Lecture 2: Paths, Cycles, Connectivity

Dr. Amitabh Trehan

amitabh.trehan@durham.ac.uk

**Based on the slides of ADS-21/22 by Dr. George Mertzios*

Reminder from last lecture

Reminder from last lecture

- A graph G is a pair $(V(G), E(G))$, where
  - $V(G)$ is a nonempty set of vertices (or nodes),
  - $E(G)$ is a set of unordered pairs uv with $u, v \in V(G)$ and $u \neq v$, called the edges of G.

Reminder from last lecture

- A graph G is a pair $(V(G), E(G))$, where
  - $V(G)$ is a nonempty set of vertices (or nodes),
  - $E(G)$ is a set of unordered pairs uv with $u, v \in V(G)$ and $u \neq v$, called the edges of G.
- When G is clear from the context, we write simply V and E.

Reminder from last lecture

- A graph G is a pair $(V(G), E(G))$, where
  - $V(G)$ is a nonempty set of vertices (or nodes),
  - $E(G)$ is a set of unordered pairs uv with $u, v \in V(G)$ and $u \neq v$, called the edges of G.
- When G is clear from the context, we write simply V and E.
- Graphs can be directed, multi-, pseudo-, weighted. By default they are not.

Reminder from last lecture

- A graph G is a pair $(V(G), E(G))$, where
  - $V(G)$ is a nonempty set of vertices (or nodes),
  - $E(G)$ is a set of unordered pairs uv with $u, v \in V(G)$ and $u \neq v$, called the edges of G.
- When G is clear from the context, we write simply V and E.
- Graphs can be directed, multi-, pseudo-, weighted. By default they are not.
- Many problems can be modelled using graphs.

Reminder from last lecture

- A graph G is a pair $(V(G), E(G))$, where
  - $V(G)$ is a nonempty set of vertices (or nodes),
  - $E(G)$ is a set of unordered pairs uv with $u, v \in V(G)$ and $u \neq v$, called the edges of G.
- When G is clear from the context, we write simply V and E.
- Graphs can be directed, multi-, pseudo-, weighted. By default they are not.
- Many problems can be modelled using graphs.
- We can draw graphs:
  - represent each vertex u by a point, and
  - each edge uv by a line or curve connecting u and v.

Reminder from last lecture

- A graph G is a pair $(V(G), E(G))$, where
  - $V(G)$ is a nonempty set of vertices (or nodes),
  - $E(G)$ is a set of unordered pairs uv with $u, v \in V(G)$ and $u \neq v$, called the edges of G.
- When G is clear from the context, we write simply V and E.
- Graphs can be directed, multi-, pseudo-, weighted. By default they are not.
- Many problems can be modelled using graphs.
- We can draw graphs:
  - represent each vertex u by a point, and
  - each edge uv by a line or curve connecting u and v.
- Paths, cycles, bipartite graphs, complete graphs, hypercubes

Contents for today's lecture

- Paths and directed paths;
- The shortest path problem;
- Connectivity and connected components;
- Eulerian and Hamiltonian cycles;
- Examples and exercises.

Walks, paths, cycles, and distances

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.
- A walk $v_0, v_1, \ldots, v_n$ with $v_0 = v_n$ is called a circuit or closed walk.

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.
- A walk $v_0, v_1, \ldots, v_n$ with $v_0 = v_n$ is called a circuit or closed walk.
- A closed walk is a cycle (or simple circuit) if all v_i's in it are distinct except $v_0 = v_n$.

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.
- A walk $v_0, v_1, \ldots, v_n$ with $v_0 = v_n$ is called a circuit or closed walk.
- A closed walk is a cycle (or simple circuit) if all v_i's in it are distinct except $v_0 = v_n$.
- If G is a directed graph then the directed paths and directed cycles are defined in a natural way, with each edge being directed from v_i to v_{i+1}.

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.
- A walk $v_0, v_1, \ldots, v_n$ with $v_0 = v_n$ is called a circuit or closed walk.
- A closed walk is a cycle (or simple circuit) if all v_i's in it are distinct except $v_0 = v_n$.
- If G is a directed graph then the directed paths and directed cycles are defined in a natural way, with each edge being directed from v_i to v_{i+1}.

- The length of a path or a cycle is the number of edges in it.

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.
- A walk $v_0, v_1, \ldots, v_n$ with $v_0 = v_n$ is called a circuit or closed walk.
- A closed walk is a cycle (or simple circuit) if all v_i's in it are distinct except $v_0 = v_n$.
- If G is a directed graph then the directed paths and directed cycles are defined in a natural way, with each edge being directed from v_i to v_{i+1}.

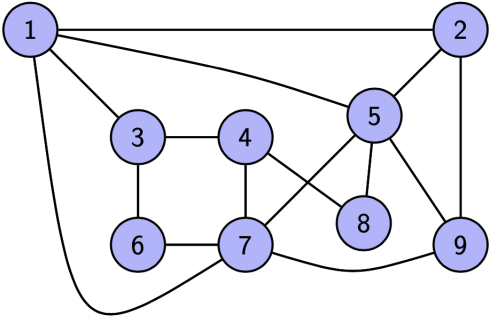
- The length of a path or a cycle is the number of edges in it.
- The distance between vertices u and v in a graph, denoted $dist(u, v)$, is the length of a shortest path from u to v if such a path exists, and ∞ otherwise.

Walks, paths, cycles, and distances

- A walk in a graph G is a sequence of edges $v_0v_1, v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$. In this case we also say that $v_0, v_1, \ldots, v_n$ is a walk in G.
- A walk $v_0, v_1, \ldots, v_n$ in G is a path if all v_i's are distinct. In this case we also say that $v_0, v_1, \ldots, v_n$ is a path in G.
- A walk $v_0, v_1, \ldots, v_n$ with $v_0 = v_n$ is called a circuit or closed walk.
- A closed walk is a cycle (or simple circuit) if all v_i's in it are distinct except $v_0 = v_n$.
- If G is a directed graph then the directed paths and directed cycles are defined in a natural way, with each edge being directed from v_i to v_{i+1}.

- The length of a path or a cycle is the number of edges in it.
- The distance between vertices u and v in a graph, denoted $dist(u, v)$, is the length of a shortest path from u to v if such a path exists, and ∞ otherwise.
- The diameter of a graph is the largest distance between two vertices in it

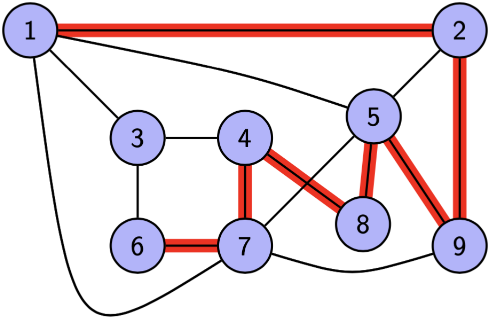
Exercise



- Does this graph contain a path of length 7?
- Does it contain a cycle of length 7?
- What is the distance from 2 to 6?
- What is the diameter of this graph?
- Answer at *PollEV.com/amitabhtrehan005*

Exercise: Solutions

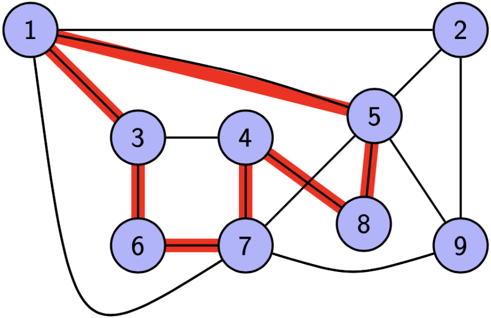
- Does this graph contain a path of length 7?



- Does it contain a cycle of length 7?
- What is the distance from 2 to 6?
- What is the diameter of this graph?

Exercise: Solutions

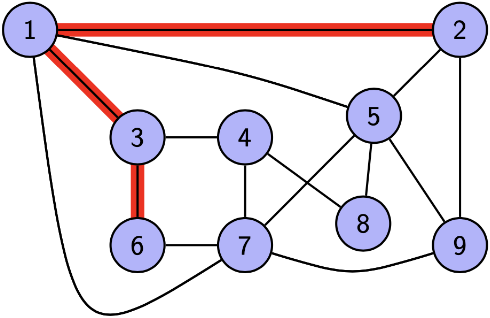
- Does this graph contain a path of length 7?
- Does it contain a cycle of length 7?



- What is the distance from 2 to 6?
- What is the diameter of this graph?

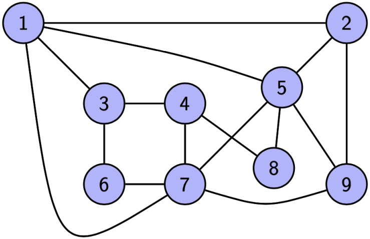
Exercise: Solutions

- Does this graph contain a path of length 7?
- Does it contain a cycle of length 7?
- What is the distance from 2 to 6?



- What is the diameter of this graph?

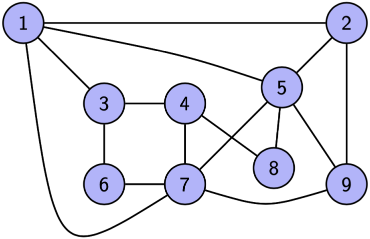
Exercise: Solutions



- Does this graph contain a path of length 7?
- Does it contain a cycle of length 7?
- What is the distance from 2 to 6?
- What is the diameter of this graph?

  The diameter is 3. Any way to prove this?

Exercise: Solutions



- Does this graph contain a path of length 7?
- Does it contain a cycle of length 7?
- What is the distance from 2 to 6?
- What is the diameter of this graph?

  The diameter is 3. Any way to prove this?

  Proof: there is a DOMINATING SET {1,5,7} which is a TRIANGLE

The acquaintance graph and six degrees of separation

The acquaintance graph:

- The vertices are all people
- There is an edge between two of them if they are acquainted

The acquaintance graph and six degrees of separation

The acquaintance graph:

- The vertices are all people
- There is an edge between two of them if they are acquainted

- Many sociologists conjectured/observed that this graph has diameter 6, after removing a tiny fraction of vertices.

The acquaintance graph and six degrees of separation

The acquaintance graph:

- The vertices are all people
- There is an edge between two of them if they are acquainted

- Many sociologists conjectured/observed that this graph has diameter 6, after removing a tiny fraction of vertices. (Can you put this in plain words?)

The acquaintance graph and six degrees of separation

The acquaintance graph:

- The vertices are all people
- There is an edge between two of them if they are acquainted

- Many sociologists conjectured/observed that this graph has diameter 6, after removing a tiny fraction of vertices. (Can you put this in plain words?)
- This is known as the "small world phenomenon".

The acquaintance graph and six degrees of separation

The acquaintance graph:

- The vertices are all people
- There is an edge between two of them if they are acquainted

- Many sociologists conjectured/observed that this graph has diameter 6, after removing a tiny fraction of vertices. (Can you put this in plain words?)
- This is known as the "small world phenomenon".
- There is a popular play (and a film) based on this, called "Six degrees of separation".

The collaboration graph and the Erdös number

The collaboration graph:

- The vertices are all people (or only scientists)
- There is an edge between two of them if they have written a joint paper

The collaboration graph and the Erdös number

The collaboration graph:

- The vertices are all people (or only scientists)
- There is an edge between two of them if they have written a joint paper

- Paul Erdös (1913-1996) was a great and extremely prolific mathematician.

The collaboration graph and the Erdös number

The collaboration graph:

- The vertices are all people (or only scientists)
- There is an edge between two of them if they have written a joint paper

- Paul Erdös (1913-1996) was a great and extremely prolific mathematician.
- The Erdös number of a person is the distance to Paul Erdös in the collaboration graph.

The collaboration graph and the Erdös number

The collaboration graph:

- The vertices are all people (or only scientists)
- There is an edge between two of them if they have written a joint paper

- Paul Erdös (1913-1996) was a great and extremely prolific mathematician.
- The Erdös number of a person is the distance to Paul Erdös in the collaboration graph.
- Btw, my Erdös number is 3. (Can you put this in plain words?)

The Hollywood graph and the Bacon number

The Hollywood graph:

- The vertices are all people (or only actors).
- There is an edge between two of them if they appeared in the same film.

The Hollywood graph and the Bacon number

The Hollywood graph:

- The vertices are all people (or only actors).
- There is an edge between two of them if they appeared in the same film.

- Kevin Bacon is a famous Hollywood actor.

The Hollywood graph and the Bacon number

The Hollywood graph:

- The vertices are all people (or only actors).
- There is an edge between two of them if they appeared in the same film.

- Kevin Bacon is a famous Hollywood actor.
- The Bacon number of a person is the distance to Kevin Bacon in the Hollywood graph.

The Hollywood graph and the Bacon number

The Hollywood graph:

- The vertices are all people (or only actors).
- There is an edge between two of them if they appeared in the same film.

- Kevin Bacon is a famous Hollywood actor.
- The Bacon number of a person is the distance to Kevin Bacon in the Hollywood graph.
- About 90% of actors have a Bacon number (i.e. the distance is not ∞)

The Erdös-Bacon number

Definition

For any person, their Erdös-Bacon number is the sum of their Erdös number and their Bacon number.

The Erdös-Bacon number

Definition

For any person, their Erdös-Bacon number is the sum of their Erdös number and their Bacon number.

There are not many people with a small Erdös-Bacon number.

The Erdös-Bacon number

Definition

For any person, their Erdös-Bacon number is the sum of their Erdös number and their Bacon number.

There are not many people with a small Erdös-Bacon number.

For example, the following people have Erdös-Bacon number 7 or less.

The Erdös-Bacon number

Definition

For any person, their Erdös-Bacon number is the sum of their Erdös number and their Bacon number.

There are not many people with a small Erdös-Bacon number.

For example, the following people have Erdös-Bacon number 7 or less.

The Erdös-Bacon number

Definition

For any person, their Erdös-Bacon number is the sum of their Erdös number and their Bacon number.

There are not many people with a small Erdös-Bacon number.

For example, the following people have Erdös-Bacon number 7 or less.

The Erdös-Bacon number

Definition

For any person, their Erdös-Bacon number is the sum of their Erdös number and their Bacon number.

There are not many people with a small Erdös-Bacon number.

For example, the following people have Erdös-Bacon number 7 or less.

- Btw, my Erdös-Bacon number is ∞, but I have a colleague of a colleague who has a co-author (Hubie Chen) with Erdös-Bacon number 5 (3+2)

Shortest-path problems

In a graph (possibly with edge weights, the problem of computing a path from a given vertex u ("source") to a given vertex v ("target") with the smallest total length (or weight) is known as the shortest-path problem.

Shortest-path problems

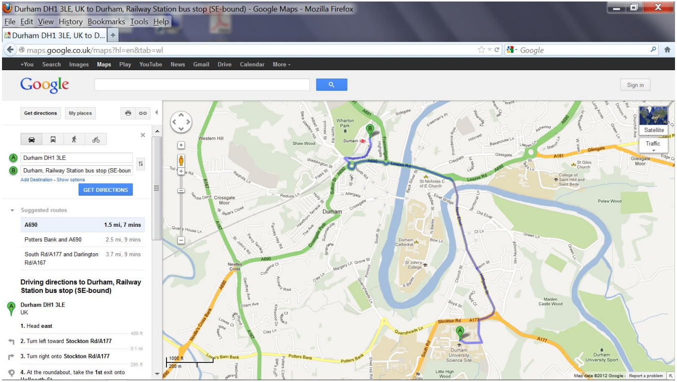
In a graph (possibly with edge weights, the problem of computing a path from a given vertex u ("source") to a given vertex v ("target") with the smallest total length (or weight) is known as the shortest-path problem.

We all often (use software applications that) compute such paths. Example?

Shortest-path problems

In a graph (possibly with edge weights, the problem of computing a path from a given vertex u ("source") to a given vertex v ("target") with the smallest total length (or weight) is known as the shortest-path problem.

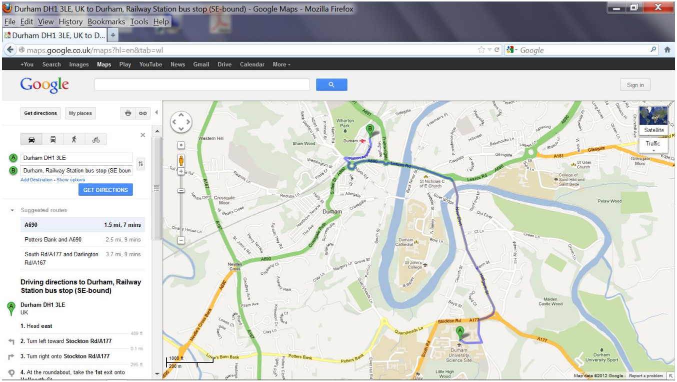
We all often (use software applications that) compute such paths. Example?



Shortest-path problems

In a graph (possibly with edge weights, the problem of computing a path from a given vertex u ("source") to a given vertex v ("target") with the smallest total length (or weight) is known as the shortest-path problem.

We all often (use software applications that) compute such paths. Example?



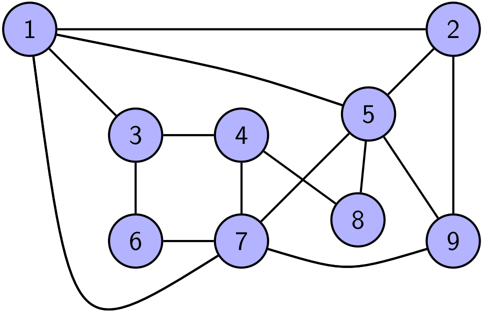
We will learn about algorithms for the (unweighted) problem in a few lectures.

Connectivity

Definition

A graph $G = (V, E)$ is called connected if, between every pair of vertices u, v, there exists at least one path in G.

A connected component of G is a maximal connected subgraph of G.



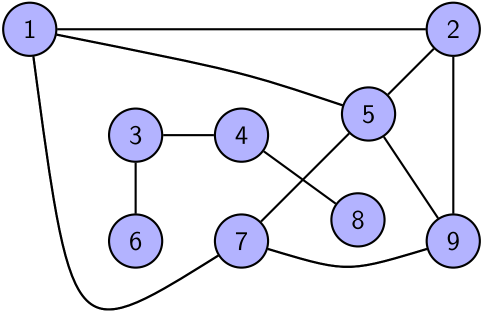
- Is this graph connected?

Connectivity

Definition

A graph $G = (V, E)$ is called connected if, between every pair of vertices u, v, there exists at least one path in G.

A connected component of G is a maximal connected subgraph of G.



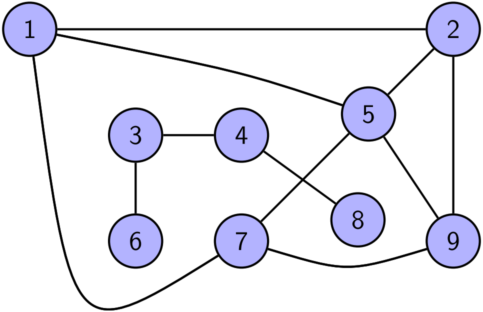
- What about this graph?

Connectivity

Definition

A graph $G = (V, E)$ is called connected if, between every pair of vertices u, v, there exists at least one path in G.

A connected component of G is a maximal connected subgraph of G.



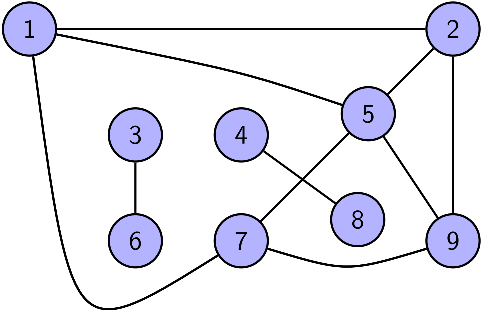
- What about this graph?

Connectivity

Definition

A graph $G = (V, E)$ is called connected if, between every pair of vertices u, v, there exists at least one path in G.

A connected component of G is a maximal connected subgraph of G.



- How many connected components does this graph have?

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Proof.

By induction over $m = |E|$.

$\square$

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Proof.

By induction over $m = |E|$.

Induction basis: $m = 0$. No edges, so only isolated vertices, i.e. $|V|$ connected components.

$\square$

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Proof.

By induction over $m = |E|$.

Induction basis: $m = 0$. No edges, so only isolated vertices, i.e. $|V|$ connected components.

Induction step: Let $G = (V, E)$ with $|E| = m + 1$. Consider an arbitrary $e \in E$, and define $E' = E \setminus \{e\}$. By induction hypothesis: $G' = (V, E')$ has at least $|V| - |E'| = |V| - m$ connected components. Two cases:

1. With e, the number of connected components does not change.
2. With e, the number of connected components ???

$\square$

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Proof.

By induction over $m = |E|$.

Induction basis: $m = 0$. No edges, so only isolated vertices, i.e. $|V|$ connected components.

Induction step: Let $G = (V, E)$ with $|E| = m + 1$. Consider an arbitrary $e \in E$, and define $E' = E \setminus \{e\}$. By induction hypothesis: $G' = (V, E')$ has at least $|V| - |E'| = |V| - m$ connected components. Two cases:

1. With e, the number of connected components does not change.
2. With e, the number of connected components decreases by 1.

$\square$

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Proof.

By induction over $m = |E|$.

Induction basis: $m = 0$. No edges, so only isolated vertices, i.e. $|V|$ connected components.

Induction step: Let $G = (V, E)$ with $|E| = m + 1$. Consider an arbitrary $e \in E$, and define $E' = E \setminus \{e\}$. By induction hypothesis: $G' = (V, E')$ has at least $|V| - |E'| = |V| - m$ connected components. Two cases:

1. With e, the number of connected components does not change.
2. With e, the number of connected components decreases by 1.

In both cases, G has at least $|V| - m - 1 = |V| - |E|$ connected components. □

Connected components

Theorem

Every graph $G = (V, E)$ contains at least $|V| - |E|$ connected components.

Proof.

By induction over $m = |E|$.

Induction basis: $m = 0$. No edges, so only isolated vertices, i.e. $|V|$ connected components.

Induction step: Let $G = (V, E)$ with $|E| = m + 1$. Consider an arbitrary $e \in E$, and define $E' = E \setminus \{e\}$. By induction hypothesis: $G' = (V, E')$ has at least $|V| - |E'| = |V| - m$ connected components. Two cases:

1. With e, the number of connected components does not change.
2. With e, the number of connected components decreases by 1.

In both cases, G has at least $|V| - m - 1 = |V| - |E|$ connected components. □

Corollary (useful in various algorithmic proofs)

If $G = (V, E)$ is connected then $|E| \geq |V| - 1$.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.
- Take two vertices u and v in different connected components.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.
- Take two vertices u and v in different connected components.
- We have that $deg(u) \geq (n-1)/2$ and $deg(v) \geq (n-1)/2$.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.
- Take two vertices u and v in different connected components.
- We have that $deg(u) \geq (n-1)/2$ and $deg(v) \geq (n-1)/2$.
- The set $N(u)$ contains none of the neighbours of v nor v itself.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.
- Take two vertices u and v in different connected components.
- We have that $deg(u) \geq (n-1)/2$ and $deg(v) \geq (n-1)/2$.
- The set $N(u)$ contains none of the neighbours of v nor v itself.
- Similarly, $N(v)$ contains none of the neighbours of u nor u itself.

Exercise

Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.
- Take two vertices u and v in different connected components.
- We have that $deg(u) \geq (n-1)/2$ and $deg(v) \geq (n-1)/2$.
- The set $N(u)$ contains none of the neighbours of v nor v itself.
- Similarly, $N(v)$ contains none of the neighbours of u nor u itself.
- Hence G contains at least $\frac{n-1}{2} + 1 + \frac{n-1}{2} + 1 = n + 1$ vertices, a contradiction.

Exercise

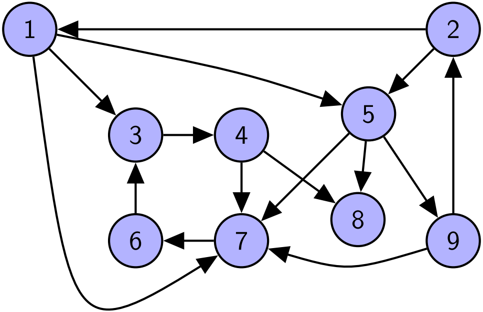
Exercise 1: Prove that if G is a graph on n vertices and $\delta(G) \geq (n-1)/2$ then G is connected.

- Proof by contradiction: Assume G is not connected, derive a contradiction.
- Take two vertices u and v in different connected components.
- We have that $deg(u) \geq (n-1)/2$ and $deg(v) \geq (n-1)/2$.
- The set $N(u)$ contains none of the neighbours of v nor v itself.
- Similarly, $N(v)$ contains none of the neighbours of u nor u itself.
- Hence G contains at least $\frac{n-1}{2} + 1 + \frac{n-1}{2} + 1 = n+1$ vertices, a contradiction.
- In fact, we even proved more: any two vertices in G are at distance at most 2 (so the diameter of G is at most 2).

Strong connectivity

Definition

A directed graph G is called (weakly) connected if the graph obtained from G by forgetting directions is connected.
A directed graph is called strongly connected if any two distinct vertices are connected by directed paths in both directions.
A strongly connected component (or simply strong component) of a digraph G is a maximal strongly connected subgraph of G.

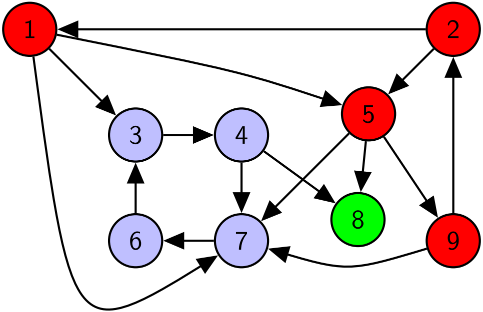


- Is this graph strongly connected?

Strong connectivity

Definition

A directed graph G is called (weakly) connected if the graph obtained from G by forgetting directions is connected.
A directed graph is called strongly connected if any two distinct vertices are connected by directed paths in both directions.
A strongly connected component (or simply strong component) of a digraph G is a maximal strongly connected subgraph of G.



- Is this graph strongly connected?

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Key: WCC = weakly connected component, SCC = strongly ...

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Key: WCC = weakly connected component, SCC = strongly ...

- Web graph: 50% SCC, 91% WCC

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Key: WCC = weakly connected component, SCC = strongly ...

- Web graph: 50% SCC, 91% WCC
- Amazon co-purchase graph: 91% SCC, 100% WCC

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Key: WCC = weakly connected component, SCC = strongly ...

- Web graph: 50% SCC, 91% WCC
- Amazon co-purchase graph: 91% SCC, 100% WCC
- Wikipedia talk graph: 48% SCC, 100% WCC

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Key: WCC = weakly connected component, SCC = strongly ...

- Web graph: 50% SCC, 91% WCC
- Amazon co-purchase graph: 91% SCC, 100% WCC
- Wikipedia talk graph: 48% SCC, 100% WCC
- Facebook graph: 100% WCC

Connectivity properties of commercially important graphs

Important large graphs tend to have similar connectivity properties:

See the Stanford Large Networks Dataset Collection:
`https://snap.stanford.edu/data/index.html`

Key: WCC = weakly connected component, SCC = strongly ...

- Web graph: 50% SCC, 91% WCC
- Amazon co-purchase graph: 91% SCC, 100% WCC
- Wikipedia talk graph: 48% SCC, 100% WCC
- Facebook graph: 100% WCC
- Google+ graph: 61% SCC, 100% WCC
- ...

Special circuits/cycles in graphs

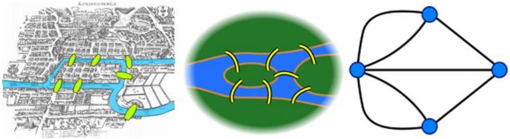
- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?

Special circuits/cycles in graphs

- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?
  - Such a circuit in G is called a Eulerian circuit, after Leonhard Euler (1707-83).

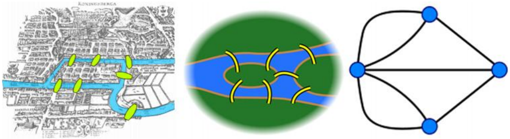
Special circuits/cycles in graphs

- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?
    - Such a circuit in G is called a Eulerian circuit, after Leonhard Euler (1707-83).



Special circuits/cycles in graphs

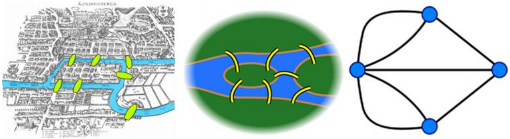
- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?
    - Such a circuit in G is called a Eulerian circuit, after Leonhard Euler (1707-83).



- Can we travel along the edges of a given graph so that we start and finish at the same vertex and visit each vertex exactly once?

Special circuits/cycles in graphs

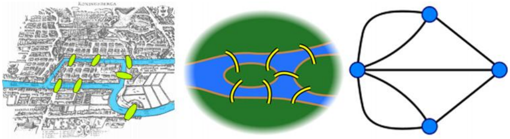
- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?
  - Such a circuit in G is called a Eulerian circuit, after Leonhard Euler (1707-83).



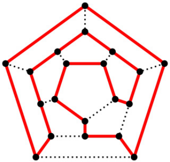
- Can we travel along the edges of a given graph so that we start and finish at the same vertex and visit each vertex exactly once?
  - Such a cycle is called a Hamiltonian cycle, after William Hamilton (1805-65).

Special circuits/cycles in graphs

- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?
  - Such a circuit in G is called a Eulerian circuit, after Leonhard Euler (1707-83).

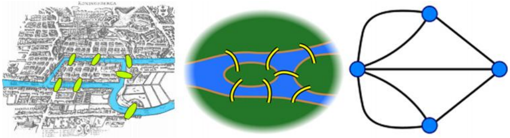


- Can we travel along the edges of a given graph so that we start and finish at the same vertex and visit each vertex exactly once?
  - Such a cycle is called a Hamiltonian cycle, after William Hamilton (1805-65).

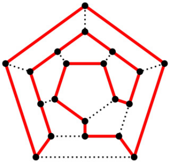


Special circuits/cycles in graphs

- Can we travel along the edges of a given graph G so that we start and finish at the same vertex and traverse each edge exactly once?
  - Such a circuit in G is called a Eulerian circuit, after Leonhard Euler (1707-83).



- Can we travel along the edges of a given graph so that we start and finish at the same vertex and visit each vertex exactly once?
  - Such a cycle is called a Hamiltonian cycle, after William Hamilton (1805-65).



- Detecting one of these two types of circuits is easy, while detecting the other is not easy at all. Which is which?

Eulerian circuits

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.
Induction base: $G = K_3$, the claim is obvious. Induction step:

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.
Induction base: $G = K_3$, the claim is obvious. Induction step:

- Start walking from any vertex u along the untraversed edges, marking off the traversed edges.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.
Induction base: $G = K_3$, the claim is obvious. Induction step:

- Start walking from any vertex u along the untraversed edges, marking off the traversed edges.
- Stop when you arrive at a vertex where you can't continue (all edges leading from it are already traversed). This vertex must be u again.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.
Induction base: $G = K_3$, the claim is obvious. Induction step:

- Start walking from any vertex u along the untraversed edges, marking off the traversed edges.
- Stop when you arrive at a vertex where you can't continue (all edges leading from it are already traversed). This vertex must be u again.
- Hence we have a circuit C. Delete it from G to obtain a smaller graph H in which all degrees are also even.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.
Induction base: $G = K_3$, the claim is obvious. Induction step:

- Start walking from any vertex u along the untraversed edges, marking off the traversed edges.
- Stop when you arrive at a vertex where you can't continue (all edges leading from it are already traversed). This vertex must be u again.
- Hence we have a circuit C. Delete it from G to obtain a smaller graph H in which all degrees are also even.
- By induction hypothesis, each connected component of H has an Eulerian circuit.

Eulerian circuits

Theorem (Rosen p.634-636)

A connected graph with at least two vertices has an Eulerian curcuit iff each of its vertices has even degree.

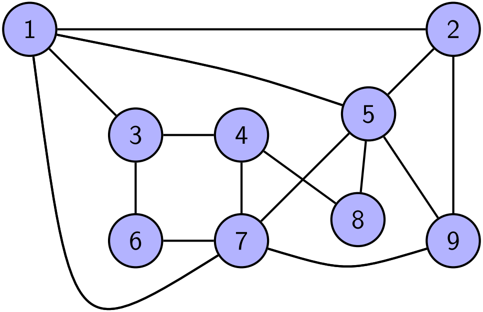
Idea of proof:

Necessity ($\Rightarrow$): each time this circuit passes through a vertex v, it contributes 2 to $deg(v)$. Since each edge is used exactly once, $deg(v)$ must be even.

Sufficiency ($\Leftarrow$): Induction on the number of vertices in G.
Induction base: $G = K_3$, the claim is obvious. Induction step:

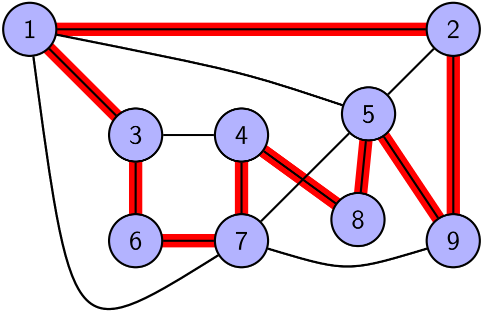
- Start walking from any vertex u along the untraversed edges, marking off the traversed edges.
- Stop when you arrive at a vertex where you can't continue (all edges leading from it are already traversed). This vertex must be u again.
- Hence we have a circuit C. Delete it from G to obtain a smaller graph H in which all degrees are also even.
- By induction hypothesis, each connected component of H has an Eulerian circuit.
- Combine C and these circuits to obtain the required circuit for G.

Hamiltonian cycles



- Does this graph have a Eulerian circuit?
- Does it have a Hamiltonian cycle?
- Answer at *PollEV.com/amitabhtrehan005*

Hamiltonian cycles



- Does this graph have a Eulerian circuit?
- Does it have a Hamiltonian cycle?
- Answer at *PollEV.com/amitabhtrehan005*

- Detecting Eulerian circuits algorithmically is easy. (How?)
- Detecting Hamiltonian cycles is hard (NP-complete).

Travelling Salesman Problem (TSP)

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point
- A (positive integer) cost $d(i, j)$ of travel between each pair (c_i, c_j) is known.

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point
- A (positive integer) cost $d(i, j)$ of travel between each pair (c_i, c_j) is known.
- Goal: find an optimal (i.e. cheapest) route for the salesman.

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point
- A (positive integer) cost $d(i, j)$ of travel between each pair (c_i, c_j) is known.
- Goal: find an optimal (i.e. cheapest) route for the salesman.

Given a graph G with set V of vertices ($|V| = n$) and set E of edges,

- for each vertex v, create a city c_v;
- for each pair of distinct $u, v \in V$, set $d(c_u, c_v) = 1$ if $uv \in E$ and $d(c_u, c_v) = 2$ otherwise.

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point
- A (positive integer) cost $d(i, j)$ of travel between each pair (c_i, c_j) is known.
- Goal: find an optimal (i.e. cheapest) route for the salesman.

Given a graph G with set V of vertices ($|V| = n$) and set E of edges,

- for each vertex v, create a city c_v;
- for each pair of distinct $u, v \in V$, set $d(c_u, c_v) = 1$ if $uv \in E$ and $d(c_u, c_v) = 2$ otherwise.

Then detecting a Hamiltonian cycle in G can be viewed as TSP:

- if G has a Hamiltonian cycle then the cycle is a route of cost exactly n.

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point
- A (positive integer) cost $d(i, j)$ of travel between each pair (c_i, c_j) is known.
- Goal: find an optimal (i.e. cheapest) route for the salesman.

Given a graph G with set V of vertices ($|V| = n$) and set E of edges,

- for each vertex v, create a city c_v;
- for each pair of distinct $u, v \in V$, set $d(c_u, c_v) = 1$ if $uv \in E$ and $d(c_u, c_v) = 2$ otherwise.

Then detecting a Hamiltonian cycle in G can be viewed as TSP:

- if G has a Hamiltonian cycle then the cycle is a route of cost exactly n.
- if there is a route of cost n then it can't use pairs with cost 2 (why?) and so goes through edges of G and hence is a Hamiltonian cycle.

Travelling Salesman Problem (TSP)

The (famous) TSP is the following problem:

- A salesman should visit cities $c_1, c_2, \ldots, c_n$ in some order, visiting each city exactly once and returning to the starting point
- A (positive integer) cost $d(i, j)$ of travel between each pair (c_i, c_j) is known.
- Goal: find an optimal (i.e. cheapest) route for the salesman.

Given a graph G with set V of vertices ($|V| = n$) and set E of edges,

- for each vertex v, create a city c_v;
- for each pair of distinct $u, v \in V$, set $d(c_u, c_v) = 1$ if $uv \in E$ and $d(c_u, c_v) = 2$ otherwise.

Then detecting a Hamiltonian cycle in G can be viewed as TSP:

- if G has a Hamiltonian cycle then the cycle is a route of cost exactly n.
- if there is a route of cost n then it can't use pairs with cost 2 (why?) and so goes through edges of G and hence is a Hamiltonian cycle.

You have just witnessed a Reduction! (from one problem to another).

Thank You!