

## 1 Build-Up Error?

Note 5

What is wrong with the following "proof"? In addition to finding a counterexample, you should explain what is fundamentally wrong with this approach, and why it demonstrates the danger of build-up error.

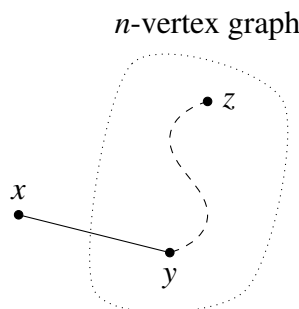
**False Claim:** If every vertex in an undirected graph has degree at least 1, then the graph is connected.

*Proof?* We use induction on the number of vertices  $n \geq 1$ .

*Base case:* There is only one graph with a single vertex and it has degree 0. Therefore, the base case is vacuously true, since the if-part is false.

*Inductive hypothesis:* Assume the claim is true for some  $n \geq 1$ .

*Inductive step:* We prove the claim is also true for  $n + 1$ . Consider an undirected graph on  $n$  vertices in which every vertex has degree at least 1. By the inductive hypothesis, this graph is connected. Now add one more vertex  $x$  to obtain a graph on  $(n + 1)$  vertices, as shown below.



All that remains is to check that there is a path from  $x$  to every other vertex  $z$ . Since  $x$  has degree at least 1, there is an edge from  $x$  to some other vertex; call it  $y$ . Thus, we can obtain a path from  $x$  to  $z$  by adjoining the edge  $\{x, y\}$  to the path from  $y$  to  $z$ . This proves the claim for  $n + 1$ .  $\square$

### Solution:

The mistake is in the argument that “every  $(n + 1)$ -vertex graph with minimum degree 1 can be obtained from an  $n$ -vertex graph with minimum degree 1 by adding 1 more vertex”. Instead of starting by considering an arbitrary  $(n + 1)$ -vertex graph, this proof only considers an  $(n + 1)$ -vertex graph that you can make by starting with an  $n$ -vertex graph with minimum degree 1. As a counterexample, consider a graph on four vertices  $V = \{1, 2, 3, 4\}$  with two edges  $E = \{\{1, 2\}, \{3, 4\}\}$ . Every

vertex in this graph has degree 1, but there is no way to build this 4-vertex graph from a 3-vertex graph with minimum degree 1.

More generally, this is an example of *build-up error* in proof by induction. Usually this arises from a faulty assumption that every size  $n + 1$  graph with some property can be “built up” from a size  $n$  graph with the same property. (This assumption is correct for some properties, but incorrect for others, such as the one in the argument above.)

One way to avoid an accidental build-up error is to use a “*shrink down, grow back*” process in the inductive step: start with a size  $n + 1$  graph, remove a vertex (or edge), apply the inductive hypothesis  $P(n)$  to the smaller graph, and then add back the vertex (or edge) and argue that  $P(n + 1)$  holds.

Let’s see what would have happened if we’d tried to prove the claim above by this method. In the inductive step, we must show that  $P(n)$  implies  $P(n + 1)$  for all  $n \geq 1$ . Consider an  $(n + 1)$ -vertex graph  $G$  in which every vertex has degree at least 1. Remove an arbitrary vertex  $v$ , leaving an  $n$ -vertex graph  $G'$  in which every vertex has degree... uh-oh! The reduced graph  $G'$  might contain a vertex of degree 0, making the inductive hypothesis  $P(n)$  inapplicable! We are stuck—and properly so, since the claim is false!

## 2 Proofs in Graphs

### Note 5

- (a) On the axis from San Francisco traffic habits to Los Angeles traffic habits, Old California is more towards San Francisco: that is, civilized. In Old California, all roads were one way streets. Suppose Old California had  $n$  cities ( $n \geq 2$ ) such that for every pair of cities  $X$  and  $Y$ , either  $X$  had a road to  $Y$  or  $Y$  had a road to  $X$ .

Prove that there existed a city which was reachable from every other city by traveling through at most 2 roads.

[Hint: Induction]

- (b) Consider a connected graph  $G$  with  $n$  vertices which has exactly  $2m$  vertices of odd degree, where  $m > 0$ . Prove that there are  $m$  walks that *together* cover all the edges of  $G$  (i.e., each edge of  $G$  occurs in exactly one of the  $m$  walks, and each of the walks should not contain any particular edge more than once).

[Hint: In lecture, we have shown that a connected undirected graph has an Eulerian tour if and only if every vertex has even degree. This fact may be useful in the proof.]

- (c) Prove that any graph  $G$  is bipartite if and only if it has no tours of odd length.

[Hint: In one of the directions, consider the lengths of paths starting from a given vertex.]

### Solution:

- (a) We prove this by induction on the number of cities  $n$ .

*Base case:* For  $n = 2$ , there’s always a road from one city to the other.

*Inductive Hypothesis:* When there are  $k$  cities, there exists a city  $c$  that is reachable from every other city by traveling through at most 2 roads.

*Inductive Step:* Consider the case where there are  $k + 1$  cities. Remove one of the cities  $d$  and all of the roads to and from  $d$ . Now there are  $k$  cities, and by our inductive hypothesis, there exists some city  $c$  which is reachable from every other city by traveling through at most 2 roads. Let  $A$  be the set of cities with a road to  $c$ , and  $B$  be the set of cities two roads away from  $c$ . The inductive hypothesis states that the set  $S$  of the  $k$  cities consists of  $S = \{c\} \cup A \cup B$ . Now add back  $d$  and all roads to and from  $d$ .

Between  $d$  and every city in  $S$ , there must be a road from one to the other. If there is at least one road from  $d$  to  $\{c\} \cup A$ ,  $c$  would still be reachable from  $d$  with at most 2 road traversals. Otherwise, if all roads from  $\{c\} \cup A$  point to  $d$ ,  $d$  will be reachable from every city in  $B$  with at most 2 road traversals, because every city in  $B$  can take one road to go to a city in  $A$ , then take one more road to go to  $d$ . In either case there exists a city in the new set of  $k + 1$  cities that is reachable from every other city by traveling at most 2 roads.

*Alternate Solution :* Alternatively, we can prove this using properties of directed graphs. Let  $c$  be the city with the largest in-degree. Note that this graph is essentially a complete graph, where each edge is a directed edge instead of an undirected edge. Therefore, the total in degree sums to  $n(n - 1)/2$ , and so does the total out degree. In addition, the in degree + out degree of any vertex must add up to  $n - 1$ .

Because the total in-degree of all vertices is  $n(n - 1)/2$ , The largest in-degree is  $d \geq (n - 1)/2$ . Let  $S$  be the these  $d$  cities that can reach  $c$  by one edge.

For any other city  $x$ , it has to have at least  $(n - 1) - d$  out-degree (because in-degree  $\leq d$ ). Notice that there are  $n$  total vertices, two of which are  $x$  or  $c$ , and  $d$  vertices that connect to  $c$  through one edge. Thus, there are  $n - 2 - d$  other vertices. Since  $x$  has out degree at least  $n - 1 - d > n - 2 - d$ , it must therefore connect to at least one vertex in  $S$  by the pigeonhole principle.

Thus, all vertices are either connected to  $c$  through 1 or 2 edges.

- (b) We split the  $2m$  odd-degree vertices into  $m$  pairs, and join each pair with an edge, adding  $m$  more edges in total. (Here, we allow for the possibility of multi-edges, that is, pairs of vertices with more than one edge between them.) Notice that now all vertices in this graph are of even degree. Now by Euler's theorem the resulting graph has an Eulerian tour. Removing the  $m$  added edges breaks the tour into  $m$  walks covering all the edges in the original graph, with each edge belonging to exactly one walk.
- (c) To prove the claim, we need to prove two directions: if  $G$  is bipartite, it contains no tours of odd length, and if  $G$  contains no tours of odd length, it must be bipartite.

Suppose  $G$  is bipartite, and let  $L$  and  $R$  be the two disjoint sets of vertices such that there does not exist any edge between two vertices in  $L$  or two vertices in  $R$ . Further, suppose there is some tour in  $G$ , and we start traversing this tour at  $v_0 \in L$ .

Since each edge in  $G$  connects a vertex in  $L$  to a vertex in  $R$ , the first edge in the tour connects the start vertex  $v_0$  to a vertex  $v_1 \in R$ . Similarly, the second edge connects  $v_1 \in R$  to  $v_2 \in L$ . In

general, it must be the case that the  $2k$ th edge connects vertex  $v_{2k-1} \in R$  to  $v_{2k} \in L$ , and the  $2k+1$ th edge connects vertex  $v_{2k} \in L$  to  $v_{2k+1} \in R$ .

Since only even numbered edges connect to vertices in  $L$ , and we started our tour in  $L$ , the tour must end with an even number of edges.

For the opposite direction, suppose  $G$  contains no tours of odd length. Without loss of generality, let us consider one connected component of  $G$ ; the following reasoning can be applied to all of the connected components of  $G$ .

Let  $v$  be an arbitrary vertex in  $G$ ; we can divide all of the vertices in  $G$  into two disjoint sets:

$$R = \{u \mid \text{the shortest path from } u \text{ to } v \text{ is even}\}$$

$$L = \{u \mid \text{the shortest path from } u \text{ to } v \text{ is odd}\}$$

We claim that no two vertices in  $L$  are adjacent. For contradiction, suppose there do exist adjacent vertices  $u_1, u_2 \in L$ . Consider the tour consisting of:

- the shortest path from  $v$  to  $u_1$  (odd length)
- the edge  $(u_1, u_2)$  (length 1)
- the shortest path from  $u_2$  to  $v$  (odd length)

This tour has odd length, and contradicts our assumption that  $G$  has no tours of odd length. This means that no two vertices in  $L$  are adjacent.

Similarly, we claim that no two vertices in  $R$  are adjacent. For contradiction, suppose there do exist adjacent vertices  $u_1, u_2 \in R$ . Consider the tour consisting of:

- the shortest path from  $v$  to  $u_1$  (even length)
- the edge  $(u_1, u_2)$  (length 1)
- the shortest path from  $u_2$  to  $v$  (even length)

This tour has odd length, and contradicts our assumption that  $G$  has no tours of odd length. This means that no two vertices in  $R$  are adjacent.

We've just shown that there are no edges between two vertices in  $L$ , and no edges between two vertices in  $R$ . If there are multiple connected components in  $G$ , the same partition can be applied to all of the components. Together, this means that  $G$  is bipartite.

### 3 Binary Trees

**Note 5** You may have seen the recursive definition of binary trees from previous classes. Here, we define binary trees in graph theoretic terms as follows (**Note:** here we will modify the definition of leaves slightly for consistency).

- A binary tree of height  $> 0$  is a tree where exactly one vertex, called the **root**, has degree 2, and all other vertices have degrees 1 or 3. Each vertex of degree 1 is called a **leaf**. The **height**  $h$  is defined as the maximum length of the path between the root and any leaf.

- A binary tree of height 0 is the graph with a single vertex. The vertex is both a leaf and a root.
- (a) Let  $T$  be a binary tree of height  $> 0$ , and let  $h(T)$  denote its height. Let  $r$  be the root in  $T$  and  $u$  and  $v$  be its neighbors. Show that removing  $r$  from  $T$  will result in two binary trees,  $L, R$  with roots  $u$  and  $v$  respectively. Also, show that  $h(T) = \max(h(L), h(R)) + 1$ .
- (b) Using the graph theoretic definition of binary trees, prove that the number of vertices in a binary tree of height  $h$  is at most  $2^{h+1} - 1$ .
- (c) Prove that all binary trees with  $n$  leaves have  $2n - 1$  vertices.

### Solution:

- (a) Since  $r$  has degree 2, removing it will break  $T$  into two connected components; suppose we call them  $L$  and  $R$ . By symmetry, we just need to prove that  $L$  is a binary tree.

Without loss of generality, suppose  $u \in L$ . Before removing  $r$ ,  $u$  must have had degree 1 or 3. If  $u$  had degree 1, then after removing  $r$ ,  $u$  is a single vertex, and would be a binary tree of height 0, and also is a root. If  $u$  had degree 3, then after removing  $r$ ,  $u$  has degree 2, and all other vertices in  $L$  have degree 1 or 3. Thus,  $L$  is a binary tree with root  $u$ ; the same reasoning shows that  $R$  would also be a binary tree with root  $v$ .

To prove that  $h(T) = \max(h(L), h(R)) + 1$ , we note that because  $T$  is a tree, any path from  $r$  to a leaf must go through either  $u$  or  $v$  but not both. This means that the maximum distance from  $r$  to any leaf is one plus either the maximum distance from  $u$  to any leaf in  $L$  (as the path cannot go back through  $r$ ) or the maximum distance from  $v$  to any leaf in  $R$ . Formally, if we define  $\mathcal{L}(L)$  and  $\mathcal{L}(R)$  to be the set of leaves in  $L$  and  $R$  respectively, and  $d(r, l)$  as the length of the path from  $r$  to some leaf  $l$ , then we have

$$\begin{aligned} h(T) &= 1 + \max \left( \max_{l \in \mathcal{L}(L)} (d(u, l)), \max_{l \in \mathcal{L}(R)} (d(v, l)) \right) \\ &= 1 + \max(h(L), h(R)) \end{aligned}$$

- (b) We proceed by induction on the height of the binary tree.

**Base Case:** a binary tree of height 0 is a singleton and so has  $2^1 - 1 = 1$  vertex.

**Inductive Hypothesis:** assume for all  $k < h$ , a binary tree of height  $k$  has at most  $2^{k+1} - 1$  vertices.

**Inductive Step:** By part (a), we can remove the root from a binary tree and obtain two binary trees  $L$  and  $R$  of height  $k$  and  $l$  respectively. Since  $h(T) = \max(h(L), h(R)) + 1$ , we know that  $k, l \leq h - 1$  so we can apply the inductive hypothesis to  $L$  and  $R$ . Thus, we have that the number of vertices in  $T$  is at most

$$\underbrace{(2^{k+1} - 1)}_{\text{from } L} + \underbrace{(2^{l+1} - 1)}_{\text{from } R} + \underbrace{1}_{\text{root}} \leq 2^{(h-1)+1} + 2^{(h-1)+1} - 1 = 2^{h+1} - 1.$$

- (c) We proceed by induction on the number of leaves in the binary tree.

**Base Case:** if a binary tree has one leaf, it is a singleton and so has  $1 = 2 \cdot 1 - 1$  vertices.

**Inductive Hypothesis:** assume for all  $k < n$ , a binary tree with  $k$  leaves has  $2k - 1$  vertices.

**Inductive Step:** For a binary tree  $T$  with  $n > 1$  leaves, remove the root,  $r$  and break  $T$  into binary trees  $L$  and  $R$ . Suppose  $L$  has  $a$  leaves and  $R$  has  $b$  leaves. Note that all the leaves of  $T$  are in  $L$  or  $R$ , as  $n > 1$  implies the root is not a leaf, which means  $a + b = n$ . By the inductive hypothesis,  $L$  has  $2a - 1$  vertices, and  $R$  has  $2b - 1$  vertices, and so the number of vertices in  $T$  is

$$\underbrace{(2a - 1)}_{\text{from } L} + \underbrace{(2b - 1)}_{\text{from } R} + \underbrace{1}_{\text{root}} = 2(a + b) - 1 = 2n - 1.$$

## 4 Touring Hypercube

Note 5

In the lecture, you have seen that if  $G$  is a hypercube of dimension  $n$ , then

- The vertices of  $G$  are the binary strings of length  $n$ .
- $u$  and  $v$  are connected by an edge if they differ in exactly one bit location.

A *Hamiltonian tour* of a graph is a sequence of vertices  $v_0, v_1, \dots, v_k$  such that:

- Each vertex appears exactly once in the sequence.
- Each pair of consecutive vertices is connected by an edge.
- $v_0$  and  $v_k$  are connected by an edge.

- (a) Show that a hypercube has an Eulerian tour if and only if  $n$  is even.
- (b) Show that every hypercube has a Hamiltonian tour.

### Solution:

- (a) In the  $n$ -dimensional hypercube, every vertex has degree  $n$ . If  $n$  is odd, then by Euler's Theorem there can be no Eulerian tour. On the other hand, the hypercube is connected: we can get from any one bit-string  $x$  to any other  $y$  by flipping the bits they differ in one at a time. Therefore, when  $n$  is even, since every vertex has even degree and the graph is connected, there is an Eulerian tour.
- (b) By induction on  $n$ . When  $n = 1$ , there are two vertices connected by an edge; we can form a Hamiltonian tour by walking from one to the other and then back.

Let  $n \geq 1$  and suppose the  $n$ -dimensional hypercube has a Hamiltonian tour. Let  $H$  be the  $n + 1$ -dimensional hypercube, and let  $H_b$  be the  $n$ -dimensional subcube consisting of those strings with initial bit  $b$ .

By the inductive hypothesis, there is some Hamiltonian tour  $T$  on the  $n$ -dimensional hypercube. Now consider the following tour in  $H$ . Start at an arbitrary vertex  $x_0$  in  $H_0$ , and follow the tour  $T$  except for the very last step to vertex  $y_0$  (so that the next step would bring us back to  $x_0$ ). Next take the edge from  $y_0$  to  $y_1$  to enter cube  $H_1$ . Next, follow the tour  $T$  in  $H_1$  backwards from  $y_1$ , except the very last step, to arrive at  $x_1$ . Finally, take the step from  $x_1$  to  $x_0$  to complete the tour. By assumption, the tour  $T$  visits each vertex in each subcube exactly once, so our complete tour visits each vertex in the whole cube exactly once.

To build some intuition, here are the first few cases:

- $n = 1$ : 0, 1

- $n = 2$ : 00, 01, 11, 10

[Take the  $n = 1$  tour in the 0-subcube (vertices with a 0 in front), move to the 1-subcube (vertices with 1 in front), then take the tour backwards. We know 10 connects to 00 to complete the tour.]

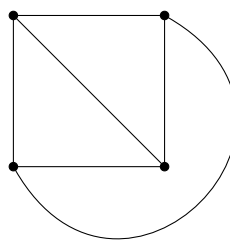
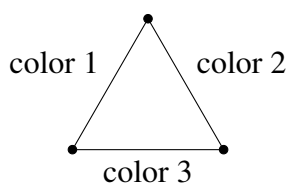
- $n = 3$ : 000, 001, 011, 010, 110, 111, 101, 100

[Take the  $n = 2$  tour in the 0-subcube, move to the 1-subcube, then take the tour backwards. We know 100 connects to 000 to complete the tour.]

The sequence produced with this method is known as a Gray code.

## 5 Edge Colorings

**Note 5** An edge coloring of a graph is an assignment of colors to edges in a graph where any two edges incident to the same vertex have different colors. An example is shown on the left.



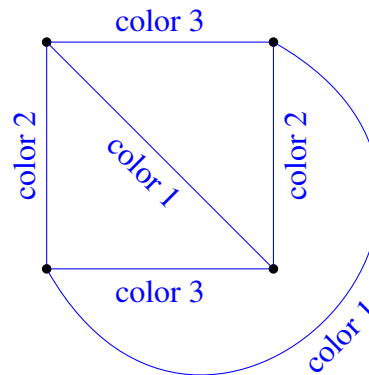
- Show that the 4 vertex complete graph above can be 3 edge colored. (Use the numbers 1,2,3 for colors. A figure is shown on the left.)
- Prove that any graph with maximum degree  $d \geq 1$  can be edge colored with  $2d - 1$  colors.
- Show that a tree can be edge colored with  $d$  colors where  $d$  is the maximum degree of any vertex.

**Solution:**

- (a) Three color a triangle  $u_1, u_2, u_3$  where  $(u_1, u_2)$  is colored 1,  $(u_2, u_3)$  is colored 2, and  $(u_3, u_1)$  is colored 3. This is a valid 3 coloring as the edges are all colored differently.

Consider adding a fourth vertex  $v$ , the incident edges must be colored differently and each incident edge  $(v, u_i)$  needs to be colored differently from the edges incident to  $u_i$ . That is, one can color  $(v, u_1)$  with 2 as it is not incident to the edge colored 2 and that color is available. Similarly one can color edge  $(v, u_2)$  with color 3 and  $(v, u_3)$  with color 1.

Another proof is simply provide a coloring which is below.



- (b) We will use induction on the number of edges  $n$  in the graph to prove the statement: If a graph  $G$  has  $n \geq 0$  edges and the maximum degree of any vertex is  $d$ , then  $G$  can be colored with  $2d - 1$  colors.

*Base case* ( $n = 0$ ). If there are no edges in the graph, then there is nothing to be colored and the statement holds trivially.

*Inductive hypothesis.* Suppose for  $n = k \geq 0$ , the statement holds.

*Inductive step.* Consider a graph  $G$  with  $n = k + 1$  edges. Remove an edge of your choice, say  $e$  from  $G$ . Note that in the resulting graph the maximum degree of any vertex is  $d' \leq d$ . By the inductive hypothesis, we can color this graph using  $2d' - 1$  colors and hence with  $2d - 1$  colors too. The removed edge is incident to two vertices each of which is incident to at most  $d - 1$  other edges, and thus at most  $2(d - 1) = 2d - 2$  colors are unavailable for edge  $e$ . Thus, we can color edge  $e$  without any conflicts. This proves the statement for  $n = k + 1$  and hence by induction we get that the statement holds for all  $n \geq 0$ .

- (c) We will use induction on the number of vertices  $n$  in the tree to prove the statement: For a tree with  $n \geq 1$  vertices, if the maximum degree of any vertex is  $d$ , then the tree can be colored with  $d$  colors.

*Base case* ( $n = 1$ ). If there is only one vertex, then there are no edges to color, and thus can be colored with 0 colors.

*Inductive hypothesis.* Suppose the statement holds for  $n = k \geq 1$ .

*Inductive Step.* Remove any leaf  $v$  of your choice from the tree. We can then color the remaining tree with  $d$  colors by the inductive hypothesis. For any neighboring vertex  $u$  of vertex  $v$ , the degree of  $u$  is at most  $d - 1$  since we removed the edge  $\{u, v\}$  along with the vertex  $v$ . Thus



its incident edges use at most  $d - 1$  colors and there is a color available for coloring the edge  $\{u, v\}$ . This completes the inductive step and by induction we have that the statement holds for all  $n \geq 1$ .

## 6 Planarity and Graph Complements

Note 5

Let  $G = (V, E)$  be an undirected graph. We define the complement of  $G$  as  $\overline{G} = (V, \overline{E})$  where  $\overline{E} = \{(i, j) \mid i, j \in V, i \neq j\} - E$ ; that is,  $\overline{G}$  has the same set of vertices as  $G$ , but an edge  $e$  exists in  $\overline{G}$  if and only if it does not exist in  $G$ .

- Suppose  $G$  has  $v$  vertices and  $e$  edges. How many edges does  $\overline{G}$  have?
- Prove that for any graph with at least 15 vertices,  $G$  being planar implies that  $\overline{G}$  is non-planar.
- Now consider the converse of the previous part, i.e., for any graph  $G$  with at least 15 vertices, if  $\overline{G}$  is non-planar, then  $G$  is planar. Construct a counterexample to show that the converse does not hold.

*Hint: Recall that if a graph contains a copy of  $K_5$ , then it is non-planar. Can this fact be used to construct a counterexample?*

### Solution:

- If  $G$  has  $v$  vertices, then there are a total of  $\frac{v(v-1)}{2}$  edges that could possibly exist in the graph. Since  $e$  of them appear in  $G$ , we know that the remaining  $\frac{v(v-1)}{2} - e$  must appear in  $\overline{G}$ .
- Since  $G$  is planar, we know that  $e \leq 3v - 6$ . Plugging this in to the answer from the previous part, we have that  $\overline{G}$  has at least  $\frac{v(v-1)}{2} - (3v - 6)$  edges. Since  $v$  is at least 15, we have that  $\frac{v(v-1)}{2} \geq \frac{v \cdot 14}{2} = 7v$ , so  $\overline{G}$  has at least  $7v - 3v + 6 = 4v + 6$  edges. Since this is strictly more than the  $3v - 6$  edges allowed in a planar graph, we have that  $\overline{G}$  must not be planar.
- The converse is not necessarily true. As a counterexample, suppose that  $G$  has exactly 15 vertices, of which five are all connected to each other and the remaining ten have no edges incident to them. This means that  $G$  is non-planar, since it contains a copy of  $K_5$ . However,  $\overline{G}$  also contains a copy of  $K_5$  (take any 5 of the 10 vertices that were isolated in  $G$ ), so  $\overline{G}$  is also non-planar. Thus, it is possible for both  $G$  and  $\overline{G}$  to be non-planar.