## Obsah 2. kapitoly (pojmy ke zvládnutí)

- optimální $(n, M, d)$-kód
  - malé $n$
  - velké $M$
  - velké $d$

- hlavní problém teorie kódování $A_q(n,d) = ?$

  $A_q(n,d) = \max M$ takové, že $q$-ární $(n, M, d)$-kód existuje.

**Věta 2.1.** $A_q(n,1) = q^n$ a $A_q(n,n) = q$

**Příklad** $A_2(5,3) = 4$ (viz 1. cvičení – kód robota)

- ekvivalence kódů
  - permutace pozic kódu
  - permutace symbolů na pevné (zafixované) pozici kódu.

**Důsledek ekvivalence**

každý $(n, M, d)$-kód je ekvivalentní $(n, M, d)$-kódu, který obsahuje nulové slovo $\bar{0} = (0, 0, \dots, 0)$

Dále jen pro binární kódy (pokud nebude řečeno jinak)

- váha slova je počet jedniček ve slově a značí se $w(\bar{x})$

**Lemma** Pro $\bar{x}, \bar{y} \in (\mathbb{F}_2)^n \Rightarrow d(\bar{x}, \bar{y}) = w(\bar{x} + \bar{y}) = w(\bar{x}) + w(\bar{y}) - 2w(\bar{x} \wedge \bar{y})$

**Věta** Pro $d$-liché,
binární $(n, M, d)$-kód existuje $\Leftrightarrow$ binární $(n+1, M, d+1)$-kód existuje

**Důsledek**
- Pro $d$-liché $A_2(n,d) = A_2(n+1, d+1)$
- Pro $d$-sudé $A_2(n,d) = A_2(n-1, d-1)$

**Příklad:** $(6, 4, 4)$-kód zkonstr z $(5, 4, 3)$-kódu. přidáním par. bitu.

Hammingova hranice (mez) ~ "sphere packing bound".

pro $q$-ární $(n, M, 2t+1)$-kód platí, že

$$M \cdot \left( \binom{n}{0} + \binom{n}{1}(q-1) + \dots + \binom{n}{t}(q-1)^t \right) \leq q^n$$

**Perfektní kód** – kód, který počtem slov $M$ dosahuje Hammingovy meze.

# Příklady řešené na cvičení

(1) Najděte dva kódy se stejnou abecedou, stejným počtem slov a stejnou vzdáleností 2, které nejsou ekvivalentní.

binární $n = 2$

00
01
10
11

$C_1$ 00 11

$C_2$ 01 10

$C_1$ a $C_2$ jsou ekvivalentní

permutace (01) v prvním sloupci

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$n = 3$ $\quad |(F_2)^3| = 8$

| $C_1$ | $C_2$ | $C_3$ |
|---|---|---|
| 000 | 000 | 000 |
| 110 | 101 | 011 |
| 011 | 110 | 101 |
| 101 | | |

$C_1, C_2$ a $C_3$ jsou ekvivalentní

jeden z druhého dostaneme vhodnou permutací pozic (sloupců)

Jsou to (3, 2, 2)-kódy binární.

(3, 3, 2)-kódy a (3, 4, 2)-kód

$n = 4$ $\quad |(F_2)^4| = 16$

| $C_1$ | $C_2$ |
|---|---|
| 0000 | 0000 |
| 1100 | 1100 |
| 0011 | 1010 |

$C_1$ není ekvivalentní s $C_2$

Žádnou permutací pozic ani permutací symbolů nelze z kódu $C_2$ získat kód, kde by se ve všech sloupcích objevily 0 i 1. V $C_1$ je ve všech sloupcích 0 i 1 a proto $C_2$ nelze permutacemi přeměnit na $C_1$, tj. $C_1$ není ekvivalentní $C_2$.

(2) Najděte dva binární kódy stejné délky, počtem slov 4 a min. vzdáleností 2, které nejsou ekvivalentní a jsou co nejkratší.

Z předchozího příkladu je zřejmé, že kódy musí mít délku $n \geq 4$. Bude to stačit?

$$C_1 = \begin{cases} 0000 \\ 1100 \\ 0011 \\ 1111 \end{cases} \qquad C_2 = \begin{cases} 0000 \\ 1100 \\ 0110 \\ 1101 \end{cases}$$

$C_1$ a $C_2$ nejsou ekvivalentní. V $C_1$ je na každé pozici kódu sudý počet jedniček. Žádnou permutací pozic $C_2$ nezměníme počet 1 v sloupci.

Permutací (01) v libovolném sloupci $C_1$ takto zůstane počet 1 sudý.

Tj. z $C_1$ nejde dostat permutacemi $C_2$, proto nejsou $C_1$ a $C_2$ ekv.

(3) Ukažte, že operace (i) a (ii) z ekvivalence kódů nezmění vzdálenost kódu.

Stačí si rozmyslet, že ani jedna z operací (i) a (ii) nezmění Hammingovu vzdálenost jakýchkoliv dvou kódových slov.

(4) Ukažte, že ternární kód $C = \begin{cases} 012 \\ 120 \\ 201 \end{cases}$ je ekvivalentní s ternárním opakovacím kódem délky 3.

Najdeme vhodné permutace, které převedou $C$ na opakovací $C'$.

1.) Na pozici druhé provedeme permutaci $\sigma_1 = (102)$

Pak $C \to C_1$ takový, že $C_1 = \begin{cases} 0\,0\,2 \\ 1\,1\,0 \\ 2\,2\,1 \end{cases}$

2.) Na pozici 3 kódu $C_1$ provedeme permutaci $\sigma_2 = (201)$

Pak $C_1 \to C'$ $\quad C' = \begin{cases} 0\,0\,0 \\ 1\,1\,1 \\ 2\,2\,2 \end{cases}$

---

Def: Váha slova (pro binární kódy = počet jedniček)

Lemma: $w(\bar{x} + \bar{y}) = dist(\bar{x}, \bar{y})$

Lemma: $dist(\bar{x}, \bar{y}) = w(\bar{x}) + w(\bar{y}) - 2(\bar{x} \wedge \bar{y})$

Věta: Že-li $d$ liché $(n, M, d)$-kód existuje $\Leftrightarrow$ existuje $(n+1, M, d+1)$-kód.

Důsledek: $A_2(n, d) = A_2(n+1, d+1)$ pro $d$-liché

2.2.1. *Sestavte následující kódy, pokud existují. a) binární $(2, 4, 1)$-kód, b) binární $(3, 4, 2)$-kód, c) binární $(7, 2, 7)$-kód, d) binární $(5, 3, 4)$-kód. Pokud takový kód neexistuje, vysvětlete proč.*

a) binární $(2, 4, 1)$-kód $\quad C_1 = \begin{cases} 0\,0 \\ 0\,1 \\ 1\,0 \\ 1\,1 \end{cases}$

$(3, 8, 1)$

b) binární $(3, 4, 2)$-kód přidáním paritního bitu k $C_1$

$(4, 8, 2)$

$C_2 = \begin{cases} 0\,0\,0 \\ 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \end{cases}$

c) binární $(7, 2, 7)$-kód – opakovací $\quad C_3 = \begin{cases} 0\,0\,0\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1\,1\,1 \end{cases}$

d) binární $(5,3,4)$ – kód – neexistuje

$$C_4 = \begin{cases} 00000 \\ 11110 \\ ? \end{cases}$$

Třetí slovo $\bar{x}_4$ s alespoň 4 jedničkami nelze přidat. Od slov se 4 jedničkami by se lišilo na nejvýše 2 pozicích což je spor s $d(C)=4$.
Třetí slovo $\bar{y}_5$ s 5 – jedničkami také nelze.
Platí, že dist dist $(\bar{x}_4, \bar{y}_5) = 1$ což je spor s $d(C) = 4$.
Tj. $(5,3,4)$-binární neexistuje.

2.2.2. *Sestavte $(7,8,3)$-kód. Pokud takový kód neexistuje, vysvětlete proč.*

$n = 7 \Rightarrow C \subset (F_2)^7 \quad |(F_2)^7| = 2^7 = 128$

Existuje? Možná, pokud $M$ nepřesáhne Hammingovu mez.

$$M \leq 2^7 / \left( \binom{n}{0}(q-t)^0 + \binom{n}{1}(q-t)^1 \right)$$

$$M \leq 2^7 / (1 + 7) = \frac{2^7}{2^3} = 2^4 = 16$$

Konstruktivní z $C_3(5,4,3)$ kódu vyrobíme $(7,4,3)$ kód $A_1$ přidáním dvou 0.

$$A_1 = \begin{cases} 0000000 \\ 1110000 \\ 0011100 \\ 1101100 \end{cases}$$

Permutací (01) vyrobíme z $A_1$ kód $A_2$, který je $(7,4,3)$-kód

$$A_2 = \begin{cases} 1111111 \\ 0001111 \\ 1100011 \\ 0010011 \end{cases}$$

Sjednocením $A_1 \cup A_2$ dostaneme $(7,8,3)$-kód

Vždy dist $(\bar{x}, \bar{y}) \geq 3$, $\forall \bar{x} \in A_1$ a $\bar{y} \in A_2$.

# Doplňující příklady

**Př 1.** Pokud existují, sestavte $(n, M, d)$-kódy s parametry $(6,2,6)$, $(3,8,1)$, $(4,8,2)$, $(5,3,4)$, $(8,30,3)$.

- $(6,2,6)$-kód, binární $q=2$: $n=6$, $d=6$, $M=2$

$$C_1 = \begin{cases} 000000 \\ 111111 \end{cases}$$

- $(3,8,1)$-kód: $n=3$, $M=8$, $d=1$

$$C_2 = \begin{cases} 000 & 111 \\ 100 & 011 \\ 010 & 101 \\ 001 & 110 \end{cases} = (F_2)^3$$

- $(4,8,2)$-kód: $n=4$, $M=8$, $d=2$

$$C_3 = \begin{cases} 0000 & 1111 \\ 1001 & 0110 \\ 0101 & 1010 \\ 0011 & 1100 \end{cases}$$

- $(5,3,4)$-kód: $n=5$, $M=3$, $4=d$

$$C_4 = \begin{cases} 00000 \\ 11110 \end{cases}$$ Neexistuje

Každý kód je ekvivalentní kódu se slovem tvořeným jen 0-mi. Aby $d=4$ musí každé další slovo obsahovat alespoň 4 1-čky. Ale žádná dvě slova délky 5 se 4-ma 1-ma se nemohou lišit na 4 pozicích.

- $(8,30,3)$-kód: $n=8$, $M=30$, $d=3$

$(F_2)^8 = 2^8 = 256 \sim (8,256,1)$

Z této množiny se lze vybrat i $(8,30,3)$-kód. (Ukážeme později)

**Př. 2.** z Přednášky

Ukažte, že existuje-li binární $(n, M, d)$ kód, pak existuje také binární $(n-1, M', d)$-kód, kde $M' \geq \frac{M}{2}$

Dále ukažte, že $2 \cdot A_2(n-1, d) \geq A_2(n, d)$

Označím $C_1 = (n, M, d)$-kód a $C_2 = (n', M', d')$

$C_2$ vybereme z $C_1$ tak, že vypustíme všechna slova, která v $C_1$ mají poslední symbol 0 (nebo 1 podle toho, kterých je více). Pak jistě $M' \geq \frac{M}{2}$.

Současně všechna vybraná slova mají zaručenou min. vzdálenost z $C_1$, tedy $d' = d$ a $C_2$ je $(n-1, M', d)$-kód, kde $M' \geq \frac{M}{2}$. $C_2$ je tzv. zkrácený kód z $C_1$

Tuto operaci lze provést s libovolným kódem. Tedy je-li:

$$A_2(n, d) = M \text{ pak } A_2(n-1, d) \geq \frac{A_2(n, d)}{2}$$

a úpravou dostaneme $A_2(n, d) \leq 2 A_2(n-1, d)$

**Př. 3.** Ukažte, že existuje-li binární $(n, M, d)$-kód, kde $d$ je sudé, pak existuje binární $(n, M, d)$-kód, kde všechna slova mají sudou váhu.

Předpokl., že $C_1$ je $(n, M, d)$ kod kde $d$ je sudé.

Podle Věty ?? a důsledku existuje $C_1$ <=> existuje $C_2 - (n-1, M, d-1)$ kde $d-1$ je liché.

Důkaz dále. Stačí najít slova $\bar{x}, \bar{y}$ ve vzdálenosti $d$ a ve všech slovech $C_1$ vymazat symbol na stejné pozici tak, aby $d(\bar{x}', \bar{y}') = d-1$

Jestliže z $C_2$ vytvoříme $C_3$ přidáním poslední cifry, která je kontrolním součtem, pak jistě platí, že $C_3$ je $(n, M, d)$ kód a $\forall \bar{x} \in C_3$ je $w(x) \equiv 0 \pmod 2$.

**Př. 4.** Ukažte, že $A_2(8,5) = 4$ a že až na ekvivalenci existuje jen jeden binární $(8,4,5)$-kód.

Můžeme předpokládat, že jedno ze slov je $\vec{x}_0 = 00000000$

$\vec{x}_1 = 11111000$

$\vec{x}_2 = 11000111$

$\vec{x}_3 = 00111111$

$d(\vec{x}_1, \vec{x}_2) = 6 \qquad d(\vec{x}_1, \vec{x}_0) = d(\vec{x}_2, x_0) = 5$

$d(\vec{x}_3, x_i) = 5$

Kód nemůže obsahovat žádné slovo váhy $1, 2, \ldots, 4$. Muselo by mít $d = 5$.

Kód nemůže obsahovat žádné slova váhy 7 nebo 8. Protože od slov váhy 5 by se tato slova lišila na nejvýše 4 pozicích.

Dále, kód může obsahovat nejvýše jedno slovo váhy 6.

Dvě slova váhy 6 se musou lišit na nejvýše 4 pozicích.

Předpokládejme, že $C = \begin{cases} 00000000 \\ ? \end{cases}$

Další 2 slova tedy musí být váhy 5.

Až na ekvivalenci můžeme předpokládat, že $\vec{x}_2 = 11111000$

$\text{např } \vec{x}_3 = 11000111$

Jediné slovo váhy 6, které k nim lze přidat aby $d$ zůstalo 5 je $\vec{x}_4 = 00111111$

Odtud $A_2(8,5) = 4$

**Věta:** Pro $d$-liché a binární kódy — Probráno na přednášce

$(n, M, d)$-kód existuje $\Leftrightarrow$ $(n+1, M, d+1)$-kód existuje

**Důkaz:**

- „$\Rightarrow$" $C$ je binární $(n, M, d)$ kód. Vytvoříme $\hat{C}$ délky $n+1$ přidáním kontrolního paritního symbolu

  tj. $\forall \tilde{x} \in C$ bude $\hat{\tilde{x}} = x_1, x_2, \ldots, x_n, x_{n+1} \in \hat{C}$, kde

  $$x_{n+1} = \sum_{i=1}^{n} x_i \pmod 2$$

  Pak $\forall \hat{\tilde{x}} \in \hat{C}$ platí, že $w(\hat{\tilde{x}})$ je sudé číslo ($w(\hat{\tilde{x}}) \equiv 0 \pmod 2$).

  Dále z „Lemma" zejmé $d(\hat{\tilde{x}}, \hat{\tilde{y}})$ je sudé $\forall \hat{\tilde{x}}, \hat{\tilde{y}} \in \hat{C}$.

  Platí $\underbrace{d = d(C)}_{\text{liché}} \leq d(\hat{C}) \leq \underbrace{d+1}_{\text{sudé}}$. Odtud $d(\hat{C}) = d+1$ a $\hat{C}$ je $(n+1, M, d+1)$-kód.

- „$\Leftarrow$" Předpokl., že $\hat{C}$ je $(n+1, M, d+1)$-kód, $d$ je liché

  Vezmeme $\hat{\tilde{x}}, \hat{\tilde{y}} \in \hat{C}$ takové, že $d(\hat{\tilde{x}}, \hat{\tilde{y}}) = d+1$ (takové $\hat{\tilde{x}}, \hat{\tilde{y}}$ jistě existují)

  a najdeme pozici $i$, na které se $\hat{\tilde{x}}$ a $\hat{\tilde{y}}$ liší.

  Vytvoříme nový kód $C$ z kódu $\hat{C}$ smazáním pozice $i$ ze všech slov $\hat{C}$.

  Pak $C$ je $(n, M, d)$-kód.

**Důsledek:** Pro binární kódy

- Je-li $d$ - liché $\quad A_2(n, d) = A_2(n+1, d+1)$
- Je-li $d$ - sudé $\quad A_2(n, d) = A_2(n-1, d-1)$

**Př. 5.** Označte $E_n$ množinu všech slov (vektorů) z $(\mathbb{F}_2)^n$, která jsou sudé váhy. Ukažte, že $E_n$ je kód, který obdržíme přidáním paritního kontrolního symbolu ke slovům kódu $(\mathbb{F}_2)^{n-1}$. Vyvoďte, že $E_n$ je $(n, 2^{n-1}, 2)$-kód.

$$E_n = \{\bar{x} : \bar{x} \in (\mathbb{F}_2)^n \wedge w(\bar{x}) \equiv 0 \pmod 2\}$$

$(\mathbb{F}_2)^{n-1}$ je $(n-1, 2^{n-1}, 1)$-kód označ. $C$

Dokážeme, že $C \subseteq E_n$ ① $\wedge$ $E_n \subseteq C$ ②

ad ① Ke slovům $C$

přidáním kontrolního paritního symbolu dostaneme $(n, 2^{n-1}, 2)$-kód, kde každé slovo $\bar{x} = x_1, x_2, \dots, x_{n-1}, x_n$ a $x_n = \sum_{i=1}^{n-1} x_i \pmod 2$ (viz lemma 2.5, 2.6 a Theorem 2.7)

Pak ale $w(x) \equiv 0 \equiv \sum_{i=1}^{n} x_i \pmod 2$ a vytvořený kód $C \subseteq E_n$.

ad ② Platí také, že $E_n \subseteq C$? Ano.

Pro každé slovo $\bar{x} \in E_n$ můžeme umazat posl. symbol a dostaneme $\bar{x}'$ takové, že $\bar{x}' \in (\mathbb{F}_2)^{n-1}$.
Zpětným přidáním posledního symbolu dostaneme zpět $\bar{x}$, a zřejmě $w(\bar{x}) \equiv 0 \pmod 2$, tj. přidaný symbol je kontrolní a $\bar{x} \in (n, 2^{n-1}, 2)$-kódu, tedy kódu $C$.

Ukázali jsme, že $|E_n| = |(\mathbb{F}_2)^{n-1}| = 2^{n-1}$, tj. počet slov sudé váhy délky $n$ je stejný jako počet všech binárních slov délky $n-1$.
Také to znamená, že $|E_n| = \frac{1}{2}|(\mathbb{F}_2)^n|$, tj. že polovina binárních slov délky $n$ je sudé váhy a polovina liché váhy.

( Jinak: Kolik je binárních řetězců s lichým/sudým počtem 1-ček délky $n$? označíme $a_n$ počet sudých )

Sestavením rekurentní rovnice

$\dots\dots - 0 \quad a_{n-1}$

$\dots\dots - 1 \quad 2^{n-1} - a_{n-1}$

$a_n = a_{n-1} + (2^{n-1} - a_{n-1}) = 2^{n-1}$

Umíte vymyslet další jiný argument?

Př: Ukažte, že existuje-li $(n, M, d)$ kód pro $d$ sudé, pak existuje $(n, M, d)$ kód, jehož všechna slova jsou sudé váhy.

Nebylo součástí cvičení

Předpokl., že $C_1$ je $(n, M, d)$ kod kde $d$ je sudé.

Podle Theorem v 2.7. a důsledku 2.8. $C_1$ existuje $\Leftrightarrow$ existuje $C_2 - (n-1, M, d-1)$ kde $d-1$ je liché.

Stačí najít slova $x, y$ ve vzdálenosti $d$ a ve všech slovech $C_1$ vymazat symbol na stejné pozici tak, aby $d(x', y') = d-1$

Jestliže z $C_2$ vytvoříme $C_3$ přidáním poslední cifry, která je kontrolním součtem, pak jistě platí, že $C_3$ je $(n, M, d)$ kód a $\forall x \in C_3$ je $w(x) \equiv 0 \pmod 2$. Viz konstrukce v důkazu Theoremu 2.7

Tj existuje $C_3 - (n, M, d)$ kód pro $d$-sudé takový, že slova jsou sudé váhy.

**Př. 6.** Ukažte, že libovolný $q$-ární $(q+1, M, 3)$-kód splňuje $M \leq q^{q-1}$

$\left|(\mathbb{F}_q)^{q+1}\right| = q^{q+1} \sim (q+1, q^{q+1}, 1)$ kód

pro $d = 3$

pro $q = 2$ $\quad (3, M, 3)$ $\quad M \leq 2^1 = 2$ $\quad = \{000, 111$

pro $q \geq 3$ $\quad (4, M, 3)$ $\quad M \leq 3^2 = 9$ $\quad = \{0000, 1110, 2220, \vdots$

Využijeme Hammingova mez (hraniol)

$$M \cdot \left\{\binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \ldots + \binom{n}{t}(q-1)^t\right\} \leq q^n$$

pro $(n, M, 2t+1)$-kód

pro $(q+1, M, 3)$-kód platí: $d = 3 \Rightarrow t = 1$

$$M \cdot \left\{\binom{q+1}{0} + \binom{q+1}{1}(q-1)\right\} \leq q^{q+1}$$

$$M \cdot \{1 + (q+1)(q-1)\} \leq q^{q+1}$$

$$M \cdot (1 + q^2 - 1) \leq q^{q+1}$$

$$M \leq q^{-2} \cdot q^{q+1} \Rightarrow \underline{\underline{M \leq q^{q-1}}}$$

## Další cvičení:

Je $(5, 4, 3)$-kód ($C_3$-kód robota) perfektní?

Ověřte: Sami

Již dříve jsme ukázali, že $A_2(5, 3) = 4$