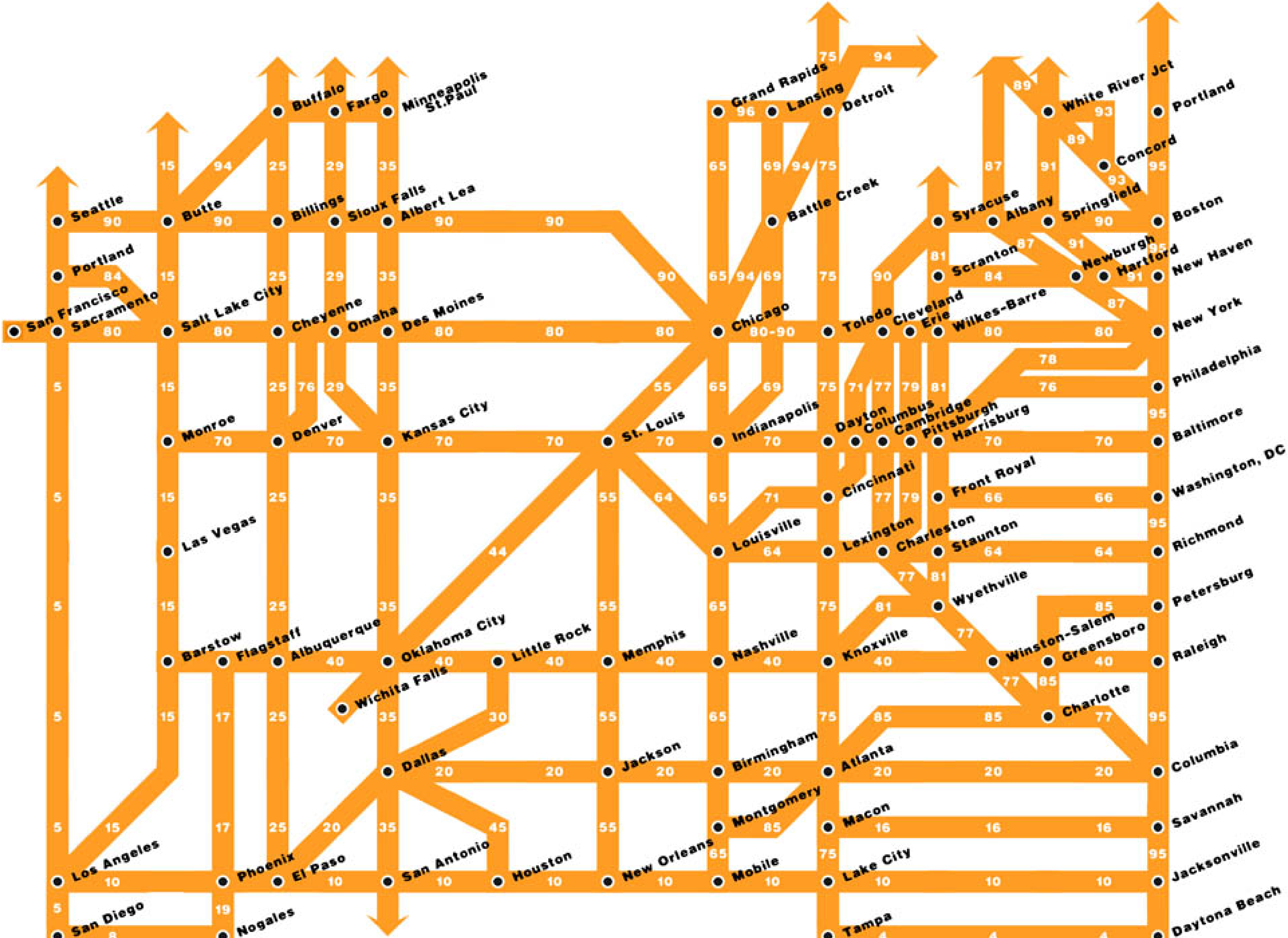
Graph Theory

For those of you who have already completed CS106B/X:



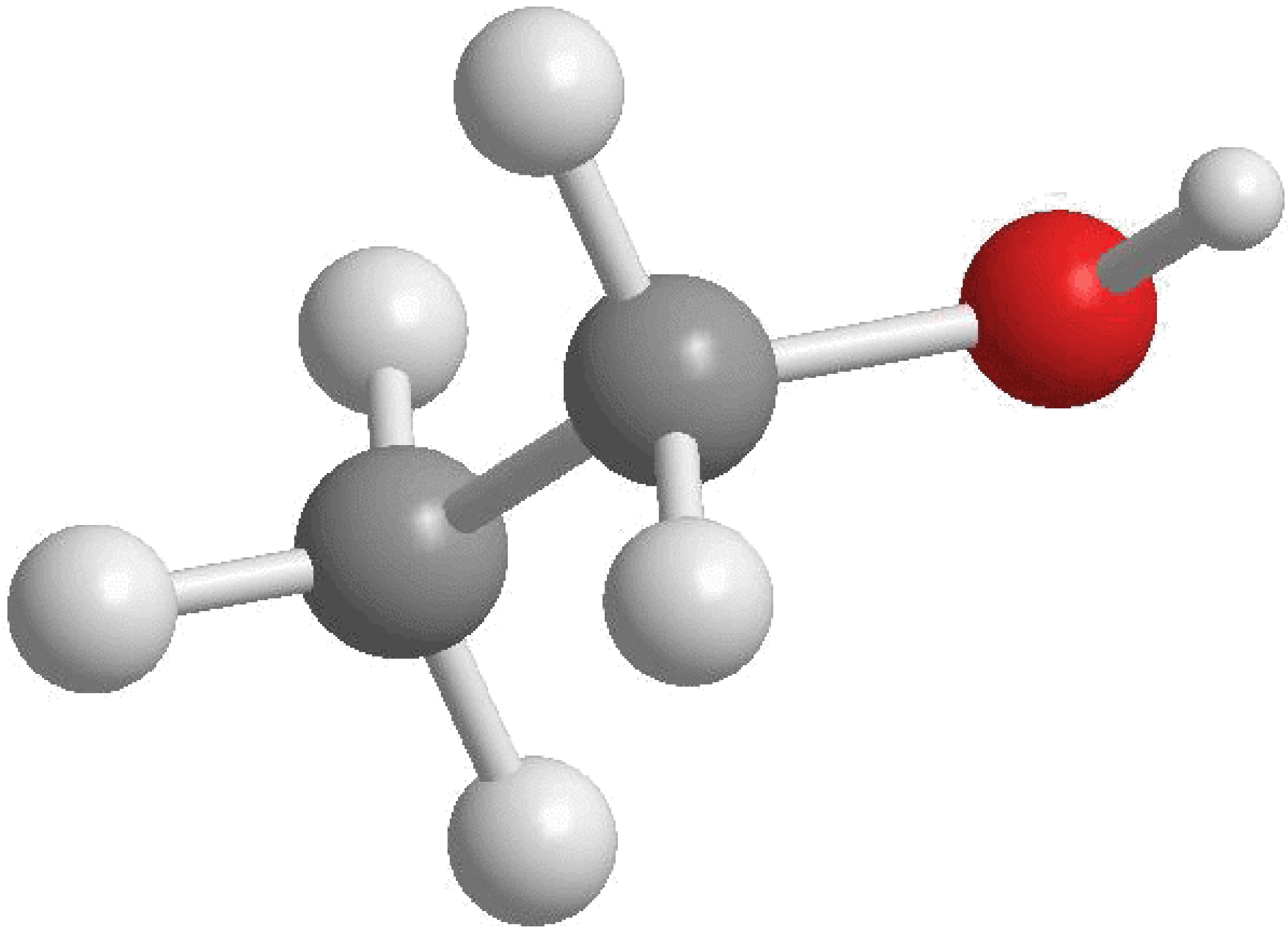
THE EISENHOWER INTERSTATE SYSTEM

(simplified)

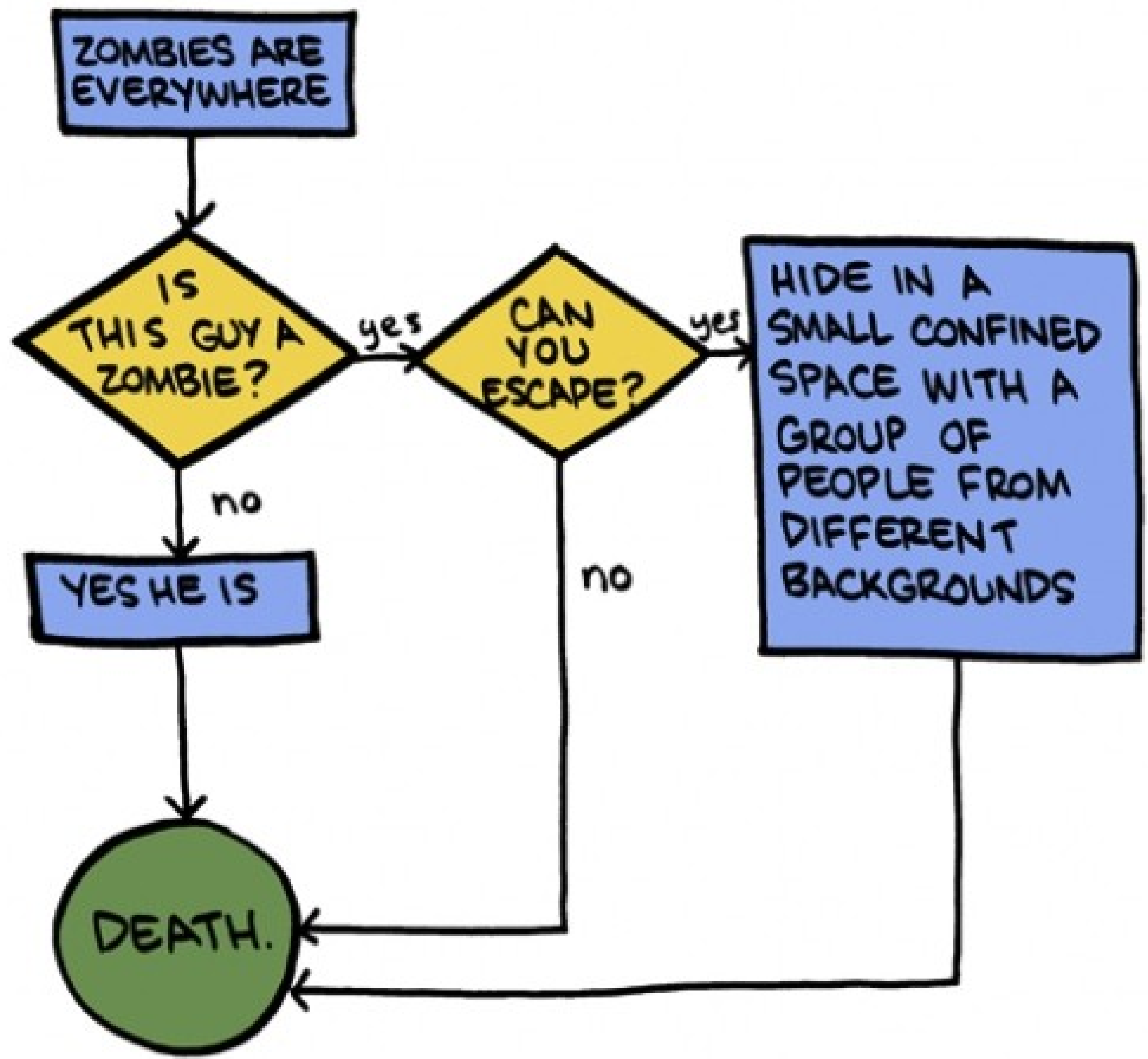
CHRIS YATES 2007

http://strangemaps.files.wordpress.com/2007/02/fullinterstatemap-web.jpg

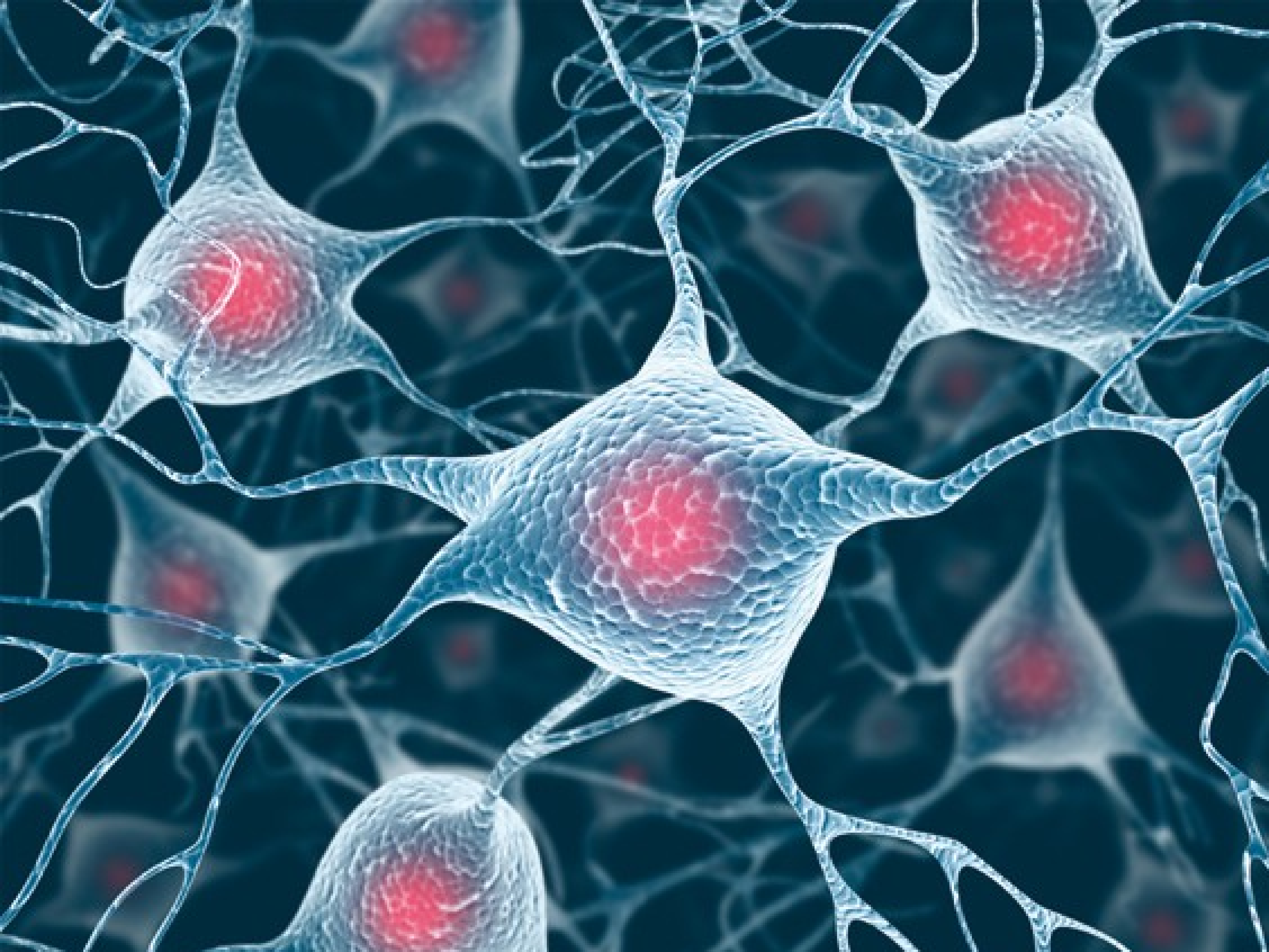
Chemical Bonds



http://4.bp.blogspot.com/-xCtBJ8lKHqA/Tjm0BONWBRI/AAAAAAAAAK4/-



http://www.toothpastefordinner.com/

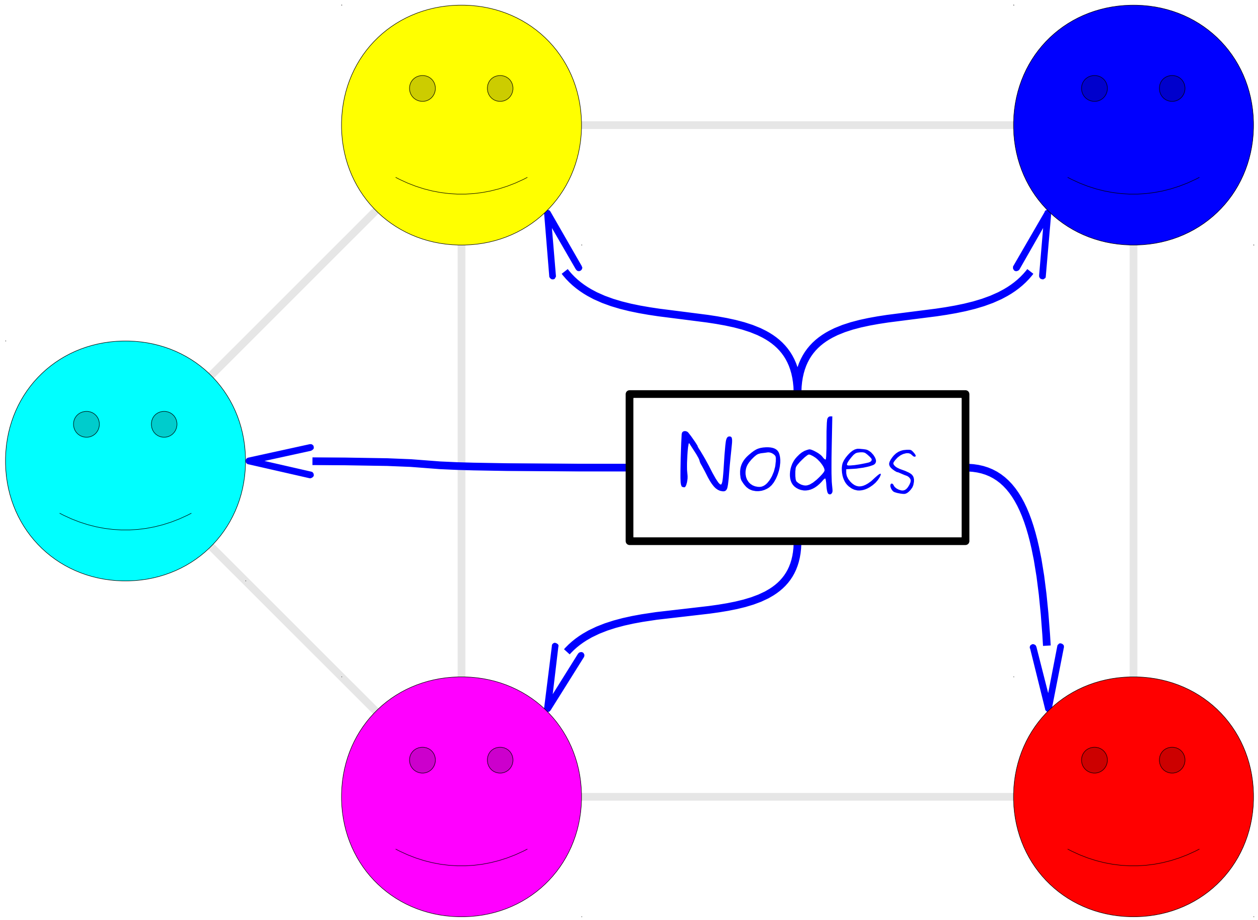


facebook®

What's in Common

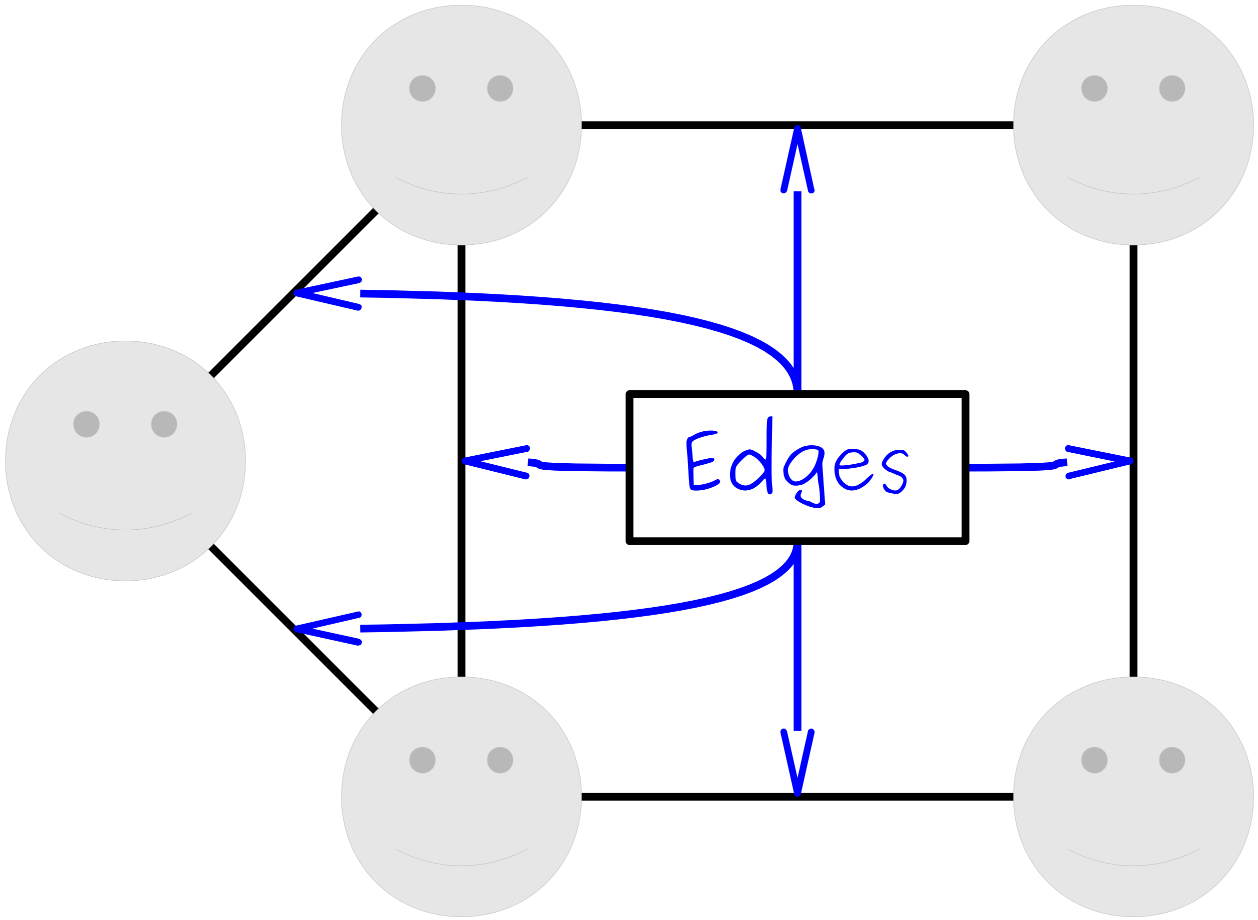
- Each of these structures consists of
  - a collection of objects and
  - links between those objects.
- ***Goal:*** find a general framework for describing these objects and their properties.

A ***graph*** is a mathematical structure for representing relationships.



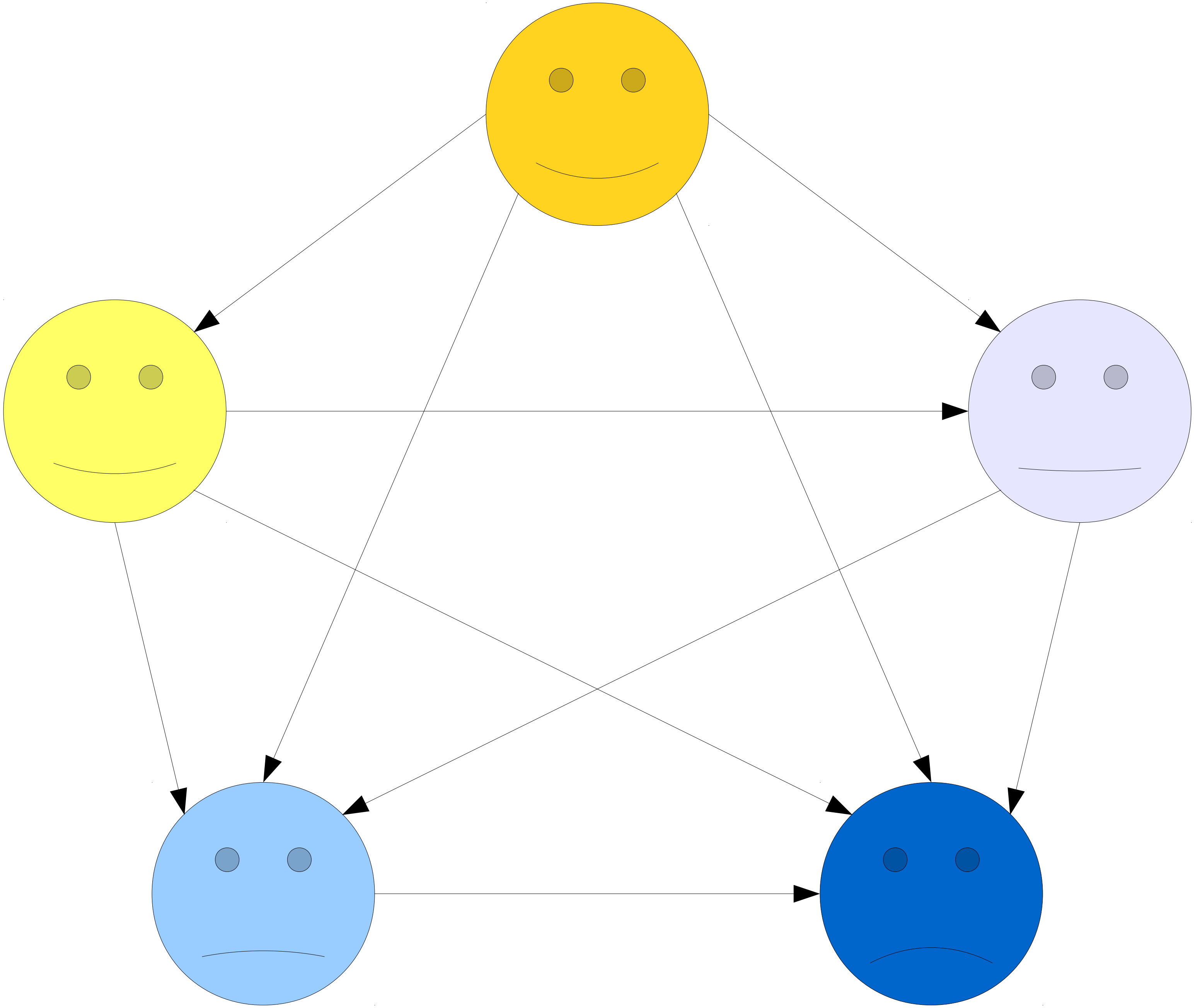
A graph consists of a set of ***nodes*** (or ***vertices***) connected by ***edges*** (or ***arcs***)

A ***graph*** is a mathematical structure for representing relationships.

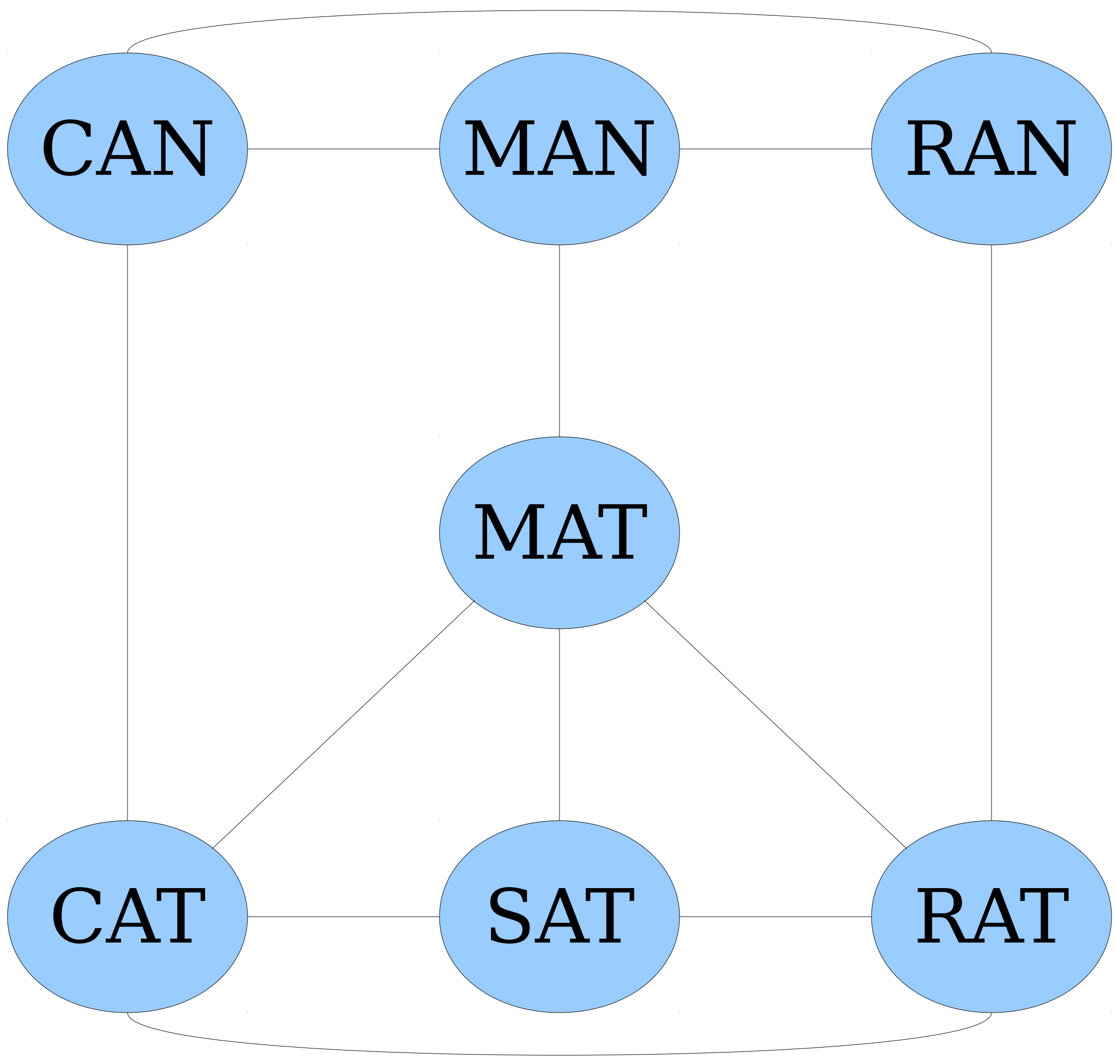


A graph consists of a set of ***nodes*** (or ***vertices***) connected by ***edges*** (or ***arcs***)

Some graphs are ***directed***.



Some graphs are ***undirected***.



Going forward, we're primarily going to focus on undirected graphs.

The term “graph” generally refers to undirected graphs with a finite number of nodes, unless specified otherwise.

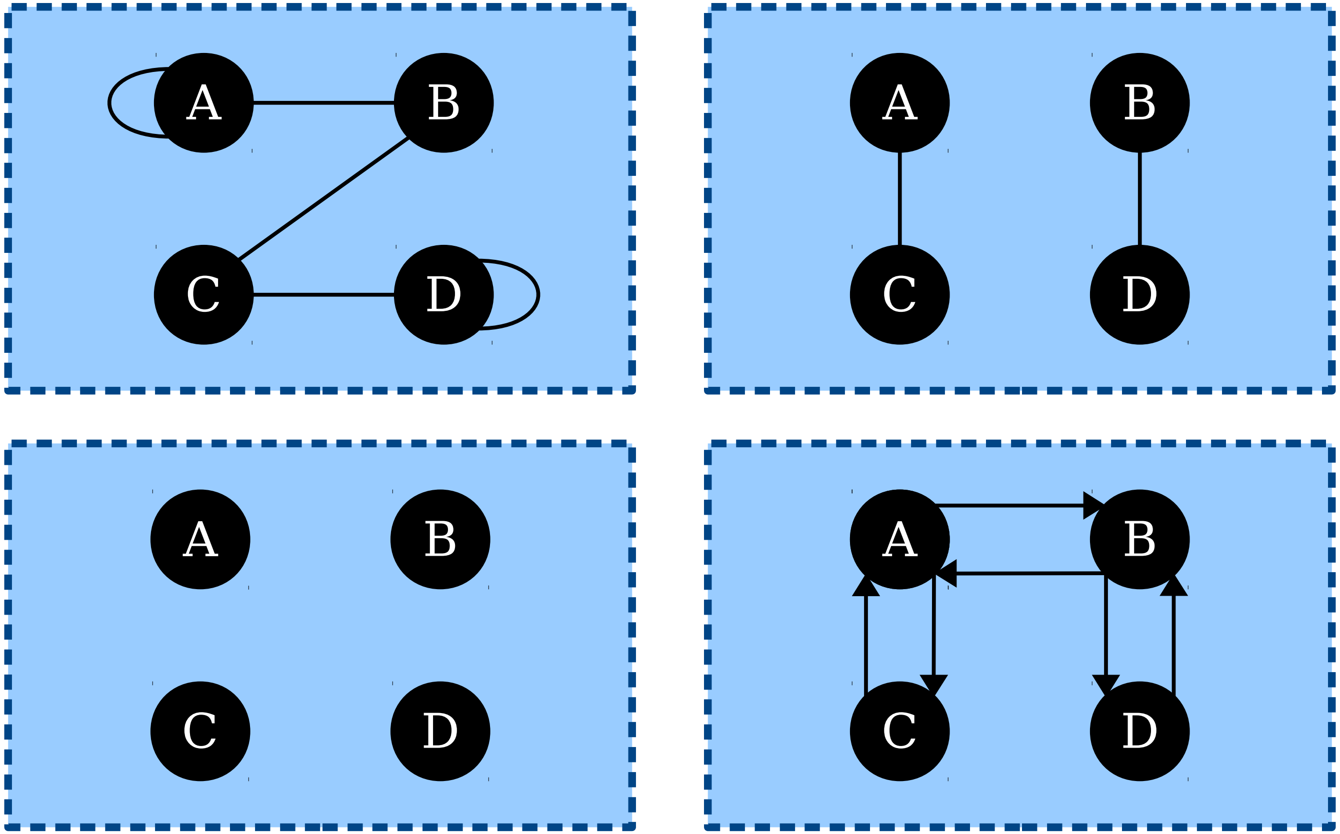
Formalizing Graphs

- How might we define a graph mathematically?
- We need to specify
  - what the nodes in the graph are, and
  - which edges are in the graph.
- The nodes can be pretty much anything.
- What about the edges?

Formalizing Graphs

- An ***unordered pair*** is a set $\{a, b\}$ of two elements $a \neq b$. (Remember that sets are unordered).
  - $\{0, 1\} = \{1, 0\}$
- An ***undirected graph*** is an ordered pair $G = (V, E)$, where
  - V is a set of nodes, which can be anything, and
  - E is a set of edges, which are unordered pairs of nodes drawn from V.
- A ***directed graph*** is an ordered pair $G = (V, E)$, where
  - V is a set of nodes, which can be anything, and
  - E is a set of edges, which are *ordered* pairs of nodes drawn from V.

- An ***unordered pair*** is a set $\{a, b\}$ of two elements $a \neq b$.
- An ***undirected graph*** is an ordered pair $G = (V, E)$, where
  - V is a set of nodes, which can be anything, and
  - E is a set of edges, which are unordered pairs of nodes drawn from V.



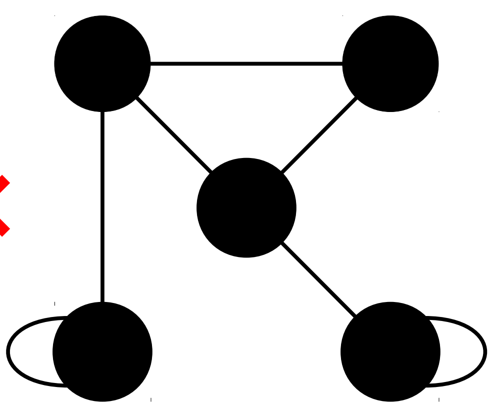
How many of these drawings are of valid undirected graphs?

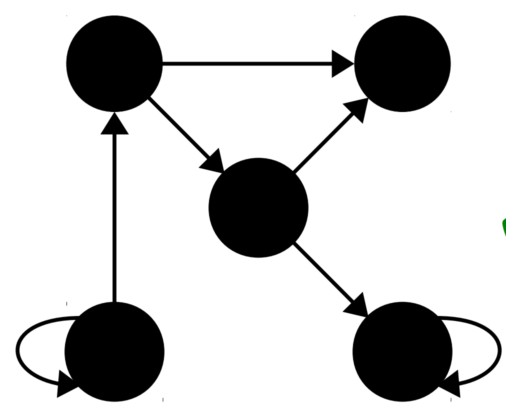
Answer at **PollEv.com/cs103** or
text **CS103** to **22333** once to join, then a number.

Self-Loops

- An edge from a node to itself is called a ***self-loop***.
- In undirected graphs, self-loops are generally not allowed.
  - Can you see how this follows from the definition?
- In directed graphs, self-loops are generally allowed unless specified otherwise.

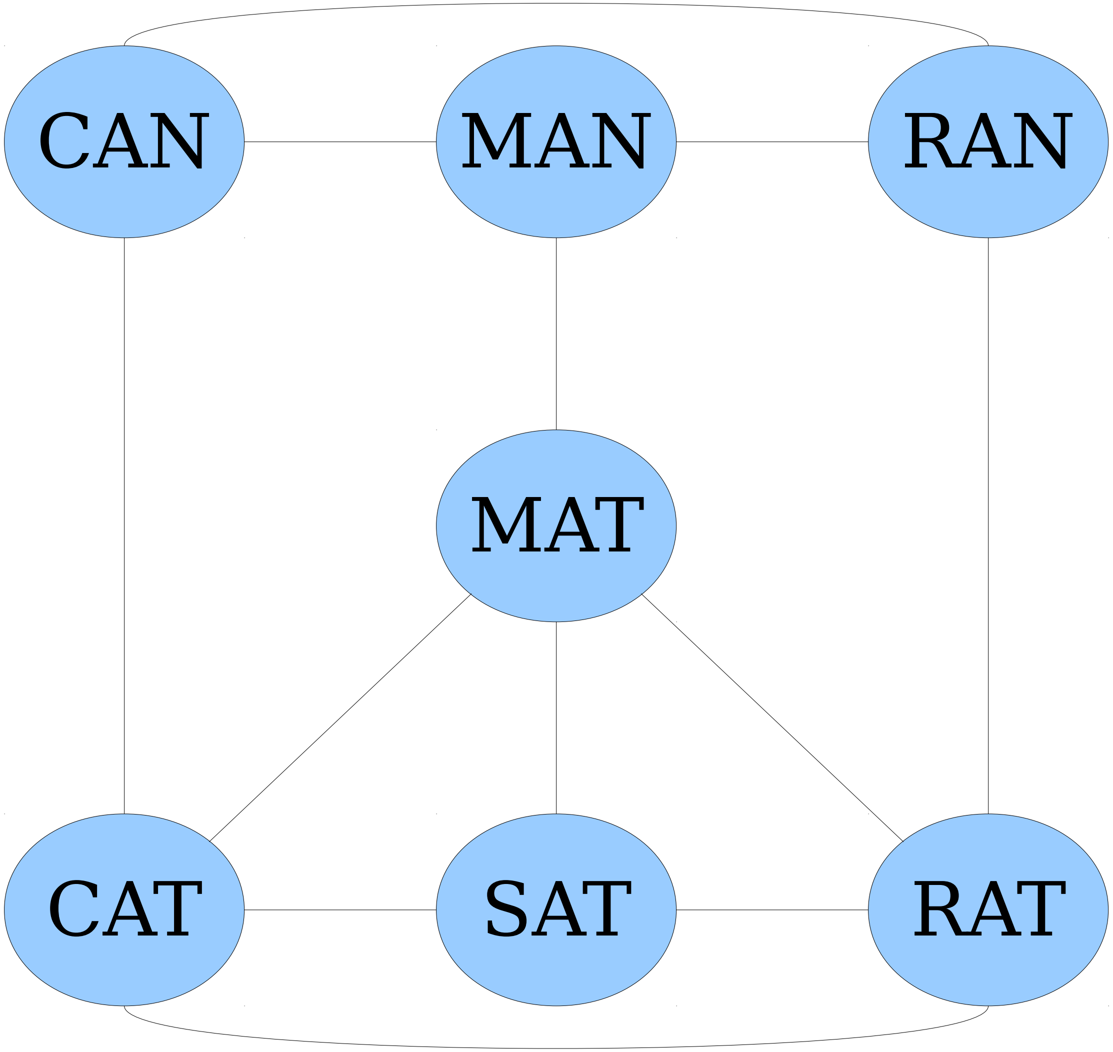
✗





✓

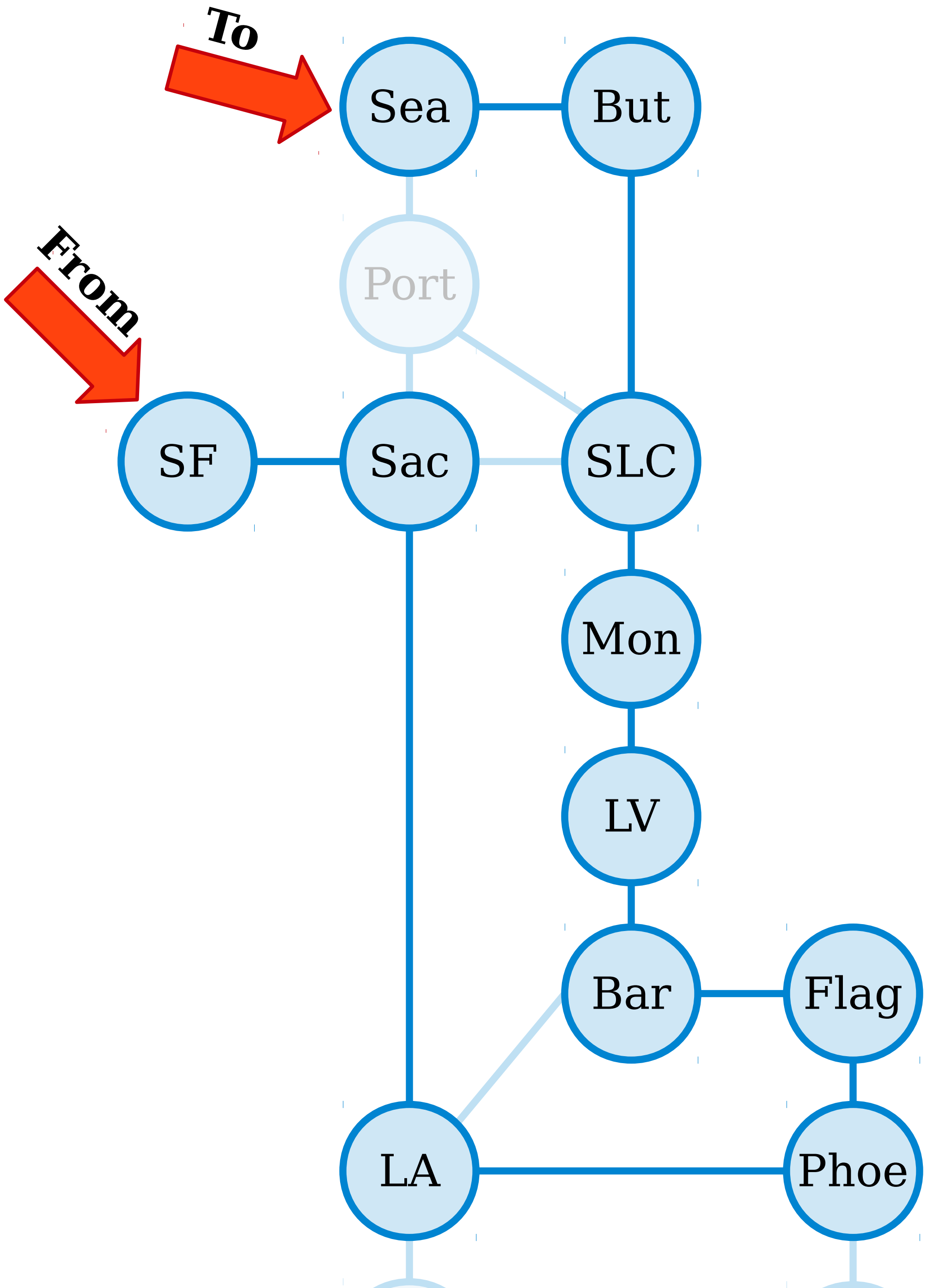
Standard Graph Terminology



Two nodes are called ***adjacent*** if there is an edge between them.

Using our Formalisms

- Let $G = (V, E)$ be a graph.
- Intuitively, two nodes are adjacent if they're linked by an edge.
- Formally speaking, we say that two nodes $u, v \in V$ are ***adjacent*** if $\{u, v\} \in E$.

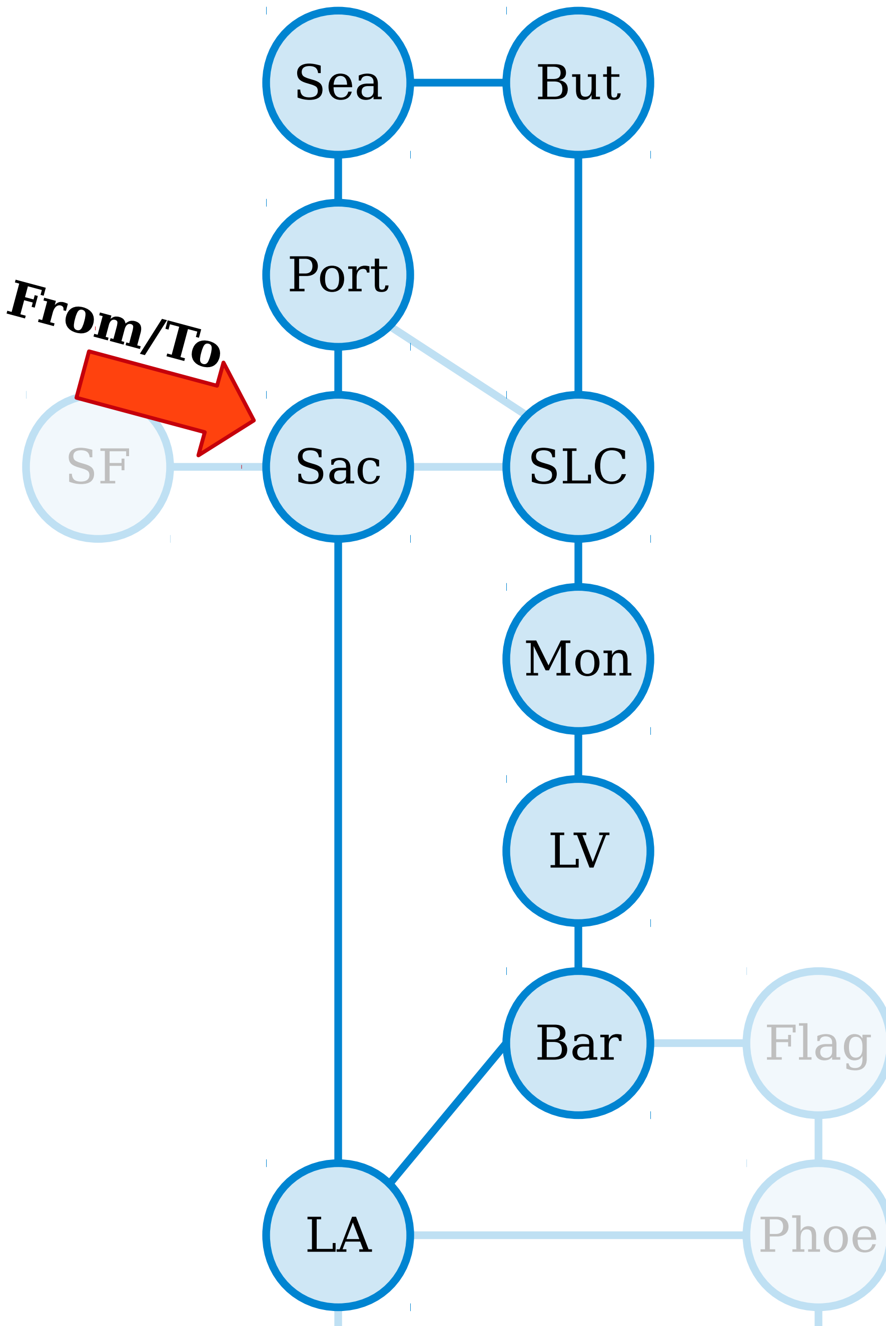


A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

The ***length*** of the path $v_1, \ldots, v_n$ is $n - 1$.

(This path has length 10, but visits 11 cities.)

SF, Sac, LA, Phoe, Flag, Bar, LV, Mon, SLC, But, Sea



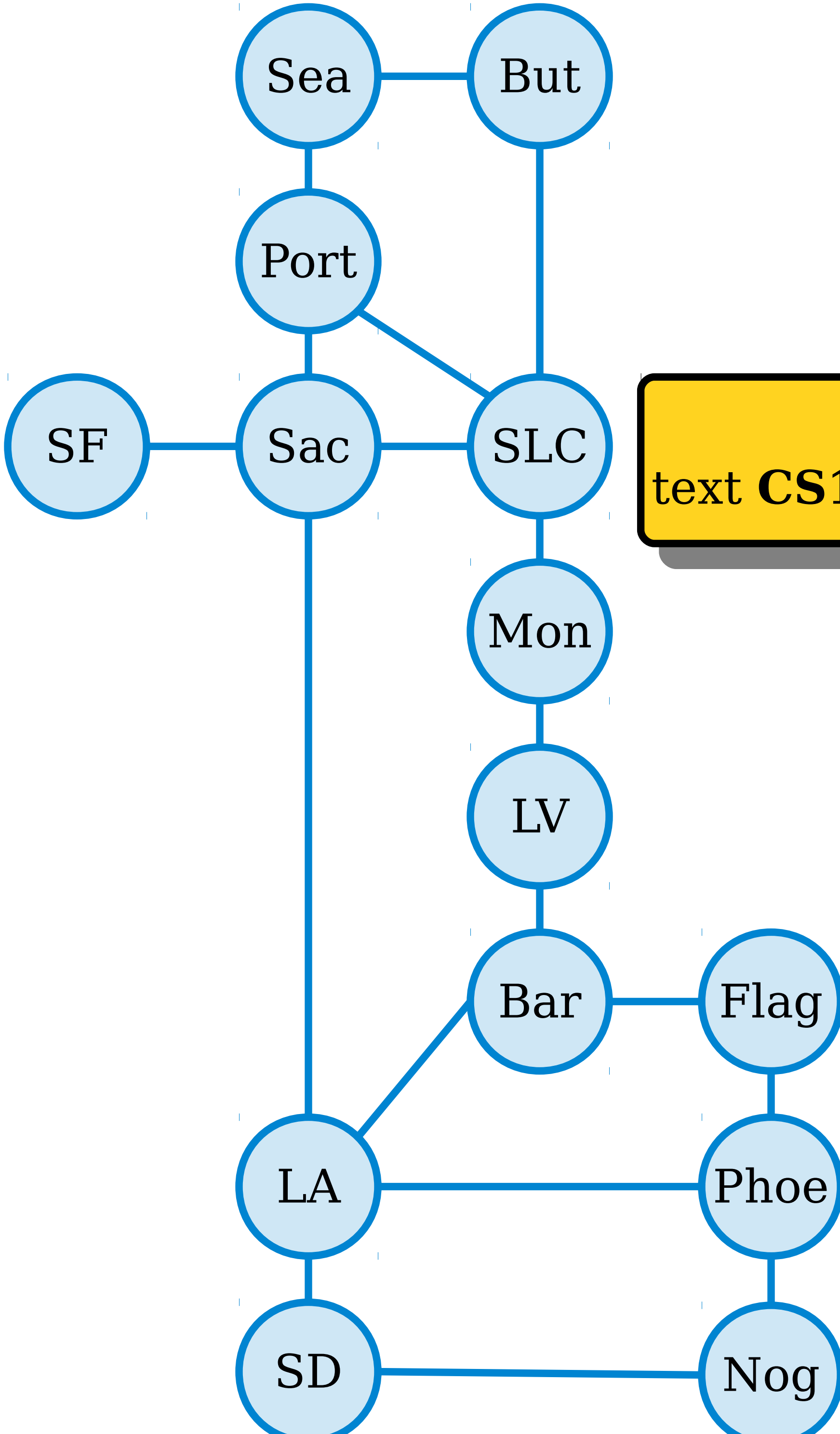
A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

The ***length*** of the path $v_1, \ldots, v_n$ is $n - 1$.

A ***cycle*** in a graph is a path from a node back to itself. (By convention, a cycle cannot have length zero.)

(This cycle has length nine and visits nine different cities.)

Sac, Port, Sea, But, SLC, Mon, LV, Bar, LA, Sac



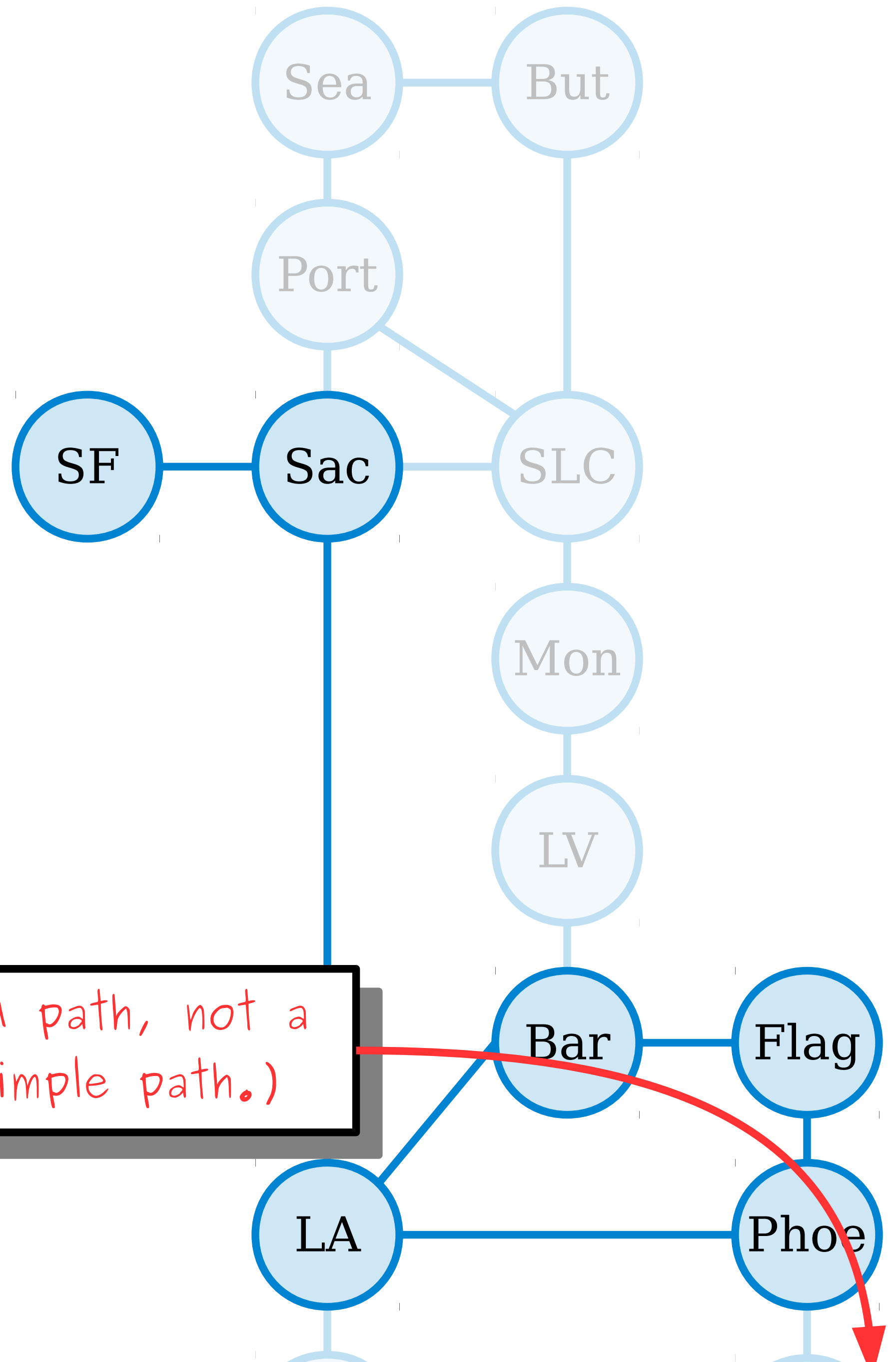
A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

Answer at **PollEv.com/cs103** or
text **CS103** to **22333** once to join, then **A**, ..., or **E**.

A ***cycle*** in a graph is a path from a node back to itself. (By convention, a cycle cannot have length zero.)

How many paths in this graph are there from SF to LA?

A. 1
B. 4
C. 10
D. 20
E. None of these.



(A path, not a simple path.)

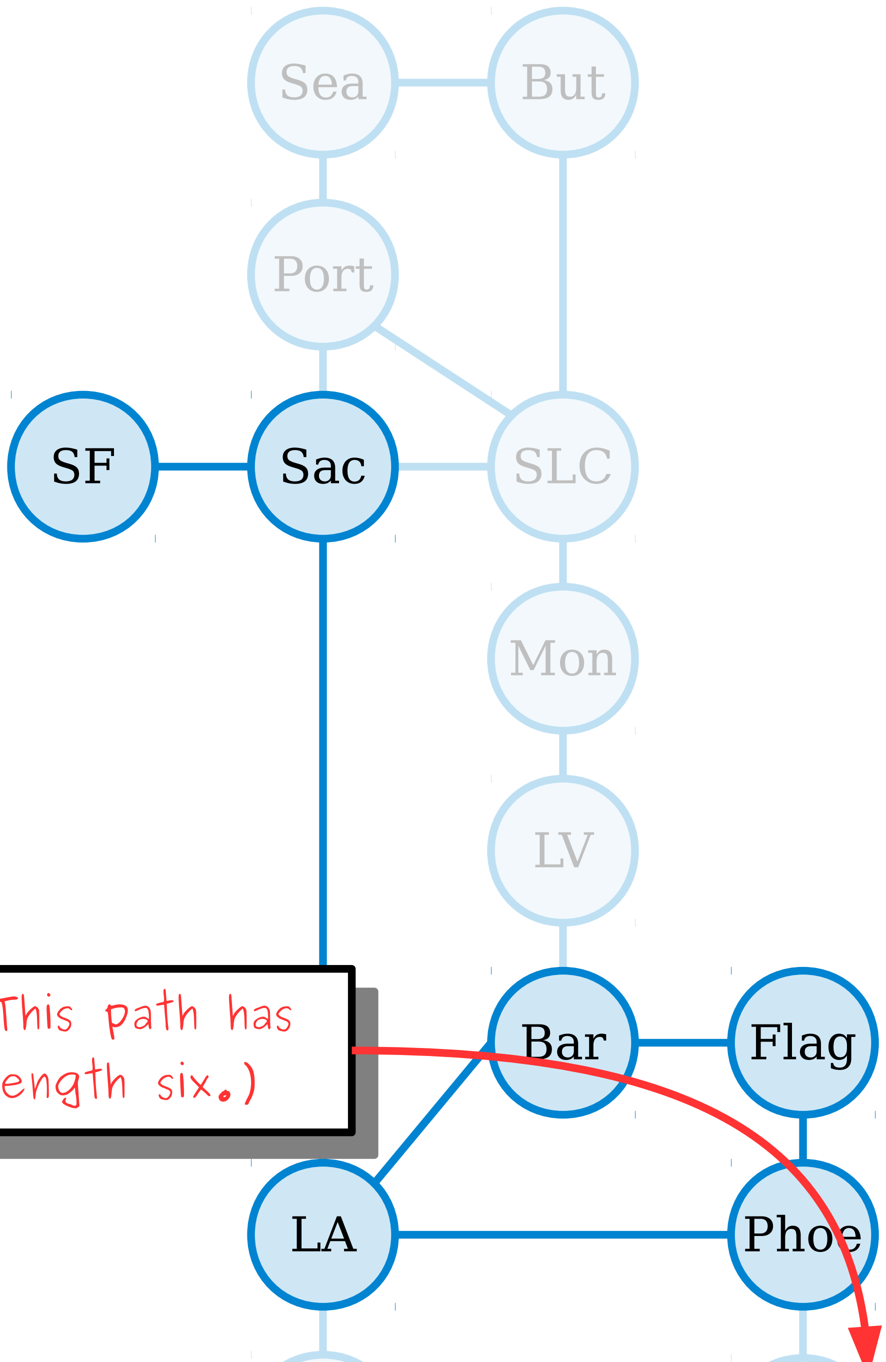
SF, Sac, LA, Phoe, Flag, Bar, LA

A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

The ***length*** of the path $v_1, \ldots, v_n$ is $n - 1$.

A ***cycle*** in a graph is a path from a node back to itself. (By convention, a cycle cannot have length zero.)

A ***simple path*** in a graph is path that does not repeat any nodes or edges.



A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

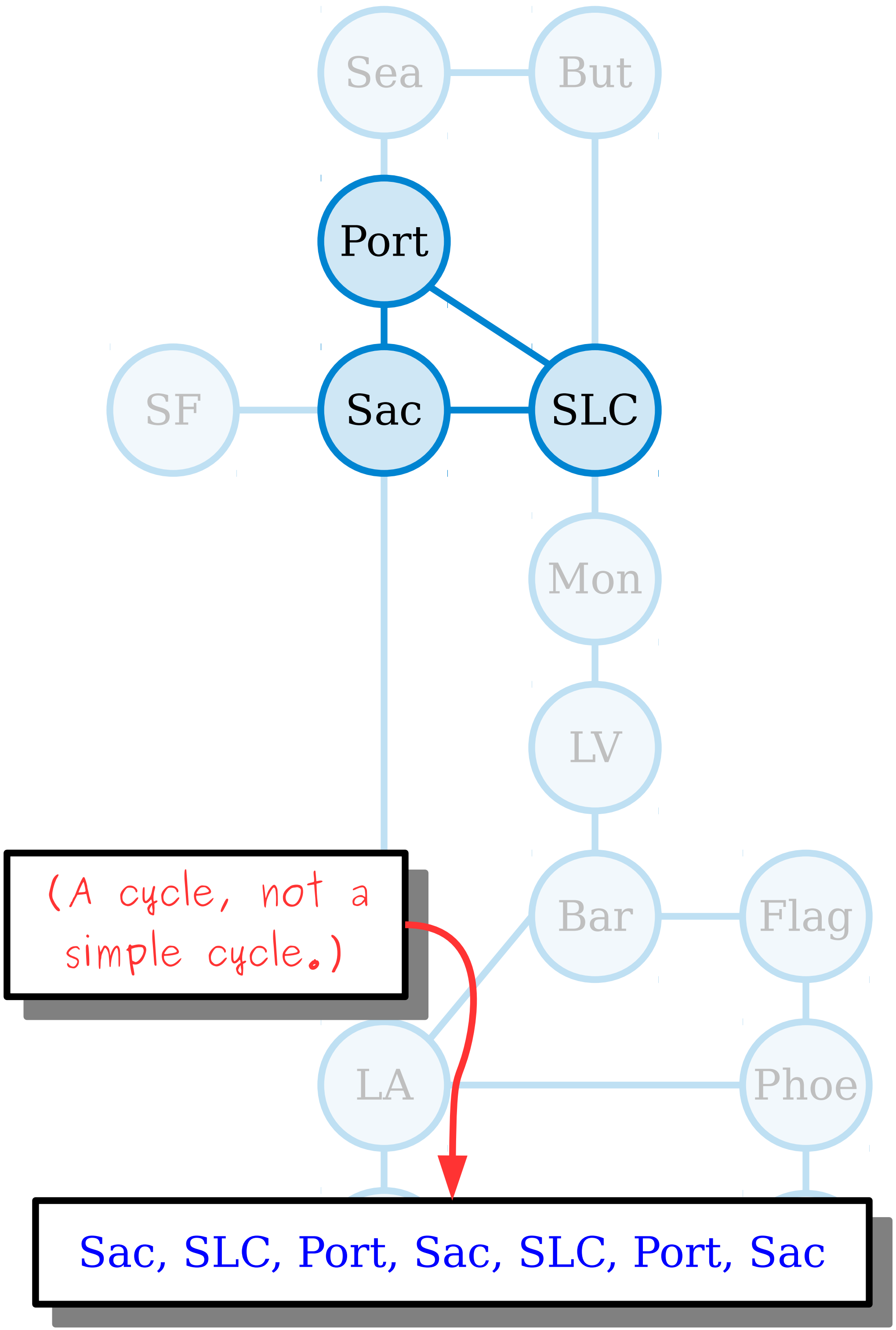
The ***length*** of the path $v_1, \ldots, v_n$ is $n - 1$.

A ***cycle*** in a graph is a path from a node back to itself. (By convention, a cycle cannot have length zero.)

A ***simple path*** in a graph is path that does not repeat any nodes or edges.

(This path has length six.)

SF, Sac, LA, Phoe, Flag, Bar, LA



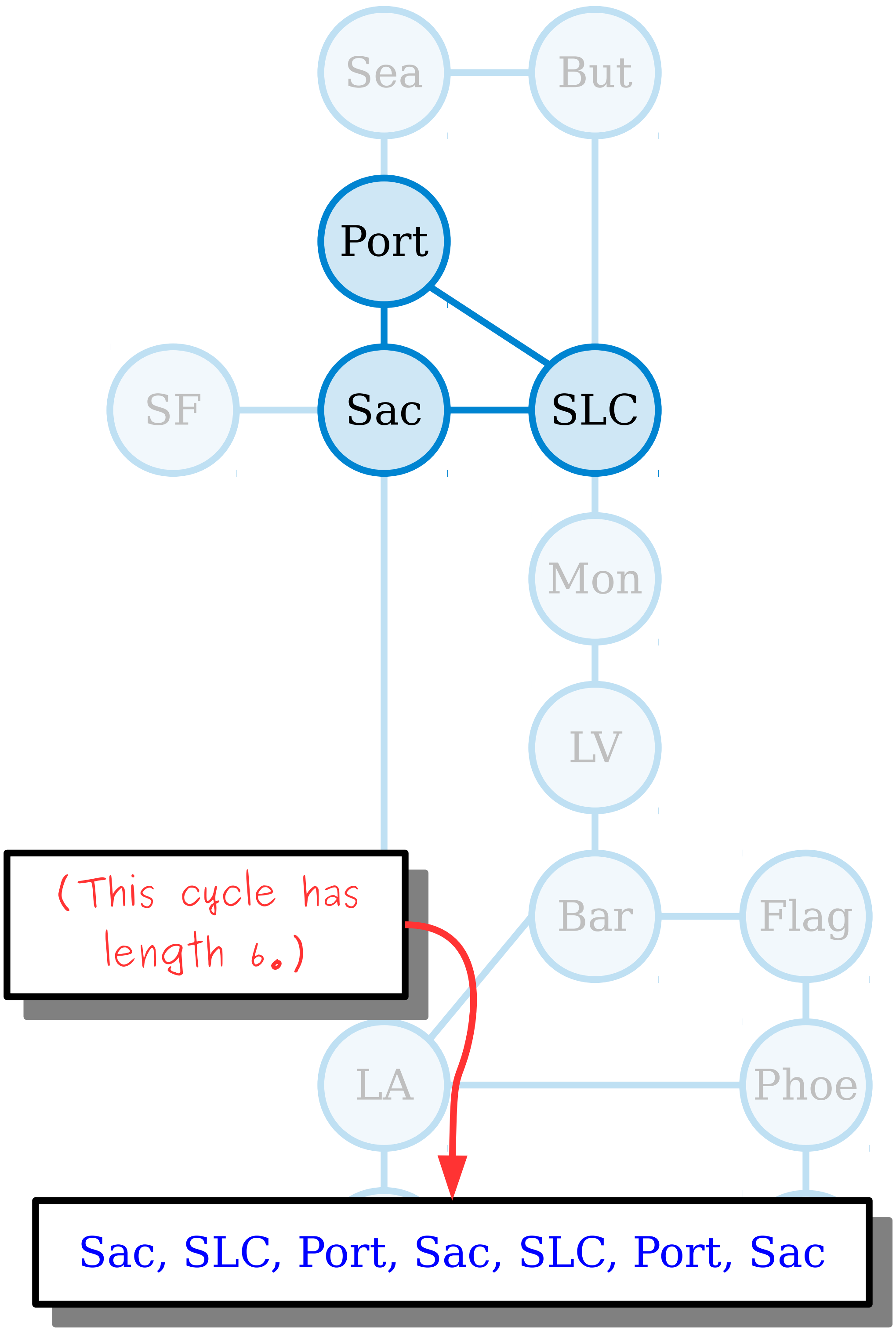
A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

The ***length*** of the path $v_1, \ldots, v_n$ is $n - 1$.

A ***cycle*** in a graph is a path from a node back to itself. (By convention, a cycle cannot have length zero.)

A ***simple path*** in a graph is path that does not repeat any nodes or edges.

A ***simple cycle*** in a graph is cycle that does not repeat any nodes or edges except the first/last node.



A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

The ***length*** of the path $v_1, \ldots, v_n$ is $n - 1$.

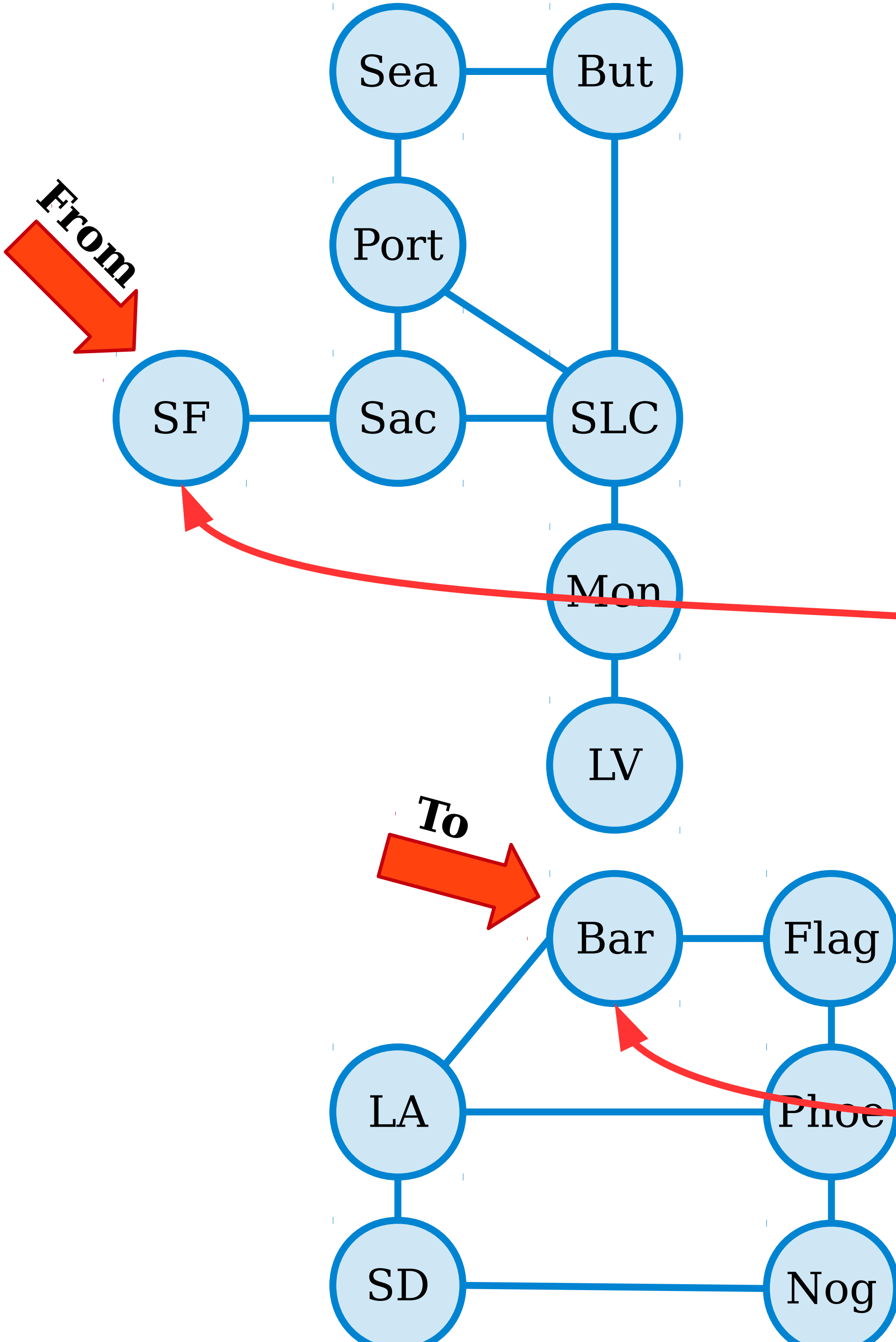
A ***cycle*** in a graph is a path from a node back to itself. (By convention, a cycle cannot have length zero.)

A ***simple path*** in a graph is path that does not repeat any nodes or edges.

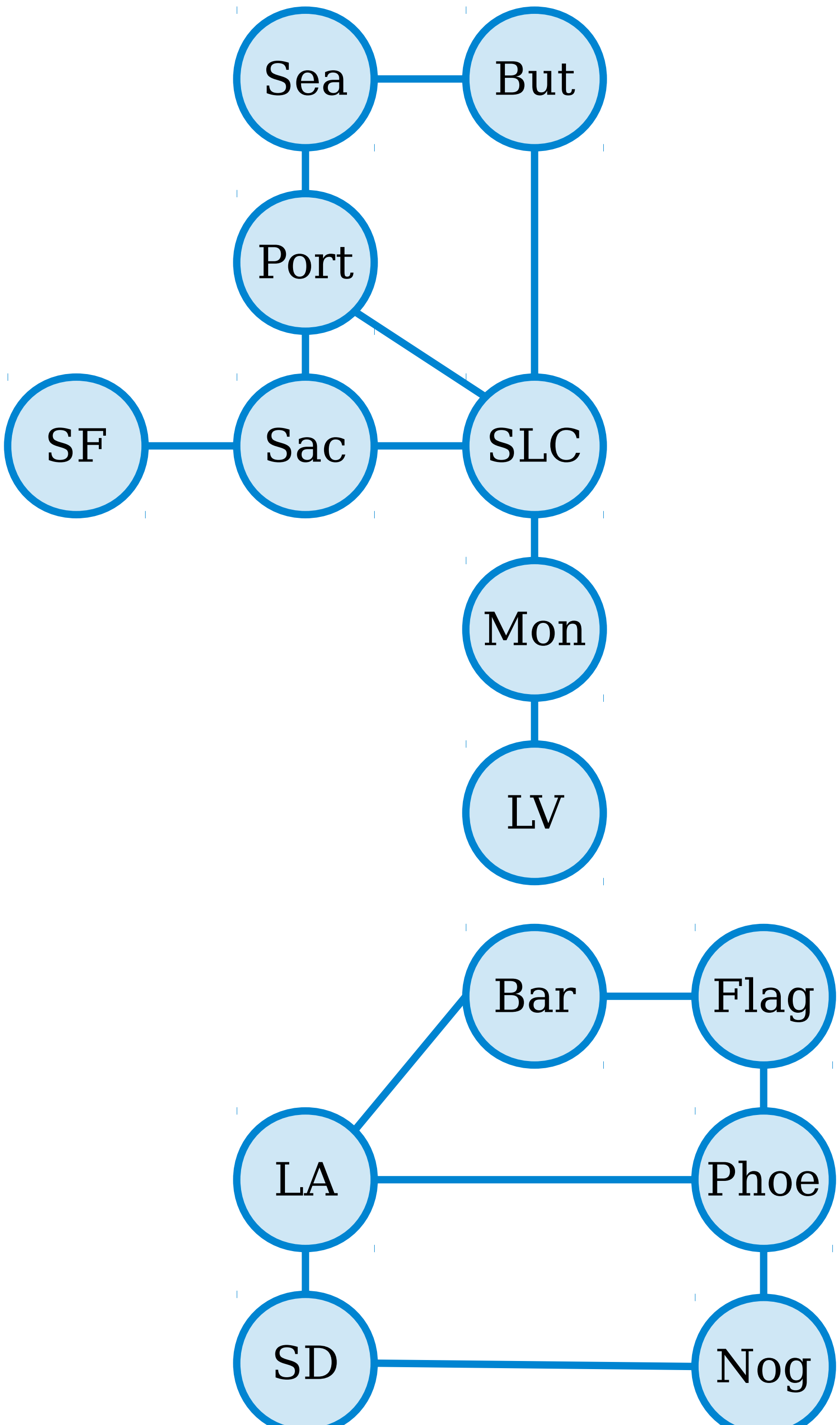
A ***simple cycle*** in a graph is cycle that does not repeat any nodes or edges except the first/last node.

A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

Two nodes in a graph are called ***connected*** if there is a path between them.



(These nodes are not connected. No Grand Canyon for you.)



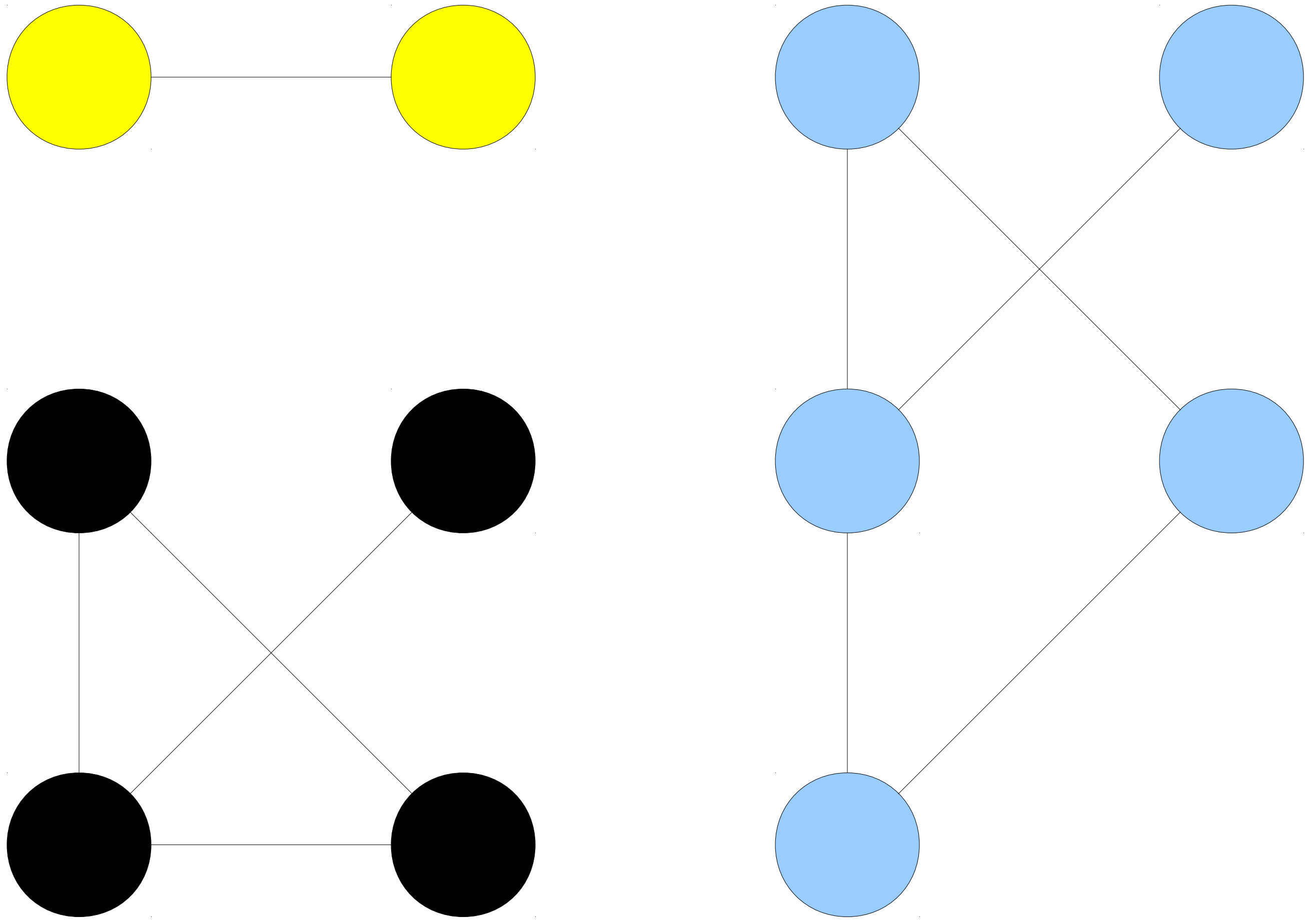
A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.

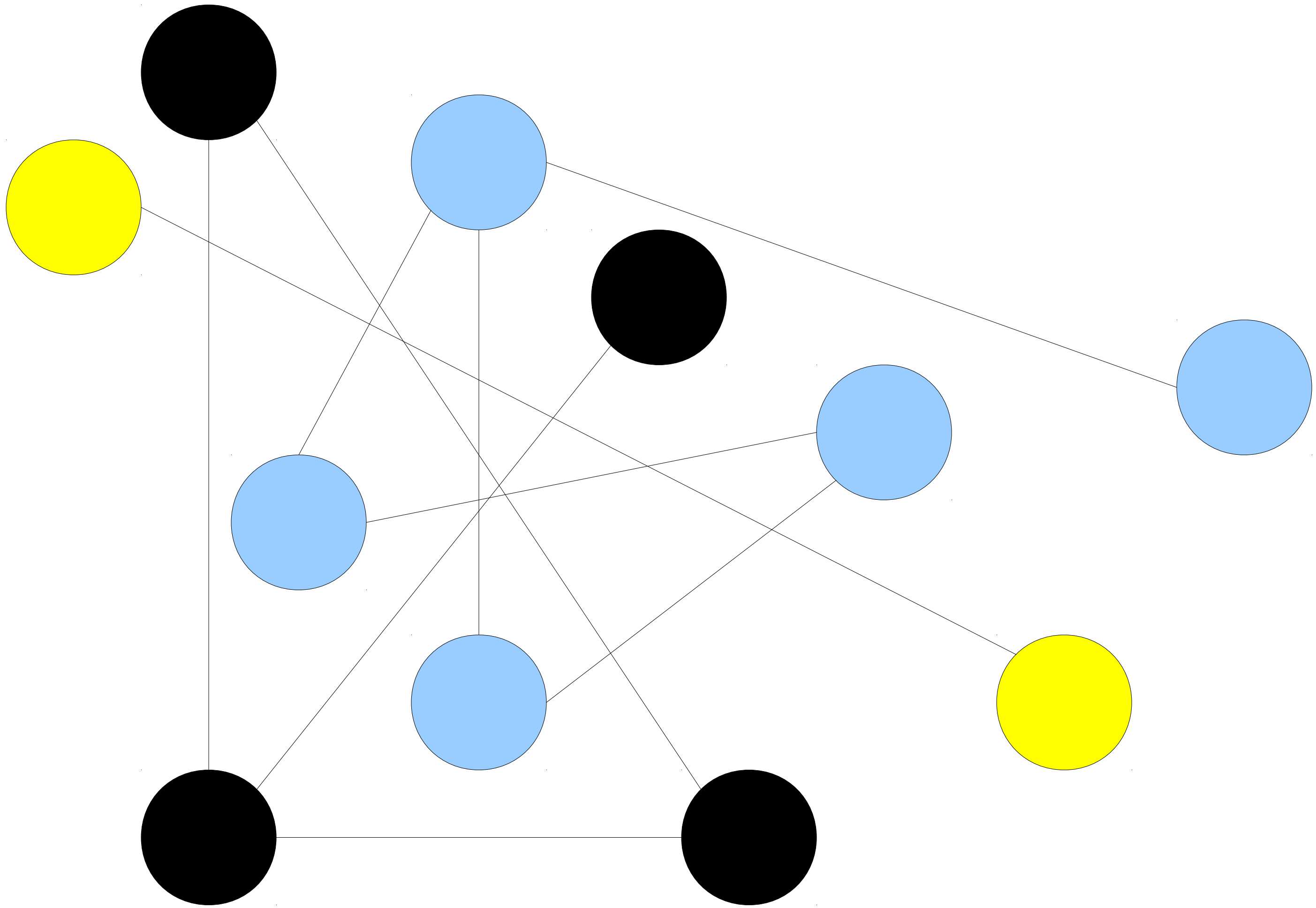
Two nodes in a graph are called ***connected*** if there is a path between them.

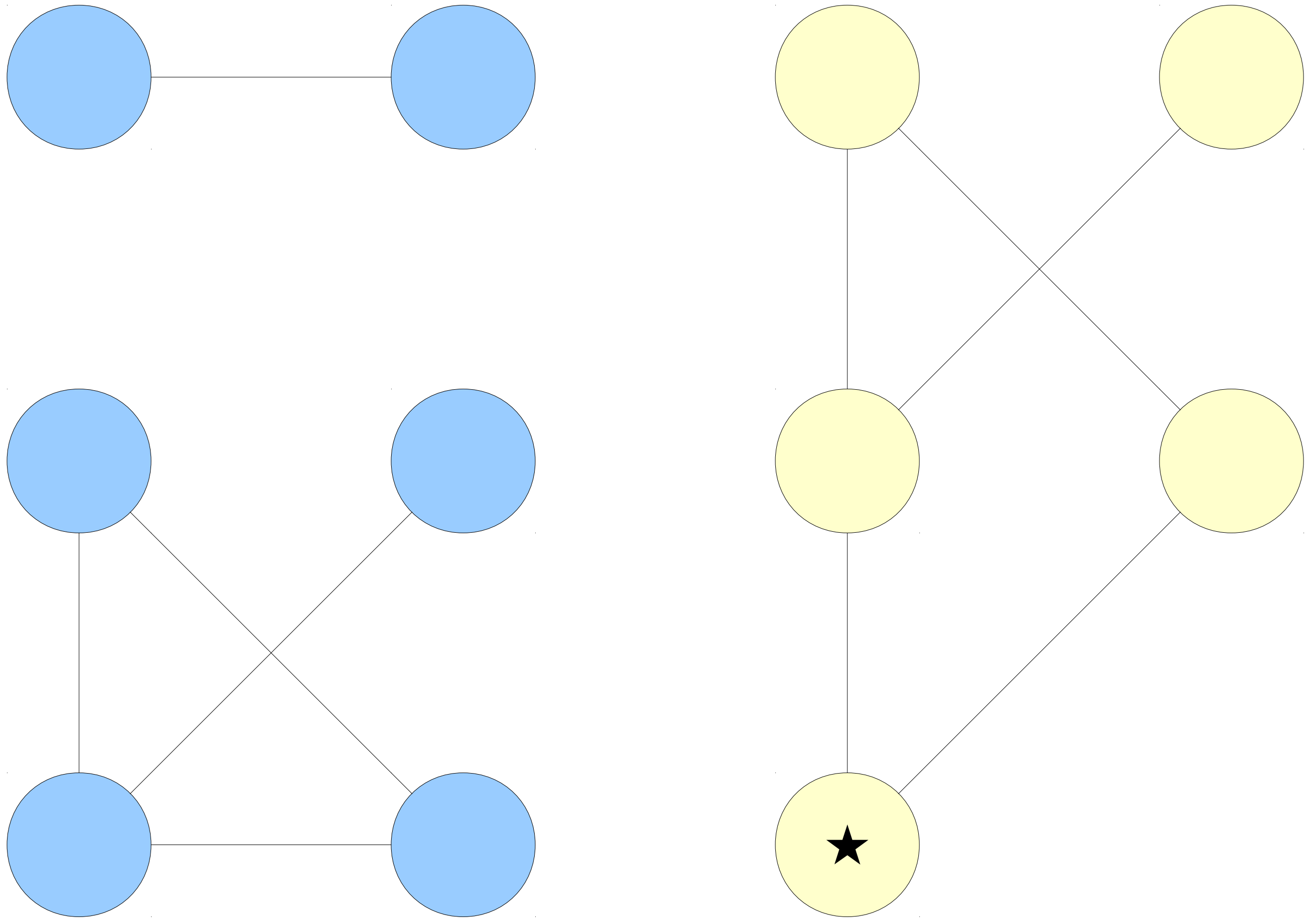
A graph G as a whole is called ***connected*** if all pairs of nodes in G are connected.

(This graph is not connected.)

Connected Components







Connected Components

- Let $G = (V, E)$ be a graph. For each $v \in V$, the ***connected component*** containing v is the set

$$[v] = \{\ x \in V \mid v \text{ is connected to } x\ \}$$

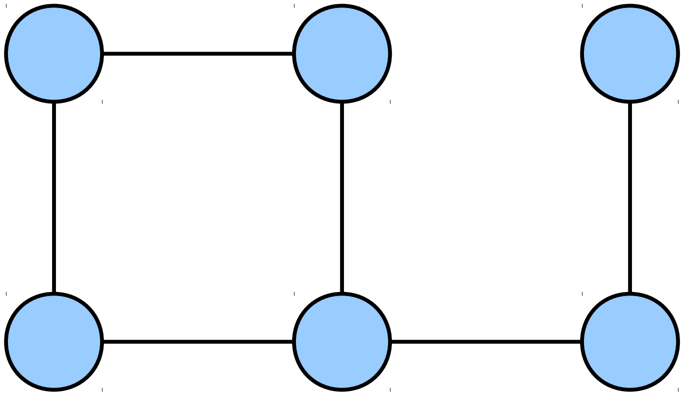
- Intuitively, a connected component is a "piece" of a graph in the sense we just talked about.
- ***Question:*** How do we know that this particular definition of a "piece" of a graph is a good one?
- ***Goal:*** Prove that any graph can be broken apart into different connected components.

We’re trying to reason about some way of partitioning the nodes in a graph into different groups.

What structure have we studied that captures the idea of a partition?

Connectivity

- ***Claim:*** For any graph G, the "is connected to" relation is an equivalence relation.
  - Is it reflexive?
  - Is it symmetric?
  - Is it transitive?



Connectivity

Claim: For any graph G, the "is connected to" relation is an equivalence relation.

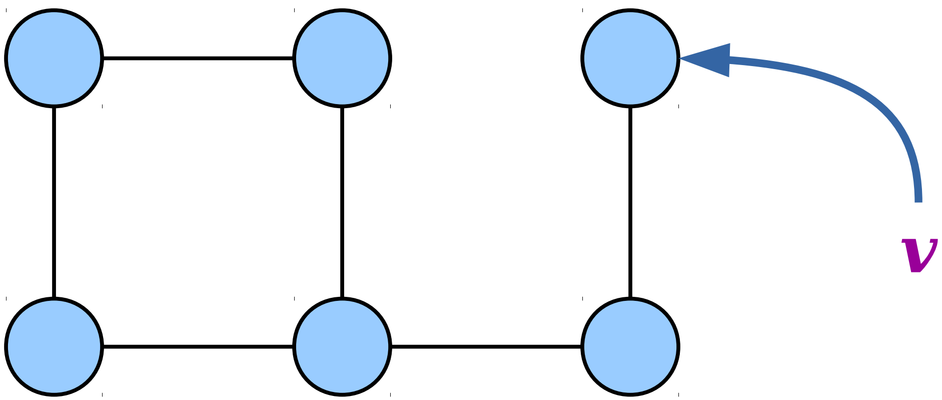
- Is it reflexive?

  Is it symmetric?

  Is it transitive?

$$\forall v \in V.\ Conn(v, v)$$

A ***path*** in a graph $G = (V, E)$ is a sequence of one or more nodes $v_1, v_2, v_3, \ldots, v_n$ such that any two consecutive nodes in the sequence are adjacent.



Connectivity

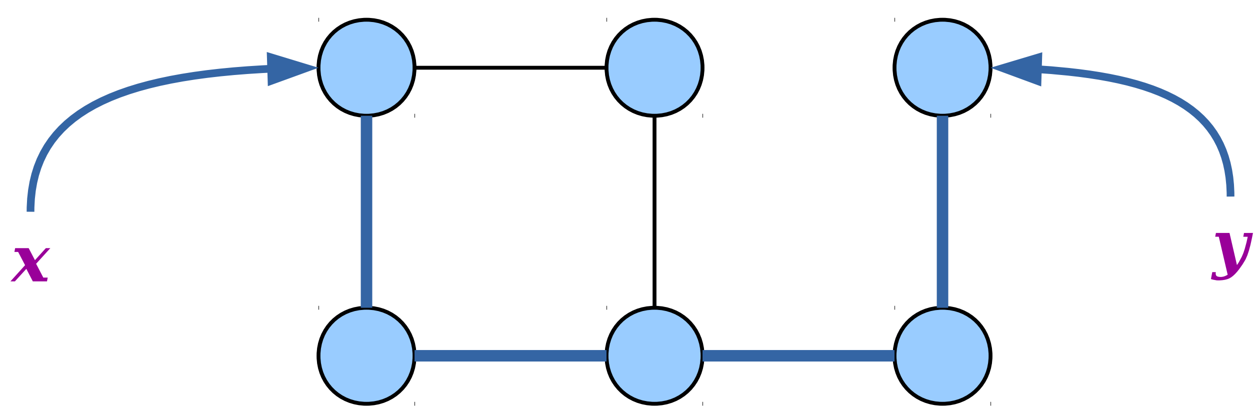
Claim: E
connecte
relation.

$\forall x \in V.\ \forall y \in V.\ (Conn(x, y) \to Conn(y, x))$

Is it reflexive?

- Is it symmetric?

Is it transitive?



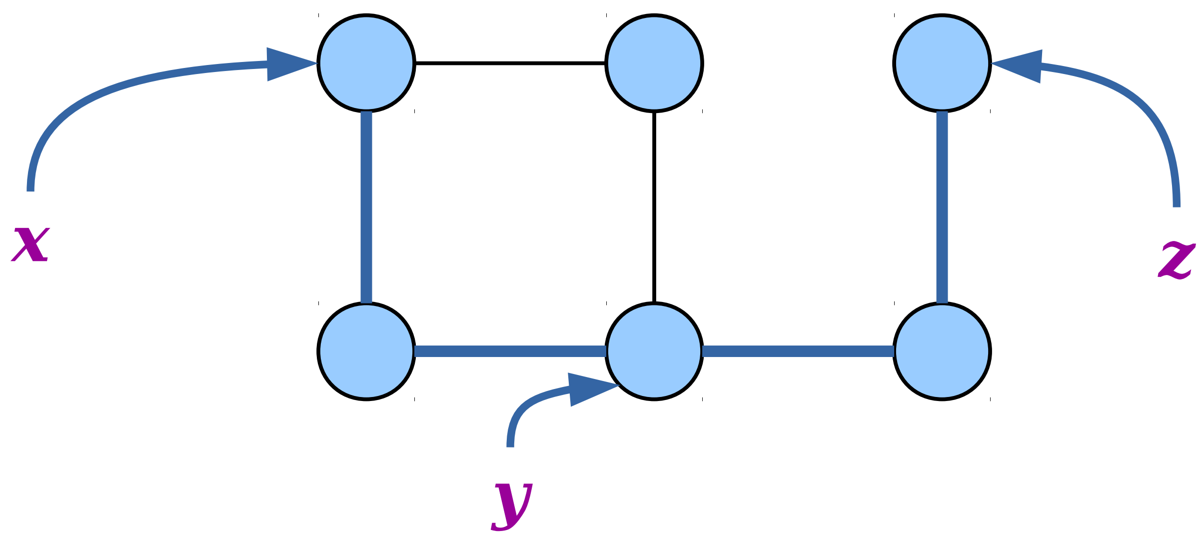
Connectivity

$$\forall x \in V.\ \forall y \in V.\ \forall z \in V.\ (Conn(x, y) \land Conn(y, z) \rightarrow Conn(x, z))$$

Is it reflexive?

Is it symmetric?

- Is it transitive?



Theorem: Let $G = (V, E)$ be a graph. Then the connectivity relation over V is an equivalence relation.

Proof: Consider an arbitrary graph $G = (V, E)$. We will prove that the connectivity relation over V is reflexive, symmetric, and transitive.

To show that connectivity is reflexive, consider any $v \in V$. Then the singleton path v is a path from v to itself. Therefore, v is connected to itself, as required.

To show that connectivity is symmetric, consider any $x, y \in V$ where x is connected to y. We need to show that y is connected to x. Since x is connected to y, there is some path $x, v_1, \ldots, v_n, y$ from x to y. Then $y, v_n, \ldots, v_1, x$ is a path from y back to x, so y is connected to x.

Finally, to show that connectivity is transitive, let $x, y, z \in V$ be arbitrary nodes where x is connected to y and y is connected to z. We will prove that x is connected to z. Since x is connected to y, there is a path $x, u_1, \ldots, u_n, y$ from x to y. Since y is connected to z, there is a path $y, v_1, \ldots, v_k, z$ from y to z. Then the path $x, u_1, \ldots, u_n, y, v_1, \ldots, v_k, z$ goes from x to z. Thus x is connected to z, as required. ■

Putting Things Together

- Earlier, we defined the connected component of a node v to be

$$[v] = \{ x \in V \mid v \text{ is connected to } x \}$$

- Connectivity is an equivalence relation! So what's the equivalence class of a node v with respect to connectivity?

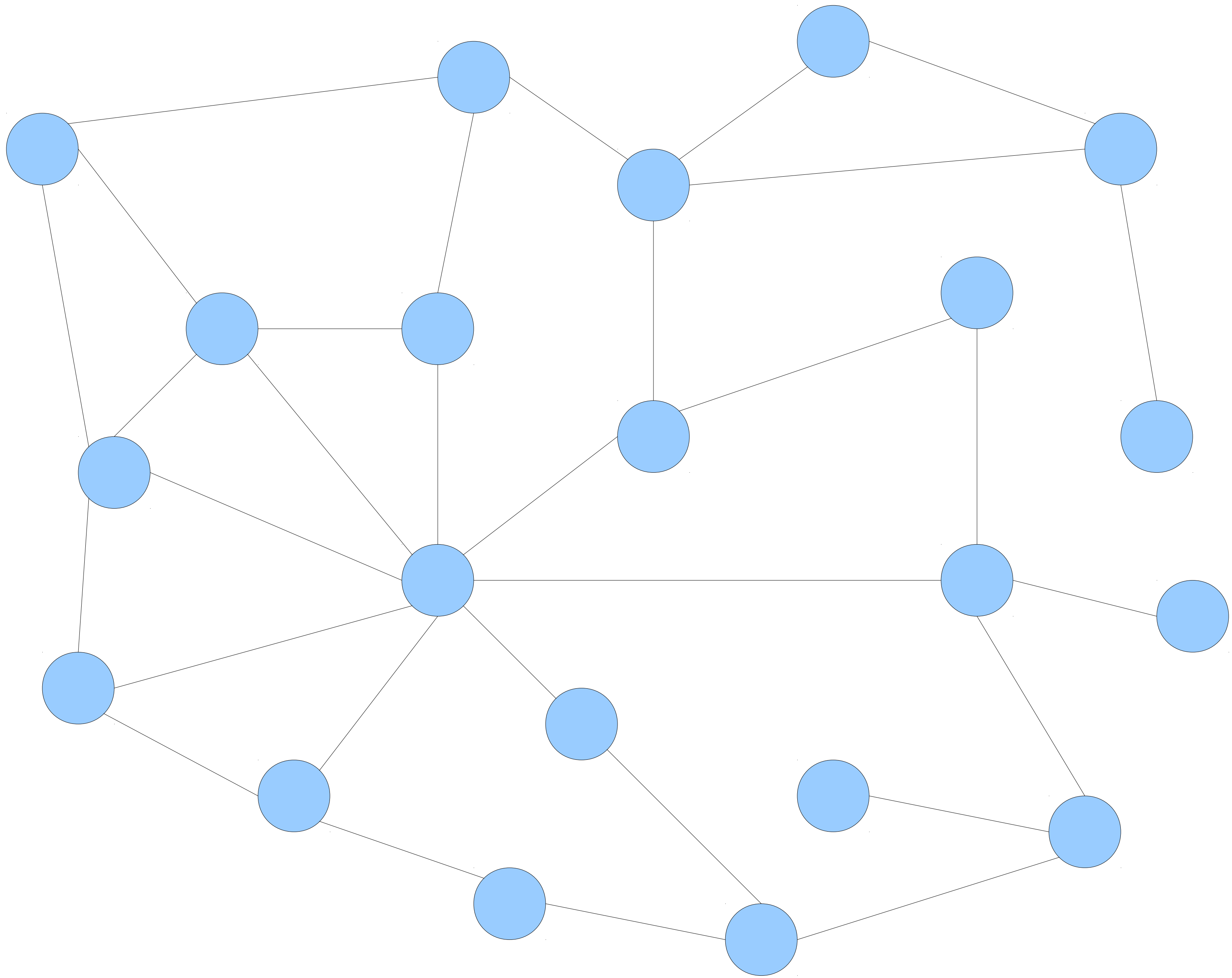
$$[v] = \{ x \in V \mid v \text{ is connected to } x \}$$

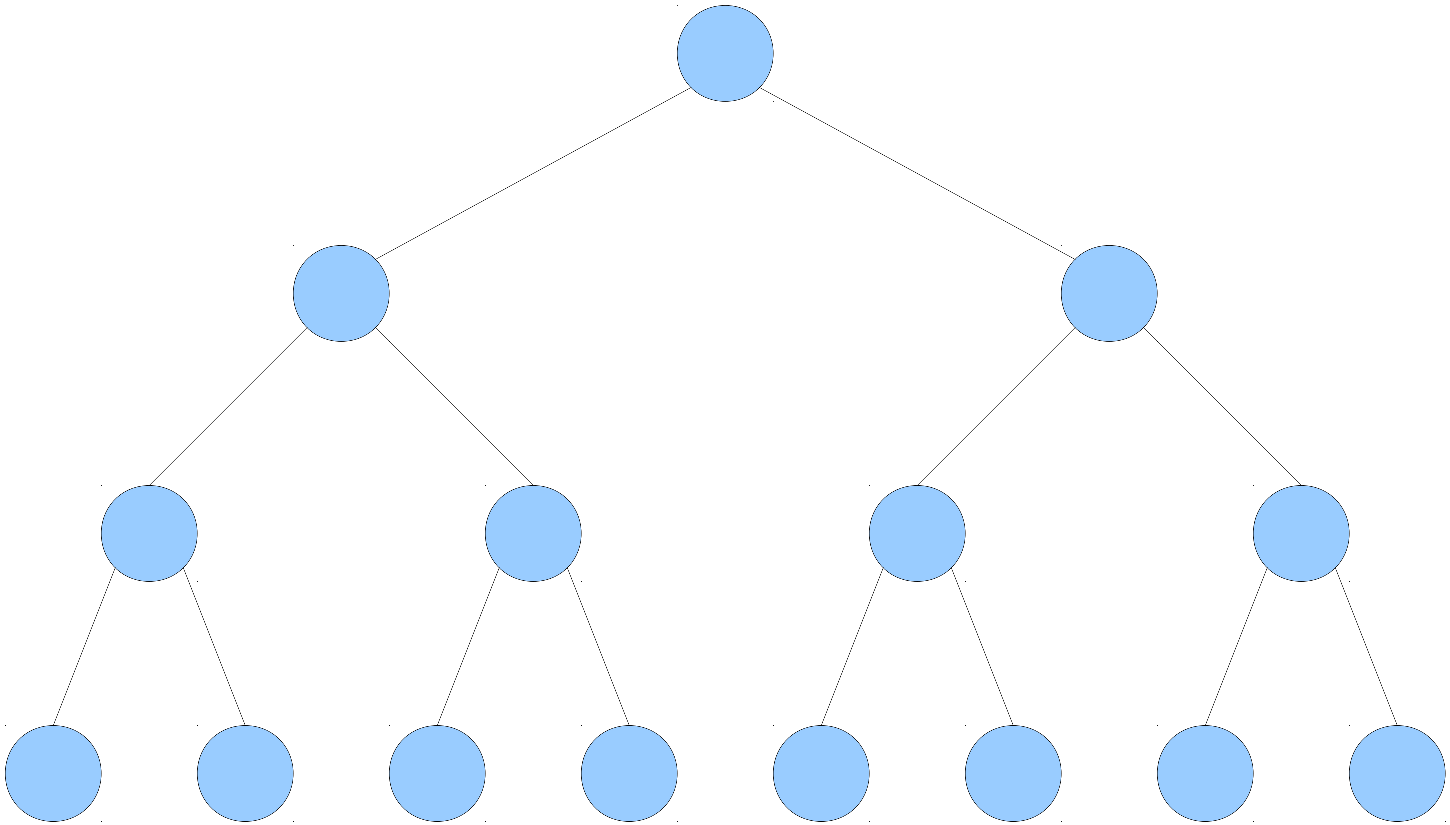
- ***Connected components are equivalence classes of the connectivity relation!***

Theorem: If $G = (V, E)$ is a graph, then every node in G belongs to exactly one connected component of G.

Proof: Let $G = (V, E)$ be an arbitrary graph and let $v \in V$ be any node in G. The connected components of G are just the equivalence classes of the connectivity relation in G. The Fundamental Theorem of Equivalence Relations guarantees that v belongs to exactly one equivalence class of the connectivity relation. Therefore, v belongs to exactly one connected component in G. ■

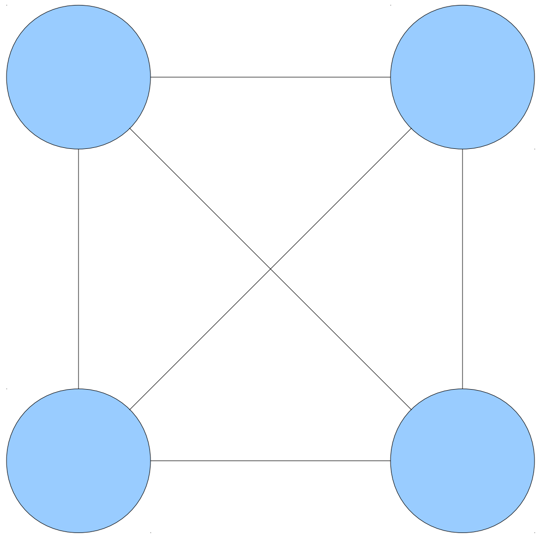
Planar Graphs





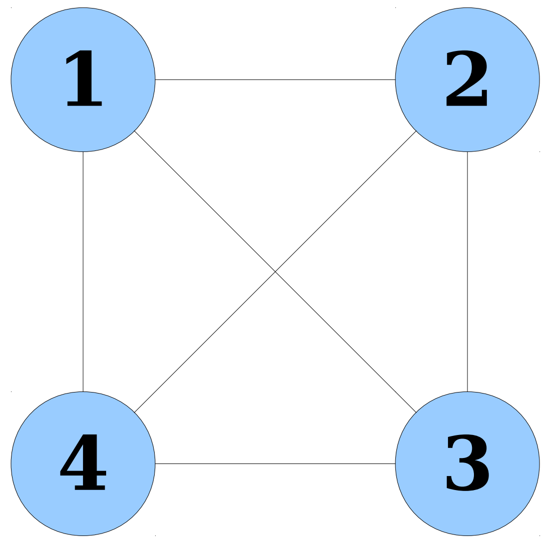
A graph is called a ***planar graph*** if there is some way to draw it in a 2D plane without any of the edges crossing.

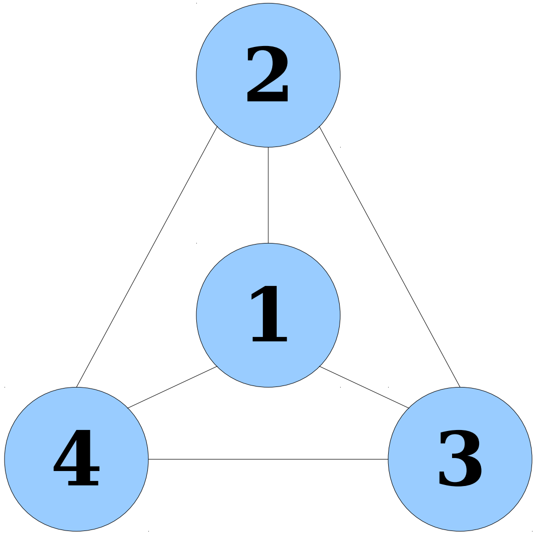
A graph is called a ***planar graph*** if there is some way to draw it in a 2D plane without any of the edges crossing.



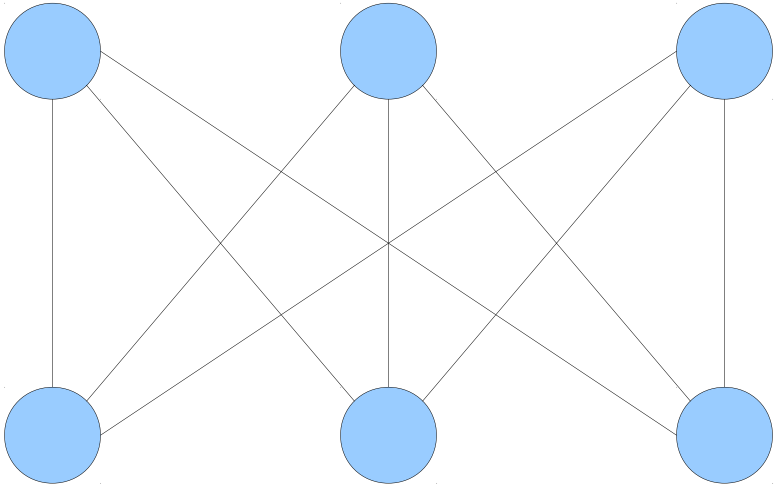
Is this graph planar?

Answer at **PollEv.com/cs103** or
text **CS103** to **22333** once to join, then **Y** or **N**.


1
2
4
3



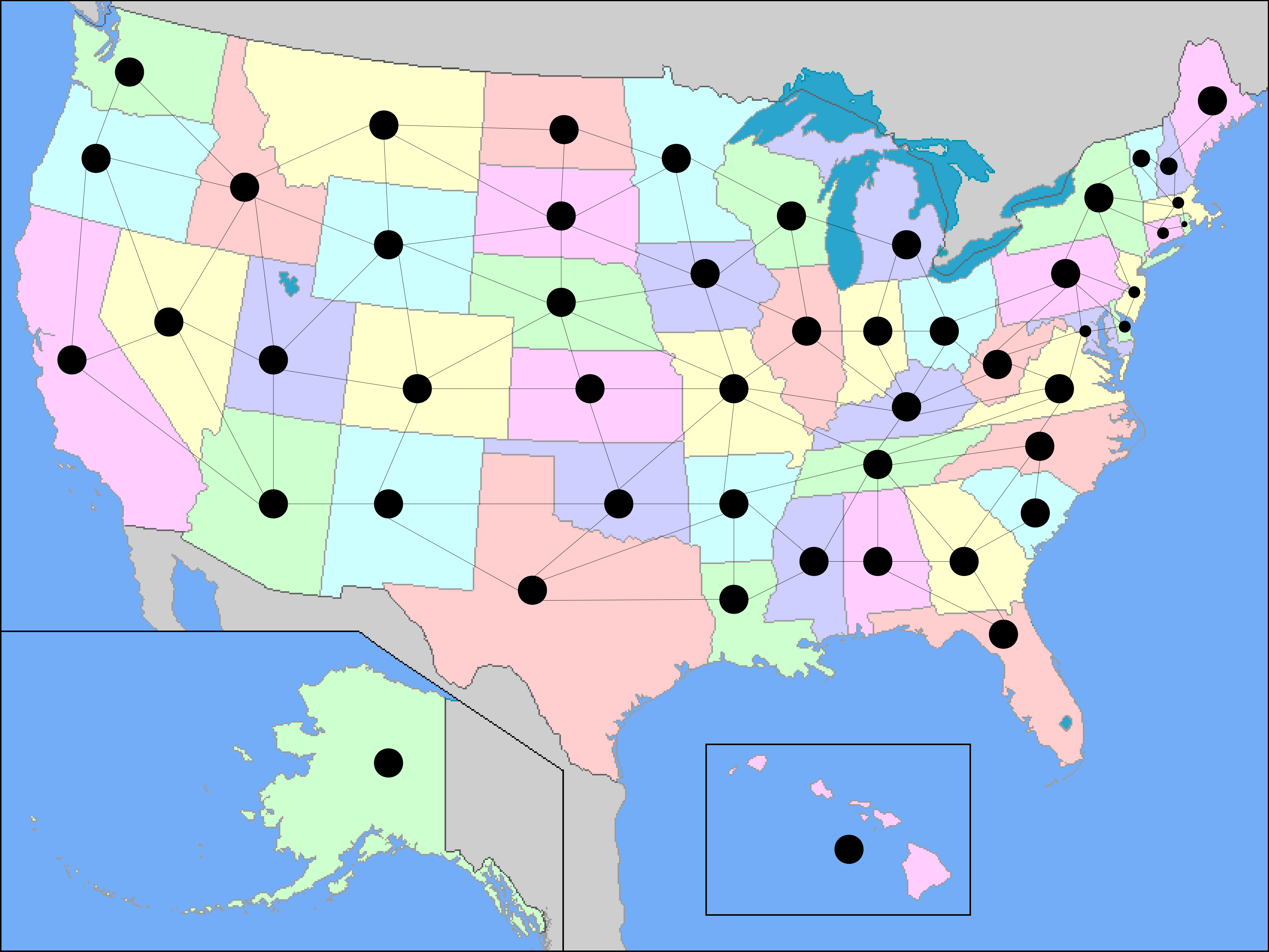
2
1
4
3

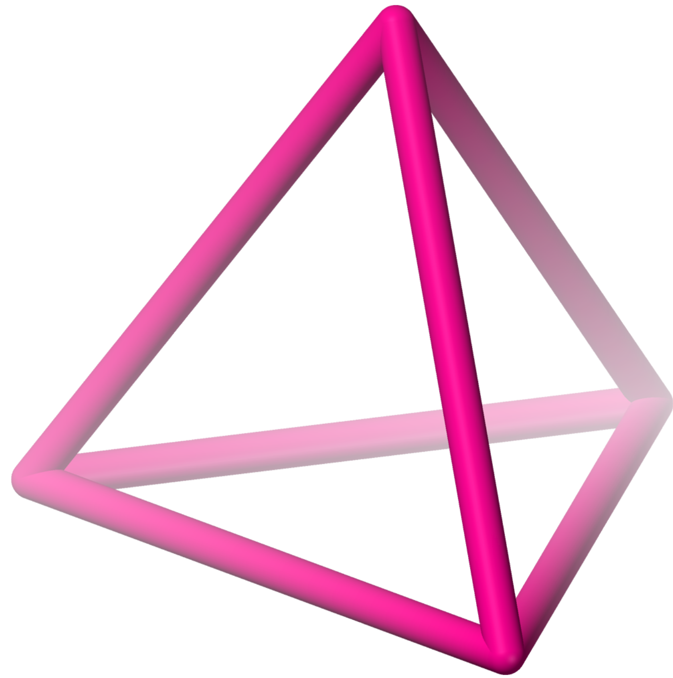


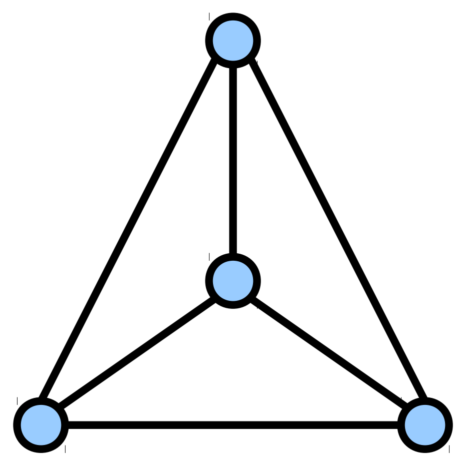
This graph is called the ***utility graph***. There is no way to draw it in the plane without edges crossing. Check out ***this video*** for an explanation!

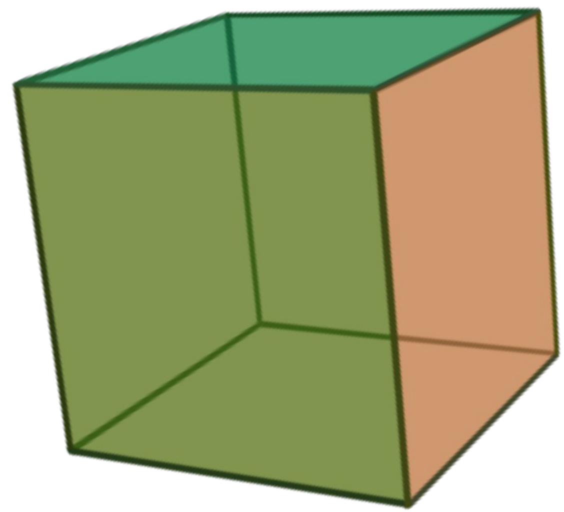
A fun game by a former CS103er:

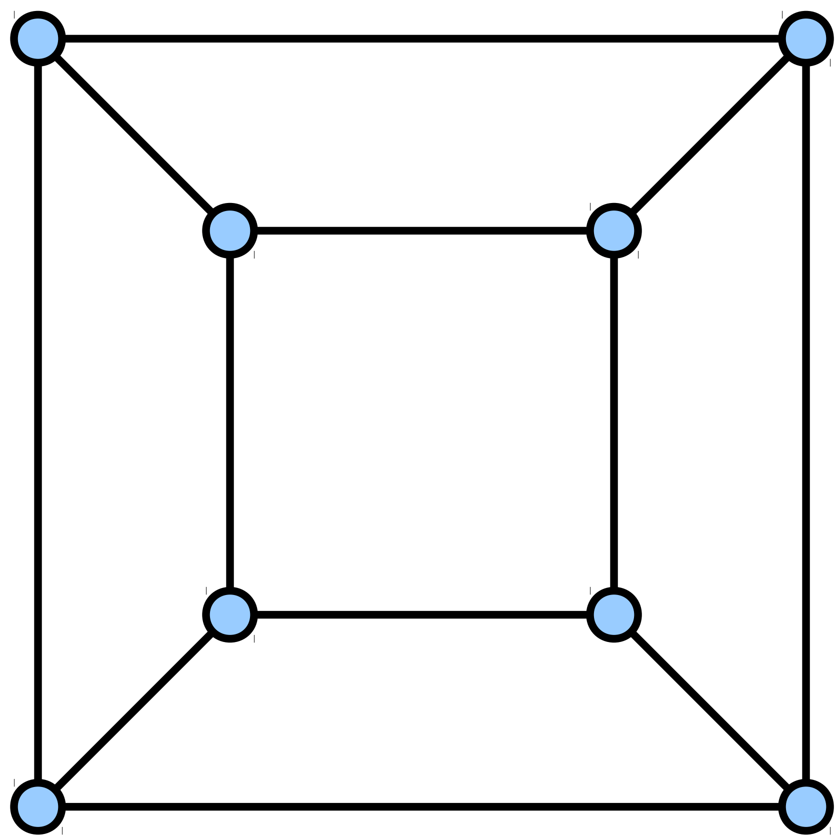
http://www.nkhem.com/planarity-knot/

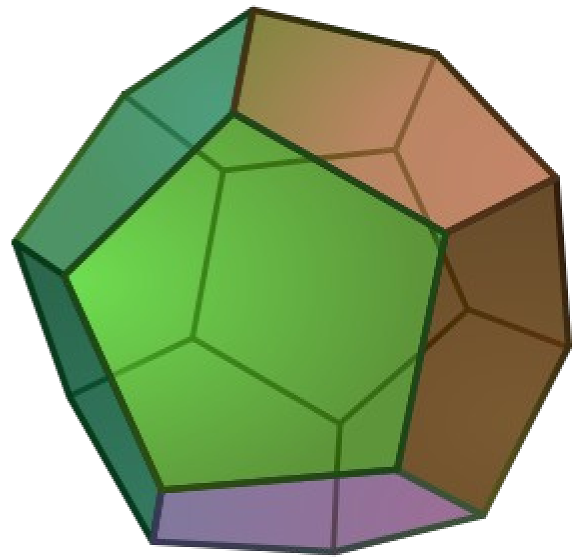


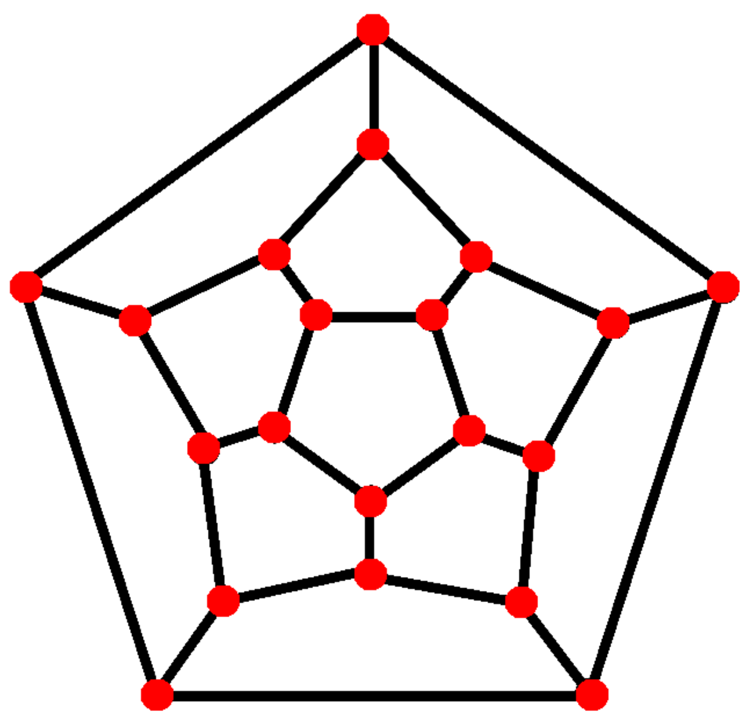




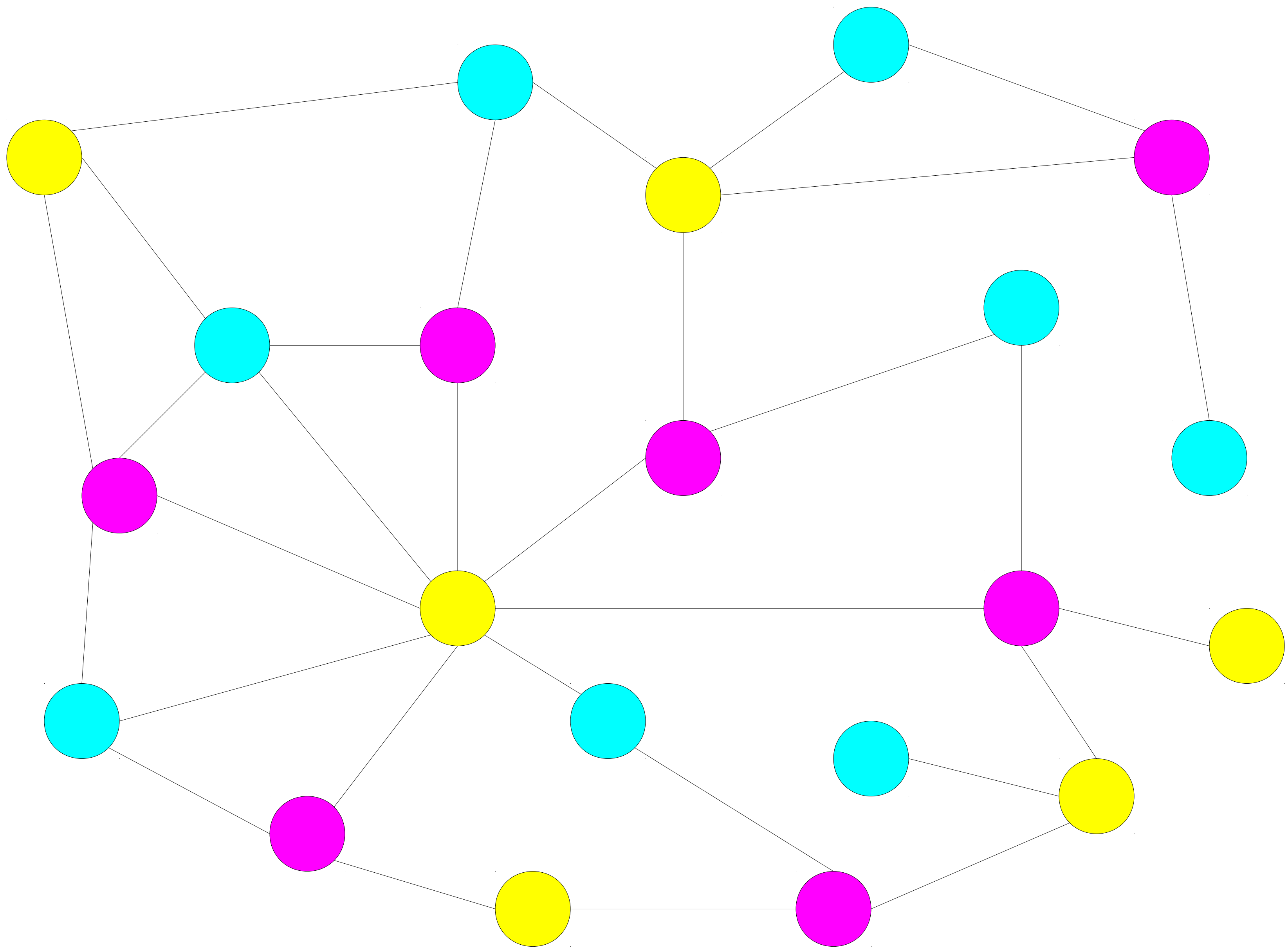


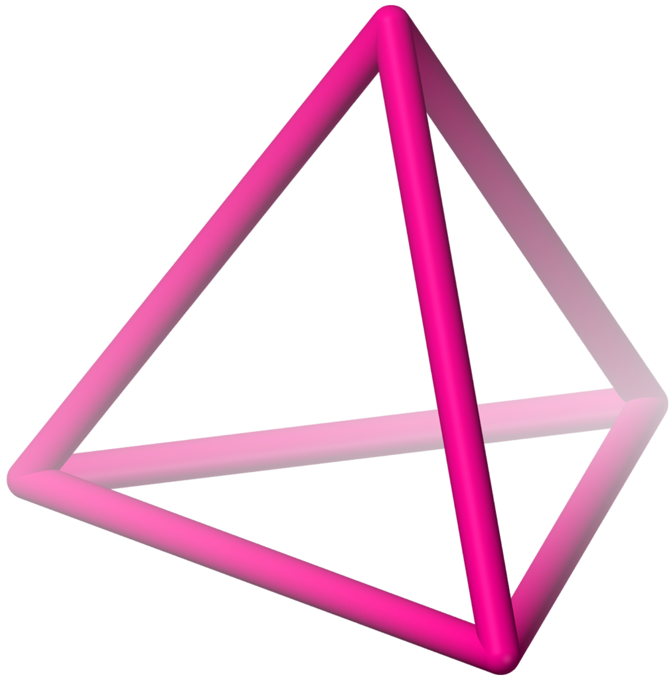


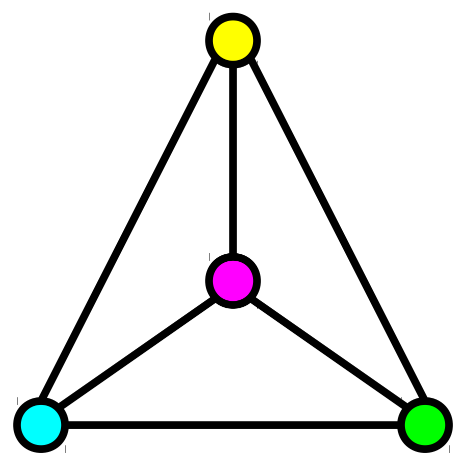


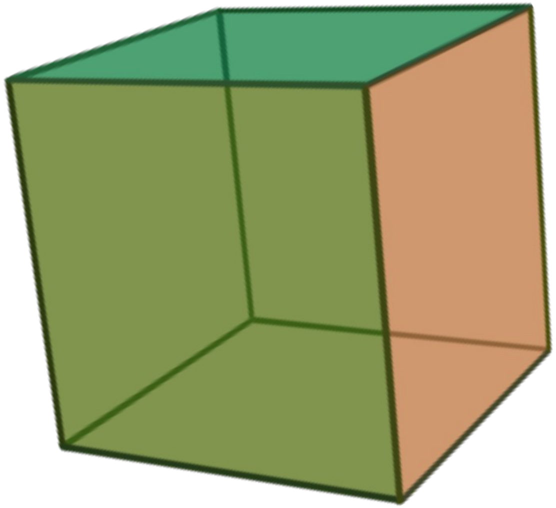


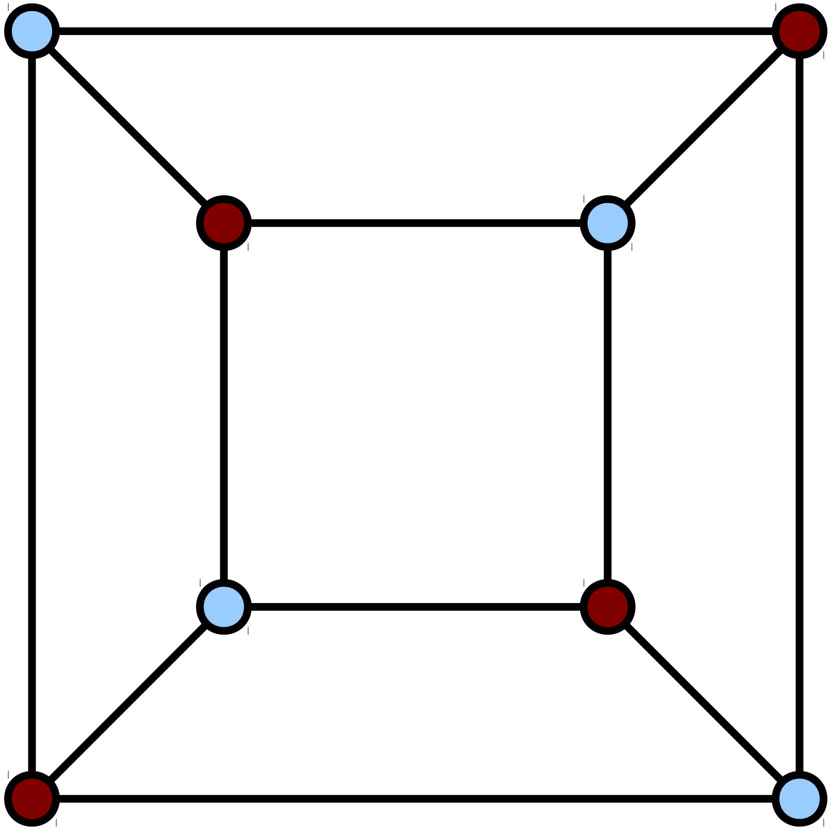
Graph Coloring











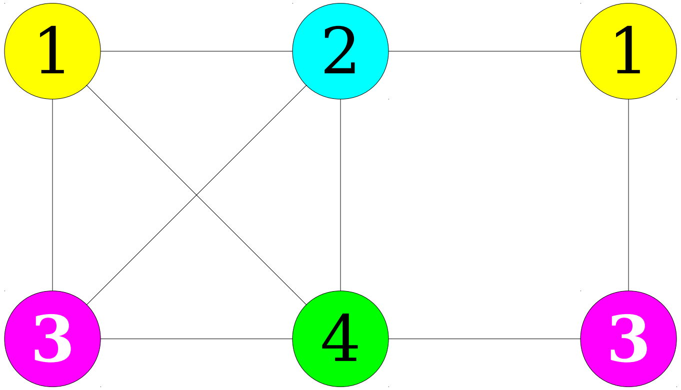
Graph Coloring

- Intuitively, a ***k-vertex-coloring*** of a graph $G = (V, E)$ is a way to color each node in V one of k different colors such that no two adjacent nodes in V are the same color.
- A ***k-vertex-coloring*** of a graph $G = (V, E)$ is a function

$$f : V \to \{1, 2, \ldots, k\}$$

  such that

$$\forall u \in V.\ \forall v \in V.\ (\{u, v\} \in E \to f(u) \neq f(v))$$



Graph Coloring

- Intuitively, a ***k-vertex-coloring*** of a graph $G = (V, E)$ is a way to color each node in V one of k different colors such that no two adjacent nodes in V are the same color.
- A ***k-vertex-coloring*** of a graph $G = (V, E)$ is a function

$$f : V \to \{1, 2, \dots, k\}$$

  such that

$$\forall u \in V.\ \forall v \in V.\ (\{u, v\} \in E \to f(u) \neq f(v))$$

Although this is the formal definition of a k-vertex-coloring, you rarely see it used in proofs. It's more common to just talk about assigning colors to nodes. However, this definition is super useful if you want to write programs to reason about graph colorings!

Graph Coloring

- Intuitively, a ***k*-vertex-coloring** of a graph $G = (V, E)$ is a way to color each node in V one of k different colors such that no two adjacent nodes in V are the same color.
- A ***k*-vertex-coloring** of a graph $G = (V, E)$ is a function

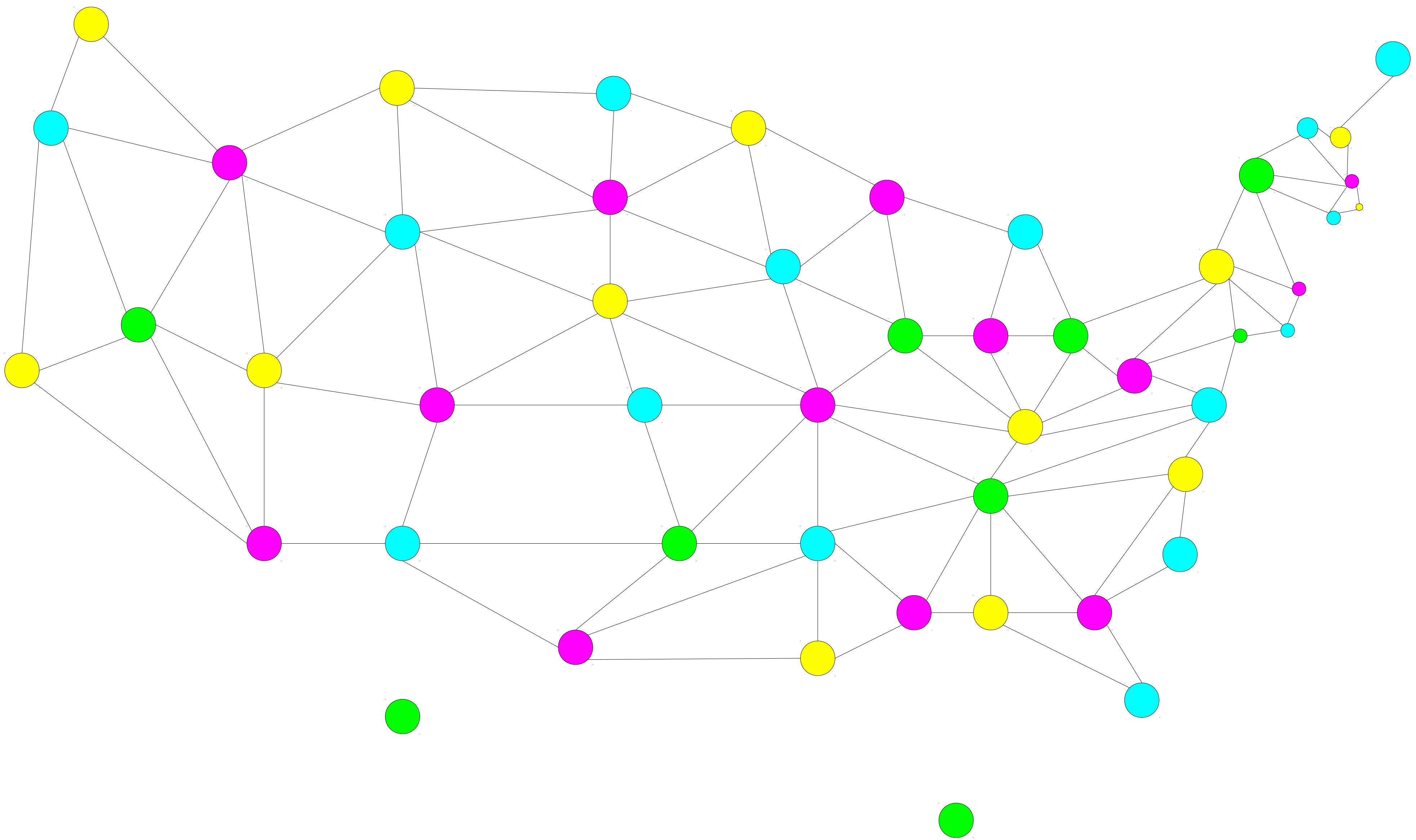
$$f : V \to \{1, 2, \ldots, k\}$$

  such that

$$\forall u \in V.\ \forall v \in V.\ (\{u, v\} \in E \to f(u) \neq f(v))$$

- A graph G is ***k*-colorable** if a k-vertex coloring of G exists.
- The smallest k for which G is k-colorable is its ***chromatic number***.
  - The chromatic number of a graph G is denoted $\chi(G)$, from the Greek *χρώμα*, meaning "color."

Graph Coloring



Theorem (Four-Color Theorem): Every planar graph is 4-colorable.

- **1850s:** Four-Color Conjecture posed.
- **1879:** Kempe proves the Four-Color Theorem.
- **1890:** Heawood finds a flaw in Kempe's proof.
- **1976:** Appel and Haken design a computer program that proves the Four-Color Theorem. The program checked 1,936 specific cases that are “minimal counterexamples;” any counterexample to the theorem must contain one of the 1,936 specific cases.
- **1980s:** Doubts rise about the validity of the proof due to errors in the software.
- **1989:** Appel and Haken revise their proof and show it is indeed correct. They publish a book including a 400-page appendix of all the cases to check.
- **1996:** Roberts, Sanders, Seymour, and Thomas reduce the number of cases to check down to 633.
- **2005:** Werner and Gonthier repeat the proof using an established automatic theorem prover (Coq), improving confidence in the truth of the theorem.

Philosophical Question: Is a theorem true if no human has ever read the proof?

A Fantastic Video on a Cool Theorem:

https://youtu.be/-9OUyo8NFZg

Next Time

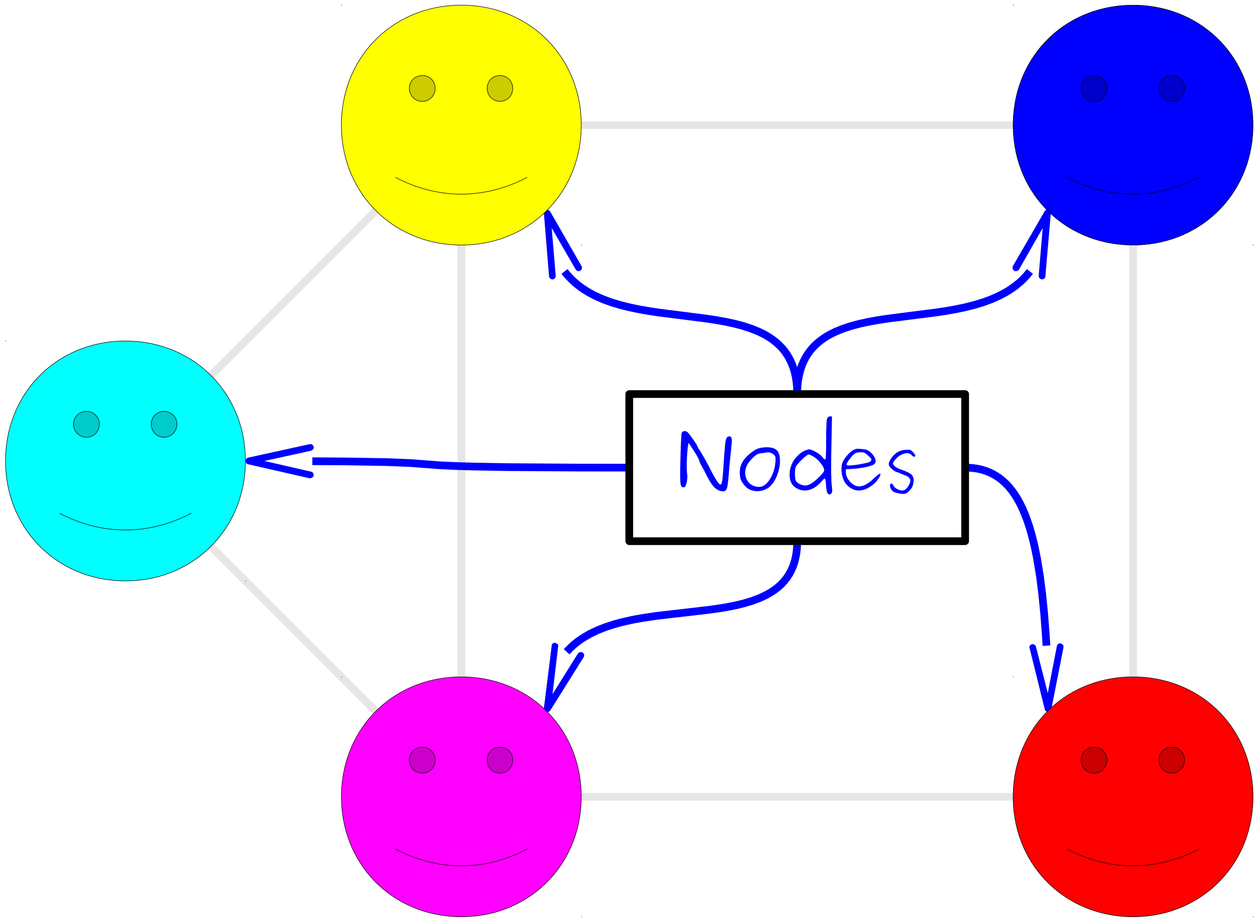
- ***The Pigeonhole Principle***
  - A simple, powerful, versatile theorem.
- ***Graph Theory Party Tricks***
  - Applying math to graphs of people!

Graph Theory

Part Two

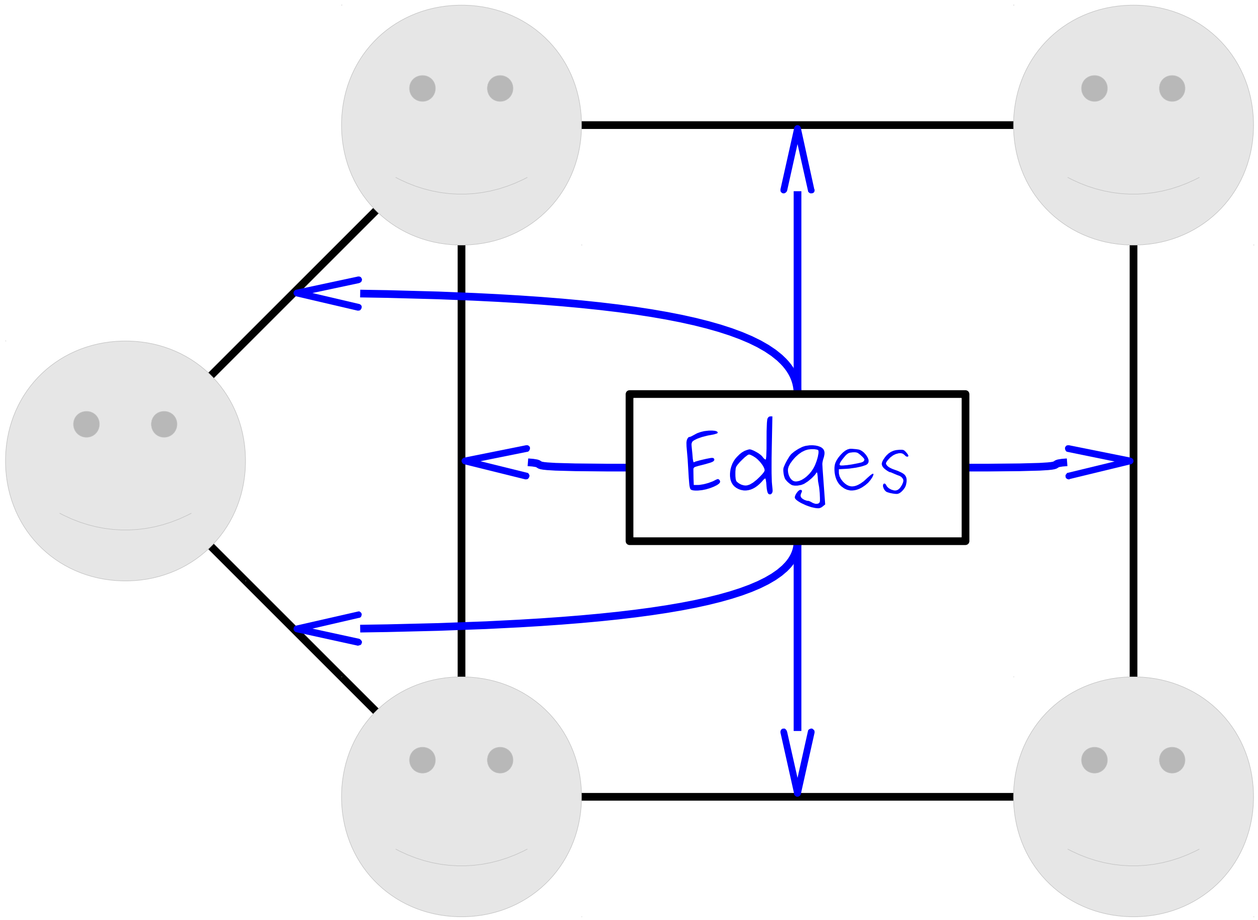
Recap from Last Time

A ***graph*** is a mathematical structure for representing relationships.



A graph consists of a set of ***nodes*** (or ***vertices***) connected by ***edges*** (or ***arcs***)

A ***graph*** is a mathematical structure for representing relationships.



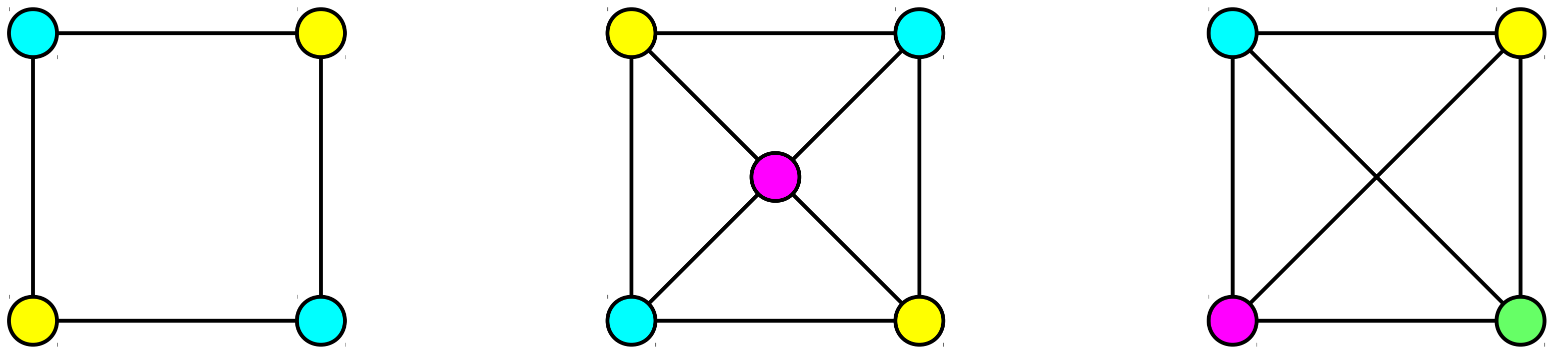
A graph consists of a set of ***nodes*** (or ***vertices***) connected by ***edges*** (or ***arcs***)

Adjacency and Connectivity

- Two nodes in a graph are called ***adjacent*** if there's an edge between them.
- Two nodes in a graph are called ***connected*** if there's a path between them.
  - A path is a series of one or more nodes where consecutive nodes are adjacent.

k-Vertex-Colorings

- If $G = (V, E)$ is a graph, a ***k-vertex-coloring*** of G is a way of assigning colors to the nodes of G, using at most k colors, so that no two nodes of the same color are adjacent.



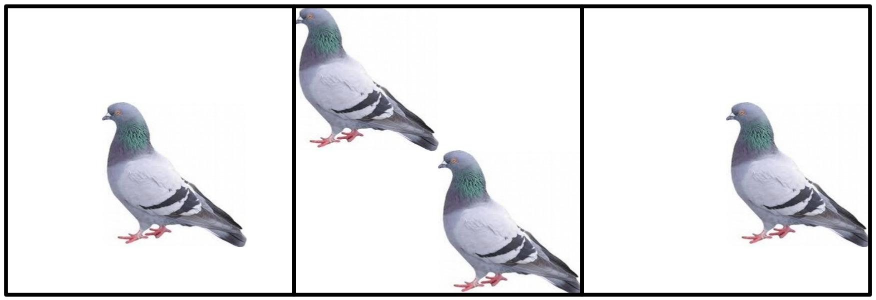
- The ***chromatic number*** of G, denoted $\boldsymbol{\chi(G)}$, is the minimum number of colors needed in any k-coloring of G.
- Today, we're going to see several results involving coloring parts of graphs. They don't necessarily involve *k-vertex-colorings* of graphs, so feel free to ask for clarifications if you need them!

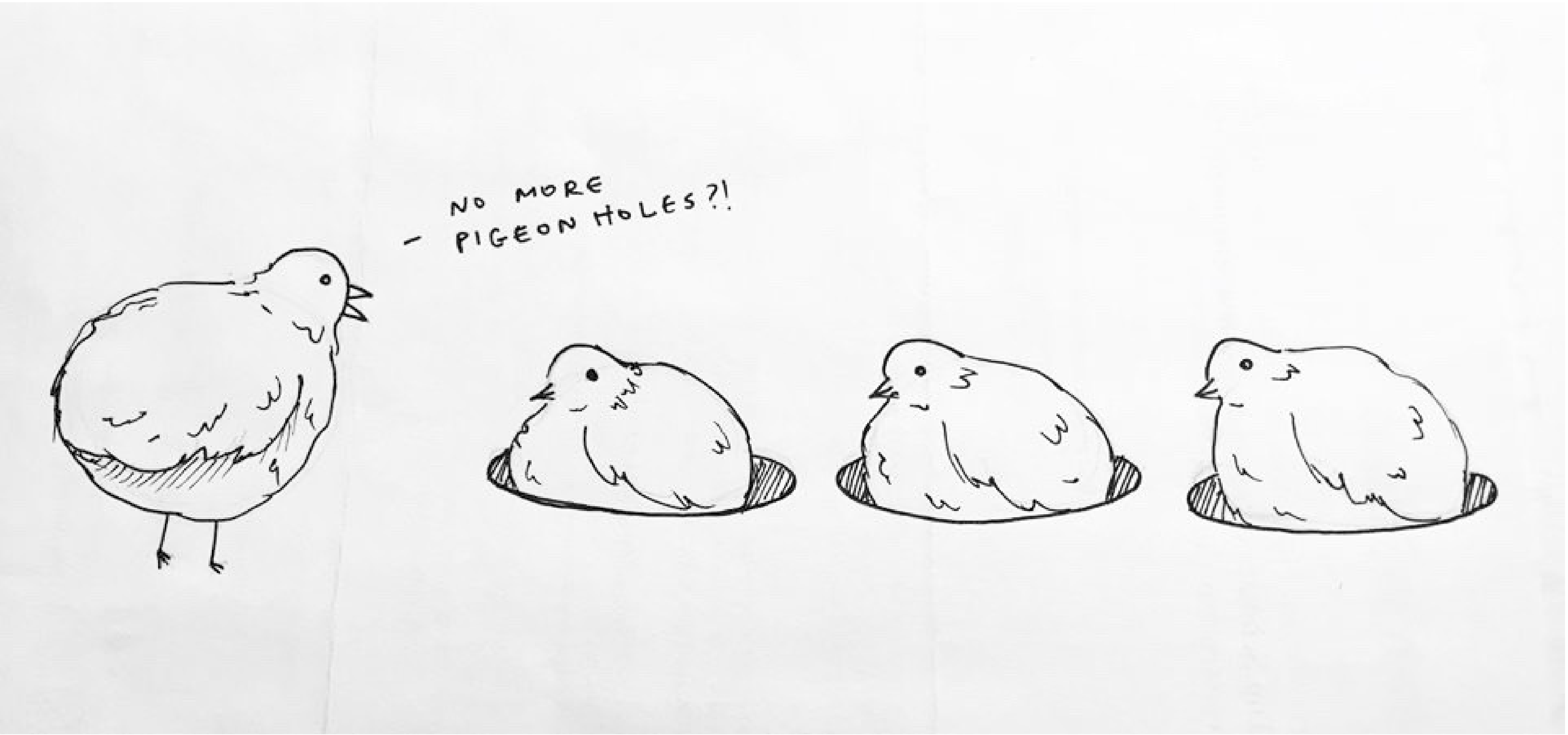
New Stuff!

The Pigeonhole Principle

The Pigeonhole Principle

- ***Theorem (The Pigeonhole Principle):*** If m objects are distributed into n bins and $m > n$, then at least one bin will contain at least two objects.





$m = 4, n = 3$

Thanks to Amy Liu for this awesome drawing!

Some Simple Applications

- Any group of 367 people must have a pair of people that share a birthday.
  - 366 possible birthdays (pigeonholes)
  - 367 people (pigeons)
- Two people in San Francisco have the exact same number of hairs on their head.
  - Maximum number of hairs ever found on a human head is no greater than 500,000.
  - There are over 800,000 people in San Francisco.

Theorem (The Pigeonhole Principle): If m objects are distributed into n bins and $m > n$, then at least one bin will contain at least two objects.

Let A and B be finite sets (sets whose cardinalities are natural numbers) and assume $|A| > |B|$. How many of the following statements are true?

(W) If $f : A \to B$, then f is injective.
(X) If $f : A \to B$, then f is not injective.
(Y) If $f : A \to B$, then f is surjective.
(Z) If $f : A \to B$, then f is not surjective.

Answer at **PollEv.com/cs103** or
text **CS103** to **22333** once to join, then a number.

Theorem: If m objects are distributed into n bins and $m > n$, then there must be some bin that contains at least two objects.

Proof: Suppose for the sake of contradiction that, for some m and n where $m > n$, there is a way to distribute m objects into n bins such that each bin contains at most one object.

Number the bins 1, 2, 3, …, n and let x_i denote the number of objects in bin i. There are m objects in total, so we know that

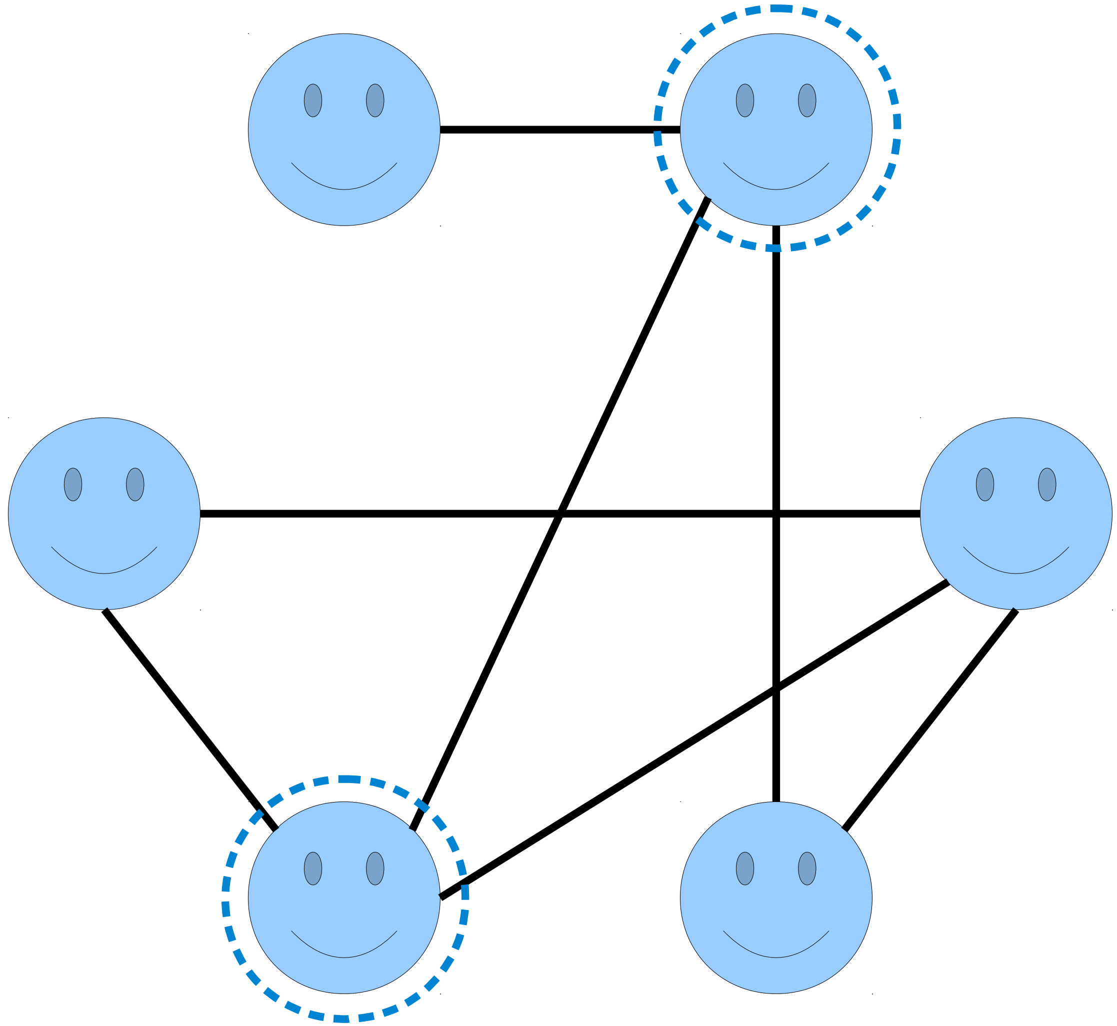
$$m = x_1 + x_2 + \ldots + x_n.$$

Since each bin has at most one object in it, we know $x_i \le 1$ for each i. This means that

$$\begin{aligned} m &= x_1 + x_2 + \ldots + x_n \\ &\le 1 + 1 + \ldots + 1 \quad \textit{(n times)} \\ &= n. \end{aligned}$$

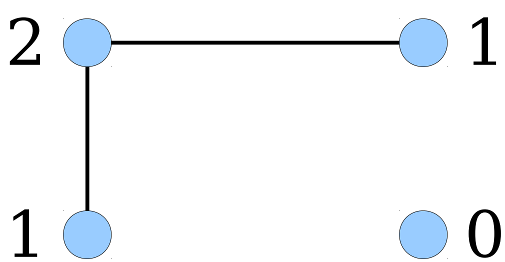
This means that $m \le n$, contradicting that $m > n$. We've reached a contradiction, so our assumption must have been wrong. Therefore, if m objects are distributed into n bins with $m > n$, some bin must contain at least two objects. ■

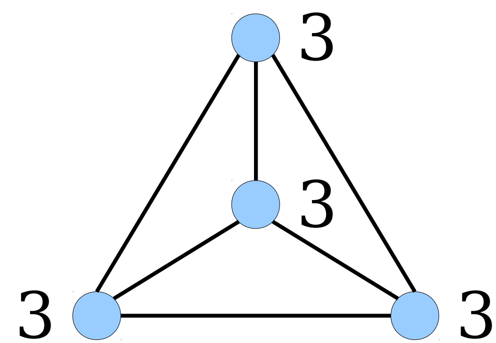
Pigeonhole Principle Party Tricks



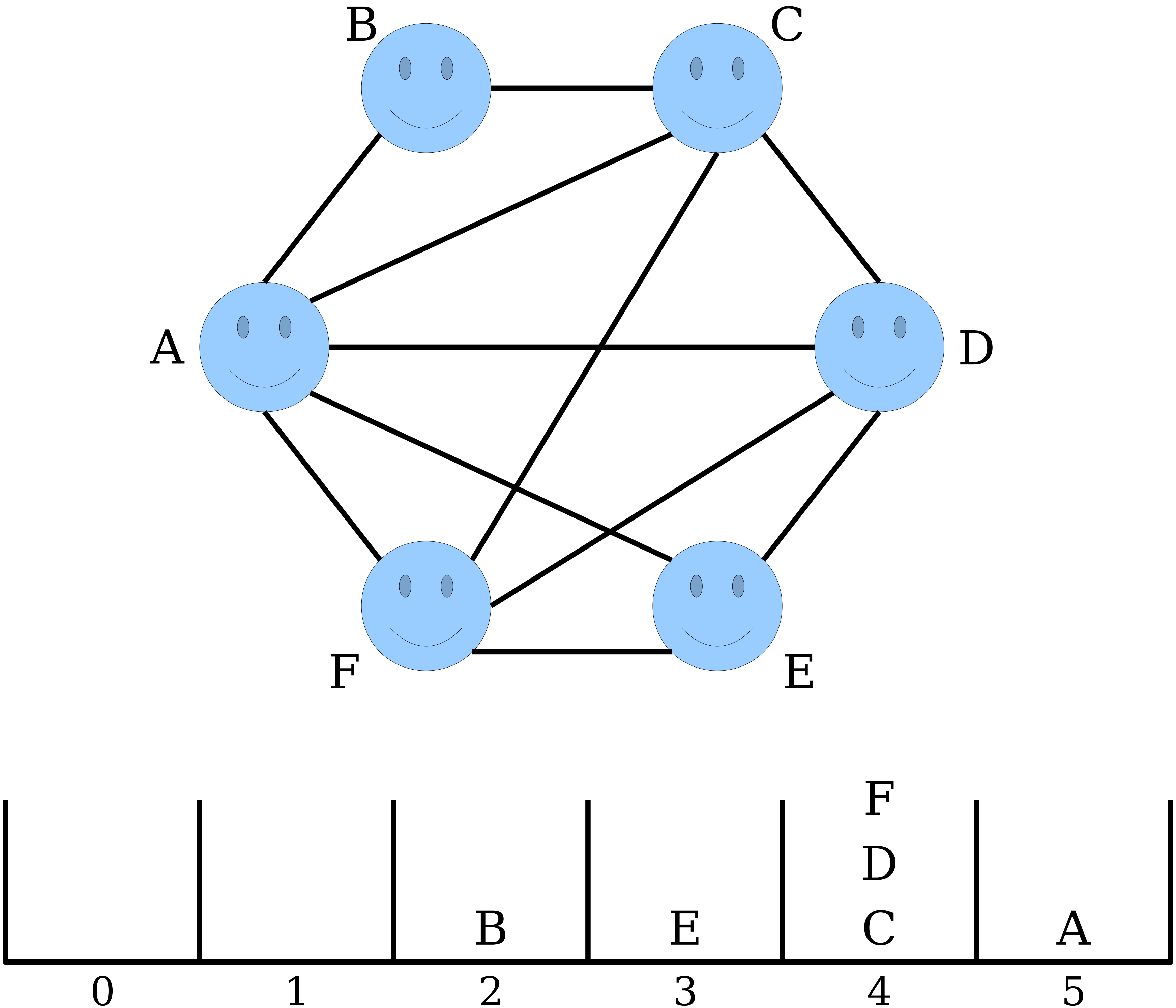
Degrees

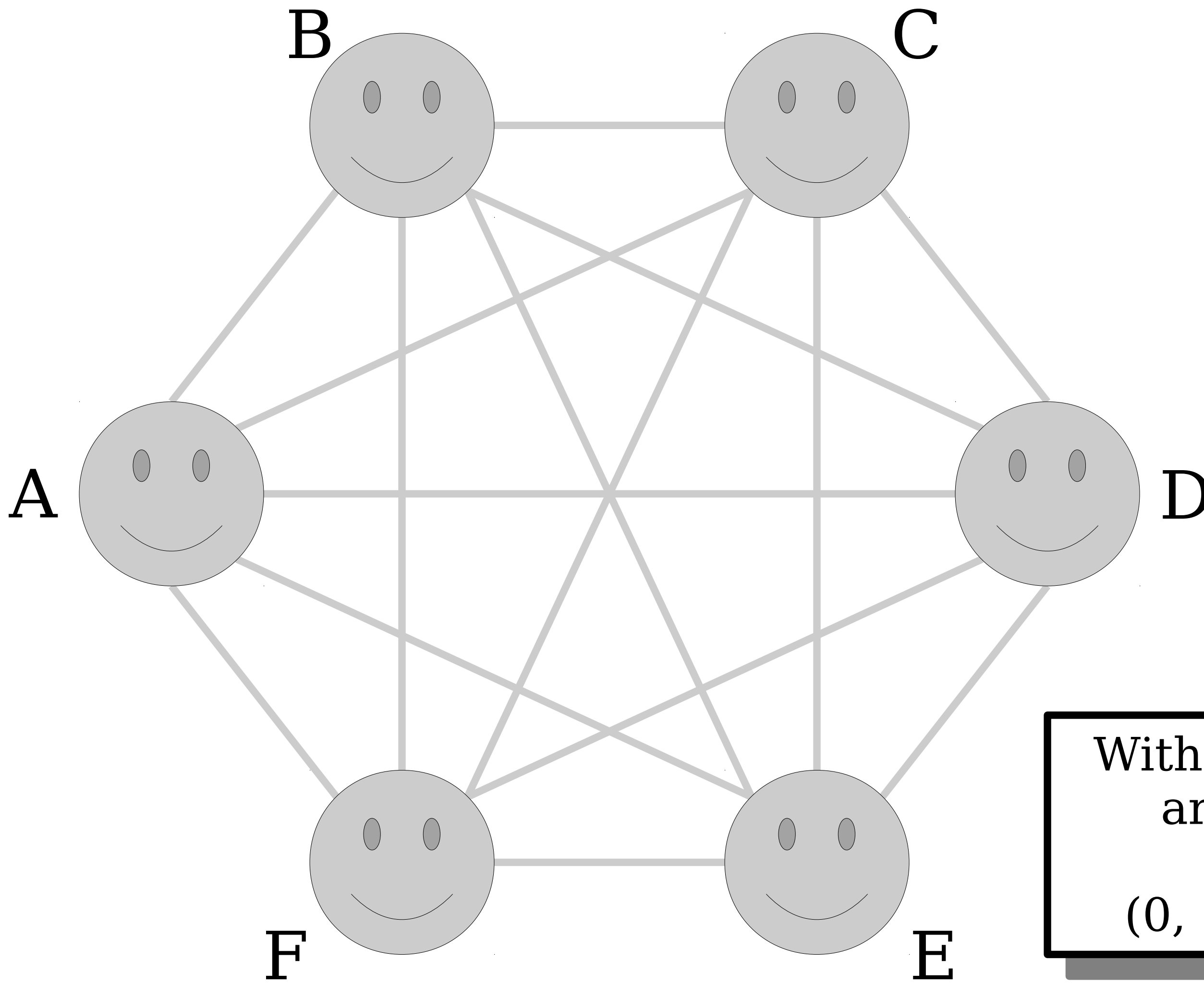
- The ***degree*** of a node *v* in a graph is the number of nodes that *v* is adjacent to.



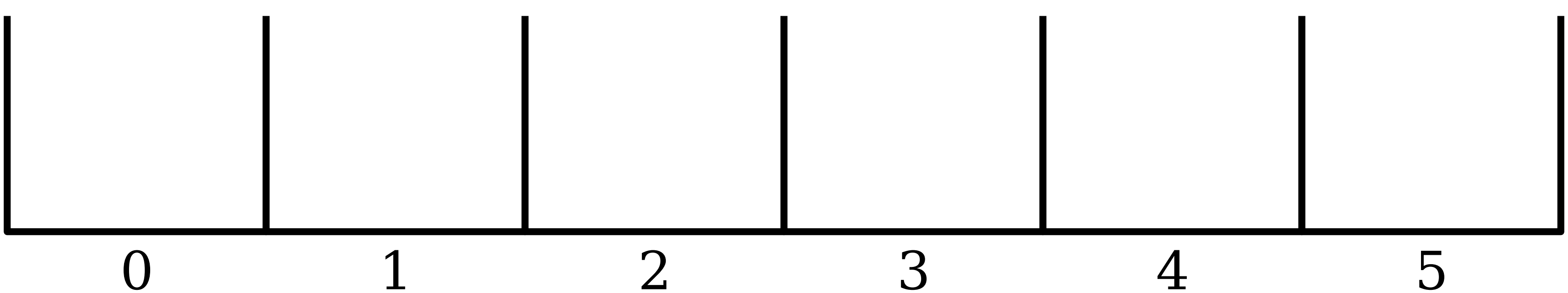


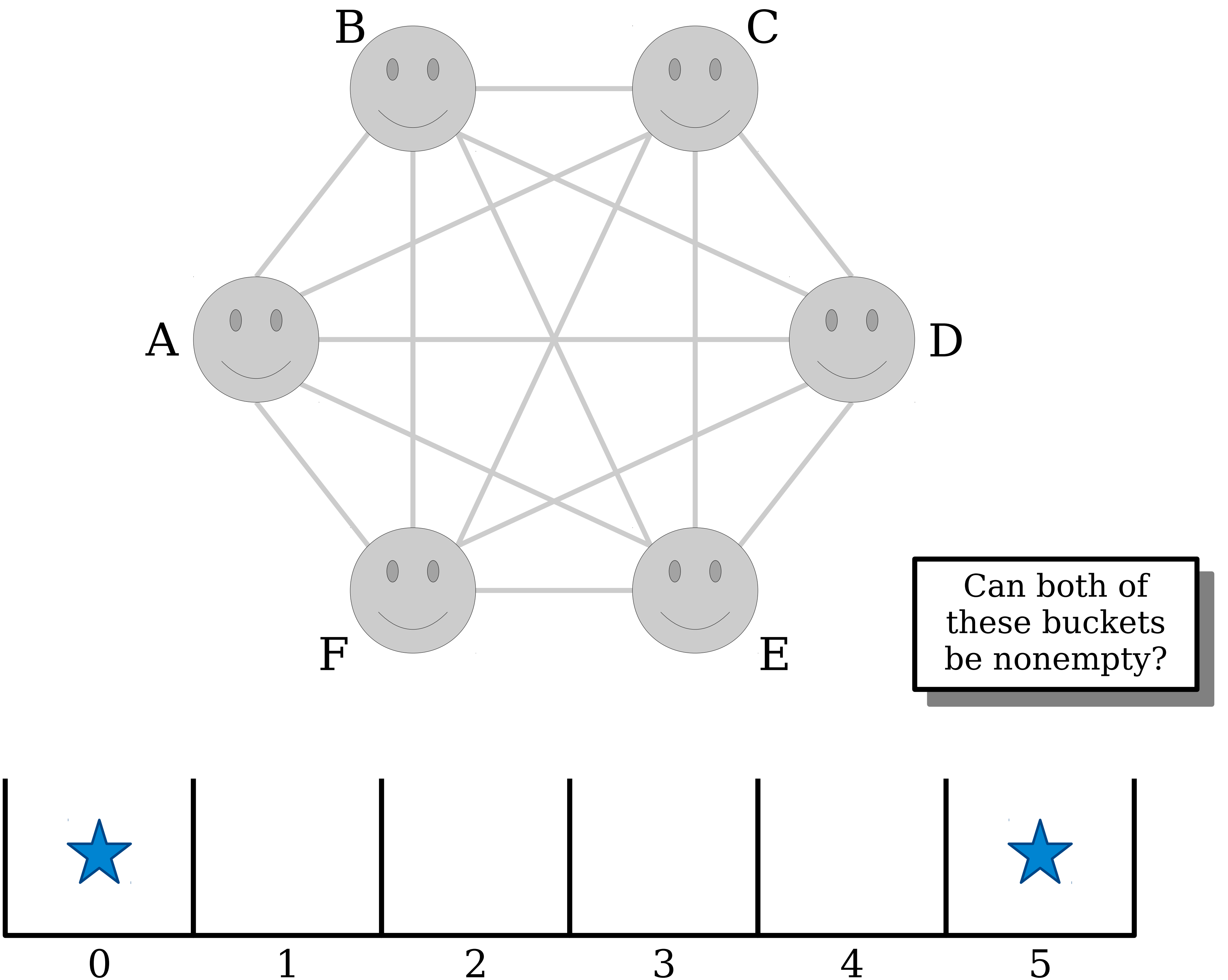
- ***Theorem:*** Every graph with at least two nodes has at least two nodes with the same degree.
    - Equivalently: at any party with at least two people, there are at least two people with the same number of Facebook friends at the party.


B
C
A
D
F
E
F
D
B
E
C
A
0
1
2
3
4
5

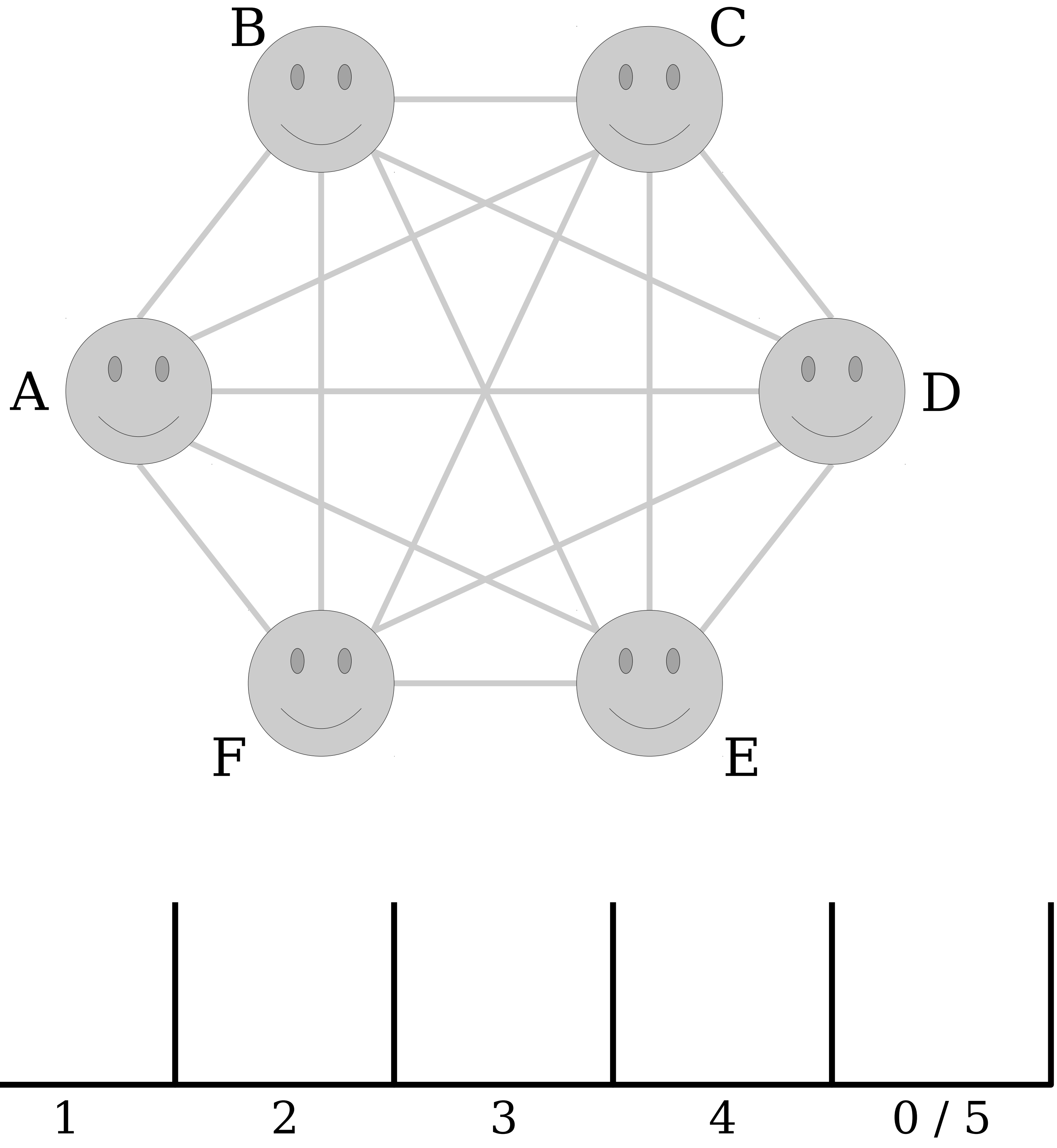


With n nodes, there are n possible degrees ($0, 1, 2, \ldots, n - 1$)





B
C
A
D
F
E
Can both of these buckets be nonempty?
0
1
2
3
4
5



B
C
A
D
F
E
1
2
3
4
0 / 5

Theorem: In any graph with at least two nodes, there are at least two nodes of the same degree.

Proof 1: Let G be a graph with $n \geq 2$ nodes. There are n possible choices for the degrees of nodes in G, namely, 0, 1, 2, …, and $n - 1$.

We claim that G cannot simultaneously have a node u of degree 0 and a node v of degree $n - 1$: if there were such nodes, then node u would be adjacent to no other nodes and node v would be adjacent to all other nodes, including u. (Note that u and v must be different nodes, since v has degree at least 1 and u has degree 0.)

We therefore see that the possible options for degrees of nodes in G are either drawn from 0, 1, …, $n - 2$ or from 1, 2, …, $n - 1$. In either case, there are n nodes and $n - 1$ possible degrees, so by the pigeonhole principle two nodes in G must have the same degree. ■

Theorem: In any graph with at least two nodes, there are at least two nodes of the same degree.

Proof 2: Assume for the sake of contradiction that there is a graph G with $n \geq 2$ nodes where no two nodes have the same degree. There are n possible choices for the degrees of nodes in G, namely $0, 1, 2, \ldots, n - 1$, so this means that G must have exactly one node of each degree. However, this means that G has a node of degree 0 and a node of degree $n - 1$. (These can't be the same node, since $n \geq 2$.) This first node is adjacent to no other nodes, but this second node is adjacent to every other node, which is impossible.

We have reached a contradiction, so our assumption must have been wrong. Thus if G is a graph with at least two nodes, G must have at least two nodes of the same degree. ■

The Generalized Pigeonhole Principle

Suppose 11 objects are distributed into 5 bins.
How many of the following statements are true?

U: The bin with the most objects must contain at least 2 objects.
V: The bin with the most objects must contain at least 3 objects.
W: The bin with the most objects must contain at least 4 objects.
X: The bin with the fewest objects must contain at most 1 object.
Y: The bin with the fewest objects must contain at most 2 objects.
Z: The bin with the fewest objects must contain at most 3 objects.

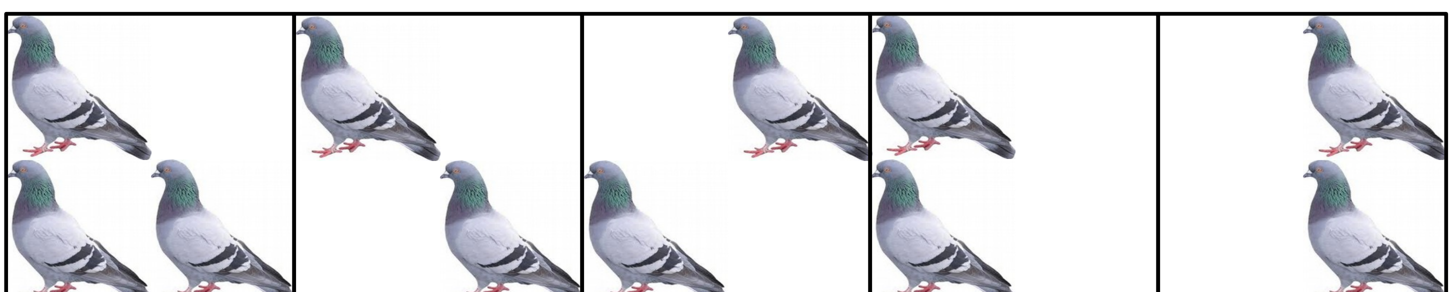
(If there are many bins tied for the most or fewest objects, you can pick any one of them)

Answer at **PollEv.com/cs103** or
text **CS103** to **22333** once to join, then a number.

A More General Version

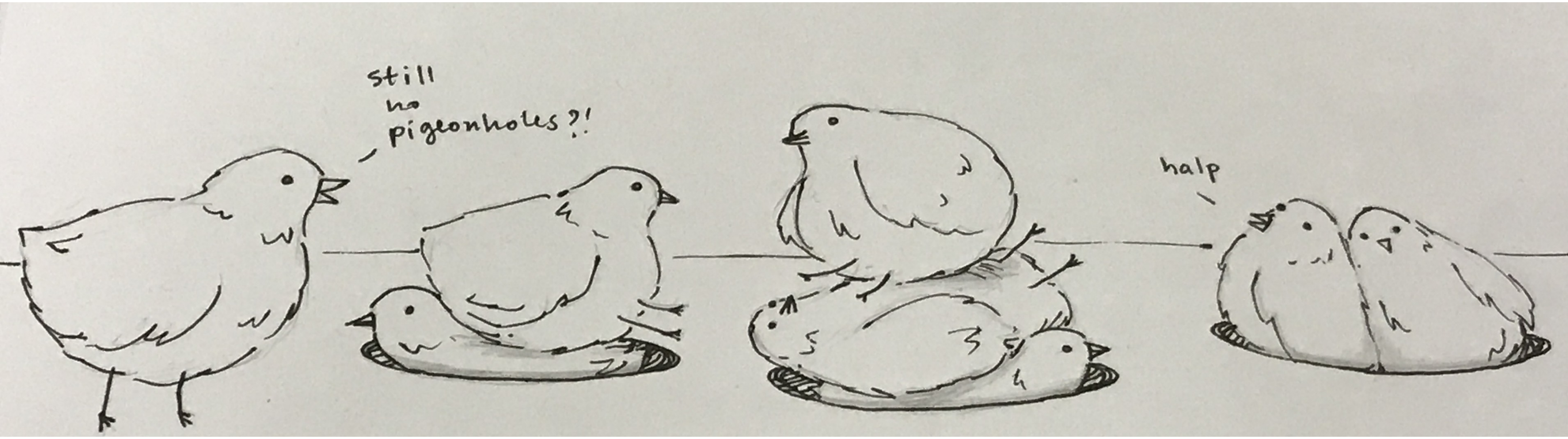
- The ***generalized pigeonhole principle*** says that if you distribute *m* objects into *n* bins, then
  - some bin will have at least $\lceil {}^{m}/_{n} \rceil$ objects in it, and
  - some bin will have at most $\lfloor {}^{m}/_{n} \rfloor$ objects in it.

$\lceil {}^{m}/_{n} \rceil$ means "${}^{m}/_{n}$, rounded up."
$\lfloor {}^{m}/_{n} \rfloor$ means "${}^{m}/_{n}$, rounded down."



$m = 11$
$n = 5$

$\lceil m / n \rceil = 3$
$\lfloor m / n \rfloor = 2$



$m = 8$, $n = 3$

Thanks to Amy Liu for this awesome drawing!

Theorem: If m objects are distributed into $n > 0$ bins, then some bin will contain at least $\lceil m/n \rceil$ objects.

Proof: We will prove that if m objects are distributed into n bins, then some bin contains at least m/n objects. Since the number of objects in each bin is an integer, this will prove that some bin must contain at least $\lceil m/n \rceil$ objects.

To do this, we proceed by contradiction. Suppose that, for some m and n, there is a way to distribute m objects into n bins such that each bin contains fewer than m/n objects.

Number the bins 1, 2, 3, …, n and let x_i denote the number of objects in bin i. Since there are m objects in total, we know that

$$m = x_1 + x_2 + \ldots + x_n.$$

Since each bin contains fewer than m/n objects, we see that $x_i < m/n$ for each i. Therefore, we have that

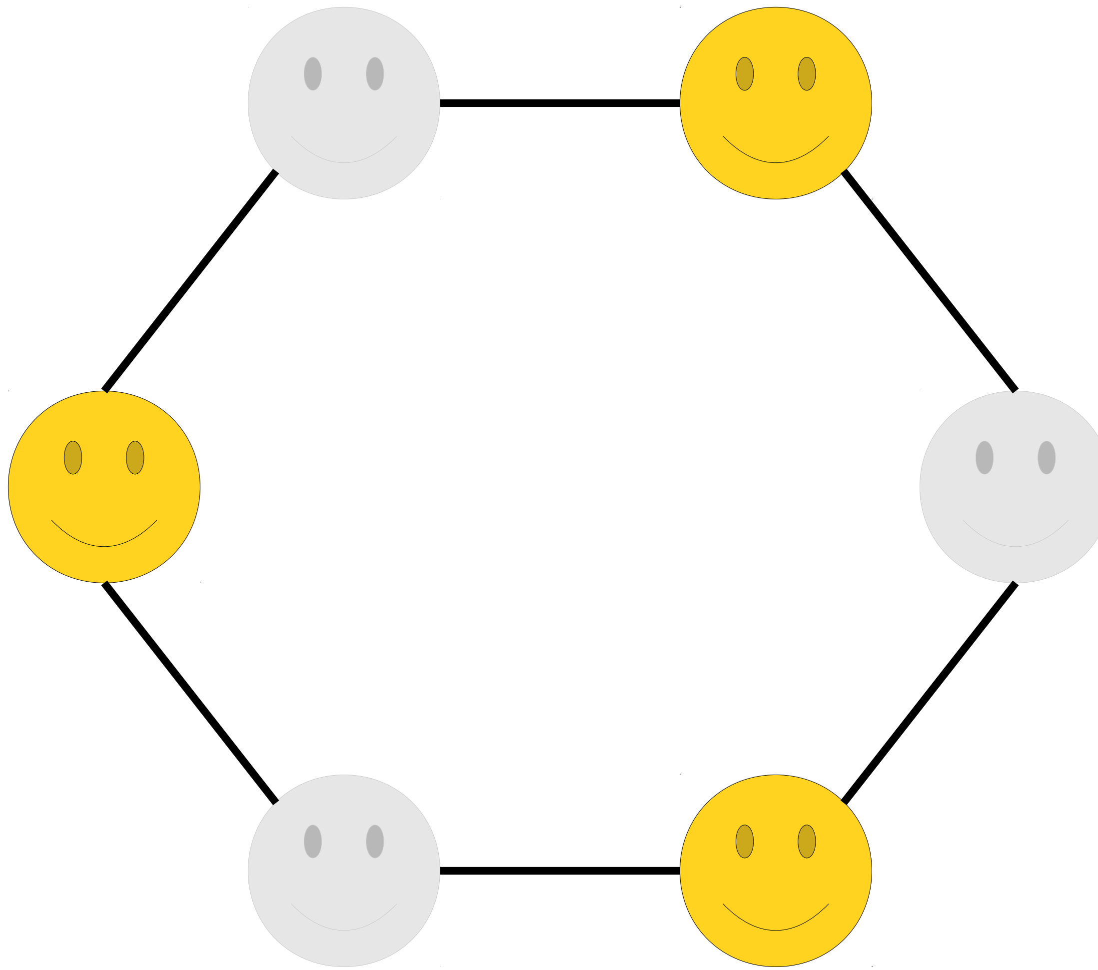
$$\begin{aligned} m &= x_1 + x_2 + \ldots + x_n \\ &< m/n + m/n + \ldots + m/n \quad (n \text{ times}) \\ &= m. \end{aligned}$$

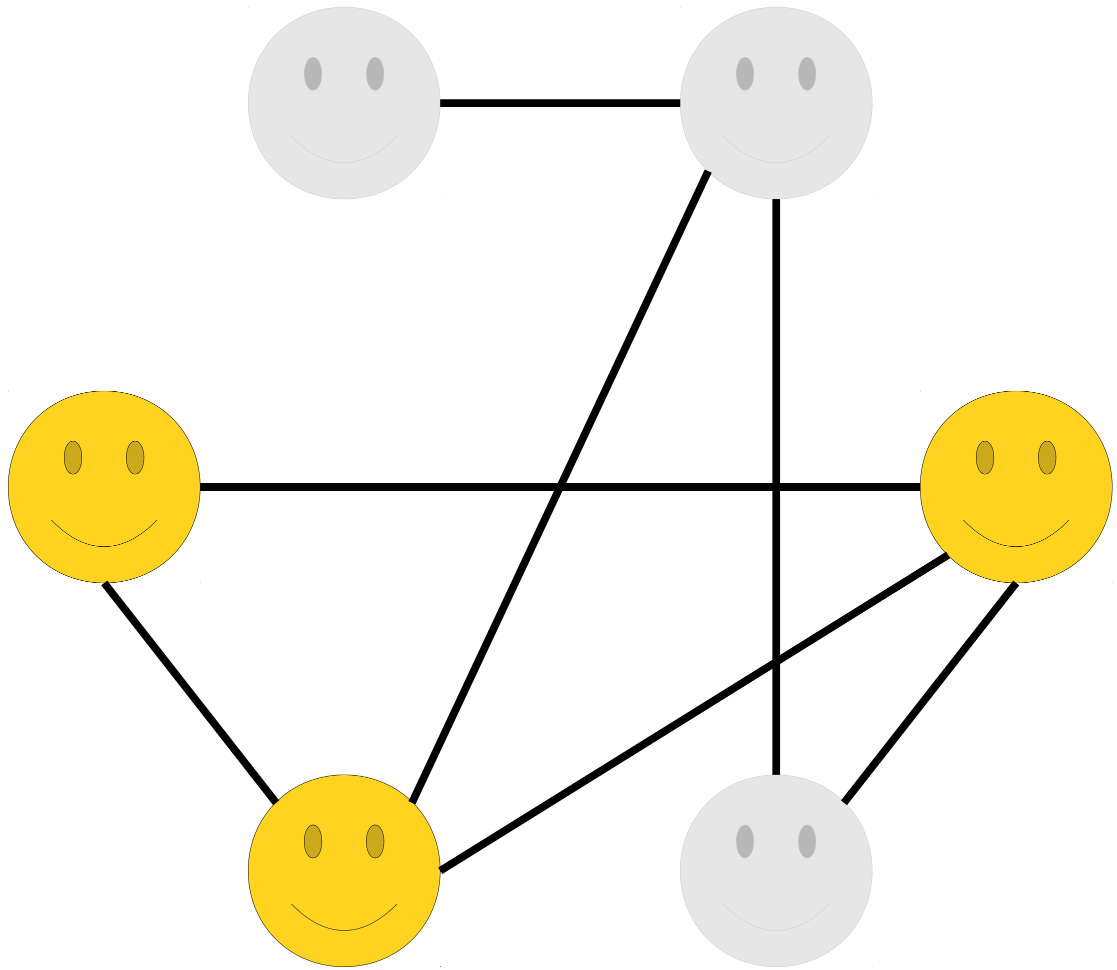
But this means that $m < m$, which is impossible. We have reached a contradiction, so our initial assumption must have been wrong. Therefore, if m objects are distributed into n bins, some bin must contain at least $\lceil m/n \rceil$ objects. ■

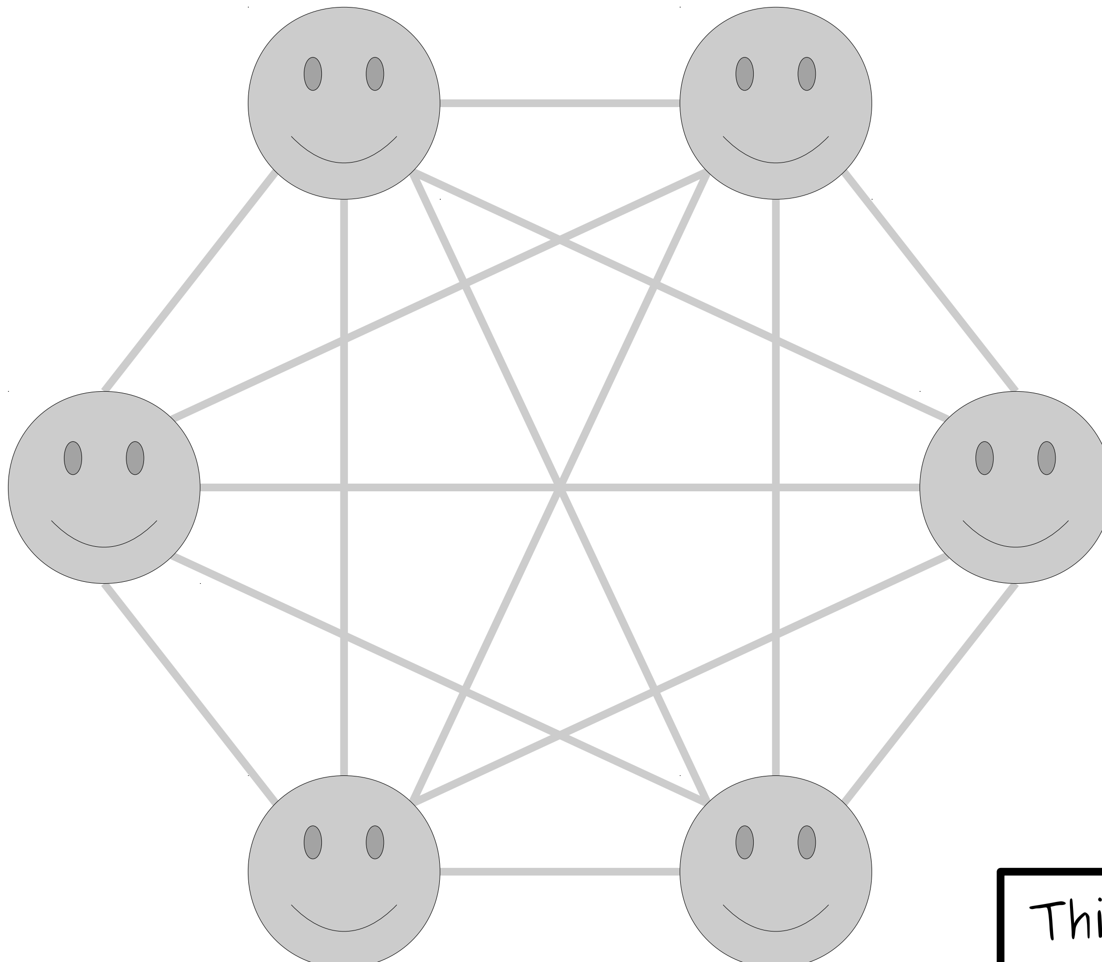
An Application: Friends and Strangers

Friends and Strangers

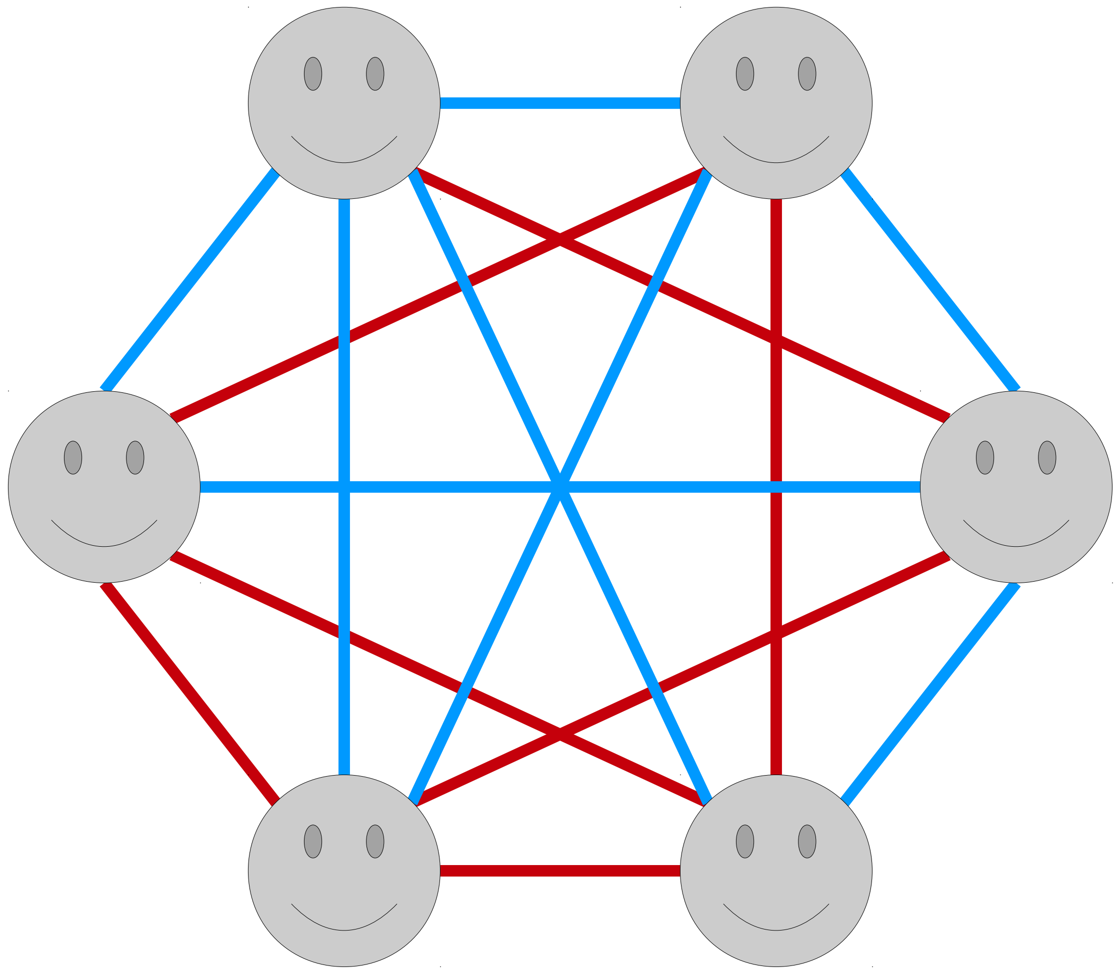
- Suppose you have a party of six people. Each pair of people are either friends (they know each other) or strangers (they do not).
- ***Theorem:*** Any such party must have a group of three mutual friends (three people who all know one another) or three mutual strangers (three people, none of whom know any of the others).

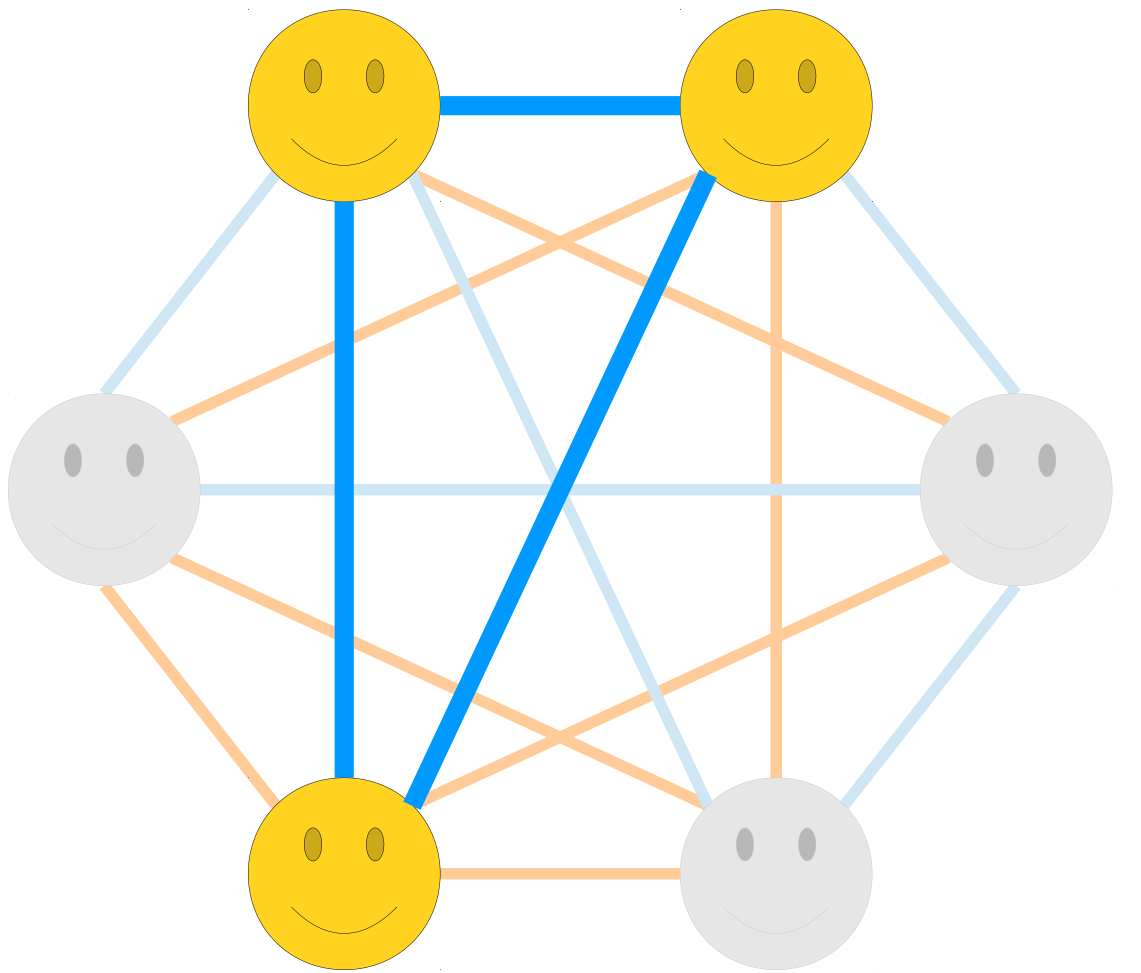


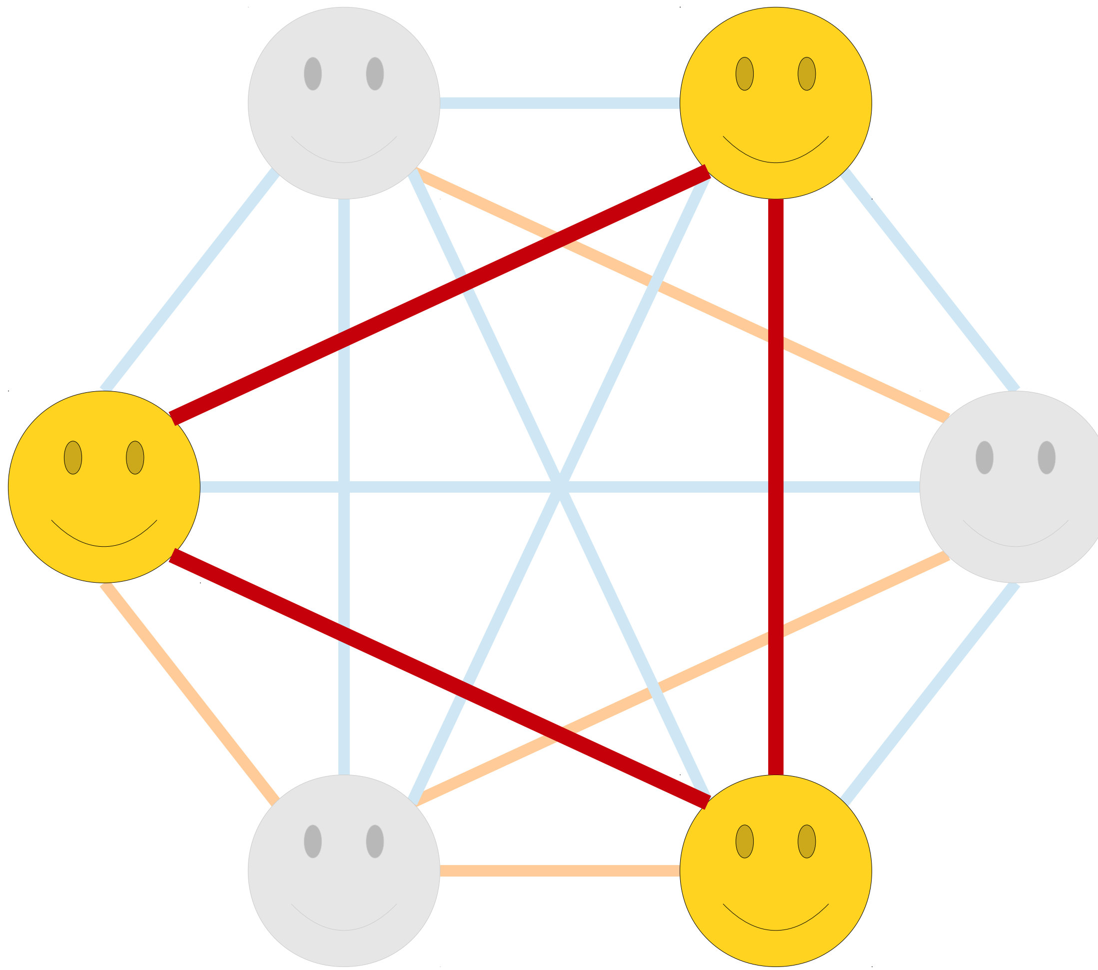




This graph is called a *6-clique*, by the way.

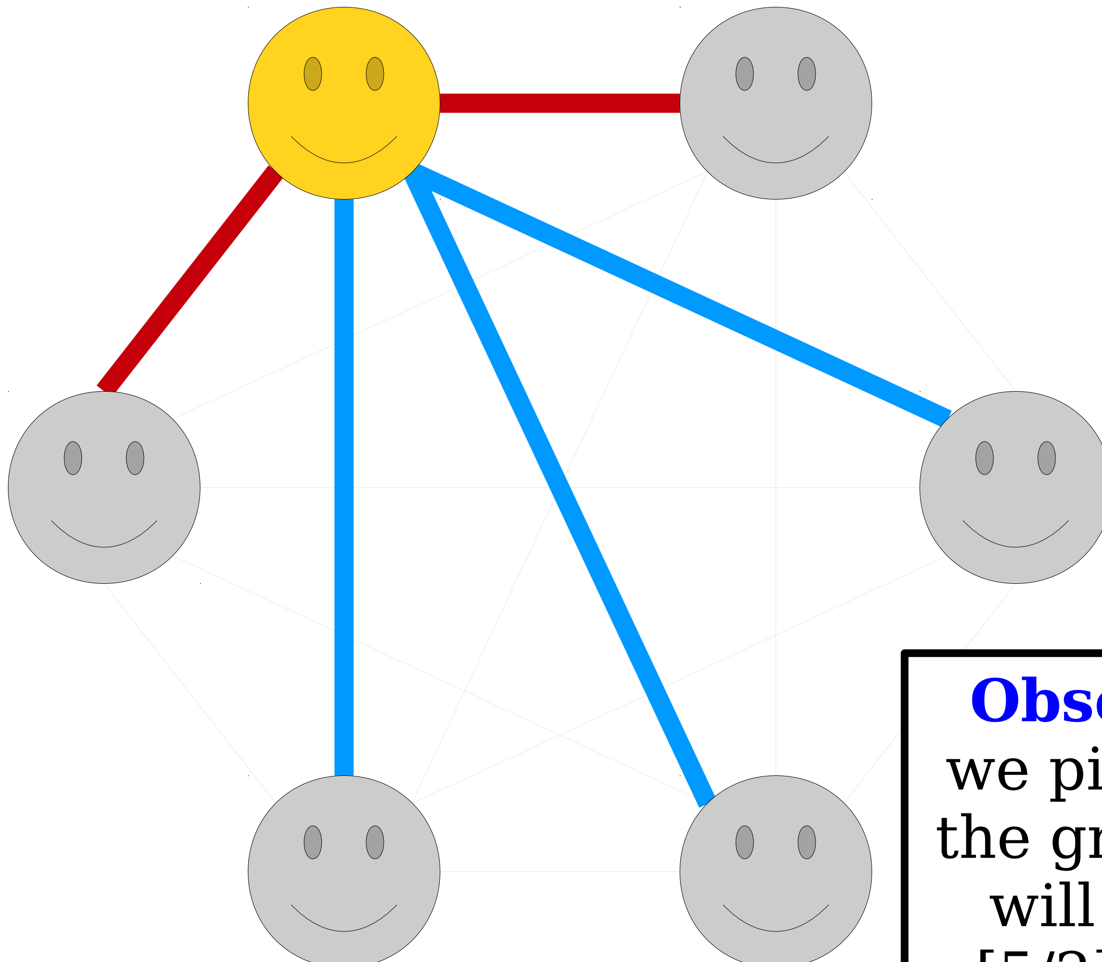




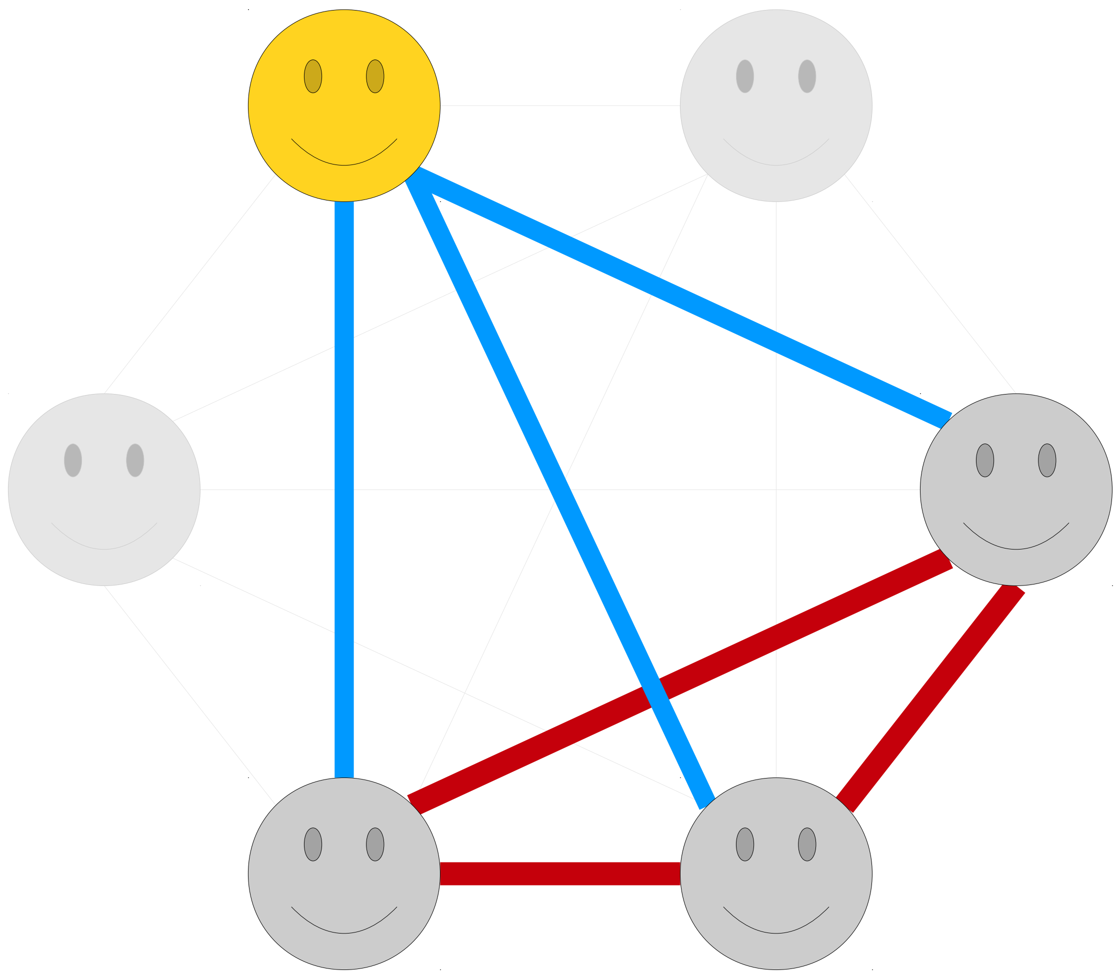


Friends and Strangers Restated

- From a graph-theoretic perspective, the Theorem on Friends and Strangers can be restated as follows:
- ***Theorem:*** Consider a 6-clique where every edge is colored red or blue. The the graph contains a red triangle, a blue triangle, or both.
- How can we prove this?



Observation 1: If we pick any node in the graph, that node will have at least $\lceil 5/2 \rceil = 3$ edges of the same color incident to it.



Theorem: Consider a 6-clique in which every edge is colored either red or blue. Then there must be a triangle of red edges, a triangle of blue edges, or both.

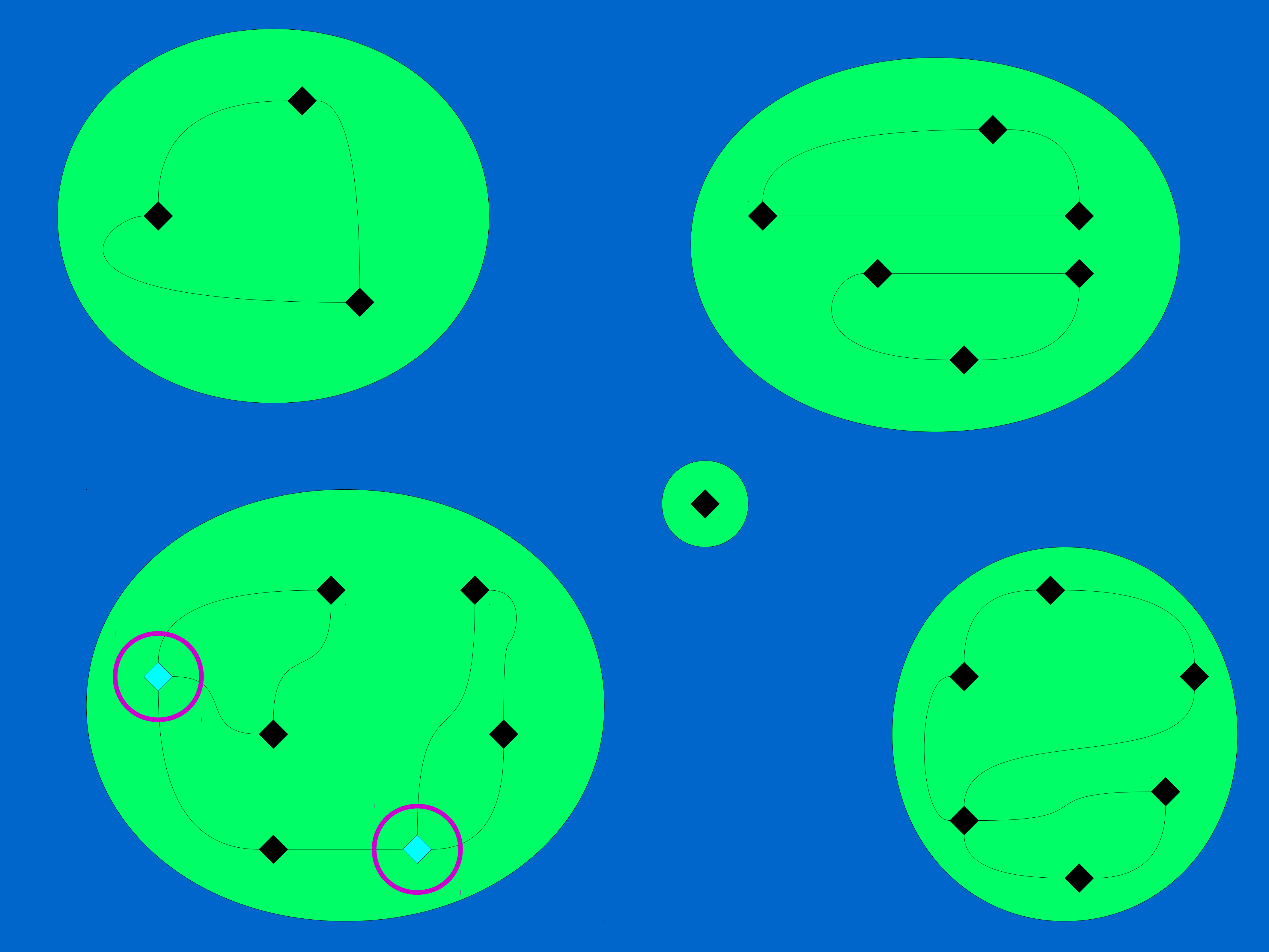
Proof: Color the edges of the 6-clique either red or blue arbitrarily. Let x be any node in the 6-clique. It is incident to five edges and there are two possible colors for those edges. Therefore, by the generalized pigeonhole principle, at least $\lceil 5/2 \rceil = 3$ of those edges must be the same color. Call that color c_1 and let the other color be c_2.

Let r, s, and t be three of the nodes adjacent to node x along an edge of color c_1. If any of the edges $\{r, s\}$, $\{r, t\}$, or $\{s, t\}$ are of color c_1, then one of those edges plus the two edges connecting back to node x form a triangle of color c_1. Otherwise, all three of those edges are of color c_2, and they form a triangle of color c_2. Overall, this gives a red triangle or a blue triangle, as required. ■

Ramsey Theory

- The proof we did is a special case of a broader result.
- ***Theorem (Ramsey's Theorem):*** For any natural number n, there is a smallest natural number $R(n)$ such that if the edges of an $R(n)$-clique are colored red or blue, the resulting graph will contain either a red n-clique or a blue n-clique.
  - Our proof was that $R(3) \leq 6$.
- A more philosophical take on this theorem: true disorder is impossible at a large scale, since no matter how you organize things, you're guaranteed to find some interesting substructure.

A Little Math Puzzle



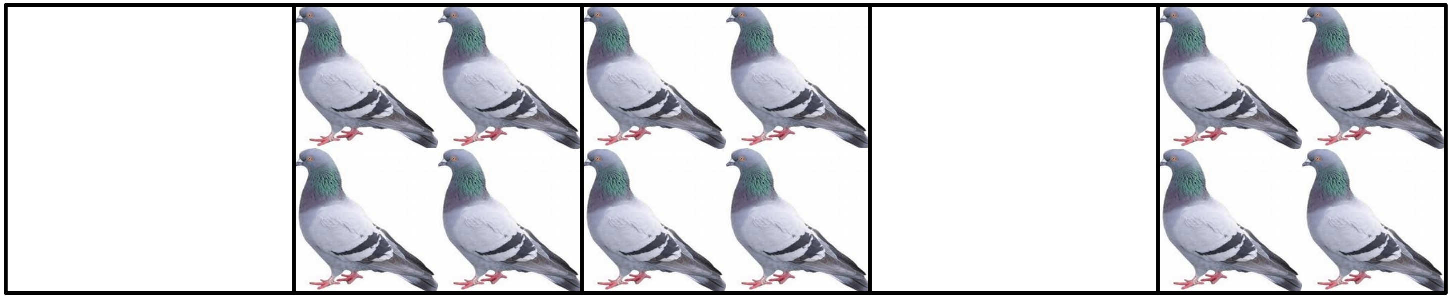
Another View of Pigeonholing

- The pigeonhole principle is a result that, broadly speaking, follows this template:

  ***m* objects cannot be distributed into *n* bins without property *X* being true.**

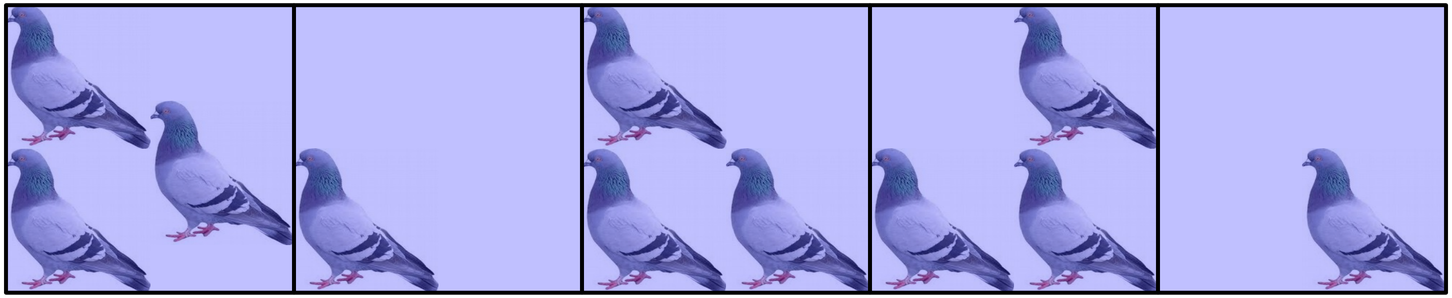
- What other sorts of properties can we say about how objects get distributed?

Observation: The number of boxes containing an odd number of pigeons seems to always be even!



$m = 12$ pigeons
$n = 5$ boxes

Observation: Now the number of boxes containing an odd number of pigeons seems to always be odd!



$m = 11$ pigeons
$n = 5$ boxes

Theorem: Suppose m objects are distributed into some number of bins. Let k be the number of bins containing an odd number of objects. Then k is even if and only if m is even.

Proof: Pick any $m \in \mathbb{N}$ and suppose m objects are distributed into bins. Let k be the number of bins containing an odd number of objects. We will prove that k is even if and only if m is even.

Begin by removing one object from each bin with an odd number of objects in it.

Let r denote the number of objects left in the boxes after the above step (not the number of objects we removed.) What can we say about r?

A. r is even.
B. r is odd.
C. r is even if and only if k is even.
D. r is even if and only if k is odd.
E. None of these, or more than one of these.

Answer at **PollEv.com/cs103** or text **CS103** to **22333** one to join, then **A**, …, or **E**.

Theorem: Suppose m objects are distributed into some number of bins. Let k be the number of bins containing an odd number of objects. Then k is even if and only if m is even.

Proof: Pick any $m \in \mathbb{N}$ and suppose m objects are distributed into bins. Let k be the number of bins containing an odd number of objects. We will prove that k is even if and only if m is even.

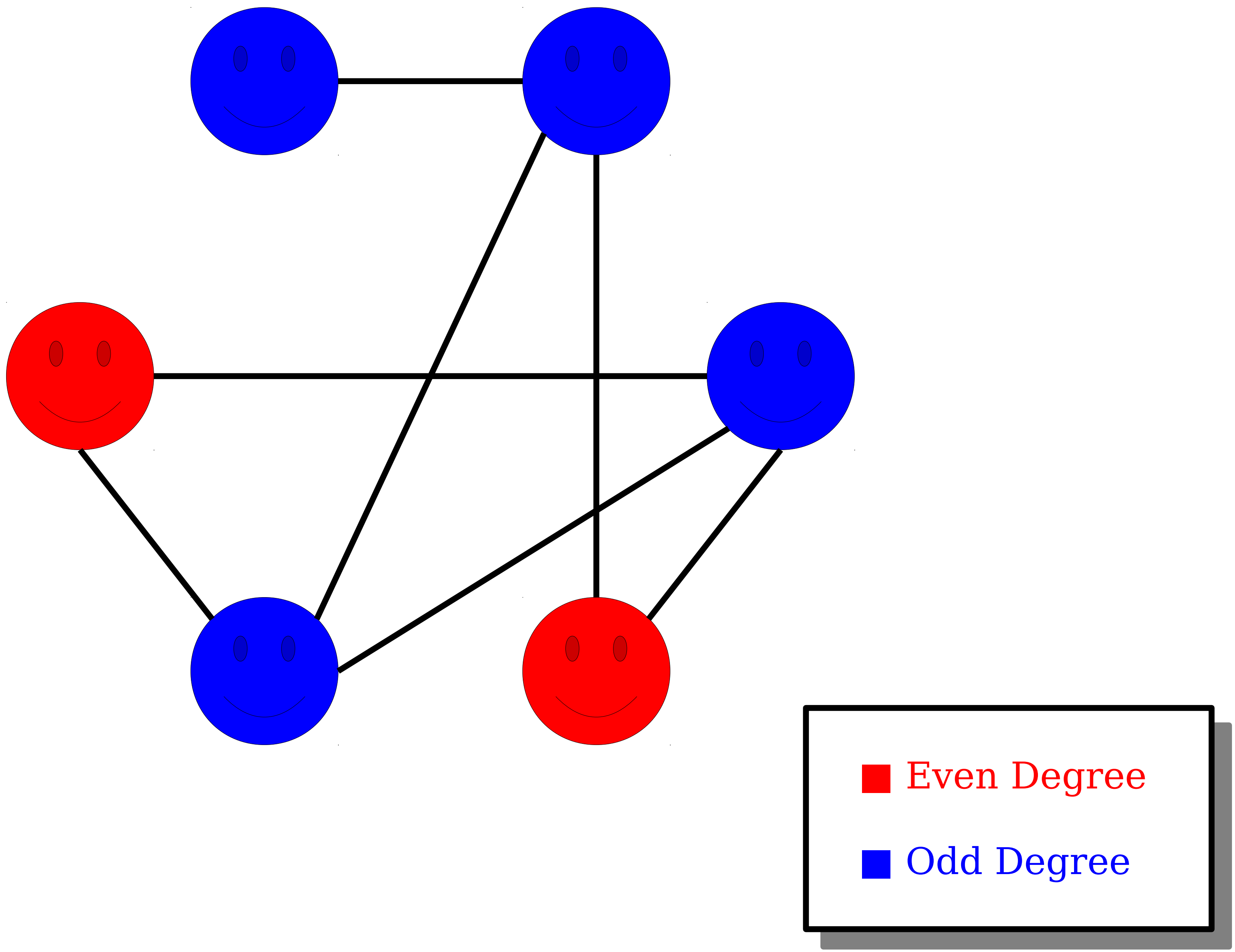
Begin by removing one object from each bin with an odd number of objects in it. Since there are m objects and were k bins containing an odd number of objects, we now have $m - k$ objects left in our bins. For notational simplicity, let $r = m - k$. This also means $m = r + k$.

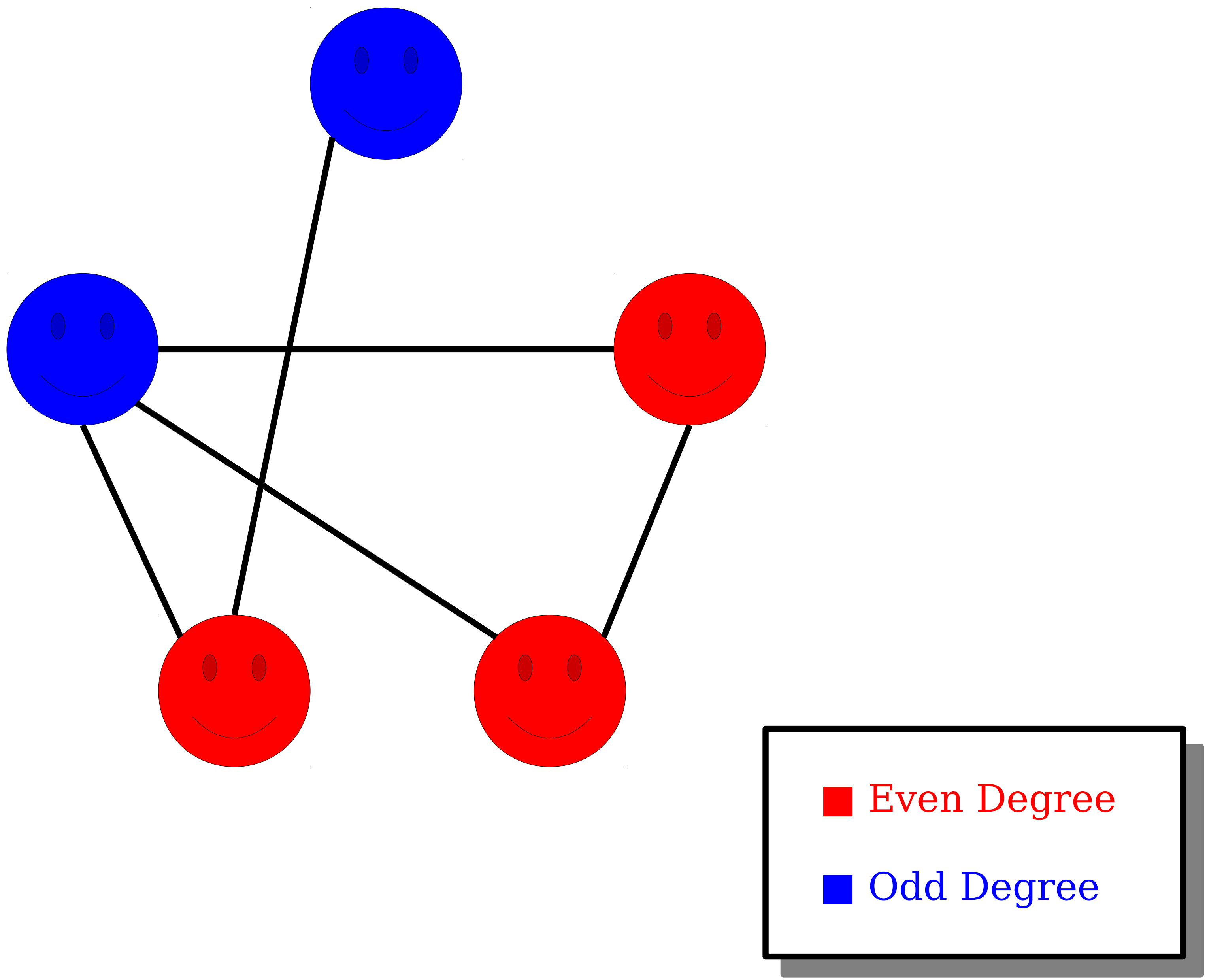
We claim that r is even. To see this, note that each bin now contains an even number of objects; each bin either started with an even number of objects, or started with an odd number of objects and had one object removed from it. This means that r is the sum of some number of even natural numbers, so r is even.

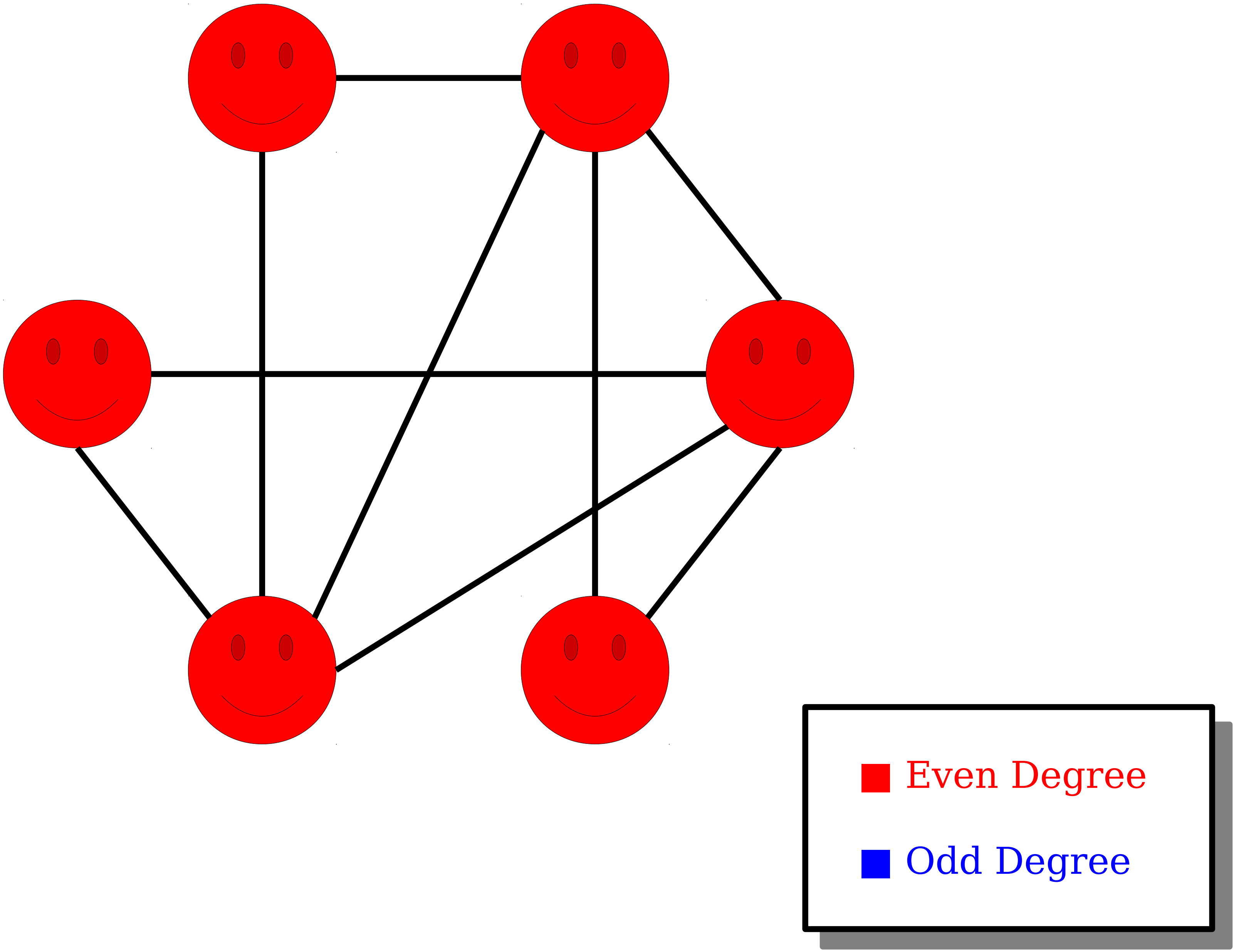
We'll now prove the theorem. First, we'll prove that if k is even, then m is even. To see this, assume k is even. Then $m = r + k$ is the sum of two even numbers, so m is even. Next, we'll prove that if k is odd, then m is odd. To see this, assume k is odd. Then since $m = r + k$ is the sum of an even number and an odd number, we see that m is odd, as required. ■

A Pretty Nifty Theorem:

The Handshaking Lemma

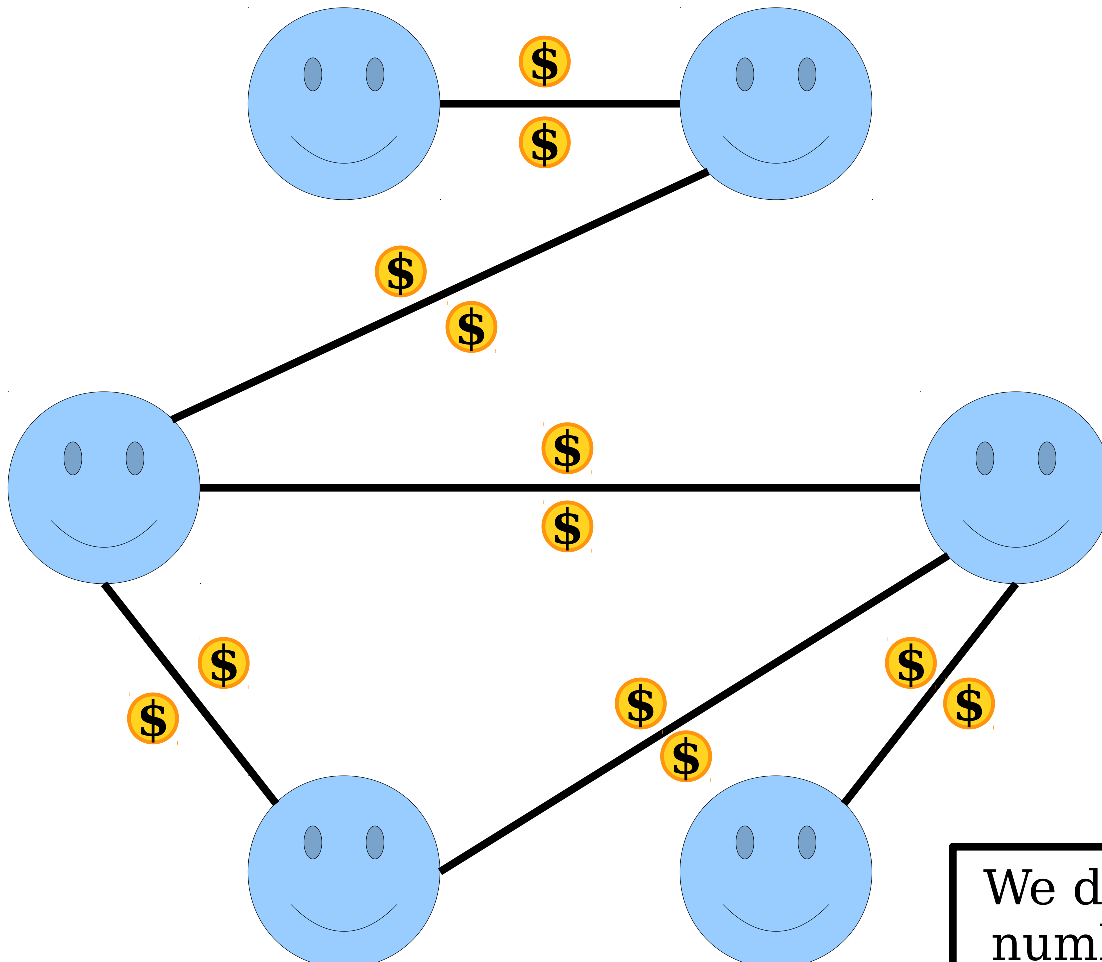

Even Degree
Odd Degree


Even Degree
Odd Degree


Even Degree
Odd Degree

Theorem (The Handshaking Lemma):
Let $G = (V, E)$ be a graph. Then each connected component of G has an even number of nodes of odd degree.

There are $2m$ total coins here, where m is the number of edges.



We distributed an even number of coins into a collection of nodes. Therefore, the number of nodes of odd degree is even!

Theorem (Handshaking Lemma): If G is a graph, then each connected component of G has an even number of nodes of odd degree.

Proof: Let $G = (V, E)$ be a graph and let C be a connnected component of G. Place one coin on each node in C for each edge in E incident to it. Notice that the number of coins on any node v is equal to $\deg(v)$.

We claim that there are an even total coins distributed across all the nodes of G. Notice that each edge contributes two coins to the total, one for each of its endpoints. This means that there are $2m$ total coins distributed across the nodes of V, where m is the number of edges adjacent to nodes in C, and $2m$ is even.

Since there are an even number of coins distributed across the nodes, our earlier theorem tells us that the number of nodes in G with an odd number of coins on them must be even. The number of coins on each node is the degree of that node, and therefore there must be an even number of nodes of odd degree. ■

A Fun Corollary

- A ***corollary*** of a theorem is a statement that follows nicely from the theorem.
- The previous theorem has this lovely follow-up:
- ***Corollary:*** If G is a graph with exactly two nodes of odd degree, those nodes are connected.

Corollary: If G is a graph with exactly two nodes of odd degree, then those two nodes are connected in G.

Proof: Let G be a graph with exactly two nodes u and v of odd degree. Consider the connected component C containing the node u. By the Handshaking Lemma, we know that C must contain an even number of nodes of odd degree. Therefore, C must contain at least one node of odd degree other than u, since otherwise C would have exactly one node of odd degree. Since v is the only node in G aside from u that has odd degree, we see that v must belong to C. Overall, this means that u and v are in the same connected component, so u and v are connected in G, as required. ■

Some Applications

- The corollary we just presented has some pretty unexpected applications:
  - The ***mountain-climbing theorem***. Suppose that two people start climbing the same mountain, beginning at any two spots they'd like. The two climbers can each choose a path to the summit such that they arrive at the same time, and have the same altitude throughout the entire journey.
  - ***Sperner's lemma***. A powerful mathematical primitive that lets you find equitable ways to split the rent in an apartment or show that no matter how you stir your coffee, there's always some particle that remains in the same place.
- And, as you saw on Problem Set Two, looking at parity is a powerful way to prove that certain objects must exist!

Next Time

- ***Mathematical Induction***
  - Proofs on stepwise processes
- ***Applications of Induction***
  - … to numbers!
  - … to data compression!
  - … to puzzles!
  - … to algorithms!