

# Graphs & Traversals

Dr.V.S.Felix Enigo

# Edges (relationships)

- A social network is a collection of sentences that describe relationships

Alice ----likes-----> Bob  
(noun) (verb) (noun)

Single consistent verb is not sufficient to describe relationship

- Alice likes Bob very much.
- Bob and Carol study together.
- Carol fights with Alice after school.

# Cond...

- Semantics of a relationship is set by software developers
- Twitter “follow” is the same as every “friend” on Facebook
- Edges can have a numeric value
- Sociologists use Likert scale, for example:
- Limitation of Likert scale - heavily skewed towards values 3,4,5
- people tend to under- or misreport negative relationships\*
- there are not enough gradations (or relevant examples) to distinguish “like” from “strongly like”

0. Don't know

1. Strongly dislike

2. Dislike

3. Neither dislike nor like

4. Like

5. Strongly like

# Objective Questions

- David Krackhardt used objective questions than subjective
- Instead of asking “do you like person X”, it asks a more objective “How often do you communicate, with X?”
- 0. Never
- 1. At most once a year
- 2. At most once a month
- 3. At most once a week
- 4. At most once a day
- Frequency of communication maps on the subjective “friendship” or “liking” scale
- If you dislike someone, you will not talk to them more often

# Advantages of Objective Questions

- Minimizes self-reporting errors
- easy to remember if one talked, once a month or once a year
- Objectively measured, considering email timestamps or blog post replies

# Adjacency Matrix

	A	B	C	D	E
A	0	1	0	1	1
B	1	0	0	1	0
C	0	0	0	1	1
D	1	1	1	0	0
E	1	0	1	0	0

- Above matrix says if there is a relationship (edge) between nodes

	A	B	C	D	E
A	0	2	0	5	5
B	2	0	0	1	0
C	0	0	0	3	4
D	5	1	3	0	0
E	5	0	4	0	0

Valued graph in frequency scale

# Cond...

- 90% of cells would be zeros
- Density = non-zero cells / zero cells is too low.
- Most online social networks have density of 0.1% or less

## Solution

- Edge-Lists and Adjacency Lists

From To Value (frequency)

A B 2

A D 5

A E 5

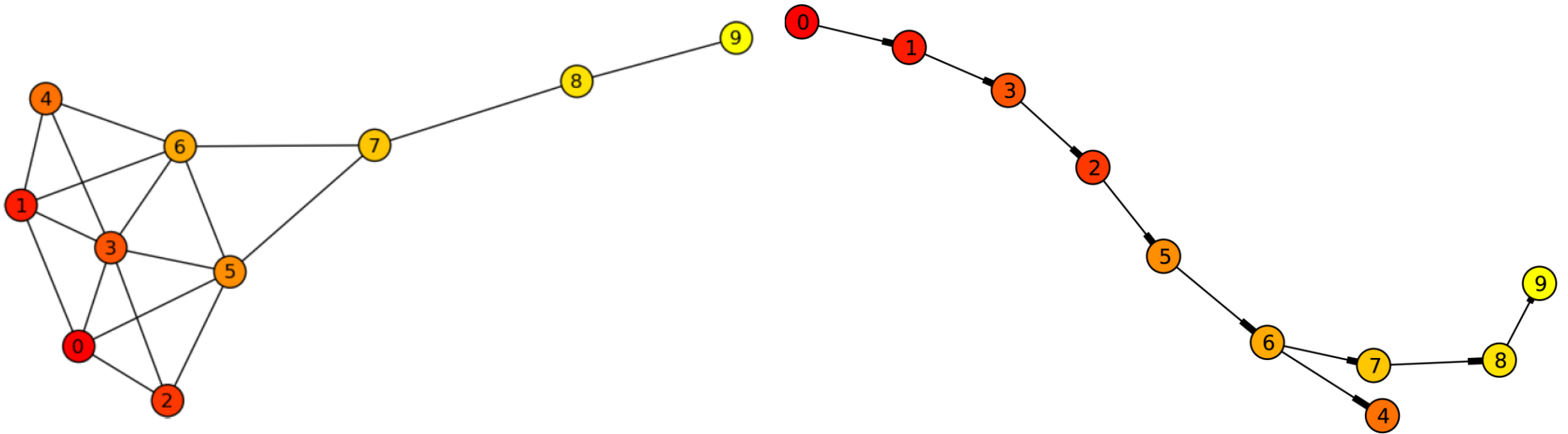
B A 2

# Graph Traversals and Distances

- Walking” or “crawling” the graph, literally means, from some starting point, follows links to its neighbors, and in turn the neighbors
- walk algorithms are designed to:
- find the shortest path from point A to point B
- walk the entire graph to understand or sample its structure
- Refer the program



# Depth-First Search



DFS involves descending down a child's child, iteratively, and then backtracking and turning to each of its siblings