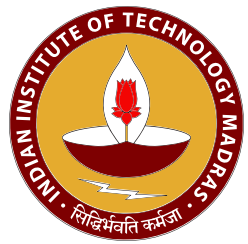
Lecture 3: Introduction to Graph Theory

Dynamics of Complex Systems: Week 3

Karthik Raman
Department of Biotechnology



Indian Institute of Technology Madras
https://home.iitm.ac.in/kraman/lab/

May 23, 2017

History of Graph Theory

The Seven Bridges of Königsberg

- Problem set in the picturesque Prussian city of Königsberg in 1735 (present day Kaliningrad, Russia), around the Pregel river
- Cities residents had a question "*Is it possible to set out from my house, cross each bridge exactly once, and return home?*"

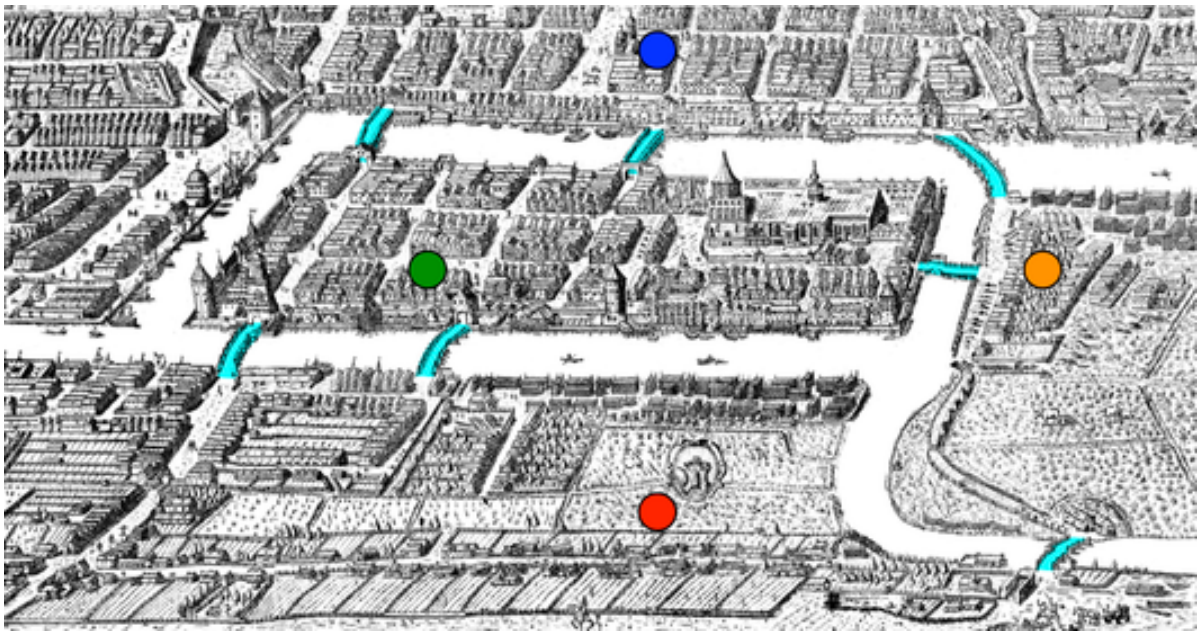
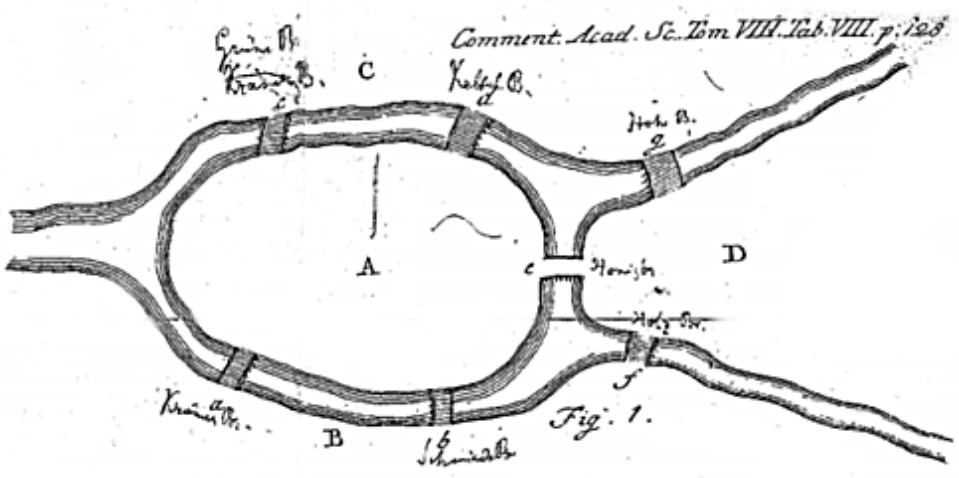


Figure Courtesy: http://rosalind.info/glossary/eulerian-cycle/

History of Graph Theory

The Seven Bridges of Königsberg

- No discussion of any math can be complete without discussing Euler!
- Euler's solution to the problem laid the foundations for graph theory!



Leonhard Euler

1707–1783

History of Graph Theory

The Seven Bridges of Königsberg

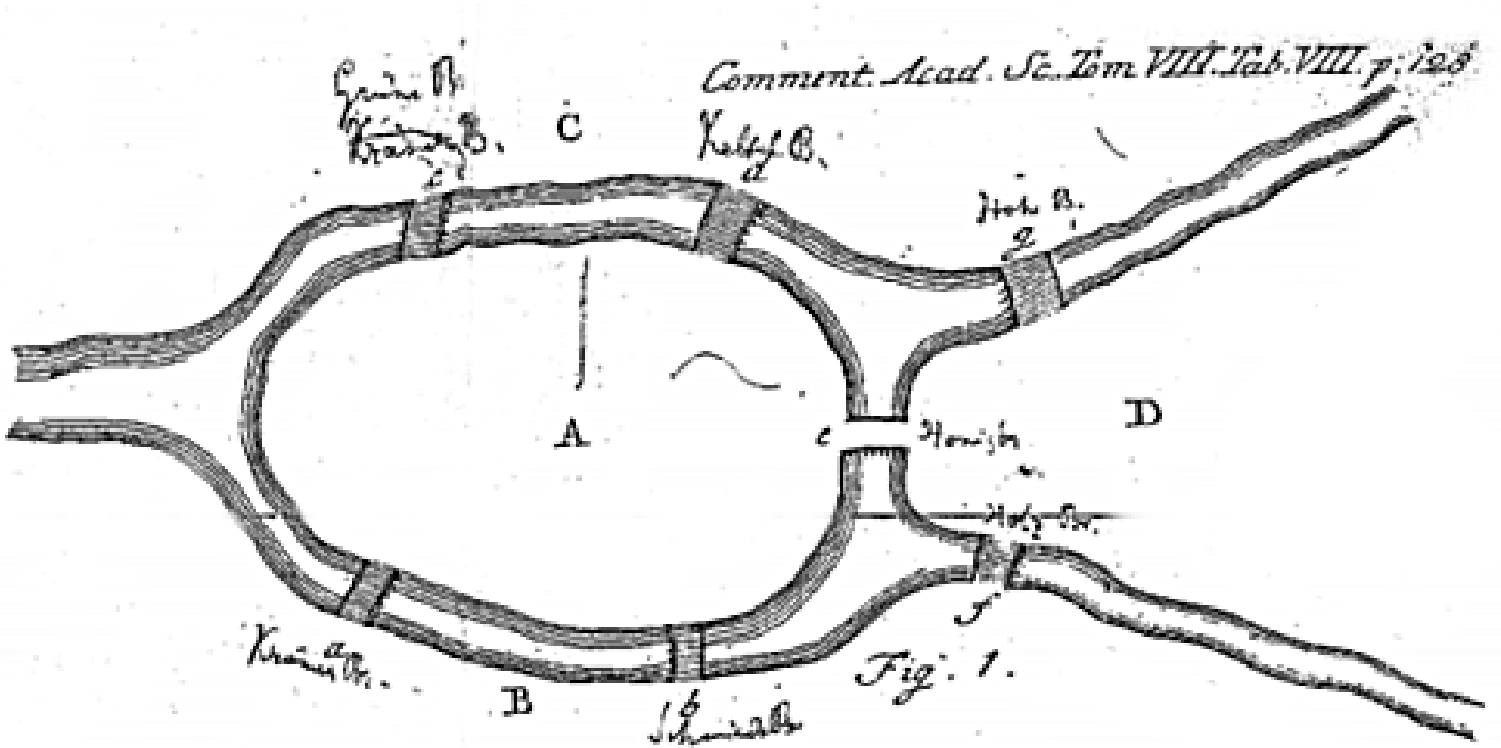
- What did Euler do?
- *Thus you see, most noble Sir, how this type of solution bears little relationship to mathematics, and I do not understand why you expect a mathematician to produce it, rather than anyone else, for the solution is based on reason alone, and its discovery does not depend on any mathematical principle. Because of this, I do not know why even questions which bear so little relationship to mathematics are solved more quickly by mathematicians than by others.*[a]
- *This question is so banal, but seemed to me worthy of attention in that [neither] geometry, nor algebra, nor even the art of counting was sufficient to solve it.*

[a] http://www.maa.org/press/periodicals/convergence/leonard-eulers-solution-to-the-konigsberg-bridge-problem

History of Graph Theory

The Seven Bridges of Königsberg

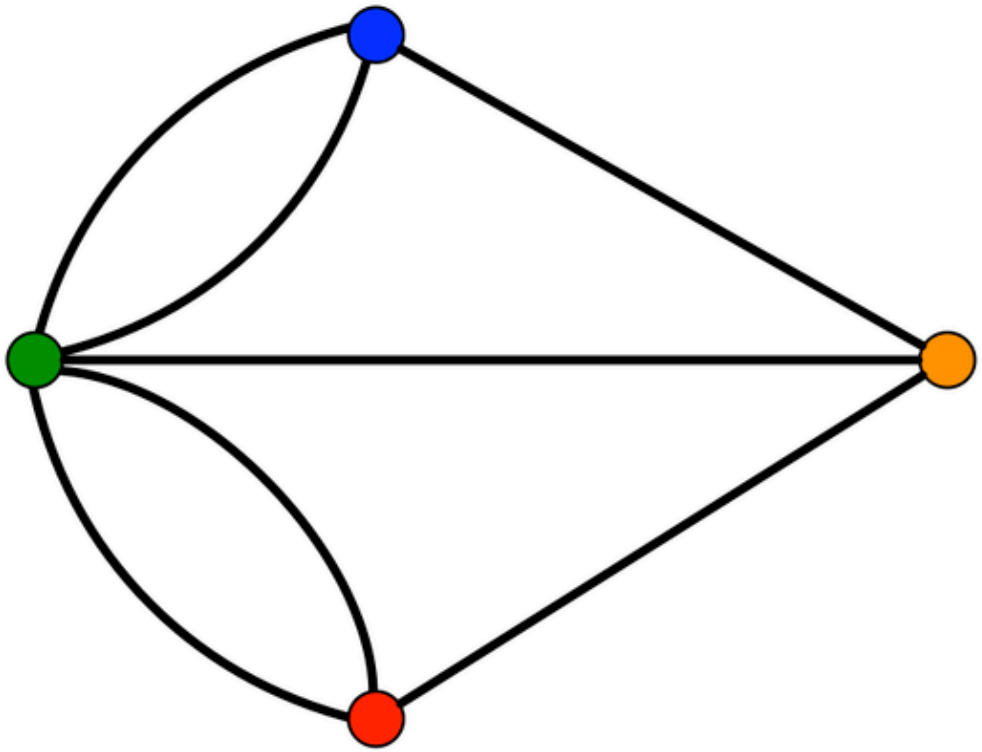
- What did Euler do?



History of Graph Theory

The Seven Bridges of Königsberg

- What did Euler do?



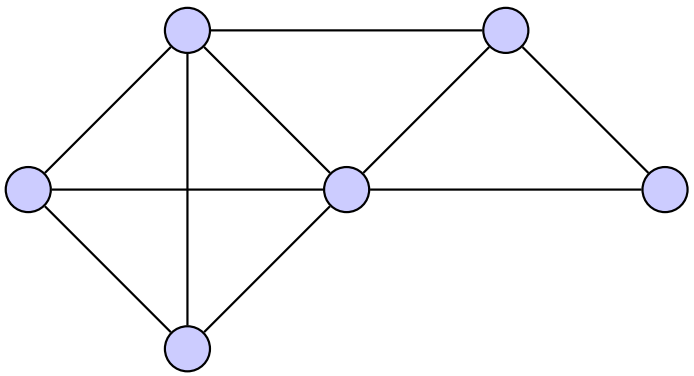
- Can you find the *walk* that the citizens were looking for?
- What did Euler prove? He proved that there is no *Eulerian circuit* in this graph!

What are Graphs?

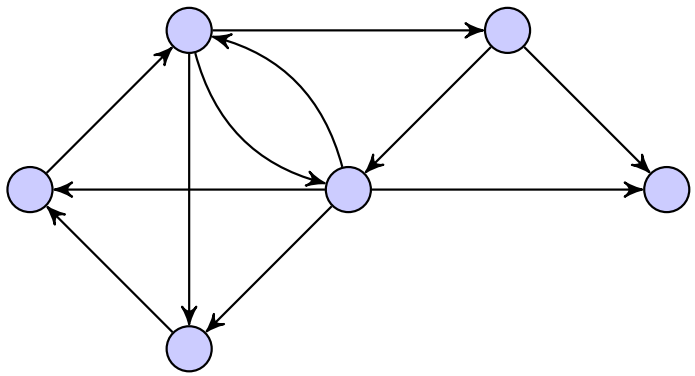
- One of the most important themes of computer science!
- A graph $G(V, E)$ is defined by a set of *vertices* V and a set of *edges* E, consisting of pairs of vertices from V
- Graphs are often referred to as networks, for example
  - Road networks
  - Social networks
  - Metabolic networks
  - Gene regulatory networks
  - Scientific citation networks
  - ...
- Graphs are classified elaborately — also influences the choice of algorithms

Directed vs. Undirected Graphs

- $G(V, E)$ is undirected if edge $(A, B) \in E$ implies that $(B, A) \in E$



Undirected graph



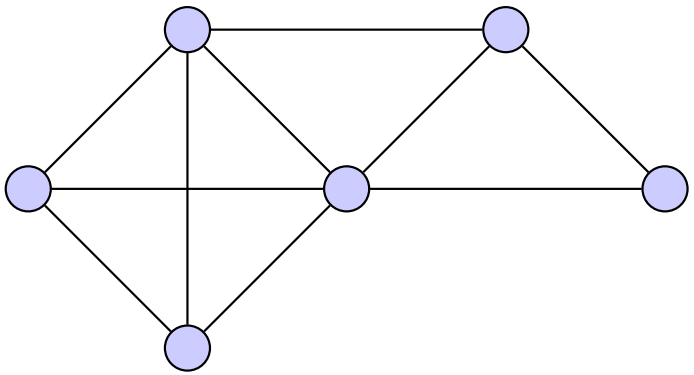
Directed graph

Examples

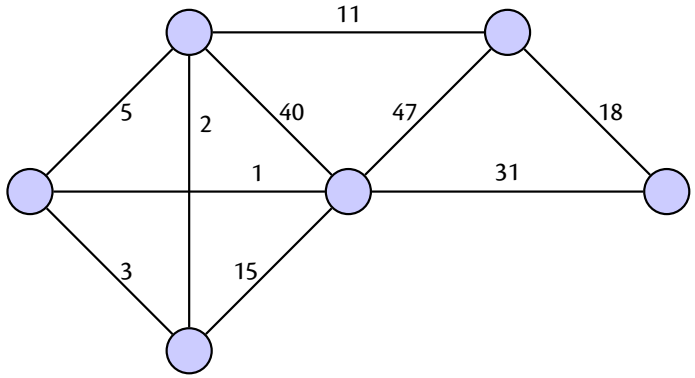
- Road networks between cities are typically undirected, while street networks within cities are often directed (why?)
- Facebook is undirected, while Twitter is directed
- Protein-interaction networks are undirected, while gene regulatory networks are directed

Weighted vs. Unweighted Graphs

- In a weighted graph, each edge is assigned a numerical value, or *weight*, often denoting a cost



Unweighted graph



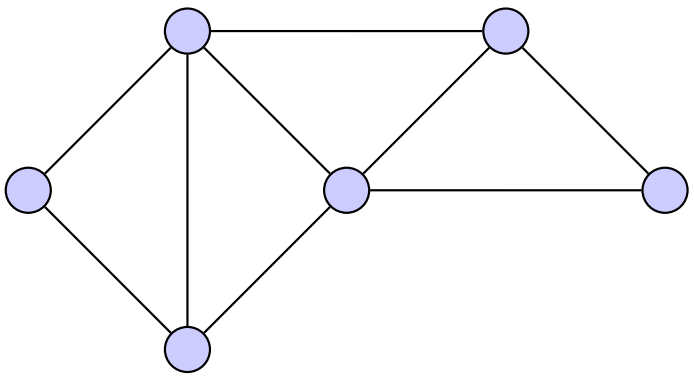
Weighted graph

Examples of weights

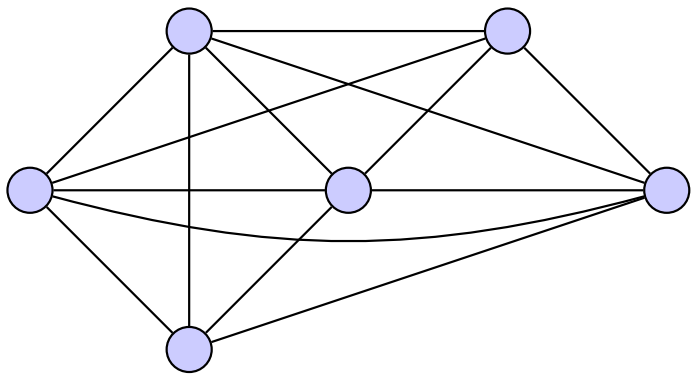
- Distance between cities
- Strength of an interaction

Sparse vs. Dense Graphs

- Graphs are sparse, when only a small fraction of the possible vertex pairs have edges defined between them



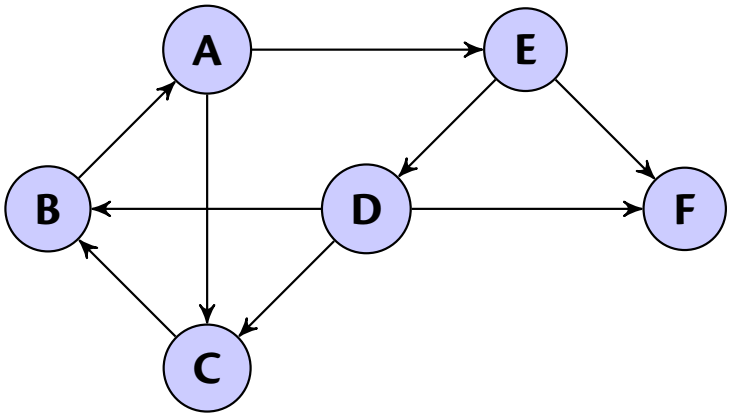
Sparse graph



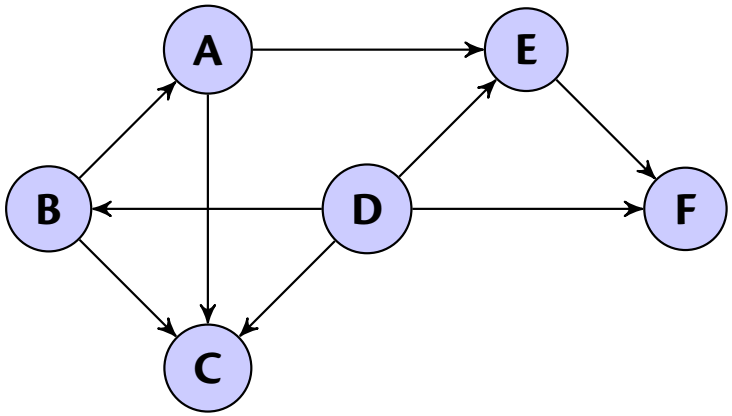
Dense graph

- Typically dense graphs have a quadratic number of edges, while sparse graphs are linear in size
- Many real graphs are usually sparse

Cyclic vs. Acyclic Graphs



Cyclic graph

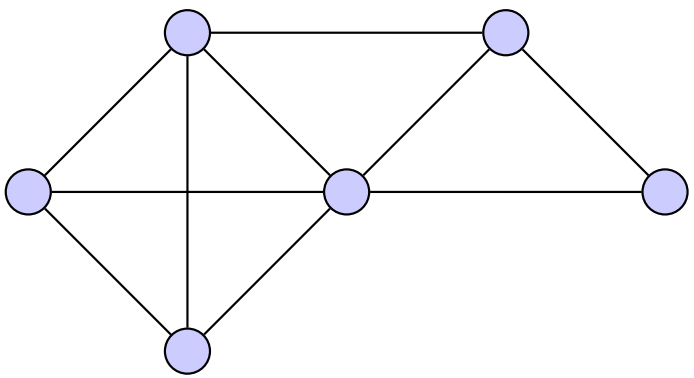


Acyclic graph

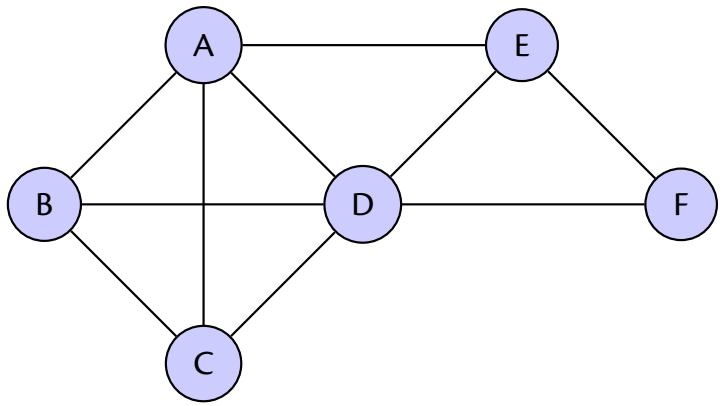
- An acyclic graph does not contain any *cycles*
- Trees are connected acyclic undirected graphs
- Directed acyclic graphs (DAGs) arise naturally in many scenarios

Labelled vs. Unlabelled Graphs

- In a labelled graph, each vertex has a unique name/label/identifier, distinguishing it from other vertices



Unlabelled graph



Labelled graph

- Important in graph alignment
- Graph isomorphism

Other Graph Types

- Implicit graphs
- Bi-partite graphs
- Hypergraphs

Other graph terminology:

- Converse/Transpose/Reverse
- Complete graph/Clique
- Walk (from A to B)

Many interesting questions can be asked of graphs

Social Networks

- Do I know someone who knows someone … who knows *X*?
  - *existence of a path*
- How long is that chain to *X*?
  - *shortest path problem*
- Is everyone in the world connected to one another?
  - *identification of connected components*
- Who has the most friends?
  - *most connected nodes/centrality analyses*

Many interesting questions can be asked of graphs

Biological Networks

- Is there a way to produce metabolite *X* from *A*?
  - *existence of a path*
- How long is that chain to *X* from *A*?
  - *shortest path problem*
- Are all proteins connected to others by a path?
  - *identification of connected components*
- Which is the most influential protein in a network?
  - *most connected nodes/centrality analyses*

Graph Algorithms

Many many problems in science and engineering can be cast back on to a graph!

- Shortest path problem
- Travelling salesperson problem
- Finding [strongly] connected components
- Graph isomorphism
- Vertex cover problem
- Minimum spanning tree problem
- Hamiltonian path problem
- Eulerian path problem
- *k*-shortest path problem
- Centrality measures

Graph Algorithms in Biology

Many biological problems map back on to graph problems

- Path finding in metabolic networks
- Identifying important proteins in networks
- Clusters of proteins in interaction networks
- Assembling reads of a genome from a next-generation sequencer
- Chemoinformatics problems

Network Jargon

- Node/Edge/Edge Weight
- Density
- Degree
- Shortest path/geodesic
- Diameter
- Characteristic path length
- Degree distribution
- Clustering coefficient
- Closeness centrality
- Betweenness centrality
- Edge betweenness
- Connected component
- Strongly connected component in directed graphs
- Acyclic graphs
- Motifs