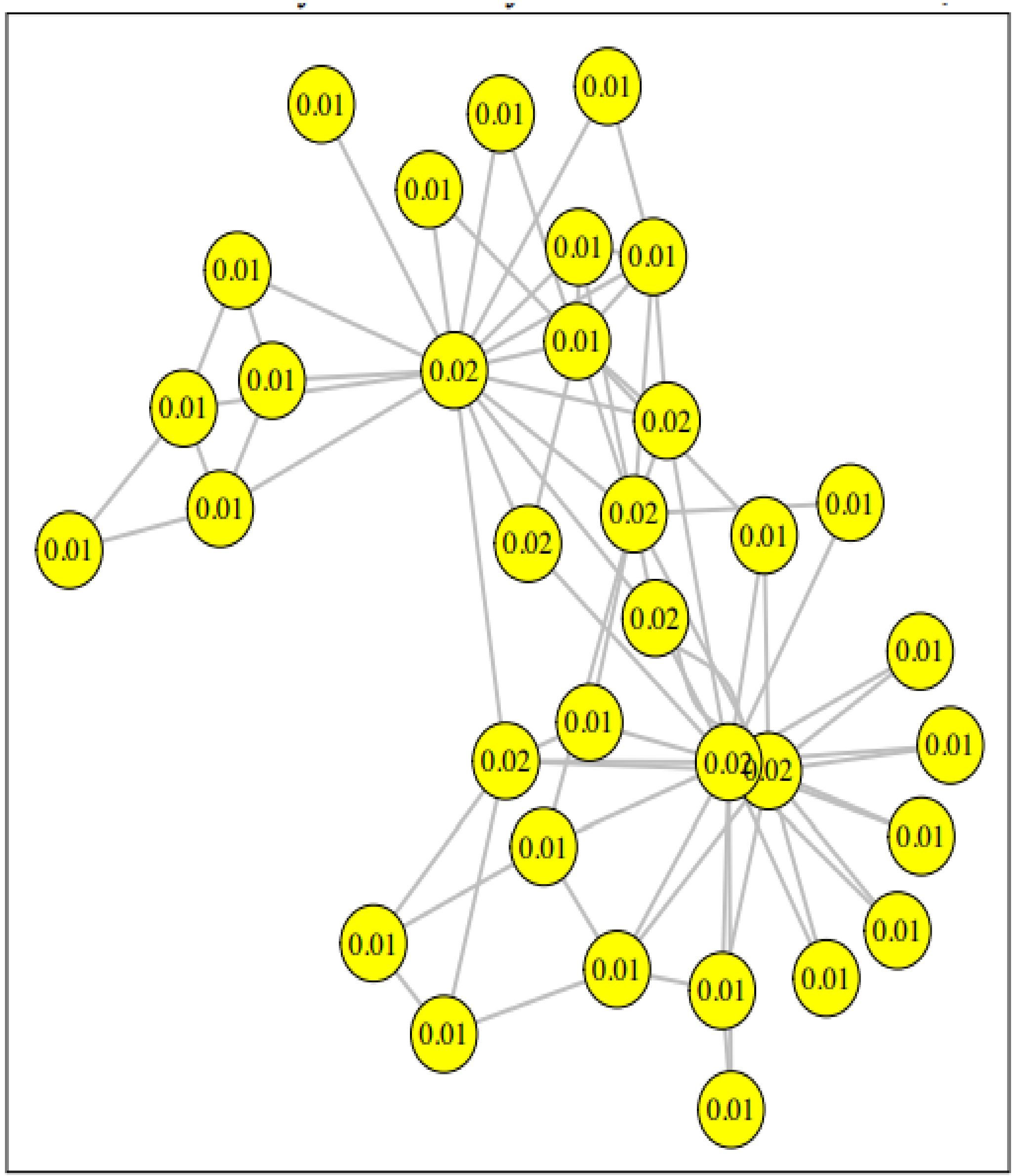
- ✓ It is roughly defined as the number of shortest paths between alters that go through a particular actor.
- ✓ More precisely, it is the **sum of [the shortest path lengths between every set of alters where the path goes through the actor**. Calculating the measure for divided by the shortest path lengths (not necessarily through the target actor) between those actors].
- ✓ This intuitively **measures the degree to which information or relationships have to flow through a particular actor and their relative importance as an intermediary in the network**.
- ✓ Betweenness scores for Karate club network are displayed in figure.



CLOSENESS CENTRALITY

Closeness Centrality

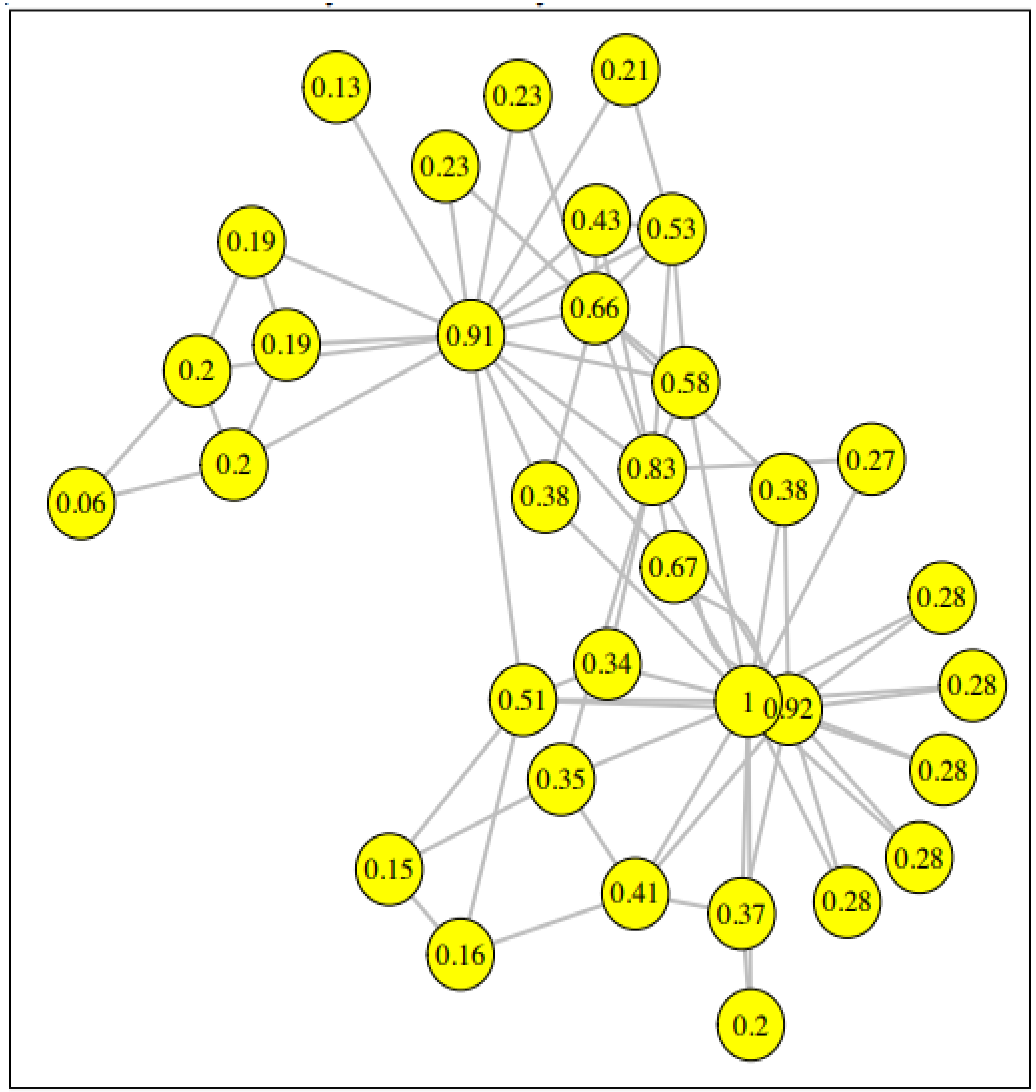
- ✓ It measures **how many steps (ties) are required for a particular actor to access every other actor in the network**.
- ✓ This is measured as 1 divided by the sum of geodesic distances from an actor to all alters in the network.
- ✓ The measure will reach its **maximum for a given network size when an actor is directly connected to all others** in the network and its **minimum when an actor is not connected to any others**.
- ✓ This captures the intuition that short path lengths between actors signal that they are closer to each other.
- ✓ Note that **this measure is sensitive to network size and is decreasing in the number of actors in the network**.
- ✓ This makes intuitive sense in many situations because it gets **more difficult to maintain close relationships with all members of the network as the network grows** but can also be corrected for by multiplying by the number of actors in the network.
- ✓ Closeness scores for Karate club network are displayed in figure.



EIGENVECTOR CENTRALITY

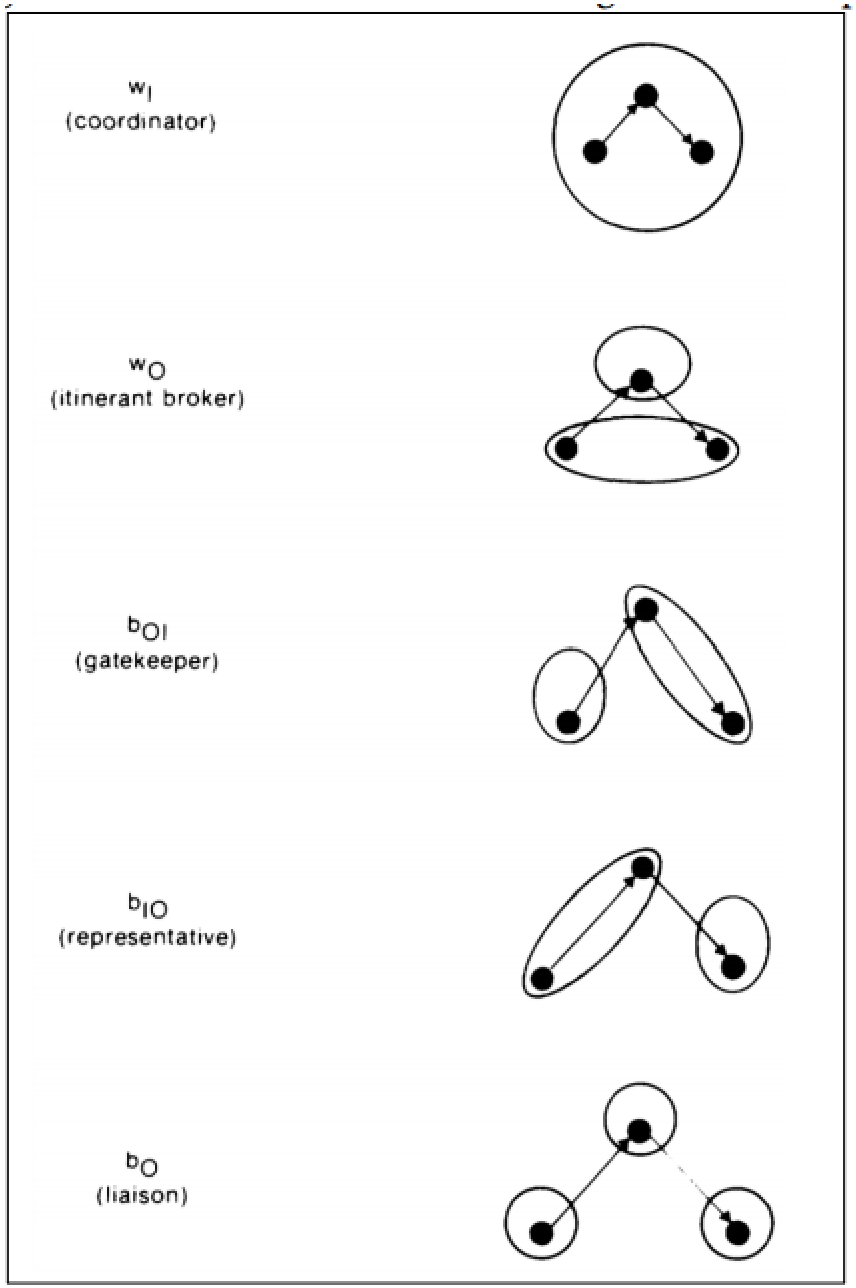
Eigenvector centrality.

- ✓ It measures the **degree to which an actor is connected to other well connected actors**.
- ✓ It takes advantage of a mathematical property of networks (represented as adjacency matrices) that allows for the easy calculation of **how well connected an actor is to other well connected actors**.
- ✓ This measure **captures the value of having a lot of friends in high places**.
- ✓ Eigenvector scores for Karate club network are displayed in figure.



Brokerage

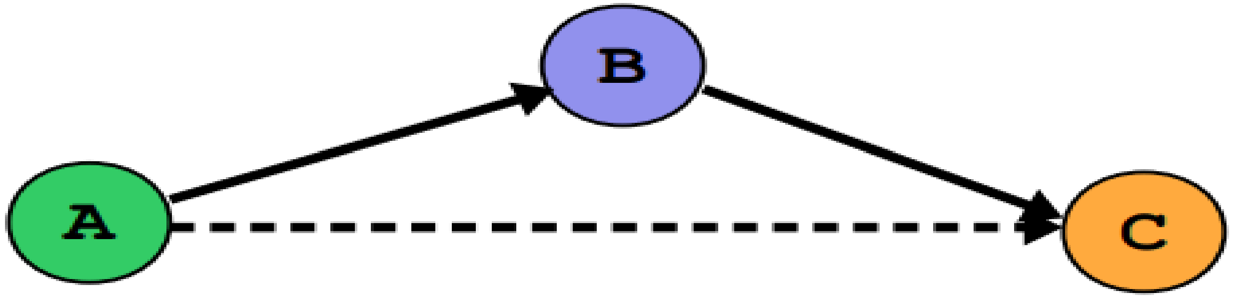
- ✓ It describes the position of actors such that they occupy an advantageous position where they can broker interactions between other actors in the network.
- ✓ Brokerage Centrality is then a measure of the degree to which an actor occupies a brokerage position across all pairs of alters.
- ✓ It is meant to capture the intuition that a broker serves as a go-between and thus can gain benefits from their position as an intermediary.
- ✓ There are five kinds of brokerage relationships, each of
    - ✓ (a) A **Coordinator** is an Actor in the same group as two alters who connects the two nodes. An example might be a graduate student who makes sure that all of the rest of their cohort is made aware of parties being hosted by anyone in their cohort.
    - ✓ (b) An **Itinerant broker** is a member of an outside group that connects two others who share group membership.
    - ✓ (c) A **Gatekeeper** is a member of the same group as the target a member of another group hopes to connect with that can control whether or not that outside actor is able to gain access to the in group member. An example might be a secretary or office manager.
    - ✓ (d) A **Representative** is a member of the same group as an Actor that wishes to connect with an actor outside of the group but has to go through an intermediary. An example is an Ambassador for a country.
    - ✓ (e) A **Liaison** is a member of a group that is distance from two actors that wish to connect but do not share group membership themselves. A delivery truck driver is a good example.



NETWORK RELATIONSHIPS AND STRUCTURES

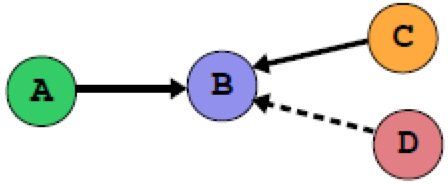
- ✓ **Reciprocity**: is the tendency for directed ties from actor i to actor j be be reciprocated and sent back from actor j to actor i.
  - ✓ This captures the classic finding that **feeling and actions tend to be reciprocated**.

- ✓ **Transitivity**: is the tendency for **friends of friends to be friends and enemies of enemies to be enemies**.
  - ✓ More generally a transitive relationship is one where two nodes being connected to a third increases the likelihood that they will connect themselves

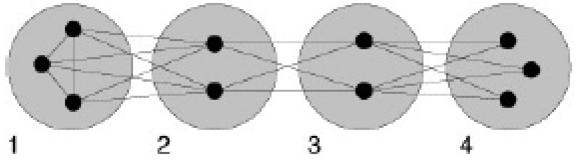


- ✓

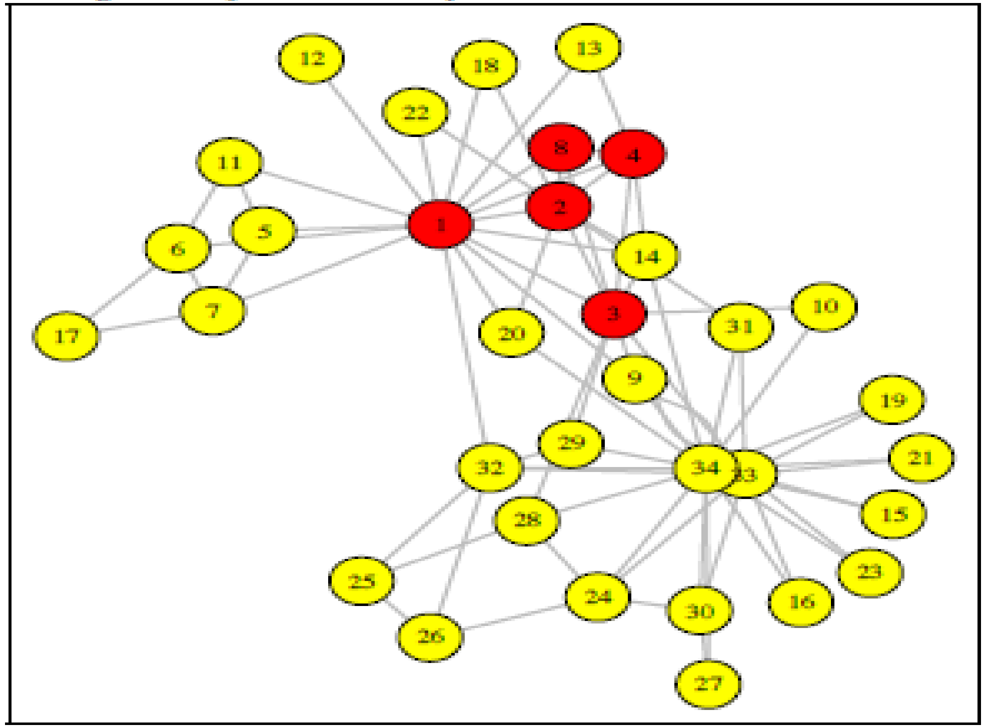
- **Preferential Attachment (Popularity):** expresses the tendency for nodes that are already central to gain more connections at a greater rate than those who are not already central.
  - This is often the case in academia where as a researcher becomes more active and collaborates more in publishing, they are more likely to attract new collaborators who want to work with them.



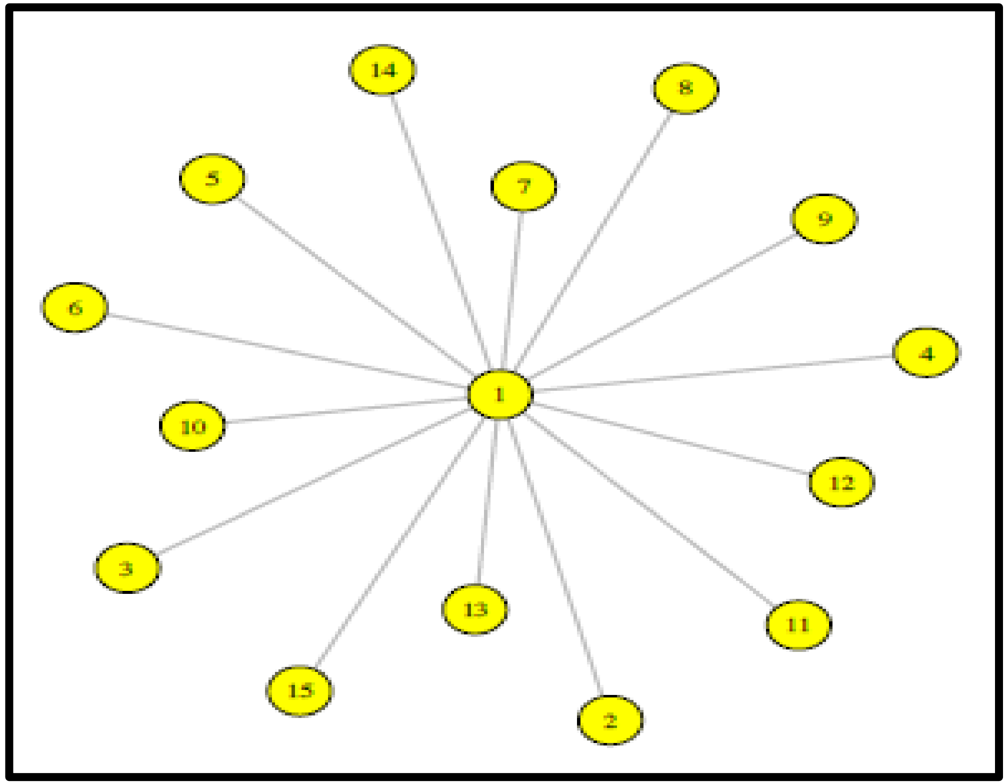
  - **Structural Equivalence**: is a concept that describes actors occupying the **same position in the network relative to all other actors**.
    - In the example figure below, each grey circle contains a set of actors that are structurally equivalent to all others.
    - This concept is important in making comparisons between nodes about their relative importance and position in a network.



- ✓ **A Clique**: is a **subset of actors in a network such that every two actors in the subset are connected by a tie**.
  - ✓ This definition follows the common English language usage of the word meaning a densely connected group.
  - ✓ A large example clique is coloured red in Figure.
  - ✓ Largest Clique in Karate Club Network.



✓**A Star**: is a network structure where all ties connect to one central node, making the shape of a star.

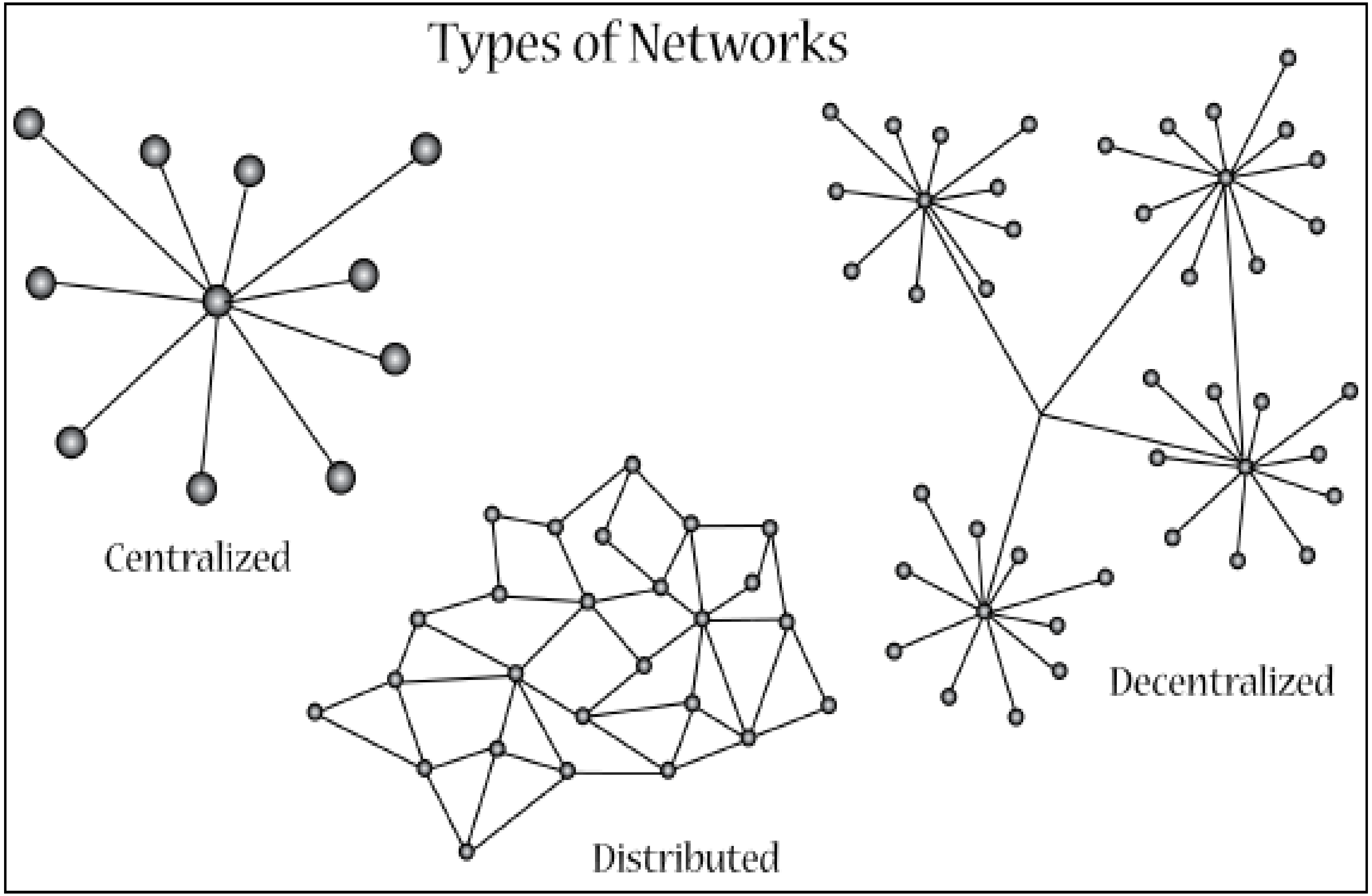


NETWORK PROPERTIES

- ✓ All of the properties discussed above refer to individual actors or subsets of actors in a network.
- ✓ While these are important characteristics to measure, we can also think about properties that a network as a whole exhibits.
- ✓ These properties are important because they impost structure on the entire space of interactions and relationships and can have profound aggregate effects on how actors in the network behave and function as a whole.
- ✓ **Centralization (Degree, Betweenness, Closeness, Eigenvector, etc.):** is a measure of the unevenness of the centrality scores of actors in a network.
  - ✓ It ranges from zero, when every actor is just as central for whatever score we are interested in, to 1, when one node is maximally central and all others are minimally central.
  - ✓ This measure is a good way to express the idea that there are couple of very powerful or important actors in a network or that power/importance is spread out evenly in one simple measure.

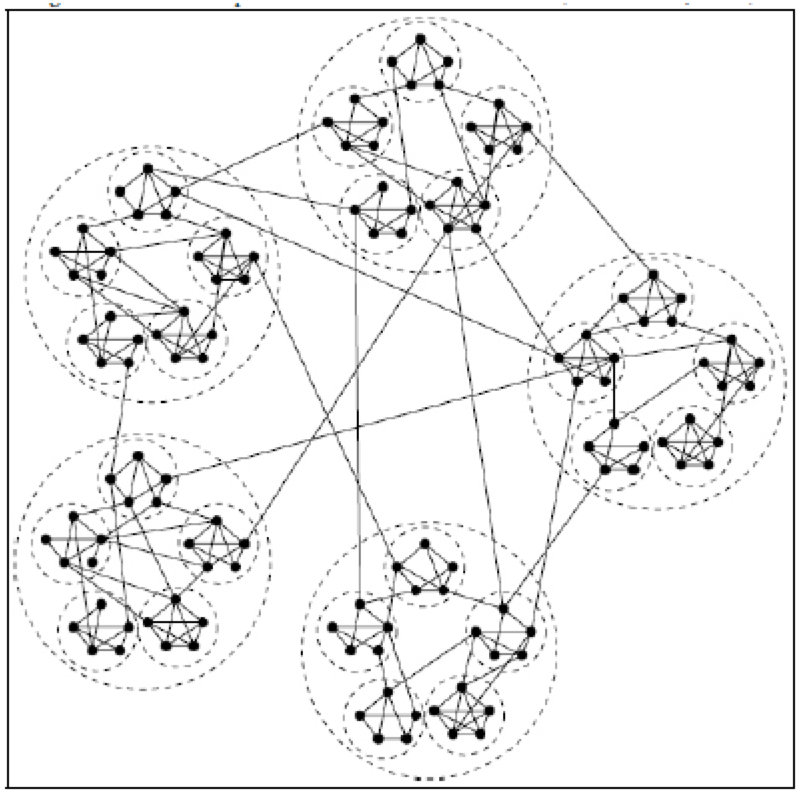
TYPES OF NETWORK

An example of a highly **centralized** network and for comparison, a **decentralized** network (small centralized components that are connected), and a **distributed** network.



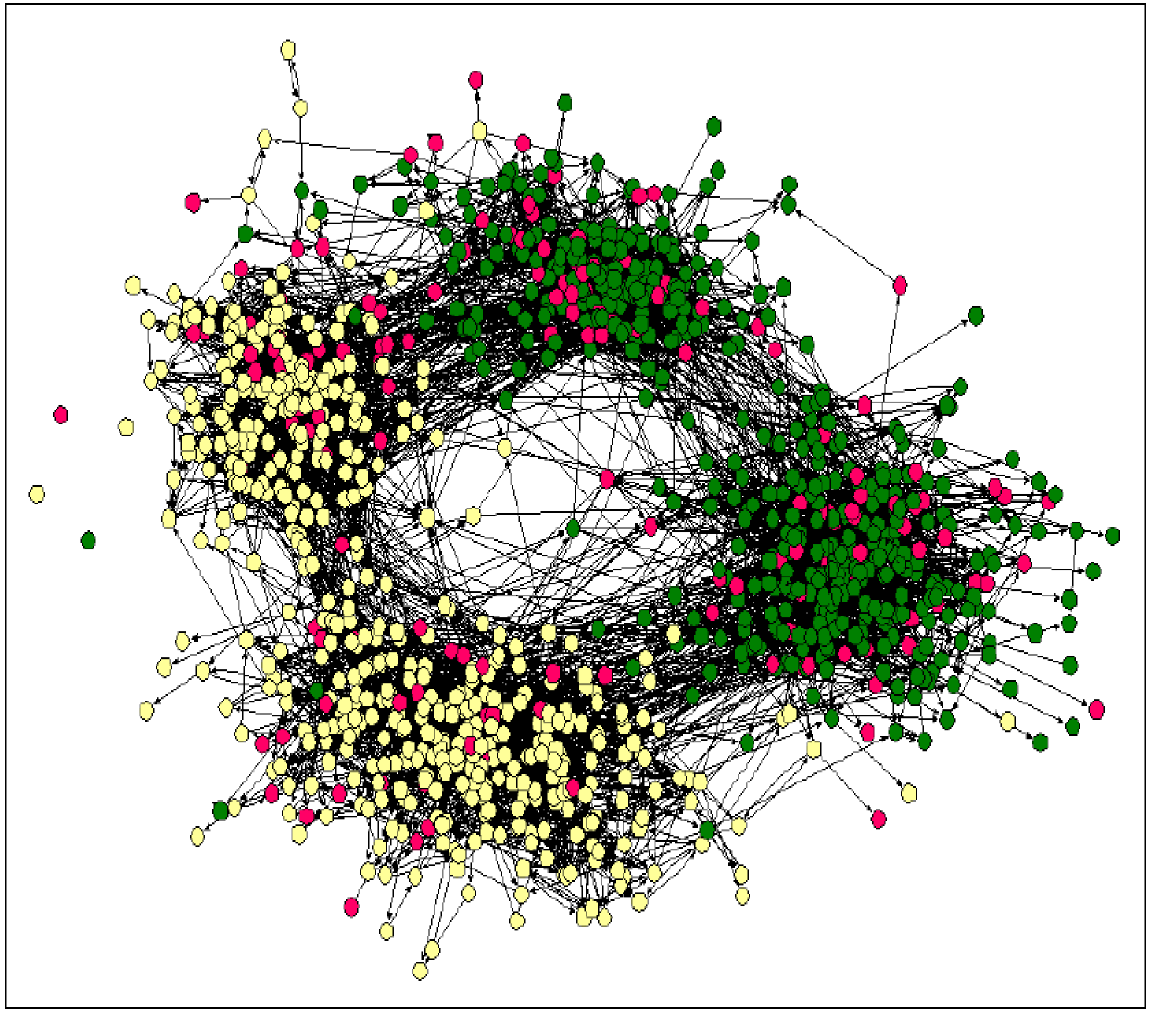
CLUSTERING

- ✓The network **Clustering Coefficient:** measures the degree to which actors form ties in dense, relatively unconnected (between groups) groups.
- ✓This measure is agnostic about why the network is clustered.
- ✓ The degree of clustering in a network is related to the efficiency with which information can diffuse over the network, as well as its robustness to disruption.
- ✓An example of clusters within a network, is shown in a figure.



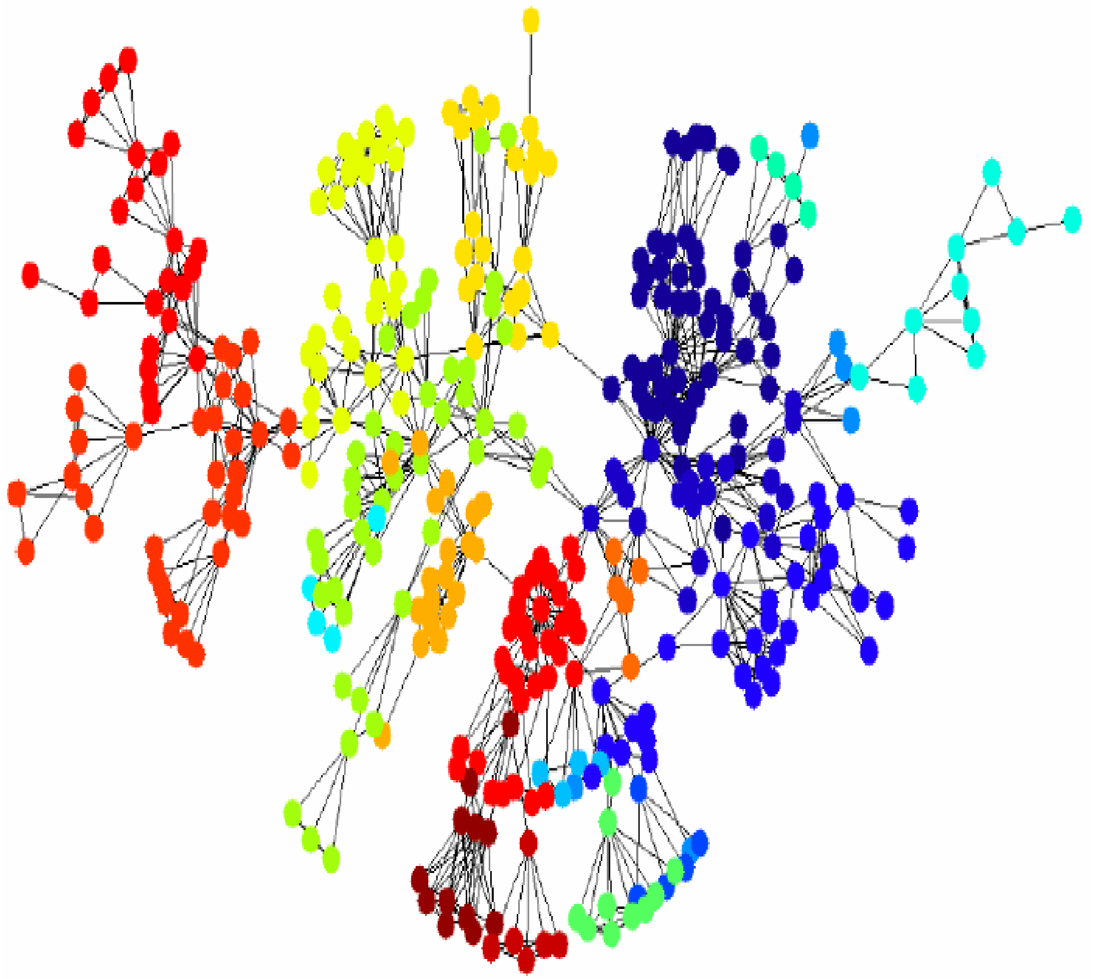
OTHER TERMONOLOGIES

- **Homophily** is a process where actors who are similar on a particular trait are more likely to form ties.
  - This process is the basis for the commonly used phrase “birds of a feather, flock together”.
- **Heterophily**, the mirror opposite of homophily, is defined as the degree to which pairs of individuals who interact are different in certain attributes".
  - An example of **heterophily** would be to individuals from different ethnic and socio-economic backgrounds becoming friends.
- Racial homophily in a secondary school friendship network is shown in a figure.
- Nodes are connected if students are friends and coloured by race with yellow and green nodes forming two distinguishable groups and even smaller minority students (red) in both main clusters.



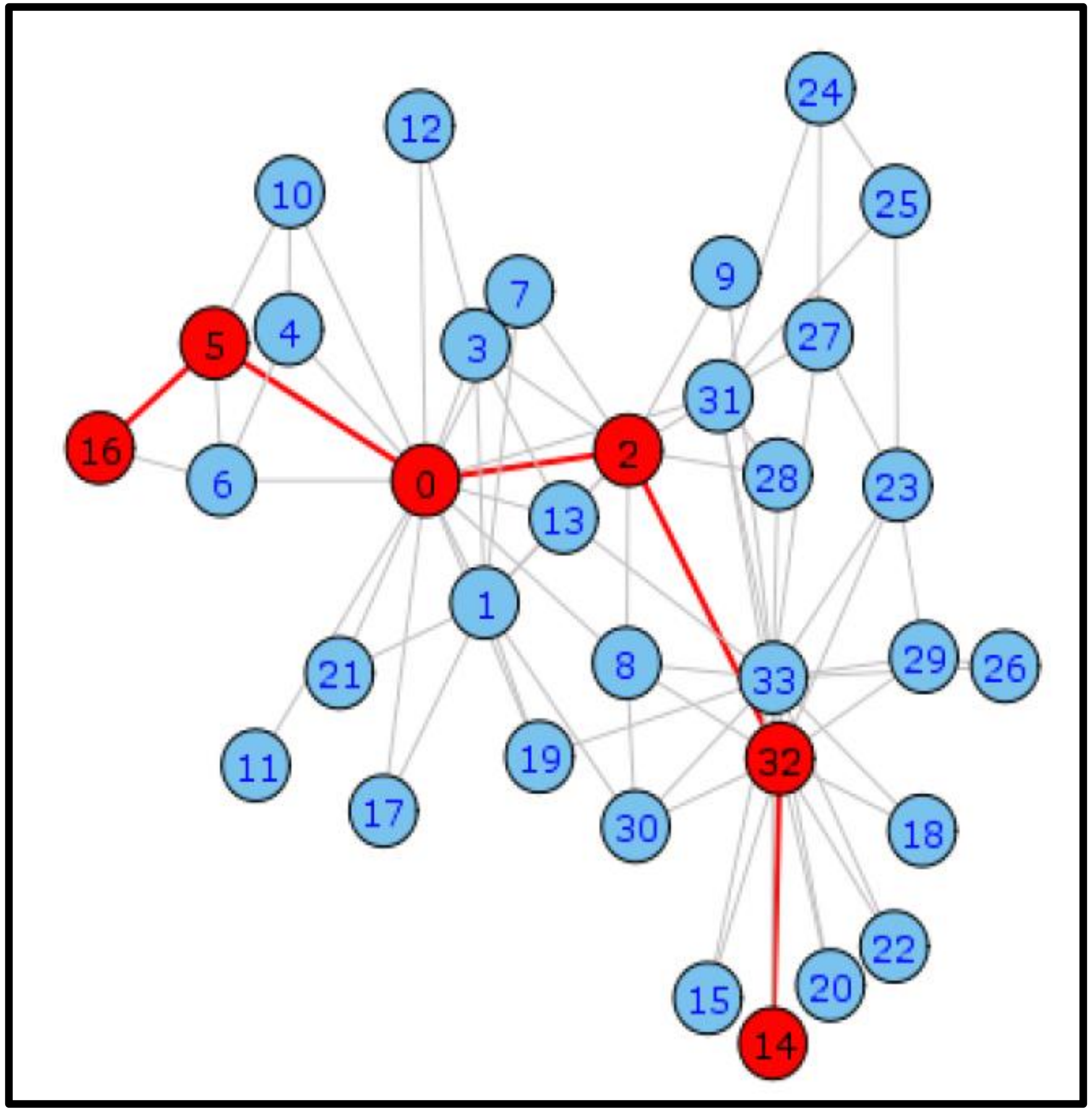
MODULARITY

- ✓ **Modularity** is a measure of the degree to which a network displays **Community Structure**, with clusters that are not densely connected to others but densely connected within cluster.
- ✓ This measure is very difficult to calculate, but provides a way to identify community structure on a network where one is unsure if such a structure exists.
- ✓ An example of community structure between authors of papers about network analysis is shown in the figure.
- ✓ The largest connected component of citation network for authors publishing on networks with actors coloured by community membership is shown on the left.



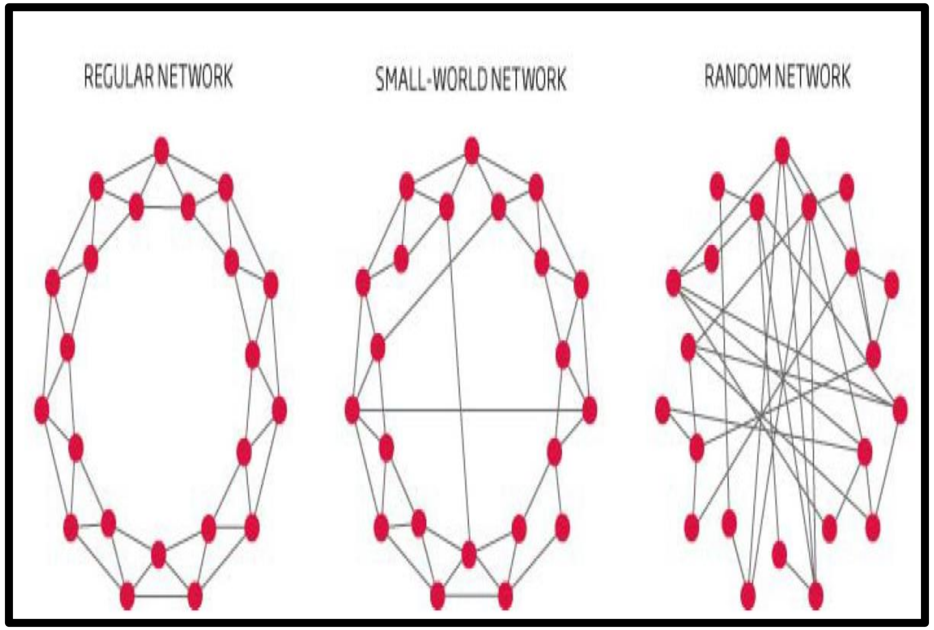
DIAMETER

- ✓ The **Diameter** of a network is defined as the longest of all the calculated shortest paths between actors.
- ✓ Network diameter gives us an idea about how easily reachable Actors are on a network.
- ✓ A very large diameter means that even though there is theoretically a way for ties to connect any two actors through a series of intermediaries, there is no guarantee that they actually will be connected.
- ✓ Diameter is thus a signal about the ability for information or disease to diffuse on the network.
- ✓ The diameter of Karate club network is displayed graphically in Figure.

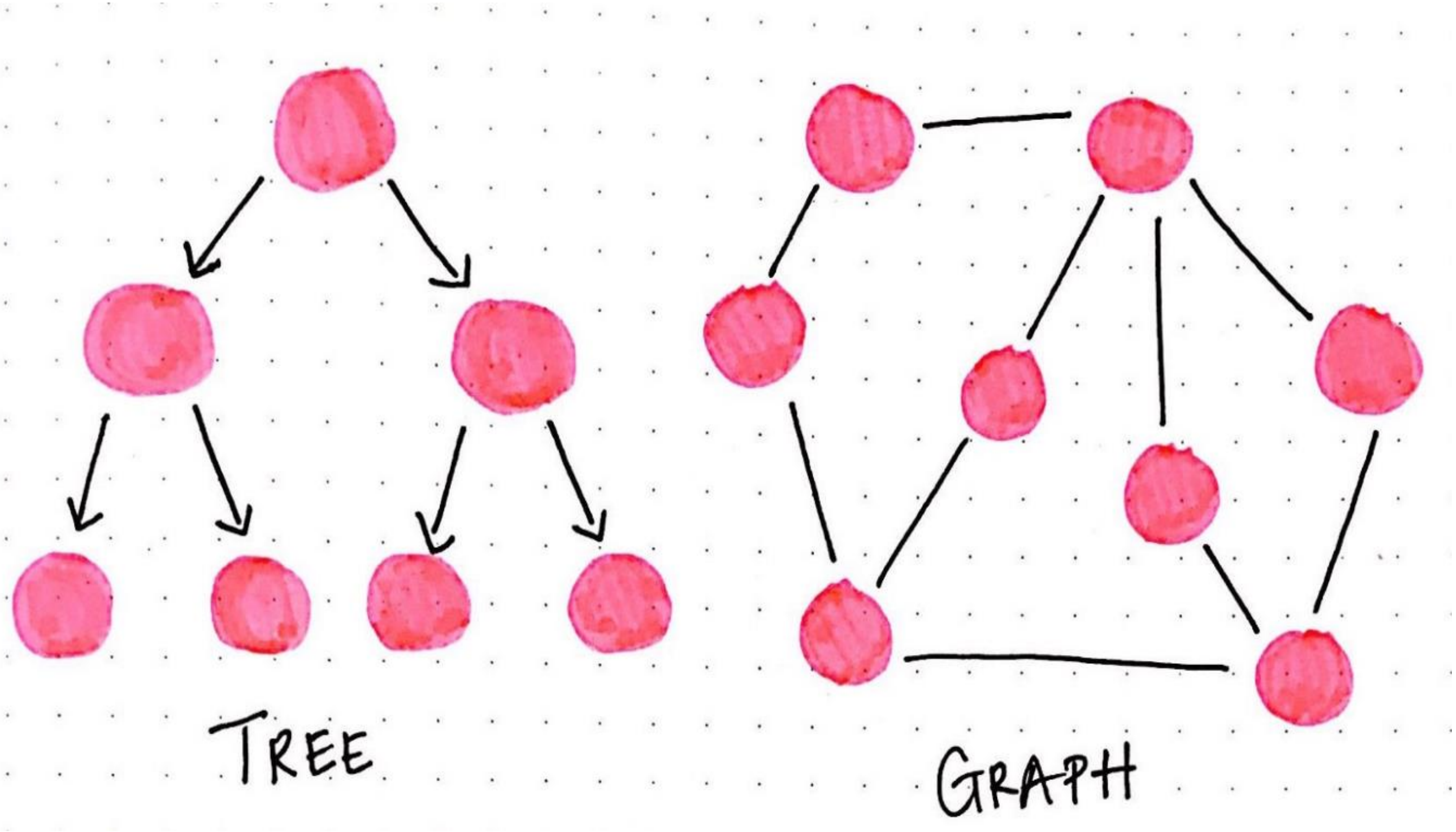


NETWORK TYPES

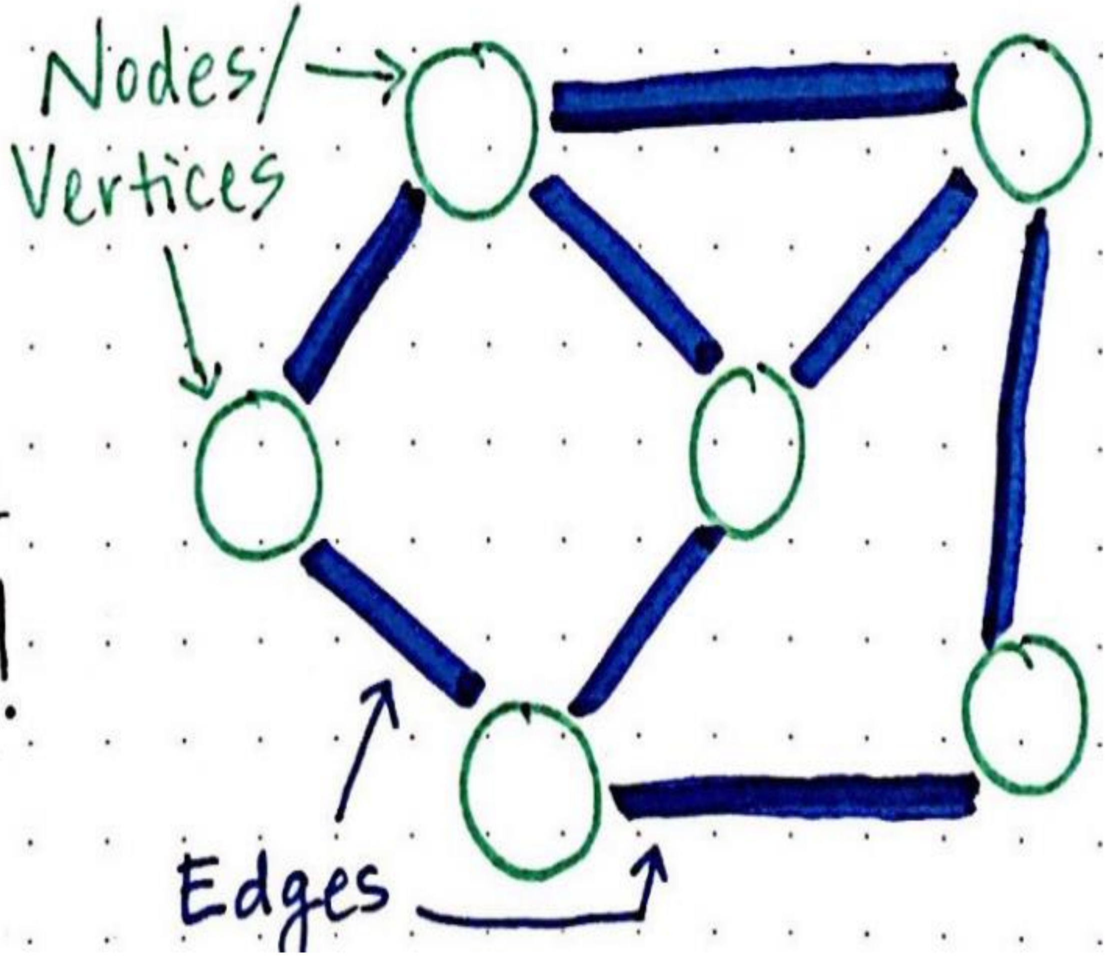
- There are a number of classic **Network Types** that can be used to characterize the stereotypical social structure in different situations.
- Regular networks are characterized by all actors having the same degree and are often a starting point for simulation studies of networks.
- Small world networks are very efficient for information transfer in that most nodes are not connected (so a high degree of clustering) but also have a relatively short average path length between actors.
- Random networks are very robust to disruptions but may be difficult for people to maintain, especially if ties are across long distances.
- An example of a Regular Network (all actors have the same degree and are structurally equivalent to each other).
- A Small World Network where dense clusters are connected by random and far reaching ties.
- Random Network, where actors are randomly connected and there is no discernible structure.



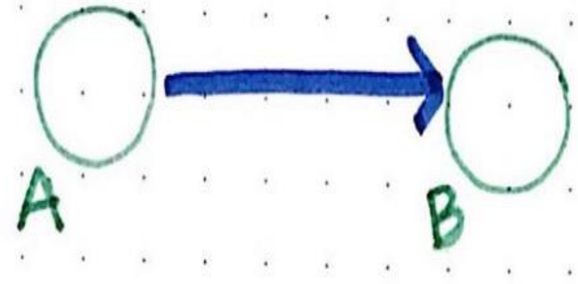
SUMMARY: GRAPH THEORY IN RELATION TO SOCIAL NETWORK PLATFORMS


TREE
GRAPH

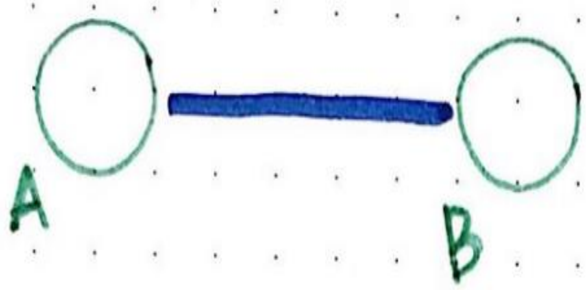
Edges can connect nodes in <u>any</u> possible way! No rules!



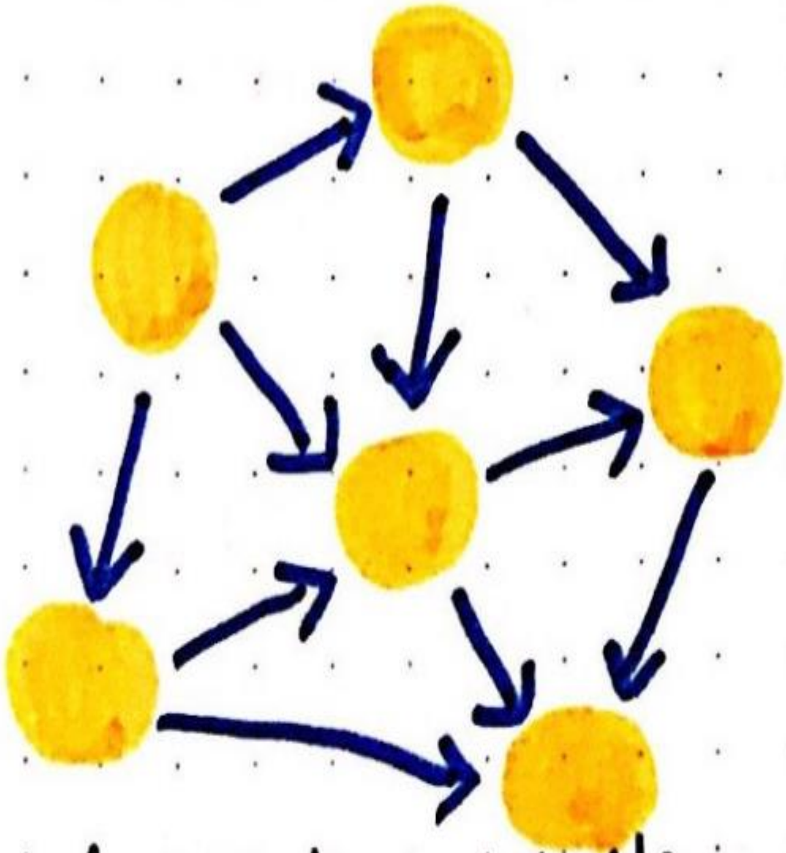
Different types of edges in graphs



directed edge: there is only a path from A, the origin, to B, the destination



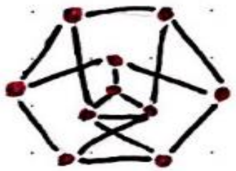
undirected edge: the path between A and B is bidirectional, meaning origin & destination are not fixed.



directed graph/digraph

undirected graph

A VERY BRIEF INTRODUCTION TO *graph theory*



- Graphs are a way to formally represent a network, or a collection of inter connected objects.
- In mathematics, graphs are defined as ordered pairs, with two parts: vertices + edges.

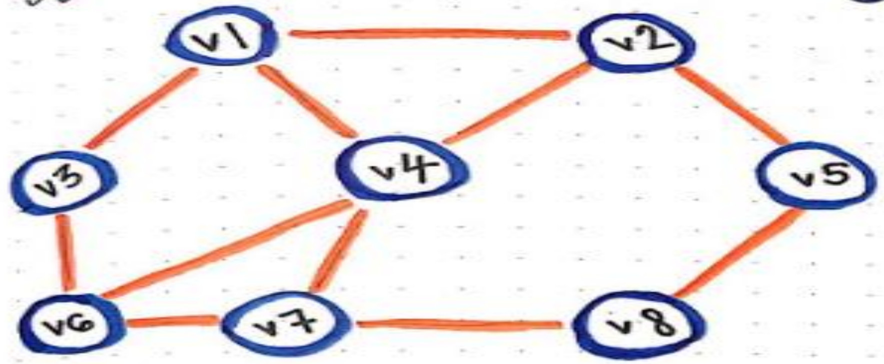
So, what's the definition of a graph?

it looks like this!

$$G = (V, E)$$

where **V** is a set of nodes, also called vertices and **E** is a set of edges, also called links.

(Formally) Defining a Graph



8 vertices/nodes

11 edges/links

$$G = (V, E)$$

$$V = \{v1, v2, v3, v4, v5, v6, v7, v8\}$$

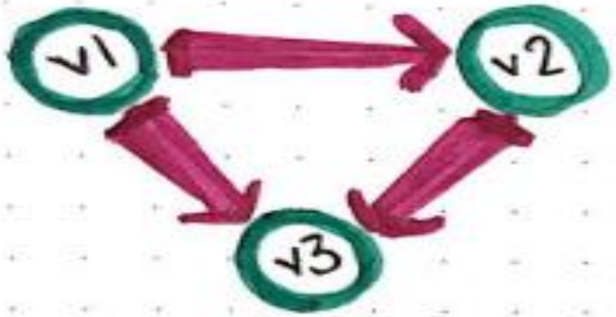
$$E = \{ \{v1, v2\}, \{v1, v3\}, \{v1, v4\}, \{v2, v4\}, \{v2, v5\}, \{v3, v6\}, \{v4, v6\}, \{v4, v7\}, \{v5, v8\}, \{v6, v7\}, \{v7, v8\} \}$$

these edge definitions are <u>unordered</u> pairs!

→ G=(V,E) is the formal mathematical notation for defining graphs.

→ A graph G is an ordered pair of a set V vertices and E, a set of edges.

→ An ordered pair is a pair of mathematical objects in which the order of objects in the pair <u>matters</u>.



But what about a directed graph?

G = (V, E)

how would our edge objects be different?

V = {
v1,
v2,
v3
}

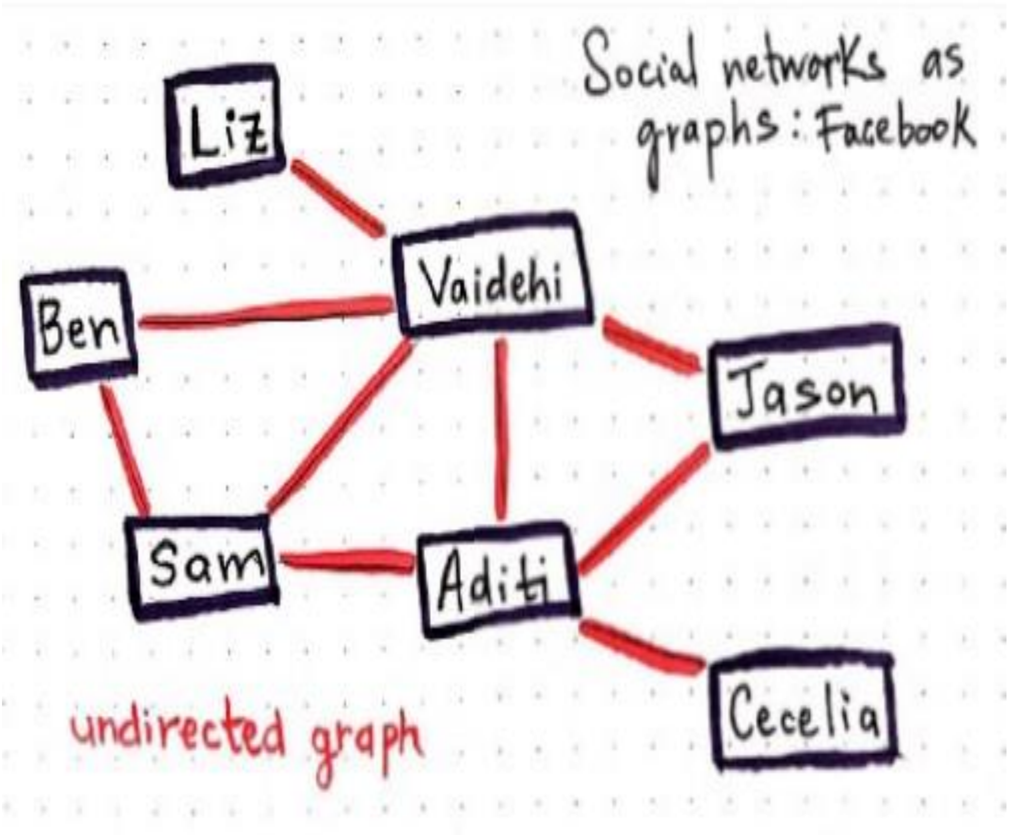
E = {
(v1, v2),
(v1, v3),
(v2, v3),
}

these edge definitions are ordered pairs, because direction matters!

SUPER SOCIAL GRAPHS

Super social graphs

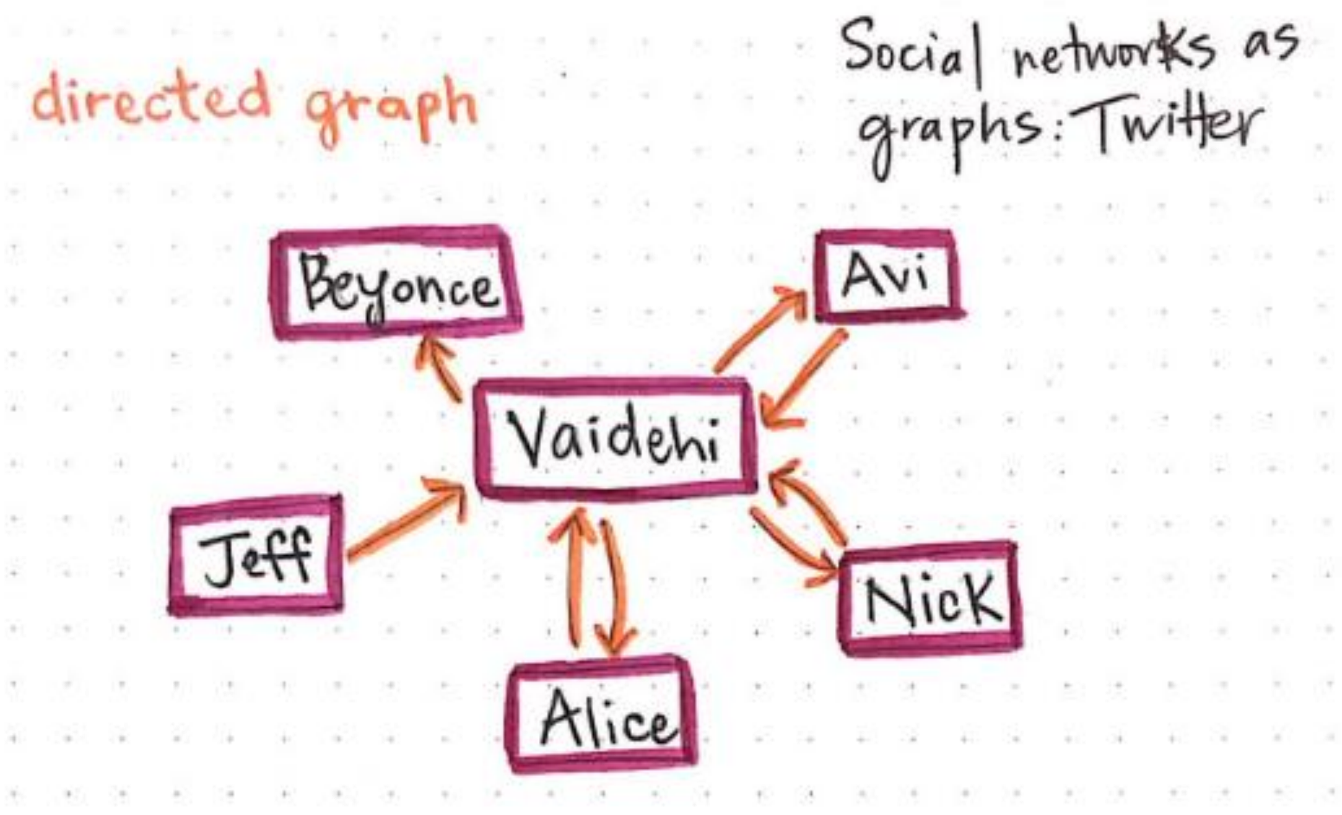
- ✓ The web is a massive graph structure. When we click between websites and navigate back and forth between URLs, we're really just navigating through a graph.
- ✓ Sometimes those graphs have nodes with edges that are undirected – one can go back and forth from one webpage to another — and others that are directed — one can only go from webpage A to webpage B, and never the other way around.
- ✓ Social Networks.
  - ✓ Facebook, a massive social network, is a type of graph.
  - ✓ On Facebook, if one adds you as a friend, you must accept the request. It's not possible for to be your friend on the network without you accepting the request.
  - ✓ The relationship between two users (read: nodes or vertices in graph terms!) is *bidirectional*.
  - ✓ There's no concept of an "origin" and a "destination" node — instead, you're friend of '**A**' and '**B**' is friend of yours.
  - ✓ Facebook relationship is *undirected graph*. Relationships are two-way, so if one has to define Facebook's friend network as a graph, its edges would all end up being unordered pairs.



Facebook as an undirected graph structure

TWITTER

- ✓ Twitter, on the other hand, works very differently from Facebook.
- ✓ One can follow you, but you might not follow him back.
- ✓ Case in point: Vaidehi follows Beyonce, but she does not follow Vaidehi.
- ✓ Twitter can be represented as a *directed* graph.
- ✓ Each edge you create represents a *one-way* relationship.
- ✓ When you follow some one on Twitter, you create an edge in the graph with your account as the origin node, and other's account as the destination node.
- ✓ So what happens when the other person follow you back? – He creates a second edge, with *his* account as the origin node and yours as the destination.



Twitter as a directed graph structure

REFERENCES

Matthew Denny, Social Network Analysis, Institute for Social Science Research, University of Massachusetts Amherst.
https://medium.com/basecs/a-gentle-introduction-to-graph-theory-77969829ead8.
Aral, Sinan and Marshall Van Alstyne. The Diversity-Bandwidth Trade-off. American Journal of Sociology, 117(1):90–171, 2011. http://www.jstor.org/stable/10.1086/661238.
Aral, Sinan and Dylan Walker. Identifying influential and susceptible members of social networks. Science (New York, N.Y.), 337(6092):337–41, July 2012.
http://www.ncbi.nlm.nih.gov/pubmed/22722253.
Aral, Sinan, Erik Brynjolfsson, and Marshall Van Alstyne. Information, Technology, and Information Worker Productivity. Information Systems Research, 23(3-part-2):849–867, September 2012. http://pubsonline.informs.org/doi/abs/10.1287/isre.1110.0408. 17
Aral, Sinan, Lev Muchnik, and Arun Sundararajan. Engineering Social Contagions: Optimal Network Seeding in the Presence of Homophily. Network Science, pages 0–43, 2013.
Callaway, Duncan S., Mark E. J. Newman, Steven H. Strogatz, and Duncan J. Watts. Network robustness and fragility:

Straits, Bruce C. Ego-net diversity: Same- and cross-sex coworker ties. Social networks, 18:29–45, 1996. http://www.sciencedirect.com/science/article/pii/0378873395002545.
Sundararajan, Arun, Foster Provost, Gal Oestreicher-Singer, and Sinan Aral. Information in Digital, Economic and Social Networks. SSRN Electronic Journal, 3, 2012. http://www.ssrn.com/abstract=2141682.
Suri, Siddharth and Duncan J Watts. Cooperation and contagion in web-based, networked public goods experiments. PloSone, 6(3):e16836, January 2011.
http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3055889&tool=pmcentrez&rendertype=abstract. Travers, J and S Milgram. An experimental study of the small world problem. Sociometry, pages 425–443, 1969.
Ward, Michael D., Katherine Stovel, and Audrey Sacks. Network Analysis and Political Science. Annual Review of Political Science,
http://www.annualreviews.org/doi/abs/10.1146/annurev.polisci.12.040907.115949.
Wasserman, Stanley and Katherine Faust. Social Network Analysis: Methods and Applications, volume 8 of Structural analysis in the social sciences, 8. Cambridge University Press, 1994. http://www.amazon.com/dp/0521387078.
Watts, Duncan J. and Steven H. Strogatz. Collective dynamics of ’small-world’ networks. Nature, 393(6684):440–2, June 1998.

http://www.ncbi.nlm.nih.gov/pubmed/9623998. 17, 17 Watts, Duncan J., Peter Sheridan Dodds, and Mark E .J. Newman. Identity and Search in Social Networks. Science (New York,
N.Y.), 296(5571):1302–5, May 2002.
http://www.ncbi.nlm.nih.gov/pubmed/12016312. 1 Zachary, Wayne W. An Information Flow Model for Conflict and Fission in Small Groups. Journal of Anthropological Research, 33(4):452–473, 1977.
http://www.maths.tcd.ie/~mnl/store/Zachary1977a.pdf. 1, 3, 6, 7, 7, 8, 9, 11, 16 Zhang, Yan, A. J. Friend, Amanda L. Traud, Mason A. Porter, James H. Fowler, and Peter J. Mucha. Community Structure in Congressional Cosponsorship Networks. Physica A, 387(7):1705–1712, 2008. http://www.sciencedirect.com/science/
article/pii/S037843710701206X. 15