Phase Transitions of Bounded Satisfiability Problems*

Delbert D. Bailey, Phokion G. Kolaitis
Computer Science Department
University of California, Santa Cruz
Santa Cruz, CA 95064, U.S.A
{dbailey,kolaitis}@cse.ucsc.edu

Abstract

We investigate phase transitions for the family of bounded satisfiability problems 3SAT(b), recently introduced by Zhang, that ask: given a 3CNF-formula, is there a truth assignment that violates no more than b of its clauses. Zhang's results were experimental and for a fixed number of variables (n = 25), and suggested that the locations of the phase transitions for 3SAT(b) are separated and move significantly as *b* increases. Analysis of these locations was posed as an open question.

We analytically show that the phase transitions of all 3SAT(6) problems must occur within a narrow region, regardless of how large the value of *b* is. Moreover, our experiments reveal that the phase transitions for these problems occur in a remarkable way. Specifically, unlike 3SAT, the probability curves for 3SAT(6) do not have a quasi-common intersection point about which they rotate as they become steeper with increasing n. Instead, they move rapidly to the left toward the narrow region that the analysis predicts.

1 Introduction and Summary of Results

A phase transition of a "system" can be described as an abrupt change in the behavior of the system that occurs when a certain "control parameter" is at or near a certain critical value. Traditionally, phase transitions have been studied by physicists working in the area of statistical mechanics. During the past decade, however, computer scientists have carried out an intensive study of phase transitions of algorithmic problems, first of NP-complete decision problems and, more recently, of decision problems that are complete for higher computational complexity classes. These investigations have shed new light on the "structure" of presumably intractable decision problems by examining them from a perspective that had been hitherto unexplored in computer science; moreover, for certain fundamental algorithmic problems, the location of the phase transition has been correlated to the peak average cost of natural algorithms for solving these problems.

* Research of the authors was partially supported by NSF Grant No. IIS-9907419

The Boolean satisfiability problems $k\mathrm{SAT}, k \geq 3$, constitute the most thoroughly investigated collection of NP-complete problems from the perspective of phase transitions. An instance of $k\mathrm{SAT}$ is a $k\mathrm{CNF}$ formula; the "control parameter" of such a formula is the ratio of the number of clauses over the number of variables occurring in the formula. Each fixed value *r* of this ratio gives rise to the family of probability spaces $F_k(n,r), n \geq 1$, where $F_k(n,r)$ is the collection of all kCNF-formulas with rn clauses generated by selecting *k* variables from n variables without replacement and then negating each variable with probability 1/2. Franco and Paul [Franco and Paul, 1983] were the first to focus on this fixed clauses-to-variables ratio model and to initiate a study of the asymptotic behavior of the probability $p_k(n,r)$ that a random formula in $F_k(n,r)$ is satisfiable. During the past decade, much of the work in this area has been motivated from the conjecture of Chvatal and Reed [Chvatal and Reed, 1992] to the effect that, for every $k \geq 3$, there is a positive real number rk such that if $r < r_k$, then $\lim_{n\to\infty} p_k(n,r) = 1$, whereas if $r > r_k$, then $\lim_{n\to\infty} p_k(n,r) = 0$.

The above conjecture asserts that, for every $k \geq 3$, a phase transition occurs in the probability $p_k(n,r)$ of a random formula in $F_k(n,r)$ being satisfiable, as $n \to \infty$. In spite of intensive efforts by several researchers, this conjecture has not been settled thus far. Nonetheless, progress toward establishing this conjecture has been made in two different fronts. On the experimental front, large-scale experiments with random Boolean formulas have provided evidence for the existence of a critical ratio r_k and have yielded estimates of its actual value. In particular, experiments by Selman, Mitchell and Levesque [Selman *et al.,* 1996] with random 3CNF-formulas and analysis of these experiments by Kirkpatrick and Selman [Kirkpatrick and Selman, 1994] indicate that r_3 is about 4.2. On the analytical front, there has been continuous progress toward establishing progressively tighter upper and lower bounds for the value of r_3. The best analytical results obtained to date assert that if *r3* exists, then $3.42 < r_3 < 4.507$ [Kaporis *et* al., 2002; Dubois *et al,* 2000]. The experiments carried out by Selman, Mitchell and Levesque [Selman *et al,* 1996] for random 3CNF-formulas also revealed that the critical ratio 4.2 appears to be the location at which the average cost of the Davis-Putnam-Logemann-Loveland (DPLL) procedure for Boolean satisfiability peaks.

Many fundamental algorithmic problems are optimization problems and not mere decision problems. In particular, if a Boolean formula happens to be over-constrained and, thus, unsatisfiable, it is natural to ask for the maximum number of clauses in the formula that can be simultaneously satisfied. By focusing on 3CNF-formulas, we obtain the optimization problem MAX 3SAT: given a 3CNF-formula φ, find the maximum number of clauses of φ that can be simultaneously satisfied. MAX 3SAT is a prototypical constraint optimization problem that is known to play a prominent role in the study of the approximability properties of NP-optimization problems. Indeed, as shown by Papadimitriou and Yannakakis [Papadimitriou and Yannakakis, 1991], MAX SAT is a MAX SNP-complete problem; this means that MAX SAT is a constant-approximable optimization problem and that every NP-optimization problem in the class MAX SNP can be reduced to it via an approximation-preserving polynomial-time reduction (see also [Papadimitriou, 1994]).

Its importance in complexity and approximability notwithstanding, MAX 3SAT had not been investigated from the perspective of phase transitions until recently. Zhang [Zhang, 2001] was the first to investigate phase transitions for MAX 3SAT by carrying out an initial set of experiments for the family of *bounded satisfiability* problems 3SAT(6), where *b* is a non-negative integer. An instance of 3SAT*(b)* is a 3CNF-formula cp and the question is: does there exist a truth assignment that violates no more than *b* clauses of φ? Equivalently, 3SAT(6) asks whether the optimum value of MAX SAT on an instance φ is at least *m* - B, where *m* is the number of clauses of φ. Thus, each 3SAT(b) is a decision problem obtained from MAX SAT by imposing a "quality bound" on the optimum value. The control parameter of each 3SAT(6) is the ratio of clauses to variables, that is, the same control parameter as the one used for 3SAT.

Zhang [Zhang, 2001] explored phase transitions for the bounded satisfiability problems 3SAT(6) by running experiments on random 3CNF-formulas with n = 25 variables and for b = 5,10,15,20. His findings suggest that there is a series of separated phase transitions corresponding to different quality bounds. The location of each phase transition appears to increase with *b* (see Figure 7); moreover, the average cost for solving the optimization problem MAX SAT appears to be the envelope of the peak average cost for solving the decision problems 3SAT(b), as *b* increases. Zhang [Zhang, 2001] did not report on the asymptotic behavior of 3SAT(b) as the number of variables increases, because he ran experiments for just a single value (n = 25) of the number of variables. He posed the analysis of phase transitions for 3SAT(b) as an open problem and, in particular, raised the question of finding the exact location of the phase transitions for these decision problems.

In this paper, we report on a systematic investigation of phase transitions for the family of bounded satisfiability problems 3SAT(6). Our investigation has produced both analytical and experimental results that yield a more complete and, at the same time, rather surprising picture of these phase transitions. As stated above, Zhang's [Zhang, 2001] initial experiments suggest that the location of the phase transition for each 3SAT(b) increases with *b.* Here, using the first-moment method, we show analytically that the phase transitions for 3SAT(b) must occur (if they occur at all) within a rather narrow region, regardless of the value of *b.* Moreover, the same behavior is exhibited by the families *kSAT(b)* of decision problems underlying the optimization problems MAX kSAT, $k \geq 3$. In particular, the locations of the phase transitions for 3SAT(b) are bounded from above by $1/(3 - \lg(7)) \approx 5.19$ and from below by the greatest known lower bound 3.42 for the location of the phase transition for 3SAT. At the experimental front, we investigated the asymptotic behavior of 3SAT(b) for the values *b* = 3, 4 and 5 by running experiments for several different values of the number *n* of variables of random 3CNF-formulas (n = 10, 15, 20, 25, 30, 35). Our most striking experimental finding is the discovery that, as *n* increases, the phase transition for each 3SAT(b) emerges in a manner that is qualitatively different from that for 3SAT, as we describe next.

In the case of 3SAT, the experiments in [Selman *et a*/., 1996] revealed that the family of curves for the probability of satisfiability of a random 3CNF-formula (one curve for each number *n* of variables) has the property that the curves become progressively steeper and every two of them intersect at a single point whose abscissa is near the location of the phase transition for 3SAT (see Figure 1). In contrast, our experiments reveal that, for each bounded satisfiability problem 3SAT(b), the family of curves for the probability of a "yes" answer of a random 3CNF-formula (one curve for each number *n* of variables) become progressively steeper but they do *not* intersect; instead, they are separated by a distance that is getting smaller as *n* increases. Figures 2, 3, and 4 depict these findings for 3SAT(3), 3SAT(4) and 3SAT(5), respectively. This is a qualitatively different pattern of emergence of a phase transition that does not seem to have been encountered in earlier investigations of phase transitions of other NP-complete problems in which the probability curves appear to have a quasi-common intersection point.

2 Bounded Satisfiability Problems

For every positive integer $k \geq 3$, MAX kSAT is the following optimization problem: given a kCNF-formula φ, find the maximum number of simultaneously satisfied clauses of φ. The study of phase transitions for an optimization problem begins by focusing on one or more decision problems that underlie that optimization problem. The most natural way to derive a decision problem from an optimization problem is to consider as input both an instance of the optimization problem and an arbitrary quality bound, and to ask whether the optimum value on this instance is bigger (or smaller) than the given quality bound. In the case of MAX kSAT, this gives rise to the following NP-complete decision problem: given a kCNF-formula φ and an integer c, is there a truth assignment to the variables of φ that satisfies at least c clauses of φ? It is also possible to derive a family of decision problems that underlie MAX kSAT and are such that the inputs to each of them are kCNF-formulas only. For each non-negative integer 6, let *kS* AT(b) be the following *bounded satisfiability* problem: given a kCNF-formula φ, is there a truth assignment to the variables of φ that violates no more than *b* clauses of

φ? Clearly, kSAT(0) is kSAT itself; moreover, it is easy to show that every bounded satisfiability problem kSAT(b), where $k \geq 3$ and $b \geq 1$, is NP-complete. Since the input to kSAT(b) is a kCNF-formula, phase transitions for the family kSAT(b) can be studied using the ratio of clauses to variables as the control parameter, that is, the same control parameter as the one used in the study of phase transitions for kSAT.

There have been earlier investigations of phase transitions for other fundamental optimization problems, including the TRAVELING SALESMAN PROBLEM [Gent and Walsh, 1996b] and NUMBER PARTITIONING [Gent and Walsh, 1996a]. It should be noted, however, that for each of these two optimization problems a single decision problem was considered (instead of a family of decision problems) because the quality bound was part of the input to the decision problem. Moreover, the control parameter for these problems was chosen in such a way that it encoded the quality bound in some direct or indirect way, and this choice was criticized in subsequent investigations [Slaney and Thiebaux, 1998]. Clearly, this criticism does not apply to the bounded satisfiability problems kSAT(b).

For every integers $k \geq 3$, $n \geq 1$, and $b \geq 0$, let $\Pr_k[n,r,b]$ be the probability that a random formula φ in the space $F_k(n,r)$ is a "yes" instance of kSAT(b), that is, there is a truth assignment that violates no more than b clauses of φ. We now have all the notation in place to formulate the following conjecture concerning phase transitions for the bounded satisfiability problems kSAT(b).

Conjecture 2.1: For every integer $k \geq 3$ and every integer $b \geq 0$, there is a positive real number $r_{k,b}$ such that

- If $r < r_{k,b}$, then $\lim_{n\to\infty} \Pr_k[n,r,b] = 1$;
- If $r > r_{k,b}$, then $\lim_{n\to\infty} \Pr_k[n,r,b] = 0$.

Note that Chvatal and Reed's conjecture [Chvatal and Reed, 1992] concerning phase transitions for kSAT, $k \geq 3$, is the special case of Conjecture 2.1 in which $b = 0$. As mentioned in Section 1, Zhang [Zhang, 2001] raised the problem of finding the exact location of the phase transition for 3SAT(b), which, in terms of the notation introduced here, amounts to first establishing Conjecture 2.1 for $k = 3$ and then determining the critical ratio $r_{3,b}$, for each b. Although far from solving these problems, the next result yields analytical upper bounds for the values of $r_{k,b}$; in particular, it demonstrates that for each $k \geq 3$, all critical ratios $r_{k,b}$ are bounded by a quantity that depends only on k.

Proposition 2.2: *Fix two integers $k \geq 3$ and $b \geq 0$.*

- *If $r > \frac{1}{k-\lg(2^k-1)}$, then $\lim_{n\to\infty} \Pr_k[n,r,b] = 0$.*
- *Consequently, if $r_{k,b}$ exists, then $r_{k,b} \leq \frac{1}{k-\lg(2^k-1)}$, regardless of the value of b.*

Proof: If X is a random variable taking non-negative values and having finite expectation $E(X)$, then Markov's inequality asserts that $\Pr[X \geq 1] \leq E(X)$. This inequality has been used in the past to obtain a coarse upper bound for the critical ratio r_k for the phase transition of kSAT. Here, we will apply the same method to obtain upper bounds for the critical ratios $r_{k,b}$ for the phase transitions of kSAT(b). In this case, however, bounding the expectation of a suitable random variable X is not as trivial as it was for kSAT, but still quite straightforward.

Let $X_k^{n,r,b}$ be a random variable on $F_k(n,r)$ such that if φ is a formula in $F_k(n,r)$, then $X_k^{n,r,b}(\varphi)$ is the number of truth assignments on n variables that violate no more than b clauses of φ. Clearly, $\Pr_k[n,r,b] = \Pr[X_k^{n,r,b} \geq 1]$, thus our goal now is to bound the expectation $E(X_k^{n,r,b})$.

For every truth assignment α on n variables, let I_α^b be the Bernoulli random variable on $F_k(n,r)$ such that $I_\alpha^b(\varphi) = 1$, if α violates no more than b clauses of φ, and $I_\alpha^b(\varphi) = 0$, otherwise. It is easy to see that the number of formulas in $F_k(n,r)$ for which α violates no more than b clauses is $\sum_{i=0}^{b} \binom{rn}{i}\binom{n}{k}^i \left((2^k-1)\binom{n}{k}\right)^{rn-i}$, which implies that

$$E(I_\alpha^b) = \frac{\sum_{i=0}^{b}\binom{rn}{i}\binom{n}{k}^i (2^k-1)^{rn-i}\binom{n}{k}^{rn-i}}{2^{krn}\binom{n}{k}^{rn}}.$$

Since $X_k^{n,r,b} = \Sigma_\alpha I_\alpha^b$, the linearity of expectation implies that $E(X_k^{n,r,b}) = \Sigma_\alpha E(I_\alpha^b) = 2^n E(I_\alpha^b)$. Thus,

$$\begin{aligned} E(X_k^{n,r,b}) &= 2^n \left(\frac{\sum_{i=0}^{b}\binom{rn}{i}\binom{n}{k}^i (2^k-1)^{rn-i}\binom{n}{k}^{rn-i}}{2^{krn}\binom{n}{k}^{rn}}\right) \\ &= \left(2\left(\frac{2^k-1}{2^k}\right)^r\right)^n \sum_{i=0}^{b}\frac{\binom{rn}{i}}{(2^k-1)^i} \end{aligned}$$

Consequently,

$$\lim_{n\to\infty} E(X_k^{n,r,b}) \leq \lim_{n\to\infty} \frac{O(n^b)}{\left(\frac{1}{2}\left(\frac{2^k}{2^k-1}\right)^r\right)^n}.$$

Thus, if $r > \frac{1}{k-\lg(2^k-1)}$, then $\lim_{n\to\infty} E(X_k^{n,r,b}) = 0$, regardless of the value of b. Now an application of Markov's inequality to $X_k^{n,r,b}$ gives the result. ∎

Concerning lower bounds for $r_{k,b}$, it is obvious that if r is a ratio such that almost all kCNF-formulas are satisfiable, then $\lim_{n\to\infty} \Pr_k[n,r,b] = 1$, regardless of the value of b. Consequently, if r_k is the critical ratio for kSAT, then $r_k \leq r_{k,b}$, regardless of the value of b. Of course, this assumes that the critical ratios r_k and $r_{k,b}$ actually exist. Until this is established, we can use the analytically derived lower bounds for r_k as lower bounds for every $r_{k,b}$. In particular, the best currently known such lower bound for r_3 is 3.42 [Kaporis *et al*, 2002]. By combining these remarks with Proposition 2.2, we obtain the following bounds for the phase transitions of the bounded satisfiability problems 3SAT(b).

Corollary 2.3: *Let $b \geq 0$ be an integer.*

- *If $r < 3.42$, then $\lim_{n\to\infty} \Pr_3[n,r,b] = 1$*
- *If $r > \frac{1}{3-\lg(7)} \approx 5.19$, then $\lim_{n\to\infty} \Pr_3[n,r,b] = 0$.*

Consequently, if $r_{3,b}$ exists, then $3.42 \leq r_{3,b} \leq 5.19$, regardless of the value of b.

3 Experimental Results

We ran experiments for random 3CNF-formulas drawn from *n* = 10, 15, 20, 25, 30 and 35 variable spaces. We implemented a suitably modified version of the Davis-Logemann-Loveland-Putnam (DPLL) procedure and tested each random formula drawn to determine whether it is a "yes" instance of the bounded satisfiability problems 3SAT(b), for *b* = 3, 4 and 5. For each (n, b)-pair, probability curves were determined by recording the average number of "yes" instances in samples of 1200 formulas, one sample for each possible ratio of clauses to variables up to a ratio of 15.

The probability curves for 3SAT(6) with *b* = 3, 4 and 5 are depicted in Figures 2, 3 and 4, respectively.

In these figures, we see that, as *r* increases, the probability curves start out in a region where the probability is 1 or close to 1 and then they transition to a region where the probability is 0 or close to 0. To make the terminology more precise, let us define the *location* of the transition for the probability of 3S AT(b) to be the ratio *r* at which the probability is 0.5, and the *width* of the transition for the probability of 3SAT(6) to be the difference in the r values between the ratios at which the probability falls from 0.9 to 0.1. As *n* increases, the probability curves move from right to left; moreover, both the location and the width of the transition appear to change with different values of *n.* For each fixed 6, the width becomes smaller with increasing *n* just as is the case for the probability curves for 3SAT in Figure 1; the location, however, moves dramatically to the left in sharp contrast with the behavior for 3SAT, where the location moves very slightly and converges to a value in the transition region where all the curves appear to intersect at one point. The magnitude of the leftward movement of the transition location with increasing *n* appears to become greater with increasing values of *b.* Thus, the phase transitions for the bounded satisfiability problems 3SAT(6) emerge in a pattern that is novel and qualitatively different from that of the phase transitions for other NP-complete problems.

We also recorded the average cost of the modified DPLL-procedure for solving 3SAT(6), where *b* = 3, 4, 5. Each of these problems is much harder on average than 3SAT; moreover, these problems are getting harder on average as *b* increases. As a matter of fact, the experiments for *b* = 5 and for n = 35 took several weeks to complete. Figure 5 depicts the average cost of solving 3SAT, while Figures 6 depicts the average cost of solving 3S AT(5).

In Figure 6, the performance curves for 3SAT(5) move from lower to higher with increasing *n* and, as has been observed with 3SAT, the peaks in cost appear to correspond to the location of the transitions. The peaks move from right to left in synchrony with the movement of the location of the transition for 3SAT(6); this movement is more dramatic than in the case of 3SAT.

4 Finite Size Scaling

The locations of the transition in the probability curves for 3SAT(6) appear to approach a limiting critical ratio r 3,b that is within the analytically derived upper and lower bounds in Corollary 2.3. Nonetheless, it is not clear what the critical ratio r3,b actually is, nor can it be estimated by visual inspec-

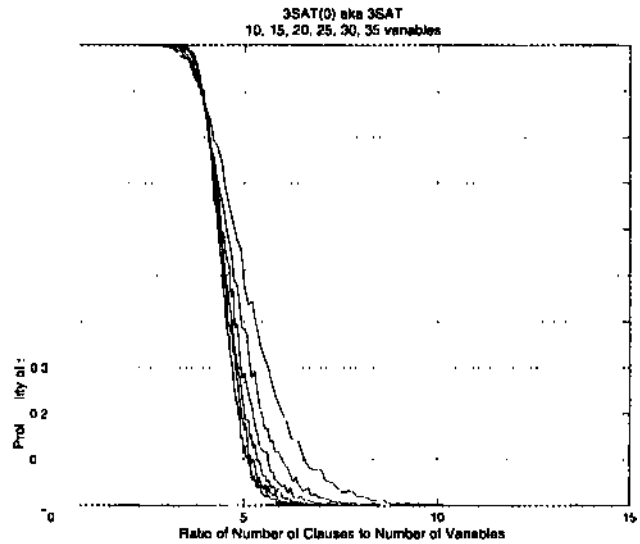

Figure 1: The probability curves for bounded satisfiability with *b* — 0. There is a curve for each of six different values of n. They show the well known 3SAT behavior. As *n* increases the transition regions of the curves become steeper and appear to have a common intersection point about which they rotate. The common hypothesis is that in the limit the curves approach a step function through this point and its ratio is the critical phase transition value for 3SAT, approximately 4.2.

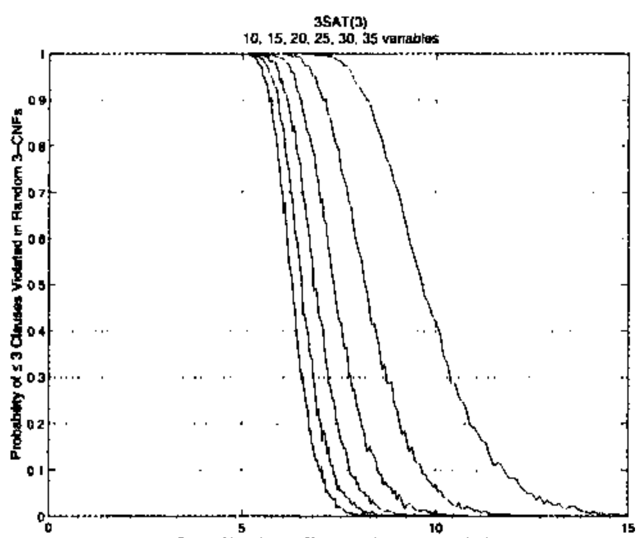

Figure 2: These arc the probability curves for 3S AT(3). There is a curve for each of the same *n* values as in Figure 1. As 71 increases, the transition regions appear to steepen and distinctly move to the left. The steepening is similar to 3SAT behavior but here there is no apparent common rotation point. The transitions appear to be approaching a step function in the limit but it is not possible to visually determine the critical ratio. We used a form of finite size scaling to estimate it as approximately 4.4

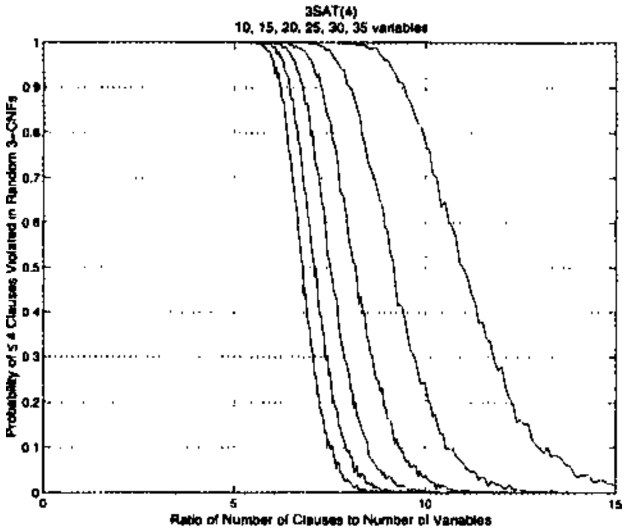

Figure 3: These are the probability curves for 3SAT(4). There is a curve for each of the same *n* values as in the previous figures. Again, as *n* increases, the transition regions appear to steepen and distinctly move to the left. The steepening is similar to 3SAT behavior but here there is no apparent common rotation point. The transitions appear to be approaching a step function in the limit but it is not possible to visually determine the critical ratio. We used a form of finite size scaling to estimate it as approximately 4.5

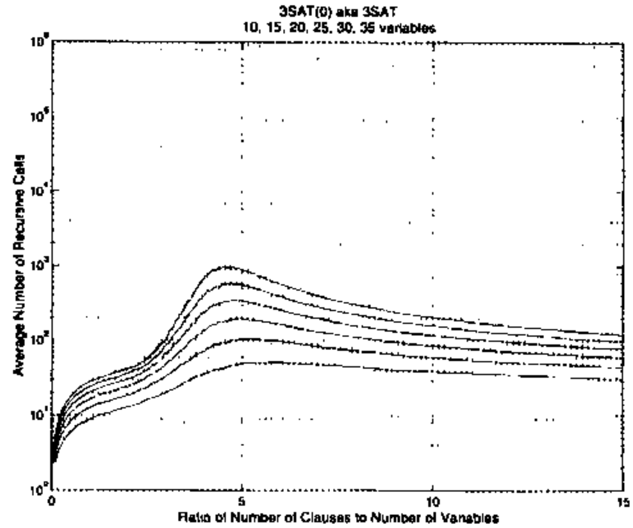

Figure 5: This set of curves plots the well known performance behavior of the DPLL procedure for 3SAT. This figure is here for comparison with the following figure for 3SAT(5).

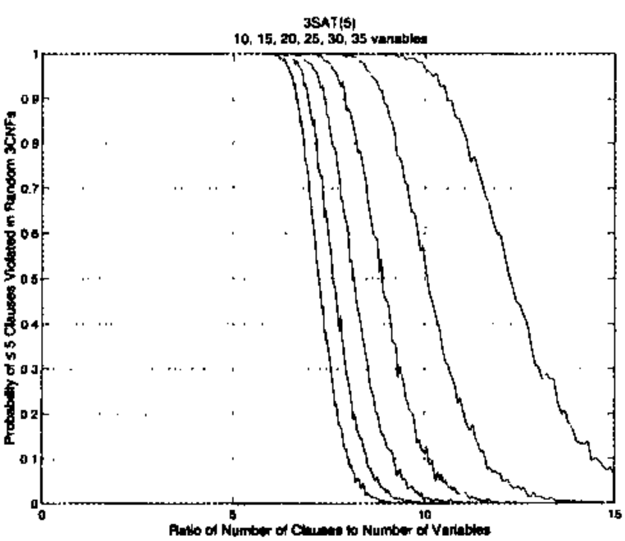

Figure 4: These are the probability curves for 3SAT(5). There is a curve for each of the same n values as in the previous figures. Again, as *n* increases, the transition regions appear to steepen and distinctly move to the left. The steepening is similar to 3SAT behavior but here there is no apparent common rotation point. The transitions appear to be approaching a step function in the limit but it is not possible to visually determine the critical ratio. We used a form of finite size scaling to estimate it as approximately 4.8. Note the analytical upper bound is approximately 5.19.

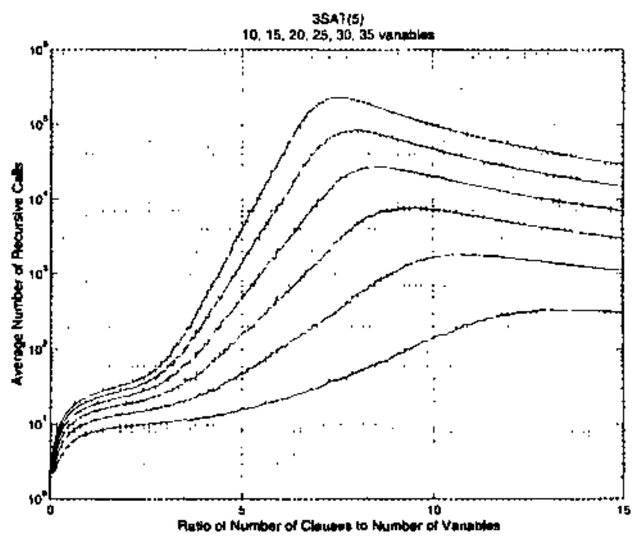

Figure 6: This set of curves plots the algorithmic costs for 3SAT(5). The behavior is similar to 3SAT in that cost increases with increasing n; moreover, cost exhibits the so-called "easy/hard/easy" pattern as the ratio increases. As would be expected the costs for bounded satisfiability arc much higher than for satisfiability. Note that the movement of the peak locations, as n increases, is quite marked and corresponds to the leftward movement of the transition regions in the probability curves.

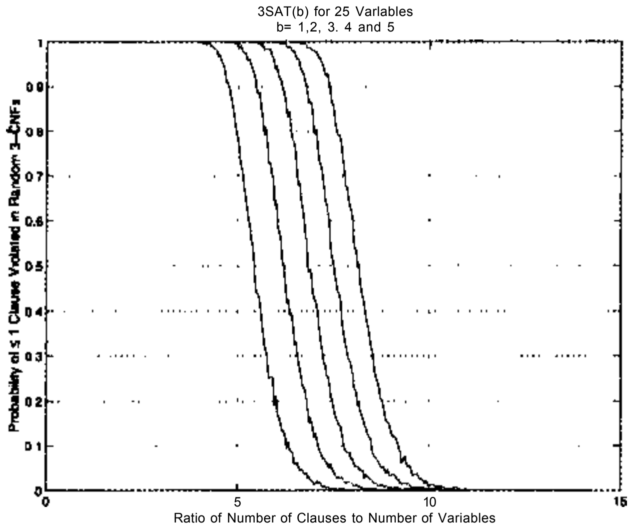

Figure 7: This figure shows the apparently separate "phase transitions" for different values of *b* when looking at probability curves for a fixed n, in this case 25 variables. The transition region moves from left to right with increasing *b.* This type of behavior was previously reported for 6 = 5,10,15, and 20 [Zhang, 2001].

tion since no two probability curves for 3S AT(b) intersect. A finite-size scaling analysis of the data, assuming a power law of the form $\frac{(r-r_{3,b})n^{\nu_{3,b}}}{r_{3,b}}$, allowed us to obtain estimates for both the critical ratio $r_{3,b}$ and for the exponent $\nu_{3,b}$, where *b* = 3,4 and 5.

In [Kirkpatrick and Selman, 1994], an estimate for the value of the critical ratio r_K was obtained by visually estimating the point at which the probability curves for *kSAT* appear to intersect, and then regression techniques were applied to determine the best exponent V k in the power law for kSAT. This cannot be done with the bounded satisfiability problems 3SAT(b), since the probability curves do not intersect, but appear to be moving to the left as the number *n* of variables increases. Nevertheless the finite-size scaling hypothesis can still be tested with a more elaborate procedure. Our goal was to test whether it is possible find values for r3,b and v3,b such that when the ratio *r* is rescaled with the above power law, the probability curves for the different values *n* of the number of variables collapse to a single curve. Using routines for interpolation and regression from Matlab, we stepped through small increments of putative r 3,b values and measured how well the curves collapsed to a single curve by calculating the sum of their pairwise squared differences when accordingly transformed. The estimated r3,b was taken to be the value that minimized this difference measure. We found this technique to be well behaved, in that it exhibited a clear minimum and gave much more precision than visually judging how well the curves appeared to collapse to a single curve. To further validate this technique, we checked it on the phase transition of 2SAT, which has been analytically determined to occur at the critical ratio r_2 = 1 [Chvdtal and Reed, 1992; Fernandez de la Vega, 1992; Goerdt, 1996]. The agreement was very good, namely, the estimate for r_2 was 0.98.

Table 1: Finite-size Scaling Results

b	$r_{3,b}$	$V_{3,b}$
3	4.40+	0.64+
4	4.48+	0.68+
5	4.84	0.77

The results of the finite-size scaling analysis for 3SAT(b), where *b* = 3, 4 and 5, are shown in Table 1. Note that these results are consistent with the analytical upper and lower bounds for r 3,b in Corollary 2.3, that the r b,k are in the region between r k for k-SAT and the Markov upper bound for k-SAT. This consistency is the most that can be reasonably concluded from the finite-size scaling analysis. The technique provides a more principled way to make the estimates than eyeballing, but no great accuracy should be attributed to the estimates in spite of their apparent precision. Although the finite-size scaling results suggest that the locations of the phase transitions of 3SAT(b) are different for different 6's, it may still be the case that they all coincide with the location for 3SAT.

5 Concluding Remarks

The results reported here advance the understanding of the phase transitions for the family of bounded satisfiability problems 3SAT(b), introduced by Zhang [Zhang, 2001 J.

Our main analytical finding is that the phase transitions of all 3SAT(6) problems must occur within a narrow region, regardless of how large the value of *b* is. Moreover, our experiments revealed that the phase transitions for these problems occur in a remarkable way. Specifically, unlike 3SAT, the probability curves for 3SAT(6) do not have a quasi-common intersection point about which they rotate as they become steeper with increasing n. Instead, they move rapidly to the left toward the narrow region that the analysis predicts.

Identifying the exact locations of the phase transitions for 3SAT*(b)* remains an open problem, which is at least as hard as identifying the location of the phase transition for 3SAT. Equally hard appears to be the problem of analytically establishing whether these locations are separated or coincide. Nonetheless, it may be possible to analytically obtain tighter bounds for these locations using some of the sophisticated techniques that have been quite successful in finding tighter bounds for the phase transition of 3SAT.

Acknowledgments: We are grateful to Victor Dalmau for stimulating and insightful discussions. We also wish to thank the reviewers for their many useful comments and suggestions.

References

[Chvatal and Reed, 1992] V. Chvatal and B. Reed. Mick gets some (the odds are on his side). In *33th Annual Symposium on Foundations of Computer Science,* pages 620-627, 1992.

[Dubois *et al.,* 2000] O. Dubois, Y. Boufkhad, and J. Mandler. Typical random 3-SAT formulae and the satisfiability threshold. In

Proc. Uth Annual ACM-SIAM Symp. on Discrete Algorithms, pages 126-127, 2000.

[Fernandez de la Vega, 1992] W. Fernandez de la Vega. On random 2-SAT. Manuscript, 1992.

[Franco and Paul, 1983] J. V. Franco and M. Paul. Probabilistic analysis of the Davis-Putnam Procedure for solving the satisfiability problem". *Discrete Applied Math.,* 5:77-87, 1983.

[Gent and Walsh, 1996a] Ian Gent and Toby Walsh. Phase transitions and annealed theories: number partitioning. *Proc. of 12th European Conf. on Artificial Intelligence, ECAI-96,* pages 170-174, 1996.

[Gent and Walsh, 1996b] Ian Gent and Toby Walsh. The TSP phase transition. *Artificial Intelligence,* 88(l-2):349-358, 1996.

[Goerdt, 1996] A. Goerdt. A threshold for unsatisfiability. *J. Comput. System Sci.,* 53(3):469-486, 1996.

[Kaporis et al.,2002] A. C. Kaporis, L. M. Kirousis, and E. G. Lalas. The probabilistic analysis of a greedy satisfiability algorithm. In R. Mohring and R. Ramen, editors, *Algorithms-ESA2002, Proceedings of the Tenth Annual European Symposium, LNCS 2461,* pages 574-585. Springer-Vcrlag, 2002.'

[Kirkpatrick and Selman, 1994] S. Kirkpatrick and B. Sclman. Critical behavior in the satisfiability of random formulas. *Science,* 264:1297-1301, 1994.

[Papadimitriou and Yannakakis, 1991] C. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Science,* 43:425-440, 1991.

[Papadimitriou, 1994] C. H. Papadimitriou. *Computational complexity.* Addison-Wesley, 1994.

[Selman *et al,* 1996] B. Sclman, D. G. Mitchell, and H. J. Levesquc. Generating hard satisfiability problems. *Artificial Intelligence,* 81 (1 -2): 17-29, 1996.

[Slancy and Thicbaux, 1998] John Slaney and Sylvic Thiebaux. On the hardness of decision and optimisation problems. In Henri Pradc, editor, *13th European Conference on Artificial Intelligence, ECAI98,* pages 224-248. John Wiley & Sons, Ltd., 1998.

[Zhang, 2001] Weixiong Zhang. Phase transitions and backbones of 3-SAT and maximum 3-SAT, In T. Walsh, editor, *Principles and Practice of Constraint Programming - CP 2001,* volume 2239 of *LNCS,* pages 153-167. Springer-Vcrlag, 2001.