HW3 - due 5/5

$c: X \to \{0,1\}^*$

Ex: r.v. $X$ & a code for $X$

$$\mathbb{E}[|c(x)|] < H(x).$$

- for prefix-free code $C$: $\mathbb{E}[|C(x)|] \geq H(X)$.

- $\forall x \neq y \Rightarrow C(x) \neq C(y)$ ... needs to be true (non-singular)

$C(x) \circ C(x') \circ C(x'') \circ$

uniformly chosen # $\{0, \dots, 1111111\}$

$X \in \{0,1\}^n$ encoding - remove the leading zero's.

$H(X) = n \qquad \mathbb{E}[|C(x)|] < n$

$\forall X, C$

(Exc) - $H(X) - 2\lg H(X) \leq \mathbb{E}[C(x)]$

$H(X) - 2 \leq \mathbb{E}[C(x)]$

code $\to$ prefix-free

length

"self-delimiting"

00 11 01

- code $C$ $[n,k,d]_q$

$C' = \{ \ \boxed{X} \ \}$

$[n,k,d]_q \to [n-1,k,d-1]_q$

$d \geq 2$ $\qquad |\mathbb{F}| = q$

- code

most frequent symbol, keep codewords with that symbol

$C' = \{ \ \boxed{X} \ \}$

ith

$[n-1, k-1, d]_q$

$\lg_q$

$[n,k,d]_q$ code $C$ $\qquad |\mathbb{F}| = (q')^s$

s

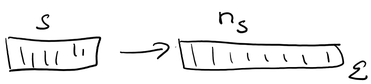
$q'$

whole string consists of $ns$ symbols form alphabet of size $q'$.

$q^k = (q'^s)^k$

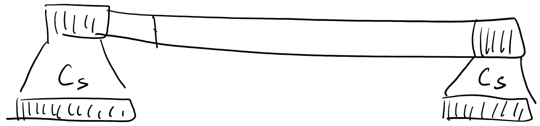
... string consists of $n_s$ symbols form ~~stg~~ alphabet of size $q'$!

$$q^k = (q'^s)^k = q'^{s\cdot k}$$




$C_s\ [n_s, s, d_s]_{q'}$

C



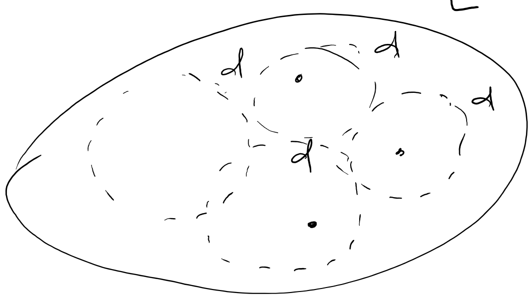

$$[n\cdot n_s,\ k\cdot \underline{\underline{s}},\ d\cdot d_s]_{q'}$$

$\Rightarrow C \circ C_s$ ("C concatenated with $C_s$")

1) $d + d_s$
2) $d \cdot d_s$ 3 ✓
3) $\frac{d}{d_s}\cdot s$

binary alphabet : $[n, \boxed{k}, \boxed{d}]_2$

can it be ?




$\{0,1\}^n$

How may codewords can we pack in $\{0,1\}^n$ ?

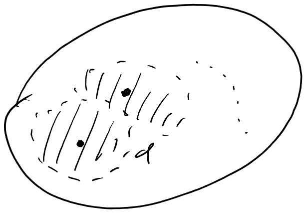
pick $n, d \rightarrow k$ ?

$$\frac{2^{H(\frac{d}{n})n}}{n} \leq Vol(n,d) \leq 2^{H(\frac{d}{n})n}$$

$$\frac{2^n}{2^{H(\frac{d}{n})n}} \leq 2^k$$

$$k \geq \left(1 - H\left(\frac{d}{n}\right)\right)n$$

greedy build the code




$\frac{2^n}{2^{H(\frac{d}{n})n}}$ exhaust the space

$k \geq 1 - H(\frac{d}{n})n$ rate $\geq 1 - H\left(\frac{d}{n}\right)$

$d = \varepsilon n$ $[n\ (1 - H(\varepsilon))n\ \ \varepsilon n]$

$d = \varepsilon n$ $\quad [n, [1-H(\varepsilon)]n, \varepsilon n]_2$

$\rightarrow$ correct $\boxed{\frac{\varepsilon n}{2}}$ errors

Shannon $[n, [1-H(\varepsilon)], \boxed{\varepsilon n}]_2$ $\approx$ correctable errors

---

types of errors:
- bit flips
- removal / insertion of bits

} edit distance

Ex: $\boxed{010101 \dots 01}$ Hamming distance $n$

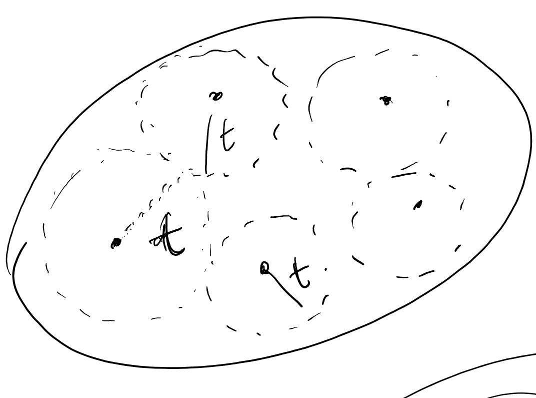
$\boxed{101010 \quad 10}$ edit distance $\leq 2$

$x, y$ ... edit distance – # bit flips & ins/del's to from $x$ to $y$.

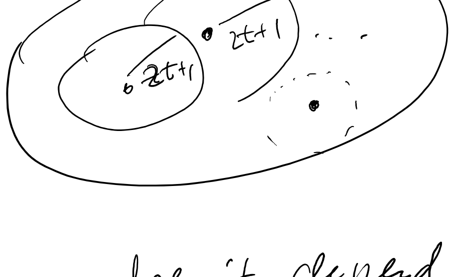
symmetric, triangle ineq.

$x \rightarrow y$

$(x \rightarrow z) + (z \rightarrow y) \geq (x \rightarrow y)$




$\{0,1\}^n$ + $2t+1$




$\{0,1\}^n$

$Vol_{edit}(n, t)$ $\quad$ • does it depend

$Vol_{edit}(n,t)$

- does it depend on starting pt? Yes
- how big it it?

$11 \xrightarrow{1} 111$ $\quad 1$

$01 \quad \cancel{1}01 \quad 0\ 1$

$\quad\quad 011$

$\begin{array}{c} 1 \quad n \\ 111111111 \\ \hline 0101010101 \\ n \quad n+n \end{array}$

$t$ op's each $\leq (2n)^t \cdot n^t \cdot n^t = 2^t \cdot 2^{3t \lg n}$

$t$ small $t \approx \sqrt{n}$

$t = \varepsilon n$ $\quad 2^{\varepsilon n} \cdot 2^{2H(\varepsilon)n} \quad 2^{H(\varepsilon)n} \cdot 2^{H(\varepsilon)n} \simeq 2^{\varepsilon n} \cdot 2^{4H(\varepsilon)n}$

$n$

$t$ places to remove

$\frac{1}{100}$

$<< 2^{(1-\delta)n}$

for some $\delta > 0$

$Vol(n,t)$

edit distance $\rightarrow [2^{\delta n}, \varepsilon n]_2$ $\quad n, \delta n,$