CS271 Randomness & Computation **Spring 2020**

Lecture 8: February 13

Instructor: Alistair Sinclair

8.1 Random 2-SAT and proof that $r_2 = 1$

We begin this lecture by completing the proof that the threshold for random 2-SAT exists and is equal to $r_2^* = 1$ (Theorem 8.3 of the previous lecture).

Proof of Theorem 8.3 (cont.): We proved one direction of Theorem 8.3 in the last lecture, by showing that $\Pr[\varphi \text{ is satisfiable}] \to 1$ for $r \leq (1-\epsilon)$. We now prove the other direction.

Assume $m = (1+\epsilon)n$. Define $t = n^{1/10}$. (We'll see the reason for this choice later.)

The approach we'll take is to define a collection of sets of clauses that are not satisfiable, and then show that with high probability, some such set appears in φ.

Definition 8.1 *A* snake *is a sequence of literals $w_1, w_2, \ldots, w_s$ with distinct variables, where $s = 2t-1$.*

We associate with a snake A a set of $s+1 = 2t$ clauses

$$F_A = (w_t \vee w_1) \wedge (\overline{w}_1 \vee w_2) \wedge \ldots \wedge (\overline{w}_{t-1} \vee w_t) \wedge (\overline{w}_t \vee w_{t+1}) \wedge \ldots \wedge (\overline{w}_{s-1} \vee w_s) \wedge (\overline{w}_s \vee \overline{w}_t)$$

Notice that this set of clauses is not satisfiable because it induces in the graph of implications a directed cycle that includes both w_t and $\overline{w}_t$.

Now we'll show that with high probability φ contains F_A for some A, which immediately implies that φ is not satisfiable.

As usual, define indicator variables and their sum:

$$X_A = \begin{cases} 1 & \text{if } \varphi \text{ contains every clause in } F_A \textit{ exactly once} \\ 0 & \text{otherwise} \end{cases}$$

and $X = \sum_A X_A$. We will use the second moment method to argue that $\lim_{n\to\infty} \Pr[X > 0] = 1$, which implies $\Pr[\varphi \text{ satisfiable}] \to 0$.

Recall our second moment calculation for cliques from the last lecture. Since the indicators X_A have the same symmetry properties as the clique indicators X_S from that proof, we know that it is sufficient to prove that, for any fixed A,

$$\sum_{B:B\sim A} \Pr[X_B = 1 | X_A = 1] = o(\mathrm{E}[X]) \text{ for any fixed } A$$

Here $B \sim A$ denotes that $A \neq B$ and X_A, X_B are not independent, which in our present context means that F_A and F_B share at least one clause.

Now observe that

$$\mathrm{E}[X_A] = \underbrace{\binom{m}{2t}}_{\text{choose } 2t \text{ clauses}} \underbrace{(2t)!}_{\text{choose ordering}} \underbrace{\left(\frac{1}{4\binom{n}{2}}\right)^{2t}}_{\text{prob. } 2t \text{ req'd clauses}} \underbrace{\left(1 - \frac{2t}{4\binom{n}{2}}\right)^{m-2t}}_{\text{remaining clauses distinct from } F_A} := h(2t).$$

Now let $B \sim A$ and let $k > 0$ be the number of clauses shared by F_A and F_B. Then we have

$$\Pr[X_B = 1 | X_A = 1] = \frac{\Pr[X_B = X_A = 1]}{\Pr[X_A = 1]} = \frac{h(4t-k)}{h(2t)}.$$

Thus

$$\sum_{B:B\sim A} \Pr[X_B = 1 | X_A = 1] = \sum_{k=1}^{2t-1} \frac{h(4t-k)}{h(2t)} P_k N,$$

where $P_k = \Pr[\text{random } F_B \text{ shares exactly } k \text{ clauses with a given } F_A]$ and N is the total number of snakes.

Now compare this with $\mathrm{E}[X] = \sum_A \mathrm{E}[X_A] = h(2t)N$:

$$\frac{\sum_{B:B\sim A} \Pr[X_B = 1 | X_A = 1]}{\mathrm{E}[X]} = \frac{\sum_{k=1}^{2t-1} h(4t-k)}{h(2t)^2} P_k. \tag{8.1}$$

Now let's examine $h(z)$:

$$h(z) = \binom{m}{z} z! \left(\frac{1}{4\binom{n}{2}}\right)^z \left(1 - \frac{z}{4\binom{n}{2}}\right)^{m-z} = (1+o(1)) \left(\frac{m}{2n(n-1)}\right)^z \qquad \text{provided } z \ll \sqrt{n},$$

and this holds uniformly in z. [**Exercise:** check this!] In our case, $t = n^{1/10} \ll \sqrt{n}$, so the above approximation is valid.

Therefore, the ratio in (**??**) can be written as

$$\frac{\sum_{B:B\sim A} \Pr[X_B = 1 | X_A = 1]}{\mathrm{E}[X]} \quad = \quad (1+o(1)) \sum_{k=1}^{2t-1} \left(\frac{2n(n-1)}{m}\right)^k P_k. \tag{8.2}$$

The final ingredient is to estimate the probabilities P_k. Using an elementary but slightly technical argument, it can be shown [CR92] that:

Fact 8.2 *The P_k are bounded as follows:*

1. $P_k < \frac{c_1 t^9}{n} \left(\frac{1}{2n}\right)^k$ *for* $1 \le k \le t-1$;
2. $P_k < c_2 t n \left(\frac{1}{2n}\right)^k$ *for* $1 \le k \le 2t$,

where c_1, c_2 are universal constants.

[Those who like combinatorial exercises may enjoy the challenge of obtaining these (or similar) bounds.]

Recall that the overall goal is to show that the quantity in (**??**) above goes to 0. So split the sum into two parts. Using part 1 of Fact **??**, the first half of the sum becomes

$$\sum_{k=1}^{t-1} \left(\frac{2n(n-1)}{m}\right)^k P_k \le \frac{c_1 t^9}{n} \sum_{k=1}^{t-1} \left(\frac{n-1}{m}\right)^k,$$

which approaches 0 as n grows by choice of t and the fact that $m = (1+\epsilon)n$.

Using part 2 of Fact **??**, the second half of the sum is

$$\sum_{k=t}^{2t-1} \left(\frac{2n(n-1)}{m}\right)^k P_k \leq c_2 tn \sum_{k=t}^{2t-1} \left(\frac{n-1}{m}\right)^k.$$

By our choice of t, the largest term of the geometric series on the right-hand side is $c^{n^{1/10}}$, where $c < \frac{1}{1+\epsilon} < 1$. This kills the term $c_2 tn = O(n^{11/10})$ as $n \to \infty$, so the second half of the sum also tends to 0.

This completes the proof of Theorem 8.3. ■

8.2 Thresholds for k-SAT

We conclude our discussion of random k-SAT by sketching a proof of the result by Achlioptas and Peres [AP04] which establishes an almost tight lower bound on the threshold. Their lower bound is $2^k \ln 2 - O(k)$, which is remarkably close to the trivial upper bound of $2^k \ln 2$ (see last lecture), and pins down the threshold exactly asymptotically in k, namely, $r_k \sim 2^k \ln 2$. (Recall from the previous lecture that the exact value of the threshold for large k is $2^k \ln 2 - \frac{1}{2}(1+\ln 2) - \varepsilon_k$, by [CP16,DSS14].) The proof is non-algorithmic, and builds on a previous breakthrough by Achlioptas and Moore [AM02] which was also non-algorithmic and gave a lower bound of the form $2^{k-1} \ln 2 - \Theta(1)$, off by a factor of 2.

Theorem 8.3 [AP04] *There exists a sequence $\delta_k \to 0$ such that for all $k \geq 3$,*

$$r_k \geq 2^k \ln 2 - (k+1)\frac{\ln 2}{2} - 1 - \delta_k.$$

In particular, $r_k \geq 2^k \ln 2 - k$ for all $k \geq 3$.

The following table presents a comparison between the known upper bounds and the lower bounds derived in [AP04]:

k	3	4	5	10	20
Upper bound	4.51	10.23	21.33	708.94	726,817
Lower bound	2.68	7.91	18.79	704.94	726,809

The proof involves a clever application of the second-moment method. We begin with a slightly more refined version of the standard second-moment bound:

Lemma 8.4 *Let X be any non-negative random variable. Then*

$$\Pr[X > 0] \geq \frac{\mathrm{E}[X]^2}{\mathrm{E}[X^2]}.$$

Proof: A useful Exercise! [Hint: Assume for simplicity that X is discrete. Start from the expression $\mathrm{E}[X^2]\Pr[X > 0] - \mathrm{E}[X]^2$ and expand each term into a sum over possible values of X.]

Exercise: Compare this bound with the analogous one derived from Chebyshev's inequality, which you should check is $\Pr[X > 0] \geq 2 - \frac{\mathrm{E}[X^2]}{\mathrm{E}[X]^2}$.

Now let $X = \sum_\sigma X_\sigma$ be the number of satisfying assignments of a random formula φ, where each X_σ is 1 if σ is a satisfying assignment and 0 otherwise.

We already know that the probability of satisfiability sharply goes from to 1 to 0 around a *sequence* of densities $r_k(n)$ (see the result of Friedgut [F99] quoted in the last lecture). Thus it is sufficient to prove that if $r < 2^k \ln k - k$ then $(\mathrm{E}[X])^2/\mathrm{E}[X^2] \geq \epsilon$ for some constant $\epsilon > 0$. For then by Lemma **??** we have that $\Pr[X > 0] \geq \epsilon$, and Friedgut's result immediately implies that the probability must in fact tend to 1.

Recall from the previous lecture that

$$\mathrm{E}[X] = \sum_\sigma \mathrm{E}[X_\sigma] = 2^n(1 - 2^{-k})^{rn}.$$

Also, we have

$$\begin{aligned}
\mathrm{E}[X^2] &= \mathrm{E}[(\sum_\sigma X_\sigma)^2] \\
&= \sum_{\sigma,\tau} \mathrm{E}[X_\sigma X_\tau] \\
&= \sum_{\sigma,\tau} \Pr[\text{both } \sigma \text{ and } \tau \text{ satisfy } \varphi] \\
&= \sum_{\sigma,\tau} \prod_{i=1}^{m} \Pr[\text{both } \sigma \text{ and } \tau \text{ satisfy the } i^{\text{th}} \text{ clause}].
\end{aligned}$$

To compute the probability that two given assignments satisfy a particular random clause, we assume that the assignments agree on exactly $z = \alpha n$ variables. (We call the set of variables on which the assignments agree the *overlap* of the assignments.) The probability that one of them does not satisfy the clause is 2^{-k}. If both the assignments do not satisfy the clause then their overlap must include all the variables in the clause; thus the probability that this happens is $2^{-k}\alpha^k$. Therefore, by inclusion-exclusion:

$$\Pr[\text{two given assignments with overlap } \alpha n \text{ both satisfy a given clause}] = 1 - 2^{1-k} + 2^{-k}\alpha^k.$$

We denote this quantity by $f(\alpha)$. [Note for future reference that $f(1/2) = (1 - 2^{-k})^2$, which means that assignments which overlap on exactly half the variables are uncorrelated.] Using this, we have

$$\mathrm{E}[X^2] = 2^n \sum_{z=0}^{n} \binom{n}{z} f(z/n)^m. \tag{8.3}$$

Using the approximation $\binom{n}{\alpha n} \approx \left(\frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\right)^n \Theta(\frac{1}{\sqrt{n}})$ for α not close to 0 or 1, we get

$$\mathrm{E}[X^2] \;\approx\; \sum_{z=0}^{n} \left(\frac{2f(\alpha)^{m/n}}{\alpha^\alpha(1-\alpha)^{1-\alpha}}\right)^n \Theta\left(\frac{1}{\sqrt{n}}\right) \qquad (\alpha = z/n).$$

(Here we are assuming that the big errors at the two ends (α close to 0 and 1) don't affect the sum very much.)

Writing as usual $r = \frac{m}{n}$, and using $\Lambda(\alpha)$ to denote $\frac{2f(\alpha)^r}{\alpha^\alpha(1-\alpha)^{1-\alpha}}$, we have

$$\mathrm{E}[X^2] \approx \sum_{z=0}^{n} (\Lambda(\alpha))^n \Theta(\frac{1}{\sqrt{n}}).$$

Note also that

$$\mathrm{E}[X]^2 = (2^n(1-2^{-k})^{rn})^2 = (4f(1/2)^r)^n = \Lambda(1/2)^n.$$

This gives us an approximation for the ratio that we want to estimate:

$$\frac{\mathrm{E}[X^2]}{\mathrm{E}[X]^2} \approx \sum_{z=0}^{n} \left(\frac{\Lambda(\alpha)}{\Lambda(1/2)}\right)^n \Theta(\frac{1}{\sqrt{n}}). \tag{8.4}$$

We want to prove that the inverse of the above ratio is greater than some constant $\epsilon > 0$. Clearly, a necessary condition for this is that $\Lambda(\alpha)$ is maximized for $0 \leq \alpha \leq 1$ at $\alpha = 1/2$; otherwise, the largest term in the above sum will grow exponentially with n, and the sum will be unbounded. Unfortunately, however, this is not the case, as can be seen from the graphs of the functions $\varepsilon(\alpha) \equiv 1/\alpha^\alpha(1-\alpha)^{1-\alpha}$ and $f(\alpha)$ in Fig. **??** and Fig. **??** respectively. Since the first function is symmetric about $\alpha = 1/2$ and the second is monotone, their product cannot have an extremum at $\alpha = 1/2$. Therefore, the vanilla second moment method is doomed in this setting.

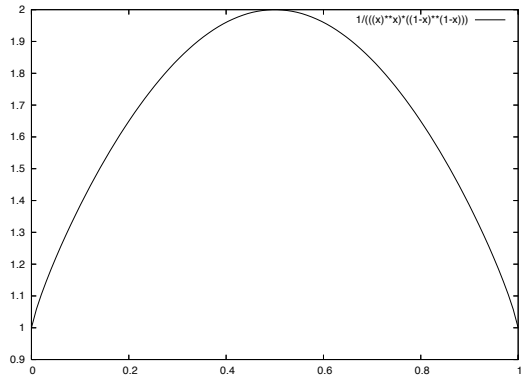

Figure 8.1: $\frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}}$

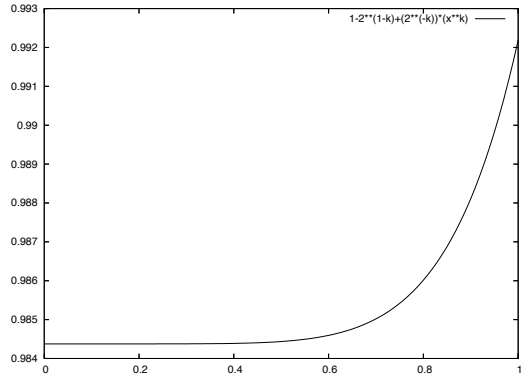

Figure 8.2: $f(\alpha)$ for $k = 7$

A useful intuition to bear in mind in what follows is that the functions $\varepsilon(\alpha)$ and $f(\alpha)$ correspond essentially to "entropy" and "correlation" respectively. Their product, which is the function of interest to us, reflects a tension between these two quantities. The entropy is always maximized at $\alpha = 1/2$. We would like to somehow alter the "correlation" function so that it also has a turning point (actually a minimum) at

$\alpha = 1/2$, thus making it possible for the product to be maximized at this point. This will involve changing the correlation function so as to boost the contribution of uncorrelated pairs of assignments in the sum for $\mathrm{E}[X^2]$, thus reducing the second moment.

The key idea in [AP04] is to apply the second moment method to a different random variable X. Note that we are free to apply the second moment to any r.v. X that is zero when the formula is unsatisfiable (for then a lower bound on $\Pr[X > 0]$ is a lower bound on the probability that φ is satisfiable). So, define a *weight function* $W(\sigma, \varphi)$ which is 0 when σ does not satisfy φ and positive otherwise. Moreover, assume that W is multiplicative over clauses, i.e.,

$$W(\sigma, \varphi) = \prod_{i=1}^{m} w(\sigma, c_i)$$

where c_i is the i^{th} clause and $w(\sigma, c_i) = 0$ unless σ satisfies c_i. Then we consider the random variable

$$X = \sum_{\sigma} \prod_{i=1}^{m} w(\sigma, c_i).$$

By analogy with (**??**) above, it is easy to see that

$$\mathrm{E}[X^2] = 2^n \sum_{z=0}^{n} \binom{n}{z} f_w(z/n)^m, \tag{8.5}$$

where $f_w(z/n) = \mathrm{E}[w(\sigma, c)w(\tau, c)]$ is the correlation between two assignments with overlap $z = \alpha n$, with respect to a single random clause c. Similarly, it is easy to check that $\mathrm{E}[X]^2 = (4f_w(1/2)^r)^n$. Hence, as in (**??**), in order to have any chance of success we once again want $f_w(\alpha)$ to have a turning point at $\alpha = 1/2$.

Now note that w.l.o.g. we may assume by symmetry that $w(\sigma, c)$ is a function only of the vector $\mathbf{v} \in \{\pm 1\}^k$, where

$$v_i = \begin{cases} +1 & \text{if } \sigma \text{ satisfies the } i\text{th literal of } c \\ -1 & \text{otherwise.} \end{cases}$$

Now the condition that $f_w'(1/2) = 0$ can be shown to be equivalent to the geometric condition

$$\sum_{\mathbf{v}} w(\mathbf{v})\mathbf{v} = 0. \tag{8.6}$$

(This is an interesting **exercise**.) Hence any successful weight function w must be *balanced*, in the sense that the weighted sum over all vectors $\mathbf{v}$ in the cube $\{\pm 1\}^k$ is zero. Note that the vanilla second moment method can be viewed as using the "flat" weight function $w(-1, \ldots, -1) = 0$ and $w(\mathbf{v}) = 1/(2^k - 1)$ for all other $\mathbf{v}$. (We assume w.l.o.g. that w is normalized to sum to 1.) This is clearly *not* balanced.

So our goal is to find a weight function w that is "as close as possible to" the original flat one while also being balanced. Thus we seek a function w on $\{\pm 1\}^k$ such that $w(-1, \ldots, -1) = 0$, $\sum_{\mathbf{v}} w(\mathbf{v})\mathbf{v} = 0$, $\sum_{\mathbf{v}} w(\mathbf{v}) = 1$, and such that the entropy over $\mathbf{v} \neq (-1, \ldots, -1)$ is maximized. (This is how we formalize the idea of being as close as possible to the flat weight function.) This entropy maximization problem is not hard to solve, once one observes that it is sufficient to make $w(\mathbf{v})$ a function only of the *height* $h(\mathbf{v})$ of $\mathbf{v}$, i.e, the number of $+1$'s in $\mathbf{v}$. (This is equivalent to making $w(\sigma, c)$ a function only of the number of literals of c satisfied by σ.) The resulting single-parameter optimization problem has solution

$$w(\mathbf{v}) = \begin{cases} 0 & \text{if } \mathbf{v} = (-1, \ldots, -1) \\ \lambda^{h(\mathbf{v})}/Z & \text{otherwise,} \end{cases}$$

where λ satisfies $(1 + \lambda)^{k-1} = \frac{1}{1-\lambda}$ and Z is a normalizing constant. (This value of λ ensures that w satisfies the balance condition (**??**); **exercise:** check this!)

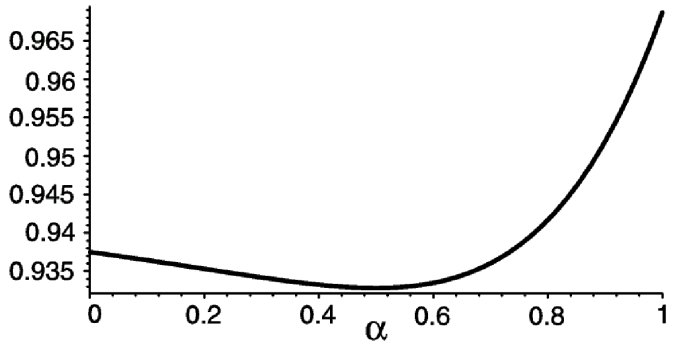

Figure 8.3: $f_w(\alpha)$ with $k = 5$ [AP04]

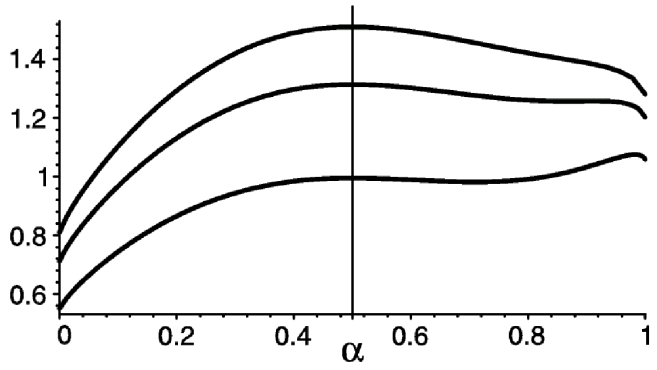

Figure 8.4: $\Lambda_w(\alpha)$ with $k = 5$ and $r = 14, 16, 20$ [AP04]

With this choice of weight function, the function $f_w(\alpha)$ takes the shape shown in Fig. **??**, which indeed has a turning point at $\alpha = 1/2$. This means that we can hope that $\Lambda_w(\alpha) = \frac{2f_w(\alpha)^r}{\alpha^\alpha(1-\alpha)^{1-\alpha}}$ is maximized at $\alpha = 1/2$, and hence that the sum in (**??**) is bounded.

Fig. **??** shows the approximate shape of $\Lambda_w(\alpha)$ for various values of r. When r is small enough, $\Lambda_w(\alpha)$ does indeed peak at $\alpha = 1/2$. However, as r gets larger the factor $f_w(\alpha)^r$ overwhelms the product and a second peak, to the right of $\alpha = 1/2$, becomes the global maximum. Detailed calculations show that the sum in (**??**) is bounded above by a constant as long as $r \leq 2^k \ln 2 - (k+1)\frac{\ln 2}{2} - 1 - \delta_k$, which gives the result claimed in Theorem **??**.

Note that $\lambda < 1$, so the weight function penalizes assignments that satisfy too many literals. The reason for the failure of the vanilla second moment here is that assignments that satisfy many literals tend to dominate, and they also tend to be correlated with one another (because all of them are correlated with the "majority" assignment, which sets each variable to its majority value over the clauses); this correlation makes the second moment too large. To counteract this effect, what we really want to do is to design a weight function that puts most of the weight in the second moment $\mathrm{E}[X^2]$ on assignments that satisfy close to half the literals (and thus have low correlation); rather than doing this explicitly, the above weight function achieves the same effect in a "smooth" way that makes the analysis easier.

This completes our sketch of the proof.

References

[AM02] D. Achlioptas and C. Moore, "The asymptotic order of the random k-SAT threshold," *Proceedings of the 43rd IEEE FOCS*, 2002, pp. 779–788.

[AP04] D. Achlioptas and Y. Peres, "The threshold for random k-SAT is $2^k \log 2 - O(k)$," *Journal of the American Mathematical Society* **17** (2004), pp. 947–973.

[CP16] A. Coja-Oghlan and K. Panagiotou, "The asymptotic k-SAT threshold," *Advances in Mathematics* **288** (2016), pp. 985–1068.

[DSS14] J. Ding, A. Sly and N. Sun. "Proof of the satisfiability conjecture for large k," ArXiv preprint arXiv:1411.0650 [math.PR], 2014.

[F99] E. Friedgut, "Necessary and sufficient conditions for sharp thresholds of graph properties," *Journal of the American Mathematical Society* **12** (1999), pp. 1017–1054.