Analysis of Boolean Functions		Februay 14, 2013
	Lecture 2	
Lecturer: Irit Dinur		*Scribe: Henry Yuen*

1 Recap

Last week, we defined the Fourier basis for Boolean functions on n variables, or equivalently $\mathbb{R}^{2^n}$, and presented it using formalism based on tensor products. We concluded that the Fourier the functions χ_S for every $S \subseteq [n]$, where

$$\chi_S(x) := \prod_{i \in S} x_i.$$

So every function $f : \{\pm 1\}^n \to \{\pm 1\}$ can be expressed as a linear combination of these *character* functions, i.e. there are unique real numbers $\widehat{f}(S)$ for every $S \subseteq [n]$ such that

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \chi_S(x).$$

We note, additionally, that χ_S is a monomial function, and the Fourier representation of f is a representation of f as a polynomial over $\mathbb{R}$.

We then discussed linearity testing, and proved that the natural test – namely, pick two points $x, y \in \{\pm 1\}^n$ uniformly and independently at random, and check that $f(x)f(y) = f(x \cdot y)$ – is complete and sound. Additionally, the analysis of this test shows that the local characterization of linearity testing (for all triples $(x, y, x \cdot y)$, $f(x)f(y) = f(x \cdot y)$) is *robust*.

We now move on to discuss the notion of *influence* of variables on Boolean functions.

2 Influence

The influence of a variable on Boolean functions was first studied by a paper of Michael Ben-Or and Nati Linial, who sought to study properties of collective coin flipping schemes [1]. Imagine the following situation: there is a collective of n persons, each who have a coin to flip. The goal is to apply a function to the outcome of each of these coin flips, and hope that the resulting function output is a fair coin. The question is, in what situations is this possible to achieve, or approximate?

It is easy to see that, in the synchronous coin flipping case, where each coin flip is independent of each other, as long as there exists a person whose coin is fair (or approximately fair), the parity function applied to the coin flip outcomes will produce an nearly uniformly random bit. However, suppose we're in the asynchronous setting, where some players may choose to wait to see the outcomes of the coin flips of others, before flipping their own. The goal of producing a fair bit now seems much more difficult. For example, suppose the function applied were PARITY and player n waited to see the outcome of all $n-1$ other players. The n player could force the output to be deterministically 0.

Ben-Or and Linial asked whether it was possible to construction a function such that f was not sensitive to such "gaming" by individual players. This led them to define the notion of the *influence* of a variable, which intuitively is the following quantity: given a function f and a variable x_i, how often can flipping its value change the value of f, when uniformly averaged over all possible assignments to x?

Definition 1 (Influence of a variable) *Let $f : \{\pm 1\}^n \to \{\pm 1\}$. Then, the influence of the variable x_i on f is*

$$\mathrm{Inf}_i(f) = \Pr_x[f(x) \neq f(x^{\oplus i})],$$

where $x^{\oplus i}$ is x except the ith bit is flipped.

Definition 2 (Total Influence of a function) *Let $f : \{\pm 1\}^n \to \{\pm 1\}$. The total influence of f is*

$$\mathrm{Inf}(f) = \sum_i \mathrm{Inf}_i(f).$$

Which functions have the smallest maximum influence (over all variables)? Well, clearly, the constant function has the smallest influence: 0. We restrict our attention to balanced Boolean functions, that is, functions f such that $\mathbb{E}_x[f(x)] = 0$. Let's look at the influence of variables on some familiar Boolean functions.

2.1 PARITY

Consider the parity function, or $f := \chi_{[n]}$. Clearly, for all i, $\mathrm{Inf}_i(f) = 1$, because flipping any single bit of an input will flip the output bit. Thus, $\mathrm{Inf}(f) = n$. Thus PARITY has the maximum total influence achievable by a Boolean function.

2.2 MAJORITY

For simplicity assume that n is odd. Consider the majority function

$$f = \mathrm{sgn}\left(\sum_i x_i\right).$$

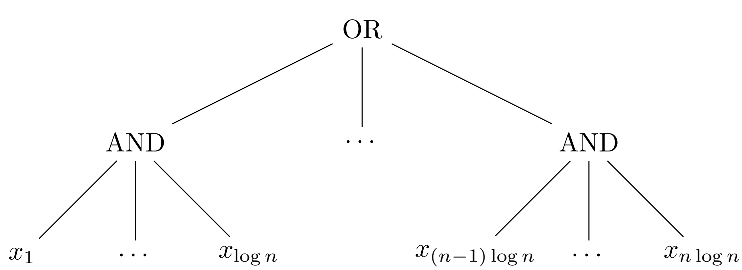
In what cases does flipping a single bit of an input alter the output of f? This happens when the input is already balanced; that is, $(n-1)/2 + 1$ bits are $+1$ and the others are -1, or vice versa. Thus for any i,

$$\mathrm{Inf}_i(f) = \frac{\binom{n}{(n-1)/2+1}}{2^n} \approx \frac{\binom{n}{n/2}}{2^n} \approx \frac{\Theta(2^n/\sqrt{n})}{2^n} = \Theta(1/\sqrt{n}).$$

Hence, the total influence is $\Theta(\sqrt{n})$.

2.3 Tribes

Well, the tribes function may not be as familiar to you. It is the OR of n ANDs of $\log n$ disjoint variables. Pictorially, we have



Thus, our function $f : \{\pm 1\}^m \to \{\pm 1\}$ can be expressed as a Boolean function on $m = n \log n$ variables. The reason this function is called the "tribes" function is because each AND cluster can be thought of as a tribe of voting members, and each AND cluster passes on whether its members all unanimously voted yes to the OR mediator.

This function was conjectured by Ben-Or and Linial to be a balanced Boolean function with minimal maximum influence (that is, an f such that $\max_i \mathrm{Inf}_i(f)$ is minimal). This conjecture was eventually proved by Kahn, Kalai and Linial [2], and the proof ultimately used Fourier analysis.

Let's compute the influence of the tribes function. Fix an i. The situation where flipping x_i will change the function output from $+1$ to -1 is where all the AND tribes members are not unaminously -1, but the tribe that x_i is in is nearly there: only x_i is $+1$. Similarly, the situation where flipping x_i will change $f(x)$ from -1 to $+1$ is the same, except x_i's tribe IS unanimously -1. We can calculate the probability of this happening. The probability that a particular tribe will be unanimous is $2^{-\log n} = 1/n$. Thus,

$$\mathrm{Inf}_i(f) \approx \left(1 - \frac{1}{n}\right)^{n-1} \frac{1}{n} = \Theta(1/n) = \Theta\left(\frac{\log m}{m}\right).$$

Hence the total influence is $\Theta(\log m)$.

3 A Graph-Theoretic Interpetation of Influence

Next we take a look at influence from a graph theoretic perspective. Recall the Boolean hypercube, which is the graph $H = (V, E)$, where $V = \{\pm 1\}^n$, and $E = \cup_i E_i$, with $E_i = \cup_x \{(x, x^{\oplus i})\}$. When defining the Boolean hypercube in this way we notice that E_i is a perfect matching of V, and the edge set E decomposes into disjoint matchings $\{E_i\}$. We call the E_i the edges in the *ith direction*.

We can view a boolean function f as defining a cut in the Boolean hypercube, namely, dividing the vertices into set $A_+(f) = \{x : f(x) = +1\}$, and $A_-(f) = \{x : f(x) = -1\}$. The influence of a variable has a very nice interpretation now: we can equivalently define the influence of the ith variable to be

$$\mathrm{Inf}_i(f) = \frac{E_i(A_+, A_-)}{2^{n-1}}$$

where $E_i(A_+, A_-)$ counts the number of edges in E_i that cross the cut (A_+, A_-). Then, the total influence of a function f can be written as

$$\mathrm{Inf}(f) = \sum_i \frac{E_i(A_+, A_-)}{2^{n-1}} = \frac{E(A_+, A_-)}{2^{n-1}}.$$

3.1 Functions with small total influence

We can ask a question related to that of Ben-Or and Linial's: What a balanced Boolean functions with the minimal *total* influence? It turns out that one can prove *dictator* functions have minimal total influence.

Definition 3 (Dictator function) *A Boolean function* $f : \{\pm 1\}^n \to \{\pm 1\}$ *is a* dictator *function if there exists an* i *such that* $f(x) = x_i$.

We show this by studying the eigenvectors of the normalized adjacency matrix of H. Let's abuse notation slightly and let H stand for this normalized adjacency matrix (i.e. all the entries are either 0 or $1/n$). We'll show that the eigenvectors of H correspond to the Fourier basis!

Let $S \subseteq [n]$. Treat χ_S as a vector in $\mathbb{R}^{2^n}$ in the natural way, indexed by $\{\pm 1\}^n$. Observe that χ_S is an eigenvector of H. Let $v = H\chi_S$. Then for all $x \in \{\pm 1\}^n$,

$$v(x) = \frac{1}{n} \sum_i \chi_S(x^{\oplus i}).$$

But observe that $\chi_S(x^{\oplus i}) = -\chi_S(x)$ if $i \in S$, and $\chi_S(x^{\oplus i}) = \chi_S(x)$ otherwise. So,

$$v(x) = \frac{1}{n} \left(\sum_{i \notin S} \chi_S(x) - \sum_{i \in S} \chi_S(x) \right) = \left(1 - \frac{2|S|}{n}\right) \chi_S(x).$$

Hence, χ_S is an eigenvector of H with eigenvalue $\lambda_S = 1 - \frac{2|S|}{n}$. Now, since the dimension of H is 2^n, and we've identified 2^n orthogonal eigenvectors of H, we've identified an orthogonal eigenbasis of H.

Claim 4 *For every* $f : \{\pm 1\}^n \to \{\pm 1\}$,

$$\mathrm{Inf}(f) = \sum_{S \subseteq [n]} \widehat{f}_S^2 \cdot |S|.$$

Proof First, consider the correlation quantity $\rho := \mathbb{E}_{(x,y)\in E}[f(x)f(y)]$. Intuitively, this measures the correlation between the value of f on adjacent vertices. This is easily seen to be related to the total influence: we're assigning a value $+1$ to adjacent vertices that agree and -1 to adjacent vertices that do not agree (with respect to f). Thus,

$$\mathrm{Inf}(f) = n \cdot \frac{1-\rho}{2}.$$

On the other hand, we can express ρ in an interesting way. We can write $\rho = \langle f, Hf \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the inner product function in $\mathbb{R}^{2^n}$. This is because

$$\begin{aligned}
\langle f, Hf \rangle &= \mathbb{E}_x[f(x)(Hf)(x)] \\
&= \mathbb{E}_x \left[f(x) \frac{1}{n} \left(\sum_i f(x^{\oplus i}) \right) \right] \\
&= \mathbb{E}_{(x,y)\in E}[f(x)f(y)] \\
&= \rho.
\end{aligned}$$

But now we invoke Fourier magic and rewrite this inner product:

$$\begin{aligned}
\rho &= \left\langle \sum \widehat{f}_S \chi_S, H \sum \widehat{f}_S \chi_S \right\rangle \\
&= \left\langle \sum \widehat{f}_S \chi_S, \sum \widehat{f}_S \lambda_S \chi_S \right\rangle \\
&= \sum_S \widehat{f}_S^2 \lambda_S \\
&= \sum_S \widehat{f}_S^2 (1 - \frac{2|S|}{n}).
\end{aligned}$$

Rearranging terms, we see that we get the final formula,

$$\mathrm{Inf}(f) = \sum_{S \subseteq [n]} \widehat{f}_S^2 \cdot |S|.$$

■

Armed with this claim, we can now show that dictator functions have minimal total influence out of all balanced Boolean functions. Observe that f balanced implies that $\widehat{f}_\emptyset = 0$. Next, we invoke Parseval's theorem, which states that $\sum \widehat{f}_S^2 = 1$. Since $\widehat{f}_\emptyset = 0$, we know that $Inf(f) \geq 1$.

On the other hand, it is easy to see that for a dictator function $f = \chi_{\{i\}}$, $Inf(f) = 1$, so we see that they achieve the minimum possible total influence.

4 Next time

Next time, we prove the famed KKL Theorem, resolving the Ben-Or and Linial conjecture, which states that Tribes function achieves the smallest maximum influence of a variable.

References

[1] Ben-Or, M. Linial, N. *Collective coin flipping*, Randomness and computation, Academic Press, new York, 1989, pp. 91-115.

[2] Kahn, J. Kalai, G. Linial, N. *The influence of variables on Boolean functions*, Foundations of Computer Science, 1988.