31.1.2008

Linearity

Consider a Boolean function $f:\{0,1\}^n \to \{0,1\}$.

Def: Such an f is linear iff $\exists a_1,\dots,a_n \in \{0,1\}$ s.t. $f(x_1,\dots,x_n) = \sum_{i=1}^{n} a_i x_i$ (or equivalently, $\exists S \subseteq [n]$ s.t. $f(x_1,\dots,x_n) = \sum_{i\in S} x_i$). Another def. $\forall x,y\ f(x+y) = f(x)+f(y)$.

Claim: These definitions are equivalent.

Proof: $1\Rightarrow 2$: $f(x+y) = \sum a_i(x_i+y_i) = \sum a_i x_i + \sum a_i y_i = f(x)+f(y)$. $2\Rightarrow 1$: Define $a_i = f(e_i)$ where $e_i = (0,\dots,0,\overset{i}{1},0,\dots,0)$, then $f(x_1,\dots,x_n) = f(\sum x_i e_i) = \sum_i x_i f(e_i) = \sum_i a_i x_i$ and also $f(0,\dots,0)=0$. ■

There are 2^n linear functions, out of 2^{2^n} functions.

Approx. Linearity

Def: f is ε-close to linear if $\exists$ a linear function g s.t. $f(x)=g(x)$ for $1-\varepsilon$ of all $x\in\{0,1\}^n$. (i.e., f is ε-close to g)

Def: f is ε-approx. linear if $f(x)+f(y)=f(x+y)$ for $1-\varepsilon$ of $(x,y\in\{0,1\}^n)$

Claim: $\forall \varepsilon, f$, if f is ε-close to linear, then f is also 3ε-approx. linear.

Proof: Let g be a linear function that is ε-close to f. If we choose $x,y\in\{0,1\}^n$ uniformly at random then by union bound, we have that with prob. $1-3\varepsilon$, $f(x)=g(x)$, $f(y)=g(y)$ and $f(x+y)=g(x+y)$, where we use that $x+y$ is also uniformly distributed. In such a case, $f(x+y)=f(x)+f(y)$. ■

The opposite direction (ε-approx $\Rightarrow$ ε-close) is not clear. Our previous proof doesn't work: take any linear function and modify its values on $e_1, e_2, \dots, e_n$.

Testing Linearity - The BLR Test [Blum, Luby, Rubinfeld 90]

We are given a black-box computing a Boolean func. f and we want to check that it is close to linear. Procedure:
- Pick $x,y\in\{0,1\}^n$ uniformly
- Query $f(x), f(y), f(x+y)$
- Accept iff $f(x+y)=f(x)+f(y)$

Claim: If f is ε-close to linear, then BLR accepts w.p. $\geq 1-3\varepsilon$ (as we saw)

Thm: [BLR 90, Bellare Coppersmith Hastad Kiwi Sudan 95] If f passes the BLR test w.p. $\geq 1-\varepsilon$ then f is ε-close to linear.

This implies that if we repeat the BLR test $O(1/\varepsilon)$ times, we can be 99% convinced that f is ε-close.

Fourier Expansion First, from now on the range of Boolean functions is $\{-1,1\}$ where $0/\text{FALSE} \mapsto 1$, $1/\text{TRUE} \mapsto -1$, addition (mod 2) $\mapsto$ multiplication on $\mathbb{R}$. So a Boolean function is $f:\{0,1\}^n \to \{-1,1\}$.

The linear parity functions are $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$. We can think of the set of functions $\{0,1\}^n \to \mathbb{R}$ as a vector space of dimension 2^n (i.e., $\mathbb{R}^{2^n}$).

Def: For $f,g:\{0,1\}^n \to \mathbb{R}$ define $\langle f,g\rangle = \frac{1}{2^n}\sum_{x\in\{0,1\}^n} f(x)\cdot g(x) = \mathbb{E}_x[f(x)g(x)]$. Also define $\|f\|_2 = \|f\| = \sqrt{\langle f,f\rangle}$.

Claim: For $f,g:\{0,1\}^n \to \{-1,1\}$, $\langle f,g\rangle \in [-1,1]$ (the correlation between f and g). Moreover, $\langle f,g\rangle = 1-\frac{2}{2^n}\mathrm{HamDist}(f,g)$. In particular f is ε-close to g iff $\langle f,g\rangle \geq 1-2\varepsilon$.

Def: For $S\subseteq[n]$, define $\hat{f}(S) = \langle f,\chi_S\rangle$. **Remark:** $\hat{f}(\phi) = \mathbb{E}_x[f(x)]$, $\hat{f}(\{1\}) = \mathbb{E}_x[f(x)\cdot\chi_{\{1\}}] = \frac{1}{2}\left(\mathbb{E}_{x_1=0}[f(x)] - \mathbb{E}_{x_1=1}[f(x)]\right)$.

Prop: The set $\{\chi_S : S\subseteq[n]\}$ is an orthonormal basis for $\{0,1\}^n \to \mathbb{R}$.

Proof: For any S, $\|\chi_S\| = \sqrt{\langle\chi_S,\chi_S\rangle} = 1$. Now take any $S\neq T$ and let $i\in S\Delta T$. Assume w.l.o.g that $i\in S\setminus T$. Then $\langle\chi_S,\chi_T\rangle = \frac{1}{2^n}\sum_x \chi_S(x)\cdot\chi_T(x)$. We can partition the sum into 2^{n-1} sums, each over two elements x,x' that differ exactly in the ith coordinate. Then since $\chi_S(x) = -\chi_S(x')$ and $\chi_T(x) = \chi_T(x')$, we get $\chi_S(x)\chi_T(x) + \chi_S(x')\chi_T(x') = 0$ so also the entire sum is 0.

Finally, there are 2^n of them so they must be a basis.

Ex: Prove that $\{2^{n/2}\cdot\delta_x : x\in\{0,1\}^n\}$ where $\delta_x(y) = \begin{cases}1 & x=y\\ 0 & \text{o.w.}\end{cases}$ is an orthonormal basis.

Thm: Every function $f:\{0,1\}^n\to\mathbb{R}$ can be uniquely expressed as a linear combination of the parity functions χ_S and this combination is given by $f = \sum_S \hat{f}(S)\cdot\chi_S$.

Examples: • Constant func: $f(x)=1$. $\hat{f}(\phi)=1$, $S\neq\phi$: $\hat{f}(S)=0$, indeed $f = 1\cdot\chi_\phi$.

• Dictatorship func: $f(x) = (-1)^{x_i} = \chi_{\{i\}}$, so $\hat{f}(\{i\}) = 1$ and all others are 0.

• $f(x) = \mathrm{AND}(x_1,x_2) = \frac{1}{2}\chi_\phi - \frac{1}{2}\chi_{\{1\}} - \frac{1}{2}\chi_{\{2\}} - \frac{1}{2}\chi_{\{1,2\}}$

• $f(x) = \mathrm{MAJORITY}(x_1,x_2,x_3) = \frac{1}{2}\chi_{\{1\}} + \frac{1}{2}\chi_{\{2\}} + \frac{1}{2}\chi_{\{3\}} - \frac{1}{2}\chi_{\{1,2,3\}}$. $f(0) =$ סכום המקדמים

Thm (Plancherel): $\forall f,g:\{0,1\}^n\to\mathbb{R}$, $\langle f,g\rangle = \sum_S \hat{f}(S)\hat{g}(S)$.

Proof: $\langle f,g\rangle = \langle\sum_S \hat{f}(S)\chi_S, \sum_T \hat{g}(T)\chi_T\rangle = \sum_{S,T}\hat{f}(S)\cdot\hat{g}(T)\cdot\langle\chi_S,\chi_T\rangle = \sum_S \hat{f}(S)\hat{g}(S)$.

Cor: (Parseval) $\forall f:\{0,1\}^n\to\mathbb{R}$, $\|f\|^2 = \sum_S \hat{f}(S)^2$, $\forall f:\{0,1\}^n\to\{-1,1\}$, $\sum\hat{f}(S)^2 = 1$.

Analysis of BLR Test

BLR Test: Choose $x,y\in\{0,1\}^n$ uniformly and accept iff $f(x)\cdot f(y)\cdot f(x+y) = 1$. (3-query local test).

Thm: (1). If $f=\chi_S$ then BLR always accepts.

(2). If f is ε-far from linear then $\Pr[\text{BLR accepts } f] \leq 1-\varepsilon$.

Lemma: $\forall f:\{0,1\}^n\to\{-1,1\}$, $\Pr[\text{BLR accepts } f] = \frac{1}{2} + \frac{1}{2}\sum_S \hat{f}(S)^3$.

Proof of (2): Assume that $\Pr[\text{BLR accepts}] \geq 1-\varepsilon$. Therefore, $\frac{1}{2}+\frac{1}{2}\sum_S\hat{f}(S)^3 \geq 1-\varepsilon \Rightarrow \sum\hat{f}(S)^3 \geq 1-2\varepsilon$.

But, $\sum\hat{f}(S)^3 \leq \max(\hat{f}(S))\cdot\sum\hat{f}(S)^2 = \max(\hat{f}(S))$. So $\exists S$, s.t. $\hat{f}(S)\geq 1-2\varepsilon$ so f is ε-close to χ_S. ■

By repeating the BLR test $\frac{2}{\varepsilon}$ times, we get that if f is ε-far from linear, then it is caught w.p. $\geq 1-(1-\varepsilon)^{2/\varepsilon} > \frac{2}{3}$.

<u>Proof Lemma</u>: $\Pr_{x,y}[\text{BLR accepts}] = \mathbb{E}_{x,y}\left[\frac{1}{2}+\frac{1}{2}f(x)f(y)f(x+y)\right] = \frac{1}{2}+\frac{1}{2}\mathbb{E}_{x,y}[f(x)f(y)f(x+y)]$.

$$\mathbb{E}_{x,y}[f(x)f(y)f(x+y)] = \mathbb{E}_{x,y}\left[\left(\sum_S \hat{f}(S)\chi_S(x)\right)\left(\sum_T \hat{f}(T)\chi_T(y)\right)\left(\sum_U \hat{f}(U)\chi_U(x+y)\right)\right] =$$

$$= \sum_{S,T,U\subseteq[n]} \hat{f}(S)\cdot\hat{f}(T)\cdot\hat{f}(U)\cdot\underbrace{\mathbb{E}_{x,y}[\chi_S(x)\cdot\chi_T(y)\cdot\chi_U(x+y)]} = \sum_S \hat{f}(S)^3 \quad \blacksquare$$

$$\mathbb{E}_{x,y}[\chi_S(x)\cdot\chi_T(y)\cdot\chi_U(x)\cdot\chi_U(y)] = \mathbb{E}_{x,y}[\chi_{S\Delta U}(x)\cdot\chi_{T\Delta U}(y)] = \underbrace{\mathbb{E}_x[\chi_{S\Delta U}(x)]}_{\substack{S=U & 1\\ \text{[illegible]} & 0}}\cdot\underbrace{\mathbb{E}_y[\chi_{T\Delta U}(y)]}_{\substack{T=U & 1\\ \text{[illegible]} & 0}} = \begin{cases}1 & \text{if } S=T=U\\ 0 & \text{o.w.}\end{cases}$$

<u>Alternative proof</u>: <u>Def</u>: For $f,g:\{0,1\}^n\to\mathbb{R}$, define $f*g:\{0,1\}^n\to\mathbb{R}$ by $(f*g)(z) = \frac{1}{2^n}\sum_{x\in\{0,1\}^n} f(x)\cdot g(z-x) = \mathbb{E}_x[f(x)g(z-x)]$.

<u>Remark</u>: $(f*g)(0) = \langle f,g\rangle$.

<u>Claim</u>: $\widehat{f*g} = \hat{f}\cdot\hat{g}$. <u>Proof</u>: $\widehat{(f*g)}(S) = \frac{1}{2^{2n}}\sum_{x,z} f(x)\cdot g(z-x)\chi_S(z) = \frac{1}{2^{2n}}\sum_{x,z} f(x)g(z-x)\chi_S(x)\cdot\chi_S(z-x) =$

$= \frac{1}{2^{2n}}\sum_x f(x)\chi_S(x)\left(\sum_z g(z-x)\chi_S(z-x)\right) = \frac{1}{2^{2n}}\sum_x f(x)\chi_S(x)\cdot\left(\sum_y g(y)\chi_S(y)\right) = \hat{f}(S)\cdot\hat{g}(S)$. $\blacksquare$

Proof 2 of the lemma: $\mathbb{E}_{x,y}[f(x)f(y)f(x+y)] = \mathbb{E}_x[f(x)\cdot\mathbb{E}_y[f(y)f(x+y)]] = \mathbb{E}_x[f(x)\cdot(f*f)(x)] =$

$= \langle f, f*f\rangle = \sum_S \hat{f}(S)\widehat{(f*f)}(S) = \sum_S \hat{f}(S)^3$. $\blacksquare$

Testing Dictatorship

<u>Def</u>: The function $\chi_{\{1\}},\chi_{\{2\}},\dots,\chi_{\{n\}}$ are the dictator functions, $f_i(x) = \begin{cases}-1 & x_i=1\\ 1 & x_i=0\end{cases}$

<u>Hastad Test</u>: Fix parameter $\delta\in[0,1]$.
- Pick $x,y\in\{0,1\}^n$ uniformly, let $z=x+y$.
- Pick $w\in\{0,1\}^n$ according to μ_δ, which is the distribution in which each bit is 1 w.p. δ and 0 o.w.
- Accept iff $f(x)\cdot f(y)\cdot f(z+w)=1$.

<u>Example</u>: $\chi_{\{1\}}$ passes test w.p. $1-\delta$.

$\chi_{\{1,2\}}$ passes test w.p. $(1-\delta)^2+\delta^2 \approx 1-2\delta$.

χ_ϕ passes test w.p. 1. (problem!)

<u>Thm</u>: $\Pr[\text{test accepts}] = \frac{1}{2}+\frac{1}{2}\sum_S (1-2\delta)^{|S|}\cdot\hat{f}(S)^3$.

7.2.2008 <u>Def</u>: The dictator functions are $\chi_{\{1\}},\dots,\chi_{\{n\}}$ where $\chi_{\{i\}}(x) = \begin{cases}-1 & x_i=1\\ 1 & x_i=0\end{cases}$

<u>Hastad test with parameter $\delta\in[0,1]$</u>.

- Pick $x,y\in\{0,1\}^n$ uniformly
- Pick $w\in\{0,1\}^n$ according to μ_δ, i.e., each bit is 1 w.p. δ.
- Accept iff $f(x)\cdot f(y)\cdot f(x+y+w)=1$.

<u>Thm</u>: $\Pr[\text{Hastad accepts}] = \frac{1}{2}+\frac{1}{2}\sum_S (1-2\delta)^{|S|}\cdot\hat{f}(S)^3$.

Let $\varepsilon<\frac{1}{100}$ be some constant, and let $\delta=\frac{3}{4}\varepsilon$. If f is a dictator then accepts

[illegible] w.p. $1-\delta = 1-\frac{3}{4}\varepsilon$. Therefore, $1-2\varepsilon \le \sum_S (1-2\delta)^{|S|}\hat{f}(S)^3 \le \left(\sum_S \hat{f}(S)^2\right)\cdot\max_S\left((1-2\delta)^{|S|}\cdot\hat{f}(S)\right) =$

$= \max_S\left((1-2\delta)^{|S|}\cdot\hat{f}(S)\right)$.

Since $(1-2\delta)^2 < 1-2\epsilon$, $1-2\epsilon \le \max_{S:|S|\le 1}\left((1-2\delta)^{|S|}\hat{f}(S)\right) \le \max_{S,|S|\le 1} \hat{f}(S)$.

Therefore, f is ϵ-close to either dictator or constant 1.

To summarize, Hastad test does the following:

- If f is a dictator or constant 1 func then accepted w.p. $\ge 1-\frac{3}{4}\epsilon$.
- If f is ϵ-far from dictator or constant 1 then accepted w.p. $< 1-\epsilon$.

Thm 1 (Chernoff) If we sample from μ_p n times and let $\tilde{p} = \frac{\#1}{n}$, then $\forall \epsilon$, $\Pr(|p-\tilde{p}|>\epsilon) < 2^{-n\epsilon^2}$.

Thm 2 (Chernoff) If X is a distribution on some bounded interval $[a,b]$ then the average of $O((\log\frac{1}{\delta})/\gamma^2)$ samples is within $\pm\gamma$ of the expectation w.p. $\ge 1-\delta$.

- Therefore, by repeating Hastad test $O(1/\epsilon^2)$ times, we can make these probabilities 90% and 10%.

How to get rid of the const 1 func?

Solution 1 - Query f on k random locations. If we see $> 3/4$ 1s then reject.

The prob of rejecting a dictator is $< 2^{-\Omega(k)}$ and similarly, prob. of accepting ϵ-close to const 1 is $< 2^{-\Omega(k)}$. For $k>100$ this is enough.

To summarize, using $O(1/\epsilon^2)$ queries we can accept directors w.p. ≥ 0.9 and reject ϵ-far from dictators w.p. ≥ 0.9.

The Noise Operator (Beckner Operator)

For $-1\le\rho\le 1$, let $p=\frac{1-\rho}{2}$ and define the operator T_ρ mapping $\{0,1\}^n \to \mathbb{R}$ to itself, by $(T_\rho f)(x) = \underset{y\sim\mu_p}{E}[f(x+y)]$.

Remark: • $T_1 f = f$ $(\forall f)$, so T_1 is the identity operator.

• $\forall f$, $T_0 f =$ const. func. $E[f]$.

• T_ρ is a linear operator on the 2^n-dim. vector space $\{0,1\}^n\to\mathbb{R}$. $\forall f,g$. $T_\rho(f+g) = T_\rho f + T_\rho g$; $\forall \alpha\in\mathbb{R}$. $T_\rho(\alpha f) = \alpha T_\rho f$

Claim: $\forall S,\rho$, $T_\rho(\chi_S) = \rho^{|S|}\cdot\chi_S$. (ובפרט χ_S הם ו"ע)

Proof: $T_\rho\chi_S(x) = \underset{y\sim\mu_p}{E}[\chi_S(x+y)] = \underset{y\sim\mu_p}{E}[\chi_S(x)\cdot\chi_S(y)] = \chi_S(x)\cdot\underset{y\sim\mu_p}{E}[\prod_{i\in S}(-1)^{y_i}] = \chi_S(x)\cdot\prod_{i\in S}\underbrace{\underset{y_i\sim\mu_p}{E}[(-1)^{y_i}]}_{1-2p=\rho} = \rho^{|S|}\chi_S(x)$ ∎

Cor: $T_\rho f = \sum_S \rho^{|S|}\hat{f}(S)\chi_S$.

Proof: $T_\rho f = T_\rho(\sum \hat{f}(S)\cdot\chi_S) = \sum_S \hat{f}(S) T_\rho(\chi_S) = \sum_S \hat{f}(S)\cdot\rho^{|S|}\cdot\chi_S$. ∎

Cor: $T_{\rho_1} T_{\rho_2} f = T_{\rho_1\rho_2} f$

Analysis of Hastad

The success prob is : $\frac{1}{2}+\frac{1}{2}\mathop{E}_{x,y,w}[f(x)f(y)f(x+y+w)]$.

$$\mathop{E}_{x,y,w}[f(x)f(y)f(x+y+w)] = \mathop{E}_{x}\Big[f(x)\mathop{E}_{y}\big[f(y)\cdot \mathop{E}_{w\sim\mu_\delta}[f(x+y+w)]\big]\Big] = \mathop{E}_{x}\Big[f(x)\cdot \mathop{E}_{y}\big[f(y)\cdot (T_{1-2\delta}f)(x+y)\big]\Big] =$$

$$= \mathop{E}_{x}[f(x)\cdot(f * T_{1-2\delta}f)(x)] = \langle f, f * T_{1-2\delta}f\rangle = \sum_S \hat{f}(S)\cdot \widehat{f * T_{1-2\delta}f}(S) = \sum_S \hat{f}(S)\cdot\hat{f}(S)\cdot(1-2\delta)^{|S|}\hat{f}(S) =$$

$$= \sum_S (1-2\delta)^{|S|}\hat{f}(S)^3. \quad \blacksquare$$

<u>Ex (Solution 2 to const func issue)</u> : Show that if we replace the test by

$f(x+(1,\dots,1))\cdot f(y)\cdot f(x+y+w) = -1$ we get a test for dictatorship.

Testing Averages

Assume we have $f_1, f_2, \dots, f_d : \{0,1\}^n \to \{-1,1\}$ and we apply the Hastad test to them by sending each of the 3 queries to one of the d functions randomly and independently.

<u>Define</u>: $h = \frac{1}{d}\sum_{i=1}^{d} f_i$. By linearity, $\hat{h} = \frac{1}{d}\sum \hat{f}_i$.

<u>Thm</u>: The prob of accepting when applying Hastad to $\{f_1,\dots,f_d\}$ is $\frac{1}{2}+\frac{1}{2}\sum_S(1-2\delta)^{|S|}\cdot\hat{h}(S)^3$.

<u>Proof</u>: The prob of success is given by $\mathop{E}_{i,j,k}\Big[\frac{1}{2}+\frac{1}{2}\mathop{E}_{x,y,w}[f_i(x)f_j(y)f_k(x+y+w)]\Big] =$

$$= \mathop{E}_{x,y,w}\Big[\frac{1}{2}+\frac{1}{2}\mathop{E}_{i}[f_i(x)]\cdot\mathop{E}_{j}[f_j(y)]\cdot\mathop{E}_{k}[f_k(x+y+w)]\Big] = \mathop{E}_{x,y,w}\Big[\frac{1}{2}+\frac{1}{2}h(x)h(y)h(x+y+w)\Big] = \dots = \frac{1}{2}+\frac{1}{2}\sum_S(1-2\delta)^{|S|}\hat{h}(S)^3.$$

<u>Def</u>: For $f:\{0,1\}^n\to\mathbb{R}$ define $f^+(x) = -f(x+(1,\dots,1))$ and also $f^{odd} = \frac{1}{2}(f+f^+)$.

As we'll see in homework, $f^{odd} = \sum_{S,\,|S|\text{ odd}} \hat{f}(S)\chi_S$

Let the Hastodd$_\delta$ test be the Hastad test applied to $\{f, f^+\}$.

<u>Cor</u>: $\Pr[\text{Hast-odd}_\delta \text{ accepts } f] = \frac{1}{2}+\frac{1}{2}\sum_S(1-2\delta)^{|S|}\cdot\widehat{f^{odd}}(S)^3 = \frac{1}{2}+\frac{1}{2}\sum_{S,\,|S|\text{ odd}}(1-2\delta)^{|S|}\cdot\hat{f}(S)^3$.

This leads to a third solution to the const func problem.

The NAE Test

3 candidates, each voter gives his ranking (one of $6=3!$ possibilities), and results are summarized using some $f:\{0,1\}^n\to\{-1,1\}$

Condorcet's Paradox! If f is MAJ then we can get irrational outcome.

A>B	1	1	1	1	1	1	0	0	0	$\xrightarrow{maj}$	A>B !
B>C	0	0	0	1	1	1	1	1	1	$\xrightarrow{maj}$	B>C !
C>A	1	1	1	0	0	0	1	1	1	$\xrightarrow{maj}$	C>A !

<u>Thm (Arrow's impossibility theorem)</u>: The only functions f leading to rational outcomes are the dictators and the anti-dictators.

<u>NAE Test</u>:
- Choose $x,y,z\in\{0,1\}^n$ by choosing each coordinate independently subject to not $x_i=y_i=z_i$ (i.e., one of the 6 possibilities).
- Accept iff $NAE(f(x), f(y), f(z))$.

Lemma: $\Pr[\text{NAE accepts } f] = \frac{3}{4} - \frac{3}{4} \cdot \sum_S \left(-\frac{1}{3}\right)^{|S|} \cdot \hat{f}(S)^2$.

Denote $W_k = \sum_{|S|=k} \hat{f}(S)^2$. Then $W_k \geq 0$, and if f is Boolean, $\sum_{k=0}^{n} W_k = 1$.

Cor: $\Pr[\text{NAE accepts } f] = \frac{3}{4} - \frac{3}{4} W_0 + \frac{1}{4} W_1 - \frac{1}{12} W_2 + \frac{1}{36} W_3 - \ldots$

Hence, if accepts w.p. $\geq 1-\varepsilon$ then $W_1 \geq 1 - \frac{9}{2}\varepsilon$.

Proof: $\Pr[\text{NAE accepts}] \leq \frac{3}{4} + \frac{2}{9} W_1 + \frac{1}{36}(W_1 + W_2 + W_3 + \ldots) \leq \frac{7}{9} + \frac{2}{9} W_1$. ∎

Does $W_1 \geq 1 - \Omega(\varepsilon)$ imply that f is ε-close to dictator?

Thm [Friedgut, Kalai, Naor]: If $f : \{0,1\}^n \to \{-1,1\}$ has $\sum_{|S|=1} \hat{f}(S)^2 \geq 1-\varepsilon$ then f is $O(\varepsilon)$-close to dictatorship or antidictatorship.

Therefore, NAE, gives a 3-query local-test for dictatorship or antidictatorship.

In fact, we can avoid [FKN] and also get rid of the antidictators.

Thm: The property of being dictator is locally testable with 3 queries.

Proof: We combine NAE+BLR: We toss a coin and apply either NAE or BLR. Dictatorship are accepted w.p. 1. Assume now f is accepted w.p. $\geq 1-\varepsilon$. Therefore, both NAE and BLR accept f w.p. $\geq 1-2\varepsilon$. The latter implies that $\exists S^*$ s.t. $\hat{f}(S^*) \geq 1-4\varepsilon$, and this S^* must be unique (since $\sum \hat{f}(S)^2 = 1$). The former implies that $\sum_{|S|=1} \hat{f}(S)^2 \geq 1-9\varepsilon$ and hence $|S^*| = 1$ (since $\sum_1 \hat{f}(S)^2 = 1$). We see that f is 2ε-close to the dictator func χ_{S^*}.

Cor: For any $P \subseteq \{1,\ldots,n\}$, the property $\{\chi_{\{i\}} : i \in P\}$ is locally testable with 3-queries.

Proof: With prob $\frac{1}{2}$ do the BLR&NAE test, and otherwise let $x \in \{0,1\}^n$ be the indicator of P and do "local decoding" of $f(x)$ and accept iff result gives -1.
Dictatorships in P accepted w.p. 1. Assume f is accepted w.p. $\geq 1-\varepsilon$. This implies that BLR&NAE accepts w.p. $\geq 1-2\varepsilon$ and hence f is $O(\varepsilon)$-close to dictator. If it is close to $\chi_{\{i\}}$ for $i \notin P$ then w.p. $\Omega(\varepsilon)$ local decoding will give 1 and we reject. ∎