Testing Monotonicity

Oded Goldreich, Shafi Goldwasser, Eric Lehman, Dana Ron, Alex Samorodnitsky

Presented by Shlomo Jozeph

Definition of Monotonicity

- For $x=(x_1x_2\dots x_n)$, $y=(y_1y_2\dots y_n)\in\{0,1\}^n$, $x<y$ if for all i, $x_i\leq y_i$, and for some j, $x_j<y_j$.
- A function $f\in\{0,1\}^n\rightarrow\{0,1\}$ is monotone if for all $x<y$, $f(x)\leq f(y)$.
- A DNF formula with no negations over $\{x_i\}$.
- A function respecting the partial order defined by a directed boolean hypercube.

Testing Monotonicity

- There is an algorithm with query complexity $O(n/\varepsilon)$ that always accepts monotone functions and rejects function that are ε-far from monotone with constant probability.
- Known lower bound - $\Omega(n^{½})$ for 1-sided error, $\Omega(\log n)$ for 2-sided error.

The Algorithm (Single Step)

For $f \in \{0,1\}^n \rightarrow \{0,1\}$:

1. Uniformly at random select $i \in \{1,\ldots,n\}$ and $x \in \{0,1\}^n$.
2. If $f(x^i(0)) \leq f(x^i(1))$ accept, otherwise reject.

Where $x^i(b) = x_1 \ldots x_i b x_{i+1} \ldots x_n$.

Definitions

- $\delta(f)$ – The probability the algorithm rejects f.
- $\varepsilon(f)$ – The distance of f from the monotone functions.
- Claim:

$$\varepsilon(f)/n \leq \delta(f) \leq 2\varepsilon(f)$$

Analysis of the Algorithm

- Trivially, the algorithm always accepts monotone functions
- Assuming the claim, $O(n/\varepsilon)$ iterations suffice.

Definitions

- $U = \{(x^i(0), x^i(1)) \mid x \in \{0,1\}^n, i \in \{1..n\}\}$, all the pairs that differ on one coordinate.
- $|U| = n2^{n-1}$.

- $\Delta(f) = \{(x, y) \in U \mid f(x) > f(y)\}$, all the pairs violating monotonicity.
- $\delta(f) = |\Delta(f)|/|U|$.

Upper Bound on δ

- In order to make *f* monotone, one output from each violating pair must be changed.
- Every string belongs to at most *n* pairs.
- The number of changes is

$$\varepsilon(f)2^n \geq |\Delta(f)|/n = \delta(f)|U|/n = \delta(f)2^{n-1}$$

- Thus, $\delta(f) \leq 2\varepsilon(f)$.

Definitions

Function $S_i(f)$:

- If $f(x_i(0)) \leq f(x_i(1))$, $S_i(f)(x) = f(x)$.
- Otherwise, $S_i(f)(x) = 1 - f(x)$.

- $D_i(f) = |\{x \mid S_i(f)(x) \neq f(x)\}|$.
- $\sum D_i(f) = 2|\Delta(f)|$.

Non Decreasing Monotonicity

- Lemma: $D_j(S_i(f)) \leq D_j(f)$.

- Let x be such that $S_i(f)(x) \neq S_j(S_i(f))(x)$.
- Define $h(a,b)=S_i(f)(x^{ij}(a,b)$.

Non Decreasing Monotonicity

- Possible values of $h(a,b)$:

a\b	0	1
0	1	0
1	1	0

a\b	0	1
0	1	0
1	1	1

a\b	0	1
0	0	0
1	1	0

a\b	0	1
0	1	0
1	1	0

- In all cases, there is a unique y with $f(y) \neq S_j(f)(y)$.

Lower Bound on δ

- By inductive application of the lemma,
  $D_i(S_{i-1}\ldots S_1(f)) \leq D_i(\mathrm{f})$.

- $g = S_n S_{n-1}\ldots S_2 S_1(f)$.
- g is monotone, so $\varepsilon(f) \leq \mathrm{dist}(f,g)$.

Lower Bound on δ

- $\delta(f) = |\Delta(f)|/|U|$.
- $\sum D_i(f)=2|\Delta(f)|$.
- $\varepsilon(f) \leq \text{dist}(f,g)$.
- $2^n\text{dist}(f,g) \leq \sum D_i(S_{i-1}\ldots S_1(f)) \leq \sum D_i(f)$.
- $\delta(f) = |\Delta(f)|/|U| = 2^{-n}\sum D_i(f)/n \geq \text{dist}(f,g)/\text{n} \geq \varepsilon(f)/n$.

Almost Tight Bounds on δ

- For $\varepsilon>0$, there are functions g and h such that:

  $\varepsilon(g), \varepsilon(h) = \varepsilon - o(\varepsilon)$

  $\delta(g) = 2\varepsilon/n$

  $\delta(h) = \varepsilon$

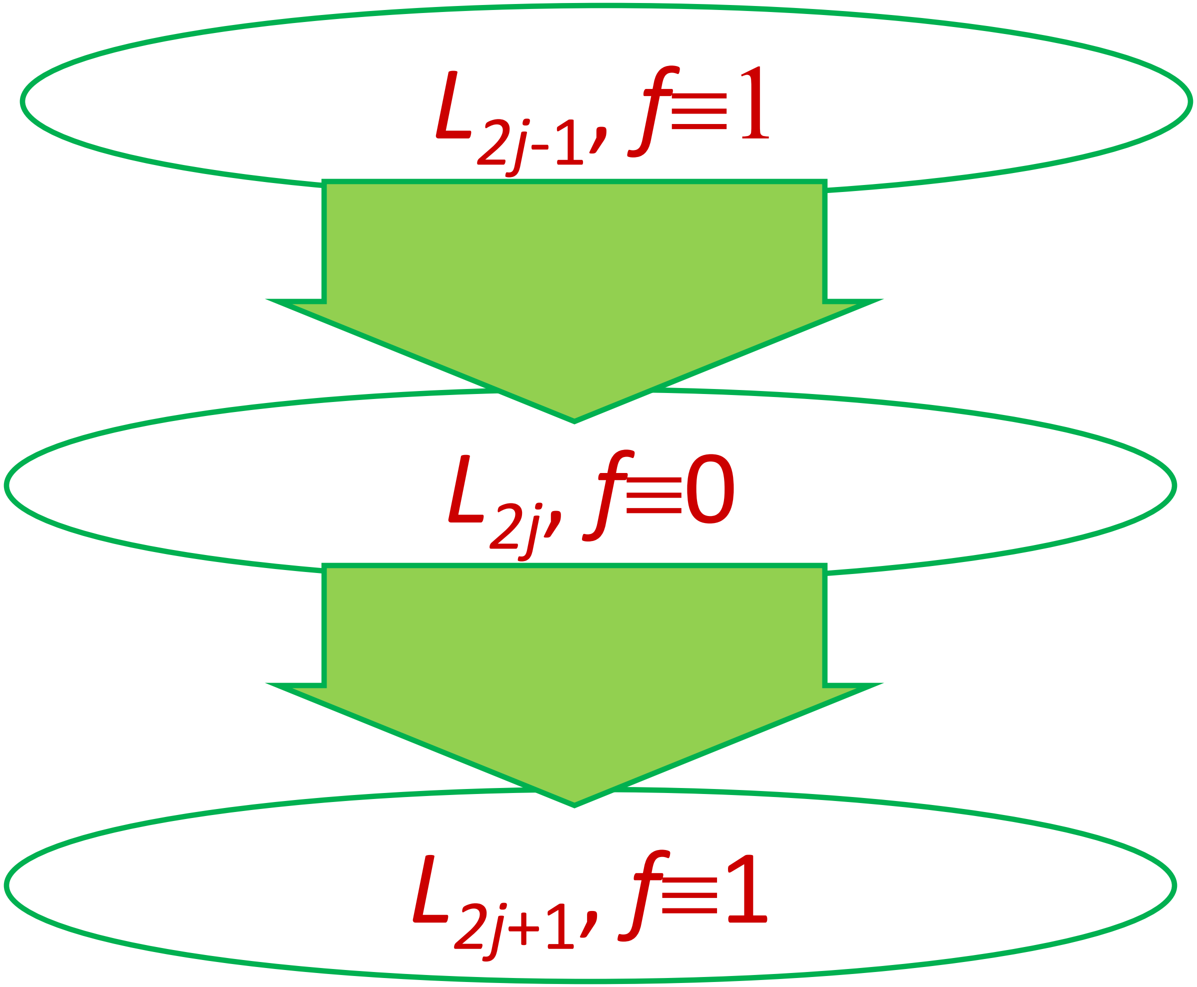
Almost Tight Bounds on δ

- Let g be an anti-dictatorship function (1 if x_1=0, 0 otherwise).
- $\delta(g) = 1/n$.
- $\varepsilon(g) = ½$, since there is a perfect matching between the set of values with x_1=0 and x_1=1, and at least one value in each pair must be modified.

Almost Tight Bounds on δ

- Consider the boolean hypercube as a directed graph, where the directed edges are from $(x_1x_2\ldots x_{i-1}0x_{i+1}\ldots x_n)$ to $(x_1x_2\ldots x_{i-1}1x_{i+1}\ldots x_n)$.
- Let L_i be the set of vertices with hamming weight i.
- There are only edges from L_i to L_{i+1}.
- Let h be the function receiving $i \bmod 2$ on L_i.
- $\delta(h) = ½$.

Almost Tight Bounds on δ



Almost Tight Bounds on δ

- Consider a pair of layers with all violating edges between them.
- Using Hall's Theorem, there is a matching containing all the vertices of the smaller layer.
- The number of unmatched vertices is at most

$$\sum \big|\,|L_{2i}|-|L_{2i-1}|\,\big| \leq 2|L_{[n/2]}| = O(2^n/\sqrt{n})$$

- $\varepsilon(h) = ½-O(1/\sqrt{n})$.

Almost Tight Bounds on δ

- These results can be extended to general values of ε, by considering only vertices with a certain suffix.

Extending the Domain

For $f \in \{1 \ldots d\}^n \rightarrow \{0,1\}$:

1. Uniformly at random select $i \in \{1,\ldots,n\}$ and $x \in \{1 \ldots d\}^n$.
2. According to some distribution p, select $a<b$.
3. If $f(x^i(a)) \leq f(x^i(b))$ accept, otherwise reject.

Extending the Domain

- There is an algorithm with query complexity $O(q_p(n,\varepsilon,d))$ that always accepts monotone functions and rejects function that are ε-far from monotone with constant probability.

Extending the Domain

- Using similar arguments, it is possible to show that

$$E_{i,y}[\delta(f \circ y^i))] \leq \delta(f)$$

$$\varepsilon(f)/2n \leq E_{i,y}[\varepsilon(f \circ y^i)]$$

- Hence, enough to lower bound $\delta(f \circ y^i)$ in terms of $\varepsilon(f \circ y^i)$.
- $f \circ y^i$ is a function from $\{1 \ldots d\}$ to $\{0,1\}$.

Distribution #1

- Uniform over all pairs $(a, a+1)$.
- If f is non monotone, There is at least (and possibly at most) one pair $(a, a+1)$ such that $f(a)>f(a+1)$.
- There are d-1 pairs and $\varepsilon(f)\leq½$.
- $2\varepsilon(f)/(d-1) \leq \delta(f)$.
- $O(dn/\varepsilon)$ repetitions suffice.

Distribution #2

- Uniform over all pairs (a, b) such that $a<b$.

f	0	0	1	1	0	1	0	0	1	0	1	1	1	1
f^*	0	0	0	0	0	0	1	1	1	1	1	1	1	1

- $2e$ difference between f and f^*.
- $\varepsilon(f) \leq 2e/d$.
- $\delta(f) \geq 2(e/d)^2 \geq \varepsilon(f)^2/2$.

Distribution #2

- $E_{i,y}[\delta(f \circ y^i)] \geq E_{i,y}[\varepsilon(f \circ y^i)^2] \geq (\varepsilon(f)/2n)^2$.
- $O(n^2/\varepsilon^2)$ repetitions suffice.

Distribution #3

- The distribution is uniform over P, where P is the set containing all pairs $\{a, b\}$ such that 2^k divides a, but 2^{k+1} does not divide a and b, and $|a-b| \le 2^k$.
- There are $O(d \log d)$ such pairs: each i is a member of at most $O(\log d)$ pairs, by considering the binary representation of i.
- Claim: there are $\Omega(d\varepsilon(f))$ violating pairs.

Distribution #3

- Consider P as directed edges on a graph, where the direction is towards the larger number.
- If $a>b$ there is a directed path of length at most 2 from b to a.
- Let i be the MSB where a and b differ. Then, $(a_1a_2\ldots a_{i-1}10\ldots0)=(b_1b_2\ldots b_{i-1}10\ldots0)$ is the middle vertex in the path.

Distribution #3

f	0	0	1	1	0	1	0	0	1	0	1	1	1	1
f^*	0	0	0	0	0	0	1	1	1	1	1	1	1	1

- $2e$ difference between f and f^*.
- $\varepsilon(f) \leq 2e/d$.
- There are least $e=\Omega(d\varepsilon(f))$ edge disjoint paths with a violating edge.
- $\delta(f) = \Omega(d\varepsilon(f)/d\log d) = \Omega(\varepsilon(f)/\log d)$.
- $O(n\log d/\varepsilon)$ repetitions suffice.

Extending the Range

For $f \in \{1 \ldots d\}^n \rightarrow \{0 \ldots c\}$:

1. Uniformly at random select $i \in \{1, \ldots, n\}$ and $x \in \{1 \ldots d\}^n$.
2. According to some distribution p, select $a < b$.
3. If $f(x^i(a)) \leq f(x^i(\mathrm{b}))$ accept, otherwise reject.

Extending the Range

- Define $f_i(x)$ to be 0 if $f(x)<i$, 1 otherwise.
- Then

$$\varepsilon(f) \leq \sum \varepsilon(f_i)$$

$$\delta(f) \geq \delta(f_i)$$

- Which implies an additional multiplicative factor of c to the query complexity.

Extending the Range

- It is possible to show $O(n \log d \log c/\varepsilon)$ queries suffice.
- A different algorithm can achieve query complexity of $O((n/\varepsilon)\log^2(n/\varepsilon))$.

Unateness

- A function $f \in \{0,1\}^n \rightarrow \{0,1\}$ is unate if there is $a \in \{0,1\}^n$ such that $f(x \oplus a)$ is monotone.
- A DNF formula where every variable is either always negated or never negated.
- Similar tester; $O(n^{1.5}/\varepsilon)$ pairs to find evidence for non unateness (using the generalized birthday paradox).

Improved Testing Algorithms for Monotonicity

Yevgeniy Dodis, Oded Goldreich, Eric Lehman, Sofya Raskhodnikova, Dana Ron, Alex Samorodnitsky

Definitions

- $S[f,a,b]$ – changes the range of f to be between a and b by changing all values that are more than b and less than a to be b and a respectively.
- $M[f]$ – arbitrary monotone function closest to f.

Definitions

- $C[f,a,b]$ – if $S[f,a,b]$ is different than $M[S[f,a,b]]$, gives the value of $M[S[f,a,b]]$, otherwise the value of f.
- $\mathrm{dist}(f, C[f,a,b]) = \varepsilon(S[f,a,b])$.

Properties of $C[f,a,b]$

- Does not add violating pairs.
- Has no violating pairs with values crossing the interval $[a,b]$.
- If (y,x) is a violating pair with $C[f,a,b](x)<C[f,a,b](y)$ then $f(x)\leq C[f,a,b](x)$, $C[f,a,b](y)\leq f(y)$.
- Proof by case analysis.

Analysis of the Algorithm

- $g_1 = S[f, c/2-1, c/2]$ $f_1 = C[f, c/2-1, c/2]$
- $g_2 = S[f_1, 0, c/2-1]$ $f_2 = C[f_1, 0, c/2-1]$
- $g_3 = S[f_2, c/2, c]$ $f_3 = C[f_2, c/2, c]$

- $\delta(f) \geq \delta(g_1)$, since S does not add violating pairs.
- $\delta(f) \geq \delta(g_2) + \delta(g_3)$, since the set of violating pairs of g_2 and g_3 is disjoint.

Analysis of the Algorithm

- $g_1 = S[f, c/2-1, c/2]$ $f_1 = C[f, c/2-1, c/2]$
- $g_2 = S[f_1, 0, c/2-1]$ $f_2 = C[f_1, 0, c/2-1]$
- $g_3 = S[f_2, c/2, c]$ $f_3 = C[f_2, c/2, c]$

- f_3 is monotone, since it has no violating pairs in the intervals (or crossing them) $[c/2-1, c/2]$, $[0, c/2-1]$, $[c/2, c]$.

Analysis of the Algorithm

- $g_1 = S[f, c/2-1, c/2]$ $f_1 = C[f, c/2-1, c/2]$
- $g_2 = S[f_1, 0, c/2-1]$ $f_2 = C[f_1, 0, c/2-1]$
- $g_3 = S[f_2, c/2, c]$ $f_3 = C[f_2, c/2, c]$

- $\varepsilon(f) \le \mathrm{dist}(f,f_3) \le \mathrm{dist}(f,f_1)+\mathrm{dist}(f_1,f_2)+\mathrm{dist}(f_2,f_3)$ $\le \varepsilon(g_1)+\varepsilon(g_2)+\varepsilon(g_3)$.

Analysis of the Algorithm

- Assume $c=2^s$.
- Then, there is K such that $\varepsilon(f) \leq Ks\delta(f)$ (for $s=1$ already proved with $K=O(n\log d)$):

$\varepsilon(f) \leq \varepsilon(g_1)+\varepsilon(g_2)+\varepsilon(g_3) \leq$

$K(\delta(g_1)+(s-1)\delta(g_2)+(s-1)\delta(g_3)) \leq$

$K(\delta(f)+(s-1)\delta(f)) = Ks\delta(f)$