CS229r: Info Theory in CS

LECTURE 7

2/19/2019

TODAY

CHANNEL CODING (cont'd.)

1. Converse Coding Theorems
2. Efficient algorithmics
3. Linear Coding + Linear Compression

Review of last time

- General Channel

specified by $P_{Y|X}$

- Capacity = max. rate of inf. transmission with error → 0
  - Capacity formula = $\max_{P_X} \{ I(X;Y) \}$
  - achieved by random coding $E_n : \{0,1\}^{Rn} \to \Omega_X^n$.

$E_n(m)_i \sim P_X$ i.i.d. over (m,i)

Today: $R \leq \underbrace{\max_{P_X} \{ I(X;Y) \}}_{C} + \epsilon \quad \forall \epsilon > 0$.

Proof: Recall

$$m \longrightarrow X^n \longrightarrow Y^n \longrightarrow \hat{m}$$

$$H(m) = nR = H(m|\hat{m}) + I(m;\hat{m})$$

$$\leq H(m|\hat{m}) + I(X^n;Y^n) \quad [\text{D.P.I.}]$$

$$\leq H(m|\hat{m}) + \underbrace{n \cdot \cancel{I(X;Y)}}_{n \cdot C} \quad [\text{Chain Rule}]$$

$$H(m|\hat{m}) \leq h(\Pr[m \neq \hat{m}]) + \Pr[m \neq \hat{m}] \cdot nR$$

$$\leq 1 + \epsilon \cdot nR$$

$$\Rightarrow \quad nR \leq \frac{n \cdot C}{1 - \epsilon} \qquad \boxtimes$$

Caveat with Proof: only rules out $o(1)$ ~~constant~~ error prob.

What about fixed constant? Error $\to 1$?

All ruled out for BSC(p). Exercise!

Efficient Algorithms?

- Currently we have
  1. Prob. Exponential time + Exp. Space algorithm for encoding
  2. Det. Exponential time + Space for decoding.

- Can we do better? ~~Today: 3 answers~~

  Historically 3 answers
  1. YES .... weakest, folklore
  2. YES .... Weakish, Forney '66
  3. YES ... Strong, Arikan, Guruswami-Xia, '08 Hassani et al 2013.

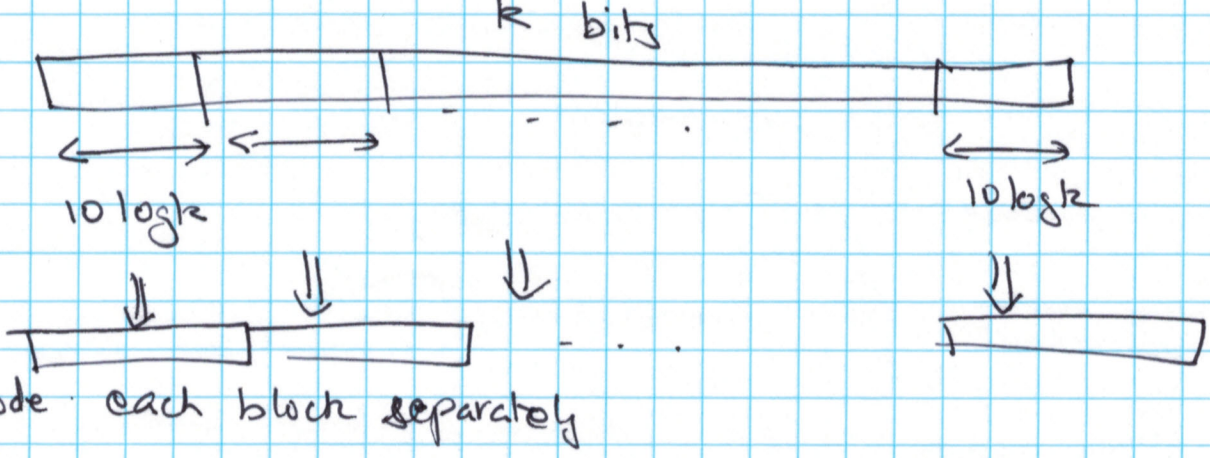
     "Polar Codes".

Will cover 1, mention 2 & skirt 3 today.

[Most ideas work more generally...
but will stick to BSC(p) here.]

Idea (1):

to encode k bits, break it up into chunks of length say $10 \log k$



$\ell = 10 \log k$. ~~find~~ ~~E_ℓ~~

& L be such that $\ell = (1 - h(p) - \epsilon) \cdot L$

find "best" $E_L : \{0,1\}^\ell \to \{0,1\}^L$

Apply E_L separately to each block.

(1) Preprocessing time : $\text{poly}(2^\ell) = \text{poly}(k)$ (randomized)

(2) Encoding : $\text{poly}(n)$ space ; ~~linear~~ $\text{poly}(n)$ time

[brute force on each block]

(3) Decoding : "

(4) Error Prob: $\frac{k}{10 \log k} \times \exp(-10 \log k) \leq \frac{1}{k^2}$ ... ☒

Bad news about (1)

(A) Error Prob $\geq \frac{1}{\text{Running time}}$

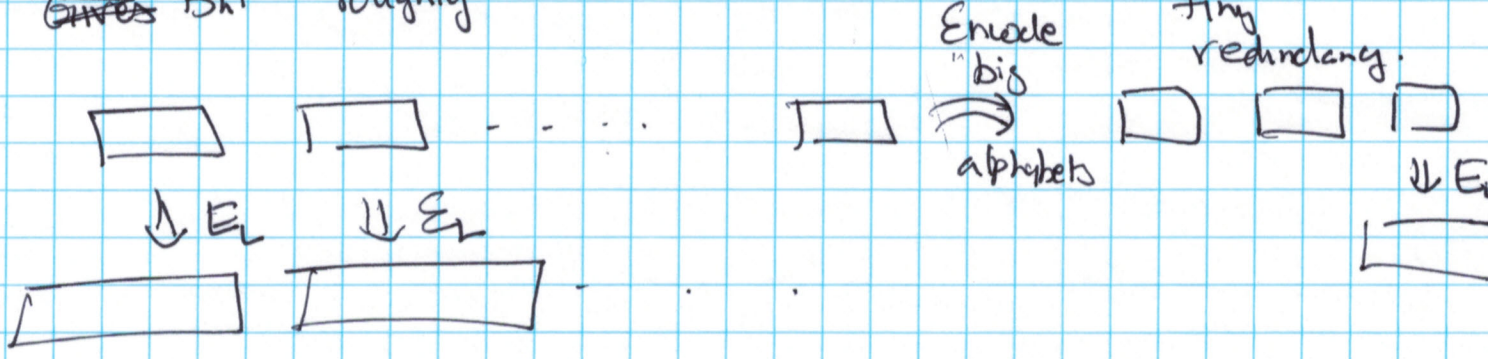
(B) Running time $\geq$ Exp$(\frac{1}{\epsilon^2})$

[to get ϵ close to capacity, need $10 \log k \geq \frac{1}{\epsilon^2}$]

Prob (A) fixed by Forney "concatenated codes"

[ideas out of scope for us...]

~~Given~~ But roughly



To get decoding wrong need linear fraction of ~~errors~~ blocks wrong

$\Rightarrow$ happens w.p. exp in # blocks.

Reduces error prob; but complexity of decoding block remains

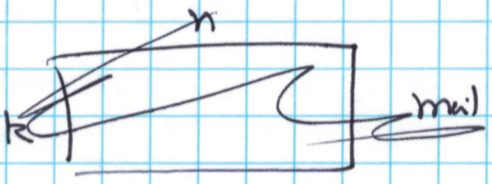
To get around $1/\epsilon^2$ problem, need to do something different.... leads to Polar Coding

---

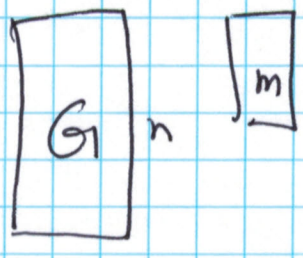
Idea 0: Linear Coding + Compression

Claim: $E_n: \{0,1\}^k \to \{0,1\}^n$ need not be totally random

can pick $E_n =$



matrix



$E_n(m) = G \cdot m$ [algebra over $\mathbb{F}_2$ = field on 2 elements (+, * mod 2)]

Claims: (1) Random E_n yield capacity also! [Exercise, uses p.w.i.]

(2) Even yield $\forall m$ as opposed to $\Pr_m[\ ]$...

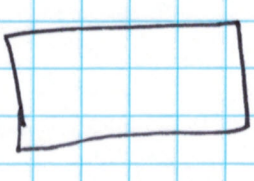
(3) other structural value too.

Simple check for errors.

Claim 1: Given $G \in \mathbb{F}_2^{n \times k}$ $\exists$ $H \in \mathbb{F}_2^{(n-k) \times n}$

full rank

full rank



s.t. $H \cdot G = 0 \in \mathbb{F}_2^{(n-k) \times k}$.

Claim 2: $X = E(m)$ iff $H \cdot X = 0$

Proof: $\Rightarrow X = G \cdot m \Rightarrow H \cdot X = H \cdot G \cdot m = 0$.

$\Leftarrow$ rank...

Will design good G by designing good H.

Will ask: if H a good compressor for $Z \sim \text{Ber}(p)^n$?

good for $\text{Ber}(p)^n$

Defn: H is a linear compressor if $\exists$ efficient decompressor D

s.t. $\Pr_{Z \sim \text{Ber}(p)^n}\left[D(H \cdot Z) \neq Z \right] \leq \delta(n) \to 0$.

~~Coding~~ Want H s.t.

(1) $m \triangleq n-k \leq (h(p)+\epsilon)\cdot n$; ~~$n = n_0$~~ $\forall n \geq n_0 = \text{poly}(1/\epsilon)$.

(2) D is efficient

———— ∞ ————

Why this helps

- Given m : encode $m \longmapsto G\cdot m$
- Decode: given $Y = G\cdot m + Z$

  - $HY = HGm + HZ = HZ$
  - w.h.p. $D(HY) = Z$

    so decoding $= Y - D(HY)$ yields $Gm \longrightarrow m$ (linear algebra).

running time, $\Pr[\text{error}]$ all depend only on $D(\cdot)$.

Next few lectures: Linear Compression with efficient decoding.

⊠