Math for CS

Ian Poon

June 2024

Selected theorems in MIT 6.042j Complete Coursenotes by Eric Lehman 2010

Contents

1 induction

Definition 1 (Principle of strong induction)

Let P be a predicate on nonnegative integers. If

- $P(0)$ is true and
- for all $n \in \mathbb{N}, P(0), \ldots, P(n)$ *together*(not just $P(n)$ like in normal induction) imply $P(n+1)$

Then $P(m)$ is true for all $m \in \mathbb{N}$

2 sets

2.1 Injections, Surjections and Bijections

Definition 2

surjection: a surjective function

When f is a surjection, we say f is an *onto function* In other words range(f) = codomain(f). Thus $\forall y \in$ codomain there exists $x \in$domain such that $f(x) = y$. Recall that the range(f) $\subseteq$ codomain(f) always. Codomain is just the target set you want to map to. Range is the output set of the map.

Definition 3

injection: an injective function

When f is an injection we say that f is a *1 to 1 function*. This means for all $x_1, x_2 \in A$, if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$

Definition 4

bijection: a function that is both a surjection and injection

When f is a bijection, we say f is a *1 to 1 onto function*

Definition 5

A relation is an equivalence relation if it is

- reflexive: $aRa, \forall a \in A$
- symmetric: $aRb \rightarrow bRa, \forall a, b \in A$
- transitive: $(aRb \wedge bRa) \rightarrow aRc, \forall a, b, c \in A$

2.2 Finite Sets

Let a $A \sim B$ denote a bijection from A onto B
Let $\mathbb{N}_k$ be defined as $\{1, 2, 3 \ldots k\}$ where $k \in \mathbb{N}^+$

Definition 6

A set is a *finite* set if $A = \emptyset$ or there exists natural number k such that $A \sim \mathbb{N}_k$ Then it has *cardinality* 0 and K respectively

Theorem 7

any set equivalent to finite non empty set A is finite and has the same cardinality as A

Proof. suppose $A \sim B$ and $A \sim \mathbb{N}_k$. By **reflexivity** of equivalence relations, $B \sim A$. By **transitivity** of equivalence relations $(B \sim A \wedge A \sim \mathbb{N}_k) \rightarrow B \sim \mathbb{N}_k$.

Lemma 8

If A is a finite set and $x \notin A$ then $A \cup \{x\}$ is finite and card($A \cup \{x\}$)=card(A)+1

Proof. Let $f : A \to \mathbb{N}_k$ be a bijection.

$$g(t) = \begin{cases} f(t) & t \in A \\ k+1 & t = x \end{cases}$$

Clearly $(g : A \cup \{x\}) \to \mathbb{N}_k$ is a bijection. We can certainly find a corresponding t for every element in g(t). If $x_1 = x$ and $x_2 \in A$, $g(x_1) \neq g(x_2)$ since $f(A) = \mathbb{N}_k$ which definitely does not include k+1. If both x_1 and x_2 are in A, then clearly $g(x_1) \neq g(x_2)$ since $f \sim \mathbb{N}_k$

Lemma 9

For each natural number m, if $A \subseteq \mathbb{N}_m$ then A is a finite set and card(A)≤m

Proof. Let our induction hypothesis be: $P(k) : A \subseteq \mathbb{N}_k \to card(A) \leq k$. Prove that P(K+1) is true. Suppose $A \subseteq \mathbb{N}_{k+1}$ Then $(A - \{k+1\}) \subseteq \mathbb{N}_k$. If $\{k+1\} \notin A$ then $card(A) \leq K < K+1$. If $\{k+1\} \in A$ then by 8, $card\left((A - \{k+1\}) \cup \{k+1\}\right) \leq k+1$

Theorem 10

If S is a finite set and A is a subset of S, then A is a finite set and card(A)≤card(S)

Proof. Since finite, there exist a bijection $f : S \to \mathbb{N}_k$ where card(S)=k. Since surjective, range equal co-domain $\mathbb{N}_k = f(S)$. Since $f(A) \subseteq f(S)$ then $f(A) \subseteq \mathbb{N}_k$ so by 9, $card(f(A)) \leq card(S)$. Due to the surjectivity of $S \sim f(S)$, $\forall y \in f(S)$ which then includes $\forall y \in f(A)$, we can find an $x \in S$ such that $f(x) = y$. But if $y \in f(A)$, we are definitely unable to find $x \in \{S - A\}$ due to the injectivity of $S \sim f(S)$. Combining these facts $\forall y \in f(A)$ we can find $x \in A$ such that f(x)=y. Another implication of this injectivity is that $\forall x_1, x_2 \in S \in A$ where $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$. Hence, $f(A) \sim A$ so $card(f(A)) = card(A) \leq card(S)$

Theorem 11

If a set is finite, then is not equivalent to any of its proper subsets

Proof. Suppose A is a proper subset of finite set B. Then there exists $x \in \{B - A\}$. Then A is a subset of $\{B - \{x\}\}$. By 8 $card(B - \{x\}) = card(B) - 1$. Thus $card(A) < card(B)$ so $B \not\sim A$

Corollary 12 (Infinite Set: definition by contrapositiv)

If set A is equivalent to one its proper subsets, set A is infinite

2.3 Infinite Sets

Theorem 13 (Infinite Set: definition by contradiction)

Let A and B be sets

1. if A is infinite and $A \sim B$ then B is infinite
2. if A is infinite and $A \subseteq B$ then B is infinite

Proof. To prove (a), suppose that if A is infinite then B is finite. Since $A \sim B$, by 7 a contradiction is implied. To prove (b) again suppose if A is infinite then B is finite. By 10, we arrive at a contradiction.

Definition 14

$$card(\mathbb{N}) = \aleph_0$$

A set A is *countably infinite* provided $A \sim \mathbb{N}$. Thus $card(A) = \aleph_0$

Remark 15. *a countably infinite set is also known as a* denumerable *set*

Definition 16 (Countable)
A set C is countable if it is *countably infinite* or *finite*

Proposition 17
Prove that the the set $\mathbb{Z}$ is countably infinite

Proof. define $f : \mathbb{N} \to \mathbb{Z}$

$$f(x) = \begin{cases} \frac{n}{2} & \text{n is even} \\ \frac{1-n}{2} & \text{n is odd} \end{cases}$$

for $y \in f(x)$

- if y>0 then $2y \in \mathbb{N}$ since

$$f(2y) = \frac{2y}{2} = y$$

- if $y \leq 0$ then $(1-2y) \in \mathbb{N}$ since

$$f(1-2y) = \frac{1-(1-2y)}{2} = \frac{2y}{2} = y$$

Thus f is a surjection. Now consider $m, n \in \mathbb{N}$ such that $f(m) = f(n)$. We need to prove the contrapositive of 2.1 which is show that then $m = n$.There are 2 cases

- f$(m), f(n) \geq 0$ so both m, n are even natural numbers

$$\frac{n}{2} = \frac{m}{2} \to n = m$$

- f$(m), f(n) < 0$ so both m, n are odd natural numbers

$$\frac{1-n}{2} = \frac{1-m}{2} \to n = m$$

Thus f is an injection too hence f is bijection.

Theorem 18
If A is a countably infinite set then $A \cup \{x\}$ is a countably infinite set.

Proof. Let $f : \mathbb{N} \to A$ be a bijection Consider the case when $\{x\} \notin A$. Then Define $g : \mathbb{N} \to A \cup \{x\}$:

$$g(n) = \begin{cases} x & n = 1 \\ f(n-1) & n > 1 \end{cases}$$

For $y \in g(n)$

- if y=x then $1 \in \mathbb{N}$ that satisfies the map
- if $y \in A$ then there exist $j \in \mathbb{N}$ such that $f(j) = y$ and

$$j = n - 1 \to n = j + 1 \in \mathbb{N}$$

Hence there exist $n > 1 \in \mathbb{N}$ that satisfies the map. The injectivity of g follows from that of f. Thus $g \sim \mathbb{N}$

Consider the case where $x \in A$. Then $A \cup \{x\} = A$. Then $f \sim g \sim \mathbb{N}$

Theorem 19

if A is countably infinite and B is a finite set, then $A \cup B$ is a countably infinite set.

Proof. Let our induction hypothesis be P(n): If set B is finite and card(B)=n then $A \cup B$ is countably infinite. The base cases P(0) is trivial as union with empty set doesnt change the original set and P(1) follows by 18. To prove P(n+1), $card(B - \{x\}) = n$ by 8. Thus since $B - \{x\}$ is finite and has cardinality n, P(n) follows.t

Theorem 20

a union of countable sets is countable

Proof. if you were to generalize the results from above, we basically finding a 1 to 1 correspondance with natural numbers. A union of sets could be seen as a union of rows. You can clearly count the number of rows but the columns must be countable too or we can use cantor diagonal to find a row not in the range of the function(cantor diagonal argument). See rudin for more. A more sophisticated proof is to use perfect sets.

Definition 21 (Power Sets)

A power set pow(A) is the set containing all the subsets of A

Example 22

if $A = \{1, 2, 9\}$ then $\text{pow}(A) = \{\emptyset, \{1\}, \{2\}, \{9\}, \{1, 2\}, \{1, 9\}, \{2, 9\}, \{1, 2, 9\}\}$

Lemma 23

Let $card(A) = n$. Then $card(pow(A)) = 2^n$

Proof. Suppose we have a set of n elements say $A = \{a_1, a_2 \ldots, a_n\}$. Suppose a possible subset of pow(A) is $\{a_1, a_5, a_7\}$. Define a sequence x_n such that $x_i = 1$ if a_n is in this subset and 0 otherwise for 1<i<n. Then this sequece corresponding to this subset will have w elements all 0s except the 1st,5th and 7th element which are 1s.In this way, all possible permutations of this binary sequence will correspond to all possible selections of a_i that can form a subset of A. Thus $f : \text{pow}(A) \to \{1, 0\}^n$ is bijective and their cardinality is 2^n

Theorem 24 (Cantor)
$f : A \to \text{pow}(A)$ is not surjective

Proof. The approach is show that there exists a set $T \in \text{pow}(A)$ that is not in the range of $f(A)$. In other words, set T in the codomain is not mapped from the domain. For each $a \in A$ its mapping to subset $f(a)$ could either contain a as a set member or doesn't. For example $f : 1 \to \{2, 9\}$ doesn't but $f : 2 \to \{1, 2, 9\}$. Thus to construct such a T we simply let it be equal A_f which we define to be $\{a \in A_f | a \notin f(a)\}$. Once again, this is well defined subset of A and hence a member of pow(A)

Why? Consider pow({1,2,3}). Say if 1 is not a member of f(1) and 2 is not a member of f(2), surely (1,2) is neither a member of f(1) or f(2).Thus it is obviously not equal to f(1) or f(2). If 3 is a member of f(3), we simply don't include it in T because there is a chance that it equals f(3) given that it contains 3. For example f(3) could be {1,2,3},{1,3},{3} etc. This implies $\forall a \in A, f(a) \neq A_f$ which is exactly what we want.

More precisely, consider the case if $\exists a \in A_f$ where $f(a) = A_f$ implies $a \in f(a)$ which is a contradiction to the definition of A_f. Now consider if $\exists a \notin A_f$ where $f(a) = A_f$. However that is impossible as $a \in f(a)$. Therefore $f(a)$ contains a which is not a member of A_f. Therefore proposing $f(a) = A_f$ leads to a contradiction.

Corollary 25
The following sets are uncountable

(a) pow($\mathbb{N}$)

(b) $\{1, 0\}^w$ which refers to all possible infinite bit strings

Proof. To prove (a) notice that pow($\mathbb{N}$) means pow({1,2,3,4,5....all other natural numbers}). Thus we know $f : \mathbb{N} \to$ pow($\mathbb{N}$) is not surjective from **Cantor's theorem**. To prove (b), from 23 we that $f : \{1, 0\}^w \to$ pow($\mathbb{N}$) is a bijective since $\mathbb{N}$ is a infinite length set

3 directed graphs & partial orders

Definition 26
A directed graph G consists of a nonempty set $V(G)$ called the **vertices** of G and a set $E(G)$ called the **edges** of G. An element of $V(G)$ is called a **vertex** or **node**. An element of $E(G)$ is called **directed edge**(or sometimes an "arrow"). A directed edge has a **head** starting at some vertex u and **tail** ending at some vertex v

$$\langle u \to v \rangle$$

Definition 27
If G is **directed graph** or **digraph** for short, the **in-degree** of a vertex is the number of arrow coming into it while the **out-degree** is the number of arrows out of it

$$\text{indeg}(v) = \{e \in E(G) | \text{head}(e) = v\}$$

$$\text{outdeg}(v) = \{e \in E(G) | \text{tail}(e) = v\}$$

Hence in a digraph

$$\sum_{v \in V(G)} \text{indeg}(v) = \sum_{v \in V(G)} \text{outdeg}(v) = |E(G)|$$

Definition 28

A sequence of edges traversed is called **walk**. A **path** is a walk which never visits a vertex more than once. For example a walk $\mathbf{v}$ is a sequence of the form

$$\mathbf{v} = v_0\langle v_0 \to v_1\rangle v_1 \dots \langle v_{k-1} \to v_k\rangle v_k$$

where $\langle v_i \to v_{i+1}\rangle \in E(G)$ for $i \in [0, k)$. The **lenght** of the walk is k and said to start and end at v_0 and v_k respectively.

A **closed walk** is a walk that begins and ends at the same vertex. A **cycle** is a positive length closed walk whose vertices are distinct except for the beginning and end vertices("closed path" in a sense except for start and end)

Example 29

Consider

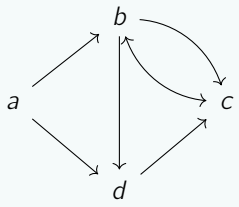


Figure 1: A 4 node directed graph with 6 edges

- (a, b, d) or simply abd is vertex-sequence description of a
- $(\langle a \to b\rangle, \langle b \to d\rangle)$ or simply $\langle a \to b\rangle\langle b \to d\rangle$ is an edge sequence description of the same lenght 2 path above
- $abcbd$ is a lenght 4 walk
- $dcbcbd$ is a lenght 5 closed walk
- $\langle c \to b\rangle\langle b \leftarrow a\rangle\langle a \to d\rangle$ is not a walk as it is not allowed to follow edges in the wrong direction

Definition 30

We denote $\mathbf{f}\hat{v}\mathbf{r}$ to be the merge of a walk $\mathbf{f}$ that ends at v with a walk $\mathbf{r}$ that starts with $v$or simply $\mathbf{f}\hat{}\mathbf{r}$ if v is not specified. So clearly

$$|\mathbf{f}\hat{}\mathbf{r}| = |\mathbf{f}| + |\mathbf{r}|$$

Definition 31

The distance $\text{dist}(u, v)$ in a graph from vertex u to v is the lenght of the shortest path from u to v so clearly

$$\text{dist}(u, v) \leq \text{dist}(u, x) + \text{dist}(x, v)$$

Definition 32

If a graph G has n vertices $v_0, \ldots, v_{n-1}$ we can then repsent it with an $n \times n$ matrix off zeroes and ones called its adjaency matrix defined by

$$(A_G)_{ij} = \begin{cases} 1 & \langle v_i \to v_j \rangle \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

It follows directly from the definition that each element of a_{ij} in (A_G) is the number of length 1 from $v_i \to v_j$. More generally the matrix $(A_G)^k$ provides the count of the number of length k walk between vertices in any digraph. We will prove this below.But you should also notice that $(A_G)^0$ the identity matrix is the number of length zero walks between elements since it represents the graph where the only element connected to it is itself.

Definition 33

The lenght k walk counting matrix for an n-vertex graph G is the $n \times n$ matrix C such that

$$C_{uv} = \text{the number of length-k walks from u to v}$$

Theorem 34

The length k counting matrix of a digraph G is $\langle A_G \rangle^k$ for all $k \in \mathbb{N}$. That is, C_{uv} corresponds to an element $a_{ij} \in A_G$ where u and v is the ith and jth vertex respectively. If $\mathbf{C}$ is the length-k walk counting matrix for a graph G and D is the length walk counting matrix then CD is the length $k + m$ walk counting matrix for G.

Proof. Consider

$$(A_G)^2 = (A_G)(A_G) = \sum_{ij} \sum_{k} a_{ik} a_{kj}$$

hence every $a_{ij} \in (A_G)^2$ is defined by $\sum_k a_{ik} a_{kj}$ which clearly counts the number of possible connected paths $v_i \langle v_i, v_k \rangle v_k \to v_k \langle v_k, v_j \rangle v_j$, enumerating through all possible k. By induction suppose $(A_G)^k$ counts the number of k lenght paths between any 2 nodes. By the same logic as for the base case $(A_G)^2$ which we have proven $(A_G)^k (A_G)$ simply counts the number of $k + 1$ paths hence proving the induction step. Therefore it follows that $(A_G)^k (A_G)^m$ counts the number of $k + m$ paths between any 2 elements.

Definition 35

For a diagraph G The **walk relation** G^* on $V(G)$ is written as

$$uG^*v = \text{there is a walk in G from u to v}$$

This could include length zero walks or self loops.

On the other hand (non-zero)**positive walk relation** G^+ on $V(G)$ is written as

$$uG^+v = \text{there is a positive length walk in G from u to v}$$

Similarly the is **length-n walk relation** G^n on $V(G)$ is written as

$$uG^+v = \text{there is a length-n walk in G from u to v}$$

We say w is **reachable** from v or v is **connected** to w if there is walk from vertex v to w.

Definition 36 (composition of relation)

Let $R : B \to C$ and $S : A \to B$ be binary relations. THen the composition of R with S $(R \circ S) : A \to C$ is defined by

$$a(R \circ S)c = \exists b \in B(aSb)\text{AND}(bRc)$$

Definition 37

A **directed acyclic graph**(DAG) is a directed graph with no cycles

4 Simple Graphs

Definition 38

Simple graphs are defined as bigraphs in which edges are **undirected**, that is they connect two vertices without pointing in either direction between vertices. So write edges $\langle v - w \rangle$ instead of $\langle v \to w \rangle$ like above. Like a directed graphs it has the nonempty sets $V(G)$ of vertices/nodes and $E(G)$ of edges. For every edge $\langle u - v \rangle$ in a DAG, we denote $\{u, v\}, u \neq v$ to be its **endpoints**

Definition 39

The two vertices in a simple graph are said to be **adjacent** if and only if they are the endpoints of the same edge and an edge is said to be **incident** to each of its endpoints. The number of edges incideent to a vertex v is called the **degree** of the vertex and is denoted by $\deg(v)$. Equivalently the degree of a vertex is the number of vertices adjacent to it

It is also clear to see the sum of degrees of the vertices in a graph equals twice the number of edges

Example 40

Consider

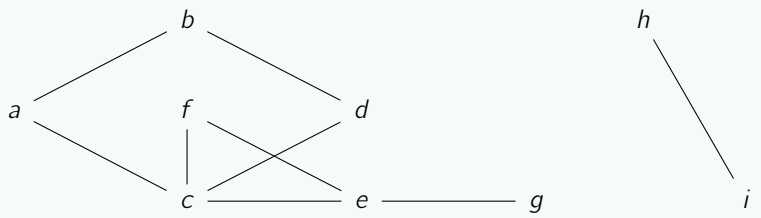


Figure 2: An example of a graph with 9 nodes and 8 deges

- vertex a is adjacent to vertex b and b is adjacent to d
- $\langle a - c \rangle$ is incident to a as well as to c(since $\{a, c\}$ are its endpoints)
- $\deg(e) = 3$

Note that is possible to have vertex with degree 0. In which case it is not adjacent to any other vertices.

We denote that "graph" refers to simple graphs for rest of the section

Definition 41

A **complete graph** K_n has n vertices and an edge between every two vertices. An **empty graph** has no edges at all. An n-node graph containing $n-1$ deges in sequence is know as the **line graph** L_n

Example 42

Consider

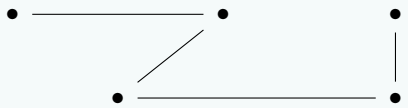


Figure 3: L_5: a 5 node line graph

Definition 43

An **isomorphism** between graphs G and H is a bijection $f : V(G) \to V(H)$ such that

$$\langle u - v \rangle \in E(G) \text{ if and only if } \langle f(u) - f(v) \rangle \in E(H)$$

for all $u, v \in V(G)$

One can see that an isomorphism is a degree preserving map. In simple examples like below, one can easily determine if they are ismorphic by inspection. However no one has found a procedure for determining whether tow graphs are isomorphhic that is guaranteed to run in polynomial time on all pairs of graphs.

Example 44

Let G and H be the C_5 graph on the left and right respectively. Let $f : V(G) \to V(H)$.

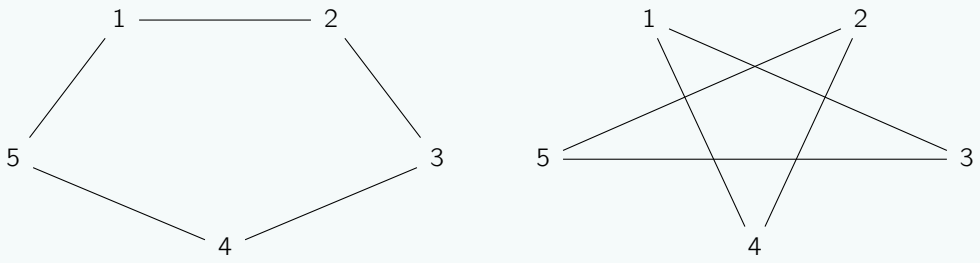


Figure 4: Isomorphic C_5 graphs

See that there exists an isomorphism

$$f(1) = 5, \quad f(2) = 2$$

$$f(2) = 2, \quad f(3) = 4$$

$$f(3) = 4, \quad f(4) = 1$$

$$f(4) = 1, \quad f(5) = 3$$

$$f(5) = 3, \quad f(1) = 5$$

where

$$\langle i, j \rangle \in E(G) \quad \Leftrightarrow \quad \langle f(i), f(j) \rangle \in E(H)$$

Definition 45

A **bipartite graph** is a graph whose vertices can be partitioned into two sets $L(G)$ and $R(G)$ such that every edge has one endpoint in $L(G)$ and the other endpoint in $R(G)$.

Definition 46

A **matching** in a graph G is a set M of edges of G such that no vertex is an endpoint of more than one edge in M. In other words every vertex can only be found in 1 edge. A matching is said to **cover** a set S of vertices if and only if each vertex in S is an endpoint of an edge of the matching. A matching is said to be **perfect** if it covers $V(G)$.

Example 47

Consider the following bipartitte graph G where the set of men M make up $L(G)$ while the set of women W make up $R(G)$

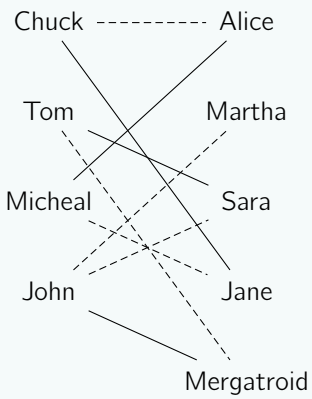


Figure 5: A graph where an edge between a man and woman denotes that the man likes the woman

- The bolded edges represent one possible matching
- This matching *covers* the set $V(G)$ – Mergatroid.
- This is not a *perfect* matching
- Notice that every subset of men likes at least as many women

Definition 48

In any graph G the set $N(S)$ of **neighbors** of some set S of vertices is the image of S under the edge relation that is

$$N(S) = \{r | \langle s - r \rangle \in E(G) \text{ for some } s \in S\}$$

S is called a **bottleneck** if

$$|S| > |N(S)|$$

Theorem 49 (Hall)

Let G be a *bipartite graph*. There is matching in G that covers $L(G)$ if and only if no subset of $L(G)$ is a bottleneck.

Proof. See your `Combinatory Analysis 18.211` notes under the graph theory section.

5 sums and asymptotics

5.1 stirling formula

Fact 50

For all $n \geq 1$

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\varepsilon(n)}$$

where

$$\frac{1}{12n+1} \leq \varepsilon(n) \leq \frac{1}{12n}$$

This can be derived by elementary calculus so we are not too interested. Just noitce that by squeeze theorem clearly $\lim_{n\to\infty} \varepsilon(n) = 0$. Therefore

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = \lim_{n\to\infty} e^{\varepsilon(n)} = 1$$

so

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

5.2 asymptotic notation

Definition 51 (Asymtotic equality)

For functions $f, g : \mathbb{R} \to \mathbb{R}$ we say f is **aysmptotically equal** to g in symbols

$$f(x) \sim g(x)$$

if and only if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$$

Definition 52 (little O)

For functions $f, g : \mathbb{R} \to \mathbb{R}$ with g nonnegative we say f is **asymtotically smaller** than g in symbols

$$f(x) = o(g(x))$$

if and only if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$$

Example 53

Consider

$$1000x^{1.9} = o(x^2)$$

because $1000x^{1.9}/x^2 = 1000/x^{0.1}$ and since $x^{0.1}$ goes to infinity as $x \to \infty$

Corollary 54

$x^a = o(x^b)$ for all nonnegative constants $a < b$.

Lemma 55

$$\log x = o(x^{\varepsilon})$$

for all $\varepsilon > 0$

Proof. Choose $\varepsilon > \delta > 0$ and let $x = z^{\delta}$ in the inequality $\log x < x$ for all $x > 1$ (recall it is the inverse of $x^1 0$ and so it the bottom reflection about $y = x$). Then we have

$$\log z < z^{\delta}/\delta = o(z^{\varepsilon})$$

which follows by 54

Definition 56 (Big O)

Given functions $f, g : \mathbb{R} \to \mathbb{R}$ with g nonnegative we say that

$$f = O(g)$$

if and only if

$$\limsup_{x \to \infty} |f(x)| / g(x) < \infty$$

We use limsup because lim doesnt always exist. Recall that limsup always exists on the extended real line from real analysis.

Lemma 57

If $f = o(g)$ or $f \sim g$ then $f = O(g)$

Proof. $\lim \frac{f}{g} = 0$ or $\lim \frac{f}{g} = 1$ implies $\lim \frac{f}{g} < \infty$

Lemma 58

If $f = o(g)$ then it is not true that $g = O(f)$

Proof. Consider

$$\lim_{x \to \infty} \frac{g(x)}{f(x)} = \frac{1}{\lim_{x \to \infty} \frac{f(x)}{g(x)}} = \frac{1}{0} = \infty$$

Definition 59

Given functions $f, g : \mathbb{R} \to \mathbb{R}$ with g nonnegative we say

$$f = O(g)$$

if and only if there exists a constant $c \geq 0$ and an x_0 such that for all $x \geq x_0, |f(x)| \leq cg(x)$

Proposition 60

$$100x^2 = O(x^2)$$

Proof. Choose $c = 100$ an $x_0 = 1$. Then for all

$$x \geq, \left|100x^2\right| \leq 100x^2$$

which is true

Proposition 61

$$x^2 + 100x + 10 = O(x^2)$$

Proof. consider

$$(x^2 + 100x + 10)/x^2 = 1 + 100/x + 10/x^2$$

so its limit as $x \to \infty$ is 1. So in fact $x^2 + 100x + 10 \sim x^2$ and so the proposition follows

Corollary 62

$$a_k x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0 = O(x^k)$$

Definition 63 (Theta)

$$f = \Theta(g)$$

if and only if

$$f = O(g) \text{ and } g = O(f)$$

Definition 64 (Omega)

Given functions $f, g : \mathbb{R} \to \mathbb{R}$ wth f nonnegative define

$$f = \Omega(g)$$

to mean

$$g = O(f)$$

Problem 65

Show that

$$\ln(n^2!) = \Theta(n^2 \ln n)$$

Hint: logarithmic functions with factorial variable inside is an immediate cue to use stirling formula...

Solution. By stirling formula we have

$$(n^2)! = \sqrt{2\pi n^2}\left(\frac{n^2}{e}\right)^{n^2} e^{\varepsilon(n^2)}$$

where $\frac{1}{12n^2+1} \leq \varepsilon(n^2) \leq \frac{1}{12n^2}$(it is clear that $\varepsilon(n^2) \to 0$ as $n \to \infty$ by squeeze theorem) Upon simplification we have

$$\ln(n^2!) = \frac{1}{2}(\ln 2\pi) + \ln(n) + n^2 \ln\left(\frac{n^2}{e}\right) + \varepsilon(n^2)\ln(\varepsilon)$$

then diving both sides by $n^2 \ln n$ and taking limits we have

$$\lim_{n\to\infty} \frac{\ln(n^2)}{n^2 \ln(n)} = \lim_{n\to\infty} \frac{\ln(2\pi)}{2n^2 \ln(n)} + \lim_{n\to\infty} \frac{1}{n^2} + 2 - \lim_{n\to\infty} \frac{1}{\ln n} + \lim_{n\to\infty} \frac{\varepsilon(n^2)}{n^2 \ln(n)} = 2$$

so we have

$$\lim_{n\to\infty} \frac{n^2}{\ln(n^2!)} = \frac{1}{2}$$

which also means

$$\lim_{n\to\infty} \frac{\ln(n^2!)}{n^2} = \frac{2}{1}$$

so by definition big theta we have the shown the relation follows

Problem 66

Without using stirling gormula show that

$$\log(n!) = \Theta(n \log n)$$

Solution. First notice that

$$\log(n!) = \sum_{i=1}^{n} \log(i) < \sum_{i=1}^{n} \log(n) = n \log(n)$$

On the other hand see that

$$\begin{aligned}
\log *N!) &= \sum_{i=1}^{n} \log(i) \\
&> \sum_{\lceil (n+1)/2 \rceil}^{n} \log(i) \\
&> \sum_{\lceil (n+1)/2 \rceil}^{n} \log(n/2) \\
&> \frac{n}{2} \cdot \log(n/2) \\
&= \frac{n \log n}{2} - \frac{n}{2} \\
&> \frac{n \log n}{2} - \frac{n \log n}{6} \quad (\text{for } n > 8) \\
&= \frac{n \log n}{3}
\end{aligned}$$

where the in second line $\sum_{\lceil (n+1)/2 \rceil}^{n}$ essentially sums the upper half of the terms. And so the third line follows because every single upper half $i > \frac{n}{2}$. The rest follow by basic properties of log. Therefore

$$\frac{1}{3} n \log n < \log(n!) < n \log(n)$$

for $n > 8$ proving the proposition

Problem 67

Show that

$$\sum_{k=1}^{n} k^6 = \theta(n^7)$$

Solution. By definition of reimann integral and that because we are using positive terms we have

$$\sum_{k=1}^{n} k^6 x \geq \int_0^n x^6 dx$$

This sum clearly takes the sup of all partitions $x_0 = 0, 1, 2, \ldots, n = x_n$ so by definition the riemann integral is the infinum of it. Now on the other hand

$$\sum_{k=0}^{n-1} k^6 \leq \int_0^n x^6 dx$$

This sum clearly takes the inf over the same partitions$x_0 = 0, 1, 2, \ldots, n = x_n$ so by definition the riemann integral is the suprenum of it. By a change of variable shifting n by 1 we immediately see the 2nd inequality.

$$\sum_{k=1}^{n} k^6 \leq \sum_{k=0}^{n} k^6 \leq \int_0^{n+1} x^6 dx$$

Then knowing that we are summing positive terms the 1st equality follows. Now comhing everything we have

$$\int_0^n x^6 dx \leq \sum_{k=1}^{n} k^6 \leq \int_0^{n+1} x^6 dx$$

so we have

$$\frac{n^7}{7} \leq \sum_{k=1}^{n} \leq \frac{(n+1)^7}{7}$$

so

$$\frac{1}{7} \leq \sum_{k=1}^{n} \leq \frac{(n+1)^7}{7n^7}$$

but notice that

$$\lim_{n\to\infty} \frac{(n+7)^7}{7n^7} = \frac{1}{7}$$

just consider the binomial expansion and think of the only term that will not have be a reciprocal of a power of n. Therefore by squeenze theorem

$$\lim_{n\to\infty} \frac{\sum_{n=1}^{n} k^6}{n^7} = \frac{1}{7}$$

and so

$$\lim_{n\to\infty} \frac{n^7}{\sum_{n=1}^{n} k^6} = \frac{7}{1}$$

This proposition clearly follows from here

Problem 68

Indicate the implications among the following 5 assertions

1. $f \sim g$
2. $f = o(g)$
3. $f = O(g)$
4. $f = \Theta(g)$
5. $f = O(g)$ AND NOT $g = O(f)$

Solution. Recall above we know (1) implies (3)

(1) implies (4): if $\lim \frac{f}{g} = 1$ then $f = O(g)$ and $g = O(f)$(just take $C = 1$) so $f = \Theta(g)$. However (4) does not imply (1): consider $x = \Theta(2x)$ but $\frac{1}{2} \neq 1$ so $x \not\sim 2x$

(2) implies (3) but (3) does not imply (2). Quite obvious

(2) and (5) are equivalent: $\lim \frac{f}{g} = 0$ if and only if $\frac{f}{g}$ is finite and $\lim \frac{g}{f} = \infty$

(4) implies (3) by definition of big-Theta but obviously (3) does not imply (4). Again consider $x = O(x^2)$ but $x \neq \Theta(x^2)$ since

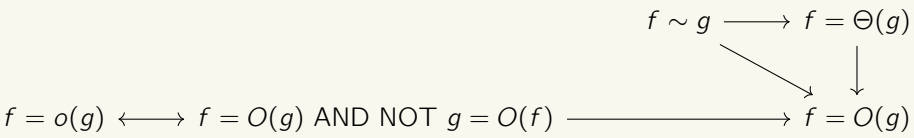
$$c_1 x^2 \leq x \leq c_2 x^2$$

you can find a c_2 but not c_1 because x^2 grows much faster than x as $x \to \infty$.

(5) clearly implies (3) but not vice versa.

Fact 69

In summary



6 Random Variables

Definition 70

A **random variable** R on a probability space is a total function whose domain is the sample space

Example 71

Random variables are actually functions. Consider the following sample space

$$\mathcal{S} = \{HHH.HHT, HTH, HTT, THHTHT, TTH, TTT\}$$

Then a random variable say C could be one that counts the number of heads that appear

$$C(HHH) = 3 \quad C(THH) = 2$$
$$C(HHT) = 2 \quad C(THT) = 1$$
$$C(HTH) = 2 \quad C(TTH) = 1$$
$$C(HTT) = 1 \quad C(TTT) = 0$$

Example 72

$$[C = 2] = \{THH, HTH, HHT\}$$

and this event has probability

$$Pr[C = 2] = Pr[THH] + Pr[HTH] + Pr[HHT] = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

Definition 73

If R is a random variable defined on a sample space $\mathcal{S}$ then the **expectation** of R is

$$Ex[R] = \sum_{w \in S} R(w) Pr[w]$$

Alternatively

$$Ex[R] = \sum_{r \in \text{range}(R)} x \cdot Pr[R = x]$$

Definition 74

The **conditional expectation** $Ex[R|A]$ of a random variable R given event A is

$$Ex[R|A] = \sum_{r \in \text{range}(R)} r \cdot Pr[R = r|A]$$

7 Recurrences

Proposition 75

$T_n = 2^n - 1$ satisfies the recurrence:

$$T_1 = 1$$
$$T_n = 2T_{n-1} + 1 \quad (\text{for } n \geq 2)$$

Proof. By induction on n