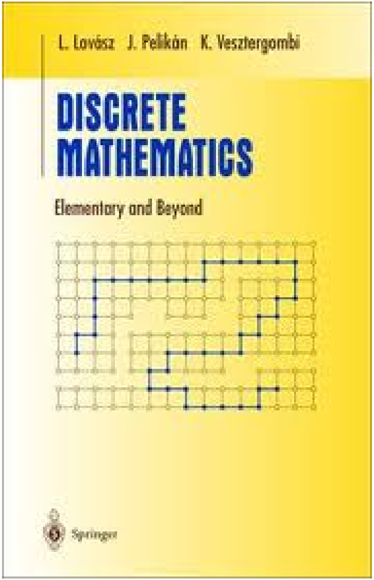
ICME Math Refresher
Discrete Math and Algorithms

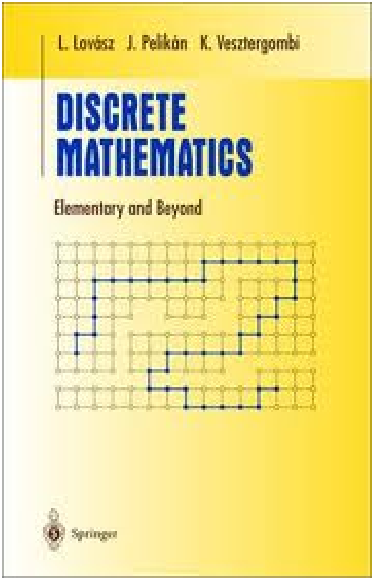
Madeleine Udell

September 22, 2011

Outline



Important chapters:
1-3,5,7-10,13,15



Important chapters:
1-3,5,7-10,13,15

Read these before taking
CME 305

Outline

A Clustering Problem

Find groups consisting of at least one person from each of these categories, and no more than four people total:

1. I've taken a class on all this in the not so distant past
2. I've seen some of this stuff before
3. I have no idea what you're talking about

A Clustering Problem

Find groups consisting of at least one person from each of these categories, and no more than four people total:

1. I've taken a class on all this in the not so distant past
2. I've seen some of this stuff before
3. I have no idea what you're talking about

Exercise

How many distinct class clusterings are possible?

Hint

Feel free to modify the question or to add extra assumptions to simplify the problem. Be explicit about what you add!

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there?

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$?

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$? $(n!)^3$

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$? $(n!)^3$
3. How many distinct groupings are there in which each group consists of three people of three different levels, if there are n people with each level?

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$? $(n!)^3$
3. How many distinct groupings are there in which each group consists of three people of three different levels, if there are n people with each level? $\frac{(n!)^3}{n!} = (n!)^2$

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$? $(n!)^3$
3. How many distinct groupings are there in which each group consists of three people of three different levels, if there are n people with each level? $\frac{(n!)^3}{n!} = (n!)^2$
4. What if the groups have sizes l, m, n with $n = \min(l, m, n)$, and we simply exclude the extra people from participating in any group?

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$? $(n!)^3$
3. How many distinct groupings are there in which each group consists of three people of three different levels, if there are n people with each level? $\frac{(n!)^3}{n!} = (n!)^2$
4. What if the groups have sizes l, m, n with $n = \min(l, m, n)$, and we simply exclude the extra people from participating in any group? $\binom{l}{n}\binom{m}{n}(n!)^2$

A Clustering Problem: Solutions

1. Suppose that we assign groups dictatorially by assigning a seating chart. If $3n$ people are in the class, how many distinct seating charts are there? $(3n)!$
2. If we have equal numbers of each of the three groups, how many seating charts are there in which only people in group i sit in seat m such that $m = i \bmod 3$ for $i = 1, 2, 3$? $(n!)^3$
3. How many distinct groupings are there in which each group consists of three people of three different levels, if there are n people with each level? $\frac{(n!)^3}{n!} = (n!)^2$
4. What if the groups have sizes l, m, n with $n = \min(l, m, n)$, and we simply exclude the extra people from participating in any group? $\binom{l}{n}\binom{m}{n}(n!)^2$

Continuous Approximation

Exercise

How lucky are each of you to be with your current partners? How probable is it that we arrived at this arrangement?

According to Stirling's Approximation,

$$n! \approx \sqrt{2\pi n} e^{-n} n^n$$

Continuous Approximation

Exercise

How lucky are each of you to be with your current partners? How probable is it that we arrived at this arrangement?

According to Stirling's Approximation,

$$n! \approx \sqrt{2\pi n} e^{-n} n^n$$

If there were 150 people in the class, that makes $(50!)^2 \approx 9.21 \times 10^{128}$ possible groupings for the class.

Continuous Approximation

Exercise

How lucky are each of you to be with your current partners? How probable is it that we arrived at this arrangement?

According to Stirling's Approximation,

$$n! \approx \sqrt{2\pi n} e^{-n} n^n$$

If there were 150 people in the class, that makes $(50!)^2 \approx 9.21 \times 10^{128}$ possible groupings for the class. A stack of papers enumerating these groupings would be taller than the observable universe.

Outline

Even and Odd

Exercise

53 students take the iCME refresher course together. Show that at least one person knows an even number of other students.

Even and Odd

Exercise

53 students take the iCME refresher course together. Show that at least one person knows an even number of other students.

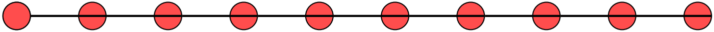
How could you approach this problem?

- Experiment on small examples?
- Generalize the problem?

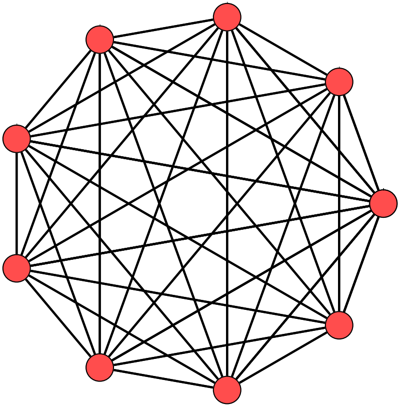
Definitions (I)

- A graph $G(V, E)$ is a set V of vertices and a set E of edges, where $E = \{(a, b), a \in V, b \in V\}$.
- Usually, we denote the number of vertices $|V|$ as n and the number of edges $|E|$ as m.
- The degree of a vertex is the number of edges incident to it, denoted d_i for vertex $v_i \in V$.

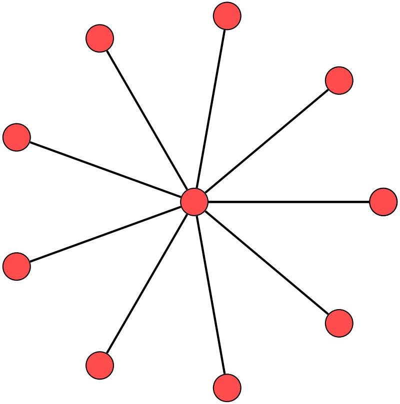
Line Graph L_N



Complete Graph K_N



Star Graph S_N



Experiment on Small Examples

Generalize the Problem

Theorem

If a graph has an odd number of nodes, then it has a node with even degree.

Generalize the Problem

Theorem

If a graph has an odd number of nodes, then it has a node with even degree.

Theorem

If a graph has an odd number of nodes, then the number of nodes with odd degree is even.

Generalize the Problem

Theorem

If a graph has an odd number of nodes, then it has a node with even degree.

Theorem

If a graph has an odd number of nodes, then the number of nodes with odd degree is even.

Theorem

In every graph, the number of nodes with odd degree is even.

Proof by Induction

Start with a graph on n nodes with 0 edges, so that the number of nodes with odd degree is 0, an even number. Add edges one at a time. Each time we add an edge,

Proof by Induction

Start with a graph on n nodes with 0 edges, so that the number of nodes with odd degree is 0, an even number. Add edges one at a time. Each time we add an edge, either

- Two nodes of odd degree become nodes of even degree
- Two nodes of even degree become nodes of odd degree, or
- An even node becomes odd and an odd node becomes even

In any case, the parity of the number of nodes with odd degree is unchanged. Hence the number of nodes of odd degree is even.

Proof by Contradiction

Suppose the number of nodes of odd degree is odd. Then the total degree of vertices in the graph,

$$d(G) = \sum_i d(i),$$

must be odd. But the sum of the degrees must be twice then number of edges in the graph — an even number! Hence we have a contradiction, and the premise we assumed, that odd numbers of nodes with odd degree can exist, must be false.

Basic Proof Techniques

- Proof by Induction: First, show that the *base case* holds, e.g. for $n = 1$. Then, *assuming* the statement is true for $n - 1$, show that the statement holds for n.
- Proof by Contradiction: In logical terms, this can be stated

  $$(A \Rightarrow B) \equiv (\bar{B} \Rightarrow \bar{A})$$

  In other words, to show that A implies B, it is equivalent to show that B is false implies that A is false.

The Pigeonhole Principle

Exercise

Are there two people in Beijing with the same number of hairs on their heads?

The Pigeonhole Principle

Exercise

Are there two people in Beijing with the same number of hairs on their heads?

- Number of people in Beijing: 2×10^7
- Maximum number of hairs: 2×10^5

The Pigeonhole Principle

Exercise

Are there two people in Beijing with the same number of hairs on their heads?

- Number of people in Beijing: 2×10^7
- Maximum number of hairs: 2×10^5

Yes

Outline

Graph Coloring

Definition

A bipartite graph is one in which nodes can be divided into two types, and no node has an edge connecting it to another node of its own type

Definition

A graph coloring is an assignment of the vertices of a graph to different classes (or colors) such that no vertex has an edge to another vertex of the same color.

Definition

The chromatic number of a graph is the minimum number of colors necessary to obtain a valid graph coloring

For example, the chromatic number of a bipartite graph is 2.

Bipartite Graph Exercises

Definition

A bipartite graph is one in which nodes can be divided into two types, and no node has an edge connecting it to another node of its own type

Exercise

1. *How could you check if a graph is bipartite?*
2. *Can you devise a condition?*
3. *Can you prove it is necessary?*
4. *Can you prove it is sufficient?*

Bipartite Graph Exercises

Definition

A bipartite graph is one in which nodes can be divided into two types, and no node has an edge connecting it to another node of its own type

Exercise

1. *How could you check if a graph is bipartite?*
2. *Can you devise a condition?*
3. *Can you prove it is necessary?*
4. *Can you prove it is sufficient?*

Hint

Try a greedy algorithm. How can you show it works?

Outline

Trees

Definition

A graph is connected if there exists a path between each pair of vertices.

Definition

A cycle is a closed path, i.e. a path combined with the edge (v_k, v_1).

Definition

A tree is a connected graph with no cycles.

Tree Identification

Which of these are trees?

Principled Tree Identification

Show the following properties :

1. Every tree (with $|V| \geq 2$) has at least two leaves.
2. A tree has $|V| - 1$ edges.
3. Any connected graph G with $|V| - 1$ edges is a tree.

Outline

Minimum Spanning Tree

Exercise

In 2050, after the infrastructure debacle of the previous half-century, a junta of discrete mathematicians stages a successful coup in California. Supported by legions of thirsty citizens, their task is to build a new network to transport water to every home and business in the state. Due to NIMBY opposition, however, their choices are somewhat restricted: a citizens' council provides them with a list of locations between which water lines can be built, and the cost of building such lines.

How can they build an adequate water system as cheaply as possible?

Minimum Spanning Tree — Graph Formulation

Definition

A Minimum Spanning Tree for a graph $G = (V, E, w)$ is a subset $T \subset E$ of the edges of G such that

- $F = (V, T)$ is a tree
- For any tree $H = (V, T')$,

$$\sum_{e \in T} w(e) \leq \sum_{e \in T'} w(e)$$

Exercise

Given a graph $G = (V, E)$, find a minimum spanning tree $T \subset E$.

Minimum Spanning Tree — Graph Formulation

Definition

A Minimum Spanning Tree for a graph $G = (V, E, w)$ is a subset $T \subset E$ of the edges of G such that

- $F = (V, T)$ is a tree
- For any tree $H = (V, T')$,

$$\sum_{e \in T} w(e) \leq \sum_{e \in T'} w(e)$$

Exercise

Given a graph $G = (V, E)$, find a minimum spanning tree $T \subset E$.

Hint

Try a greedy algorithm. How can you show it works?

Minimum Spanning Tree: Kruskal's Algorithm

Solution

Initialize by setting $T = \emptyset$. While there are still some nodes not connected to the graph, choose the cheapest edge that does not create a loop if added to the tree and place it in T.

Exercise

1. *Show that this procedure does produces a spanning tree.*
2. *Show that this procedure produces the minimum spanning tree.*
3. *What is the complexity of this algorithm?*

Directed Graphs: Definitions

- A directed edge is an ordered pair of vertices.
- Directed graphs permit directed edges.

Minimum Spanning Tree: Prim's Algorithm

Solution

Initialize by selecting a node and placing it into set S*, and setting* $T = \emptyset$*. While there are still some nodes not in* S*, choose the cheapest edge* $e = (u, v)$ *from* $u \in S$ *to* $v \in S^C$*. Place* e *in* T *and* v *in* S*.*

Exercise

1. *Show that this procedure does produces a spanning tree.*
2. *Show that this procedure produces the minimum spanning tree.*
3. *What is the complexity of this algorithm?*

Outline

Traveling Salesman Problem

Exercise

Continuing their research into optimal water transportation, the junta realizes that by creating a cyclical water network, they can recycle the water that circulates through the system rather than allowing it to fester at the leaves of the tree. How can they build a water system through all nodes that consists of a single cycle as cheaply as possible?

Traveling Salesman Problem

Exercise

Continuing their research into optimal water transportation, the junta realizes that by creating a cyclical water network, they can recycle the water that circulates through the system rather than allowing it to fester at the leaves of the tree. How can they build a water system through all nodes that consists of a single cycle as cheaply as possible?

Hint

The greedy algorithm doesn't work! Can you provide a counterexample?

Traveling Salesman Problem

Exercise

Continuing their research into optimal water transportation, the junta realizes that by creating a cyclical water network, they can recycle the water that circulates through the system rather than allowing it to fester at the leaves of the tree. How can they build a water system through all nodes that consists of a single cycle as cheaply as possible?

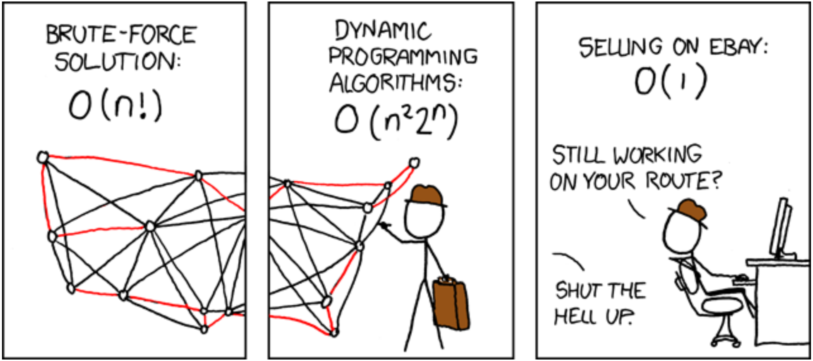
Hint

The greedy algorithm doesn't work! Can you provide a counterexample?

Hint

What's a surefire, if time-consuming, way to find the solution?

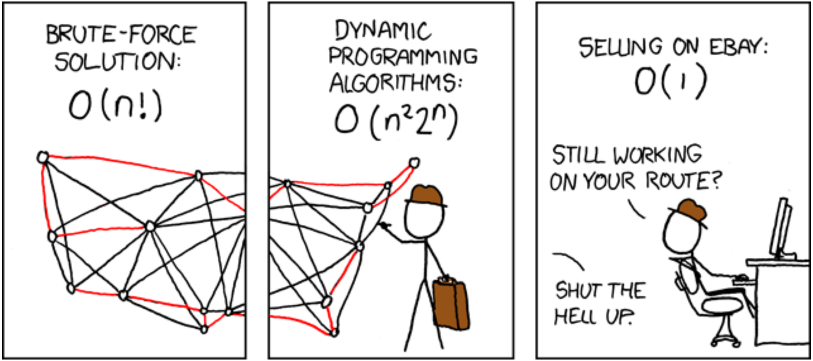
NP Hardness



Up to constant factors, we don't know any algorithm that can find the exact solution faster than by checking all possible routes.

- How many is that, if our community has 200 homes?

NP Hardness



Up to constant factors, we don't know any algorithm that can find the exact solution faster than by checking all possible routes.

- How many is that, if our community has 200 homes?
  $(199)! \approx 4 \times 10^{372}$
  This would take longer to process than the universe is old.

NP Hardness: Definitions

Definition

A problem is in NP if a solution to the problem can be exhibited in polynomial time.

Exercise

How could you exhibit a solution to the traveling salesman problem?

If the problem can also be solved in polynomial time, we say the problem is in P

NP Hardness: Definitions and Technicalities

- Decision Problem: Yes/No answer
  Is there a circuit in the graph with total length less than k?
- Optimization Problem: Numerical answer
  What is the length of the shortest circuit in the graph?

Definition

An oracle is a device that returns the answer to a problem, with no insight into how that answer was computed (also called a black box).
A problem P is called NP-hard if one can compute the solution to some decision problem in NP using a polynomial number of calls to the oracle.

NP Hardness: Definitions and Technicalities

- Decision Problem: Yes/No answer
  Is there a circuit in the graph with total length less than k?
- Optimization Problem: Numerical answer
  What is the length of the shortest circuit in the graph?

Definition

An oracle is a device that returns the answer to a problem, with no insight into how that answer was computed (also called a black box).
A problem P is called NP-hard if one can compute the solution to some decision problem in NP using a polynomial number of calls to the oracle.

Exercise

1. *Show how to find the solution to the decision problem above using an oracle for the optimization problem. How many oracle calls are required?*
2. *Show how to find the solution to the optimization problem above using an oracle for the decision problem. How many*

But we need to solve it anyway

But we need to solve it anyway

Instead, we use heuristics:

- Backtracking
- Simulated Annealing
- ...

Steiner Tree

Exercise

Deciding that democratic procedures are not in the best interests of society at large, the junta sends the NIMBYs to forced labor camps in Siberia. Freed of constraints, they find that the cost of building a water system is just the length of the total pipes laid down. They figure that the best way to do this is add artificial nodes to the graph of nodes already given, and to construct a minimum spanning tree on this enlarged graph. Where should the extra nodes be built?

Steiner Tree

Exercise

Deciding that democratic procedures are not in the best interests of society at large, the junta sends the NIMBYs to forced labor camps in Siberia. Freed of constraints, they find that the cost of building a water system is just the length of the total pipes laid down. They figure that the best way to do this is add artificial nodes to the graph of nodes already given, and to construct a minimum spanning tree on this enlarged graph. Where should the extra nodes be built?

Recommended Reading: Scott Aaronson, NP-complete Problems and Physical Reality.

How to solve NP-complete problems using:

- Soap Bubbles
- The Anthropic Principle
- Time Travel
- Quantum Gravity
- ...