Pigeonhole Principle

Q: Are there two people in Berlin with the SAME number of strands of hair? $\rightsquigarrow$ YES

Informal PHP (Schubfach Prinzip)

When $n+1$ pigeons sit in n pigeonholes then $\exists$ a pigeonhol with at least 2 pigeons.

Formal PHP X, Y finite sets with $|X| > |Y|$

$$\forall \underbrace{f : X}_{\text{pigeons}} \longrightarrow \underbrace{Y}_{\text{pigeonholes}} \quad \exists\, x_1, x_2 \in X,\ x_1 \neq x_2 \quad \text{s.t.}\ f(x_1) = f(x_2)$$

Examples / Applications

1. Among 13 people there are two who were born in the same month.

2. $\mathcal{B} \subseteq [200]$, $|\mathcal{B}| = 101 \implies \exists\, x \neq y \in \mathcal{B}$ s.t. $x \mid y$

Remark: best possible, as it is false with $|\mathcal{B}| = 100$, e.g. when $\mathcal{B} = \{101, 102, \ldots, 200\}$.

Proof What should be the pigeonholes?

$x \in [200] \rightsquigarrow x = 2^{s_x} \cdot b_x$ with $s_x \in \mathbb{N}_0$, b_x odd.

Put $\underset{\text{pigeons}}{X = \mathcal{B}}$, $\underset{\text{pigeonholes}}{Y = \text{odd numbers in } [200]}$

$$f : X \longrightarrow Y \qquad |X| > |Y| \implies \begin{array}{l}\exists\, x_1 \neq x_2 \in \mathcal{B} \\ \text{with } b_{x_1} = b_{x_2}\end{array}$$
$$x \longmapsto b_x$$

Say $x_1 < x_2 \implies s_{x_1} < s_{x_2} \implies x_1 = 2^{s_{x_1}} \cdot b_{x_1} \mid 2^{s_{x_2}} \cdot b_{x_2} = x_2$

as $\frac{x_2}{x_1} = 2^{s_{x_2} - s_{x_1}}$ ∎

3. **Thm**

Chinese Remainder Theorem

$\forall\, m, n \in \mathbb{N}$ $\quad \gcd(m, n) = 1$

$\forall\, a, b \in \mathbb{N}_0$ $\quad 0 \leq a < m$, $0 \leq b < n$

$\exists\, x \in \mathbb{Z}$ s.t. $x \equiv a \pmod m$ and $x \equiv b \pmod n$
$(\exists\, q \in \mathbb{Z}: x = q \cdot m + a)$ $(\exists\, r \in \mathbb{Z}: x = r \cdot n + b)$

Example

$\exists^?\, x \in \mathbb{N}$ s.t. $x \equiv 26 \pmod{63}$, $x \equiv 13 \pmod{40}$

$\Longrightarrow$ YES! (since $\gcd(63, 40) = 1$) e.g. $x = 1853$

Proof (CHRT)

$A = \{a, m+a, 2m+a, \dots, (n-1)\cdot m + a\}$

Then $|A| = n$ and every element of A is $\equiv a \pmod m$

Case 1: $\exists\, x \in A$ s.t. $x \equiv b \pmod n$ ✓

Case 2: $\forall\, x \in A$ $x \not\equiv b \pmod n$

$$\left.\begin{array}{l} \underset{\text{pigeons}}{X} = A, \quad \underset{\text{pigeonholes}}{Y} = \{0, 1, \dots, n-1\} \setminus \{b\} \\ f: X \longrightarrow Y \\ \quad x \longmapsto \text{remainder} \pmod n \end{array}\right\} \Longrightarrow \begin{array}{l} \exists\, x_1, x_2 \in A \\ x_1 \neq x_2 \\ x_1 \equiv x_2 \pmod n \end{array}$$

$\exists\, i_1 < i_2$, $i_1, i_2 \in \{0, 1, \dots, n-1\}$ s.t. $x_1 = i_1 \cdot m + a$, $x_2 = i_2 \cdot m + a$

$\Longrightarrow i_1 \cdot m + a \equiv i_2 \cdot m + a \pmod n \Longrightarrow (i_2 - i_1)\, m = 0 \pmod n$

$\Longrightarrow n \mid (i_2 - i_1) \cdot m \Longrightarrow$ as $\gcd(m, n) = 1$: $n \mid i_2 - i_1$

However $0 \leq i_2 - i_1 < n \rightsquigarrow$ we must have $i_1 = i_2$ $\lightning$ ■

Pigeonhole Principle - General form

Let $q_1, \dots, q_n \in \mathbb{N}$. When $1 + \sum_{i=1}^{n} (q_i - 1)$ objects are placed in n (numbered) boxes, then $\exists\, i \in [n]$ s.t. Box i contains at least q_i objects.

Proof. By contradiction: If the statement were not true, then $\forall i \in [n]$ Box i would contain less then q_i objects $\Rightarrow$ together we could place at most $\sum_{i=1}^{n}(q_i-1)$ objects $\lightning$. ■

Remark:
When applied, either write out the setting in detail or provide the proof.

Special case $q_1 = \dots = q_n = q \in \mathbb{N}$

- when n boxes together hold $n(q-1)$ objects then $\nexists$ a Box with q objects.
- Reformulation (Averaging): when n boxes together hold Q objects, then
  - $\exists$ Box with at least $\lceil \frac{Q}{n} \rceil$ objects
  - $\exists$ Box with at most $\lfloor \frac{Q}{n} \rfloor$ objects

! PHP is only an EXISTENCE proof technique. The algorithmic problem is much harder, no general method known.

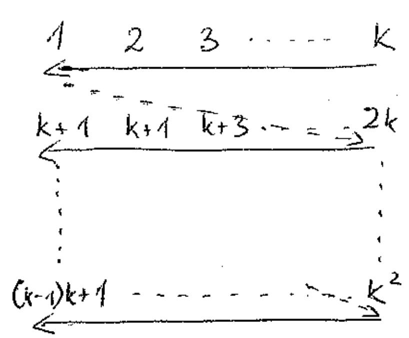
Qestion Given a sequence of length n, how long monotone subsequence can we find in it? (can we guarantee?)

E.g.

n	2	3	4	5	6	7	8	9
	2	2	2	3	3	3	3	3

- we can assume that all elements are distinct (equality just "helps")

Construction for $n=k^2$:



Read rows backwards one after the other.

Claim Longest monotone subsequence is of length k.

Proof → at least k ✓

→ at most k: • every increasing subsequence can have at most one element from each row ⇝ length is at most k.

• every decreasing subsequence can have at most one element from each column ⇝ length is at most k. ∎

Now we prove that this construction is best possible:

Thm (Erdős - Szekeres, 1930's)

Every sequence of k^2+1 reals contains a monotone subsequence of length $k+1$.

Proof $x_1, x_2, \dots, x_{k^2+1}$

For x_j let i_j be the length of the longest increasing subsequence ending at x_j.

<u>Case 1</u> $\exists x_j$ s.t. $i_j \geq k+1 \leadsto \checkmark$

<u>Case 2</u> $\forall x_j$ we have $i_j \leq k$.

Objects (pigeons) : $x_1, \dots, x_{k^2+1}$

Boxes (pigeonholes) : $I_p = \{x_j : i_j = p\}$ $1 \leq p \leq k$

By averaging $\exists$ box I_r s.t.

$$|I_r| \geq \left\lceil \frac{k^2+1}{k} \right\rceil = k+1$$

Claim: Elements of I_r form a decreasing sequence of length $k+1$.

Indeed, $\forall x_\ell, x_j \in I_r$, $\ell < j$ we must have $x_\ell \geq x_j$, otherwise $r = i_\ell < i_j = r$ $\lightning$. ∎

Further (easy) application of the PHP:

$\forall$ coloring of $2k-1$ points with 2 colors contains k points with the same color.

(this is also best possible)

In many applications there are two element relations, so the question also makes sense if we color 2-element subsets.

<u>Q:</u> How many points do we need to have to guarantee k points such that all pairs from them are of the same color?

<u>Example</u> $k=3 \leadsto 5$ not enough, 6 is enough:

<u>Def</u> Let V be a set. Given a Red-Blue coloring of $\binom{V}{2}$, a subset $K \subseteq V$ is called

monochromatic (m.c.) if every set in $\binom{k}{2}$ has the same color.

<u>Prop</u> Given a Red-Blue coloring of $\binom{[6]}{2}$
$\exists$ m.c. $M \subseteq [6]$, $|M| = 3$.

<u>Proof</u> Look at 12, 13, 14, 15, 16.
By the PP (averaging): $\exists \lceil \frac{5}{2} \rceil = 3$ of the same color.

$2 \le i < j < k \le 6$ $1i, 1j, 1k$ are all say Red.
Now look at ij, ik, jk.
Case 1: $\exists$ pair among them of color Red.
↳ say $ij \Rightarrow M = \{1, i, j\}$ is m.c.
Case 2: $\forall$ pair among them is of color Blue.
$\rightsquigarrow M = \{i, j, k\}$ is m.c. ▨

How about $k = 4$? 18 enough, 17 not
(much harder)

In general:

<u>Def</u> Ramsey number, $k \geq 2$

$$R(k) = \min \left\{ N : \forall c : \binom{[N]}{2} \to \{R, B\} \ \exists \text{ m.c } K \subseteq [N] \ |K| = k \right\}$$

<u>Example</u>

$R(2) = 2$, $R(3) = 6$

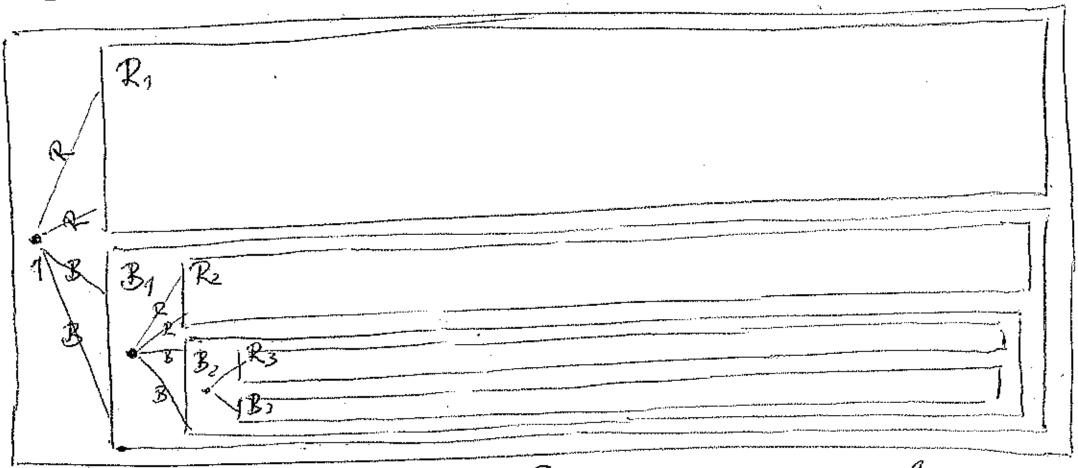
<u>Thm</u> (Ramsey) $R(k) < 4^k$

In particular: $R(k) < \infty$

Proof Let $N = 2^{2k-1}$, $c : \binom{[N]}{2} \longrightarrow \{R, B\}$ arbitrary

We need to find $K \subseteq [N]$, $|K| = k$ m.c.

We will use the idea from the case $k = 3$.



$i_1 := 1$, $V_1 := [N]$, $R_1 := \{ j \in V_1 : c(1j) = \text{Red} \}$

$B_1 := \{ j \in V_1 : c(1j) = \text{Blue} \}$

By averaging either $|R_1|$ or $|B_1|$ is $\geq \left\lceil \frac{2^{2k-1} - 1}{2} \right\rceil = 2^{2k-2}$

Let V_2 be this (larger) set. $\rightsquigarrow |V_2| \geq 2^{2k-2}$

and all pairs $1j$, $j \in V_2$ have the same color C_1

In general we aim to construct:

$$1 = i_1 < i_2 < \cdots < i_{2k} \quad \text{s.t.} \quad i_j \in V_j \;\; \forall j$$

$$[N] = V_1 \supseteq V_2 \supseteq \cdots \supseteq V_{2k} \qquad |V_j| \geq 2^{2k-j} \;\; \forall j$$

and $\forall j$ $\exists$ color C_j s.t. $c(i_j x) = C_j$ $\forall x \in V_{j+1}$

Given the construction for j:

$R_j = \{ x \in V_j : c(i_j x) = \text{Red} \}$

$B_j = \{ x \in V_j : c(i_j x) = \text{Blue} \}$

By averaging either $|B_j|$ or $|R_j|$ is

$$\geq \left\lceil \frac{|V_j|-1}{2} \right\rceil \geq \left\lceil \frac{2^{2k-j}-1}{2} \right\rceil = 2^{2k-(j+1)}$$

Let V_{j+1} be this (larger) set. $\leadsto |V_{j+1}| \geq 2^{2k-(j+1)}$

Let $i_{j+1} = \min\{x \in V_{j+1}\}$, $C_{j+1} = c(i_j i_{j+1})$.

At the end we have a sequence:

$C_1, \ldots, C_{2k-1} \in \{\text{Red}, \text{Blue}\}$

By averaging either Red or Blue appears at least $\left\lceil \frac{2k-1}{2} \right\rceil = k$ times. Say

$C_{j_1} = C_{j_2} = \cdots = C_{j_k} = \text{Red}$ $\quad 1 \leq j_1 \leq \cdots \leq j_k \leq 2k-1$

Then $\{i_{j_1}, \ldots, i_{j_k}, \underset{\substack{\| \\ j_{k+1}}}{i_{2k}}\}$ is a Red $\{k+1\}$-set

For $\quad \ell < \ell' \leq k+1 \qquad c(i_{j_\ell}, i_{j_{\ell'}}) = C_{j_\ell} = \text{Red}$

as $i_{j_{\ell'}} \in V_{j_{\ell'}} \subseteq V_{j_\ell+1}$ and $\forall x \in V_{j_\ell+1}$

we have $c(i_{j_\ell} x) = C_{j_\ell} = \text{Red}$.

This shows: $\quad N = 2^{2k-1} \geq R(k+1)$ ■

What we actually showed: $R(k) \leq 2^{2(k-1)-1} = \frac{4^k}{8}$

What is known?

$R(2) = 2$, $R(3) = 6$, $R(4) = 18$

$43 \leq R(5) \leq 48$

$102 \leq R(6) \leq 165$

$205 \leq R(7) \leq 540$

$\vdots$

$798 \leq R(10) \leq 23556$

Every time when there is an upper bound, we want to know how good it is, i.e. we need a lower bound as well. ⇝ Need a construction of a coloring of $\binom{[N]}{2}$ with Red / Blue without a m.c. set of size k ⇝ $R(k) > N$.

<u>First attempt</u>:
- k-1 sets U_i of size k-1
- pairs inside U_i are Red
- pairs between U_i an U_j are Blue

$\Longrightarrow R(k) > (k-1)^2$ (Quite far from 4^k)

Can we do better?

<u>Thm</u> (Erdős) 1947 $R(k) > \sqrt{2}^k$

- exponential lower bound
- only existence is proved, no construction

<u>Proof</u> Consider ALL Red/Blue - colorings
⇝ There are $2^{\binom{N}{2}}$.
Enumerate all "bad" ones: containing a m.c. k-set. ⇝ "Hope" it is strictly less than $2^{\binom{N}{2}}$ ⇝ Then $\exists$ at least one "good" col.

Fix a k-set $K \subseteq [N]$. Let:

$$\mathcal{B}_K = \{ c : \binom{[N]}{2} \to \{R, B\} : K \text{ is m.c. w.r.to } c \}$$

All "bad" colorings: $\mathcal{B} = \bigcup_{K \in \binom{[N]}{k}} \mathcal{B}_K$

We need $|\mathcal{B}| < 2^{\binom{N}{2}}$ (then there is a "good" coloring)

We will perform very-very rough estimates

$$|\mathcal{B}| = \left| \bigcup_{K \in \binom{[N]}{k}} \mathcal{B}_K \right| \leq \sum_{K \in \binom{[N]}{k}} |\mathcal{B}_K| = \binom{N}{k} \cdot 2^{\binom{N}{2} - \binom{k}{2} + 1}$$

all pairs pairwise col

For us enough if this is $< 2^{\binom{N}{2}}$, i.e.

$$\binom{N}{k} < 2^{\binom{k}{2} - 1} = 2^{\frac{k \cdot (k-1)}{2} - 1}$$

We make further estimates:

$$\binom{N}{k} \leq \left(\frac{N \cdot e}{k} \right)^k < 2^{\frac{k(k-1)}{2} - 1}$$

we want this

$$\iff N < 2^{\frac{k-1}{2}} \cdot \frac{1}{2^{1/k}} \cdot k \cdot \frac{1}{e} \Rightarrow \frac{\sqrt{2}}{e} \cdot k \cdot \sqrt{2}^{k} \cdot (1 + o(1))$$

as $k \to \infty$

$\blacksquare$