The threshold for random k-SAT is $2^k \ln(2) - O(k)$

Dimitris Achlioptas and Yuval Peres

1 Introduction

Let $V = \{X_1, X_2, \ldots, X_n\}$ be a set of n Boolean variables. A k-CNF instance $I_k(n, m)$ is a propositional formula in the form $I_k(n, m) = c_1 \wedge c_2 \wedge \ldots c_m$, where each k-clause c_i is a disjunction of k literals over a set V, besides a literal and its complement cannot appear in the same clause.

Let $C_k(V)$ denote the set of all $2^k n^k$ possible k-clauses on V. A random k-CNF formula $F_k(n, m)$ is formed by selecting uniformly, independently and with replacement m clauses from $C_k(v)$ and taking their conjunction. The interesting question is: Given a formula chosen uniformly at random, Is there an assignment of $X_1, X_2, \ldots, X_n$ such that $F_k(n, m)$ evaluates to true? if that is the case, then find such a satisfying assignment.

Is known by different methods in statistical mechanics that in the random constrain satisfaction problem there exist a SAT and UNSAT (satisfiable and unsatisfiable) phase transition at a critical value of its density (the number of constraints per variable, i.e. m/n), however the exact value for this critical point is still unknown. In this paper, Achlioptas and Peres show an upper bound for the "random k-satisfaction problem" which is the best rigorous bound founded until now.

It will say that a sequence of events ε_n occurs with hight probability (w.h.p.) if $\lim_{n\to\infty} P(\varepsilon_n) = 1$ and with uniformly positive probability if $\liminf_{n\to\infty} P(\varepsilon_n) > 0$.

Throughout the paper k is arbitrarily large but fixed, while $n \to \infty$. For each $k \geq 2, let$

$$
\begin{aligned}
r_k &\equiv sup\{r : F_k(n, rn) \text{ is satisfiable w.h.p}\}, \\
r_k^* &\equiv inf\{r : F_k(n, rn) \text{ is unsatisfiable w.h.p}\}.
\end{aligned}
$$

Clearly, $r_k \leq r_k^*$. The *Satisfiability Threshold Conjecture* asserts that $r_k = r_k^*$ for all $k \geq 3$. The main result in this paper establishes an asymptotic form of this conjecture.

Theorem 1 *As $k \to \infty$*

$$r_k = r_k^*(1 - o(1)).$$

Theorem 2 *There exist a sequence $\delta_k \to \infty$ such that, for all $k \geq 3$,*

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

In 1983 franco and Paull ([1]) observed by the first moment argument that $r_k^* \leq 2^k \ln(2)$. Thus, Theorem 2 implies that this bound is asymptotically tight.

2 Main Ideas

In 1999 Friedgut ([2]) proved the existence of a *non-uniform* threshold given by the next theorem.

Theorem 3 *For each $k \geq 2$, there exist a sequence $r_k(n)$ such that for every $\epsilon > 0$,*

$$\lim_{n\to\infty} P[F_k(n, rn) \text{ is satisfiable}] = \begin{cases} 1 & \text{if } r = (1-\epsilon)r_k(n) \\ 0 & \text{if } r = (1+\epsilon)r_k(n) \end{cases}$$

The method used to prove Theorem 2 is known as "the second moment argument" It uses the next lemma which establish that for any non-negative random variable X one can get a lower bound on $P(X > 0)$ by the following inequality.

Lemma 1 *For any non-negative random variable X,*

$$P(x > 0) \geq \frac{E[X]^2}{E[X^2]}. \tag{1}$$

In particular, if X denotes the number of satisfying assignments of a random formula $F_k(n, rn)$, one can get a lower bound on the probability of satisfiability by applying (1) to X. Thus, the following immediate corollary of Theorem 3 implies that if $P(X > 0) > 1/C$ for any constant $C > 0$, then $r_k \geq r$.

Corollary 1 *Fix $k \geq 2$. If $F_k(n, rn)$ is satisfiable with uniformly positive probability, then $r_k \geq r$.*

Unfortunately, this is never the case: for every constant $r > 0$, there exists $\beta = \beta(r) > 0$ such that $E[X^2] > (1+\beta)^n E(x)^2$. For this reason, the authors use a delicate application of the second moment method. They introduce a scheme that weights satisfying truth assignments according to their number of satisfied literal occurrences.

References

[1] J. Franco, M. Paull (1983) Probabilistic analysis of the Davis-Putman procedure for solving the satisfiability problem. *Discrete Appl. Math.* **Vol.5,Number 1** 77–87. MR0678818 (84e:68038)

[2] E. Friedgut (1999) Necessary and sufficient conditions for sharp threshold of graph properties, and the k-SAT problem. *J. Amer. Math. Soc.* **Vol.12** 1017–1054.