

5.3. Словарные методы сжатия

• LZ'77

окно = словарь + буфер

1) словарь = \emptyset , буфер содержит начало сообщения

2) ищем в словаре максим. фрагмент, совп. с началом содержащего буфера

3) записываем код (i, j, s) ,

i - номер позиции в словаре, с ктв нач. совп. фрагмент

j - длина совп. фрагм.

s - первый симв. буфера, след. за фрагм.

X-ки: - степень сжатия: 2-4

- тип данных: любые, но лучше разнородные

- симв-ти: 10:1

• LZ'78

нет скольз. окна

1) словарь = \emptyset , в буфер помещ. все сообщ-е

2) ищем в словаре макс. фрагмент, совп. с началом буфера

3) записываем код (i, s) :

i - номер слова в словаре, с ктв совп. нач. буфера

s - перв. символ, след. за фрагментом

X-ки: степень сжатия: 2-3

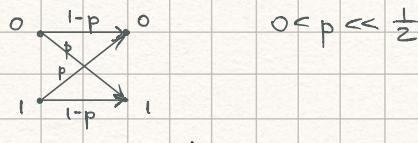
тип данных: любые, но лучше однородные

симв-ти: 3:2

• LZW - модиф. LZ'78

6. Введение в теорию кодирования

6.1. Дв. симметр. канал связи.



источник приемник

$$u = (u_1 \dots u_n) \xrightarrow{n \geq k} x = (x_1 \dots x_n) \xrightarrow{\text{канал связи}} x' = (x'_1 \dots x'_n) \xrightarrow{\text{декод.}} u' = (u'_1 \dots u'_k)$$

шар. ошибок кодов. слова

$n \cdot p \cdot (1-p)^{n-1}$ - вер-ть того, что при передаче по каналу связи произв. дв. вект. ошибки n в итоге произв. ровно 1 ошибка.

$$C_n^2 p^2 (1-p)^{n-2} \sim 2 \text{ ошибки}$$

Пусть мы используем такое код-е, что Число ошибок $\leq t$ ио скорректировано при декодир-ии

Тогда опр-и вер-ть ошибки при дек-ии:

$$P_{err} = 1 - \sum_{i=0}^t C_n p^i (1-p)^{n-i}$$

кода, исправл. + ошибок

6.2. Булев куб. Метрика Хэмминга.

Опн E^n - булев куб разм-ти n - зерн. спеч. вида:

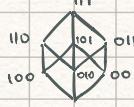
$$V = \{x = (x_1, \dots, x_n) \mid x_i \in \mathbb{Z}_2\} \quad |V| = 2^n$$

$E = \{(x, y) \mid x \neq y \text{ разн. ровно в 1 коор-те}\}$

$n=1$

$n=2$

$n=3$



$$|E| = n \cdot 2^{n-1}$$

Опн Носитель вектора — $\text{supp}(x) = \{i \mid x_i = 1\}$

Вес — $\text{wt}(x)$ - число единиц в векторе = $|\text{supp}(x)|$

ЛБ супд булевого куба — все векторы веса k

Опн $x \oplus y = (x_1 \oplus y_1, \dots, x_n \oplus y_n)$, т.е. $\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$ (xor)

Опн Расстояние Хэмминга: $d(x, y)$ - число коор-т, в к-р они различ-ся

$$d(x, y) = \text{wt}(x \oplus y)$$

Ex $x = \begin{array}{c} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{array}$ $y = \begin{array}{c} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{array}$ $x \oplus y = \begin{array}{c} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{array}$ $d(x, y) = 2$

Опн Сфера — $S_r(x) = \{y \mid d(x, y) = r\}$
Шар — $B_r(x) = \{y \mid d(x, y) \leq r\}$

$$|S_r(x)| = C_n^r$$
$$|B_r(x)| = \sum_{i=0}^r C_n^i$$

Опн Скл. произв-е вект. x, y — это число $\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n$ (аналог толькo)

Опн Линейное подпр-во в булевом кубе — подмн-во булева куба, замкн. относ-я скл*-я векторов

Линейное подпр-во в E^n имеет мощн-ть 2^k где подкоо. k , т.е.
 k - размерность лин. подпр-ва

$$k=0 \quad L = \{0\} \quad k=1 \quad L = \{0, x\} \quad k=2 \quad L = \{0, x, y, x \oplus y\}$$

Опн Ортогон. подпр-во $L^\perp = \{y \in E^n \mid \langle x, y \rangle = 0 \quad \forall x \in L\}$

Сл-во: $\dim L + \dim L^\perp = n$

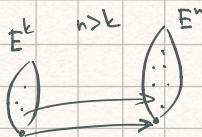
6.3. линейные коды

Опк Код — способ представления букв. языка E^n

Опк Мин. код — способ минимизированного представления букв. языка E^n

$$u = (u_1 \dots u_k) \rightarrow x = (x_1 \dots x_n)$$

нагр. слово



$$C \subseteq E^n \quad |C| = 2^k$$

Если код линейный, то каждое из x_i линейными образом зависит от $u_1 \dots u_k$

Ex $k=2 \quad n=3$

$$\begin{matrix} u_1 & u_2 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{matrix} \rightarrow \begin{matrix} x_1 & x_2 & x_3 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{matrix}$$

$$\begin{aligned} x_1 &= u_1 \\ x_2 &= u_2 \\ x_3 &= u_1 \oplus u_2 \end{aligned}$$

иначе чистота

Код только обнаруживает ошибку, но не исправляет



$$(01)(11)(00)$$

$$(011)(110)(000)$$

$$(011)(100)(000)$$

$$(011) \downarrow \text{ок} \quad (100) \downarrow \text{не ок} \quad (000) \downarrow \text{ок}$$

$$(011) \quad : \quad (100) \quad (000)$$

$$000 \quad 110 \quad 101 \quad \text{если } 1 \text{ ошибка}$$

сторонам, \rightarrow 011
т.к. наим. вероятн.

Ex $k=2$

$n=5$

$$\begin{aligned} x_1 &= u_1 \\ x_2 &= u_2 \\ x_3 &= u_1 \\ x_4 &= u_1 \oplus u_2 \\ x_5 &= u_2 \end{aligned}$$

$$\begin{array}{c|ccccc} 00 & 000 \\ 01 & 011 \\ 10 & 110 \\ 11 & 101 \end{array}$$

$$u \rightarrow x \xrightarrow{k \text{ об.}} x' = 01111$$

$$\text{01011}$$

иначем наим. вер. сущ.

$$u \rightarrow x \xrightarrow{k \text{ об.}} x' = 11110$$

$$\text{10110}$$

$$\text{10}$$

В этом коде всегда
ошибку исправишь

Параметры мин. кода: $[n, k, d]$

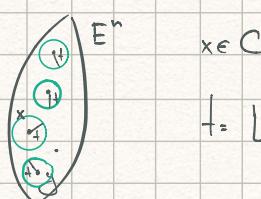
n — длина кода (размерностьпр-ва, в ктр рассм. код)

k — разм-ть кода

d — мин. расст-е Хэмминга между код. словами

Из Ex1: $[3, 2, 2]$

Из Ex2: $[5, 2, 3]$



$$t = \left\lfloor \frac{d-1}{2} \right\rfloor - \text{тогда ошибки испр-ся}$$

Опк Порождающая ма-ца мин. кода G $k \times n$, ее строки — базис мин. кода

$$u \cdot G = x \quad u \in E^k \quad - \text{генерация кода}$$

Опр Проверочная матрица кода H $(n-k) \times n$, её строки — базис ортогон. кода.

$$Hx^\top = 0 \iff x \in L$$