CS 880: Advanced Complexity Theory 1/25/2008

Lecture 2: Basics of Harmonic Analysis

Instructor: Dieter van Melkebeek Scribe: Priyananda Shenoy

In the last lecture, we mentioned one structural property of Boolean functions, namely the influence of variables and a few of its applications based on Harmonic Analysis. In this lecture, we will see some more structural properties which have a close connection to Harmonic Analysis. We will then discuss the basics of Harmonic Analysis.

1 Structural properties of Boolean functions (continued)

1.1 Influence of variables

Recall that we defined the influence of the ith variable on a function f as $I_i(f) = \Pr_x[f(x) \neq f(x^{(i)})]$ and the total influence of the function as $I(f) = \sum_i I_i(f)$. We saw some applications of this property in the last lecture. Here is another one.

1.1.1 Application in Social Choice Theory

One interpretation we saw of a Boolean function in a Social Choice context, was that the function determines which of the two candidates wins an election, where the variables represent the choices of the n voters. Some of the properties we might want such a function to have are:

1. *Balanced*: We want the function to be "fair" and not biased towards one candidate.
2. *Monotonicity*:

   Definition 1. *A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is said to be* monotone *if $f(a) \leqslant f(b)$ whenever $\forall i \in [n]$ $a_i \leqslant b_i$ where $a, b \in \{0,1\}^n$.*

   In no case should it happen that if a single vote is changed from 0 to 1, the function value changes from 1 to 0.
3. *Small Influences*: We would like to ensure that no one voter is able to change the outcome of the election drastically, i.e., the influence of each voter should be small.

Let us consider how the various schemes we have seen till now fulfill these properties:

1. Majority: As we have seen, the Majority function is balanced. It is also trivially monotone. But individual influences are quite high, namely $\Theta(\frac{1}{\sqrt{n}})$.
2. Dictator: The Dictator functions are balanced but the influence of the dictating variable equals 1.
3. Tribes: The Tribes function is essentially balanced, and monotone. Also, the influences are small, namely $\Theta(\frac{\log n}{n})$.

The trivial lower bound for the influence of a single variable of a balanced function is $\frac{1}{n}$, since the lower bound for the total influence is 1. It turns out that it is not possible to achieve this bound and Tribes, with an influence of $\Omega(\frac{\log n}{n})$, is optimal.

Theorem 1. *If f is a balanced Boolean function, then $\exists i$ such that $I_i(f) = \Omega(\frac{\log n}{n})$.*

If we are only interested in the properties listed above, then Tribes is the best election scheme. But it turns out that this is a bad scheme, for a different reason. This is because one of the candidates can guarantee a win by bribing $O(\log n)$ voters.The following figure illustrates this.

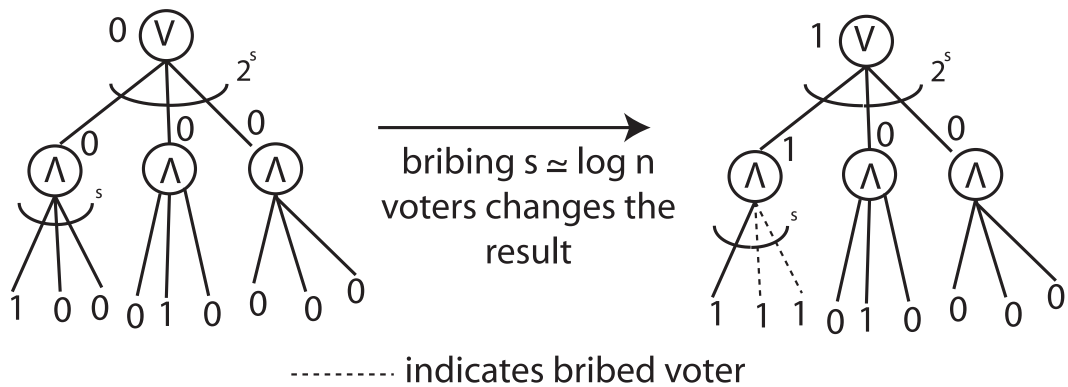

Fig 1: Tribes is susceptible to bribery.

In general, the Tribes function is succeptible to errors (noise), i.e., a miscount or error in a small number of bits can have a drastic effect on the result. The next structural property we will examine formalizes this susceptibility to noise.

1.2 Noise Sensitivity

Until now we were considering how a single bit flip affects the value of the function. Now we will consider the case where each bit of the input may have been flipped with some probability. This is a way to model noise or random errors in the system.

Definition 2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function and $\varepsilon > 0$. Then the* Noise Sensitivity *of f at level ε is is defined as*

$$NS_\varepsilon(f) = \Pr_{\substack{x \\ y \sim_\varepsilon x}} [f(x) \neq f(y)],$$

where $y \sim_\varepsilon x$ denotes that y is obtained by flipping each bit in x independently with probability ε.

The Noise Sensitivity and the Total Influence of a function have the following relationship.

Claim 1. *For any function f, $NS_\varepsilon(f) \leqslant \varepsilon I(f)$.*

For many functions, the gap can be quite large, i.e., the noise sensitivity is much smaller than $\varepsilon I(f)$.

1.2.1 Applications

1. Social Choice Theory: We saw that Tribes voting scheme was not desirable, in spite of being balanced, monotone and having small influences. If we also require that the voting scheme have low noise sensitivity (i.e., the scheme is "stable" in the face of miscounted votes), then it can be shown that the Majority scheme is the most suitable one. The precise statement of this result, known as *Majority is stablest*, will be defined later in the course.

2. Hardness Amplification within NP: Functions like Tribes can be used to boost the hardness of average case hard functions in NP.

3. Hardness of Approximation: The best known approximation algorithm for Max-Cut has an approximation ratio of 0.878. Under the Unique Games Conjecture, it can be proved that this is optimal. The proof of this inapproximability result hinges on the "Majority is Stablest" result.

1.3 Closeness to linearity

Linear functions are an important class of functions. Let us first define the notion of Linearity formally.

Definition 3. *A function* $f : \{0,1\}^n \to \{0,1\}$ *is said to be* linear *if* $\forall x, y \in \{0,1\}^n$

$$f(x+y) = f(x) + f(y).$$

Sometimes we would like to know how "close" a function is to being linear. We can define this notion of "closeness" formally.

Definition 4. *The* distance *between a function* f *and a set* $\mathcal{F}$ *of functions over the same domain as* f *is defined as*

$$D(f, \mathcal{F}) = \min_{g \in \mathcal{F}} [\textit{relative Hamming distance between } f \textit{ and } g].$$

Harmonic Analysis allows us to analyze $D(f, LIN)$ where LIN denotes the set of linear functions. This has a number of applications, of which a few are listed below.

1.3.1 Applications

1. PCP Theorem: Linearity testing is an essential part of the proof of the PCP Theorem.

2. Hardness of Approximation: The parameters of the linearity test determine the strength of the inapproximability result.

2 Basics of Harmonic Analysis

The main topic for this course is Harmonic Analysis of Boolean functions. Rather than delving into Boolean functions directly, it is instructive to see how Harmonic Analysis is developed over the reals and then generalize it to the Boolean domain. We start with the simplest case, analyzing periodic functions over the reals.

Let $f : \mathbb{R} \to \mathbb{R}$ be a periodic function with period 2π. Under some mild conditions, we can write f as a superposition of sinusoidal functions (called *harmonics*), each characterized by an integral frequency k:

$$f(x) = \sum_{k=0}^{\infty} a_k \cos(kx) + b_k \sin(kx).$$

Since $\cos(x) = \frac{e^{ix}+e^{-ix}}{2}$ and $\sin(x) = \frac{e^{ix}-e^{-ix}}{2}$ we can write f as

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{1}$$

where the kth *Fourier coefficient* c_k can be computed as $c_k = \langle f, e^{ikx} \rangle$ and the *inner product* of two functions is defined as

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{g(x)}dx$$

.

The expression for c_k follows by taking the inner product with e^{ikx} on both sides of (1), as the functions e^{ikx} are orthonormal w.r.t $\langle \cdot, \cdot \rangle$.

The framework can be generalized to include non-periodic functions as well. We describe the generalization without going into the details of the formalism.

Let $f : \mathbb{R} \to \mathbb{C}$ be a function. Then under some mild constraints f can be expressed in the following way.

$$f(x) = \int \hat{f}(k) e^{ikx} dk.$$

Unlike in the previous case, where the "frequency" k was always integral, here k varies over the real line.

The function $\hat{f} : \mathbb{R} \to \mathbb{C}$ is termed as the *Fourier Transform* of f and can be computed as follows:

$$\hat{f} = \langle f, e^{ikx} \rangle,$$

where

$$\langle f, g \rangle = \frac{1}{2\pi} \int f(x)\overline{g(x)}dx.$$

The Fourier Transform is the transform which maps $f : \mathbb{R} \to \mathbb{C}$ to $\hat{f} : \mathbb{R} \to \mathbb{C}$. We can interpret this as a mapping from the representation of the function in the standard basis to a representation in the basis formed by the harmonics e^{ikx}.

2.1 Properties of Fourier Transforms

1. Linearity: If $f, g : \mathbb{R} \to \mathbb{C}$ are two functions and α is a scalar then

   $$\widehat{f+g} = \hat{f} + \hat{g}$$

   and

   $$\widehat{\alpha f} = \alpha \hat{f}.$$

2. (Almost) Involution: The inverse transform, which maps $\hat{f}$ to f, is computed similar to the transform.

$$\hat{f}(x) = \frac{1}{2\pi} \int f(x)e^{-ikx}dx$$

   Applying the transform twice results in the original function up to a constant factor.

3. (Almost) Isometry: Angles and distances are preserved by the transform, up to constant factors

   - *Plancherel's Equality* $\langle \hat{f}, \hat{g} \rangle = 2\pi \langle f, g \rangle$.
   - *Parseval's Equality* $\langle \hat{f}, \hat{f} \rangle = 2\pi \langle f, f \rangle$.

   In fact, we can redefine the inner product to have the constant factor as $\frac{1}{\sqrt{2\pi}}$, in which case we have a perfect isometry.

4. Convolution: Convolution is an operation which occurs naturally in many scenarios.

   Definition 5. *The convolution of two functions $f, g : \mathbb{R} \to \mathbb{C}$ denoted as $f * g$ is defined as*

$$(f * g)(k) = \frac{1}{2\pi} \int f(y)g(k - y)dy.$$

   The following theorem, termed the *Convolution Theorem* is of great practical importance.

   Theorem 2. *The Fourier transform of a convolution is the point wise multiplication of the transforms of each of the functions, i.e.,*

$$\widehat{(f * g)}(k) = \hat{f}(k) \cdot \hat{g}(k).$$

   The convolution operation occurs in a wide range of applications, of which we mention a couple.

   - In *signal processing*, convolutions occur when a filter is being applied to a signal. A filter is an operation which changes the signal, for example a filter might smoothen the signal to remove small rapid changes. The filter operation is represented by a convolution of the input signal and a filter function.

     Since convolution is a complex operation, it makes sense to take the Fourier transform of the signal and the filter, do simple multiplication and then take the inverse transform. This is useful when multiple filters are to be applied on the signal, so that the transform can be done once, multiple filters are applied on the signal and then the inverse transform is applied at the end.
   - In *probability theory*, if f and g describe the distributions of two independent variables X and Y respectively, then the distribution representing $X + Y$ is given by $f * g$.

2.2 Generalizing Harmonic Analysis

There are several ways to generalize Fourier Analysis. In the continuous setting, Fourier Analysis arose as a tool to solve certain types of differential equations. In the discrete setting, harmonic analysis can be developed in a similar way in the context of solving corresponding difference equations.

Another way to generalize the concepts of Fourier analysis is to allow the domain of the function to be groups rather than reals. This is the way we will follow. The next table lists the concepts used in the generalization.

"classical" Harmonic Analysis	Generalized Harmonic Analysis
The function is defined over reals $f : \mathbb{R} \to \mathbb{C}$	The function is defined over a locally compact Abelian group $f : G \to \mathbb{C}$
The inner product $\langle , \rangle$ is based on the *Lebesgue* measure	The inner product $\langle , \rangle$ is based on the *Haar* measure
The harmonics or basis functions are of the form e^{ikx}	The basis functions are characters (defined below)

It is beyond the scope of this class, and also irrelevant for us, to go into the details of the full generalization. The relevant notion for us is that of a character.

Definition 6. *A* character $\chi : G \to \mathbb{C}$ *of a group* G *is a continuous homomorphism from* $(G, +)$ *to* $(\mathbb{C}, \cdot)$, *i.e. a continuous mapping such that* $\forall x, y \in G$

$$\chi(x + y) = \chi(x) \cdot \chi(y).$$

Since we are interested in finite groups, the continuity constraint disappears.

2.2.1 General Properties of characters

1. The range of χ is the unit circle on the complex plane. We will show this for finite groups, but the result holds true for the infinite case as well, where it follows from the continuity requirement. Let $(G, +)$ be a finite group, then $\forall x \in G$

   $$x^{|G|} = 0.$$

   Since χ is a homomorphism

   $$(\chi(x))^{|G|} = 1.$$

   Hence χ maps elements to unit-magnitude complex numbers, namely the points of the regular n-gon inscribed in the unit circle with 1 being one of its points.

2. (*Trivial character*) The neutral element for $(\mathbb{C}, \cdot)$, namely $\chi \equiv 1$, is a character, called the trival character.

3. (*Inverse*) If χ is a character, then $(\chi)^{-1} = \overline{\chi}$ is also a character.

4. (*Closure*) If χ and ψ are characters, then so is $\chi \cdot \psi$.

From the above properties it is clear that the set of characters form a group under point-wise multiplication. This group $\hat{G}$ is called the *dual group* of G.

3 Next Time

In the next lecture, we will prove that characters form an orthonormal set. In the case of abelian groups, we will show that the group and its dual are isomorphic. In particular, the characters form an orthonormal basis, which allows us to define the Fourier transform.