# An analysis of search landscape neutrality in scheduling problems

**Andrew M. Sutton**
sutton@cs.colostate.edu
Department of Computer Science
Colorado State University
Fort Collins, CO, 80521

Scheduling problems are often amenable to stochastic search techniques such as genetic algorithms, simulated annealing, tabu search, and different variants of iterated local search. The behavior of these algorithms can be described as a dynamic process evolving on a *search landscape*.

Many scheduling domains produce a landscape feature known as *neutrality*. Neutral moves on a landscape are characterized by search step transformations that have no effect on the objective function under consideration. The behavior of the search process is vastly different in the presence of neutrality than on many classical combinatorial problems, rendering established search hardness measures ineffective (Barnett 1998).

Several research projects have quantified static neutral features in terms of "plateaus" and "benches" or "ledges" (Frank, Cheeseman, & Stutz 1997; Hoos & Stützle 2004). However, little is understood about the dynamic behavior of different search algorithms in the presence of high neutrality. Interestingly, in many cases, very simple biased random-walk algorithms have been found to produce competitive results on high-neutrality landscapes (Forrest & Mitchell 1993; Barnett 2001; Barbulescu *et al.* 2006). One important example is the WalkSAT algorithm developed for maximum Boolean satisfiability (Selman, Kautz, & Cohen 1996). These results seem counterintuitive since we would expect simple local search algorithms to fail in the absence of local move gradients to direct search progress. In this research we seek to gain an understanding of the effect of neutrality on search.

## Scheduling problems and neutrality

Given a set $X$ of candidate solutions, a *neighborhood* relation $N : X \mapsto \mathcal{P}(X)$ imposes a topology on $X$. Together with an objective map $f : X \mapsto \mathbb{R}$, we have a discrete *landscape*. A neutral move is simply defined as a pair $x, y \in X$ such that $y \in N(x)$ and $f(x) = f(y)$. A high neutrality landscape has a relatively large proportion of such moves. Neutral moves can be somewhat vacuous in the sense that they provide no gradient information about motion towards the optimal solution (see Figure 1 for an illustration).

In scheduling problems, neutral moves may occur when there is an inherent symmetry in the objective that can arise



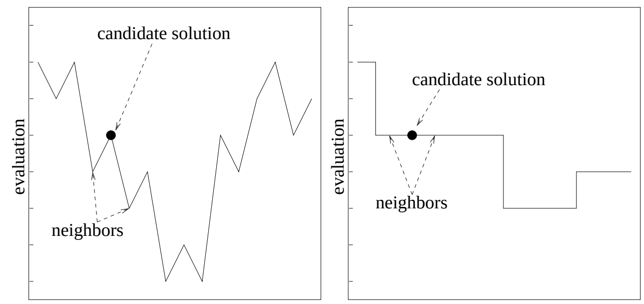


Figure 1: Simplified example of a non-neutral search landscape (left) and a high neutrality landscape (right).

from several potential issues. Using an *intermediate representation* may in fact embed the space of candidate solutions into a schedule space that is too large (Reidys & Stadler 2002). Candidate solutions can often be characterized by some intermediate representation which facilitates the application of search operators. For example, many schedules can be represented by *task orderings* or *activity graphs*. In these cases, the set $X$ of candidate solutions may not cover the entire set of feasible schedules. It may also be possible that $X$ contains infeasible schedules. In spite of this, often a concise representation makes $X$ easier to work with in practice, and also makes theoretical results more accessible.

We focus on the particular case of *task orderings*. In this case, a candidate solution can be represented by a total order on $n$ tasks. The order is transformed into a schedule by a "schedule builder": a deterministic (often greedy), problem-specific scheduling algorithm. This approach reduces the original optimization problem to a search over orderings on $n$ points and has been successfully employed in many scheduling domains (Davis 1985; Whitley, Starkweather, & Fuquay 1989; Taillard 1990; Barbulescu *et al.* 2006). In this case, several candidate solutions may be mapped to identical schedules or to differing schedules with the same objective value. If neighboring solutions have identical values, the link between them is neutral.

Neutrality can also emerge from a discretization of the objective function itself. Many scheduling objectives map to an integral codomain with a cardinality that is significantly lower than the number of candidate solutions. For example, evaluations such as makespan, number of tardy tasks, and number of conflicting tasks are all objectives with a (relatively) small number of integral possible values. Two neighboring solutions are more likely to be neutral in these cases.

Interestingly, there appears to be a controversy in the literature about the *benefits* of neutrality. Nor is it clear exactly *why* or *under what conditions* neutrality can be advantageous. For instance, Ebner et al. (2001) suggest that adding redundancy to a search space in order to *force* neutrality can improve convergence. However, Knowles and Watson (2002) give empirical support *against* this, suggesting that added redundancy might inflate the search space and slow down convergence. In some cases, removing neutrality has shown to be detrimental. In a satellite scheduling problem, Roberts et al. (2005) showed that removing neutral moves by restricting the neighborhood hurt performance of a local search algorithm.

## Modeling the dynamics of search in the presence of neutrality

Most research on neutral landscapes involves quantifying "static" features either exhaustively or statistically. For example Frank et al. (1997) explore several static characteristics that arise in the presence of neutrality such as "plateaus" (regions of connected neutral neighbors) and "benches" (plateaus with exits). Reidys and Stadler (2001) explore neutrality as a variable of a "random field" landscape. On the other hand, the "dynamic" behavior of search on highly neutral landscapes is not as well understood. Barnett (2001) studies the statistical dynamics of "netcrawler": a simple stochastic local search process. However, these results hold under rather heavy simplifying restrictions. For example, no suboptimal plateaus exist from which there is no escape and each improving move accesses only the next fittest level.

In general, the behavior of search can be described by a (possibly stochastic) discrete dynamical process evolving on the underlying landscape. The characteristics of this process govern how effective one particular method may be over another. A rich amount of information becomes available as the process evolves that could potentially be exploited in order to influence search behavior.

When the search process is Markovian, we can model it as a Markov process on $X$. Using the theory of Markov chains, we can have explicit access to probabilities and waiting times by computing the stationary distribution. Each state of the chain corresponds to a state $x \in X$. Of course, such a computation is only tractable for small solution spaces. For larger instances, we propose to sample the solution space to construct a "sublandscape" that may have similar properties to the larger space. We will employ both a uniform random sampling and a Metropolis-Hastings technique to focus on regions of the space.

We conjecture that simple stochastic local search algorithms might perform well in the presence of neutrality because of an inherent increase in *mutual accessibility* under high neutrality. In other words, distinct solutions that would otherwise be inaccessible on a non-neutral landscape could become mutually reachable under neutral moves. This increase in accessibility will affect convergence and stagnation probabilities: an effect that will be reflected in the dynamic model. To study this, we create a large number of random landscapes controlling for neutrality (we measure neutrality as the proportion of neutral moves on the landscape).

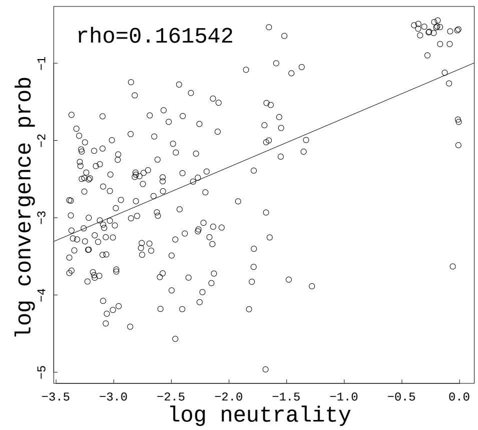


Figure 2: Log-log relationship of neutrality and convergence probabilities with least-squares regression line.

**Preliminary results.** We generated 100 random landscapes controlling for neutrality by specifying the number of allowed unique values in the codomain of $f$. The neighborhood $N$ was defined to be the scheduling *shift* operator (Taillard 1990) on permutations of length 6 (i.e., this gives a solution space cardinality of $|X| = 720$). For each landscape we created a Markov chain to model a simple hill-climbing search (that accepts neutral moves). The Markov chains are produced by masking out strictly disimproving moves in the adjacency matrix of $N$ and normalizing the resulting matrix to make it stochastic. For each matrix we calculate the stationary distribution and find the state probability associated with the global optimum. For these experiments, we fixed the number of global optima at 1.

Our preliminary results suggest that convergence probability is (weakly) positively correlated with neutrality. A Spearman rank correlation test gives a correlation estimate of $\rho \approx 0.16$ with $p < 0.0001$. We report these results in Figure 2. Note that we have introduced a log-log transformation to the data to stabilize variance.

We also found that stagnation probability *decreases* with neutrality ($p < 0.0001$) supporting our hypothesis that mutual accessibility increases over all states. However this accessibility comes at a price. Convergence probabilities in the stationary distribution hold under the assumption of unbounded time. Finite time performance will undoubtedly decrease since plateaus become intractably large. This suggests a crossover point which we seek to capture.

We have also found that, on some scheduling problems, neutrality might be subspace dependent. In other words, only a certain subset of tasks are responsible for creating neutral moves. This effect suggests the presence of "backbone" like structures similar to those found in maximum Boolean satisfiability problems (Monasson *et al.* 1999).

## Ongoing work

The goal of this research is to study how neutrality affects the dynamics of search processes. Many questions remain unanswered about this surprisingly pervasive search space characteristic. Plateaus can hinder many search algorithms by offering no gradient information to direct motion toward global optima. However, neutrality may also prove beneficial. All other things being equal, a neighborhood structure with *more* neutral moves will tend to have more acceptable neighbors. This will increase connectivity between candidate solutions and could potentially eliminate many local optima. This may even result in a reduction in stagnation likelihood.

We intend to study the mechanism by which a simple stochastic local search method interacts with neutrality. Our preliminary results suggest that the mean convergence probability tends to be positively correlated with neutrality. We hypothesize that this is evidence of an effect related to the increased solubility observed by stochastic local search. We also believe that attractor basins tend to weaken as neutrality increases. We intend to rigorously explore how neutrality distinctly affects convergence.

Barbulescu et al. (2006) discovered that a simple algorithm called Attenuated Leap Local Search (ALLS) achieves competitive performance on a satellite scheduling problem that features a highly neutral landscape. ALLS produces local neighbors by simply taking an unbiased random walk of length $k$ on the landscape. As the search progresses, $k$ is attenuated to take shorter leaps. The authors conjecture that algorithms that can "move quickly" across plateaus tend to perform best. Despite these results, it is still not clear what constitutes "efficient" motion on a plateau. Is a simple random walk in fact the optimal strategy to clear a plateau? How does the initial setting and scheduling of the $k$ parameter affect convergence? Do these results generalize to other important scheduling domains?

Neighborhood structure and size can directly transform the underlying landscape. We seek to investigate the role of neighborhood size in neutrality. If increasing neighborhood size affects neutrality, is the expense of higher neighborhood costs and potential search space inflation justified by favorable search dynamics?

Finally, many recent results in landscape analysis depend on the fact that certain landscape features are "statistically isotropic." Roughly speaking, this means that the features are homogeneous with respect to their "location" on the landscape. We conjecture that neutrality is not isotropic in practice, and in fact depends on the evaluation of the objective function. If neutrality is not isotropic, we might expect the search process to behave differently as it neared the globally optimal solution.

Answers to these questions will provide insight into the dynamics of search processes in the presence of neutrality: a common feature in many scheduling problems. This work can serve to 1) further our understanding about what landscape features make a problem amenable to search 2) inform search algorithm design to more efficiently handle landscapes with a large number of neutral neighbors.

## References


Barbulescu, L.; Howe, A. E.; Whitley, L. D.; and Roberts, M. 2006. Understanding algorithm performance on an oversubscribed scheduling application. *Journal of Artificial Intelligence Research* 27:577–615.

Barnett, L. 1998. Ruggedness and neutrality: the NKp family of fitness landscapes. In *ALIFE: Proc. of the 6th International Conf. on Artificial Life*, 18–27. Cambridge, MA, USA: MIT Press.

Barnett, L. 2001. Netcrawling – optimal evolutionary search with neutral networks. In *Proc. of the 2001 Congress of Evolutionary Compuation (CEC2001)*, 30–37.

Davis, L. 1985. Job shop scheduling with genetic algorithms. In *International Conf. on Genetic Algorithms and their Applications*.

Ebner, M.; Shackleton, M.; and Shipman, R. 2001. How neutral networks influence evolvability. *Complexity* 7(2):19–33.

Forrest, S., and Mitchell, M. 1993. Relative building-block fitness and the building-block hypothesis. In Whitley, L. D., ed., *Foundations of Genetic Algorithms 2*. San Mateo, CA: Morgan Kaufmann. 109–126.

Frank, J.; Cheeseman, P.; and Stutz, J. 1997. When gravity fails: Local search topology. *Journal of Artificial Intelligence Research* 7:249–281.

Hoos, H. H., and Stützle, T. 2004. *Stochastic Local Search: Foundations and Applications*. Morgan Kaufman Publishers.

Knowles, J. D., and Watson, R. A. 2002. On the utility of redundant encodings in mutation-based evolutionary search. In *Proc. of the Seventh International Conference on Parallel Problem Solving from Nature (PPSN-2002)*.

Monasson, R.; Zecchina, R.; Kirkpatrick, S.; Selman, B.; and Troyansky, L. 1999. Determining computational complexity from characteristic 'phase transitions'. *Nature* 400:133–137.

Reidys, C. M., and Stadler, P. F. 2001. Neutrality in fitness landscapes. *Applied Mathematics and Computation* 117:321–350.

Reidys, C. M., and Stadler, P. F. 2002. Combinatorial landscapes. *SIAM Review* 44:3–54.

Roberts, M.; Whitley, L. D.; Howe, A. E.; and Barbulescu, L. 2005. Random walks and neighborhood bias in oversubscribed scheduling. In *Multidisciplinary International Scheduling: Theory and Applications (MISTA 05)*.

Selman, B.; Kautz, H.; and Cohen, B. 1996. Local search strategies for satisfiability testing. In Johnson, D. S., and Trick, M. A., eds., *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, volume 26. AMS.

Taillard, E. 1990. Some efficient heuristic methods for the flow shop sequencing problem. *European Journal of Operations Research* 47:65–74.

Whitley, D.; Starkweather, T.; and Fuquay, D. 1989. Scheduling problems and traveling salesmen. In *Proc. of the Third International Conf. on Genetic Algorithms*.