Phase Transition in Random K-SAT

Ying-Yu Ho*
Department of Physics, Massachusetts Institute of Technology
(Dated: May 18, 2012)

This paper introduces *the satisfiability problem*, which lies at the heart of computation theory, from a statistical mechanical view. The emphasis is on the average behavior of random SAT ensembles, especially the abrupt transition in the ratio of satisfiable Boolean formulas.

1. THE SATISFIABILITY PROBLEM

Satisfiability (SAT) is the problem of determining whether a given Boolean formula allows an assignment of variables that makes the formula evaluate to TRUE (satisfiable) or, conversely, whether a Boolean formula is identically FALSE (unsatisfiable).

SAT is central to computational complexity theory for it serves as a prototype of combinatorial optimization problems and, more importantly, it is *NP-complete*. Roughly speaking, the latter means that SAT represents all problems that may not be decidable in polynomial time but allow proofs (e.g. true-assignments of variables) to be verified in polynomial time.

It is conventional to consider only Boolean formulas in *conjunctive normal form* (CNF). For instance,

$$F = C_1 \wedge C_2 \wedge C_3 \cdots \wedge C_M, \tag{1}$$

and the *clauses* C_i could be

$$C_1 = (x_i \vee \bar{x}_j \vee \cdots),\ C_2 = \cdots \tag{2}$$

Here $\bar{x}_i$ is the negation of the *variable* x_i. Both x_i and $\bar{x}_i$ are called *literals*.

K-SAT is the subset of SAT with each clause containing exactly K literals. While K-SAT with $K \geq 3$ is NP-complete and essentially as hard as SAT, simple polynomial-time algorithms for 1-SAT and 2-SAT are known [1].

1.1. Random K-SAT

There are $\binom{N}{K}2^K$ ways to form a K-literal clause from N variables. Let $\mathrm{SAT}_K(N, M)$ denote the ensemble generated by including in the formula each of the $\binom{N}{K}2^K$ different clauses independently with probability $M2^{-K}/\binom{N}{K}$ so that the mean number of clauses in a formula is M. When we fix the ratio $\alpha \equiv M/N$ and consider the limit $N \to \infty$, interesting phenomena resembling phase transitions in statistical physics occur.

2. PHASE TRANSITION IN RANDOM K-SAT

Intuitively, we expect that a Boolean formula is less likely to be satisfiable as more clauses are added. From the view of statistical physics, it is natural to suspect that some kind of phase transition will occur in the large N limit. This section discusses phase transition as the jump between an almost satisfiable random SAT ensemble and an almost unsatisfiable one.

2.1. $K = 2$

The $K = 2$ case is discussed first because it is more tractable but not trivial. Random 2-SAT undergoes a phase transition between a satisfiable phase and an unsatisfiable phase at $\alpha_c = 1$ [3][4]. More precisely, let $P_K(N, \alpha)$ be the portion of satisfiable formulas in $\mathrm{SAT}_K(N, \alpha N)$. Then

$$\lim_{N\to\infty} P_2(N, \alpha) = \begin{cases} 1 & \text{if } \alpha < 1; \\ 0 & \text{if } \alpha > 1. \end{cases} \tag{3}$$

In addition to a mathematical proof omitted here, it is enlightening to guess the same conclusion through an (loose) RG analysis.

Consider eliminating a variable from the ensemble $\mathrm{SAT}_2(N, M)$ so that

$$N' = N - 1. \tag{4}$$

Each formula should be transformed without changing its Boolean function. For example, if x_4 is being eliminated, then the formula

$$(x_1 \vee x_4) \wedge (\bar{x}_1 \vee \bar{x}_4) \wedge (\bar{x}_2 \vee x_4) \wedge (x_3 \vee \bar{x}_4) \tag{5}$$

should become

$$\begin{aligned} & (x_1 \wedge \bar{x}_2) \vee (\bar{x}_1 \wedge x_3) \\ = \ & (x_1 \vee \bar{x}_1) \wedge (x_1 \vee x_3) \wedge (\bar{x}_2 \vee \bar{x}_1) \wedge (\bar{x}_2 \vee x_3) \end{aligned} \tag{6}$$

There are certainly infinite ways to write down another equivalent boolean formula, but the above transformation is most straightforward and yields a new ensemble *closest* to a random one. The mean number of clauses in the transformed ensemble is

$$M' = M - 2\alpha + \alpha^2. \tag{7}$$

*Electronic address: yingyu@mit.edu

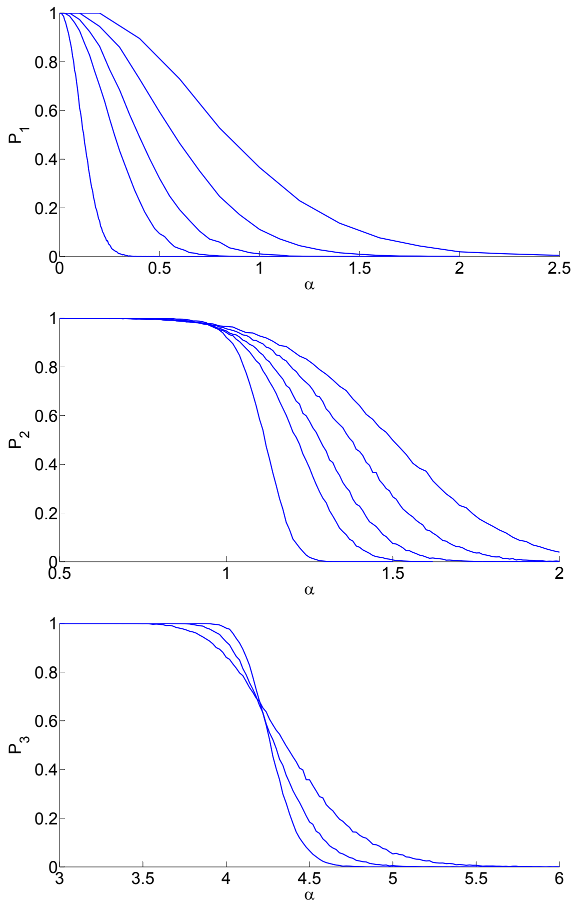

FIG. 1: The probability that a formula from the random ensemble $\mathrm{SAT}_K\,(N, M = \alpha N)$ is satisfiable. *Up*: $K = 1$, $N = 5, 10, 20, 40, 200$ (from *right* to *left*); *middle*: $K = 2$, $N = 50, 100, 200, 400, 2000$ (from *right* to *left*); *bottom*: $K = 3$, $N = 50, 100, 200$ (increasing steepness). Each data point is obtained by averaging over 10^4 random formulas using the SAT-solver core described in [2].

The possibility of eliminating clauses such as $(x_4 \vee \bar{x}_4)$ is not considered because their $\mathcal{O}(M/N^2)$ contribution is negligible in the large N limit. Pretend that the new ensemble is indeed random and call it $\mathrm{SAT}_2\,(N', M')$ for the time being. Then the new clause-to-variable ratio is

$$\alpha' = \frac{M'}{N'} = \alpha \cdot \left(1 + \frac{\alpha - 1}{N - 1}\right). \tag{8}$$

We can immediately identify the critical value

$$\alpha_c = 1, \tag{9}$$

which coincides with the exact value in Eq. (3).

The transformation Eq. (8) fails to predict the eigenvalue at $\alpha = \alpha_c = 1$, as indicated by Fig. 2, in which applying the transformation on $N = 500$ ensembles 100 times does not give the correct description of $N = 400$ ensembles. The correct scaling of the transition region in α is $\Delta\alpha \sim N^{-1/3}$ (Fig. 3) instead of $\Delta\alpha \sim N^{-1}$ predicted by Eq. (8), although the proof is involved [4]. Looking carefully at the example Eq. (6), we can find that the transformed clauses tend to contain similar literals when $\alpha > 1$, and therefore the new ensemble has less information content than the random ensemble with M given by Eq. (7). On the other hand, if the transformation results in mostly mergers of two clauses with opposite literals x_i and $\bar{x}_i$, the new ensemble will be close to a random one, which is the case at the critical point $\alpha_c = 1$.

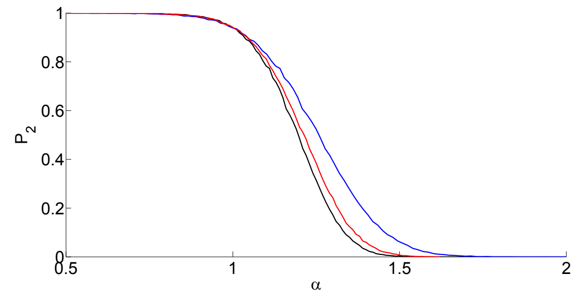

FIG. 2: $P_2\,(N, \alpha)$ versus α. *Black*: $N = 500$; *red*: $N = 400$; *blue*: apply Eq. (8) 100 times on the $N = 500$ ensemble.

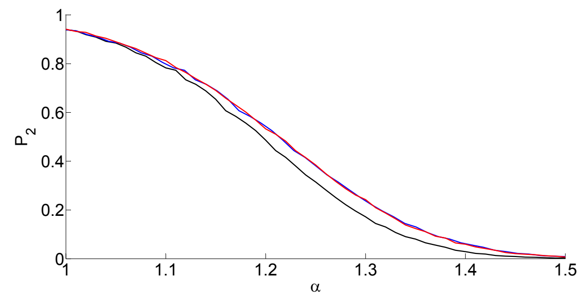

FIG. 3: $P_2\,(N, \alpha)$ versus α. *Black*: $N = 500$; *red*: $N = 400$; *blue*: apply Eq. (8) 33 times on the $N = 500$ ensemble.

2.2. $K = 1$

With N variables and $M = \alpha N > 0$ one-literal clauses, it is almost certain that both x_i and $\bar{x}_i$ appear in the same formula for some i in the $N \to \infty$ limit. Thus a random formula is almost always unsatisfiable for all $\alpha > 0$.

An "RG" analysis similar to the $K = 2$ case is not applicable because there is no obvious way to eliminate a variable x_i from a formula such as $x_i \wedge \bar{x}_i \cdots$ while preserving its Boolean function.

2.3. $K = 3$ or more

Similar to Eq. (3), it is conjectured that for any $K \geq 2$, there exists $\alpha_c(K)$ such that

$$\lim_{N \to \infty} P_K\,(N, \alpha) = \begin{cases} 1 & \text{if } \alpha < \alpha_c(K); \\ 0 & \text{if } \alpha > \alpha_c(K). \end{cases} \tag{10}$$

For example, numerical simulations give $\alpha_c(3) \approx 4.17$ (Fig. 1) [5]. While numerical and theoretical studies favor the conjecture, a definitive proof is lacking [3].

It is difficult to find a procedure that simplifies 3-SAT formulas like Eq. (6); otherwise it would be a polynomial-time algorithm for an NP-complete problem, the existence of which is still a big open problem [1].

3. REMARKS

The connection between statistical physics and computation is not unfamiliar, and there are other ways in which the SAT problem can be treated physically. For example, SAT can be formulated as spin glasses, a prototypical family of disordered systems. Deciding satisfiability then corresponds to finding the ground state (annealing), and random SAT ensemble is mapped to quenched disorder [5].

Interestingly, there is an analogy between the complexity K-SAT and K-dimensional Ising model. Both problems are readily solvable at $K = 1$, not quite trivial at $K = 2$, but they present great challenges when $K = 3$. In fact, it has been proven that 3-D Ising model is NP-complete by relating 3-SAT and 3-D Ising model [6]. Nevertheless, it should be noted that 4-D Ising model becomes easy again due to applicability of mean-field theory while 4-SAT is not less intractable.

[1] M. Sipser, *Introduction to the Theory of Computation* (Thomson Course Technology, 2006), 2nd ed.

[2] N. Eén and N. Sörensson (2003), URL http://www.minisat.se/downloads/MiniSat.pdf.

[3] M. Mézard and A. Montanari, *Information Physics, and Computation* (Oxford University Press, 2009).

[4] W. F. de la Vega, Theoretical Computer Science **265**, 131 (2001).

[5] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing* (Oxford University Press, 2001).

[6] S. Istrail, in *Proceedings of the thirty-second annual ACM symposium on Theory of computing* (ACM, New York, NY, USA, 2000), STOC '00, pp. 87–96, ISBN 1-58113-184-4, URL http://doi.acm.org/10.1145/335305.335316.