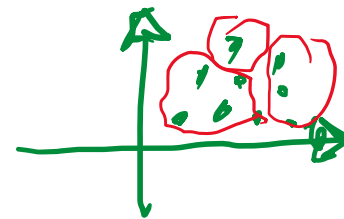
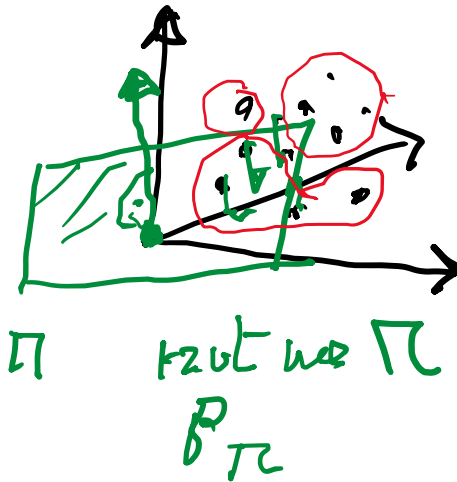
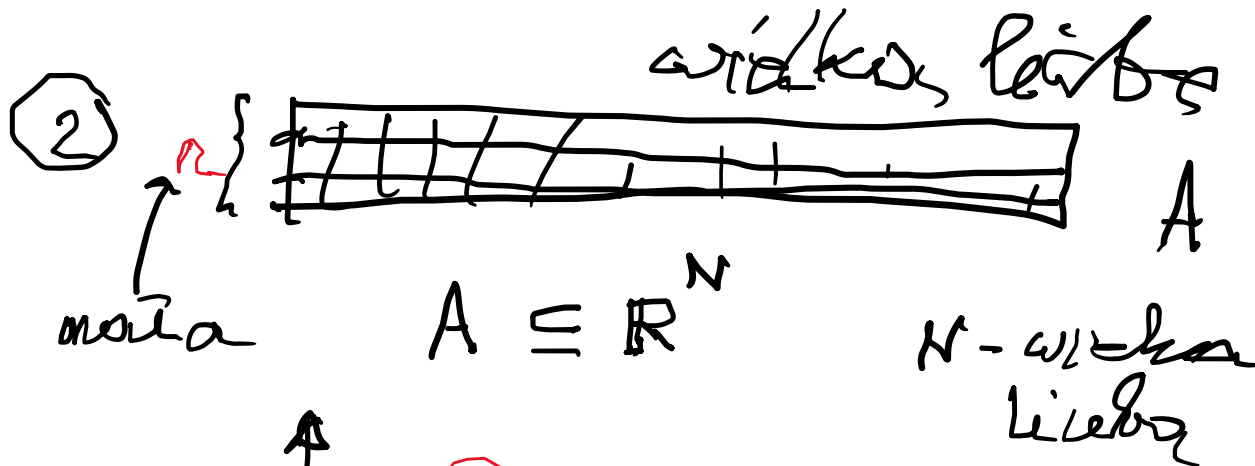
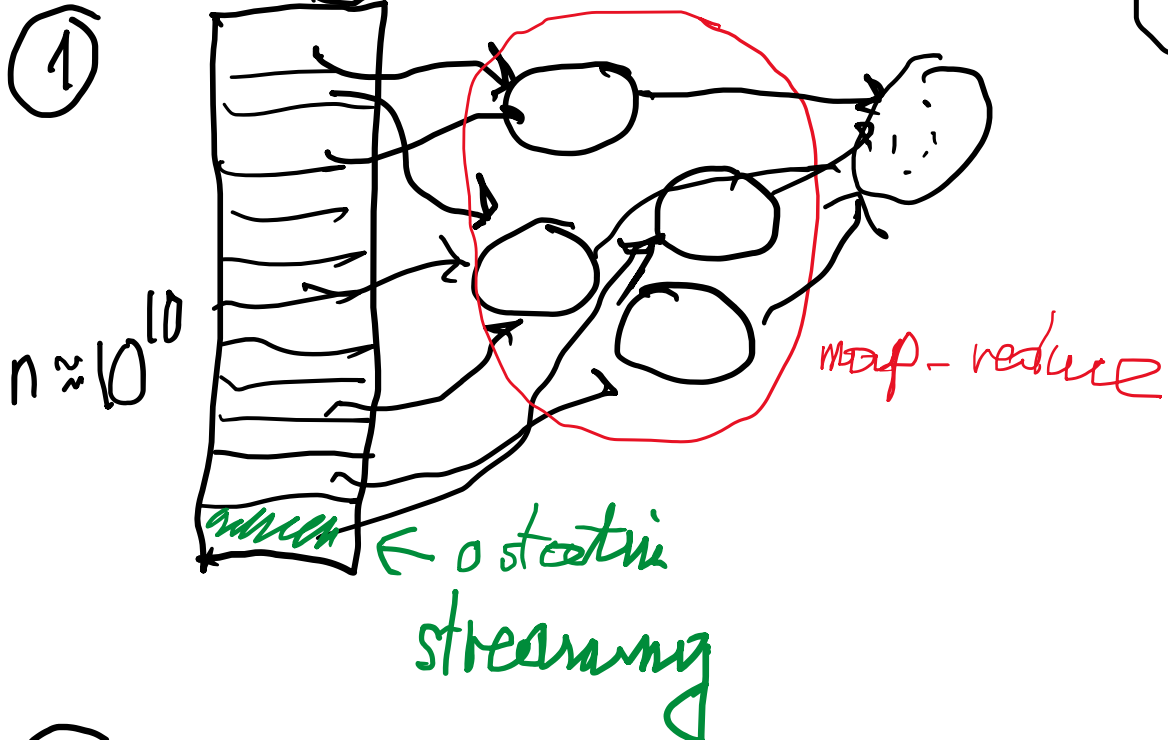
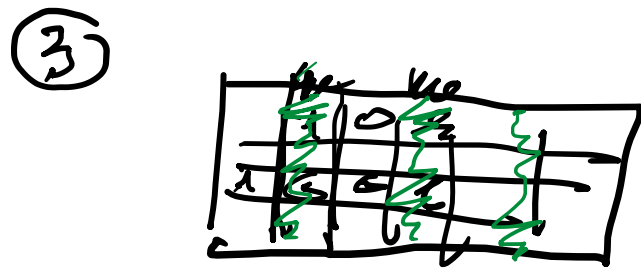


# 3 - schematy



"rzut na losowo, punkt jest OK"



$A \subseteq \{1, \dots, 10^5\}$  wymiar

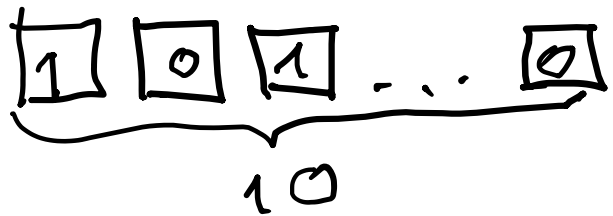
$$P(\{1, \dots, 10^5\}) = 2^{10^5} \approx \ln(n) \frac{1}{\epsilon}$$

$\epsilon$  - zadana precyzja

niebýt dwie

# KOINCIDENCE

• 1950 - 65: ±Rzecz



• 10000 osob

▶  $\approx 10$  osob "jasnowidzow"

kolejny etap.

▶ 0 osob

$$P[\text{suksesu}] = P[S_1 \wedge S_2 \wedge \dots \wedge S_{10}] = P[S_1] \cdot \dots \cdot P[S_{10}] = \left(\frac{1}{2}\right)^{10} \approx \frac{1}{1000}$$

$$E[L_{\text{pasu}}] \approx 10^4 \cdot \frac{1}{10^3} = 10 !$$

$$E[L_{\text{pasobis}}] \approx 10 \cdot \frac{1}{10^3} = \frac{1}{100}$$

Q: how many person "see cat" & "see black cat" job

ZADANIE:  $E \times E \rightarrow R$ .

Polska wyrodek program  $\approx 20 \cdot 10^6$

średni czas pracy  $\approx 10$  lat

$\frac{1}{2}$  of them are hidden

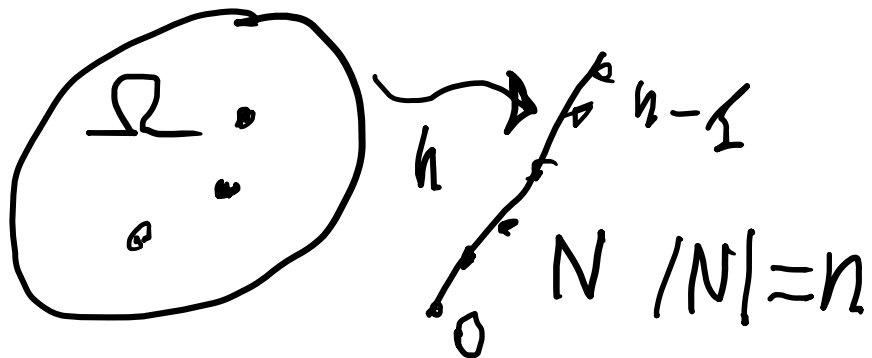
$p = \text{prob. that a given person is in a given stage of job.}$

$q = \text{prob. that a person see "black cat"}$

$\approx \frac{1}{100}$

# Hash functions

## Example

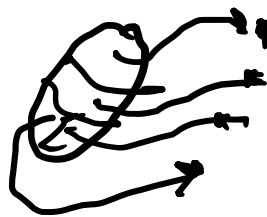


$\vec{a} = (a_0 a_1 a_2 \dots a_r) \leftarrow$  string of ASCII chars

$$f(\vec{a}) = \left( \sum_{i=0}^{r-1} a_i \cdot 256^i \right) \bmod n$$

$|\Omega| = M \gg n$

$f: \text{ASCII}^* \rightarrow \{0, \dots, n-1\}$

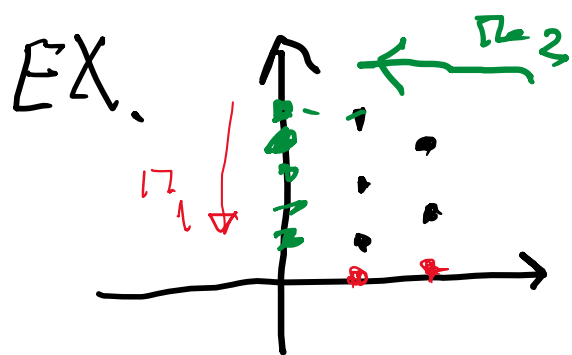


$h: \Omega \rightarrow N$

$\Rightarrow \left( \exists i \in \{0, \dots, n-1\} \mid \left| h^{-1}(\{i\}) \right| \geq \frac{M}{n} \right)$

$\Omega = \bigcup_{a \in \{0, \dots, n-1\}} h^{-1}(\{a\})$

$\left| h^{-1}(\{a\}) \right| < \frac{M}{n} \Rightarrow |\Omega| = \sum_{a=0}^{n-1} \left| h^{-1}(\{a\}) \right| < \sum_{a=0}^{n-1} \frac{M}{n} = n \frac{M}{n} = M$   
 contradiction



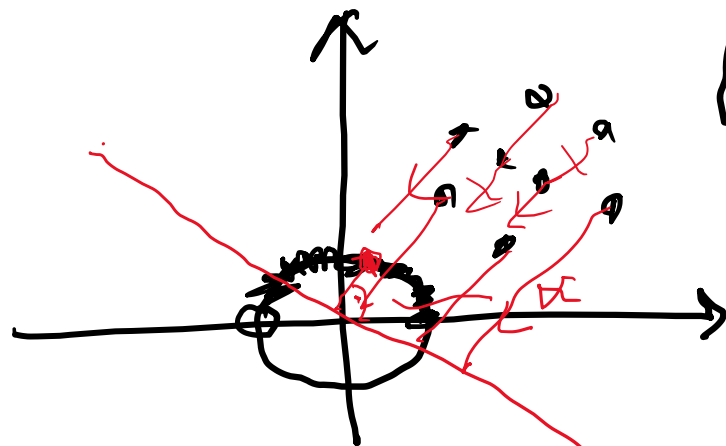
$\pi_1(x, y) = x \leftarrow \text{bad}$

$\pi_2(x, y) = y \leftarrow \text{good}$

SOLUTION: Take random  $\alpha \in [0, 2\pi)$ ;  $\Pr[\pi_\alpha \text{ is good}] = 1/2$ .

$\forall x \in A \quad h(x) \approx 2$

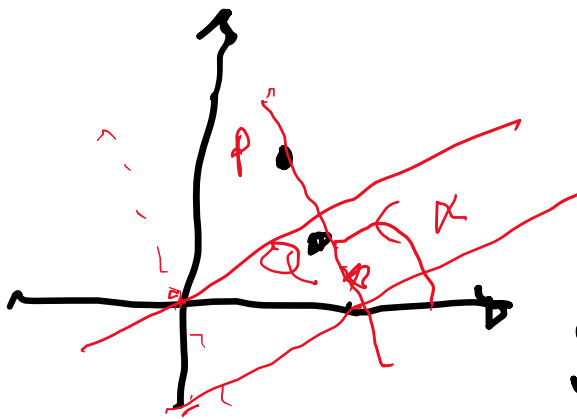
$\pi_\alpha(x, y) = x \cdot \cos \alpha + y \cdot \sin \alpha$



$\alpha \in [0, 2\pi)$

$\pi_\alpha$ : orth. proj.

$\alpha_{pq} = \text{bad angle}$

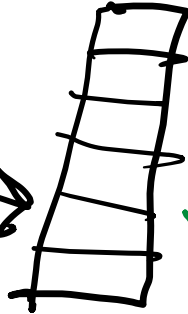
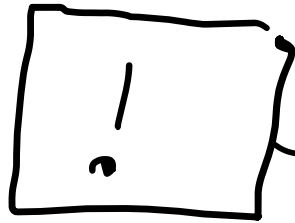


$P = \{p_1, \dots, p_r\}$

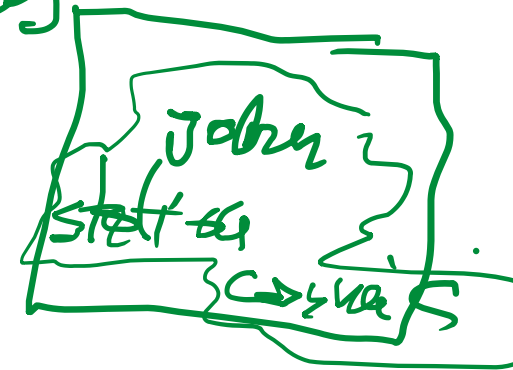
$|\{\alpha_{p_i p_j} : 1 \leq i < j \leq r\}| = \binom{r}{2} = \frac{r(r-1)}{2}$

# FIRST PROGRAMS

hello-world program

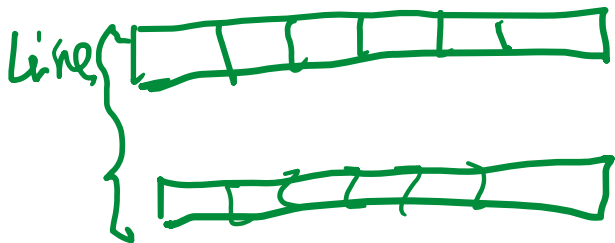


cloud of words



list of 100 most frequent words

\*.txt  
utf-8 encoding



$\{w_1, w_2, \dots, w_k\}$

[John, cat, score, John, ...]

sort

[[John, 75], [cat, 23], ...]

stop words : in English :  
the, that, and, or, ...

cleaning

① eliminate stopwords

scala : x.filterNot(isAStopList)

② john, ask the cat, ...

[john, ', ask, ...]

eliminate artifacts : ', ,, ;, ;