CSCE 411: Design and Analysis of Algorithms

Week 7: Finishing DFS; Minimum Spanning Trees

Date: February 24, 2026 *Nate Veldt, updated by Samson Zhou*

Course Logistics

- Graph algorithms: Chapter 22
- Homework 4 out, due this Friday

1 Depth First Search: Motivating Problems

Depth first search is used in several applications for analyzing directed graphs. We will now take a closer look at these applications.

Directed graph reminders

1.1 Reachability and Connected Components

Reachability. Given a graph $G = (V, E)$ and node set $S \subseteq V$, node $v \in S$ is *reachable* from node $u \in S$ if ____.

Connected components. For an undirected graph $G = (V, E)$ a connected component is a maximal subgraph in which every node in is ________________________________

Weakly Connected components If $G = (V, E)$ is directed, a *weakly connected component* is ____

Strongly Connected components If $G = (V, E)$ is directed, a *strongly connected component* is subgraph $S \subseteq V$ in which there is ____________________________________

Question 1. *How many weakly connected components and strongly connected components are there in the following graph, respectively?*

- **A** *1 and 3*
- **B** *1 and 2*
- **C** *0 and 1*
- **D** *2 and 3*

1.2 Directed Acyclic Graphs

A *cycle* in a directed graph is a directed path ________________________________

A *Directed acyclic* graph is a directed graph that ____________________.

Examples

1.3 Topological Sorting

A topologically ordering of a directed acyclic graph $G = (V, E)$ is an ordering of nodes so that:

2 Application 1: Checking if G is a DAG

Theorem 2.1. *G is a DAG $\iff$ a DFS yields no back edges. Equivalently:*

Proof First, ($\implies$) we show that if DFS yields a back edge, G is not a DAG.

Next ($\Leftarrow$) we show that if G is not a DAG there will be a back edge.

3 Application 2: Topological Sort

Given a directed acyclic graph $G = (V, E)$, a topological sort of G is an ordering of nodes such that for any $(u, v) \in E$, u comes before v in the ordering.

We can use the following procedure to solve the topological sort problem:

1.

2.

Theorem 3.1. *Ordering nodes in a directed acyclic graph $G = (V, E)$ by reversed finish times will produce a topological sort of G.*

Proof.

1. Let (u, v) be an edge in G

2. Our goal is to show that

   ____

3. When (u, v) is explored, there are three different possibilities for the status of v:

   - **Case 1**: v.status $== U$. This means v becomes a descendant of u.

     Thus, $v.F < u.F$. Reason: _______________________

   - **Case 2**: v.status $== E$, then we also have $v.F < u.F$.

     Reason:

   - **Case 3**: v.status $== D$, this means that v is an ancestor of u, so (u, v) is a back edge.

     But this is impossible. Reason: _______________________

4. In all cases that are possible, ____________________

$\square$

4 The transpose graph and connected component graph

If $G = (V, E)$ is a graph, a *strongly connected component* is maximal subgraph $S \subseteq V$ in which every node is reachable from every other node by following paths in S.

Let $G = (V, E)$ be a graph and assume that $\{C_1, C_2, \ldots, C_k\}$ represent its strongly connected components.
The *connected component graph* $G^{\text{scc}} = (V^{\text{scc}}, E^{\text{scc}})$ is defined as follows:

- There is a node $v_i \in V^{\text{scc}}$ for each component C_i
- There is an edge $(v_i, v_j) \in E^{\text{scc}}$ if and only if there is a directed edge between C_i and C_j

Lemma 4.1. *The connected component graph is* ______________________________

The *transpose graph* of G is $G^T = (V, E^T)$ where

$$E^T = \{(u, v) \colon (v, u) \in E\}$$

Lemma 4.2. *G and G^T have* ______________________________

5 Strongly Connected Components

The following algorithm will compute the strongly connected components of a graph $G = (V, E)$:

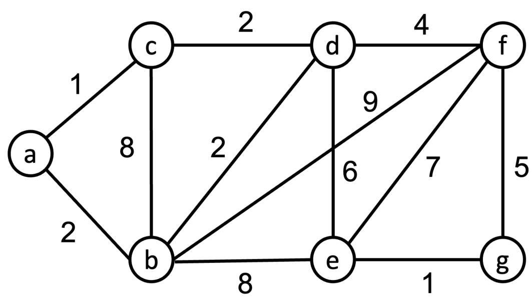
STRONGLY-CONNECTED-COMPONENTS(G)

1. Find a DFS for G to get finish times $u.F$ for each $u \in V$.
2. Compute the *transpose graph* $G^T = (V, E^T)$
3. Find a DFS for G^T, but in the main loop of DFS, always visit nodes based on the reverse order of finish times from the DFS of G.
4. Output the vertices of each tree in the DFS of G^T.

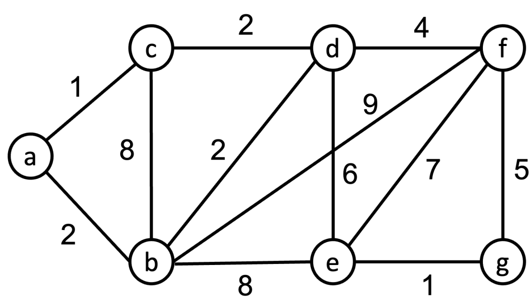
What is the key to making this work? In the second DFS, we essentially visit all of the nodes in the connected components graph in topologically sorted order.

6 Minimum Spanning Trees

Let $G = (V, E, w)$ be an undirected, connected, weighted graph where $w\colon E \to \mathbb{R}^+$ maps each edge to a nonnegative weight.

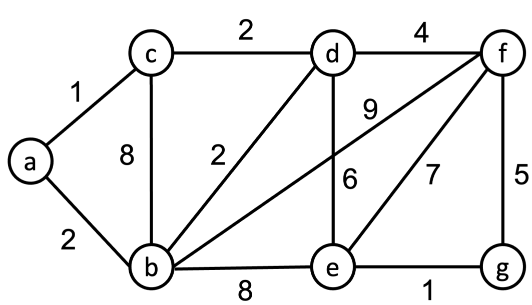


A *spanning tree* of G is a subset of edges $E_T \subseteq E$ such that $T = (V, E_T)$ ______________________



A *minimum spanning tree* of G is a spanning tree T^* that minimizes

over all spanning trees of G.



6.1 Minimum Spanning Tree Terminology

Safe Edges. Let A be a subset of edges that is guaranteed to be in some minimum spanning tree of $G = (V, E)$. An edge $(u, v) \in E$ is *safe* for A if $A \cup \{(u, v)\}$ is contained in some minimum spanning tree of G.

GenericMST(G, w)

1. Set $A = \emptyset$
2. While A is not a spanning tree
3. find a safe edge (u, v) for A
4. $A \leftarrow A \cup \{(u, v)\}$

The key to implementing this method is *finding a safe edge* (u, v) *at each step*.

Cut terminology Let $G = (V, E)$ and A be a set of its edges.

- If $S \subseteq V$, we call the partition $(S, V - S)$ a ______
- An edge $(u, v) \in E$ ____________ the cut $(S, V-S)$ if ________________________________
- A cut ______________ A if no edge in A is cut
- An edge (u, v) is a _________________ if is has minimum weight among all cut edges.

6.2 Generic greedy strategy

We define the following *greedy* strategy for choosing a *safe* edge to add to A
GenericFindSafe(G, A, w)

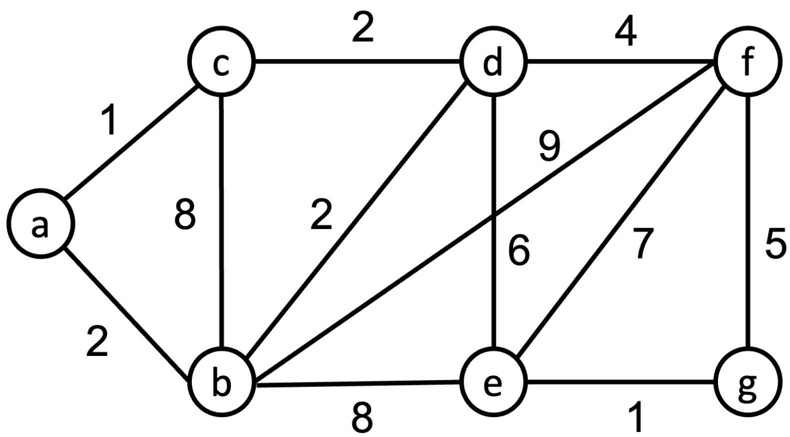
1. Let $(S, V - S)$ be a cut that *respects* A
2. Let (u, v) be a light edge in $(S, V - S)$
3. Return (u, v)

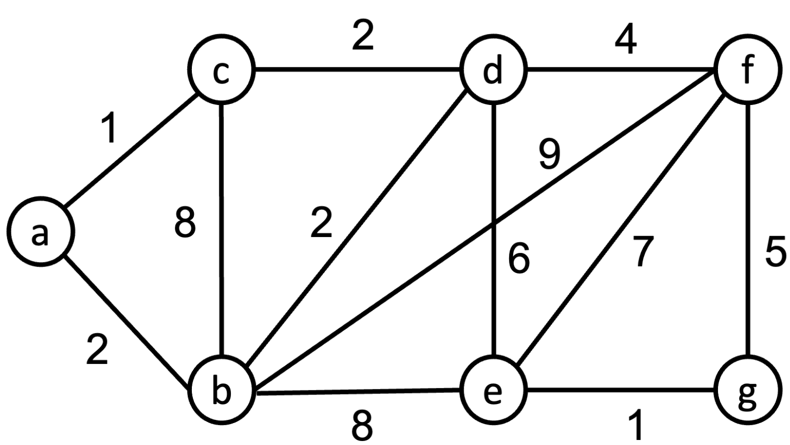
Lemma 6.1. *If A is a subset of the edges in a minimum spanning tree T of G, then* GenericFindSafe *will return a safe edge for A*

6.3 Algorithms of Kruskal and Prim

Kruskal algorithm and Prim's algorithm are two strategies for creating an MST of $G = (V, E)$.

Strategy	**What is A?**	At each step we add:
Kruskal		
Prim		





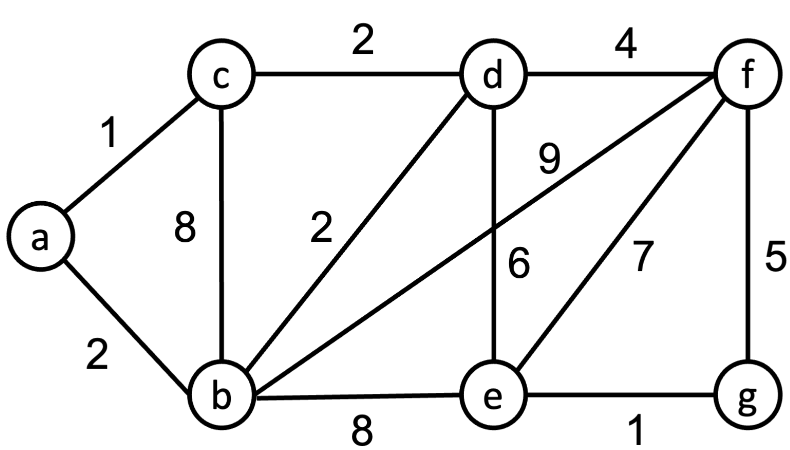
6.4 Prim's Algorithm in More Depth

During Prim's algorithm, we grow out a tree from an arbitrary starting node r. Each node has the following attributes:

- v.parent = parent node of v in the tree we are constructing
- v.key = minimum weight of an edge connecting v to the tree

We will use a min-priority queue Q to store the nodes *not in the tree* with their keys, which allows us to find the node with the smallest key in $O(\log V)$ time.

```
MST-PRIM(G, r)
  for u ∈ V do
      u.key = ∞
      u.parent = NIL
  end for
  r.key = 0
  Q = G.V
  while |Q| > 0 do
      u = EXTRACTMINQ
      if u ≠ r then
          Add edge (u, u.parent) to A
      end if
      for v ∈ Adj[u] do
          if v ∈ Q and w(u, v) < v.key then
              v.parent = u
              v.key = w(u, v)
          end if
      end for
  end while
  Return A
```



Question 2. *What is the runtime of Prim's algorithm, knowing that it takes $O(\log V)$ time to extract the minimum element from Q or update the key of an element in Q?*

A $\Theta(\log V)$

B $\Theta(\log E)$

C $\Theta(V \log V)$

D $\Theta(E \log V)$

E $\Theta(V^2)$