Lower bounds

via

Communication Complexity

recall:

Linear functions:

f "linear" iff $\forall x,y \quad f(x)+f(y)=f(x+y)$

last time we used n instead of d

Today consider $f: \{0,1\}^d \to \{0,1\}$

linear fctns $= \{ f \mid \exists S \subseteq [d] \text{ s.t. } f(x)=\bigoplus_{i\in S} x_i \}$

"parity fctns on d vars"

equivalently $\exists b$ s.t. $f(x) = (x_1 \cdots x_d)\begin{pmatrix} b_1 \\ \vdots \\ b_d \end{pmatrix}$

where inner product is $\sum x_i b_i \pmod 2$

note that role of x & b is symmetric (can cause confusion)

New definition:

f is "k-linear" if

(1) linear

(2) depends on k vars

i.e. $|S|=k$

also called "k-junta" fctn

For $f: \{0,1\}^d \to \{0,1\}$

$$k\text{-linear fctns} = \{ f \mid \exists\, S \subseteq [d],\ |S|=k,\ \& \ f(x) = \bigoplus_{i \in S} x_i \}$$

related to testing if fctn is k-junta (depends only on k vars), low Fourier degree, computable by small depth decision trees ...

A tester ("learns" f) wlog assume $f(\bar{0})=0$

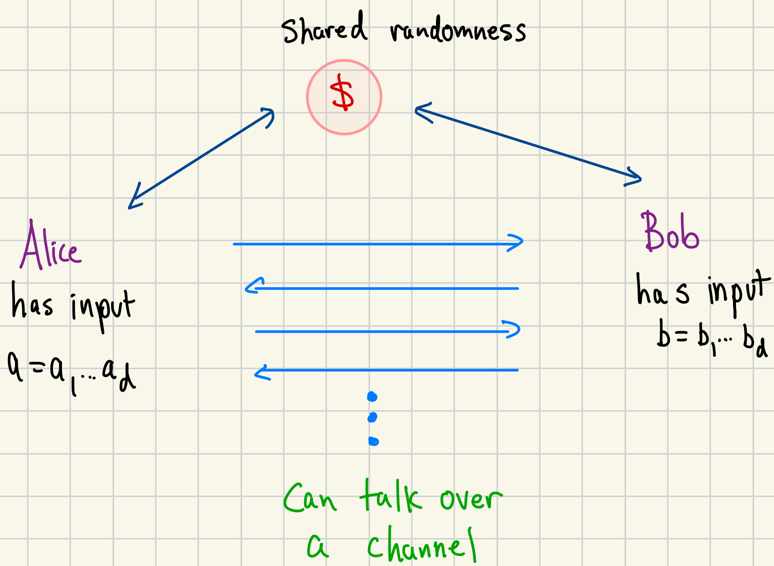
else not linear

- Query f on all $e_i = (000 \cdots 010 \cdots 0)$ for $i = 1..d$

  i^{th} position
- $S \leftarrow \{ i \mid f(e_i) = 1 \}$
- if $|S| \neq k$ fail
- else test if $f(x) = \bigoplus_{i \in S} x_i$ for most x via sampling

$O(d)$ queries. Can we do better?

What is Communication Complexity?



Goal: Compute $f(a, b)$

what does "compute" mean?

do both Alice + Bob need to know f?

how many bits, rounds required?

examples:

1) $f(a,b) = \left(\bigoplus_i a_i\right) \oplus \left(\bigoplus_i b_i\right)$

2 round 2 bit protocol: $A \to B$: $\bigoplus_i a_i$
$B \to A$: $\bigoplus_i b_i$ (or $f(a,b)$)

2) $f(a,b) = \sum a_i + \sum b_i$

integer addition

2 round, $O(\log n)$ bit protocol: $A \to B$: $\sum a_i$
$B \to A$: $\sum b_i$ (or $f(a,b)$)

3) $f(a,b) = \begin{cases} 1 & \text{if } a=b \\ 0 & \text{o.w.} \end{cases}$

requires $\Theta(\log d)$ bits with shared randomness

2 round protocols based on polynomial identity testing, Chinese remainder theorem ...

4) $f(a,b) = \begin{cases} 1 & \text{if } \exists i \text{ s.t. } a_i = b_i = 1 \\ 0 & \text{o.w.} \end{cases}$

requires $\Theta(d)$ bits of communication we have these!!

Communication Complexity lower bounds

$\Downarrow$

Property testing lower bounds

Idea: give reduction from C.C. problem to P.T. problem

lots of great work done in this area

$\Longrightarrow$ lower bnd for C.C. problem yields lower bnd for PT problem

so we get this almost for free!

Example: A hard C.C. problem:

SET DISJOINTNESS:

bit vectors represent membership in sets

Alice

$a \in \{0,1\}^d$

Bob

$b \in \{0,1\}^d$

do A & B agree on any 1-bit?

$$\text{DISJ}(a,b) = \bigvee_{i=1}^{d} (a_i \wedge b_i)$$

Known lower bound: $\Omega(d)$

SPARSE SET DISJOINTNESS:

a & b have at most k 1's

lower bound: $\Omega(k)$

easy upper bnd: send k indices. $O(k \log d)$

(even if guaranteed that x & y intersect ≤ 1 time!)

how can we use this to lower bound our property testing problems?

Reduction from sparse set disjointness to property tester for 2k-linearity

both Alice + Bob can query

$\boxed{\text{Shared Randomness}}$

Alice

Given a bit vector $a \in \{0,1\}^d$ with exactly $\underbrace{\text{k ones}}$

set $A = \{i \mid a_i = 1\}$

defines fctn $f(x) = \bigoplus_{i \in A} x_i$

Bob

Given a bit vector $b \in \{0,1\}^d$ with exactly $\underbrace{\text{k ones}}$

set $B = \{i \mid b_i = 1\}$

defines fctn $g(x) = \bigoplus_{i \in B} x_i$

Question does $h \equiv f \oplus g$ have 2k-linearity property?

notice:

$$\left.\begin{array}{l} \text{if } A \cap B = \emptyset \text{ then } h \text{ is } 2k\text{-linear} \\ \text{if } A \cap B \neq \emptyset \text{ then } h \text{ is } j\text{-linear for } j \le 2k-2 \end{array}\right\}$$

answer is "Yes" exactly when $A \cap B = \emptyset$

<u>example</u> (1) if $a = \{1,1,0,0\}$ $\quad b = \{0,0,1,1\}$

$A = \{1,2\}$ $\quad B = \{3,4\}$

$$A \cap B = \phi$$

$f(x) = x_1 \oplus x_2$ $\quad g(x) = x_3 \oplus x_4$

$h(x) = x_1 \oplus x_2 \oplus x_3 \oplus x_4$ $\leftarrow$ 4-linear

(2) if $a = \{1,1,0,0\}$ $\quad b = \{0,1,1,0\}$

$A = \{1,2\}$ $\quad B = \{2,3\}$

$$A \cap B = \{2\}$$

$f(x) = x_1 \oplus x_2$ $\quad g(x) = x_2 \oplus x_3$

$h(x) = x_1 \oplus \underbrace{x_2 \oplus x_2}_{=1} \oplus x_3 = x_1 \oplus x_3$

2-linear

<u>Observe</u>

for each $i \in A \cap B$, get $x_i \oplus x_i = 1$

so <u>two</u> variables drop out of h

$\Rightarrow$ h is $(2k - 2 \cdot |A \cap B|)$-linear.

So $|A \cap B| > 0 \Rightarrow$ not linear, but how <u>far</u> from linear?

Not 2k-linear $\Rightarrow$ far from 2k-linear:

<u>Fact</u> if $h_1 \neq h_2$ are 2 linear fctns (for any k)

then
$$\frac{\#\, x \text{ s.t. } h_1(x) \neq h_2(x)}{2^d} = \frac{\#\, x \text{ s.t. } h_1(x) = h_2(x)}{2^d}$$
$$= 1/2$$

(will prove soon)

$\Longrightarrow$ if $A \cap B \neq \emptyset$, h is $\frac{1}{2}$-far from 2k-linear

$\uparrow$
h is linear but not 2k linear

why is this interesting?

will demonstrate protocol for testing 2k-linearity using q queries

(above reduction)
$\Longrightarrow$ c.c. protocol for set-disjointness of A, B

$\boxed{\text{shared random string which contains random bits for A's queries}}$ $\}$ R

Alice runs prop test algorithm. When A needs to query $h(x) \equiv f(x) \oplus g(x)$:

1) A computes $f(x)$
2) asks Bob for $g(x)$
3) $h(x) \leftarrow f(x) \oplus g(x)$

"What is answer to my next question?" + $f(x)$ $\longrightarrow$

$\xleftarrow{g(x)}$

⋮

Bob simulates A's run on R

Bob computes x + then $g(x)$

Note: Alice doesn't need to send x's just $f(x)$, since Bob can compute x's from R

x's are d bits

$\leftarrow$ $f(x)$ is only 1 bit

Total communication:

$2q$ bits

so, reduction of set disj to k-lin testing

$\Rightarrow$ set disjointness needs only $2q$ bits $\leftarrow$ but we know l.b. of $\Omega(k)$

$\Rightarrow q = \Omega(k)$

Thm k-linearity testing requires $\Omega(k)$ queries

(but linearity testing only requires $O(1)$ queries!)

Remains to prove this:

<u>Fact</u> if $h_1 \neq h_2$ are 2 linear fctns (for <u>any</u> k)

then $\dfrac{\#\ x \text{ s.t. } h_1(x) \neq h_2(x)}{2^d} = \dfrac{\#\ x \text{ s.t. } h_1(x) = h_2(x)}{2^d}$

$= 1/2$

(for general domains/ranges we get $\geq 1/2$)

<u>Proof</u>

Given $h_1(x) = \bigoplus_{x \in S_1} x_i$ & $h_2(x) = \bigoplus_{x \in S_2} x_i$

if $h_1 \neq h_2$, $\exists$ i s.t. $i \in S_1 \Delta S_2$ ← $\equiv (S_1 \setminus S_2) \cup (S_2 \setminus S_1)$ "symmetric difference"

wlog, assume $i \in S_1 \setminus S_2$

pair inputs $x, x' \in \{0,1\}^d$

s.t. $x = x' \oplus \underbrace{(0 \ldots 010 \ldots 0)}_{e_i}$ ← i^{th} bit (so x & x' differ <u>only</u> on i^{th} bit)

note $\forall$ pairs, $h_1(x) \neq h_1(x')$ ← since only i^{th} bit differs in x, x' & $i \in S_1$, $i \notin S_2$

but $h_2(x) = h_2(x')$

↑ since $i \notin S_2$

So exactly one of

$$(h_1(x) = h_2(x)) \;\&\; (h_1(x') = h_2(x')) \text{ hold}$$

$$\Rightarrow \frac{\#\, x \text{ s.t. } h_1(x) \neq h_2(x)}{2^d} = \frac{1}{2}$$

$\blacksquare$