SLS Algorithms for SAT: Irregular Instances, Search Stagnation, and Mixture Models

(Extended Abstract)

Holger H. Hoos
University of British Columbia
Computer Science Department
hoos@cs.ubc.ca

1 Introduction and Background

For the past decade, various types of stochastic local search (SLS) methods have been applied very successfully to the propositional satisfiability problem (SAT). These include the GSAT and WalkSAT families of algorithms [19, 4, 14], as well as several other algorithms based on similar ideas [7, 20, 21, 16, 17]. GSAT and WalkSAT algorithms have been extensively studied in the literature, and include some of the best-performing SAT algorithms known to date [11, 17]. Compared to other state-of-the-art SAT algorithms (*e.g.*, Satz [13]), 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.

As recently shown, the run-time behaviour of these algorithms when applied to hard SAT instances and when using sufficiently high noise parameter settings, is typically characterised by exponential run-time distributions (RTDs) [8, 10]. (Here, "sufficiently high" includes the range in which optimal performance, as reflected in minimal mean run-time, is achieved.) However, as we 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 [1]), 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 discussed in more detail in the next section, it reflects 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 understanding, predicting, and improving the performance and behaviour of SLS algorithms for SAT.

In this work, we show that the run-time behaviour of GWSAT and WalkSAT for these irregular instances can be empirically characterised by mixtures of exponential distributions with a small number of mixture components. This generalises the previous characterisation of these algorithms' behaviour and appears to capture the observed behaviour on all problem instances that we tested so far (including SAT-encoded instances of other combinatiorial problems). We investigate various possible explanations for this phenomenon, including the competing influence of multiple solutions, the impact of randomised search initialisation, and occurrence of search stagnation caused by "traps", *i.e.*, highly attractive non-solution areas of the underlying search space which are difficult for GWSAT or WalkSAT to escape from.

Our study provides substantial evidence that the observed irregular behaviour reflects search stagnation caused by traps. This hypothesis is confirmed by means of a construction method that combines instances with very atttractive solutions and artificially designed traps that leads to SAT instances on which GWSAT shows the same stagnation behaviour as previously observed for hard Uniform Random-3-SAT instances. This leads to the formulation of a simplistic abstract model for this irregular SLS behaviour based on branched Markov chains that empirically shows exactly the same irregular behaviour observed for GWSAT and WalkSAT/SKC.

Finally, we also present some empirical evidence indicating that 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. Even more surprisingly, we observe the same behaviour (characterised by exponential mixture RTDs) for Satz_RAND, one of the best known randomised systematic search algorithms for SAT [6].

2 Preliminaries

Our study focusses primarily on two rather simple, but very well known SLS algorithms for SAT, GWSAT and WalkSAT/SKC [19, 14]. To date, these algorithms are outperformed by newer, slightly more intricate SLS algorithms for SAT, some of which are variants of the same basic methods. Nevertheless, their peformance is known to be tightly correlated with some of the best known SLS algorithms for SAT [8, 11]; and their behaviour is easier to understand and relate to properties of the problem instances they are applied to [3].

Furthermore, we focus here on SAT instances from the widely studied phase transition region of the Uniform Random-3-SAT distribution [1]. In the past, problem instances of this type have been used extensively for benchmarking and studying SAT algorithm performance, and performance improvements on this instance class would typically translate into performance improvements for other types of SAT instances, such as SAT-encodings of other combinatorial problems [11]. The Uniform Random-3-SAT instance distribution also offers the advantage that at least two of the major factors determining instance hardness are well-known and relatively straight-forward to control for: the clauses-per-variable ratio (or constrainedness) and the number of solutions (or more precisely, the solution density) [1, 2]. The former is controlled by construction of the formulae, the second by using variants of complete SAT algorithms for determining the number of solutions of any given problem instance and filtering out instances which do not have the desired number of solutions (this is a computationally expensive process that somewhat limits the size of problem instances we can deal with in the context of this study). All problem instances used in the context of this study are satisfiable, since the SLS algorithms investigated here are incomplete in the sense that they cannot determine unsatisfiability.

3 Irregular Instances and Mixture RTDs

Using the same techniques as in [8, 10], we shows that on certain Random-3-SAT instances, WalkSAT (using close-to-optimal noise settings) shows irregular RTDs, *i.e.*, RTDs which significantly deviate from the generalised exponential distribution model introduced in [8]. Instances for which we find such irregular RTDs show the same effect for GWSAT as well as for WalkSAT. Our empirical results indicate that such “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.

For all irregular instances encountered in the context of this study, the corresponding RTDs can be well approximated by mixtures of exponential distributions with only two components. (Notice that this parametric family of distributions has only three parameters — hence, considering the high resolution of the RTDs studied here, these approximations do not merely reflect the fact that the fitted model is overly general.) Interestingly, such finite mixtures of exponential distributions, also known as hyperexponential distributions, have numerous applications in statistical modelling and reliability theory. Although no attempt was made to specifically fit the mixture distribution to the right tail of the irregular RTDs, in all cases the mixture model approximates the right tail of the corrsponding empirical RTD very well. In particular, there is no indication for “heavy tails”, as reported for the RTDs of certain high-performing randomised complete SAT algorithms [6].

4 Multiple Competing Solutions?

Perhaps the most obvious explanation for the observed mixture RTDs is based on the idea that in the presence of multiple solutions with distinct basins of attraction, the random search initialiation in SLS algorithms like WalkSAT or GWSAT leads to the selection of one such basin of attraction. To test this hypothesis, we designed two different computational experiments; however, the results from both experiments do not support the proposed explanation.

An alternate explanation is based on the interpretation of the observed mixture RTDs as evidence for stagnation behaviour of the corresponding SLS algorithm. It assumes that search progress can be impeded by the simultaneous influence of the basins of attraction corresponding to multiple solutions. (This idea has been proposed by Bart Selman as a basis for generating extremely hard SAT instances [18].) We investigated this hypothesis using specifically designed combinations of single-solution SAT instances; but overall, our experimental results suggest that competing or conflicting solutions are not the primary cause for the earlier observed irregular behaviour of WalkSAT and GWSAT.

5 Building Traps

This leads us to an explanation of the observed irregular behaviour 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. (We will discuss the actual nature of such traps and their effect on algorithms like GWSAT or WalkSAT in more detail later.) If this hypothesis were correct, we should observe mixture RTDs and high search cost for SAT instances containing such traps.

This hypothesis is confirmed by means of a construction method that combines instances with very atttractive solutions and artificially designed traps that leads to SAT instances on which GWSAT shows the same stagnation behaviour as previously observed for hard Uniform Random-3-SAT instances. This construction combines two single-solution SAT instances as in the experiment mentioned in the previous section, but then converts one solution into a trap by "plugging" it in a way that, besides eliminating that solution, causes almost no side-effects on the combination instance. This way, we obtain combination instances that have one solution and one trap each, where the trap is as attractive for the SLS algorithm as the corresponding second solution before it was plugged.

The results of our analysis of SLS performance on plugged combination instances indicate clearly that traps, *i.e.* attractive areas of the search space that do not contain solutions, lead to high search cost indicative of search stagnation. In the case of GWSAT, these traps also cause the same type of irregular behaviour previously observed on outlier Random-3-SAT instances. For WalkSAT, the effect of the traps is masked by an extreme form of search stagnation caused by the impact of a simpe syntactical feature of the combination instances on the underlying search mechanism.

6 A Simple Markov Model for SLS Behaviour

The trap-based explanation of the previously observed, irregular SLS behaviour gives rise to a simplified abstract formal model that exhibits the same type of irregular behaviour as the SLS algorithms studied before. Our starting point is the general idea 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. Inuitively, 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 [3, 22, 8]. The transitions between the states thus defined correspond to the conditional probabilities of reaching a specific state from a given current state. We show that even under fairly strong simplifying assumption, 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 model can be easily extended to cases where the problem instances contain the kind of traps 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, 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. The RTDs for these branched Markov chain models are remarkably similar to those observed for the irregular SAT instances and the plugged combination instances studied before. Depending on the length of the trap and solution branches and their respective transition probabilities, we get the same type of mixture distribution as previously observed for GWSAT and WalkSAT.

We show how this model of SLS behaviour provides explanations for some of our previously reported empirical observations, including the impact of search initialisation from a fixed point on WalkSAT's run-time behaviour.

7 SLS Behaviour for Lower-than-Optimal Noise

It has been previously observed that low noise values lead to stagnation behaviour for GWSAT and various WalkSAT variants [8, 11]. The corresponding RTDs have heavier right tails than observed for optimal or higher-than-optimal noise settings and can not be approximated by exponential (or generalised exponential) distributions.

We present empirical evidence indicating that the same characterisation using simple mixtures of exponentials also applies to the run-time behaviour of GWSAT and WalkSAT for smaller-than-optimal noise values on randomly generated instances as well as on SAT-encoded instances from other domains. Hence, our abstract model of SLS behaviour as well as the precise characterisation of RTDs using simple mixtures of exponential distributions appears to be applicable for arbitrary noise settings. Our study provides an intuitive and plausible explanation for this observed behaviour in the light of the branched Markov chain model described in the last section.

8 Directions for Future Work

This study raises a number of interesting questions, some of which we are currently investigating. The first of these concerns our branched Markov chain model of SLS behaviour, which is based on the intuition that the search process implemented by procedures such as GWSAT or WalkSAT progresses through discrete stages, each of which has a characteristic "distance" to the nearest solution. 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 [3, 22, 9]. It is furthermore likely that at least one type of traps corresponds to the "failed clusters" observed by Parkes [15]. 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.

Obviously, the simple Markov model is only an approximation of the behaviour of SLS algorithms such as GWSAT or WalkSAT in the multi-dimensional, complext 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.

There is some preliminary experimental evidence suggesting that the RTD characterisations and the abstract Markov model presented here might be rather broadly applicable 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; as far as other SLS algorithms are concerned, they are plausible and might indicate the fact that the view of stochastic local search behaviour presented here is sufficiently general to apply to a wide range of algorithms and problems.

In the context of randomised systematic search behaviour, we believe that if the RTD characterisation using mixtures of exponentials applies in general, the underlying reasons will likely differ from those captured in the abstract Markov model for SLS behaviour presented here. Nevertheless, it is interesting to note that earlier experiments with Satz_RAND, one of the two randomised systematic search algorithms for SAT presented in [6], mixtures of exponentials with relatively few componentes seemed to give significantly better approximations of the observed RTDs than the heavy-tailed distributions used by Gomes *et al.* [5, 12]. Furthermore, in the systems literature, mixtures of exponentials are often used to approximate experimentally observed "fat-tailed" or heavy-tailed response or waiting time distributions. A more detailed investigation of the run-time behaviour of randomised systematic search algorithms for SAT is currently underway.

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.

References

[1] Peter Cheeseman, Bob Kanefsky, and William M. Taylor. Where the Really Hard Problems Are. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, IJCAI-91, Sidney, Australia*, pages 331–337, 1991.

[2] David A. Clark, Jeremy Frank, Ian P. Gent, Neven Tomov Ewan McIntyre, , and Toby Walsh. Local search and the number of solutions. Unpublished Manuscript, from WWW, February 1997.

[3] Jeremy Frank, Peter Cheeseman, and John Stutz. When Gravity Fails: Local Search Topology. *(Electronic) Journal of Artificial Intelligence Research*, 7:249–281, 1997.

[4] Ian P. Gent and Toby Walsh. Towards an understanding of hill–climbing procedures for SAT. In *Proceedings of AAAI'93*, pages 28–33. MIT press, 1993.

[5] C. Gomes, B. Selman, N. Crato, and H. Kautz. Heavy-Tailed Phenomena in Satisfiability and Constraint Satisfaction Problems. *J. Automated Reasoning*, 24:67–100, 2000.

[6] Carla P. Gomes, Bart Selman, and Henry Kautz. Boosting Combinatorial Search Through Randomization. In *Proceedings of AAAI'98*, pages 431–437. AAAI Press / The MIT Press, Menlo Park, CA, 1998.

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

[8] H.H. Hoos. *Stochastic Local Search - Methods, Models, Applications*. PhD thesis, TU Darmstadt, Germany, 1998.

[9] H.H. Hoos. SAT-encodings, search space structure, and local search performance. In *Proc. IJCAI-99*, pages 296–302. Morgan Kaufmann, 1999.

[10] H.H. Hoos and T. Stützle. Towards a characterisation of the behaviour of stochastic local search algorithms for sat. *Artificial Intelligence*, 112:213–232, 1999.

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

[12] Holger H. Hoos. Heavy-tailed behaviour in randomised systematic search algorithms for SAT? Technical Report TR-99-16, Computer Science Department, The University of British Columbia, 1999.

[13] Chu Min Li and Anbulagan. Look-ahead versus look-back for satisfiability problems. In *Proceedings of CP'97*, pages 341–355. Springer Verlag, 1997.

[14] David McAllester, Bart Selman, and Henry Kautz. Evidence for invariants in local search. In *Proceedings of IJCAI-97*, 1997.

[15] Andrew J. Parkes. Clustering at the Phase Transition. In *Proceedings of AAAI'97*, pages 340–345, 1997.

[16] D. Schuurmans and F. Southy. Local search characteristics of incomplete sat procedures. *Artificial Intelligence*, 132:121–150, 2000.

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

[18] B. Selman. Personal communication, 1999.

[19] Bart Selman, Henry A. Kautz, and B. Cohen. Noise strategies for improving local search. In *Proceedings of AAAI'94*, volume 1, pages 337–343, 1994.

[20] B.W. Wah and Y. Shang. Discrete Lagrangian-Based Search for Solving MAX-SAT Problems. In *Proceedings of IJCAI'97*, pages 378–393. Morgan Kaufmann Publishers, 1997.

[21] Z. Wu and B.W. Wah. An Efficient Global-Search Strategy in Discrete Lagrangian Methods for Solving Hard Satisfiability Problems. In *Proceedings of AAAI 2000*, pages 310–315, 2000.

[22] Makoto Yokoo. Why Adding More Constraints Makes a Problem Easier for Hill-Climbing Algorithms: Analyzing Landscapes of CSPs. In *Proceedings of CP'97*, number 1330 in LNCS, pages 357–370. Springer Verlag, 1997.