**CS 880: Advanced Complexity Theory** 3/26/2008

# Lecture 23: Threshold Phenomena

Instructor: Dieter van Melkebeek Scribe: Seeun Umboh

In the last lecture, we developed the necessary machinery to discuss threshold phenomena. We are interested in the behavior of monotone, non-trivial (that means non-constant in this setting) Boolean functions $f$, when inputs are drawn from a $p$-biased distribution. That is, we want to consider the probability

$$\mu_p(f) = \Pr_{x \sim p}[f(x) = -1]$$

where the subscript $x \sim p$ means each bit of the input $x$ is set to -1 with probability $p$ independently. Last lecture we argued that $\mu_p(f)$ is a nondecreasing function in $p$. We also defined the function's critical probability $p_c$ as the least $p$ such that $\mu_p(f)$ is at least $\frac{1}{2}$. In fact, in the case of finite domains we showed that $\mu_p(f)$ is increasing from 0 to 1 when $p$ goes from 0 to 1, so $p_c$ is the unique $p \in (0, 1)$ for which $\mu_p(f) = \frac{1}{2}$.

## 1 Background

We are mainly interested in the setting where $f$ represents a graph property. For a graph $G = (V, E)$ with $m = |V|$ vertices, a graph property is a Boolean function $f$ on $n = \binom{m}{2}$ variables where each variable represents whether a particular edge is in the graph or not. We want to study $\mu_p(f)$, the probability that a random graph, where each edge is in the graph with probability $p$ independently, satisfies the property $f$ as a function of $p$. Today we will argue that for every non-trivial monotone graph property, $\mu_p$ has a sharp threshold. By a sharp threshold, we mean that around the critical probability, $\mu_p$ has a steep slope. In particular, $\mu_p$ goes from almost 0 to almost 1 in a short interval.

The study of sharp thresholds originated in physics, where they are known as *phase transitions* instead. Physicists studying various systems noticed that for some properties, as a function of some parameter such as temperature, the probability of observing the property in the system jumped from almost 0 to almost 1 at a certain point. Examples include transitions from the solid to the liquid state, or from a non-ferromagnetic to a ferro-magnetic regime. The systems are often modeled using graphs and the graph properties are in turn captured using Boolean functions.

## 2 Infinite Graphs

In the case of infinite graphs, threshold phenomena are not all that surprising. In this setting the thresholds are often as sharp as they can be. That is, at the critical probability, $\mu_p(f)$ jumps suddenly from 0 to 1. This happens in the case of infinite-dimensional systems as a result of the Kolmogorov 0-1 law. Roughly, the law applies to probability spaces that are infinite-dimensional product spaces, where each component is set independently. In particular, it applies to properties of infinite graphs since each variable is set independently, and there are infinitely many of them. The law says that for every property that is invariant under changes to a finite number of components, the probability that the property holds is either 0 or 1. Thus, for every fixed $p$, $\mu_p$ is either 0 or 1. This means that at the critical probability, $\mu_p$ suddenly jumps from 0 to 1.

An an example, consider a random instance of the infinite 2-dimensional grid, where each edge is in the grid with probability $p$ independently. The property we are interested in is whether or not the graph has a connected component of infinite size. Clearly, this property does not change if we only flip a finite number of variables. So, the Kolmogorov 0-1 law applies here and by the above argument, $\mu_p$ goes directly from 0 to 1 at the critical probability. In fact, the critical probability for this graph property is 1/2.

This is related to percolation theory, in which we have some porous material that may have cavities in certain positions and there is oil somewhere in the material. The question of interest is if we drill a hole at a random vertex, whether or not the oil will flow away or remain contained within the material, over time. If the component that contains the hole is infinite, then the oil will flow away, otherwise it will be contained in that component. The relationship with the above is that below threshold, the component containing the vertex must be finite. Above the threshold, with positive probabiilty, the component containing the particular vertex is infinite. Otherwise, by symmetry, the probability of any component being infinite would be 0, and therefore the probability of the graph having an infinite component would be 0, as well.

# 3 Finite Graphs

In this lecture, we will concentrate on finite-dimensional systems and random graphs. Finite random graphs cannot exhibit such sharp threshold behavior as infinite ones, but the change can still be quite sudden. The latter is typically the case when the probabilistic method is used – a sharp threshold exists whenever the probability that a random object has the desired property is either extremely high or extremely low. We consider two concrete examples before embarking on a more general study.

## 3.1 Connectivity and Cliques

In this setting we have a fixed number of vertices, and each edge is in the graph with probability $p$ independently. We are interested in the probability that the graph is connected as a function of $p$. For very small $p$, we expect mostly isolated vertices, and for large $p$ we get connectivity with high probability. It turns out that

$$p_c \sim \frac{\ln m}{m}. \tag{1}$$

This is a standard application of the probabilistic method and we only give a sketch of the proof here. We show that if $p$ is slightly below the critical probability, that is $p \leq (1-\epsilon)p_c$, then $\mu_p \to 0$ as $m \to \infty$. If we are slightly above, that is $p \geq (1+\epsilon)p_c$, then $\mu_p \to 1$ as $m \to \infty$.

*Proof Sketch.* We first argue that if $p$ is small, then the probability of having an isolated vertex is non-zero. The probability that a given vertex is isolated is $(1-p)^{m-1}$ since it has $m-1$ neighbors. So we have

$$\mathrm{E}[\#\text{ of isolated vertices}] = m(1-p)^{m-1} \sim m \cdot e^{-pm}.$$

If $p \leq (1-\epsilon)p_c$, then $m \cdot e^{-pm} \geq m \cdot m^{\epsilon-1} \to \infty$ as $m \to \infty$. We can also show that the variance of the number of isolated vertices is small compared to its expectation. Chebyshev then tells us that with high probability the number of isolated vertices is close to its expectation. Since the expectation approaches $\infty$ as $m \to \infty$, the probability that there are no isolated goes to 0, which is what we wanted to show.

In the case where $p \geq (1+\epsilon)p_c$, we can count the number of $k$-subsets $S$ of vertices such that there are no edges between $S$ and $\bar{S}$. There are $\binom{m}{k}$ such subsets and each of the $k$ vertices have $m-k$ neighbors outside of $S$. Let $\binom{V}{k}$ denote the set of all $k$-subsets of the vertices, then

$$\begin{aligned} \mathrm{E}\left[\left|\left\{S \subseteq \binom{V}{k} : \text{no edges between } S \text{ and } \bar{S}\right\}\right|\right] &= \binom{m}{k}(1-p)^{k(m-k)} \\ &\leq \left(\frac{e\cdot m}{k}\right)^k (1-p)^{k(m-k)} \\ &= \left(\frac{e\cdot m}{k}(1-p)^{m-k}\right)^k, \end{aligned} \tag{$*$}$$

where the second inequality is a standard bound on binomial coefficients. Taking a geometric sum of $(*)$, we have that the expected number of subsets $S$ of $V$ of size at most $m/2$ such that there are no edges between $S$ and $\bar{S}$ is $o(1)$,

$$\mathrm{E}\left[\left|\left\{S \subseteq \bigcup_{k=1}^{m/2}\binom{V}{k} : \text{no edges between } S \text{ and } \bar{S}\right\}\right|\right] \leq \sum_{k=0}^{m/2}\left(\frac{e\cdot m}{k}(1-p)^{m-k}\right)^k. \tag{2}$$

Note that for small $k$, say $k \leq \frac{\epsilon}{2}\cdot m$, the ratio of the sum on the right-hand side of (2) can be upper bounded by $\frac{e\cdot m}{k}(1-p)^{m-k} \leq e\cdot m\cdot e^{-p(1-\frac{\epsilon}{2})m} \leq e\cdot m\cdot m^{-1(1+\frac{\epsilon}{2})} \to 0$, for $m\to\infty$. If $k \geq \frac{\epsilon}{2}\cdot m$, the ratio is $\frac{e\cdot m}{k}(1-p)^{m-k} \leq \frac{2e}{\epsilon}\cdot e^{-p\frac{m}{2}} \leq \frac{2e}{\epsilon}\cdot m^{-(1+\epsilon)/2} \to 0$, when $m\to\infty$. Thus, the right-hand side of (2) is $o(1)$ when $m\to\infty$. if $p \geq (1-\epsilon)p_c$.

If the graph is disconnected, then there is at least one subset $S$ of the above type since there must be a subset that is disconnected from the rest of the graph, and either it or its complement has size at most $m/2$. Now, since the expectation of the number of subsets of size at most $m/2$ such that there are no edges between $S$ and $\bar{S}$ is $o(1)$, by Markov's inequality, with all but $o(1)$ probability, the number of such subsets is 0, and so with $1-o(1)$ probability the graph is connected. □

In the connectivity example the critical probability is extreme in the sense that it converges to 0. An example in which the critical probability remains bounded away from 0 and 1 is the existence of a clique of size $\lfloor 2\log m\rfloor$. In this case, we can show a sharp threshold with $p_c \sim 1/2$. In particular, if $p > \frac{1}{2}+\epsilon$, then the probability that a clique exists is very large, and if $p < \frac{1}{2}-\epsilon$, then the probability is very small. This is a classical result in Ramsey theory and a prime illustration of the probabilistic method, but we will not go into the argument in this class.

We note that the arguments in the preceding proof sketches are very specific. We now show an argument that works for *all* non-trivial monotone graph properties, albeit with weaker parameters.

## 3.2 Every Monotone Graph Property Has a Sharp Threshold

First of all let us properly define sharp thresholds.

**Definition 1** (Weak Sharp Threshold)**.** *For all $\epsilon > 0$, the size of the interval on which $\mu_p$ goes from $\epsilon$ to $1-\epsilon$ is $o(1)$ in terms of $n$.*

Both examples satisfy this. But this does not say much about connectivity since in that case (1) $\to 0$ very quickly anyway as $n\to\infty$, and so naturally the size of the interval is $o(1)$. The following stronger definition corrects this shortcoming by comparing the size of the interval to the critical probability.

**Definition 2** (Strong Sharp Threshold)**.** *For all $\epsilon > 0$, the size of the interval on which $\mu_p$ goes from $\epsilon$ to $1-\epsilon$ is $o(\min(p_c, q_c))$ in terms of $n$, where $q_c = 1 - p_c$.*

Note that Definitions 1 and 2 are equivalent in cases where the critical probability is bounded away from 0 and 1, as in the clique example. Otherwise, Definition 2 is more strict. Our analysis of the connectivity threshold shows that the stricter Definition 2 applies there.

In our generic treatment, we will also consider those two cases. In the case of extreme critical probability, we show that the weaker Definition 1 applies to *every* nonconstant monotone $f$, but we make no claims about the stronger Definition 2. In the case of bounded critical probability, we exhibit a sharp threshold for nonconstant monotone functions $f$ that are weakly symmetric; graph properties give a natural class of examples of weakly symmetric functions.

### 3.2.1 The Case of Extreme Critical Probability

If $p_c$ is extreme, we can assume wlog that that $p_c$ is close to 0. By symmetry, our argument also holds for $p_c$ close to 1.

It turns out that weak thresholds hold for any monotone function $f$ in this case and the argument is similar to that of the Kolmogorov 0-1 law for inifinite graphs. Let $p_0$, $p_1$ be such that $\mu_{p_0} = \epsilon$ and $\mu_{p_1} = 1 - \epsilon$, respectively. Our goal is to argue that $p_1$ is not much larger than $p_0$.

Since $\mu_{p_0} = \epsilon$, if we sample strings $x_1, \ldots, x_k$ from a $p_0$-biased distribution, we have

$$\Pr_{x_1,\ldots,x_k \sim p_0}\left[\bigvee_{j=1}^{k} f(x_j) = -1\right] = 1 - (1-\epsilon)^k. \tag{+}$$

Since $f$ is monotonic, if $f(x_j) = -1$ for at least one $j$, then $f\left(\bigvee_{j=1}^{k} x_j\right) = -1$ where the $\bigvee$ denotes bitwise-OR.. So, (+) is at most

$$\Pr_{x_1,\ldots,x_k \sim p_0}\left[f\left(\bigvee_{j=1}^{k} x_j\right) = -1\right] = \Pr_{y \sim 1-(1-p_0)^k}[f(y) = -1],$$

where the equality is because $\left(\bigvee_{j=1}^{k} x_j\right)_i$ is 1 if and only if $(x_j)_i = 1$ for all $j$. This happens with probability $(1-p_0)^k$, so sampling $\left(\bigvee_{j=1}^{k} x_j\right)_i$ is equivalent to sampling a string from a $(1-(1-p_0)^k)$-biased distribution.

Combining these facts, we have

$$\Pr_{y \sim 1-(1-p_0)^k}[f(y) = -1] \geq 1 - (1-\epsilon)^k. \tag{++}$$

If the right-hand side of (++) is at least $1-\epsilon$, then by monotonicity of $\mu_p$ in $p$ and the definition of $p_1$, we must have $1 - (1-p_0)^k \geq p_1$. It is sufficient to choose $k \geq \frac{\log \epsilon}{\log(1-\epsilon)}$ to have the right-hand side of (++) $\geq 1-\epsilon$.

We have $p_1 \leq 1 - (1-p_0)^k \leq kp_0$ because $(1-p_0)^k \geq 1 - kp_0$. So, by choosing $k = \frac{\log \epsilon}{\log(1-\epsilon)}$, we have the following theorem.

**Theorem 1.** *For any nonconstant monotone function $f$, let $(p_0, p_1)$ be the interval on which $\mu_p(f)$ goes from $\epsilon$ to $1-\epsilon$. Then*

$$p_1 \leq \frac{\log \epsilon}{\log(1-\epsilon)} \cdot p_0.$$

Recall that to argue for sharp thresholds our goal is to establish that for all $\epsilon > 0$, the size of the interval in which $\mu_p$ goes from $\epsilon$ to $1-\epsilon$, that is $p_1 - p_0$, is $o(1)$. In the preceding, we have shown that $p_1$ is at most a constant factor, depending on $\epsilon$, larger than $p$. So, if $p_c$ is small, then $p_0$, and hence $p_1$, is small.

This gives a weak sharp threshold if $p_c = o(1)$. By symmetry, if $q_c = o(1)$ then we have a weak sharp threshold at the other extreme. So, we have shown that if the critical probability of any nontrivial monotone graph property approaches 0 or 1 as $n$ grows, then the property has a weak sharp threshold.

Note that even though this theorem applies to all non-trivial monotone functions, it only implies a weak sharp threshold at best and does not say anything for functions whose critical probability is bounded away from 0 and 1.

### 3.2.2 The Case of Bounded Critical Probability

In this case we cannot hope to have a result as general as for properties with extreme critical probabilities. For example, dictators have $p_c = 1/2$ and do not have a sharp threshold, as $\mu_p$ goes from 0 to 1 in a straight line. So we need an extra condition on $f$ to ensure a sharp threshold. Essentially, the function's behavior should not be local to a proper subset of variables. For example, a dictator function's behavior is local since only one variable determines the outcome. We will argue that it suffices for the function $f$ to be *weakly symmetric*.

Recall that a function is symmetric if it is invariant under all permutations of its variables. An example is *MAJORITY*. The "weak" qualifier means that instead of being invariant under all permutations, the function only needs to be invariant under some *transitive* permutation group $G$, where "transitive" means that any position can be reached from any other position using a permutation from $G$. Here is the formal definition.

**Definition 3** (weak symmetry)**.** *A group $G$ of permutations on $[n]$ is transitive if for every $i, j \in [n]$, there exists a permutation $\pi \in G$ such that $\pi(i) = j$. A function $f$ on $n$ variables is weakly symmetric if there exists a transitive group of permutations on $[n]$ such that for each $x$ in the domain of $f$ and each $\pi \in G$, $f(x) = f(x \circ \pi)$.*

In the above definition, $x \circ \pi$ denotes the action of $\pi$ on $x$ as we defined it in earlier lectures: the $i$th component of $(x \circ \pi)$ is the $\pi(i)$the component of $x$.

An example of a function that is weakly symmetric but not symmetric is Tribes. The function does not change when we switch variables within each $\bigwedge$ or swap the $\bigwedge$'s. Thus, it is invariant under the group generated by those permutations, which is transitive.

We now argue that if a monotone, nonconstant function is weakly symmetric, then it has a sharp threshold. The reason weak symmetry helps is that the influences of the individual variables of a weakly symmetric function are all the same, i.e., $I_i^{(p)}(f) = I_j^{(p)}(f)$ for all $i, j \in [n]$. This follows from the fact that the $p$-biased distribution is invariant under permutations: By definition of weak symmetry, if we consider any two variables $i$ and $j$, there exists a permutation $\pi \in G$ such that $\pi(i) = j$. Since both the distribution and $f$ are invariant under $\pi$, $I_i^{(p)}(f) = I_{\pi(i)}^{(p)}(f) = I_j^{(p)}(f)$.

In the previous lecture, we mentioned that every Boolean function $f$ on $n$ variables has a variable $i$ such that

$$I_i^{(p)}(f) \geq \frac{c \cdot \sigma_p^2(f)}{pq \log(1/(pq))} \cdot \frac{\log n}{n},$$

where $c > 0$ is a universal constant. Note that $\frac{c}{pq \log(1/(pq))} \geq d$ for some universal constant $d > 0$. Since the influences of the individual variables are equal by weak symmetry, we have

$$\begin{aligned} I^{(p)}(f) &= \sum_{i=1}^{n} I_i^{(p)}(f) \\ &\geq n \cdot \left( d \cdot \sigma_p^2(f) \cdot \frac{\log n}{n} \right) \\ &= d \cdot \mu_p(f)(1 - \mu_p(f)) \cdot \log n. \end{aligned}$$

By the corollary to the main theorem of the previous lecture, we know that for every monotone function $f$, $\frac{d\mu_p}{dp} = I^{(p)}(f)$. So, for $\epsilon \leq \mu_p \leq 1/2$,

$$\frac{d\mu_p}{dp} \geq \frac{d}{2} \cdot \mu_p \cdot \log n,$$

which we can rewrite as

$$\frac{d}{dp}(\ln \mu_p) \geq \frac{d}{2}. \tag{3}$$

By integrating (3 over $[p_0, p_c]$ where $p_0$ is the unique value of $p$ where $\mu_p(f) = \epsilon$, we get

$$\int_{p_0}^{p_c} \frac{d}{dp}(\ln \mu_p) dp = \ln \left( \frac{\mu_{p_c}}{\mu_{p_0}} \right) = \ln(\frac{1}{2\epsilon}) \geq \frac{d}{2} \cdot (\log n) \cdot (p_c - p_0).$$

We conclude that $p_c - p_0 \leq \frac{2 \ln(1/(2\epsilon))}{d \log n}$. A similar bound for $p_1 - p_c$ follows by symmetry. This gives us the following theorem.

**Theorem 2.** *Let $f$ be a weakly symmetric, nonconstant, monotone Boolean function on $n$ variables, and let $(p_0, p_1)$ be the interval on which $\mu_p(f)$ goes from $\epsilon$ to $1 - \epsilon$. Then $p_1 \leq p_0 + \frac{a \log(1/\epsilon)}{\log n}$, where $a$ is a universal constant.*

A special case of weakly symmetric functions are graph properties. The variables of graph properties represent the presence of edges in the graph. Graph properties are invariant under relabeling of vertices, or more precisely, under permutations of edges induced by relabeling vertices. Since these permutations can map any given edge to any other given edge, the group they form is transitive. Thus, we can apply the above theorem to monotone graph properties and conclude the following.

**Corollary 1.** *Every nonconstant monotone graph property has a sharp threshold in the sense of Definition 1.*

# 4 Next Lecture

We will discuss harmonic analysis over groups with more than two elements.