A Mixture-Model for the Behaviour of SLS Algorithms for SAT

Holger H. Hoos
University of British Columbia
Computer Science Department
2366 Main Mall
Vancouver, BC, V6T 1Z4, Canada
hoos@cs.ubc.ca

Abstract

Stochastic Local Search (SLS) algorithms are amongst the most effective approaches for solving hard and large propositional satisfiability (SAT) problems. Prominent and successful SLS algorithms for SAT, including many members of the WalkSAT and GSAT families of algorithms, tend to show highly regular behaviour when applied to hard SAT instances: The run-time distributions (RTDs) of these algorithms are closely approximated by exponential distributions. The deeper reasons for this regular behaviour are, however, essentially unknown. In this study we show that there are hard problem instances, *e.g.*, from the phase transition region of the widely studied class of Uniform Random 3-SAT instances, for which the RTDs for well-known SLS algorithms such as GWSAT or WalkSAT/SKC deviate substantially from exponential distributions. We investigate these irregular instances and show that the respective RTDs can be modelled using mixtures of exponential distributions. We present evidence that such mixture distributions reflect stagnation behaviour in the search process caused by "traps" in the underlying search spaces. This leads to the formulation of a new model of SLS behaviour as a simple Markov process. This model subsumes and extends earlier characterisations of SLS behaviour and provides plausible explanations for many empirical observations.

Introduction and Background

The propositional satisfiability problem (SAT) is a model combinatorial problem whose conceptual simplicity facilitates the design and analysis of algorithms for other hard combinatorial problems. For the past decade, various types of stochastic local search (SLS) methods have been applied very successfully to SAT. These include the GSAT and WalkSAT families of algorithms (Selman, Kautz, & Cohen 1994; Gent & Walsh 1993; McAllester, Selman, & Kautz 1997), as well as several other algorithms based on similar ideas (Gu 1992; Wah & Shang 1997; Wu & Wah 2000; Schuurmans & Southy 2000; Schuurmans, Southy, & Holte 2001). GSAT and WalkSAT algorithms have been extensively studied in the literature, and include some of the best-performing SAT algorithms known to date (Hoos & Stützle 2000a; Schuurmans, Southy, & Holte 2001). Compared to other state-of-the-art SAT algorithms, such as Satz (Li & Anbulagan 1997), these methods are rather simplistic and it is not well understood how they can solve many classes of large and difficult SAT instances surprisingly efficiently. It is also largely unclear under which conditions (*i.e.*, on which types of instances, and for which parameter settings) these SLS algorithms work well.

The run-time behaviour of GSAT and WalkSAT algorithms when applied to hard SAT instances and when using sufficiently high noise parameter settings, is typically characterised by exponential run-time distributions (RTDs) (Hoos 1998; Hoos & Stützle 1999). Here, "sufficiently high" includes the range in which optimal performance, as reflected in minimal mean run-time, is achieved. These RTD characterisation can be extended to easier SAT instances by using a generalised class of exponential distributions that supports modelling the initial search phase (as reflected in the left tail of a run-time distribution), during which the success probability increases faster than for a memory-less search process characterised by an exponential RTD (Hoos 1998).

As we will show in this study, for a small but significant number of hard instances, *e.g.*, from the widely studied "phase transition region" of the Uniform Random-3-SAT instance distribution (Cheeseman, Kanefsky, & Taylor 1991), SLS algorithms such as GWSAT or WalkSAT/SKC show a behaviour that cannot be captured by these models. This irregular behaviour is interesting for at least two reasons: Firstly, as will become clear later, it can be seen as a type of stagnation behaviour that, if present, appears to severely degrade SLS performance as the search progresses beyond a certain point. Clearly, a sufficient understanding of this phenomenon is likely to be the key towards eliminating the undesirable behaviour. Secondly, the irregularities provide a basis for refining previous models of SLS behaviour; such models are valuable for purely scientific as well as for practical reasons, as they improve our ability to understand, to predict, and to improve the performance and behaviour of SLS algorithms for SAT.

In the following, we investigate this irregular SLS behaviour in detail, focussing on GWSAT and WalkSAT/SKC, two of the most widely studied SLS algorithms for SAT, and the prominent class of Uniform Random-3-SAT "phase tran-

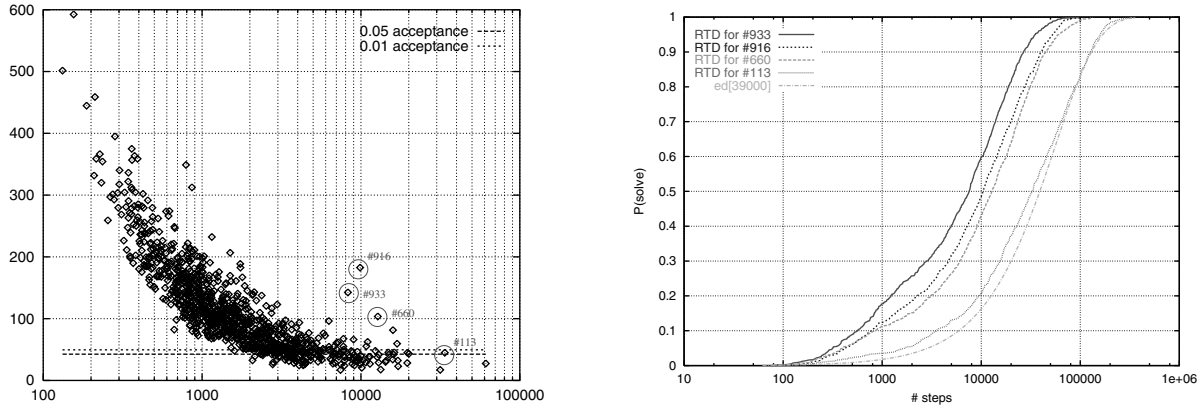


Figure 1: *Left:* Correlation between instance hardness (mean run-time for WalkSAT with approx. optimal noise measured in variable flips over 100 successful runs/instance; horizontal) and χ^2 values (vertical) from testing the RTDs of individual instances versus a best-fit exponential distribution for SATLIB test-set `uf100-430`; the horizontal lines indicate the acceptance thresholds for the 0.01 and 0.05 acceptance levels of the χ^2 test. *Right:* Irregular WalkSAT RTDs for outlier instances; note the deviation from a typical exponential distribution, indicated by the right-most curve.

sition" instances.[1] We show that the corresponding RTDs can be characterised by simple mixtures of exponential distributions, and provide evidence that this characterisation appears to apply to all hard problem instances for which irregular SLS behaviour is observed. Empirical evidence is presented for an explanation of this phenomenon based on search stagnation caused by "traps" in the underlying search spaces. Based on the insights from our analysis of irregular instances and search stagnation, we developped a conceptually simple Markov chain model which shows exactly the same behaviour as observed for GWSAT and WalkSAT/SKC on the irregular SAT instances studied here. This model provides simple and straight-forward explanations for our empirical results. Furthermore, it suggests several intuitive and testable connections between SLS behaviour and the structure of SAT instances.

Irregular Instances and Mixture Models

Our investigation starts with the observation that when studying the RTDs for WalkSAT (using approx. optimal noise settings) on sets of critically constrained Uniform Random-3-SAT instances, there are hard instances (as indicated by a high expected number of search steps for finding a solution) for which the search behaviour appears to deviate substantially from the typical memory-less behaviour reflected in exponential RTDs. Figure 1 shows the correlation between instance hardness for WalkSAT/SKC and the deviation of the corresponding RTD from a best-fit exponential distribution for test-set `uf100-430`, a set of critically constrained Uniform Random-3-SAT instances obtained from SATLIB[2] (Hoos & Stützle 2000b). This data was obtained using the same method as described in (Hoos & Stützle 1999). All RTDs reported in this study are based on at least 100 runs of the algorithm using cutoff parameter settings high enough to guarantee that a solution was found in each run without using any kind of restart mechanism. As noted by Hoos & Stützle (2000a), the high χ^2 values consistently observed for easy instances are due to effects of the initial search phase. In the present study, we largely ignore the effect of the initial search phase, which has been previously discussed and characterised in the literature (Hoos & Stützle 2000a).

The deviations reflected by high χ^2 values for hard instances, some of which are highlighted in Figure 1, are of a different nature. Closer inspection reveals that these irregular RTDs have an untypically high coefficient of variation (stddev/mean); all of them can be well aproximated by mixtures of exponential distributions of the form

$$\sum_{i=1}^{k-1} w_i \cdot ed[m_i] + \left(1 - \sum_{i=1}^{k-1} w_i\right) ed[m_k],$$

where $ed[m](x) = 1 - 2^{-x/m}$ is the cumulative distribution of an exponential distribution with median m and the w_i are the mixture weights. It should be noted that while for large k such mixtures can approximate any cumulative distribution function arbitrarily well, all approximations presented in this study use two components only and are hence significantly more restricted. Since the approximated empirical RTDs are generally based on at least 1,000 runs each, good approximations with this restricted mixture model reflect a rather surprising regularity of the underlying SLS behaviour rather than an overfitting effect due to an overly flexible model.

Additional experiments showed the same type of "outlier instances" for SATLIB test-sets `uf50-218` and `uf20-91`; in all cases, WalkSAT and GWSAT showed RTDs that could be well approximated by 2-component mixtures of exponential distributions. (See, *e.g.*, Figure 2; these results are reported in more detail in the extended version of this

[1] Algorithm outlines for GWSAT and WalkSAT/SKC, as well as a detailed description of the Uniform Random-3-SAT instances used in this study can be found in (Hoos & Stützle 2000a).

[2] www.satlib.org

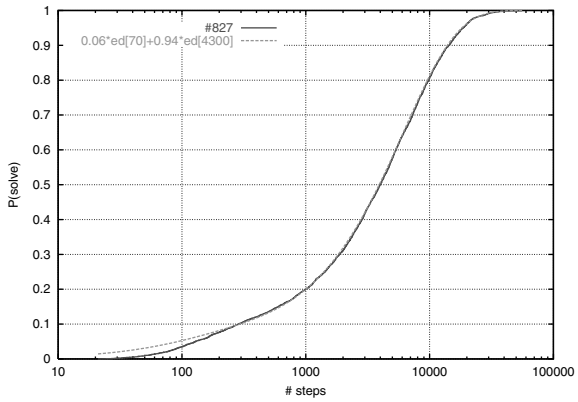


Figure 2: RTD for WalkSAT(noise=0.55) on hard irregular instance from test-set `uf50-218` and approximation by 2-component mixture of exponential distributions.

paper.) Overall, using 2-component mixtures of the previously mentioned generalised exponential distributions, all observed RTDs could be perfectly approximated to the degree supported by the sample size underlying the empirical RTD data. (The quality of these approximations can be seen, in a slightly different context, in Figure 5.) Interestingly, the extreme tails of all irregular RTDs are extremely well approximated by a model fitted to the whole distribution. In particular, there is no indication for so-called "heavy tails", as have been reported for the RTDs of certain high-performing randomised complete SAT algorithms (Gomes, Selman, & Kautz 1998).

Multiple Competing Solutions?

Perhaps the most obvious explanation for the observed mixture RTDs is based on the following idea: For instances with multiple solutions, one could assume that each solution (or cluster of solutions) has its own "basin of attraction", and that the attractivity of these basins might sometimes differ widely between various solutions. If conditional of being pulled into one given basin, the RTD of GWSAT or WalkSAT were an exponential distribution, then a biased random selection of the respective basin at the beginning of the search process would lead to the observed exponential mixture RTDs. Such a selection could be the result of the fact that GWSAT and WalkSAT both start the search at a randomly chosen assignment.

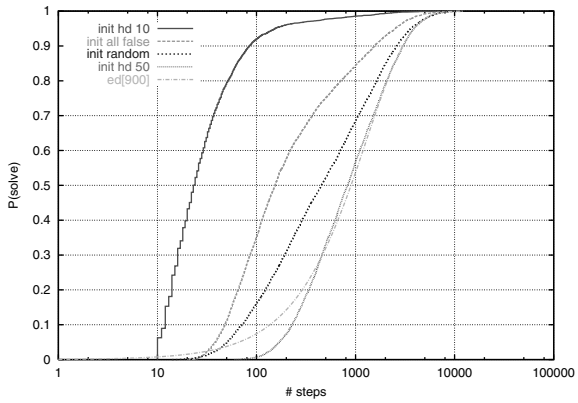
There are two ways of investigating the validity of this explanation: The first is based on a modification of the algorithms such that the search process is no longer initialised randomly, but at a specific variable assignment. If the proposed explanation of the irregular search behaviour were correct, using the fixed initialisation for the irregular instances from above should result in regular RTDs which, depending on the fixed initial assignment chosen, correspond to the components of the mixture obtained for random initialisation. A second validation experiment uses the unmodified algorithms (with random initialisation) and studies their RTDs on single-solution instances. If the attractivity of different solutions were the sole cause of mixture RTDs, these should not be observed on single solution instances.

For the first approach, we measured RTDs for a modified version of WalkSAT that always starts at a specific assignment applied to one of the irregular instances from test-set `uf50-218-1000`. Figure 3 (left) shows the RTD for WalkSAT/SKC with the standard, randomised initialisation as well as RTDs for a WalkSAT/SKC variant that always starts the search from the same given initial assignment. The specific initial assignments used here were the following: one at Hamming distance 10 from one of the instance's 48 solutions, one setting all variables to false, and one at Hamming distance 50 to a specific solution. With the exception of this last case, the resulting RTDs are mixture distributions rather than pure (generalised) exponentials, an observation that does not support the explanation proposed above. It is interesting to note that for this instance, the maximal Hamming distance between any two solutions is only 16, while the mean Hamming distance between solutions is 7. Hence, it appears that only when the search is initialised Hamming distant from the loosely clustered solutions, WalkSAT shows a simple exponential RTD. (This result is further confirmed by the RTDs for additional initial starting assignments, not shown here.)

For our investigation of the second approach, we generated sets of single-solution Uniform Random-3-SAT phase transition instances. This was done by generating Uniform Random-3-SAT instances in the usual (unbiased) way and subsequently checking for each instance whether it has exactly one solution.[3] For the three test-sets thus obtained, WalkSAT/SKC RTDs were measured (using approx. optimal noise) and fitted with exponential distributions, as described in the previous section. As can be seen in Figure 3 (right), the same kind of outlier instances as for the standard Uniform Random-3-SAT test-sets can be detected. The RTDs for these outlier instances are very similar to those shown in Fig. 1 and can be equally well approximated by mixtures of exponential distributions. These results indicate that single solution instances can exhibit the same irregular SLS behaviour, characterised by mixture RTDs, as instances with multiple solutions. Furthermore, it may be noted that test-sets of single-solution instances show a variability in search cost between the instances similar to the respective unrestricted test-sets. This clearly indicates that factors other than solution density have an important impact on the performance of SLS algorithms like WalkSAT. (Similar results were obtained for test-sets of critically constrained single-solution instances with 50 and 20 variables.)

It may be noted that the observations from the first of the two experiments described above still allow for an explanation in which the attraction areas of several or all solutions (or solution clusters) overlap at most or all locations in the given search space. While consistent with the nature of the randomised iterative improvement search process underlying WalkSAT/SKC and GWSAT, this modified hypothesis would still not explain the occurrence of mixture RTDs on

[3]This test was performed using REL_SAT, version 2.00 (Bayardo & Pehoushek 2000).



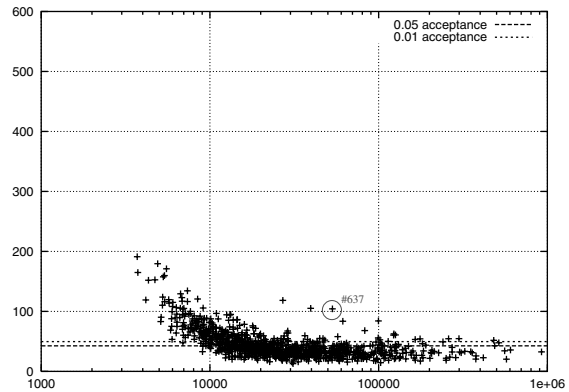


Figure 3: *Left:* RTDs for WalkSAT with fixed initialisation for irregular instance `#168` from test-set `uf50-218-1000`. *Right:* Correlation between mean run-time for WalkSAT (horizontal) and χ^2 values (vertical) from testing the RTDs of individual instances versus a best-fit exponential distribution for test-set `uf100-430-1000-s1`; the horizontal lines indicate the acceptance thresholds for the 0.01 and 0.05 acceptance levels of the χ^2 test.

single-solution instances.

Overall, the evidence from the two experiments does not support our initial hypothesis that mixture RTDs are simply caused by the presence of multiple solutions and respective basins of attraction.

Traps and Search Stagnation

An alternate explanation of the observed irregular behaviour is based on the assumption that for the respective problem instances, the local search process somehow gets trapped in regions of the search space that are attractive yet do not contain solutions. Intuitively, once trapped in such a region, it might take quite long before an SLS algorithm manages to escape from this region and find a route that finally leads to a solution. In this case, the mixture RTDs observed for the previously identified irregular instances reflect a stagnation of the search process caused by such traps. If this explanation were correct, we should be able to observe mixture RTDs and high search cost for SAT instances containing such traps.

To investigate this hypothesis, we first devised a way of combining two single-solution instances into a new SAT instance that contains one solution and a trap: For a single-solution instance F over n variables, $x_1, \ldots, x_n$, let $M(F) = (m_1, \ldots, m_n)$ denote the unique model of F, *i.e.*, F is true under the variable assignment $x_1 := m_1, \ldots, x_n := m_n$. Then for given single-solution instances F, G, we define the *plugged combination instances* $CP1[F, G]$ and $CP2[F, G]$ as follows:

$$CP1[F,G] = \bigwedge_{i=1}^{l}(\neg\hat{x} \vee \bigvee_{j=1}^{k} p_{ij}) \wedge \bigwedge_{i=1}^{m}(\hat{x} \vee \bigvee_{j=1}^{k} q_{ij}) \wedge \bigvee_{j=1}^{n} \neg m_j$$

where $M(F) = (m_1, \ldots, m_n)$ is the unique model of F; and

$$CP2[F,G] = CP1[G,F].$$

This construction uses a discriminator variable $\hat{x}$ to "switch" between the two component instances. Furthermore, the solution corresponding to one of the component instances is plugged by adding a single clause of length n. Note that adding this clause does not affect the objective function value (number of unsatisfied clauses) of any assignment other than the plugged solution; this implies that the difference between $C[F, G]$ and $CP1[F, G]$ is only visible to GWSAT or WalkSAT when the respective search process has reached the immediate neighbourhood of $M(F)$.

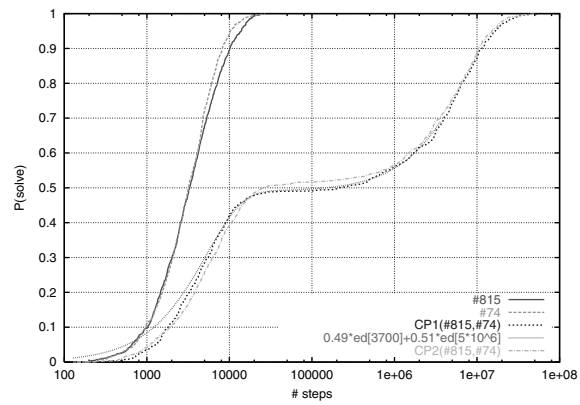


Figure 4: RTDs for GWSAT (noise=0.55) on easy instances from test-set `uf100-430-1000-s1` and plugged combinations and approximation with 2-component mixtures of exponential distributions. (The RTD for `CP2(#815,#74)` can be equally well approximated with a 2-component mixture of exponentials.)

We now assume that single-solution instances that are extremely easy for a given SLS algorithm are made easy by the fact that their single solution is very attractive for the algorithm. Based on this assumption, plugged combinations of easy single-solution instances would contain a very attractive trap, which should render them substantially more difficult to solve than the respective component instances.

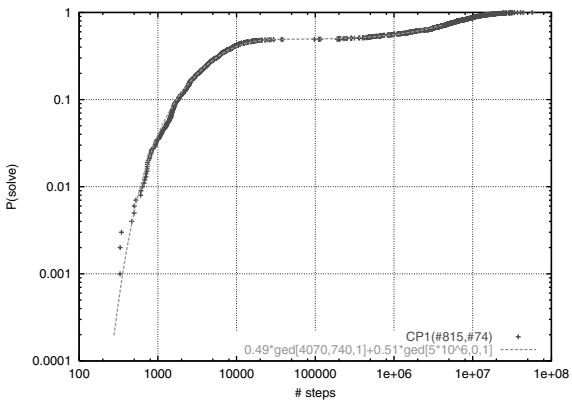


Figure 5: Approximation of GWSAT RTD for plugged combination of easy instances from test-set `uf100-430-1000-s1` with a mixture of two generalised exponential distributions; the empirical RTD data is the same as shown in Figure 4. The fit is equally good in both tails of the empirical distribution.

This conjecture was confirmed experimentally. Figure 4 shows a typical result, illustrating the hardness of plugged combinations of easy single-solution instances as well as the irregular RTDs obtained by solving these instances with GWSAT, which can be very well approximated by two-component mixtures of exponential distributions. When using mixtures of generalised exponential distributions[4] to model the initial search phase, we obtain perfect approximations (see Figure 5). Analogous results were obtained in numerous similar experiments using other component instances and test-sets. Overall, this confirms our hypothesis that traps, *i.e.*, attractive areas of the search space that do not contain solutions, can lead to search stagnation and the same type of irregular behaviour as previously observed for "outlier" Random-3-SAT instances.[5]

Based on this explanation, we now present a simpe abstract model for the observed SLS behaviour. Note that the behaviour of an SLS algorithm for SAT, such as GWSAT or WalkSAT, applied to a given SAT instance can be modelled as a Markov chain. Intuitively, the states of this chain represent areas of the search space, *i.e.*, sets of variable assignments that are considered equivalent in a certain sense. Simple examples for such sets of equivalent states are all assignments at a certain Hamming distance from the nearest solution, all assignments that satisfy a certain number of clauses, or all assignments that belong to a specific certain plateau region (Frank, Cheeseman, & Stutz 1997; Yokoo 1997; Hoos 1998). The transitions between the states thus defined correspond to the conditional probabilities of reaching a specific state from a given current state. Note that these transition probabilities depend on the problem instance as well on the SLS algorithm applied to it.

Here, we will consider a simplified version of such a model of SLS behaviour. Our model consists of a Markov chain with k states $s_1, \ldots, s_k$ (see Figure 6a). Let $p_{i,j}$ be the probability for a transition from state i to state j. We make the following assumptions:

$$p_{1,1} = 1 \quad (1)$$

$$p_{k,k-1} = 1 \quad (2)$$

$$\forall i; 1 < i < k : p_{i,i+1} = p^+ > 0 \quad (3)$$

$$\forall i; 1 < i < k : p_{i,i-i} = p^- > 0 \quad (4)$$

$$p^- = 1 - p^+ \quad (5)$$

The first assumption reflects the fact that state s_1 is an absorbing state representing the solution(s) of the given problem instance; SLS algorithms for SAT typically terminate as soon as a solution is found. Assumption (2) states that s_k is a reflecting boundary; it captures the intuition that any measure of distance to a solution modelled by this Markov chain will have a finite upper bound. The primary purpose of assumptions (3), (4), and (5) is to keep the model as simple as possible while allowing it to represent differences in problem size (reflected by k) and the attractivity of solutions (reflected by p^+ and p^-).

Interestingly, this simple Markov chain model shows precisely the same type of behaviour as GWSAT or WalkSAT applied to typical SAT instances for sufficiently high noise parameter settings. This can be seen empirically by comparing the respective RTDs, where an RTD for the model is defined as the distribution of the number of transitions needed to reach the solution state s_1 for the first time, starting from s_k (see Figure 6). It is worth noting that the same family of generalised exponential distributions introduced in (Hoos 1998; Hoos & Stützle 2000a) for accurately modelling the full RTDs of GWSAT and various WalkSAT variants can also be used to perfectly approximate the RTDs for the Markov chain model presented here. Unfortunately, so far it could not be formally proven that the RTDs for the model are always approximable by this family of distributions.

This Markov chain model can be easily extended to cases where the problem instances contain the kind of trap described in the previous section. In particular, the plugged combinations instances defined above can be modelled in a straight-forward way: We just combine the two models corresponding to the component instances into a branched chain model, as illustrated in Figure 6, where one of the two solution states is transformed into a reflecting boundary of the model (this state corresponds to the plugged solution), while the other becomes the single solution state of the branched model.

[4]This class of distribution is characterised by the cumulative distribution function

$$ged[m, \gamma, \delta](x) = 1 - 2^{-(x/m)^{1+(\gamma/x)^\delta}};$$

in most cases, empirical RTDs can be excellently approximated with a special case of this distribution for which $\delta = 1$.

[5]It should be noted that WalkSAT's behaviour on the plugged combination instances is slightly different from GWSAT's. This difference is due to the variable selection mechanism in WalkSAT and the occurrence of the discriminator variable in all clauses of a plugged combination instance; a detailed discussion can be found in the extended version of this paper.

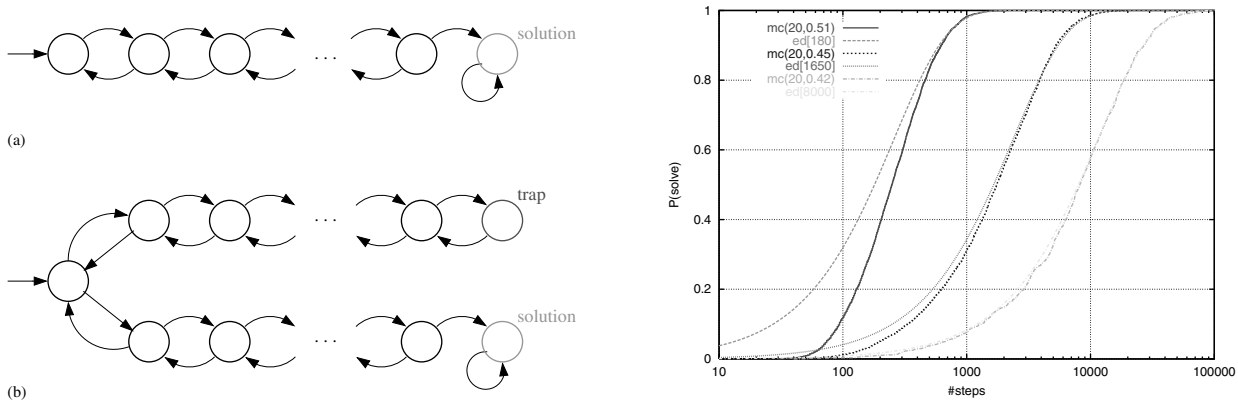


Figure 6: *Left:* Structure of simple Markov chain model (a) and branched model with trap (b); *right:* RTD for unbranched model (a) with $k = 20$, $p^+ = 0.52$, and $p^- = 0.48$.

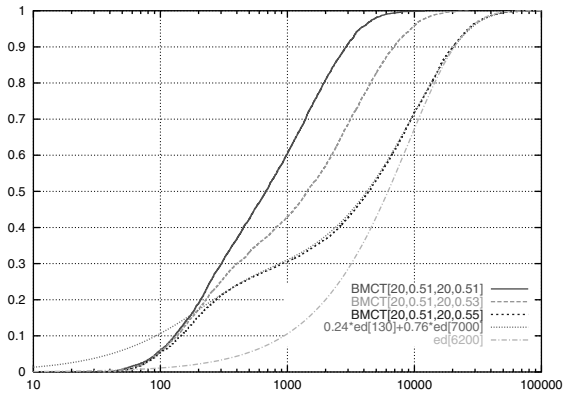


Figure 7: RTDs for branched Markov model with trap using different parameter settings can be approximated by mixtures of exponential distributions.

The RTDs for these branched Markov chain models are remarkably similar to those observed for the irregular SAT instances and for the plugged combination instances studied before. Depending on the length of the trap and solution branches and their respective transition probabilities p_t^- and p_s^-, we get the same type of mixture distribution as previously observed for GWSAT and WalkSAT/SKC. Consistent with the intuition behind the model and previous results for plugged combination instances, the two exponential components of the mixture RTD for the branched Markov chain model are more prominent for longer and more attractive trap branches (see Figure 7).

In the light of this model, the mixture distributions that are characteristic for the irregular instances reported earlier in this study are likely caused by prominent traps in the underlying search spaces. This hypothesis is consistent with the fact that many of the irregular instances are relatively hard, while none were detected amongst the easiest 10–15% of the instances within each of the respective test-sets. The model is also consistent with our observations on the behaviour of WalkSAT when using fixed initialisation from various points in the search space. When modelling an irregular instance by a branched Markov chain with a trap, it is clear that depending on the state at which the Markov process is initialised, we will observe the same qualitative differences in the resulting RTDs as observed for WalkSAT with fixed initialisation. In particular, when initialising at or near the trap state, the resulting RTD will show little or no irregular behaviour, but an increased search cost for all but the right tail of the distribution. Note that having the search space regions corresponding to the trap and solution states at high Hamming distance will maximise the area in which the attraction of either one dominates the behaviour of the search process and will thus lead to more prominent irregular SLS behaviour. Hence, it is reasonable to assume that for a prominently irregular instance, initialising Hamming distant from the solutions should be equivalent to initialising close to a prominent trap.

Conclusions and Future Work

Our study has shown that the run-time behaviour of two well-known SLS algorithms, GWSAT and WalkSAT/SKC, can be empirically characterised by mixtures of exponential distributions with a small number of mixture components. This extends previous empirical results to instances on which deviations from the typical, memory-less behaviour characterised by exponential distributions are observed; these "irregular" instances are not uncommon in the phase transition region of Uniform-Random-3-SAT and tend to be hard when compared to other instances from the same problem distribution.

As we have seen, the occurence of mixture RTDs can be explained based on a trap-based model of search stagnation. Somewhat surprisingly, we found that the empirically observed behaviour of the search process can generally be modelled by a very simple abstract model based on branched Markov chains. The model is based on the intuition that the search process implemented by procedures such as GWSAT or WalkSAT/SKC progresses through discrete stages, each of which has a characteristic "distance" to the nearest so-

lution. It is not entirely clear if and how these stages are explicitly manifested in the form of easily identifiable search space features; our current understanding of SLS behaviour suggests that the search stages might correspond to extensive plateau regions (Frank, Cheeseman, & Stutz 1997; Yokoo 1997; Hoos 1999). Furthermore, it is likely that at least one type of trap corresponds to the "failed clusters" observed by Parkes (1997). We currently investigate this hypothesis using advanced search space analysis techniques as well as the RTD characterisations and abstract search model developed in this study. Furthermore, it appears to be interesting to explore potential connections between traps and the factors underlying the hardness of Random-3-SAT instances studied by Singer *et al.* (2000), in particular backbone robustness.

Obviously, the simple Markov model is only an approximation of the behaviour of SLS algorithms such as GWSAT or WalkSAT in the multi-dimensional, complex search spaces corresponding to the SAT instances studied here. This approximation, however, seems to capture the essential features for the observed behaviour; therefore, it appears that by establishing the relation between it and identifiable features of the respective instances, considerable progress can be made towards a characterisation of the factors underlying the hardness of problem instances *w.r.t.* SLS algorithms. (It is worth noting that a slightly modified Markov chain model, where the probabilities of staying within the same state are not zero, *i.e.*, $p^- + p^+ < 1$, shows exactly the same type of RTDs as the simpler model studied here.)

There is some preliminary experimental evidence suggesting that the RTD characterisations and the abstract Markov model presented here might be rather broadly applicable. Apparently, the stagnation behaviour typically observed for GWSAT and WalkSAT when using lower-than-optimal settings of the noise parameter can be characterised and modelled analogous to the behaviour observed on irregular instances. It appears also likely that our characterisation generalises to other SLS algorithms for SAT (such as WalkSAT/TABU, Novelty$^+$, and R-Novelty$^+$), to randomised systematic search algorithms for SAT (such as Satz_RAND), and to SLS algorithms for other hard combinatorial problems (such as Iterated Local Search for MaxSAT or the Travelling Salesperson Problem). These observations and hypothesis are currently under further investigation.

Another direction for future research is of a more theoretical nature: It appears that relatively simple probabilistic models such as the branched Markov chain model for SLS behaviour presented here, should be amenable to theoretical analysis, such that the full RTDs for these models can be characterised analytically rather than experimentally, as was done in this study. Unfortunately, for the model proposed here, so far we have not been able to find in the literature or to derive analytic characterisations of the corresponding RTDs. Further questions of theoretical interest, such as under which conditions the RTDs of a Markov process can be characterised by mixtures of exponentials, appear to be also currently unanswered.

Acknowledgements

This research is supported by NSERC Individual Research Grant #238788. We gratefully acknowledge helpful comments and suggestions by Ian P. Gent, Henry Kautz, and Bart Selman as well as by the anonymous reviewers.

References

Bayardo, R., and Pehoushek, J. D. 2000. Counting Models using Connected Components. In *Proc. AAAI-2000*, 157–162.

Cheeseman, P.; Kanefsky, B.; and Taylor, W. M. 1991. Where the Really Hard Problems Are. In *Proc. IJCAI-91*, 331–337.

Frank, J.; Cheeseman, P.; and Stutz, J. 1997. When Gravity Fails: Local Search Topology. *(Electronic) J. of Artificial Intelligence Research* 7:249–281.

Gent, I. P., and Walsh, T. 1993. Towards an understanding of hill–climbing procedures for SAT. In *Proc. AAAI-93*, 28–33.

Gomes, C. P.; Selman, B.; and Kautz, H. 1998. Boosting Combinatorial Search Through Randomization. In *Proc. AAAI-98*, 431–437.

Gu, J. 1992. Efficient local search for very large-scale satisfiability problems. *ACM SIGART Bulletin* 3(1):8–12.

Hoos, H., and Stützle, T. 1999. Towards a Characterisation of the Behaviour of Stochastic Local Search Algorithms for SAT. *Artif. Intelligence* 112:213–232.

Hoos, H., and Stützle, T. 2000a. Local search algorithms for SAT: An empirical evaluation. *J. Automated Reasoning* 24:421–481.

Hoos, H., and Stützle, T. 2000b. SATLIB: An Online Resource for Research on SAT. In I.P. Gent, H. M., and Walsh, T., eds., *SAT 2000*, 283–292. IOS Press.

Hoos, H. 1998. *Stochastic Local Search - Methods, Models, Applications*. Ph.D. Dissertation, TU Darmstadt, Germany.

Hoos, H. 1999. SAT-encodings, search space structure, and local search performance. In *Proc. IJCAI-99*, 296–302.

Li, C. M., and Anbulagan. 1997. Look-ahead versus look-back for satisfiability problems. In *Proc. CP'97*, 341–355. Springer Verlag.

McAllester, D.; Selman, B.; and Kautz, H. 1997. Evidence for invariants in local search. In *Proc. IJCAI-97*, 321–326.

Parkes, A. J. 1997. Clustering at the Phase Transition. In *Proc. AAAI-97*, 340–345.

Schuurmans, D., and Southy, F. 2000. Local search characteristics of incomplete SAT procedures. *Artif. Intelligence* 132:121–150.

Schuurmans, D.; Southy, F.; and Holte, R. 2001. The exponentiated subgradient algorithm for heuristic boolean programming. In *Proc. IJCAI-01*.

Selman, B.; Kautz, H. A.; and Cohen, B. 1994. Noise strategies for improving local search. In *Proc. AAAI-94*, 337–343.

Singer, J.; Gent, I. P.; and Smaill, A. 2000. Backbone fragility and the local search cost peak. *J. of Artificial Intelligence Research* 12:235–270.

Wah, B., and Shang, Y. 1997. Discrete Lagrangian-Based Search for Solving MAX-SAT Problems. In *Proc. IJCAI-97*, 378–393.

Wu, Z., and Wah, B. 2000. An Efficient Global-Search Strategy in Discrete Lagrangian Methods for Solving Hard Satisfiability Problems. In *Proc. AAAI-2000*, 310–315.

Yokoo, M. 1997. Why Adding More Constraints Makes a Problem Easier for Hill-Climbing Algorithms: Analyzing Landscapes of CSPs. In *Proc. CP-97*, number 1330 in LNCS, 357–370. Springer Verlag.