Extending Deep Structure

Colin P. Williams and Tad Hogg

Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304, U.S.A.
CWilliams@parc.xerox.com, Hogg@parc.xerox.com

Abstract

In a previous paper we defined the "deep structure" of a constraint satisfaction problem to be that set system produced by collecting the nogood ground instances of each constraint and keeping only those that are not supersets of any other. We then showed how to use such deep structure to predict where, in a space of problem instances, an abrupt transition in computational cost is to be expected. This paper explains how to augment this model with enough extra details to make more accurate estimates of the location of these phase transitions. We also show that the phase transition phenomenon exists for a much wider class of search algorithms than had hitherto been thought and explain theoretically why this is the case.

1. Introduction

In a previous paper (Williams & Hogg 1992b) we defined the "deep structure" of a constraint satisfaction problem (CSP) to be that set system produced by collecting the nogood ground instances of each constraint and keeping only those that are not supersets of any other. We use the term "deep" because two problems that are superficially different in the constraint graph representation might in fact induce identical sets of minimized nogoods. Hence their equivalence might only become apparent at this lower level of representation. We then showed how to use such deep structure to predict where, in a space of problem instances, the hardest problems are to be found. Typically, this model led to predictions that were within about 15% of the empirically determined correct values. Whilst this model allowed us to understand the observed abrupt change in difficulty (in fact a phase transition) in general terms in this paper we identify which additional aspects of real problems account for most of the outstanding numerical discrepancy. This is particularly important because as larger problems are considered, the phase transition region becomes increasingly spiked. Hence, an acceptable error for small problems could become unacceptable for larger ones.

To this end we have identified 2 types of error; modelling approximations (such as the assumption that the values assigned to different variables are uncorrelated or that the solutions are not clustered in some special way) and mathematical approximations (such as the assumption that, for a function $f(x)$, $\langle f(x) \rangle \approx f(\langle x \rangle)$, known as a mean-field approximation). In addition we also widen the domain of applicability of our theory to CSPs solved using algorithms such as heuristic repair (Minton et al. 1990), simulated annealing (Johnston et al. 1991, Kirkpatrick 1983) and GSAT (Selman, Levesque & Mitchell 1992), that work with sets of complete assignments at all times.

In the next section we summarize our basic deep structure theory. Following this, we shall show how to make quantitatively accurate predictions of the phase transition points for 3–COL, 4–COL (graph colouring) and 3–SAT. Our results are summarized in Table 1 where the best approximations are highlighted. Finally in Section 4 we present experimental evidence for a phase transition effect in heuristic repair and adapt our deep structure theory to account for these observations.

2. Basic Deep Structure Model

Our interest lies in predicting where, in a space of CSP instances, the harder problems typically occur, more or less regardless of the algorithm used. Because the exact difficulty of solving each instance can vary considerably from case to case it makes more sense to talk about the average difficulty of solving CSPs that are drawn from some pool (or ensemble) of similar problems. This means we need to know something about how the difficulty of solving CSPs changes as small modifications are made to the structure of the constraint problem.

There are many possible types of ensemble that one could choose to study. For example, one might restrict consideration to an ensemble of problems each of whose member instances are guaranteed to have at least one solution. Alternatively, one could study an ensemble in which this requirement is relaxed and each instance may or may not have any solutions. Similarly one could choose whether the domain sizes of each variable should or should not be the same or whether the constraints are all of the same size etc. The possibilities are endless. The best choice of ensemble cannot be determined by mere

cogitation but depends on what the CSPs arising in the "real world" happen to be like and that will inevitably vary from field to field. Lacking any compelling reason to choose one ensemble over another, we made the simplest choice of using an ensemble of CSPs whose instances are not guaranteed to be soluble and having variables with a uniform domain size, b.

Given an ensemble of CSPs, then, the deep structure model allows us to predict which members will typically be harder to solve than others. The steps required to do this can be broken down into:

1. CSP → Deep Structure
2. Deep Structure → Estimate of Difficulty

CSP → Deep Structure

The first step consists of mapping a given CSP instance into its corresponding deep structure. We chose to think of CSPs that could be represented as a set of constraints over μ variables, each of which can take on one of b values. Each constraint determines whether a particular combination of assignments of values to the variables are consistent ("good") or inconsistent ("nogood"). Collecting the nogoods of all the constraints and discarding any that are supersets of any other we arrive at a set of "minimized nogoods" which completely characterize the particular CSP. By "deep structure" we mean exactly this set of minimized nogoods.

Unfortunately, reasoning with the explicit sets of minimized nogoods does not promote understanding of generic phenomena or assist theoretical analysis. We therefore attempt to summarize the minimized nogoods with as few parameters as possible and yet still make reasonably accurate quantitative predictions of quantities of interest such as phase transition points and computational costs. As we shall see, such a crude summarization can sometimes throw away important information e.g. regarding the correlation between values assigned to tuples of variables. Nevertheless, it does allow us to identify which parameters have the most important influence on the quantities of interest. Moreover, one is always free to build a more accurate model, as in fact we do in Section 3.

In our basic model, we found that the minimized nogoods could be adequately summarized in terms of their number, m, and average size, k. Thus we crudely characterize a CSP by just 4 numbers, (μ, b, m, k).

Deep Structure → Estimate of Difficulty

Having obtained the crude description of deep structure we need to estimate how hard it would be to solve such a CSP. The actual value of this cost will depend on the particular algorithm used to solve the CSP. In our original model we assumed a search algorithm that works by extending partial solutions (either in a tree or a lattice) until a complete solution is found. However, the important point is not so much the actual value of the cost but in predicting where it will attain a maximum as this corresponds to the point of greatest difficulty. In Section 4 we extend our model to cover the possibility of solving the CSP using an algorithm that works with complete states e.g heuristic repair, simulated annealing or GSAT which requires a different cost measure (still related to the minimized nogoods) to be used.

Parameter	Meaning
μ	number of variables
b	number of values per variable
m	number of minimized nogoods
k	average size of minimized nogoods

Fig. 1. A coarse description of a CSP.

To obtain a definite prediction, we defined "difficulty" to be the cost to find the first solution or to determine there are no solutions, C_s. Analytically, this is a hard function to derive and in the interests of a more tractable analysis we opted to use a proxy instead that was the cost to find all solutions divided by the number of solutions (if there were any) or else the cost to determine there were no solutions, which we approximated as[1]:

$$\langle C_s \rangle \approx \begin{cases} \langle C \rangle / \langle N_{soln} \rangle & \text{if there are solutions} \\ \langle C \rangle & \text{otherwise} \end{cases} \quad (1)$$

We analyzed what happens to this cost, on average, as the number of minimized nogood ground instances, $m = \beta\mu$, is increased. Note that we merely write m like this to emphasize that the number of minimized nogoods will grow as larger problems are considered (i.e. as μ increases). The upshot of this analysis was the prediction that, as $\mu \to \infty$, the transition occurs where $\langle N_{soln} \rangle = 1$ and so the hardest problems are to be found at a critical value of β given by:

$$\beta_{crit} = -\frac{\ln b}{\ln (1 - b^{-k})}. \quad (2)$$

In other words, if all we are told about a class of CSPs is that there are μ variables (with $\mu \gg 1$), each variable takes one of b values and each minimized nogood is of size k then we expect the hardest examples of this class to be when there are $m_{crit} = \beta_{crit}\mu$ nogoods.

[1] N.B. this approximation will fail if the solutions are tightly clustered.

3. More Accurate Predictions

We have tested this formula on two kinds of CSPs: graph colouring and k-SAT and compared its predictions against experimental data obtained by independent authors. Typically this formula gave predictions that were within about 15% of the empirically observed values. The remaining discrepancy can be attributed to one of two basic kinds of error: First, there can be errors in the model (e.g. due to assuming that the values assigned to different variables are uncorrelated). Second there can be errors due to various mathematical approximations (e.g. the mean-field approximation that $\langle C/N_{soln}\rangle \simeq \langle C\rangle/\langle N_{soln}\rangle$). Interestingly, graph colouring is more affected by errors in the model whereas k-SAT is more affected by errors in the mean-field approximation. These two CSPs then will serve as convenient examples of how to augment our basic deep structure model with sufficient extra details to permit a more accurate estimation of the phase transition points.

Graph Colouring

A graph colouring problem consists of a graph containing μ nodes (i.e. variables) that have to be assigned certain colours (i.e. values) such that no two nodes at either end of an edge have the same colour. Thus the edges provide implicit constraints between the values assigned to the pair of nodes they connect. Therefore, if we are only allowed to use b colours, then each edge would contribute exactly b nogoods and every nogood would be of size 2, so $k = 2$. Plugging these values into equation 2 gives the prediction that the hardest to colour graphs occur when $\beta_{crit} = 9.3$ (3–COL) and $\beta_{crit} = 21.5$ (4–COL) in contrast to the experimentally measured values of 8.1 ± 0.3 and 18 ± 1 respectively. This approximation isn't too bad, nevertheless, we will now show how to make it even better by taking more careful account of the structure of the nogoods that arise in graph colouring.

Imprecision due to Model

The key insight is to realize that in our derivation of formula 2 we assume the nogoods are selected independently. Thus each set of m nogoods is equally likely. However, in the context of graph colouring this is not the case because each edge introduces nogoods with a rather special structure. Specifically, each edge between nodes u and v introduces b minimal nogoods of the form $\{u = i,\ v = i\}$ for i from 1 to b, which changes, for a given number of minimized nogoods, the expected number of solutions, as follows.

Consider a state at the solution level, i.e., an assigned value for each of μ variables, in which the value i is used c_i times, with $\sum_{i=1}^{b} c_i = \mu$. In order for this state to be a solution, none of its subsets must be among the selected nogoods. This requires that the graph not contain an edge between any variables with the same assignment. This excludes a total of $\sum_{i=1}^{b} \binom{c_i}{2}$ edges. With random graphs with e edges, the probability that this given state will be a solution is just

$$p(\{c_i\}) = \frac{\binom{\binom{\mu}{2} - \sum_{i=1}^{b}\binom{c_i}{2}}{e}}{\binom{\binom{\mu}{2}}{e}} \tag{3}$$

By summing over all states at the solution level, this gives the expected number of solutions:

$$\langle N_{soln}\rangle = \sum_{c_1 \ldots c_b} \binom{\mu}{c_1 \ldots c_b} p(\{c_i\}) \tag{4}$$

where the multinomial coefficient counts the number of states with specified numbers of assigned values.

For the asymptotic behaviour, note that the multinomial becomes sharply peaked around states with an equal number of each value, i.e., $c_i = \mu/b$. This also minimizes the number of excluded edges $\sum \binom{c_i}{2}$ giving a maximum in $p(\{c_i\})$ as well. Thus the sum for $\langle N_{soln}\rangle$ will be dominated by these states and Stirling's approximation[2] can be used to give

$$\ln \langle N_{soln}\rangle \sim \mu\left[\ln b + \frac{\beta}{b}\ln\left(1 - \frac{1}{b}\right)\right] \tag{5}$$

because the number of minimal nogoods is related to the number of edges by $m = \beta\mu = eb$.

With this replacement for $\ln\langle N_{soln}\rangle$ our derivation of the phase transition point proceeds as before by determining the point where the leading term of this asymptotic behaviour is zero, corresponding to $\langle N_{soln}\rangle = 1$, hence:

$$\beta_{crit} = -\frac{b \ln b}{\ln\left(1 - \frac{1}{b}\right)} \tag{6}$$

which is different from the prediction of our basic model as given in equation 2. This result can also be obtained more directly by assuming conditional independence among the nogoods introduced by each edge (Cheeseman, Kanefsky & Taylor 1992). For the cases of 3 and 4–colouring, equation 6 now allows us to predict $\beta_{crit} = 8.1$ and 19.3, respectively, close to the empirical values given by Cheeseman et al.

[2] i.e. $\ln x! \sim x \ln x - x$ as $x \to \infty$.

k-SAT

Empirical studies by Mitchell, Selman & Levesque (Mitchell, Selman & Levesque 1992) on the cost of solving k-SAT problems using the Davis-Putnam procedure (Franco & Paull 1983), allow us to compare the predictions of our basic model against a second type of CSP.

In k-SAT, each of the μ variables appearing in the formula can take on one of two values, *true* or *false*. Thus there are $b = 2$ values for each variable. Each clause appearing in the given formula is a disjunction of (possibly negated) variables. Hence the clause will fail to be true for exactly one assignment of values to the k variables appearing in it. This in turn gives rise to a single nogood, of size k. Distinct clauses will give rise to distinct nogoods, so the number of these nogoods is just the number of distinct clauses in the formula.

Thus, using equation 2, our basic model, with $b = 2$, $k = 3$ predicts the 3–SAT transition to be at $\beta_{crit} = 5.2$ which is above empirically observed value of 4.3. However, as we show below, the outstanding error is largely attributable to the inaccuracy of the mean-field approximation and there is a simple remedy for this.

Imprecision due to Mean-field Approximation

Cheeseman et al. observed that the phase transition for graph colouring occurred at the point when the probability of having at least one solution fell abruptly to zero. In a longer version of this paper (Williams & Hogg 1992a) we explain why this is to be expected. One way of casting this result, which happens to be particularly amenable to mathematical analysis, is to hypothesize that the phase transition in cost should occur when $\left\langle \frac{1}{1+N_{soln}} \right\rangle$ transitions from being near zero to being near 1. In order to estimate this point, we consider the Taylor series approximation (Papoulis 1990, p129):

$$\left\langle \frac{1}{1+N_{soln}} \right\rangle \approx \frac{1}{1+\langle N_{soln}\rangle} + \frac{\mathrm{var}(N_{soln})}{(1+\langle N_{soln}\rangle)^3} \qquad (7)$$

In figure 2 we plot measured values of $\left\langle \frac{1}{1+N_{soln}} \right\rangle$ together with its truncated Taylor series approximation versus β for increasing values of μ. This proxy sharpens to a step function as $\mu \to \infty$ apparently at the same point as that reported by Cheeseman et al. Fortunately although the truncated Taylor series approximation overshoots the true value of $\left\langle \frac{1}{1+N_{soln}} \right\rangle$ before finally returning to a value of 1 at high β, it nevertheless is accurate in the vicinity of the phase transition point as required and may therefore be used. Hence, the true transition point can be estimated as the value of β at which the right hand side of equation 7 equals $\frac{1}{2}$. As the true transition point precedes the old one (predicted using equation 2), i.e. $\beta_{crit}^{true} < \beta_{crit}$ and as

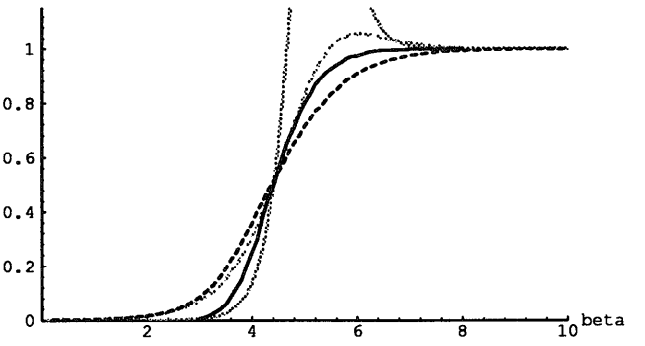

Fig. 2. Behaviour of $\langle 1/(1+N_{soln})\rangle$ vs β for $b = 2$, $k = 3$ (3–SAT). The dark curves show empirical data for $\mu = 10$ (dashed) and $\mu = 20$ (solid). The light curves show the corresponding two-term Taylor series approximation to $\langle 1/(1+N_{soln})\rangle$.

there are exponentially many solutions for $\beta < \beta_{crit}$ the first term in the above approximation must be negligible at the true transition for large enough values of μ. In this case we can estimate β_{crit}^{true} as the value of β at which

$$\mathrm{var}(N_{soln}) = \frac{1}{2}(1+\langle N_{soln}\rangle)^3. \qquad (8)$$

By the same argument as that in (Williams & Hogg 1992b) we can show,

$$\langle N_{soln}\rangle = b^{\mu}\frac{\binom{\binom{\mu}{k}b^k - \binom{\mu}{k}}{\beta\mu}}{\binom{\binom{\mu}{k}b^k}{\beta\mu}} \qquad \mathrm{var}(N_{soln}) = \langle N_{soln}^2\rangle - \langle N_{soln}\rangle^2 \qquad (9)$$

with

$$\langle N_{soln}^2\rangle = b^{\mu}\sum_{r=0}^{\mu}\binom{\mu}{r}(b-1)^{\mu-r}\frac{\binom{\binom{\mu}{k}b^k - 2\binom{\mu}{k}+\binom{r}{k}}{\beta\mu}}{\binom{\binom{\mu}{k}b^k}{\beta\mu}} \qquad (10)$$

The $\langle N_{soln}^2\rangle$ term is obtained by counting how many ways there are of picking $m = \beta\mu$ nogoods of size k such that a given pair of nodes at the solution level are both good and have a prescribed overlap r weighted by the number of ways sets can be picked such that they have this overlap. Finally, this is averaged over all possible overlaps. With these formulae the phase transition can be located as the fixed point solution (in β) to equation 7. For $\mu = 10$ or 20 this gives the transition point at $\beta = 4.4$. Asymptotically, one can obtain an explicit formula for the new critical point by applying Stirling's formula to equations 9 and 10, approximating equation 10 as an integral with a single dominant term and factoring a coefficient as a numerical integral. This gives a slightly higher critical point of

Problem	Expt	Basic Theory(2)	Basic + Correlations(6)	Basic + Corrections(7)
3-col	8.1 ± 0.3	9.3	8.1	N/A
4-col	18 ± 1	21.5	19.5	N/A
3-sat	4.3	5.2	N/A	4.5

Table 1. Comparisons of our basic theory and various refinements thereof with empirical data obtained by other authors. The numbers in the column headings refer to the equations used to calculate that columns' entries.

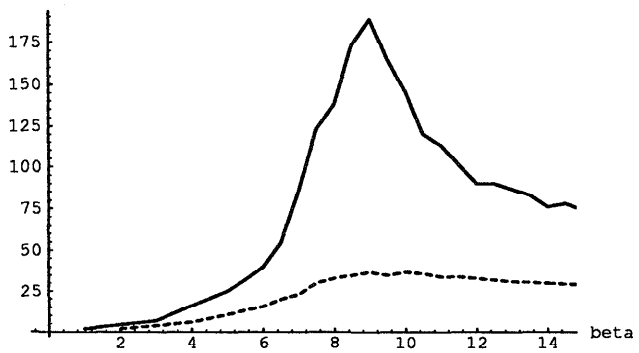

Fig. 3. Median search cost for heuristic repair as a function of β for the case $b = 3$, $k = 2$. The solid curve, for $\mu = 20$, has a maximum at $\beta = 9$. The dashed curve, for $\mu = 10$, has a broad peak in the same region.

$\beta_{crit} = 4.546$ and predicts the new functional form for the critical number of minimized nogoods as $m_{crit} = \beta_{crit}\mu + \text{const} + O\left(\frac{1}{\mu}\right)$ with const $= 3.966$.

The results for our basic model, the correlation model (for graph colouring) and the correction to mean field model (for k-SAT) are collected together in Table 1 where the best results are highlighted.

4. Heuristic Repair Has Phase Transition Too

The above results show that the addition of a few extra details to the basic deep structure model allows us to make quantitatively accurate estimates of the location of phase transition points. However, the question of the applicability of these results to other search methods, in particular those that operate on complete states, such as heuristic repair, simulated annealing and GSAT remains open. In this section we investigate the behaviour of such methods and show theoretically and empirically that they also exhibit a phase transition in search cost at about the same point as the tree based searches.

In figure 3 we plot the median search cost for solving random CSPs with $b = 3$, $k = 2$ versus our order parameter β (the ratio of the number of minimized nogoods to the number of variables) using the heuristic repair algorithm. As for other search algorithms we see a characteristic easy-hard-easy pattern with the peak sharpening as larger problem instances are considered.

To understand this recall that heuristic repair, simulated annealing and GSAT all attempt to improve a complete state through a series of incremental changes. These methods differ on the particular changes allowed and how decisions are made amongst them. In general they all guide the search toward promising regions of the search space by emphasizing local changes that decrease a cost function such as the number of remaining conflicting constraints. In our model, the number of conflicting constraints for a given state is equal to the number of nogoods of which it is a superset. A complete state is minimal when every possible change in value assignment would increase or leave unchanged the number of conflicts.

These heuristics provide useful guidance until a state is reached for which none of the local changes considered give any further reduction in cost. To the extent that many of these local *minimal* or *equilibrium* states are not solutions, they provide points where these search methods can get stuck. In such situations, practical implementations often restart the search from a new initial state, or perform a limited number of local changes that leave the cost unchanged in the hope of finding a better state before restarting. Thus the search cost for difficult problems will be dominated by the number of minimal points, $N_{minimal}$, encountered relative to the number of solutions, N_{soln}. Thus our proxy is:

$$\left\langle \frac{N_{minimal}}{N_{soln}} \right\rangle \approx \frac{\langle N_{minimal} \rangle}{\langle N_{soln} \rangle} \tag{11}$$

with $\langle N_{minimal} \rangle = b^{\mu} p_{minimal}$ where $p_{minimal}$ is the probability that a given state (at the solution level) is minimal. This in turn is just given by the ratio of the number of ways to pick m nogoods such that the given state is minimal to the total number of ways to pick m nogoods. Of course, we should be aware that the mean-field approximation will again introduce some quantitative error.

In figure 4 we plot this cost proxy and the mean-field approximation to it, for $\mu = 10$, $b = 3$, $k = 2$. This predicts that the hardest problems occur around $\beta = 9.5$. Compare this with empirical data, in figure 3. We see that heuristic repair does indeed find certain problems harder than others and the numerical agreement between predicted and observed critical points is quite good, suggesting that $\langle N_{minimal}/N_{soln} \rangle$ is an adequate proxy for the true cost. Thus our deep structure theory applies to sophisticated search methods beyond the tree search algorithms considered previously.

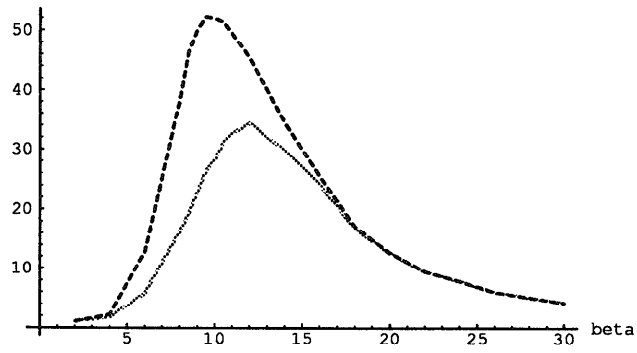

Fig. 4. Ratio of number of minimal points to number of solutions vs. β for the case of $\mu = 10$, $b = 3$, $k = 2$ (dashed curve, with maximum at $\beta = 9.5$) and its mean-field approximation (grey, with maximum at 12).

5. Conclusions

The basic deep structure model (Williams & Hogg 1992) typically led to predictions of phase transition points that were within about 15% of the empirically determined values. However, both empirical observations and theory suggest that the phase transition becomes sharper the larger the problem considered making it important to determine the location of transition points more precisely. To this end, we identified modelling approximations (such as neglecting correlations in the values assigned to different variables) and mathematical approximations (such as the mean field approximation) as the principal factors impeding proper estimation of the phase transition points. We then showed how to incorporate such influences into the model resulting in the predictions reported in Table 1. This shows that the deep structure model is capable of making quantitatively accurate estimates of the location of phase transition points for all the problems we have considered. However, we again reiterate that the more important result is that our model predicts the qualitative existence of the phase transition at all as this shows that fairly simple computational models can shed light on generic computational phenomena. A further advantage of our model is that it is capable of identifying the coarse functional dependencies between problem parameters. This allows actual data to be fitted to credible functional forms from which numerical coefficients can be determined, allowing scaling behaviour to be anticipated.

Our belief that phase transitions are generic is buoyed by the results we report for heuristic repair. This is an entirely different kind of search algorithm than the tree or lattice-like methods considered previously and yet it too exhibits a phase transition at roughly the same place as the tree search methods. We identified the ratio of the number of minimal states to the number of solutions as an adequate cost proxy which can be calculated from the minimized nogoods. Moreover, our experience suggests the exact form for the proxy is not that critical, provided it tracks the actual cost measure faithfully.

6. References

Cheeseman, P.; Kanefsky, B.; and Taylor, W. M. 1991. Where the Really Hard Problems Are. In Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, 331–337, Morgan Kaufmann.

Cheeseman, P.; Kanefsky, B.; and Taylor, W. M. 1992. Computational Complexity and Phase Transitions. In Proc. of the Physics of Computation Workshop, IEEE Computer Society.

Franco and Paull 1983. Probabilistic Analysis of the Davis Putman Procedure for Solving Satisfiability Problems, Discrete Applied Mathematics 5:77–87.

Huberman, B.A. and Hogg, T. 1987. Phase Transitions in Artificial Intelligence Systems, *Artificial Intelligence*, 33:155–171.

Johnson, D., Aragon, C., McGeoch L., Schevon, C., 1991. Optimization by Simulated Aneealing: An experimental evaluation; part ii, graph coloring and number partitioning, *Operations Research*, 39(3):378–406, May-June

Kirkpatrick S., Gelatt C., Vecchi M., 1983. Optimization by Simulated Annealing. Science 220:671–680.

Minton S., Johnston M., Philips A., Laird P. 1990. Solving Large-scale Constraint Satisfaction and Scheduling Problems using a Heuristic Repair Method. In Proc. AAAI-90, pp17–24.

Mitchell D., Selman B., Levesque H., 1992. Hard & Easy Distributions of SAT Problems. In Proceedings of the 10th National Confernce on Artificial Intelligence, AAAI–92, pp459–465, San Jose, CA.

Morris P., 1992. On the Density of Solutions in Equilibrium Points for the N-Queens Problem, In Proceedings of the 10th National Confernce on Artificial Intelligence, AAAI–92, pp428–433, San Jose, CA.

Papoulis A., 1990. Probability & Statistics, p129, Prentice Hall

Selman B., Levesque H., Mitchell D., 1992. A New Method for Solving Hard Satisfiability Problems, In Proceedings of the 10th National Confernce on Artificial Intelligence, AAAI–92, pp440–446, San Jose, CA.

Williams, C. P. and Hogg, T. 1991. Typicality of Phase Transitions in Search, Tech. Rep. SSL-91–04, Xerox Palo Alto Research Center, Palo Alto, California (to appear in *Computational Intelligence* 1993)

Williams, C. P. and Hogg, T. 1992a. Exploiting the Deep Structure of Constraint Problems. Tech. Rep. SSL-92–24, Xerox Palo Alto Research Center, Palo Alto, CA.

Williams, C. P. and Hogg, T. 1992b Using Deep Structure to Locate Hard Problems, in Proc 10th National Conf. on Artificial Intelligence, AAAI–92,pp472–477,San Jose CA.