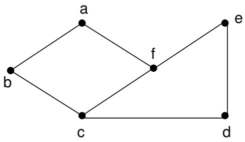
CSI 445/660 – Part 1
(Graph Theory Basics)

Ref: Chapter 2 of [Easley & Kleinberg].

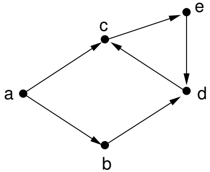
Types of Graphs

- **Undirected** and **Directed**.



Undirected graph:

- **Example:** Friendship relation among people.
- A **symmetric** relationship.

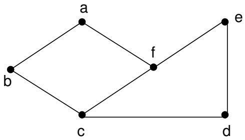


Directed graph:

- **Example:** Follower relationship in Twitter.
- May not be symmetric.

Undirected Graphs: Notation and Definitions

Example:



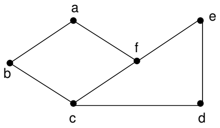
Notation: $G(V, E)$

$$V = \{a, b, c, d, e, f\} \quad (\textbf{nodes} \text{ or } \textbf{vertices})$$

$$E = \{ \{a,b\}, \{a,f\}, \{b,c\}, \{c,d\}, \{c,f\}, \{d,e\}, \{e,f\} \} \quad (\textbf{edges})$$

$|V|$ = No. of nodes = 6 $\qquad$ $|E|$ = No. of edges = 7

Notation and Definitions (continued)



Definition: The **degree** of a node v is the number of edges **incident on** v.

Example: Degree of a $=$ 2, degree of f $=$ 3.

Some observations:

- Sum of the degrees of all the nodes
  $$\begin{aligned} &= \text{Degree(a)} + \text{Degree(b)} + \ldots + \text{Degree(f)} \\ &= 2 + 2 + 3 + 2 + 2 + 3 = 14 \ (\textbf{even}) \\ &= 2 \times \text{No. of edges.} \end{aligned}$$

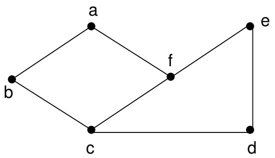
- Nodes with **odd** degree $= \{$c, f$\}$; thus, the number of nodes of odd degree is **even**.

Notation and Definitions (continued)

Theorem: **[First Theorem of Graph Theory]**

In any undirected graph, **the sum of the degrees of all the nodes is equal to twice the number of edges**.

Corollary: In any undirected graph, the **number of nodes of odd degree is even**.

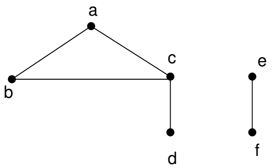


Examples of paths in graph G:

- **a – f – e – d**
- **a – b – c – f – e – d**

- There is a path between every pair of nodes.
- Graph G is **connected**.

Notation and Definitions (continued)

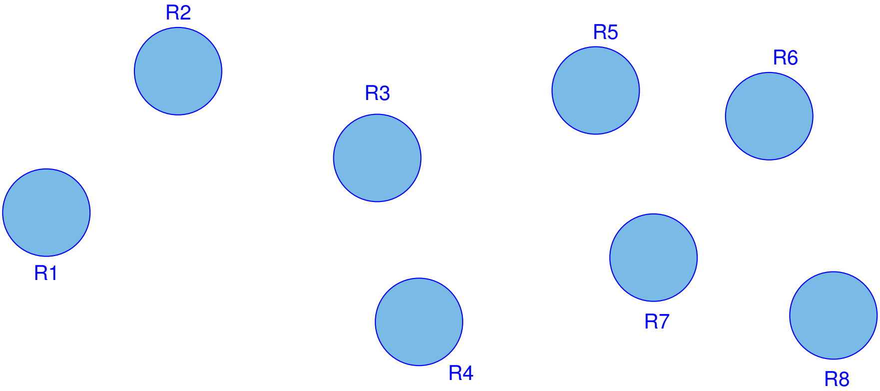


- **Disconnected** graph.
- Has two **connected components**.

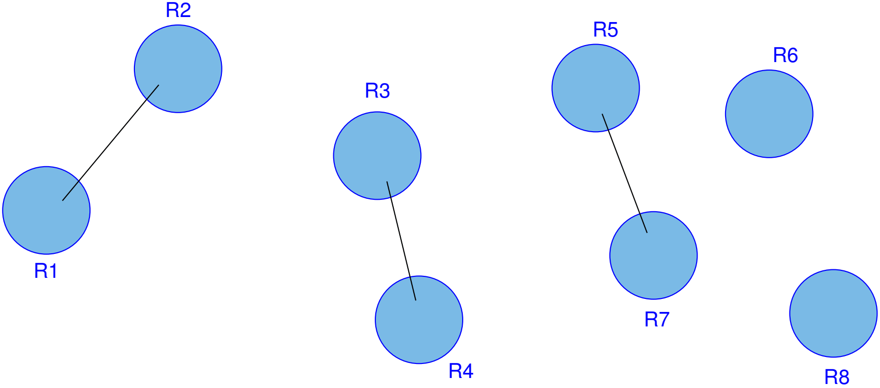
Evolution of a large social network: Imagine the following global friendship graph.

- One node per person in the world
  (No. of nodes $\approx$ 7.3 billion).
- An edge between each pair of friends.

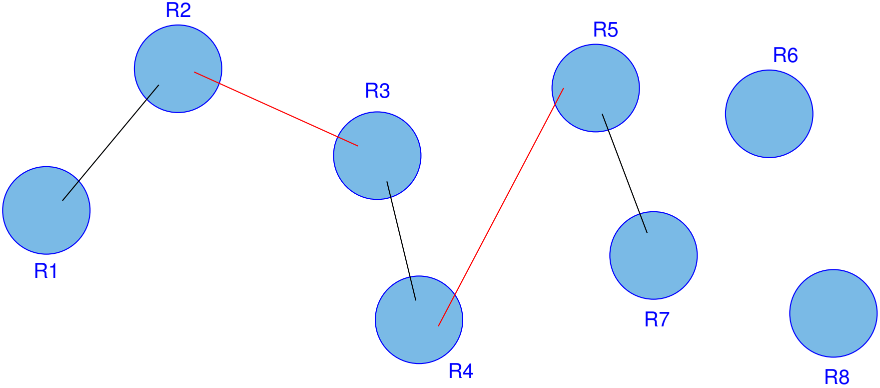
Friendship Network Evolution



Friendship Network Evolution (continued)



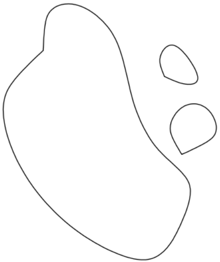
Friendship Network Evolution (continued)



Friendship Network Evolution (continued)

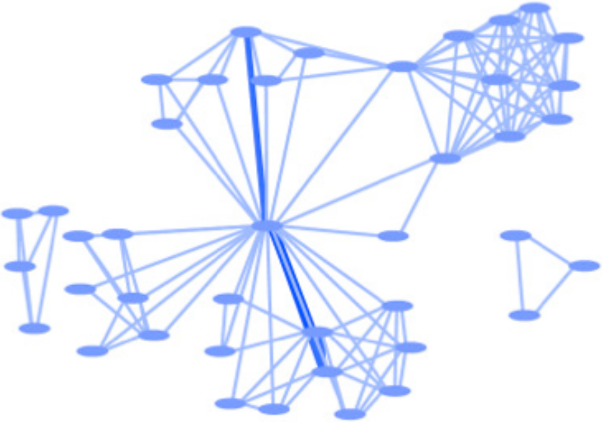
- Components get **merged** over time.
- The graph is likely to contain paths between people in remote parts of the world.
- A large subset of the nodes are in one component, called the **giant component**. (This is typical of many social networks arising in practice.)

An Illustration by Prof. Alistair Sinclair (UC Berkeley):

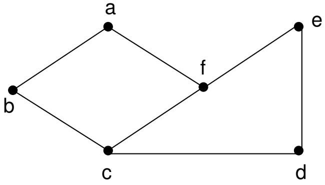


Giant Component: Another Example

Collaboration graph at a research center (from [EK]):



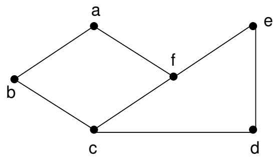
Shortest Paths

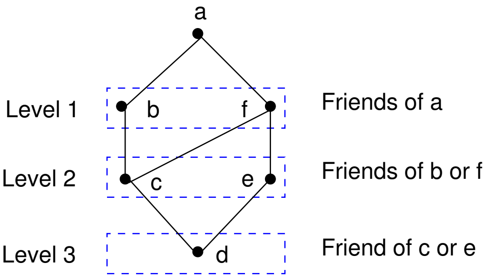


Paths between a **and** e:

- **a – f – e** : Length = 2 (No. of edges)
- **a – b – c – f – e** : Length = 4
- There is no path between a and e with length < 2.
- So, **a – f – e** is a **shortest path** between a and e.
- Shortest paths can be found using a procedure called **breadth-first-search** (BFS).

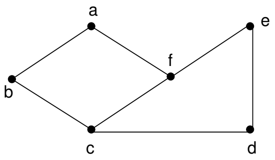
Breadth-First-Search: Example I

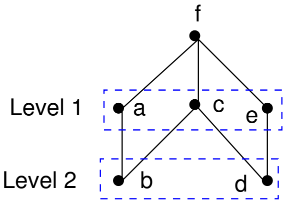




Observation: Each node is within a distance of 3 from node a.

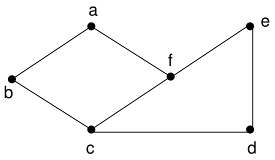
Breadth-First-Search: Example II





Observation: Each node is within a distance of 2 from node f.

Definition of Diameter



Shortest Path Lengths: (Partial list)

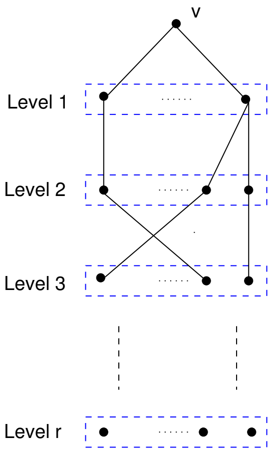
Node pair	Shortest Distance
a, b	1
a, c	2
a, d	**3**
⋮	⋮
b, e	**3**
⋮	⋮
e, f	1

- **Diameter**: Maximum among the shortest path lengths.
- Diameter of the above graph $=$ 3.

Some Notes About Diameter

- Diameter is meaningful only for **connected** graphs. (Some references use ∞ as the diameter of a disconnected graph.)
- If a graph is disconnected, one needs to consider the diameter each connected component.
- For a connected graph with n nodes, the diameter is at most $n - 1$.
- In communication networks, diameter gives an indication of the worst-case delay for message delivery.
- Typically, giant components of social networks have small diameters (**small world phenomenon**).

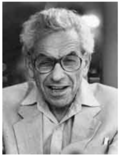
BFS and Diameter



Observation: For any **connected** graph, if a BFS produces r levels, then the diameter of the graph is at most $2r$.

Small World Phenomenon

Example: Erdős Collaboration Network



- Paul Erdős (1913 – 1996)
- Hungarian Mathematician

- Each node is a researcher and edge $\{x, y\}$ means that researchers x and y co-authored at least one paper.
- **Level 0**: Node corresponding to Erdős.
- **Level 1**: Nodes corresponding to researchers who have published a paper with Erdős.

Erdős Collaboration Network (continued)

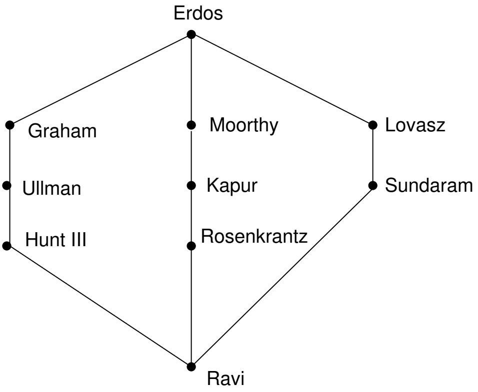
- **Level 2**: Nodes corresponding to researchers who have published a paper with some researcher in Level 1.

$$\vdots \qquad \vdots$$

- **Level j**: Nodes corresponding to researchers who have published a paper with some researcher in Level $j-1$.
- **Erdős Number** of a researcher: The level number in the graph for the node corresponding to the researcher.

Largest known Erdős Number = 8.

Example for Erdős Number



- Ravi's Erdős Number ≤ 3.
- Erdős Numbers of Teri Harrison, Catherine Dumas and Dan Lamanna ≤ 4.