Graphs, G is a pair of sets (V, E) where V is a nonempty set of items called vertices or nodes, and E is a set of 2 item subsets of V, called edges Ex: $G=(V,E)$, $V=\{x_1,x_2,x_3\}$, $E=\emptyset$, x_i and x_j are adjacent if connected by a node

an edge is incident to its endpoints. # edges incident to a node is the degree of a node.

Simple graph has no loops or multiple edges loop: mult. edge

Graph coloring Given graph G and K colors, assign colors to each node s.t. adjacent nodes get different colors

Def: minimum K is the Chromatic number of G (χ_G) - no fast algorithm for this - exponential

NP complete if you solve 1 you solve them all, central problem in CS theory (also #1 million problems)

Basic coloring algo: 1) order nodes $V_1, \dots, V_n$ 2) order colors $C_1, \dots$ 3) For $i=1,\dots,n$ assign lowest legal color

LD If every node in G has degree $\le d$, basic alg uses at most $d+1$ colors for G.

Generally induction do for nodes, then edges, then degree if previous doesn't work

Def: Graph $G=(V,E)$ is bipartite if V can be split into V_L and V_R st all edges connect a node in V_L to V_R can be in 2 colors

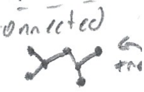
Minimum spanning tree: Walk: sequence of vertices that are connected by edges $V_0 - V_1 - \dots - V_k$ start end } length k

Paths: a walk where all vertices are different, Lemma if there is a walk from V_1 to V_k, there is a path from V_1 to V_k LD shortest walk is a path.

Def: u and v are connected if there is a path from u to v.) graph is connected if every pair of vertices are connected

Closed walk is a walk that starts and ends at the same vertex $(V_0 - V_1 - V_2 - V_1 - V_0)$

LD if $k \ge 3$ and $V_0, V_1, \dots V_{k-1}$ are all different then this is a cycle



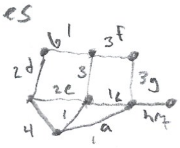
Tree: graph that is connected and does not have any cycles → Leaf node w degree 1 on a tree. taking a smaller part of something w no cycles can't have one

LD any connected subgraph of a tree is a tree. (induction, remove leaf from n+1)

LD tree w n vertices has n-1 edges

Spanning tree: (ST) of a connected graph is a subgraph that is a tree and spans all vertices

LD Every connected graph has a spanning tree.



LD Every edge has a weight. This also works

Weighted Spanning tree: spanning tree where each edge has smallest possible sum of edge weights

LD Min. Span. tree of weighted graph is ST st it has smallest possible sum of edge weights choose in alg

Algo: grow subgraph 1 edge at a time at each step; add the minimum weight edge that keeps the graph acyclic, ex → tree of n verts has n-1 edges.

LD For any connected weighted graph G, algo produces a min. weight. span. tree

If less than n-1 edges have been placed, ∃ an edge that when placed does not create a cycle.

6.3 Communication Networks □ = terminal: source and destination of data. D=6

Complete Binary Tree ○ = switch: direct packets through network

Latency: time required for packet to travel from input to output. , Diameter, longest route from input to output

Switch size: # inputs x # outputs

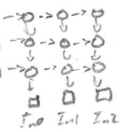
A permutation is function $\pi: \{0,\dots,N-1\} \to \{0,\dots,N-1\}$ st. no two numbers are mapped to the same value $(\pi(i)=\pi(j) \iff i=j)$

Permutation Routing Problem: $\forall i$, direct packets at In_i to $Out_{\pi(i)}$: path taken denoted by $P_{i,\pi(i)}$

The congestion of paths $P_{0,\pi(0)}, \dots, P_{N-1,\pi(N-1)}$ is equal to the largest # of paths that pass through a single switch

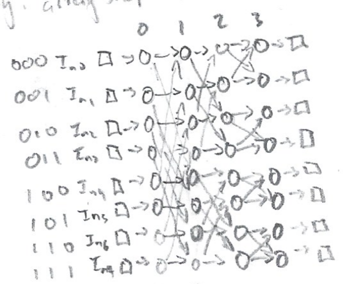
Max congestion (worst case): max over all permutations , min would be solution to routing problem.

2D Array: Thm: congestion of N-input array is 2 (path follow row of input to column of output)



Butterfly: array shape w free flow at each switch, path is either to top or bottom butterfly network.

Which row does it point to? 1) same row 2) change column th # in binary



0th col	1st col	2nd col
$000 \to 000$, $000 \to 100$	$000 \to 000$, $000 \to 010$	$000 \to 000$, $000 \to 001$
$101 \to 101$, $101 \to 001$	$101 \to 101$, $101 \to 111$	$101 \to 101$, $101 \to 100$

$(b_1, \dots, b_{\log N}, l)$ bits in binary row, level → straight $(b_1, \dots, b_{\log N}, l+1)$; other: $(b_1, \dots, \bar{b}_{l+1}, \dots, b_{\log N}, l+1)$

$(x_1, \dots, x_{\log n}, 0) \to (y_1, x_2, x_3, \dots, x_{\log n}, 1) \to \dots (y_1, \dots, y_k, x_{k+1}, \dots, x_{\log n}, k) \to \dots (y_1, y_2, \dots, y_{\log n}, \log n)$

$\downarrow$

$(y_1, \dots, y_{\log n}, \log n)$

$In \to 1 \to 2 \to 3 \to 2 \to 1 \to out$ eliminates congestion.

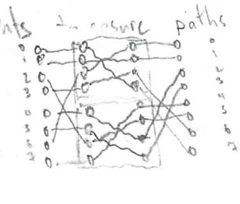
Benes Network: butterfly but adding opposite after, so butterfly rows inside them (good for induction showing congestion = 1) ... the smaller B

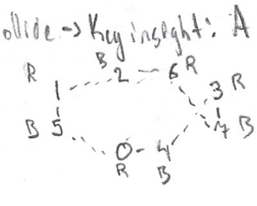
LD big Benes networks have smaller ones nested inside them

→ Thm: congestion of N-input Benes network is 1, when $N = 2^a$ for $a \ge 1$

Need constraints to ensure paths don't collide → Key insight: A 2-coloring of constraint graph.

Permutation
$\pi(0) \to 7$ $\pi(4) \to 2$
$\pi(1) \to 1$ $\pi(5) \to 3$
$\pi(2) \to 5$ $\pi(6) \to 0$
$\pi(3) \to 6$ $\pi(7) \to 4$





— Based on inputs

.... based on outputs

(summary chart next page)

(1,5) can't both go to same switch so make an edge between them

2nd last row can feed 2 diff. outer make sure each one only has 1 packet (output are cong. module

Matching Problems

Def: Given graph $G(V,E)$ a matching is a subgraph where every node has degree 1.

↳ **Def**: matching is perfect if it has size $|V|/2$ (everything paired)

Sometimes some matches are more desireable (weighted matching), generally low weight → more desirable.

'Goal is to cre[ate] a stable perfe[ct] match

Def: Weight of a matching M is the sum of weights on its edges.

Def: Min-weight matching is a perfect match w/ minimum weight.

Sometimes priority instead of weights

Def In matching M, x and y form a rogue couple if they both prefer each other to their mates

Def: matching is stable if there aren't any rogue couples

Thm ends in N^2+1 days, any day it doesn't end, 1 person crossed o[ff]

The matching algo, serenade under the balcony...

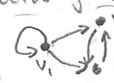
Def: optimal mate is their favourite from the real of possibility (no rogue mates)

Thm: TMA marries every male w/ his optimal mate, and every female w/ her pessimal mate

Def Euler Tour: walk that traverses every edge exactly once and starts & finishes at the same vertex

Thm: A connected graph has an Euler Tour if every vertex has even degree.

Directed Graph (digraph)



tail → head

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array}\begin{array}{c} v_1\ v_2\ v_3 \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\end{array} \quad A^3 = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

P^3_{32}

Thm: $G=(V,E)$ be n-node graph w/ $V=\{v_1,\dots,v_n\}$, let $A=\{a_{ij}\}$ denote the adjacency matrix for G. That is: $a_{ij} = \begin{cases} 1 & \text{if } \exists \text{edge } v_i \to v_j \\ 0 & \text{otw} \end{cases}$

Thm: Let $P^{(k)}_{ij}$ = # directed walks of length k from v_i to v_j $\Rightarrow A^k = \{P^{(k)}_{ij}\}$

$$\begin{array}{c} \text{from} \\ v_1 \\ v_2 \\ v_3 \end{array}\begin{array}{c} \text{to } v_1\ v_2\ v_3 \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\end{array} \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$\#v_1 \to v_1$ in 2's:
$(\#v_1\to v_1)(\#v_1\to v_1)$
$+(\#v_1\to v_2)(\#v_2\to v_1)$
$+(\#v_1\to v_3)(\#v_3\to v_1)$

A digraph is strongly connected if $\forall u,v \in V$, $\exists$ a directed path from u to v in G

Def A directed acyclic graph (DAG) is a directed graph with no cycle

Def: Directed Hamiltonian path: directed walk that visits every vertex exactly once

not a cycle

Directed Hamiltonian Path.

Thm: Every tournament graph contains a directed Hamiltonian Path.

↳ (induction) if nth is not either wins all (front), loses all (back), or you are guaranteed to be able to jam it into the path somewhere (cool to think about)

u beats v
$u \to v$
v beats u
$v \to u$

Tournament where chicken u pecks chick v or chick v pecks chick u.

u virtually pecks v if: $u \to v$ or $\exists w$ st $u \to w \to v$ (King if virtually pecks everyone else)

↑only 1 chicken

Thm: Chicken w/ highest outdegree is king.

Relations: Relation from a set A to set B is subset $R \subseteq A \times B$ ($(a,b)\in R$ or aRb or a[...])

A relation on A is a subset $R \subseteq A \times A$ (ex: $A=\mathbb{Z}$: xRy iff $x \equiv y \pmod 5$)

Set A together w/ R is a directed graph ($V=A$, $E=R$)

Properties: Relation R on a: • reflexive if $xRx\ \forall x \in A$ • symmetric if $xRy \Rightarrow yRx\ \forall x,y \in A$

• antisymmetric: if $xRy \wedge yRx \Rightarrow x=y$ • transitivity: if $xRy \wedge yRz \Rightarrow xRz$

Ex:

	Reflexive	Symmetric	Antisym
$x \equiv y \pmod 5$	✓	✓	x
$x \mid y$	✓	x	✓
$x \leq y$	✓	x	✓

equiv — partial order

Equivalence Relation: it is reflexive, symmetric and transitive (ex: equality, modulo $\equiv$)

↳ The equivalence class of $x \in A$ is the set of all elements in A related to x by R, denoted $[x]$ $[x] = \{y : xRy\}$

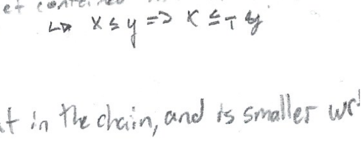
↳ Equivalence classes partition sets **Thm**: The equivalence class of an equivalence relation on a set A form a partition of A

(Weak) Partial Order: relation reflexive, antisymmetric and transitive (denoted w/ $\leq$) $(A, \leq)$

↳ Poset (Part. Order set) is a directed graph w/ vertex set A and edge set $\leq$

Hesse Diagram for poset $(A,\leq)$ is the directed graph w/ vertex set A, edge set $\leq$ but without:

1) all self loops 2) edges implied by transitivity



Thm: Poset has no directed cycles other than self loops.

↳ Deleting self-loops from a poset makes a directed acyclic graph (DAG).

Partial order can have items that are incomparable, comparable $\Rightarrow a \leq b$ or $b \leq a$

↳ Total order: Partial order in which every pair of elements is comparable. ... (other set contained in total order)

Topological sort: of a poset $(A,\leq)$ is a total order $(A, \leq_T)$ such that $\leq \subseteq \leq_T$ ↳ $x \leq y \Rightarrow x \leq_T y$

↳ Every finite poset has a topological sort.

↳ $x \in A$ is minimal of A if $\neg\exists y \in A, y \neq x$ st. $y \leq x$, $x \in A$ is maximal .... $x \leq y$

Parallel Scheduling

Def: A chain is a sequence $a_1 \leq a_2 \leq \dots \leq a_t$, $a_i \neq a_j$, st each item is comparable to the next in the chain, and is smaller w[...]

↳ Length of the chain is t, # of elements in chain.

↳ We can partition A into t subsets

↳ Total amount of parallel time needed to complete is same length as longest chain (critical path)

Def antichain is set of elements in poset is a set st. any 2 elements in the set are uncomparable.

Cor: if largest chain is t, A can be partitioned into t antichains

Lemma (Dilworth): $\forall t>0$ every partially ordered set w/ n elements must have must have $\underbrace{\text{chain leng.} > t}_{\text{bins}}$ or $\underbrace{\text{antichain} \geq n/t}_{\text{\# per bin}}$

Sums | Perturbation Method:

$S = 1 + x + x^2 + \cdots + x^{n-1}$

$- xS = \quad x + x^2 + \cdots + x^{n-1} + x^n$

$(1-x)S = 1 - x^n$

$\Rightarrow S = \frac{1-x^n}{1-x}$

Geometric Series $\sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}$, $n = \infty$ $\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}$, $|x|<1$ (top term goes to 0) mortgage example (fixed rate)

$\sum_{i=0}^{n-1} i x^i = \frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2}$, $\sum_{i=0}^{\infty} i x^i = \frac{1}{1-x^2}$, $|x|<1$ (payment growing every yr)

both using $x = \frac{1}{1+r}$, so discounting TVMoney each payment i yrs in the future

generally summing something to the n will give n a 1 higher than what is in the sum

Derivative Method:

$\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x}$

$\sum_{i=0}^{n} i x^{i-1} = \frac{1-(n+1)x^n + nx^{n+1}}{(1-x)^2}$

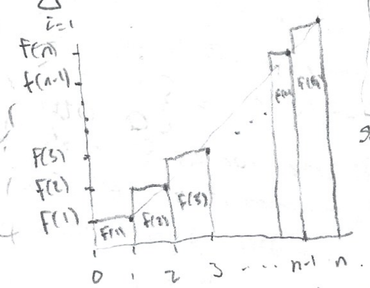
$\sum_{i=0}^{n} i x^i = \frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2}$

Def: $g(x) \sim h(x)$ means $\lim_{x\to\infty} \frac{g(x)}{h(x)} = 1$

Arithmetic Series: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$, $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$ ex: $\sum_{i=1}^{n} i^2$ summing n^2 to n $= n^3$

Integration Bounds For $\sum_{i=1}^{n} f(i)$ when $f(i)$ is positive increasing function

$\sum_{i=1}^{n} \sqrt{i}$



$\sum_{i=1}^{n} f(i) \geq F(1) + \int_1^n F(x)\,dx$

$\sum_{i=1}^{n} f(i) \leq f(n) + \int_1^n f(x)\,dx$

so example $F(i) = \sqrt{i}$

$\int_1^n \sqrt{x}\,dx = \left.\frac{x^{3/2}}{3/2}\right|_1^n$

$= \frac{2}{3}(n^{3/2} - 1)$

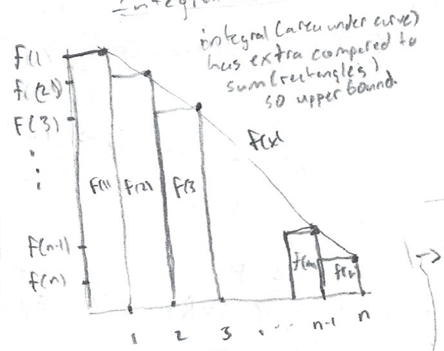
$\rightarrow \sqrt{1} + \frac{2}{3}(n^{3/2}-1) \leq \sum_{i=1}^{n} \sqrt{i} \leq \sqrt{n} + \frac{2}{3}\ldots$

bounds only differ by $\sqrt{n} - 1$ while sum is growing w/ $n^{3/2}$

We can write $\sum_{i=1}^{n} \sqrt{i} = \frac{2}{3}n^{3/2} + \delta(n)$, $\frac{1}{3} \leq \delta(n) \leq \sqrt{n}$

$\sum_{i=1}^{n} \sqrt{i} \sim \frac{2}{3} n^{3/2}$

Integration bounds for $\sum f(i)$, positive decreasing function



$\sum_{i=1}^{n} f(i) \leq f(1) + \int_1^n f(x)\,dx$

$\geq f(n) + \int_1^n f(x)\,dx$

decreasing function now has the bounds swapped

Factorial! $n! = \prod_{i=1}^{n} i$

$\ln(n!) = \ln(1 \cdot 2 \cdot 3 \cdots n)$

$= \ln(1) + \ln(2) + \cdots + \ln(n)$

$= \sum_{i=1}^{n} \ln(i)$

now increasing, use integration bounds, get bounds then e^(both sides)

$\frac{n^n}{e^{n-1}} \leq n! \leq \frac{n^{n+1}}{e^{n-1}}$

Harmonic Sum: $H_n = \sum_{i=1}^{n} \frac{1}{i}$ nth harmonic

$\hookrightarrow H_1 = 1$, $H_2 = 1 + \frac{1}{2} = \frac{3}{2}$

$\rightarrow$ no closed form solution, so use integration bounds

$f(n) + \int_1^n f(x)\,dx \leq H_n \leq f(1) + \int_1^n f(x)\,dx$

$\Rightarrow \frac{1}{n} + \ln(n) \leq H_n \leq 1 + \ln(n)$

homework: $H_n \sim \ln(n)$, $H_n = \ln(n) + \delta + \frac{1}{2n} + \frac{1}{12n^2} + \ldots$

δ = Euler's const: 0.57721…

Stirling's Formula (very good estimate for factorial)

$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\, e^{\varepsilon(n)}$, $\frac{1}{12n+1} \leq \varepsilon(n) \leq \frac{1}{12n}$

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

Note → turn products into sums by taking a ln()

$f(x) = O(g(x))$ iff $g(x) = \Omega(f(x))$

Note: never use asymptotic notation for induction proofs! Fixing n makes functions now scalars so function were O(n) are now O(1).

Asymptotic notation

1) tilde : $f(x) \sim g(x)$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$ — f(x) grows at least as slow as g(x)

2) Oh, big oh : $f(x) = O(g(x))$ if $\lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty$ (finite)

$\hookrightarrow$ used as upper bound.

3) omega : $f(x) = \Omega(g(x))$ if $\lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| > 0$

$\hookrightarrow$ upper bound

4) theta : $f(x) = \Theta(g(x))$ if $0 < \lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty$

$\hookrightarrow$ both theta and omega, equality

5) little oh : $f(x) = o(g(x))$ if $\lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| = 0$ strictly less than

little omega : $f(x) = \omega(g(x))$ if $\lim_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| = \infty$ strictly greater than

Divide and concur recurance

Towers of Hanoi - minimum time to move tower w/ n disks is T_n also $T_n = 2T_{n-1} + 1$

Methods for getting closed form solution (equations) for reccurances:

1) Guess and Verify (substitution) - w/ induction. $P(n) = T_n = 2^n - 1$, Base: $T_1 = 1 = 2^1 - 1 = 1$ ✓ Induct given $T_n = 2^n - 1$, show $T_{n+1} = 2^{n+1} - 1$
   ↳ good method bug requires a divine guess

don't collapse things along the way, helps to see pattern.

2) Plug & Chug: $T_n = 1 + 2T_{n-1} \Rightarrow T_n = 1 + 2(2T_{n-2} + 1) = 1 + 2 + 2T_{n-2} + 1 + 2 + 2(2T_{n-3} + 1) = 1 + 2 + 4 + T_{n-3}$
   notice a pattern, $1 + 2 + 4 + \dots + 2^{n-2} + 2^{n-1}T_1 \Rightarrow 2^n - 1$
   ↳ go till the end or notice a pattern

Merge Sort

To sort $n \geq 1$ $x_1, x_2, \dots, x_n$ (n = power of 2) recursively

1) sort $x_1, x_2, \dots, x_{n/2}$ & $x_{n/2+1}, x_{\frac{n}{2}+2}, \dots, x_n$

2) merge.

Ex: $\{10, 7, 23, 5, 2, 4, 3, 9\}$

1) $\{5, 7, 10, 23\}$, $\{2, 3, 4, 9\}$

2) work through comparing lowest

$\{2, 3, 4, 5, 7, 9, 10, 23\}$

Let $T(n)$ = # comparisons used by merge sort to sort n #'s (worst case)
↳ merging takes $n-1$ comparisons.
↳ $2T(n/2)$ comparisons for recursive sorting

$T(n) = 2T(n/2) + n - 1$

try guessing but cant... try plug & chug

$T(n) = n - 1 + 2T(\frac{n}{2})$
$= n - 1 + 2(\frac{n}{2} - 1 + 2T(\frac{n}{4})) = n - 1 + n - 2 + 4T(\frac{n}{4})$
$= n - 1 + n - 2 + 4(\frac{n}{4} - 1 + 2T(n/8))$
$= n - 1 + n - 2 + n - 4 + 8T(\frac{n}{8})$
$= n - 1 + n - 2 + n - 3 + \dots + n - 2^{i-1} + 2^i T(n/2^i) \to 0$
$= n - 1 + n - 2 + \dots + n - 2^{\log n - 1} + 2^{\log n} T(1)$
$= \sum_{i=0}^{\log n - 1} (n - 2^i) = \sum_{i=0}^{\log n - 1} n - \sum_{i=0}^{\log n - 1} 2^i$
$= \underbrace{n \log n}_{\text{sum } n \text{ log n times}} - (2^{\log n} - 1) = n \log n - n + 1$

Divide and conquer is $T(n)$ can be written in term of $aT(x)$ where $a \geq 1$ and x is a factor smaller then n; Splitting into many smaller problems

(look up actuall gross def if you want).

Thm (Akra and Bazzi) Set value of p st $\sum_{i=1}^{k} a_i b_i^p = 1$
$\Rightarrow$ Then $T(x) = \Theta\left(x^p + x^p \int_1^x \frac{g(u)}{u^{p-1}} du\right)$

Thm If $g(x) = \Theta(x^t)$ for $t \geq 0$ & $\sum_{i=1}^{k} a_i b_i^t < 1$
$\Rightarrow$ then $T(x) = \Theta(g(x))$

Linear Recurrences: reccurance of the form
$f(n) = a_1 f(n-1) + a_2 f(n-2) + \dots + a_d f(n-d)$
$= \sum_{i=1}^{d} a_i f(n-i)$, for fixed a_i, d = "order"

Fact: w/out boundary conditions, if $f(n) = \alpha_1^n$ & $f(n) = \alpha_2^n$ are solutions to linear reccurence
$\Rightarrow f = c_1 \alpha_1^n + c_2 \alpha_2^n$ is also a solution ∀ const. c_1, c_2.

Characteristic equation
↳ ?

Tricky case: if α is a root of characteristic eqn, repeated r-times $\Rightarrow \alpha^n, n\alpha^n, n^2\alpha^n, \dots, n^{r-1}\alpha^n$, are sol's to the reccurance

Non-homogenous.

Counting Set: unordered collection of elements, Cardinality ($|S|$) = # of elements

Sequence: ordered collection of elements (not necessarily distinct.

Permutation: sequence that contains every element in the set, permutation of set $\bar{c}$ n elements = $n!$

Function: $F: X \to Y$, relation between X and Y, every element of X mapped to 1 element of Y.

Surjective: Every element in Y is mapped to at least once, Injective: Element of Y mapped to at most once

Bijective: Injective + Surjective.

Mapping Rule: surjective $\Rightarrow |X| \geq |Y|$, injective $\Rightarrow |X| \leq |Y|$, bijective $\Rightarrow |X| = |Y|$

Generalized Pigeonhole Principle: If $|X| > k|Y| \Rightarrow \forall_{f:X \to Y} \exists_{k+1}$ different elements of X that are mapped to same element in Y.

Def: K to 1 function maps exactly k elements of X to every element of Y, for bijection 1-1

Sum rule: disjoint sets $A_1, \dots, A_n \Rightarrow |A_1 \cup A_2 \cup \dots \cup A_n| = |A_1| + \dots + |A_n|$

$|M \cup E| = |M| + |E| - |M \cap E|$, $|M| = |M \cap E| + |M \cap E^c|$

Book keeper rule: distinct # letters $l_1, l_2, \dots$

unique orders $\dfrac{(l_1 + l_2 + \dots + l_n)!}{l_1! l_2! \dots l_n!}$

Counting Guidelines:

1) For $f: A \to B$, check # elements of A mapped to each element of B, then apply the division rule.

2) Try solving problem in another way

Ex: # 2 pair hands $\binom{13}{1}\binom{4}{2}\binom{12}{1}\binom{4}{2}\binom{11}{1}\binom{4}{1}$ however this would count $\{QH, QS, KH, KS, 2D\}$ and $\{KH, KS, QH, QS, 2D\}$ — count $\{K, Q\}$ and $\{Q, K\}$ as the

↳ Thus it double counts pairs (2 to 1) so divide by 2. 2 pairs

⇒ To avoid this, choose values of pairs in the same turn ie $\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{11}{1}\binom{4}{1}$

↳ This is the difference between $\binom{13}{2}$ and $\binom{13}{1}\binom{12}{1}$

* Thm $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$ how many teams of n $\bar{w}$ k players? Guy named bob, teams $\bar{w}$ Bob + teams $\bar{w}$/o Bob = total teams $= \binom{n-1}{k-1} + \binom{n-1}{k}$

Thm $\sum_{r=0}^{n} \binom{n}{r}\binom{2n}{n-r} = \binom{3n}{n}$ 3n balls, n red, 2n blue, r red chosen. Ways to choose ball is # of ways to chose 0 red + w/ 1 red + tw... ways to pick the sets.

↳ tough way to prove this figuring out ways to pick the sets.

Prob: $P(E) = \dfrac{|E|}{|S|}$, $\sum_{w \in S} P(w) = 1$, $P(E) = \sum_{w \in E} P(w)$, $P(A^c) = 1 - P(A)$, disjoint events never independent

$P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ Bayes $P(B|A) = \dfrac{P(A|B)P(B)}{P(A)}$; Independence: $P(A|B) = P(A)$, K-way independent iff every set of K elements from set are independent

Independent iff $P(A \cap B) = P(A) \cdot P(B)$

↳ does not imply mutual independence.

Sampling values can give misleading results ex: sampling latency times gives an answer for expected value but if we know $Pr[D=i] = \begin{cases} 0 & i=0 \\ \frac{k}{i} - \frac{k}{i+1} & i \in \mathbb{N}^+ \end{cases} \Rightarrow \sum_{i \in \mathbb{N}^+} i\left(\frac{k}{i} - \frac{k}{i+1}\right)$ gives an expected value that is infinite

↳ sampling can be misleading!

Law Tot Exp: Suppose $A_1, \dots, A_n$ partition $S \Rightarrow E(R) = \sum_{A_i} E[R|A_i] \Pr(A_i)$

Functions $Ex[F(R)] = \sum_{w \in S} F(R(w)) P(w)$

Linearity of expectation: $E[R_1 + R_2] = E[R_1] + E[R_2]$, $E[\alpha_1 R_1 + \alpha_2 R_2] = \alpha_1 E[R_1] + \alpha_2 E[R_2]$

infinite: if $\sum_{i=0}^{\infty} E[|R_i|]$ converges $\Rightarrow \sum_{i=0}^{\infty} E[R_i] = E\left[\sum_{i=0}^{\infty} R_i\right]$

For independent vars: $E[R_1 \cdot R_2] = E[R_1] \cdot E[R_2]$ otherwise $E[R_1] \cdot E[R_2] \neq E[R_1 \cdot R_2]$

$Var(R) = E[R^2] - E^2[R]$ $Var(aX + b) = a^2 Var(X)$, $Var(X+Y) = Var(X) + Var(Y)$

Markov's Thm: If R is non-neg. rand. var $\Rightarrow \forall x>0$ $P(R \geq x) \leq \frac{Ex[R]}{x}$

Cor: R non-neg. r.v. $\Rightarrow \forall c>0$ $P(R \geq c\,Ex(R)) \leq 1/c$

Cor: If $R \leq u$ for some $u \in \mathbb{R}$ (upper bound) then $\Rightarrow \forall x<u.\ P(R \leq x) \leq \frac{u-Ex[R]}{u-x}$ (markov's thm is a shift) bounded shift by lowest value (same then markov's)

Chebyshev's Thm $\forall x>0$ & rand. var R: $P(|R-Ex[R]| \geq x) \leq \frac{Var(R)}{x^2}$

Cor: $P(|R-Ex[R]| \geq c\sigma(R)) \leq \frac{Var(R)}{c^2\sigma(R^2)} = \frac{1}{c^2}$

more precise bounds then markov because it requires info about the dist to the var.

Thm For any R.V. R: $P(R-Ex(R) \geq c\sigma(R)) \leq \frac{1}{c^2+1}$, $P(R-Ex(R) \leq -c\sigma(R)) \leq \frac{1}{c^2+1}$ } slight improvement on previous thm if you only want higher or lower.

Thm (Chernoff Bound): Let $T_1, \dots, T_n$ be mutually ind. R.V.'s st. $\forall i$ $0 \leq T_i \leq 1$

Let $T = \sum_{i=1}^{n} T_i$; then $\forall c>1$, $P(T \geq c\,Ex(T)) \leq e^{-z Ex[T]}$ *bound on sums of rnd var*

where $z = c\ln(c) + 1 - c > 0$

Intuition: probability you are high is exponentially small.

Chernoff will give much better bounds then Markov, this is because it requires the variables to be independent

Gambler's Ruin

win 1\$ w prob p, lose \$1 w prob $(1-p)$

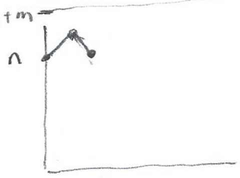
start w n dollars, each bet

Play until win \$m or lose \$n

start w \$1000 trying to win \$100 $(m=100,\ n=1000)$ ↳ almost guaranteed to lose.

Roulette: $p = \frac{18}{38} = 0.473$

Random walk: prob up move $= p$, prob down move $= 1-p$ } mut. independent of past moves ↳ martingale.



if $p = 1/2$ "unbiased", $p \neq 1/2$ "biased"

Def. W^* = event hit \$$T = n+m$ before 0

D = # dollars at start

$X_n = P(W^* \mid D=n)$

Claim $X_n = \begin{cases} 0 & n=0 \\ 1 & n=T \\ pX_{n+1} + (1-p)X_{n-1} & \text{if } 0<n<T \end{cases}$

E = event you win first

$\bar{E}$ = lose first bet

$X_n = P(W^* \cap E \mid D=n) + P(W^* \cap \bar{E} \mid D=n)$

$= P_r(E \mid D=n)P(W^* \mid E \cap D=n) + P(\bar{E} \mid D=n)P(W^* \mid D=$

$= p \cdot P(W^* \mid D=n+1) + (1-p) \cdot P(W^* \mid D=n-1)$

$\boxed{X_n = p \cdot X_{n+1} + (1-p)X_{n-1}}$ $\Rightarrow X_n - pX_{n+1} - (1-p)X_{n-1} =$ ↳ homogeneous, d

Char eqn: $pr^2 - r + (1-p) = 0 \Rightarrow r = \underbrace{\frac{1-p}{p} \text{ or } 1}_{\text{roots are the same for } p=1/2}$

$\Rightarrow$ If $p \neq 1/2 \Rightarrow X_n = A\left(\frac{1-p}{p}\right)^n + B(1)^n$

$= A\left(\frac{1-p}{p}\right)^n + B$

Boundary conditions $0 = X_0 = A + B \Rightarrow B = -A$

$1 = X_T = A\left(\frac{1-p}{p}\right)^T - A$

If $p < \frac{1}{2} \Rightarrow \left(\frac{1-p}{p}\right) > 1$

$\Rightarrow A = \frac{1}{\left(\frac{1-p}{p}\right)^T - 1}$, $B = \frac{-1}{\left(\frac{1-p}{p}\right)^T - 1}$ $\Rightarrow X_n = \frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^T - 1}$

$\leq \left(\frac{1-p}{p}\right)^{n-T} = \left(\frac{p}{1-p}\right)^m$

Thm: if $p<1/2$, then $\Rightarrow P(\text{win } \$m \text{ before lose } \$n) \le \left(\frac{p}{1-p}\right)^m$

What if you have a fair game? $p=\frac{1}{2}$: $\frac{p}{1-p} = 1$ so double root in characterist...

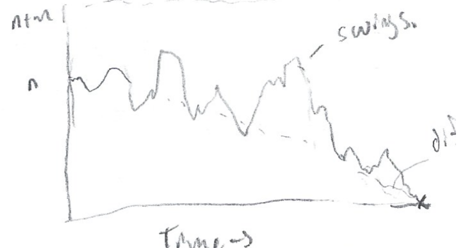
↳ Now characteristic becomes

$$X_n = (An+B)(1)^n$$

w/ boundary cond': $0 = X_0 = B \Rightarrow B=0$

$1 = X_T = AT + B = AT \Rightarrow A = 1/T$

$$X_n = \frac{n}{T} = \frac{n}{n+m}$$



In random walks drift outweighs

Thm: If $p=\frac{1}{2}$ then $P(\text{win } \$m \text{ before lose } \$n) = \frac{n}{n+m}$

$1\cdot p - 1(1-p)$ $\quad 1-2p$ win $\Rightarrow 2p-1$ loss

After x steps, drifted $(1-2p)x$ linear } swing dominated by drift.

Swing is $\Theta(\sqrt{x})$ root

How long do we play?; Def: S = # steps until we hit a boundary. , $E_n = E[S \mid D=n]$

Claim: $E_n = \begin{cases} 0 & n=0 \\ 0 & n=T \\ 1+pE_{n+1}+(1-p)E_{n-1} & \text{if } 0<n<T \end{cases}$

$\Rightarrow pE_{n+1} - E_n + (1-p)E_{n-1} = -1$, $E_0=0, E_T=0$ (Inhomogeneous).

1) First step inhomogeneous is solve homogeneous, we get $E_n = A\left(\frac{1-p}{p}\right)^n + B$ for $p \ne 1/2$.

2) Particular solution: Guess $E_n = a$ Fails!, Guess $E_n = an+b \Rightarrow a = \frac{-1}{2p-1}$ $\quad b=0$

with some solving bdy con....

Gen sol $\Rightarrow E_n = A\left(\frac{1-p}{p}\right)^n + B + \frac{n}{1-2p}$

$$\Rightarrow E_n = \frac{n}{1-2p} - \frac{T}{1-2p}\,\frac{\left(\frac{1-p}{p}\right)^n-1}{\left(\frac{1-p}{p}\right)^T-1}$$

Ex: $m=100$, $n=1000$, $T=1100$, $p=9/19$ $\Rightarrow E[\#\text{ bets to hit boundary}] \approx 19000-0.56=18499.$ long time!

intuition would be $\text{time} = \frac{\text{money start}}{\text{exp loss each bet}}$

as $m\to\infty$, $E_n \sim \frac{n}{1-2p}$ $\quad$ $m\to\infty$ means playing untill you lose all money.