STO 8 LECTURE 10 3/10/08

Note Title 3/10/2008

**TODAY:** ALGORITHMS IN CODING THEORY

- ALGORITHMIC PROBLEMS
- ERASURE DECODING
- DECODING REED SOLOMON CODES
- ABSTRACTING FOR OTHER ALGEBRAIC SETTINGS

## ALGORITHMIC PROBLEMS

1. ENCODING:
2. DETECT ERRORS:
3. CORRECT ERASURES

Easy? Subtle!

4. CORRECT ERRORS - What? How? How many?

generator matrix

Complexity of 1., 2., 3.

- "Easy" for linear codes, given $G$

**ENCODE:** $m \in \{0,1\}^k$

$\downarrow$

$mG \in \{0,1\}^n$

**DETECT:** Compute $H$ s.t. $G \cdot H = 0$

Given $y$ output OK if $y \cdot H = 0$

**ERASURE CORRECT:** Given $r \in \{0,1,?\}^n$,

let $G' = G$ with ?'ed columns deleted.

$r' = r$ ⟵ " ⟶

Then $m$: $mG' = r'$

**Claim:** if $\#\{?\} < \Delta(C)$ then $m$ is uniquely determined.

# SUBTLETIES ?

1. Works only for linear codes...
2. Assumes $G$ known.
3. Assumption 2 not always valid. E.g. GV bound (Valid when you prove it! Construct $G$ uniformly, efficiently)
4. Is problem well defined for non-linear codes? My "defn.": Code is "constructive" if can construct encoding circuit $C: \{0,1\}^k \to \{0,1\}^n$ in poly time.

# DECODING?

- Not so simple ... can't seem to handle generic $G$
- Even for special $G$'s -- can only correct limited # errors ....
- But these may be sufficient for

Hamming / Shannon / (Elias)

<u>Shannon Problem</u>: Given $y$ find $m$ that maximizes $\Pr[y/m]$.

O.K. to be wrong on some $y$ ....
Provided $\Pr[y]$ (exponentially) small.

[Average case complexity / Worst-case]

**Hamming Problem**: Given $y$ find $m$ that minimizes $\Delta(y, mG)$.

O.K. to be "wrong" if

$$\min_m \{\Delta(y, mG)\} > \underbrace{\frac{\Delta(C)}{2}}$$

or some other $t$

But if # errors $\leq t$, must get it right!

[Worst-Case Complexity]

[if $t > \frac{\Delta(C)}{2}$ .... can produce small list including every $m$ s.t. $\Delta(y, mG) \leq t$]

Seems hard, but can be done ..... e.g. for RS Codes

[Peterson, Berlekamp, Massey, S., Guruswami-S.]

# REED-SOLOMON DECODING

## PROBLEM

Given: RS Code = $(\mathbb{F}, \alpha_1, \ldots \alpha_n, k)$

$r = (r_1 \ldots r_n) \in \mathbb{F}^n$

Output: (List of all) Deg. $k-1$ poly $p(x) = \sum_{i=0}^{k-1} m_i x^i$

$\#$ errors $= \{ i \mid p(\alpha_i) \neq r_i \} \leq t$

[Today: $t \leq \frac{n-k}{2}$; So $0 \leq$ list-size $\leq 1$.]

Algorithm: PETERSON 1960 : Defined "P"

WELCH-BERLEKAMP 1986

GEMMELL-SUDAN 1992

[Kinder, gentler...]

KEY IDEA: "ERROR LOCATOR POLYNOMIAL".

Define: $Err \triangleq \{ i \mid p(\alpha_i) \neq r_i \}$

(WARNING: Don't know $p(\cdot)$ & so don't know $Err$! But still ....)

$$E(x) \triangleq \prod_{i \in Err} (x - \alpha_i)$$

(EXTENDED WARNING: Don't know $E(x)$ either ....)

$E(x)$ has nice properties

## Properties of $E(x)$

1. $\forall i, \quad p(\alpha_i)\cdot E(\alpha_i) = r_i \cdot E(\alpha_i)$
2. Deg $E \le t$ ; $E \ne 0$
3. $N(x) \triangleq p(x).E(x)$ is a poly of deg $N \le t+k-1$

1'. $\forall i \quad N(\alpha_i) = p(\alpha_i)\cdot E(\alpha_i) = r_i \cdot E(\alpha_i)$

Algorithm: Ignore all references to "p" above & find $(N, E)$ !

Step 1: find "any" $N, E$ s.t.

(i) $\forall i \quad N(\alpha_i) = r_i \cdot E(\alpha_i)$

(ii) deg $N \le k+t-1$ ; deg $E \le t$; $E \ne 0$

Step 2: Output $p(x) = \dfrac{N(x)}{E(x)}$ (if deg $k-1$ poly)

**Analysis**: Correctness? Efficiency?

Efficiency: **Step 1**: just a big linear system

**Step 3**: Ratios .... Long Division.

Correctness:

**Lemma 1**: $\exists (N, E)$ satisfying (i), (ii) provided # errors is $\leq t$; with $N/E = p$.

**Proof**: Take $E$ to be error locator; $N \triangleq E \cdot p$.

**Lemma 2**: if $\exists$ two pairs $(N_1, E_1)$, $(N_2, E_2)$ satisfying (i), (ii); then $\frac{N_1}{E_1} = \frac{N_2}{E_2}$ (provided $n \geq k+2t$) $\Leftrightarrow N_1 \cdot E_2 = N_2 \cdot E_1$

**Proof**: $\forall i$

$$\left.\begin{aligned} N_1(\alpha_i) E_2(\alpha_i) &= r_i E_1(\alpha_i) \cdot E_2(\alpha_i) \\ &= E_1(\alpha_i) \cdot N_2(\alpha_i) \end{aligned}\right\} \text{But both are deg. } k+2t-1$$

Identical $\Leftarrow$ agree at $n > k+2t-1$ points $\Leftarrow$ polys.

## Abstraction :

Key property of polynomials :

Product of $d_1, d_2$ deg polys is $d_1+d_2$ poly.

**Abstraction** : for $U, V \in \mathbb{F}^n$

- let $U \star V \triangleq (U_1V_1, U_2V_2, \ldots, U_nV_n)$

  = coordinate-wise product.

- For sets $S, T \subseteq \mathbb{F}^n$

  $S \star T \triangleq \{ U \star V \mid U \in S, V \in T \}$

**Key Property** : $\left.\begin{array}{l} - S = \text{RS Code, dim } k \\ - T = \text{RS Code, dim } t+1 \end{array}\right\} S \star C \leq \underset{k+t}{\text{dim}}$

(Generically ... expect dim $k \cdot t$)

# Abstract Decoding

Given linear code $C = [n, k, d]$

$(\mathcal{E}, \mathcal{N})$ for a $t$ - error locating pair

if (i) $\dim(\mathcal{E}) > \text{large}_1$

(ii) $\text{distance}(\mathcal{N}) > \text{large}_2$

(iii) $\mathcal{E} * C \subseteq \mathcal{N}$

(iv) $\text{dist}(\mathcal{E}) > \text{large}_3$

Exercise: Fill these values in

---

**Algorithm**: Given: $r = (r_1 \ldots r_n)$ ;

(i) Find $E \in \mathcal{E}$ ; $N \in \mathcal{N}$ ; $E \neq 0$ s.t.

$$E * r = N$$

(ii) Let $y_i = r_i$ if $E_i \neq 0$ ; & $y_i = ?$ o.w.

ERASURE-DECODE$(y)$.

**Claim:** Yields decoding algorithms for all algebraic codes; Corrects roughly $d/2$ errors. E.g. even for BCH codes !! (Exercise)

**Notes:**

1. Good News: Can correct $d/2$ errors in RS codes.
(Algorithm not totally intuitive ---- will see more later.)

2. Actually correct $s$ erasures & $t$ errors, simultaneously provided $s + 2t < d$

3. Can now imm. correct $\frac{d_1}{2} \cdot \frac{d_2}{2}$ errors in concatenated codes if outer = RS of dist $d_1$ & inner = dist. $d_2$ code

Will show this & do better next time.