Funkcje haszujące.

Ogólna definicja: $\begin{cases} U, V \text{ dwa zbiory}; |U| \geq |V| \\ f: U \to V \end{cases}$

Typowe zastosowanie:

$V = \{0, \dots, n-1\} = [n]$

↑ zbiór indeksów

$T[0, \dots, n-1]$ ← hash table

$f: U \to [n]$

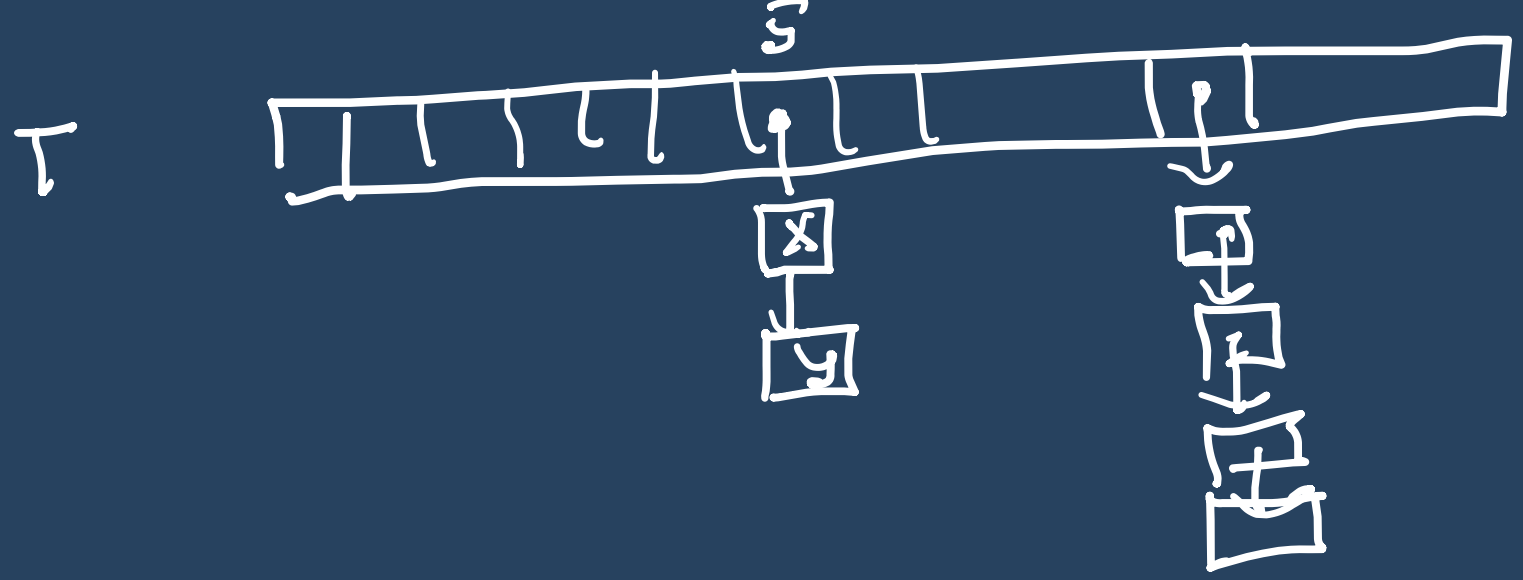
$X \subseteq U$ (np. ASCII^{256})

element $x \in X$ wkładamy do $T[f(x)]$

Q: $x \in X$ ODP: 1) oblicz $i = f(x)$

2) sprawdź zawartość $T[i]$

CZAS: $O(1)$ (±)



- $f(x) = 5$
- $f(y) = 5$

$x \neq y$

KOLIZJA

czas ODP: 1) wyliczenie $i = f(x)$

2) przejrzenie odpowiedniej listy : czas ~ długość listy

KIEPSKIE ROZW:

$f(x) \equiv const.$

NATURALNE wymagania:

- wartości $\{f(x) : x \in X\}$ powinny być
- w miarę jeden. rozłożone w $V = \{0, \dots, M-1\}$

Optymalne upakowanie:

w każdym pudełku ($y \in V$) mamy $\approx \frac{|X|}{|V|}$ ($= \alpha$ współcz. upakowania)

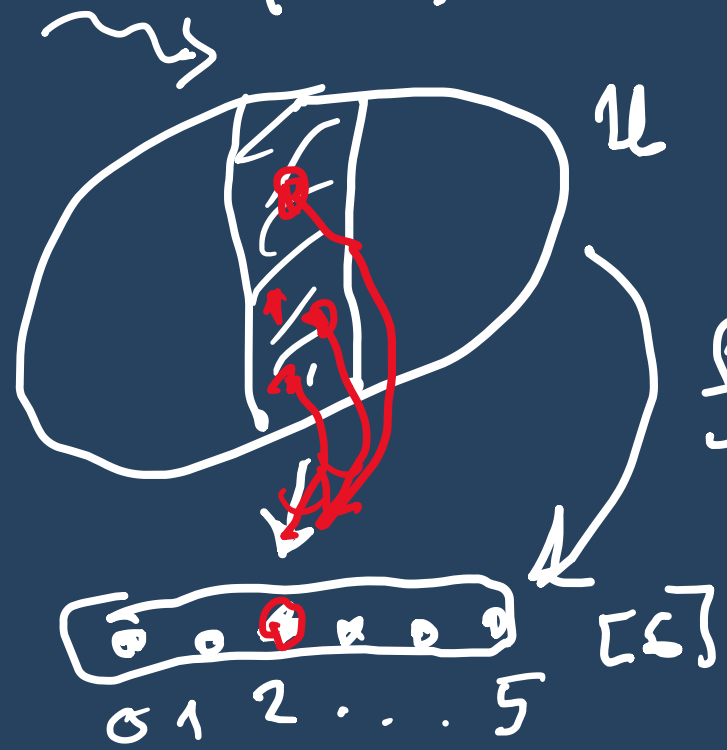
ZŁA WIADOMOŚĆ:

$f : U \longrightarrow [n] \Rightarrow$ istnieje $X \subseteq U$

taki że 1) $|X| \geq \frac{|U|}{n}$

2) $(\exists i \in [n])\ (\forall x \in X)\ (f(x) = i)$.

$\frac{|U|}{n}$



$\bigcup_{i=0}^{5} f^{-1}(i) = U$

$\sum_{i=0}^{5} |f^{-1}(i)| = |U|$

$|f^{-1}(i)| < \frac{|U|}{6}$ $\rightarrow$ sprz.

PRZYKŁAD :

$\mathcal{U} = \mathbb{R}^2$
$V = \mathbb{R}$

$f : \mathbb{R}^2 \to \mathbb{R}$

$X \subseteq \mathbb{R}^3$
$|X| < \infty$

1) $f_1((x,y)) = x$

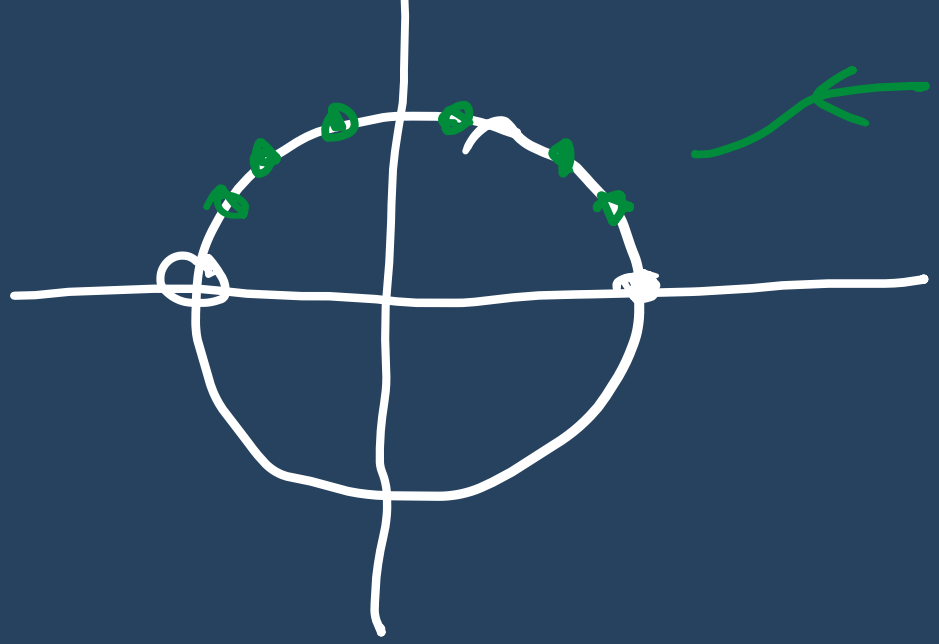
LIPA : mamy kolizje

2) $f_2((x,y)) = y$ LIPA

$(P,Q) \in X^2$, $P \neq Q$

↳ jeden zły kierunek
zły kąt

TYCH PAR MAMY SKOŃCZ. WIELE



← ile kierunków

$\approx [0, \pi)$ ∃K, K-skończ.

Jeśli $\alpha \in [0, \pi)$ jest wybr. zgodnie z rozkł jeden na $[0, \pi)$ to

$$Pr[\alpha \notin K] = 1$$

!!!

STRATEGIA:

1) wybierz $\alpha \in_u [0, \pi)$

2) weź funkcję $\varphi_\alpha(x, y) = (x, y) \circ (\cos\alpha, \sin\alpha)$

$= x \cdot \cos\alpha + y \cdot \sin\alpha$

wtedy

$$\Pr_{\alpha}\left[(\exists x,y\in X)(x\neq y \wedge \varphi_x(x)=\varphi_x(y)\right] = 0$$

INNY JĘZYK:

- mam $f: \mathcal{U} \to V$;
  odwersarz jest w stanie dobrać $X \subseteq \mathcal{U}$ tzn $(f(x): x\in X)$ ma dużo kolizji.
- ~~odwersarz~~ wybiera X.
  a my WYBIERAMY f.

DEF Niech $|U| \geq n$ i $V = \{0, \ldots, n-1\}$.
Rodzina funkcji $\mathcal{H} \subseteq V^U$ jest **k-uniwersalna** jeśli

$$\left(\forall x_1 \ldots x_k \in U\right)\left(\text{jeśli } x_1 \ldots x_k \text{ są parami różne} \rightarrow \right.$$

$$\left.(*) \rightsquigarrow \Pr_{h \in \mathcal{H}}\left(h(x_1) = \ldots = h(x_k)\right) \leq \frac{1}{n^{k-1}}\right)$$

$h \in \mathcal{H}$: wybór $h \in \mathcal{H}$ z rozkładu jednostajnego.

$$(*) \simeq \frac{\left|\{h \in \mathcal{H} : h(x_1) = \ldots = h(x_k)\}\right|}{|\mathcal{H}|} \leq \frac{1}{n^{k-1}}$$

2-uniwers.

$$x \neq y \rightsquigarrow \Pr_h[h(x) = h(y)] \leqslant \frac{1}{n}$$

Uwaga:

X, Y : zmienne losowe o wartościach w $\{0, \ldots, n-1\}$, niezależne, o tym samym rozkładzie.

$$\Pr[X = Y] = \sum_{i=0} \Pr[X = i \wedge Y = i]$$

$$= \sum_{i=0}^{n-1} \Pr[X = i] \cdot P[Y = i] = \sum_{i=0}^{n-1} P_i^2 \quad ; \; P_i = P[X = i].$$

Q: dla jakiego rozkładu $(P_0, \ldots, P_{n-1})$ to jest najmniej?

OPP : min $\equiv$ $p_i = \frac{1}{n}$ (no-bit coded protocol)

Intuitively

$$P[X=Y] = \sum_{i=0}^{n-1} \left(\frac{1}{n}\right)^2 = n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

(P) $\mathcal{H} \subseteq [n]^{\mathcal{U}}$ 2-uniwersalna. $X \subseteq \mathcal{U}$ duże.

$x \in \mathcal{U}$: $h(x) \in \{0 \dots n-1\}$.

$x \notin X$ ile jest $y \in X$ t.że $h(x) = h(y)$?

$X = \{y_1, \dots, y_k\}$

$Y_i = [\![h(x) = h(y_i)]\!]$

$L = \sum_{i=1}^{k} Y_i$.

$$E_h[L] = E_h\left[\sum_{i=1}^{k} Y_i\right] = \sum_{i=1}^{k} E_h[Y_i] = \sum_{i=1}^{k} Pr_h(h(x)=h(y_i))$$

$$\leq \sum_{i=1}^{k} \frac{1}{n} = k/n = \frac{|Y|}{n} \quad (= \alpha)$$

(P) 2-uniwersalna.

$$\begin{cases} V = \{0, \ldots, n-1\} \\ U = \{0, \ldots, p-1\} \end{cases}$$

p - liczba pierwsza

$p > n$.

(P) rozważamy łańcuchy dług. 255 znaków.

$L = (a_1 a_2 \ldots a_{255})$ $a_i \in$ ASCII.

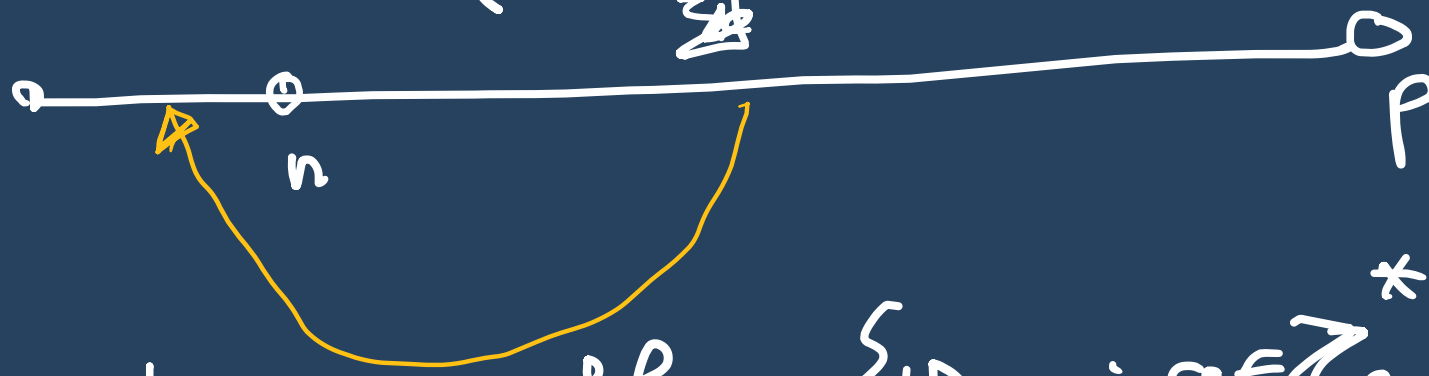
$$\varphi(L) = \sum_{i=1}^{255} \operatorname{ascii}(a_i) \cdot 256^{i-1} \in \mathbb{Z}$$

$\rightsquigarrow$ 255^{255}

p - wielka liczba pierwsza

Dla $a \in \mathbb{Z}_p^*$, $b \in \mathbb{Z}_p$:

$(x \in \mathbb{Z}_p)$ $\varphi_{a,b}(x) = \big(\underbrace{(ax+b) \bmod p}_{\in \mathbb{Z}_p^*} \big) \bmod n$.



Fakt: rodzina $\mathcal{H} = \{\varphi_{a,b} : a \in \mathbb{Z}_p^*, b \in \mathbb{Z}_p\}$ jest 2-uniwersalna.

$$|\mathcal{H}| = (p-1) \cdot p.$$

Ustalmy $x, y \in \mathbb{Z}_p$, $x \neq y$.

[obliczamy w $\mathbb{Z}_p$.

$$\begin{cases} ax + b = u \\ ay + b = v \end{cases}$$

- $(u, v) \leftrightarrow (a, b)$
- ? $ax + b = ay + b$?

$ax + b = ay + b$

$ax = ay \rightarrow ax - ay = 0$

$\rightarrow a(x - y) = 0$

$a \neq 0$, $x \neq y$, $x - y \neq 0$ sprz.

$\mathbb{Z}_p$ jest ciałem

- $u \neq v$.

wtt. różnowartościowe: a, b

- $ax - ay = u - v$

$a(x - y) = u - v$

$a = (u - v) \cdot (x - y)^{-1}$

- $b = u - a \cdot x$.

Mamy jest różnych par (u, v) t.że

- $u \neq v$
- $u \bmod p = v \bmod u$

- ustalmy u (na p sposobów)
- $r = u \bmod n$
- $v = r + l \cdot u$

$u - r = k \cdot u$ dla pewnego k.

$\left(\left\lceil \frac{p}{u} \right\rceil - 1\right) \cdot u$

$\left\lceil \frac{p}{u} \right\rceil \cdot u$

Możliwych l mamy $\leq \left(\left(\left\lceil \frac{p}{u} \right\rceil - 1\right) + 1\right) - 1$

$\left\lceil \frac{p}{u} \right\rceil - 1$

↑ z pozostałych u

CZYLI: tych par wazmy

$$\leq p \cdot \left(\left\lceil \frac{p}{n} \right\rceil - 1 \right)$$

Zadanie: $\left\lceil \frac{p}{n} \right\rceil - 1 \leq \frac{p-1}{n}$

Wskazówka: zapiszmy $p = k \cdot n - r$, $1 \leq r \leq n-1$,

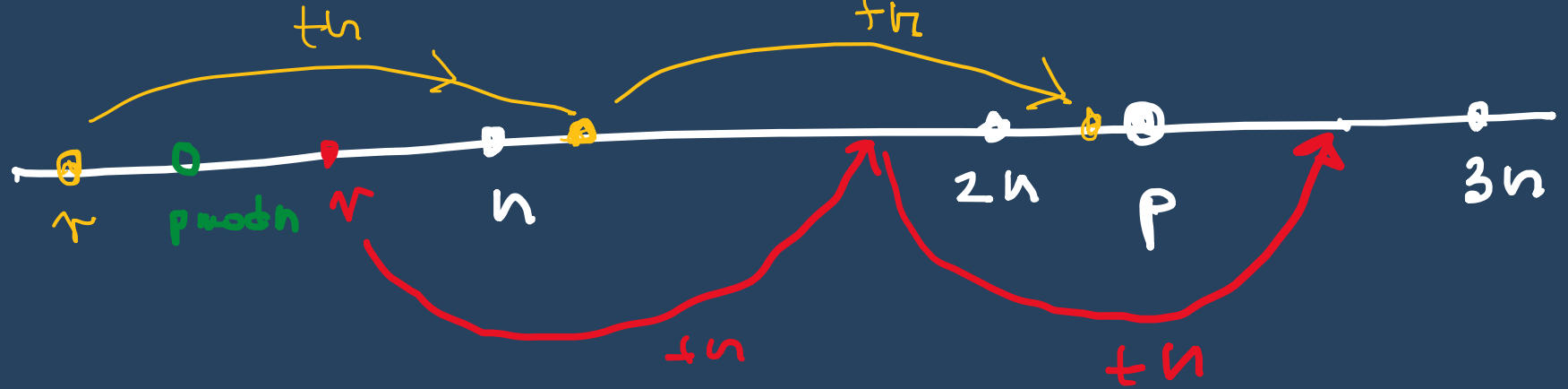
ZATEM: tych par wazmy

$$\leq p \cdot \frac{p-1}{n}.$$

ZATEM:

$$\Pr\left[\varphi_{a,b}(x) = \varphi_{a,b}(y)\right] \leq \frac{p \frac{p-1}{n}}{p(p-1)} = \frac{1}{n}$$

∎



Q: Czy nie fajniej by było gdyby $n \mid p-2$

$[\frac{p}{n}] = \frac{p}{n}$.

Def. Rodzina $\mathcal{H} \subseteq [n]^{[n]}$ jest **SILNIE** **k-uniwersalna** jeśli dla **dowolnych** parami różnych $x_1, \dots, x_k \in [n]$ oraz **dowolnych** $y_1, \dots, y_k \in [n]$ mamy

$$\Pr_{h \leftarrow \mathcal{H}}\left[\bigwedge_{i=1}^{k} h(x_i) = y_i\right] = \frac{1}{n^k}.$$

Uwaga: $Z_i = h(x_i)$; $Z_i \in [n]$; mają się rozkł. jednak.

niezależne $\Leftarrow$ wtedy

$$\Pr[Z_1 = y_1 \wedge \dots \wedge Z_k = y_k] = \Pr[Z_1 = y_1] \cdot \dots \cdot \Pr[Z_k = y_k] = \frac{1}{n^k}.$$

(P) $\mathcal{H}$ - 2-universal, $x \neq y$, $x, y \in \mathcal{U}$

$$Pr_h[h(x) = h(y)] = \sum_{i=0}^{n-1} Pr_h[h(x) = i \wedge h(y) \neq i]$$

$$= \sum_{i=0}^{n-1} \frac{1}{n^2} = \frac{1}{n}.$$

2-universal $\Rightarrow$ 2-univs.