Chapter 1

Pigeonhole Principle

Some Resource: Sami Assaf (432 videos)

Theorem 1

Let n, k be positive integers with $n > k$, Put n balls into k boxes in some what. Then must have at least 1 box with 2 balls *Proof*: Proceed by contradiction: Assume that each of the k boxes has at least 1 ball. Then the number of balls is at most k But we assumed there there are n balls and $n > k$. Contradiction

***Example* (doesn't use pigeonhole principle)** Consider the numbers $1, 2, ..., 2n$. Take any $n + 1$ of them. Then among them there must be two which are relatively prime. *Proof (sketch)*: If have $n + 1$ numbers from $1, 2, 3, ..., 2n$, then must have two consecutive numbers these are relatively prime.

***Example* (does use pigeonhole principle)** Let A be a set of $n+1$ numbers chosen from $1, 2, 3, ..., 2n$. Then there are always two numbers in A such that one divides the other. *Soln* Write each $a \in A$ as $a = 2^k m$ where m is odd and at most $2n - 1$. Call m the "odd part of a". Since A has $n+1$ elements, and there are only m possible odd parts, the pigeonhole principle implies that there are 2 numbers in A with the same odd part, so one of these numbers divides the other.

Example Consider a sequence $a_1, a_2, ..., a_{mn+1}$ of $mn + 1$ distinct real numbers. Then there exists an increasing subsequence $a_{i_1} < a_{i_2} < ... < a_{i_{m+1}}$ (here $i_1 < i_2 < ... < i_{m+1}$) of length $m + 1$ ***OR*** A decreasing subsequence $a_{j_1} > a_{j_1} > ... > a_{j_{n+1}}$ (here $j_1 < j_2 < ... < j_{m+1}$) (or both) *Example* Consider the sequence $4, 1, 3, 5, 7, 8, 2, 6$ The $3, 7, 8$ is an increasing subsequence $4, 3, 2$ is a decreasing subsequence

Remarks For a permutation Π let $I(\Pi)$ = the longest increasing subsequence of Π

1. If Π is a random permutation, then average of $I(\Pi)$ is asymptotic to $2\sqrt{n}$
2. $I(\Pi)$ has applications:
    - patience sorting
    - airline boarding

$n = 3$	$I(\Pi)$
$1, 2, 3$	3
1, 3, 2	2

...

So the average of $I(\Pi) = 2$ for $n = 3$

Claim $\lim_{n\to\infty} \frac{\text{average of } I(\Pi)}{2\sqrt{n}} = 1$

Proof of claim For each number a_i let $t_i =$ the length of the longest increasing subsequence starting at a_i. If $t_i \geq m+1$ for some i, then have an increasing subsequence of length $m+1$ and we're done.

Claim So can assume $t_i \leq m$ for all i, so the function $f : a_i \to t_i$ maps $\{a_i, ..., a_{mn+1}\} \to \{1, ..., m\}$

By generalize pigeonhole principle, there is some $s \in \{1, ..., m\}$ so that $f(a_i) = s$ for $n+1$ numbers a_i. Let $a_{j_1}, a_{j_2}, ..., a_{j_n+1}$ ($j_1 < j_2 < ... < j_{m+1}$) be these numbers. Claim $a_{j_1} > a_{j_2} > ... > a_{j_{n+1}}$ is a decreasing subsequence of length $n+1$. To see this, assume to contrary that $a_{j_i} < a_{j_i+1}$. Then we'd have a length s increasing subsequence starting from $a_{j_{i+1}}$ so a length $s+1$ subsequence starting at a_{j_i}. Contradiction! Thus $a_{j_1} > a_{j_2} > ... > a_{j_{n+1}}$, and we're done

Remark The above proof used the "generalized pigeonholde principle" which states let n, m, r be positive integers such that $n > rm$. Put n balls into m boxes. Then some box must have at least $r+1$. (Pigeonhole principle is the special case where $r = 1$) *Proof* Assume the contrary that each box has at most r balls. Then total number of balls is at rm but $n > rm$, a contradiction.

Who Cares?

n cards $1, 2, ...n$ in some order so say: $4, 2, 3, 6, 5, 1, 7$ We can cut into piles....

Last time: $I(\Pi)$: the longest increasing subsequence of a permutation of Π

For example: $\Pi = 4, 2, 5, 1, 6, 7, 9, 8, 3$, then $I(\Pi) = 5$ (Subsequence: 2, 5, 6, 7, 9).

Who Cares?

Reason 1: Patience Sorting

Have cards $1, ..., n$. Deck is shuffled giving you a permutation Π. Cards are turns up one at a time and placed according to rule:

- a low card may be placed on top of a higher card or else can start a new pile to right of existing piles.

Goal: have as few piles as possible

ex

$$\Pi = 4, 2, 3, 6, 5, 1, 7$$

Piles
4, 2, 1

Piles
3
6 5
7

Greedy Strategy Place cards as far to the left as possible

Theorem

1. Greedy strategy is optimal
2. with greedy strategy, the number of piles is $I(\Pi)$

Remarks So we now have a way of computing a fast way for $I(\Pi)$

Reason 2: Airline Boarding

Consider the following model:

1. airplane has 1 seat per row
2. Contribution to boarding time is: time it takes to store luggage. Assume this takes 1 unit of time.
3. Passengers are very thin and move quickly compared to storage time.
4. The plane is booked (n passengers, n rows)

ex

$$4, 2, 3, 6, 5, 1, 7$$

Seat No.	Seats
1	
2	
3	
4	
5	
6	
7	

time 1)

- 4 moves to his seat
- 2 moves to the seat (blocking, 3, 6, 5)
- 1 moves to seat (blocking 7)
- 4, 2, 1 store their luggage

Seat No.	seat found
1	*
2	*
3	
4	*
5	
6	
7	

time 2)

- 3 moves to seat (blocking 6, 5, 7)
- 3 stores luggage

Seat No.	seat found
1	*
2	*
3	*
4	*
5	
6	
7	

time 3)

- 6 moves to seat
- 5 moves to seat (blocking 7)
- 5, 6 store their luggage

Seat No.	seat found
1	*
2	*
3	*
4	*
5	*
6	*
7	

time 4)

- 7 moves to seat and stores luggage

Seat No.	seat found
1	*
2	*
3	*
4	*
5	*
6	*
7	*

Boarding Time of Π*:* 4

Theorem The boarding time of Π is $I(\Pi)$.

Example

Have a group of n people. Some handshaking takes place. No pair shakes hands more than once. Show that there must be 2 people who have shaken the same number of hands.

Proof (by contradiction) Assume there aren't 2 people who have shaken the same number of hands. So must have

Person	# of handshakes
Alice	0
Jason	1
...	...
Bob	n-1

To see that this is impossible, ask *have Alice and Bob shaked hands?* Answer is no. Alice shakes 0 hands. And answer is yes. Bob shaked everyone's hand. Contradiction because the answer can't be both no and yes.

Example

Theorem For any n positive integers, there is a subset of them whose sum is divisible by n.

Proof Let the numbers be $a_1, a_2, ..., a_n$. Consider the "boxes" 0 - $n-1$. Consider the subsets

$$\{a_1\}, \{a_1, a_2\}, ..., \{a_1, ..., a_n\}$$

and put each subset in the box corresponding to remainder when you divide the sum of elements in subset by n.

Note: if any of the subsets goes into box 0, then the sum of elements in the subset are divisible by n, and we're done.

If none of them go to box 0, Then we have n subsets in $n-1$ boxes. So by the pigeonhole principle, one of these boxes corresponds to two subsets, call them

$$\{a_1, ..., a_r\}, \{a_1, ..., a_s\}$$

where $r < s$.

Thus, $a_1 + ... + a_r$, $a_1 + ... + a_s$ have the same remainder when you divide by n. So $(a_1 + ... + a_s) - (a_1 + ... + a_r)$ is a multiple of n. So $a_{r+1} + a_{r+2} + ... + a_s$ is a multiple of n as needed.