Northwestern University

Graph Theory

MATH 308

just connect the dots bro :)

Author:
Elliott Yoon

29 April 2024

1 Introduction

1.1 Graphs and Graph Models

Definition 1.1. A **graph** G consists of a finite nonempty set V of object called **vertices** and a set E of 2-element subsets of V called **edges**. The sets V and E are the **vertex set** and **edge set** of G, respectively. So a graph G is an ordered pair of two sets V and E, written $G = (V, E)$. Two graphs G and H are called **equal** if $V(G) = V(H)$ and $E(G) = E(H)$, in which case we write $G = H$.

Definition 1.2. We will refer to the number of vertices of a graph as its **order**, and the number of edges as its **size**. A graph with exactly one vertex is called a **trivial graph** (and every graph of order at least 2 is non-trivial).

1.2 Connected Graphs

Definition 1.3. A graph H is called a **subgraph** of a graph G, written $H \subseteq G$, if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. (We can extend the proper subset terminology from set theory to similarly define a proper subgraph.) If a subgraph of G has the same vertex set of G, then it is called a **spanning subgraph** of G.

Definition 1.4. A subgraph F of G is called an **induced subgraph** of G if $u, v \in F$ and $uv \in E(G)$ implies that $uv \in E(F)$. On the other hand, if S is a nonempty set of vertices of a graph G, then the **subgraph of** G **induced by** S is the induced subgraph with vertex set S, denoted $G[S]$.

Definition 1.5. A $u - v$ **walk** W in G is a sequence of vertices in G, beginning with u and ending at v such that the consecutive vertices in the sequence are adjacent. Thus, we can express W as

$$W = (u = v_0, v_1, \dots, v_k),$$

where $k \geq 0$ and v_i and v_{i+1} are adjacent for $i = 0, 1, 2, \dots, k - 1$.

1. If $u = v$, then we say that the walk W is **closed**, while
2. if $u \neq v$, then W is open.

Then number of edges encountered in a walk (counting each multiple occurrence of the duplicate edges) is called the **length** of the walk. (We call a walk of length 0 a **trivial walk**!)

Let's put some parental restrictions on these walks:

Definition 1.6. A $u - v$ **trail** in a graph G is a $u - v$ walk in which no *edge* is traversed more than once. (Note that the repetition of vertices *is* allowed in trails!)

Definition 1.7. A $u - v$ **path** is a $u - v$ walk in which no *vertex* is repeated.

Theorem 1.8

If a graph G contains a $u - v$ walk of length l, then G contains a $u - v$ path of length at most l.

Proof. Let $P = (u = u_0, u_1, \dots, u_i = v)$ be a $u - v$ walk of smallest length k (thus $k \leq l$). Suppose for contradiction that P is not a path. Then there exist some $0 \leq i < j \leq k$ such that $u_i = u_j$. However, we can thus delete the vertices $u_{i+1}, u_{i+2}, \dots, u_j$ from P to obtain the $u - v$ walk

$$(u = u_0, u_1, \dots, u_{i-1}, u_i = u_j, u_{j+1}, \dots, u_k = v),$$

whose length is less than k. □

Definition 1.9. A **circuit** in a graph G is a *closed trail* of length ≥ 3. A **cycle** is a circuit that doesn't repeat any vertices except for the first and last; a k−cycle is a cycle of length k.

Definition 1.10. If there exists a $u - v$ path in G then we say u and v are **connected**. (Note, however, that two vertices do not necessarily need to be adjacent to be connected). Thus, a graph G is said to be **connected** if every two vertices of G are connected. As expected, a graph that is not connected is said to be **disconnected**.

Definition 1.11. A **component** of G is a connected subgraph of G that is not a proper subgraph of any other connected subgraph of G (i.e. a maximal one).

Theorem 1.12

Let R be the relation defined on the vertex set of a graph G by uRv, where $u, v \in V(G)$ if u is connected to v. Then R is an equivalence relation.

Proof. Reflexivity and symmetry are immediate. Let $u, v, w \in V(G)$ such that uRv and vRw. Then we can concatenate the path from u to v with the path from v to w to obtain a $u - w$ walk, which can be reduced to a $u - w$ path by Theorem 1.8. Thus uRw. □

> **Remark 1.13.** The aforementioned equivalence class partitions the vertex set of G, which give the following vibes: *Each vertex and each edge of a graph G belong to exactly one component of G. Then if G is a disconnected graph, and u and v are vertices belonging to different components of g, then $uv \notin E(G)$.*

Theorem 1.14

Let G be a graph of order ≥ 3. If $u, v \in V(G)$ are distinct and both $G - u$ and $G - v$ are connected, then G itself is connected.

Proof. Let $x, y \in V(G)$.

1. If $\{x, y\} \neq \{u, v\}$, then $u \notin \{x, y\}$. Thus $x, y \in V(G - u)$. Since $G - u$ is connected, there exists a $x - y$ path in it, and thus in G.
2. Without loss of generality, say $x = u$ and $y = v$. Since the order of G is greater than 2, there exists some $w \notin \{u, v\}$ in G. Since $G - u$ is connected, there exists some $w - v$ path in it, and since $G - v$ is connected, there exists some $u - w$ path in it. Thus, the concatenated paths give a $u - w - v$ walk in G, which we can reduce to a $u - v$ path by Theorem 1.8.

Thus, arbitrary x and v are connected in G, so G is connected. □

Definition 1.15. Let G be a connected graph, and $u, v \in V$. The **distance** between u and v, often written $d_G(u, v)$ or $d(u, v)$ is the smallest length of any $u - v$ path in G. If $d(u, v) = k$, then we call any $u - v$ path in G of length k a $u - v$ **geodesic**.

Proposition 1.16

If a path $P = (u = v_0, v_1, \ldots, v_k = v)$ is a $u - v$ geodesic, then $d(u, v_i) = i$ for every $0 \leq i \leq k$.

Definition 1.17. We define the diameter of G to be the longest distance between two distinct vertices in G, often written

$$\text{diam}(G) = \max_{u,v \in V(G), u \neq v} d(u, v).$$

Theorem 1.18

If G is a connected graph of order ≥ 3, then G contains two distinct vertices u and v such that $G - u$ and $G - v$ are connected.

Proof. Let $u, v \in V(G)$ such that $d(u, v) = \text{diam}(G)$. Without loss of generality, suppose for contradiction that $G - v$ is disconnected. Thus, there exist distinct $x, y \in V(G - v)$ such that x and y are not connected in $G - v$. However, G *is* connected, so there exist $u - x$ and $u - y$ paths in G. Now, let P be a $u - x$ geodesic and Q be a $u - y$ geodesic, both in G. Since $d_G(u, v) = \text{diam}(G)$, the vertex v can't be in either P or G, so both are paths in $G - v$, which can be concatenated to give an $x - y$ walk, and thus an $x - y$ path in $G - v$, a contradiction! □

Combining Theorem 1.14 and Theorem 1.18 give the following sufficient and necessary condition for connectedness:

Theorem 1.19

Let G be a graph of order ≥ 3. Then G is connected if, and only if, G contains two distinct vertices u and v such that $G - u$ and $G - v$ are connected.

1.3 Common Classes of Graphs

Definition 1.20. We can extend the notion of paths and cycles to categorize certain graphs:

- If the vertices of a graph G of order n can be labeled (or relabeled) $v_1, v_2, \ldots, v_n$ such that its edges are $v_1v_2, v_2v_3, \ldots, v_{n-1}v_n$, then G is called a **path**, written P_n; whereas
- if the vertices of a graph G of order $n \geq 3$ can be labeled (or relabeled) similarly such that its edges are $v_1v_2, v_2v_3, \ldots, v_{n-1}v_n, v_nv_1$, then G is called a **cycle**, written C_n.

Definition 1.21. A graph G of order n is **complete** if every two distinct vertices of G are adjacent, denoted K_n.

> **Remark 1.22.** Note that K_n has the maximum possible size for a graph of order n, which is $\binom{n}{2} = \frac{n(n-1)}{2}$.

Definition 1.23. The **complement** $\overline{G}$ of a graph G is the graph whose vertex set is $V(G)$ and for each pair of distinct $u, v \in V$, $uv \in E(\overline{G})$ if and only if $uv \notin E(G)$. Thus, we can write $\overline{G} = (V(G), \overline{E(G)})$.

> **Remark 1.24.** Note that if G is a graph of order n and size m, then the size of $\overline{G}$ is $\binom{n}{2} - m$. Thus, the graph $\overline{K_n}$ has n vertices and zero edges, and is often called the **empty graph** of order n.

Theorem 1.25

If G is a disconnected graph, then $\overline{G}$ is connected.

Proof. Since G disconnected, it contains at least two separate components. Let $u, v \in \overline{G}$. If u, v belong to separate components, then $uv \in \overline{G}$. Now, assume u, v are in the same component of G, and let $w \in V(G)$ be a vertex which belongs to a different component. Then $uw, vw \notin E(G)$, so $uv, vw \in E(\overline{G})$. Thus, $u - w - v$ is a $u - v$ path in $\overline{G}$. □

Definition 1.26. A graph G with a vertex set $V(G)$ that can be partitioned into two subsets U and W, called **partite sets**, such that every edge of G joins a vertex in U with a vertex in W is called a **bipartite graph**.

Theorem 1.27

A nontrivial graph G is bipartite if, and only if, G contains no odd cycles.

Proof. This one's just a lot of bookkeeping, and I don't feel like writing it out. If you're really curious, just start counting and you'll get there. □

Definition 1.28. If G is bipartite and *every* two vertices in different partite sets are joined by an edge, then G is a **complete k-partite graph**, or also referred to as **complete multipartite graph**. If $|V_i| = n_i$ for $1 \leq i \leq k$, then we denote the complete k−partite graph by $K_{n_1,n_2,\ldots,n_k}$. A complete bipartite graph where one of the bipartite sets has order 1 is called a **star**.

Remark 1.29. If $n_i = 1$ for every i, then $K_{n_1,n_2,\ldots,n_k}$ is the complete graph K_k.

Definition 1.30. If G, H are graphs, the **join** $G + H$ consists of $G \cup H$ and all edges joining a vertex of G and a vertex of H. The **Cartesian product** $G \times H$ has vertex set $V(G \times H) = V(G) \times V(H)$, that is, every vertex of $G \times H$ is an ordered pair (u, v) where $u \in V(G)$ and $v \in V(H)$. Two distinct vertices $(u, v), (x, y)$ are adjacent if either

1. $u = x$ and $vy \in E(H)$, or
2. $v = y$ and $ux \in E(G)$.

Definition 1.31. Define

$$Q_1 := K_2,$$

and for $n \geq 2$, define

$$Q_n := Q_{n-1} \times K_2.$$

The graphs Q_n are called **n−cubes**, or **hypercubes**. They can also be defined as the graph whose vertex set is the set of ordered n−tuples of 0s and 1s (n-bit strings), and where two vertices are adjacent if their ordered n−tuples differ in exactly one position (or coordinate).

1.4 Multigraphs and Digraphs

Definition 1.32. A **multigraph** M consists of a finite nonempty set V of vertices and a set E of edges, where every two vertices of M are joined by a finite number of edges. If two or more edges join the same pair of distinct vertices, then these edges are called **parallel edges**. In a **pseudograph**, not only are parallel edges permitted but an edge is also permitted to join a vertex to itself, with such an edge being called a **loop**.

Definition 1.33. A **digraph** (or **directed graph**) D is a finite nonempty set V of objects called vertices together with a set E of ordered pairs of distinct vertices. The elements of E are called **directed edges** or **arcs**.

2 Degrees

2.1 The Degree of a Vertex

Definition 2.1. The **degree of a vertex** v in a garph G is the number of edges incident with v and is denoted by $\deg_G v$, or simply by $\deg\ v$. Note: $\deg\ v$ is also the number of vertices adjacent to v!

1. We call a vertex of degree 0 an **isolated vertex**, and
2. a vertex of degree 1 an **end-vertex** or a **leaf**.

The **minimum degree** of G is the mininmum degree among the vertices of G, denoted by $\delta(G)$; the **maximum degree** of G is defined similarly and denoted $\Delta(G)$.

Remark 2.2. If G is a graph of order n and $v \in V(G)$, then

$$0 \le \delta(G) \le \deg\ v \le \Delta(G) \le n - 1.$$

Theorem 2.3 (The First Theorem of Graph Theory)
If G is a graph of size m, then

$$\sum_{v \in V(G)} \deg\ v = 2m.$$

Proof. When computing the aggregate sum of degrees of vertices of G, each edge of G is counted twice, once for each of its two incident vertices. □

Corollary 2.4
Every graph has an even number of odd vertices.

Proof. Let G be a graph of size m. We can partition $V(G) = V_1 \cup V_2$, where V_1 consists of all even vertices and V_2 consists of all odd vertices. Clearly, the sum of degrees over vertices in V_1 is even, so by the First Theorem of Graph Theory, we have that

$$\sum_{v \in V_2} \deg v = 2m - \sum_{v \in V_1} \deg v,$$

which implies that the LHS is even, hence the sum of all degrees is also even. □

Theorem 2.5
Let G be a graph of order n. If

$$\deg u + \deg v \ge n - 1,$$

for every nonadjacent vertices $u, v \in V(G)$, then G is connected and $\operatorname{diam}(G) \le 2$.

Proof. Let $x, y \in V(G)$. If $xy \in E(G)$, then x and y are clearly connected by a path of length at most 2. Hence, assume $xy \notin G$. Then, since $\deg x + \deg y \ge n - 1$, we have from the pigeonhole principle that there must exist some vertex $w \in V(G)$ adjacent to both x and y. □

Corollary 2.6

If G is a graph of order n with $\delta(G) \geq \frac{n-1}{2}$, then G is connected.

Proof. Let $u, v \in V(G)$ be nonadjacent vertices. Then

$$\deg u + \deg v \geq \frac{n-1}{2} + \frac{n-1}{2} = n - 1.$$

□

Example 2.7

Suppose $n = 2k$ for some $k \in \mathbb{Z}^+$, and consider the graph $G = 2K_k$, that is, G is the disconnected graph with two components each of which is K_k. If u, v are nonadjacent vertices in G, then they are in different components, each of degree $k-1$, so

$$\deg u + \deg v = (k-1) + (k-1) = 2k - 2 = n - 2.$$

But G is disconnected! Thus, the bound in Theorem 2.5 is sharp.

Remark 2.8. Generally, if G has k components, then the order of some component of G is at most n/k.

2.2 Regular Graphs

Definition 2.9. If $\delta(G) = \Delta(G)$, then the vertices of G have the same degree and we call G **regular**. More specifically, if $\deg v = r$ for every $v \in V$, where $0 \leq r \leq n-1$, then G is r−**regular**.

Theorem 2.10

Let $r, n \in \mathbb{Z}$ with $0 \leq r \leq n-1$. There exists an r−regular graph of order n if, and only if, at least one of r and n is even.

Proof. Since every graph has an even number of odd vertices, there is no r−regular graph of order n if both r and n are odd. Thus, let $r, n \in \mathbb{Z}$ be integers with $0 \leq r \leq n-1$ such that at least one of r and n is even. We will construct an r−regular graph $H_{r,n}$ of order n. Let $V(H_{r,n}) = \{v_1, v_2, \ldots, v_n\}$:

1. First, assume r is even. Then $r = 2k \leq n-1$ for some $0 \leq k \leq \frac{n-1}{2}$. For each $1 \leq i \leq n$, we will join v_i to $v_{i+1}, v_{i+2}, \ldots, v_{i+k}$ and to $v_{i-1}, v_{i-2}, \ldots, v_{i-k}$. Thus, each vertex is adjacent to the k vertices that immediately follow it and the k vertices that immediately precede it. Hence, $H_{r,n}$ is r−regular.

2. Now, assume r is odd. Then $n = 2l$ is even, and $r = 2k+1 \leq n-1$ for some $0 \leq k \leq \frac{n-2}{2}$. Similarly, we will join each v_i to the $2k$ vertices vertices as above, as well as to v_{i+l}. Thus, we are joining each vertex to the k vertices immediately following it, the k vertices immediately preceding it, and the unique vertex *opposite* it. Thus, $H_{r,n}$ is r−regular.

We call the graphs $H_{r,n}$ constructed above **Harary graphs**! □

Theorem 2.11

For every graph G and every integer $r \geq \Delta(G)$, there exists an $r-$regular graph H containing G as an induced subgraph.

Proof. If G is $r-$regular, we're done. Thus, assume G is not an $r-$regular graph. Let G be a graph with $V(G) = \{v_1, \ldots, v_n\}$, and let G' be another copy of G with vertices $\{v'_1, \ldots, v'_n\}$. We will construct a graph G_1 from G and G' by adding edges vv' for all vertices v_i where $\deg v_i < r$ Then, G is an induced subgraph of G_1 and $\delta(G_1) = \delta(G) + 1$. If G_1 is $r-$regular, we're done. Otherwise, repeat this process until we arrive at an $r-$regular graph G_k, where $k = r - \delta(G)$. □

2.3 Degree Sequences

Definition 2.12. A sequence s listing the degrees of vertices of a graph G (in non-decreasing order) is called a **degree sequence** of G.

Definition 2.13. Now, suppose we were given a finite sequence of non-negative integers. We say this sequence is **graphical** if it is a degree sequence of some graph.

Theorem 2.14

A non-increasing sequence $s = \{d_1, d_2, \ldots, d_n\}$ of non-negative integers, where $n \geq 2$, $d_1 \geq 1$ is graphical if, and only if the sequence $s_2 = \{d_2 - 1, d_3 - 1, \ldots, d_{d_1+1} - 1, d_{d_1+2}, \ldots, d_n\}$ is graphical.

Proof. Not doing this one–whole lot of counting and stuff. Look it up if you're *that* curious. (Nerd.) □

3 Isomorphic Graphs

3.1 The Definition of Isomorphism

Informally, two graphs are isomorphic if they differ only in the way they're drawn or labeled. This gives a notion of identifying sameness in structure.

Definition 3.1. Two (labeled) graphs G and H are **isomorphic** if there exists a bijection $\phi : V(G) \to V(H)$ such that $uv \in E(G)$ if, and only if, $\phi(u)\phi(v) \in E(H)$. We call ϕ an **isomorphism** from G to H, say G is **isomorphic to** H, and write $G \cong H$.

Theorem 3.2

Two graphs G and H are isomorphic if, and only if, their complements $\overline{G}$ and $\overline{H}$ are isomorphic.

Remark 3.3. A graph and its complement may be isomorphic! In fact, there's a term for it: a graph G is **self-complementary** if $G \equiv \overline{G}$. This can only occur if G and $\overline{G}$ have the same size, namely

$$\frac{1}{2}\binom{n}{2} = \frac{n(n-1)}{4}.$$

Thus we must have $n \equiv 0 \mod 4$ or $n \equiv 1 \mod 4$.

Theorem 3.4

If G and H are isomorphic graphs, then the degrees of the vertices of G are the same as the degrees of the vertices of H.

Proof. Since G and H are isomorphic, there exists an isomorphism $\phi : V(G) \to V(H)$. Let u be a vertex of G and suppose that $\phi(u) = v \in V(H)$. Suppose u is adjacent to $x_1, \dots, x_k \in V(G)$ and not adjacent to $w_1, \dots, w_l \in V(G)$. Thus, $|V(G)| = k + l + 1$. We have that $\phi(u) = v$ is adjacent to $\phi(x_1), \dots, \phi(x_k) \in V(H)$ and not adjacent to $\phi(w_1), \dots, \phi(w_l) \in V(G)$. Thus, $\deg_H v = k = \deg_G u$. □

Theorem 3.5

Let G and H be isomorphic graphs. Then

1. G is bipartite if, and only if, H is biparte and
2. G is connected if, and only if, H is connected.

3.2 Isomorphism as a Relation

Theorem 3.6

Isomorphism is an equivalence relation on the set of all graphs.

Proof. Equivalence relation proofs are so tedious and unenlightening, and this proof is no different. I'll spare you the details, but just know that it relies on three fundamental properties of bijective functions: (1) Every identity function is bijective, (2) the inverse of every bijective function is also bijective, and (3) the composition of two bijective functions is bijective. □

What do we get from this result? Isomorphisms being an equivalence relation on a set of graphs produces a partition of this set into equivalence classes which are **isomorphism classes**. Basically, if two graphs are in the same isomorphism class, they share all the same properties and structure. Conversely, if they differ in an important property or structure, they're probably (definitely) not isomorphic.

4 Trees

4.1 (Also) Trees

Definition 4.1. A graph G is called **acyclic** if it has no cycles.

Definition 4.2. A **tree** is an acyclic connected graph.

> **Remark 4.3.** Sometimes, it's convenient to select a vertex of a tree T and designate this vertex as the **root** of T. Upon doing so, we then can refer to T as a **rooted tree**.

Definition 4.4. A **forest** is an acyclic graph. Thus, a tree can be alternatively defined as a connected forest. Furthermore, each component of a forest is a tree!

Theorem 4.5

A graph G is a tree if, and only if, every two vertices of G are connected by a unique path.

Proof. ($\Rightarrow$) Let G be a tree. By definition, G must be connected. Suppose that two vertices of G are connected by two unique paths. Note that some (or all) of the edges of these two paths form a cycle. ($\Leftarrow$) Suppose that every two distinct vertices of G are connected by a unique path. Clearly, G is connected. Assuming G has a cycle, let $u \neq v$ be two vertices of G. Then we can obtain two distinct $u - v$ paths using edges from the cycle. □

Theorem 4.6

Every nontrivial tree has at least two leaf nodes.

Proof. Let T be a non-trivial tree and P be the path of greatest length in T. Suppose P is a $u - v$ path, i.e. $P = (u = u_0, u_1, \ldots, u_k = v)$, where $k \geq 1$. If u or v was adjacent to a vertex not in P, then a path of greater length would exist. Since T is acyclic, neither u nor v is adjacent to any other vertices in P, so $\deg u = \deg v = 1$. □

> **Remark 4.7.** A very useful consequence of this result is that if T is a tree of order $k + 1 \geq 2$, then for each leaf node v of T, the subgraph $T - v$ is a tree of order k. Seems useful for, say, induction...

Theorem 4.8

Every tree of order n has size $n - 1$.

Proof. Proceed by induction on n. Up to isomorphism, there is only one tree of order 1: K_1 (which has size 0). Now, assume for some $k \in \mathbb{N}$ that every tree of order k has size $k - 1$. Then, if T is a tree of order $k + 1$, we can choose one of the ≥ 2 leaves that exist, say, v. Then $T - v$ is a tree of order k, which by the induction hypothesis has order $k - 1$. Since v is connected to $T - v$ by exactly one edge, T thus has size $(k - 1) + 1 = k$. □

Corollary 4.9

Every forest of order n with k components has size $n - k$.

Proof. Suppose that the size of a forest F is m, and let $G_1, \dots, G_k$ be the components of F, where $k \geq 1$. Further, let G_i have order n_i and size m_i for $1 \leq i \leq k$. Then

$$n = \sum_{i=1}^{k} n_i$$

and

$$m = \sum_{i=1}^{k} m_i.$$

Since each component G_i is a tree, it follows from Theorem 4.6 that $m_i = n_i - 1$, thus

$$m = \sum_{i=1}^{k} m_i = \sum_{i=1}^{k} (n_i - 1) = n - k.$$

□

Theorem 4.10

The size of every connected graph of order n is at least $n - 1$.

Proof. Let G be a connected graph. This theorem is trivial for $n = 1, 2, 3$, so let $n \geq 4$ and G is the connected graph of smallest order n whose size is at most $n - 2$, i.e. $m + 2 \leq n$. First, suppose that $\delta(G) \geq 2$. Then

$$2m = \sum_{v \in V(G)} \deg v \geq 2n,$$

so

$$m \geq n \geq m + 2,$$

an obvious contradiction. Thus G must have at least one leaf node. Now, let v be a leaf node of G. Since G is connected and has order n and size $m \leq n - 2$, it follows that $G - v$ is connected with order $n - 1$ and size $m - 1 \leq n - 3$, a contradiction with our initial assumption that G was the smallest such connected graph with size at most 2 less than its order. □

Theorem 4.11

Let G be a graph of order n and size m. If G ;satisfies any two of the properties:

1. G is connected,
2. G is acyclic,
3. $m = n - 1$,

then G is a tree.

Proof. Immediate from the previously mentioned theorems. □

Theorem 4.12

Let T be a tree of order k. If G is a graph with $\delta(G) \geq k-1$, then T is isomoprhic to some subgraph of G.

Proof. Proceed by induction on k. Getting kicked out of the library because it's midnight, so I trust you can figure this one out on your own ;) □

4.2 The Minimum Spanning Tree Problem

Definition 4.13. Recall that a subgraph $H \subseteq G$ is a a spanning subgraph of G if H contains every vertex of G. A spanning subgraph H of a connected graph G such that H is a tree is called a **spanning tree** of G.

Theorem 4.14

Every connected graph contains a spanning tree.

Proof. Just construct it, baby. Bob the Builder type shit! □

Definition 4.15. Let G be a a connected graph each of whose edges is assigned a number (called the **cost** or **weight** of the edge); denote the weight of an edge $e \in E(G)$ by $w(e)$. For each subgraph $H \subseteq G$, the **weight** $w(H)$ of H is defined as the sum of the weights of its edges, that is,

$$w(H) = \sum_{e \in E(H)} w(e).$$

Definition 4.16. The spanning tree of a graph G whose weight is minumum among all spanning trees of G is called a **minumum spanning tree**.

There are two (similar) algorithms that are pretty standard in every algorithms class to construct minumum spanning trees:

Algorithm 4.17 (Kruskal) — For a connected weighted graph G, a spanning tree T of G is constructed as follows: For the first edge e_1 of T, we select any edge of G of minimum weight and for the second edge e_2 of T, we select any remaining edge of G of minimum weight. For the third edge e_3 of T, we choose any remaining edge of G of minimum weight that does not produce a cycle with the previously selected edges. We continue in this manner until a spanning tree is produced.

Theorem 4.18

Kruskal's Algorithm produces a minimum spanning tree in a connected weighted graph.

Proof. TODO: but also just look at an algorithms textbook in the meantime. you can prove this with induction on the number of chosen edges, comparing the cost of the subproblem with that of the edges chosen by the optimal solution. □

Algorithm 4.19 (Prim) — For a connected weighted graph G, a spanning tree T of G is constructed as follows: For an arbitrary vertex u for G, an edge of minimum weight incident with u is selected as the first edge e_1 of T. For subsequent edges $e_2, e_3, \ldots, e_{n-1}$, we select an edge of minimum weight among those edges having exactly one of its vertices incident with an edge already selected.

Theorem 4.20

Prim's Algorithm produces a minimum spanning tree in a connected weighted graph.

Proof. See: proof of Kruskal. □