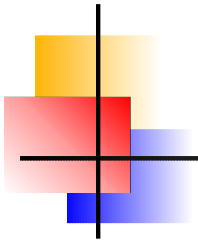
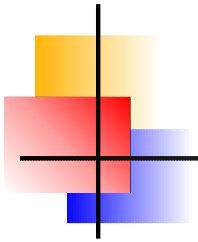


Objavovanie a opravovanie chýb

Stanislav Palúch

Fakulta riadenia a informatiky, Žilinská univerzita

27. marca 2012



Všeobecná teória samoopravných kódov

Definícia

Guľa $G_t(\mathbf{c})$ **o strede** $\mathbf{c} \in A^n$ **a polomere** t *je množina*

$$G_t(\mathbf{c}) = \{\mathbf{x} \mid \mathbf{x} \in A^n,\ d(\mathbf{x}, \mathbf{c}) \leq t\}.$$

Guľa $G_t(\mathbf{c})$ je množina všetkých takých slov, ktoré vznikli zo slova $\mathbf{c}$ nanajvýš t jednoduchými chybami.

Všeobecná teória samoopravných kódov

- Samotné slovo $\mathbf{c}$ je tiež prvkom gule $G_t(\mathbf{c})$ a prispieva k počtu jej prvkov číslom $1 = \binom{n}{0}.(r-1)^0$,
- $\binom{n}{1}.(r-1)$ – počet slov, ktoré sa líšia od $\mathbf{c} \in A^n$ práve na jednom mieste,
- $\binom{n}{2}.(r-1)^2$ – počet slov, ktoré majú od slova $\mathbf{c}$ vzdialenosť práve 2
-
- $\binom{n}{i}.(r-1)^i$ – počet slov, ktoré majú vzdialenosť od slova $\mathbf{c}$ rovnú i

Počet slov v $G_t(\mathbf{c})$ je teda

$$|G_t(\mathbf{c})| = \sum_{i=0}^{t} \binom{n}{i}.(r-1)^i \ . \qquad (1)$$

Všeobecná teória samoopravných kódov

Počet prvkov gule $G_t(\mathbf{c})$ nezávisí na tom, aké slovo $\mathbf{c}$ sme vybrali za jej stred – všetky gule o rovnakom polomere t majú rovnakú mohutnosť (1).

Definícia

Hovoríme, že **kód** $\mathcal{K}$ **opravuje** t **jednoduchých chýb**, *ak pre slovo* $\mathbf{y}$, *ktoré vzniklo z niektorého kódového slova nanajvýš* t *jednoduchými chybami, existuje jediné slovo* $\mathbf{x}$ *také, že* $d(\mathbf{x}, \mathbf{y}) \leq t$.

Ak $\mathbf{b} \in G_t(\mathbf{c}_1) \cap G_t(\mathbf{c}_2)$, potom slovo $\mathbf{b}$ mohlo vzniknúť nanajvýš t jednoduchými chybami z oboch slov $\mathbf{c}_1$, $\mathbf{c}_2$.

Ak má teda kód $\mathcal{K}$ opravovať t chýb, musí byť pre ľubovoľnú dvojicu $\mathbf{c}_1$, $\mathbf{c}_2$ rôznych kódových slov

$$G_t(\mathbf{c}_1) \cap G_t(\mathbf{c}_2) = \emptyset \ . \tag{2}$$

Naopak. Ak pre ľubovoľnú dvojicu rôznych kódových slov kódu $\mathcal{K}$ platí (2), potom kód $\mathcal{K}$ opravuje t chýb.

Všeobecná teória samoopravných kódov

Predpokladajme, že kód $\mathcal{K} \subseteq A^n$ opravuje t jednoduchých chýb. Keďže $|A^n| = r^n$, pre počet kódových slov $|\mathcal{K}|$ vzhľadom na (1) a (2) platí

$$\sum_{i=0}^{t} \binom{n}{i} .(r-1)^i \ . \ |\mathcal{K}| \leq r^n \ . \tag{3}$$

Definícia

Hovoríme, že **kód** $\mathcal{K} \subseteq A^n$ **je** t**-perfektný kód**, *ak*

$$\forall \mathbf{a},\ \mathbf{b} \in A^n, \quad \mathbf{a} \neq \mathbf{b} \quad G_t(\mathbf{a}) \cap G_t(\mathbf{b}) = \emptyset \ ,$$
$$\bigcup_{\mathbf{a} \in \mathcal{K}} G_t(\mathbf{a}) = A^n \ .$$

Veta

Kód $\mathcal{K}$ *opravuje* t*-násobné chyby práve vtedy, keď*

$$\Delta(\mathcal{K}) \geq 2t + 1 \ , \tag{4}$$

kde $\Delta(\mathcal{K})$ *je minimálna vzdialenosť kódu* $\mathcal{K}$.

Príklad

Majme abecedu $A = \{a_1, a_2, \ldots, a_r\}$.
Opakovací kód dĺžky k je blokový kód, ktorého kódové slová pozostávajú z k rovnakých znakov, t. j.

$$\mathcal{K} = \{a_1 a_1 \ldots a_1,\ a_2 a_2 \ldots a_2,\ \ldots\ ,\ a_r a_r \ldots a_r\}.$$

Minimálna vzdialenosť opakovacieho kódu dĺžky k je $\Delta\mathcal{K} = k$ *a takýto kód opravuje t-násobné chyby pre* $t < k/2$.

Špeciálne pre $r = 2$ *(t. j. pre binárnu abecedu A) a k nepárne, t. j.* $k = 2t + 1$ *je opakovací kód t-perfektný.*

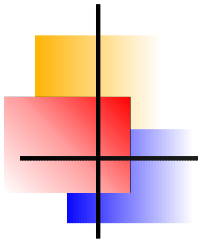
Príklad

Kód s kontrolou parity má minimálnu vzdialenosť 2, a preto neopravuje ani jednu jednoduchú chybu.

Príklad

Kód dvojrozmernej kontroly parity.

101	0 ← kontrola parity riadku
000	0
001	1
010	1
111	1
111	1
000	0
kontroly parity stĺpcov → 110	0 ← celková kontrola parity



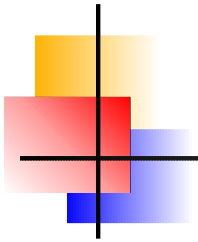
Všeobecná teória samoopravných kódov

Definícia

Dekódovanie *kódu $\mathcal{K}$ je ľubovoľné zobrazenie δ, ktorého obor hodnôt je $\mathcal{K}$, ktorého definičný obor $\mathcal{D}(\delta)$ je podmnožinou množiny A^n všetkých slov dĺžky n a obsahuje $\mathcal{K}$ a pre ľubovoľné $\mathbf{a} \in \mathcal{K}$ je $\delta(\mathbf{a}) = \mathbf{a}$.*

$$\mathcal{K} \subset \mathcal{D}(\delta) \subseteq A^n, \qquad \delta : \mathcal{D}(\delta) \to \mathcal{K}\ , \qquad \forall \mathbf{a} \in \mathcal{K}\ \ \delta(\mathbf{a}) = \mathbf{a}\ .$$

Ak $\mathcal{D}(\delta) = A^n$, hovoríme, že dekódovanie δ je **úplné**, *inak hovoríme, že dekódovanie δ je* **čiastočné**.



Všeobecná teória samoopravných kódov

Definícia

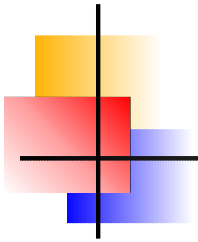
Nech $\mathcal{K} \subseteq A^n$ je blokový kód dĺžky n. Hovoríme, že **kód $\mathcal{K}$ má k informačných a $n-k$ kontrolných znakov**, *ak existuje vzájomne jednoznačné zobrazenie $\phi : A^k \leftrightarrow \mathcal{K}$. Zobrazenie ϕ nazveme* **kódovanie informačných znakov**.

Príklad

Opakovací kód dĺžky 5 s abecedou $A = \{0,1,2,3,4,5,6,7,8,9\}$ má jeden informačný znak a 4 znaky kontrolné, pretože zobrazenie ϕ definované

$$\begin{array}{lllll} \phi(0) = 00000 & \phi(1) = 11111 & \phi(2) = 22222 & (3) = 33333 & \phi(4) = 44444 \\ \phi(5) = 55555 & \phi(6) = 66666 & \phi(7) = 77777 & \phi(8) = 88888 & \phi(9) = 99999 \end{array}$$

je vzájomne jednoznačné zobrazenie $\phi : A^1 \leftrightarrow \mathcal{K}$.



Všeobecná teória samoopravných kódov

Príklad

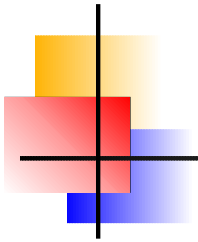
Zdvojovací kód dĺžky $2n$ má n informačných a n kontrolných znakov. Kódovanie informačných znakov $\phi : A^n \leftrightarrow \mathcal{K}$ definujeme predpisom

$$\phi(a_1 a_2 \dots a_n) = a_1 a_1 a_2 a_2 \dots a_n a_n.$$

Príklad

Kód dva z piatich vôbec nemá oddelené informačné a kontrolné znaky.

Počet kódových slov tohoto kódu je 10 – nie je mocninou čísla 2, a preto nemôže existovať vzájomne jednoznačné zobrazenie množiny $\{0, 1\}$ na množinu kódových slov mohutnosti 10.



Všeobecná teória samoopravných kódov

Definícia

Blokový kód $\mathcal{K}$ je **systematický**, *ak pre každé slovo $a_1 a_2 \dots a_k \in A^k$ existuje práve jedno kódové slovo $\mathbf{a} \in \mathcal{K}$ také, že*

$$\mathbf{a} = a_1 a_2 \dots a_k, a_{k+1} \dots a_n \ .$$

Príklad

Opakovací kód je systematický s $k = 1$.

Binárny kód s kontrolou parity dĺžky 8 je systematický s $k = 7$.

Kód medzinárodného čísla vozňa je systematický s $k = 11$.

Príklad

Zdvojovací kód s dĺžkou $2n$ väčšou ako 2 nie je systematický.

Veta

Nech $\mathcal{K}$ je systematický kód s k informačnými a $n-k$ kontrolnými znakmi. Potom pre minimálnu vzdialenosť $\Delta\mathcal{K}$ platí

$$\Delta\mathcal{K} \leq n-k+1\ . \tag{5}$$

Dôkaz.
Zvoľme dve slová $\mathbf{a} = a_1 a_2 \ldots a_{k-1} a_k \in A^k$, $\overline{\mathbf{a}} = a_1 a_2 \ldots a_{k-1} \overline{a}_k \in A^k$ líšiace sa len v poslednom k-tom znaku.
Pretože kód $\mathcal{K}$ je systematický, ku každému z takýchto slov existuje práve jedno slovo $\mathbf{b}$ resp. $\overline{\mathbf{b}}$ kódu $\mathcal{K}$, také, že $\mathbf{a}$ je prefixom $\mathbf{b}$, resp. $\overline{\mathbf{a}}$ je prefixom $\overline{\mathbf{b}}$:

$$\mathbf{b} = a_1 a_2 \ldots a_{k-1} a_k a_{k+1} \ldots a_n\ ,$$
$$\overline{\mathbf{b}} = a_1 a_2 \ldots a_{k-1} \overline{a}_k \overline{a}_{k+1} \ldots \overline{a}_n\ .$$

Keďže sa slová $\mathbf{b}$, $\overline{\mathbf{b}}$ zhodujú na $k-1$ miestach môžu sa nezhodovať najviac na $n-(k-1) = n-k+1$ miestach.
Je $d(\mathbf{b}, \overline{\mathbf{b}}) \leq n-k+1$ a teda $\Delta\mathcal{K} \leq n-k+1$. □

Dôsledok. Kód $\mathcal{K}$ s k informačnými a $n-k$ kontrolnými znakmi môže opravovať najviac $\left[\frac{n-k}{2}\right]$ chýb (kde $[x]$ je celá časť čísla x).

Príklad

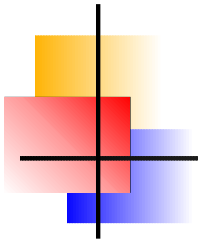
Pre zdvojovací kód dĺžky $n = 2t$ je $k = t$, $n - k = t$, ale minimálna vzdialenosť tohoto kódu je 2, čo je pre veľké t hlboko pod odhadom (5), ktorý pre náš prípad dáva $\Delta\mathcal{K} \leq 2t - t + 1 = t + 1$.

Definícia

Nech $\mathcal{K}$ je kód s k informačnými a $n - k$ kontrolnými znakmi. Pomer

$$R = \frac{k}{n} \tag{6}$$

nazveme **informačný pomer**.



Všeobecná teória samoopravných kódov

Pri navrhovaní samoopravných kódov sa snažíme zabezpečiť sa proti čo najväčšiemu počtu chýb, čo vedie k zvyšovaniu počtu kontrolných znakov.

Druhou prirodzenou požiadavkou je dosiahnuť čo najväčší informačný pomer, čo je v rozpore so zvyšovaním počtu kontrolných znakov.

Navyše na príklade 9 vidíme, že nie každé zvyšovanie počtu kontrolných znakov musí viesť k zväčšovaniu minimálnej vzdialenosti kódu.