Testing of Boolean functions

Pascal Huber

May 5, 2014

1 Motivation

2 Reminder

Boolean functions We consider functions from the hypercube $\Omega_n := \{-1,1\}^n$ into the real numbers. A function $f : \Omega_n \to \mathbb{R}$ is called *Boolean function* if it takes values only in $\{-1,1\}$.

We recall two important types of Boolean functions:

1. Let $[n] := \{1,2,\dots,n\}$ and let $S \subseteq [n]$. Then the *parity function over S* is defined by
$$\chi_S(x) := x_S := \prod_{i\in S} x_i.$$
Note that the product over the empty set is equal to one: $\prod_{i\in\emptyset} x_i = 1$.

2. The parity functions over the sets $S = \{i\}$ are called *dictator functions.* They depend solely on the i-th bit, i.e.
$$\chi_i := \chi_{\{i\}} = x_i.$$

(Note over the Great notational switch)?

Correlation of Boolean functions The measurable space $(\Omega_n, \mathcal{P}(\Omega_n))$ is equipped with the uniform probability measure

$$\mathbf{P}_{x\in\{-1,1\}^n} = \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1\right)^{\otimes n}.$$

Furthermore we define the corresponding L^2-inner product (called *correlation*) for two functions $f, g : \Omega_n \to \mathbb{R}$ on this space by

$$\langle f, g\rangle := \mathbf{E}_{x\in\{-1,1\}^n}\left[f(x)g(x)\right].$$

Note that if f and g are Boolean functions then one always gets $\langle f, g\rangle \in [-1,1]$ and $\|f\| := \sqrt{\langle f, f\rangle} = 1$.

Results from Fourier analysis of Boolean functions Before introducing the notions of testability of Boolean functions we recall the most important results of the Fourier analysis of Boolean functions.

- The parity functions $(\chi_S)_{S\subseteq[n]}$ form an orthonormal basis on the space of all functions from Ω_n to the real numbers, i.e.

$$\langle \chi_S, \chi_T \rangle = \begin{cases} 1 & S = T \\ 0 & \text{else} \end{cases}$$

- Consequently every Boolean function f has a unique representation of the form

$$f = \sum_{S\subseteq[n]} \hat{f}(S)\chi_S. \tag{1}$$

  The coefficients $\hat{f}(S)$ are called *Fourier coefficients* and (1) is called the *Fourier expansion* of f.

- *Plancharel's theorem:* Let $f, g : \Omega_n \to \mathbb{R}$. Then $\langle f, g\rangle = \sum_{S\subseteq[n]} \hat{f}(S)\hat{g}(S)$.

- *Parseval's theorem:* Let $f : \Omega_n \to \mathbb{R}$. Then $\langle f, f\rangle = \sum_{S\subseteq[n]} \hat{f}(S)^2$. Especially, if f is Boolean one has the identity $\sum_{S\subseteq[n]} \hat{f}(S)^2 = 1$.

3 Linearity of Boolean functions and basics of property testing

3.1 Linearity of Boolean functions

As a motivation we will first define what it means for a Boolean function to be linear. We introduce two possible definitions:

Definition 1 (Linearity of Boolean functions)**.** A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is called *linear* iff one of the following statements hold:

(i) For all $x, y \in \{0,1\}^n$ one has $f(x+y) = f(x) + f(y)$.

(ii) f is a parity function, i.e. there exists $S \subseteq [n]$ such that for all $x \in \{0,1\}^n$ one has $f(x) = \chi_S(x)$.

In multiplicative notation the definition reads as follows:

(i) For all $x, y \in \{-1,1\}^n$ one has $f(x \circ y) = f(x) \cdot f(y)$, where $x \circ y := (x_1y_1, \dots, x_ny_n)$.

(ii) There exists $S \subseteq [n]$ such that for all $x \in \{-1,1\}^n$ one has $f(x) = \prod_{i\in S} x_i$.

In order to make sure that Definition **??** is well defined we have to show that (i) and (ii) are equivalent:
$\underline{\text{(ii)} \Rightarrow \text{(i)}}$: is straightforward since

$$\chi_S(x+y) = \sum_{i\in S}(x_i + y_i) = \sum_{i\in S} x_i + \sum_{i\in S} y_i = \chi_S(x) + \chi_S(y).$$

(i) $\Rightarrow$ (ii): uses the representation $x = x_1e_1 + \cdots + x_ne_n$ where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 at the i-th bit. Then by iterating (i) one gets

$$f(x) = f\left(\sum_{i=1}^{n} x_ie_i\right) = \sum_{i=1}^{n} x_if(e_i) = \sum_{i \in S} x_i,$$

where $S \coloneqq \{i \in [n] : f(e_i) = 1\}$.

Thus we see that the linear Boolean functions are exactly the parity functions. (Note that we showed this using the additive notation, but this is of course also true when we use the multiplicative representation.)

3.2 Approximate linearity and basic notions of property testing

In the following we illustrate the basic notions of property testing using the example of linearity testing.

Approximate linearity The requirement to be linear on $\{0,1\}^n$ is rather expensive to check (one has to do all 2^n queries) and hence often too restrictive. That's why one considers a more relaxed concept of linearity:

A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is called *approximate linear* iff

(i') for "most" pairs $x, y \in \{0,1\}^n$ one has $f(x+y) = f(x) + f(y)$.

(ii') there exist $S \subseteq [n]$ s.t. for "most" $x \in \{0,1\}^n$ one has $f(x) = \chi_S(x)$.

(We will define more precisely what is meant by "most" in a second.)

As before we would like both definitions to be equivalent and one easily sees by copying the above argument that (ii') indeed implies (i'). The other implication on the other hand is much less obvious.
A main result of this course will be that in fact (i') implies (ii'). In order to show this we will use the results of the Fourier analysis extensively. But before dwelling on the proof we want to put the problem in the more general context of property testing. Therefore we will need some definitions.

Definition 2 (Property of Boolean functions)**.** A *property of Boolean functions* is a subset $\mathcal{P}$ of the set of all Boolean functions. We say that a Boolean function f *has property* $\mathcal{P}$ if $f \in \mathcal{P}$.

Definition 3 (ε-far and ε-close)**.** (i) Two Boolean functions f, g are ε-close if they agree on a $(1-\varepsilon)$-fraction of $\{0,1\}^n$, i.e.

$$\mathbf{P}_{x \in \{0,1\}^n}\left[f(x) = g(x)\right] \geq (1-\varepsilon).$$

Otherwise they are ε-far.

(ii) A Boolean function f is ε-close to having property $\mathcal{P}$ if there exists some function $g \in \mathcal{P}$ such that f and g are ε-close.

In our case the property we are interested in is to being linear, i.e. $\mathcal{P}_{lin} := \{\chi_S : S \subseteq [n]\}$.
Thus (ii') can be reformulated as being ε-close to being linear.

How can (i') be understood in this context? It is favorable to interpret it as a property test:
Multiplicative notation of BLR test !!

Definition 4 (BLR linearity test (Blum, Luby, Rubinfeld))**.** Given blackbox access to a Boolean function f do the following steps:

1. Pick x and y independently and uniformly at random from $\{0,1\}^n$.
2. Query f on x, y and $x+y$.
3. "Accept" iff $f(x)+f(y)=f(x+y)$.

Using this definition (i') says that the probability of BLR accepting the function f is large, more precisely we want to have

$$\mathbf{P}_{x,y\in\{0,1\}^n}\left[\mathsf{BLR}(f) \text{ accepts }\right] \leq (1-\varepsilon).$$

Hence if we can prove that (i') implies (ii') then we have shown that for the linearity property there exists a querying algorithm making only 3 queries such that whenever f is ε-far from being linear then the algorithm accepts f with probability of at most $1-\varepsilon$.
The existence of such a querying algorithm can also be shown for other properties and motivates the following definition:

Definition 5 (Locally testable property)**.** A property $\mathcal{P}$ of Boolean functions is called *locally testable* if there exists a randomized querying algorithm $\mathcal{T}$ making at most $\mathcal{O}(1)$ queries such that:

(i) If $f \in \mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f) \text{ accepts}] = 1$.

(ii) If f is ε-far from having property $\mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f) \text{ accepts}] \leq 1-\Omega(\varepsilon)$.

A more general definition of testability which relates the closeness to the property with the number of queries is due to Rubinfeld and Sudan:

Definition 6 (Testable property)**.** A property $\mathcal{P}$ of Boolean functions is *testable with $q(\varepsilon)$ queries* if there exists a randomized algorithm $\mathcal{T}$ (which gets ε as input) such that for all $\varepsilon > 0$ it makes $q(\varepsilon)$ queries and satisfies:

(i) If $f \in \mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f) \text{ accepts}] \geq \frac{2}{3}$.

(ii) If f is ε-far from $\mathcal{P}$ then $\mathbf{P}[\mathcal{T}(f) \text{ accepts}] \leq \frac{1}{3}$.

Note about the 1/3 thing...

The aim of today's course will be to show that the property $\mathcal{P}_{lin}$ of being linear is

- is locally testable (this is the implication (i') ⇒ (ii'))
- is testable with $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ queries.

4 Linearity testing of Boolean functions

4.1 Testability of linearity

We will prove the (local) testability of linearity in three steps:

1. Express the "acceptance probability" of the BLR-test for an arbitrary Boolean function f in terms of its Fourier coefficients.
2. Prove local testability of linearity using this representation.
3. Prove testability by executing the BLR-test multiple times.

Since we want to use the Fourier expansion of Boolean functions we will now switch to the multiplicative notation.

Lemma 1 (Acceptance probability of the BLR-test)**.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a Boolean function. Then the following holds:*

$$\mathbf{P}_{x,y\in\{-1,1\}^n}[\mathit{BLR}(f) \text{ accepts}] = \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[n]} \hat{f}(S)^3. \tag{2}$$

Proof. The proof consists primarily in writing the probability term in a suitable way and using the Fourier expansion of f.
We write the indicator function $\mathbb{1}_{\{\mathsf{BLR}(f) \text{ accepts}\}}$ in the following way for $x, y \in \{-1,1\}^n$

$$\mathbb{1}_{\{\mathsf{BLR}(f) \text{ accepts}\}}(x,y) = \frac{1}{2} + \frac{1}{2}f(x)f(y)f(x\circ y).$$

Hence we can express the right-hand side of (2) by

$$\mathbf{P}_{x,y\in\{-1,1\}^n}[\mathsf{BLR}(f) \text{ accepts}] = \frac{1}{2} + \frac{1}{2}\mathbf{E}_{x,y}[f(x)f(y)f(x\circ y)].$$

Using the Fourier expansion of f we find by linearity:

$$\begin{aligned}
&\mathbf{P}_{x,y\in\{-1,1\}^n}[\mathsf{BLR}(f) \text{ accepts}] \\
&\quad= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U\subseteq[n]} \Big(\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x\circ y)]\Big) && (3)\\
&\quad= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U\subseteq[n]} \Big(\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x)\chi_U(y)]\Big) && (4)\\
&\quad= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U\subseteq[n]} \Big(\hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}[\chi_{S\triangle U}(x)\cdot\chi_{T\triangle U}(y)]\Big), && (5)
\end{aligned}$$

where we used the linearity of parity functions and the relation $\chi_S(x)\cdot\chi_T(x) = \chi_{S\triangle T}(x)$.
Since in the product measure $\mathbf{P}_{x,y}$ functions of x, y are independent we finally

get

$$\begin{aligned}
&\mathbf{P}_{x,y\in\{-1,1\}^n}[\mathsf{BLR}(f) \text{ accepts}] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U\subseteq[n]} \Big(\hat{f}(S)\hat{f}(T)\hat{f}(U) \cdot \mathbf{E}_x[\chi_{S\triangle U}(x)]\mathbf{E}_y[\chi_{T\triangle U}(y)]\Big) && (6)\\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U\subseteq[n]} \Big(\hat{f}(S)\hat{f}(T)\hat{f}(U) \cdot \langle\chi_S,\chi_U\rangle \cdot \langle\chi_T,\chi_U\rangle\Big) && (7)\\
&= \frac{1}{2} + \frac{1}{2} \sum_{S\subseteq[n]} \hat{f}(S)^3, && (8)
\end{aligned}$$

since the parity functions are a orthonormal basis of all Boolean functions which means in particular

$$\langle\chi_S,\chi_U\rangle \cdot \langle\chi_T,\chi_U\rangle = \begin{cases} 1 & \text{if } S = U = T,\\ 0 & \text{else.}\end{cases}$$

□

Theorem 1 (Local testability of linearity)**.** *The property of being linear is locally testable.*

Here the proof is still missing! State that $\mathbf{P}[\mathit{BLR}(f) \textit{ accepts}] < 1-\varepsilon$

Proof. If $f = \chi_S$ for some $S \subseteq [n]$ then by Definition 1 and the subsequent discussion $f(x \circ y) = f(x) \cdot f(y)$ for all $x, y \in \{-1,1\}^n$.
Let now f be ε-far from being linear and assume for the sake of contradiction

$$\mathbf{P}_{x,y}[\mathsf{BLR}(f) \text{ accepts}] \geq (1-\varepsilon).$$

Then using Lemma 1 we have $1-\varepsilon \leq \frac{1}{2} + \frac{1}{2}\sum_{S\subseteq[n]}\hat{f}(S)^3$ and consequently

$$\begin{aligned}
1-2\varepsilon &\leq \sum_{S\subseteq[n]} \hat{f}(S)^3 && (9)\\
&\leq \left(\max_{S\subseteq[n]} \hat{f}(S)\right) \cdot \sum_{S\subseteq[n]} \hat{f}(S)^2 && (10)\\
&= \left(\max_{S\subseteq[n]} \hat{f}(S)\right) \cdot 1, && (11)
\end{aligned}$$

where we used Parseval's theorem in (11).
Hence there exists a subset $T \subseteq [n]$ such that

$$1-2\varepsilon \leq \mathbf{E}_{x\in\{_1,1\}^n}\left[f(x)\chi_T(x)\right].$$

But this implies

$$1-\varepsilon \leq \mathbf{P}_{x\in\{_1,1\}^n}[f(x) = \chi_T(x)],$$

which is a contradiction of f being ε-far from being linear. □

Theorem 2 (Testability of linearity)**.** *The property of being linear is testable with* $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ *queries.*

Proof. We define the randomized query algorithm taking $\mathcal{T}(f, \varepsilon)$ taking f and ε as arguments as follows:

- Run $\mathsf{BLR}(f)$ $\frac{2}{\varepsilon}$ times, independently.
- Accept iff every $\mathsf{BLR}(f)$ accepts.

Note that if f is linear then $\mathcal{T}(f)$ accepts with probability one, since $\mathsf{BLR}(f)$ accepts.
Assume now that f is ε-far from being linear. Since the BLR queries are independent from each other, we get from Theorem 1

$$\mathbf{P}\left[\mathcal{T}(f) \text{ accepts}\right] < (1-\varepsilon)^{\frac{2}{\varepsilon}} \tag{12}$$

$$= \left(1 - \frac{1}{(1/\varepsilon)^{\frac{1}{\varepsilon}}}\right)^2. \tag{13}$$

Since the last term converges for $\varepsilon \to 0$ to e^{-2} for ε small enough (e.g. $\varepsilon \leq \frac{1}{2}$) one gets

$$\mathbf{P}\left[\mathcal{T}(f) \text{ accepts}\right] < \frac{1}{3}.$$

□

4.2 Local decodability of linearity

Theorem 2 gives us a recipe to test if a given Boolean function is ε-close to a parity function. But how can we guess the right parity function using a minimum number of queries? The following result tells us that there is an algorithm making only two queries and predicts with high probability the right parity function.

Theorem 3 (Local docadiblity of linearity)**.** *The property $\mathcal{P}_{lin}$ of being a parity function is* locally docodable *with 2 queries, i.e. there exists a randomized 2-query algorithm $\mathcal{T}$ which has access to a function $f : \{-1,1\}^n \to \{-1,1\}$ and takes strings $x \in \{-1,1\}^n$ as arguments such that if f is ε-close to a parity function χ_S, then for every (!) $x \in \{-1,1\}^n$ one has*

$$\mathbf{P}\left[\mathcal{T}(x) = x_S\right] \geq 1 - 2\varepsilon. \tag{14}$$

Remark 1. Note that the Theorem 3 does not say for *almost every* $x \in \{-1,1\}^n$ $\mathcal{T}(x)$ computes the right value (this would be $\mathbf{P}_{x\in\{-1,1\}^n}[\mathcal{T} = \chi_S] \geq 1 - 2\varepsilon$). Instead, the theorem states that for *every (fixed)* $x \in \{-1,1\}^n$ $\mathcal{T}(x)$ computes the right value with high probability (the probability comes from the internal randomization of the algorithm $\mathcal{T}$).

Proof. Let $x \in \{-1,1\}^n$ be fixed. We define $\mathcal{T}$ as follows:

- Pick $y \in \{-1,1\}^n$ uniformly at random.
- Return $f(y)f(x \circ y)$.

We make the following observation: First, since y is uniformly distributed and f is ε-close to χ_S one has

$$\mathbf{P}_{y\in\{-1,1\}^n}[f(y)=\chi_S(y)]\geq 1-\varepsilon. \tag{15}$$

Similary $x\circ y$ is again uniformly distributed (but not independent of y) and

$$\mathbf{P}_{y\in\{-1,1\}^n}[f(x\circ y)=\chi_S(x\circ y)]\geq 1-\varepsilon. \tag{16}$$

Using the union bound and (15) and (16) we find

$$\begin{aligned}
\mathbf{P}_y[f(x\circ y)=\chi_S(x\circ y), f(y)=\chi_S(y)] & \qquad (17)\\
=\mathbf{P}_y[f(x\circ y)=\chi_S(x\circ y)]+\mathbf{P}_y[f(y)=\chi_S(y)] & \qquad (18)\\
-\mathbf{P}_y[\{f(x\circ y)=\chi_S(x\circ y)\}\cup\{f(y)=\chi_S(y)\}] & \qquad (19)\\
\geq(1-\varepsilon)+(1-\varepsilon)-1 & \qquad (20)\\
=1-2\varepsilon. & \qquad (21)
\end{aligned}$$

But in this case that booth equalities hold we have

$$f(y)f(x\circ y)=y_Sx_Sy_S=x_S,$$

which finishes the proof. □

5 Dictator testing of Boolean functions