Fitness Landscape and the Behavior of Heuristics

C. Fonlupt and D. Robilliard and P. Preux*

April 30, 1997

Abstract

This paper focuses on using the fitness landscape paradigm in order to gain a better understanding of the behavior local search heuristics in order to sove NP-hard problems. We concentrate ourselves on the *TSP* and we compare experimental facts about landscapes associated with the well-known 2opt-move and city-swap operators. We measure the "fitness-distance correlation" (*FDC*), i.e. the correlation between the fitness of sampled points and their distance to a global optimum. We show that the 2opt-move landscapes have greater *FDC* values than the city-swap. We relate this with the fact that 2opt hill-climbers walks terminate on solutions quite close to the global optimum, whereas city-swap walks are unable to reduce the distance to the optimum (distances are measured in terms of operator metric). This confirms former propositions about the importance of a good *FDC*, and examplify the existence of non artificial problems exhibiting this property. The difference in *FDC* also seems more convincing than the usual auto-correlation measures when it comes to explain the difference in performance between these two operators. Furthermore, a good *FDC* value points to the possible use of intensifying search techniques.

1 Introduction

There has been a recent interest in computational search techniques like *Hill-Climbers (HC)*, *Genetic Algorithms (GA)* or *Tabu Search (TS)* for combinatorial optimization problems. *GA*s are known to work efficiently in the case of numerical optimization, but results are less favorable when applied to combinatorial issues. They all implement *local search*: the size of the set of solutions being too big for exhaustive exploration, the search is restricted to manageable subsets of the combinatorial space. This restriction is generally obtained via a neighborhood notion: starting out from a solution (or subset of solutions) generated randomly or via some algorithm, each solution has a reasonably small set of neighbors, among which the exploration will be pursued. Usually the neighborhood can be defined in terms of applying one or many operators to the point (or set of points) of the search space, for example flipping one bit of its binary representation. These methods are also called *adaptive*, because the exploration usually goes through one of the bests of all neighbors, in a way that reflects Darwinian adaptation. So *HC*, *GA* and *TS* are indeed meta-heuristics, in the sense

*Université du Littoral, Laboratoire d'Informatique du Littoral, BP 719, 62228 Calais Cedex, France, fonlupt@lil.univ-littoral.fr

that they yield different results when choosing different neighborhood operators or a different selection process for obtaining the most promiseful neighbor.

For all these techniques, there is a real need for a better understanding of their intrinsic behavior. An interesting approach to gain some intuition on these matters is the *fitness landscape* paradigm. To biologists, this notion is not a new one, having been defined by Sewall Wright [1] who used it as a model for Darwinian evolution. Intuitively speaking, a fitness landscape is obtained by associating a fitness value with each point of a combinatorial space where a neighborhood notion is defined. Then a walk can be made along neighboring points, considering the fitness values encountered as altitude values. Along this walk, one will wander through plateaus and plains, will climb hills or peaks, and so on. Areas of high fitness may be viewed as mountainous regions whereas those with low fitness may be considered as valleys. This provides a track for an intuitive and qualitative reasoning, even if one must be aware of not relying too much on this view.

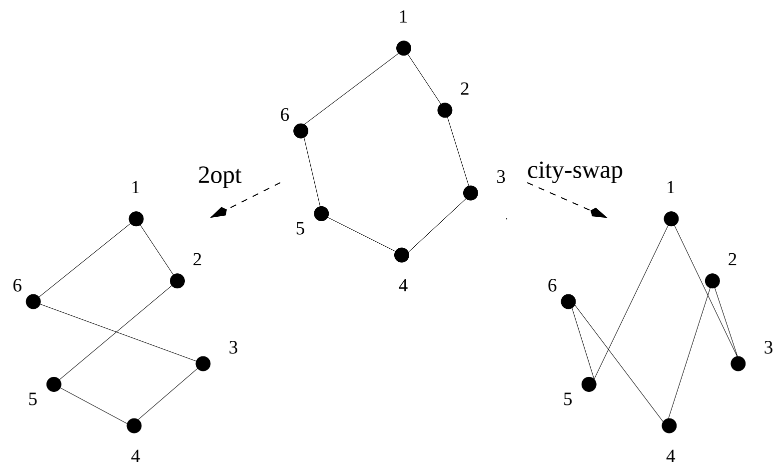
Our main issue is using the fitness landscape paradigm in order to understand the behavior of the search methods mentioned above. We currently focus on the well-known *Traveling Salesman Problem (TSP)*, in its planar euclidean symmetric version. In this paper, we concentrate on *HC* algorithms, using two neighborhood operators: the so-called *2-opt move* and the *city swap*. These operators are frequently encountered as mutation operators in *GA*, hence we expect that results obtained through the study of *HC* models will also be usable for *GA*, which are difficult to study from the fitness landscape point of view. We investigate some statistical features of reference instances of the *TSP*. These measures follow the current trend for statistical analysis of fitness landscapes [2, 3, 4]. We give confirmation to Jones and Forrest's proposition [4] that fitness-distance correlation (*FDC*) is a crucial point when caracterizing a problem difficulty, and we give a positive answer to their interrogation about the existence of real, non-artificial problems with large positive *FDC*. We exhibit *HC* walks that illustrate the impact of this correlation and open new ways for intensifying search methods.

2 Definition of the *TSP* Landscape

To solve a *TSP* instance, one has to find a shortest tour linking every cities of the instance, and coming back to the first one at the end. The two operators we study are well known in the literature ([2, 5, 3]):

1. The *2opt-move* reverses a sequence of consecutive cities in the tour (or, equivalently, it "exchanges" 2 edges) as shown in figure 1. Each tour has $\frac{n(n-3)}{2}$ neighbors.
2. The *city-swap* swaps any two cities in the tour as shown in figure 1. Each tour has $\frac{n(n-1)}{2}$ neighbors.

A very natural and often used fitness function consists in using the length of tours. We will use this standard, though it implies that the reader has to remember we are not climbing uphill as said in the general introduction, but on the opposite we are going downhill : better fitness means lower fitness value and so shorter tour.



The tour on the left is obtained by a 2opt-move applied on edges $(2,3)$ and $(5,6)$: these edges are removed and there is only one way to rebuild a hamiltonian tour with two new edges. The tour on the rigth is obtained by a city-swap applied on cities 1 and 4 : the tour $(1,2,3,4,5,6)$ becomes $(4,2,3,1,5,6)$.

Figure 1: Effect of 2opt-move and city-swap operators.

In the next sections we will work with reference instances taken from the tsplib [6], where we know a global optimum. In order to examine the behavior of *HC* walks with respect to the global optimum, we will use different notions of distance, based on the neighborhood operators. Let t_1 and t_2 be two tours:

1. Let $\delta_c(t_1, t_2)$ denote the smallest number of the city swap operator steps needed to obtain t_1 from t_2.
2. Let $\delta_2(t_1, t_2)$ denote the smallest number of the 2-opt operator steps needed to obtain t_1 from t_2.
3. Let $\delta_e(t_1, t_2)$ denote the number of edges from t_2 not present in t_1.

We don't know how to quickly compute δ_2, but we notice that a good linear approximation is given by δ_e, which, by the way, has a sensible semantic: the more edges in common between two tours, the smaller the distance. When studying a given instance with known global optimum g, we will note: $\Delta_e(t) = \delta_e(g,t)$ and $\Delta_c(t) = \delta_c(g,t)$, the distance of tour t to the optimum.

3 Correlation

A particular notion of *correlation* in a fitness landscape has been recently investigated in order to estimate the difficulty of a given problem or the adequation of neighborhood operators [2, 5, 7, 3, 8, 9]. Basically, this correlation indicates to which extent a relationship exists between a point and its neighbors. Otherly stated, during a walk in the fitness landscape, is it possible to have a probabilistic prediction for the fitness of the next point, from the fitnesses of those encountered previously? Statistical measures have been proposed, usually based on the notion of Box-Jenkins analysis of time series [10]:

- *The autocorrelation function* is a measure of the correlation between the fitness of two points separated by i random steps. This results in a number between -1 and 1. The closer to 1 the absolute value, the larger the correlation between the points.
- *The correlation length of the landscape* indicates the largest distance in number of steps for which there is still a correlation between the start point and the end point.

Experiments have been done on the *TSP*, notably in [3, 9]. The neighborhood was given by applying the 2-opt move or the city swap operators. The results showed that the landscape is more correlated for 2-opt move than for city swap. Furthermore it is well-known that *HC* and *GA* yield better results with the 2-opt move.

One could think that the above remarks settle the problem. However one should notice that there is no clear explanation, in these works, for why a good correlation implies good results for a search method. Indeed, this correlation is measured for *markovian* walks, where each step is made independently of the fitness values encountered, picking up a neighbor at random as the next point of the walk. This is different from the fitness values encountered in a *HC* walk that uses a steepest descent rule, choosing its best neighbor for its next step. It is unclear wether the assumption of isotropy of landscape still holds when comparing such different walks.

Another point worth mentioning is that the correlation length is given in [3, 11] as a function of the size of the instance regardless of any other parameter. On the one hand, we acknowledge that being able to characterize the overall performance of a particular operator for any instance of a given size, is really interesting. On the other hand, there certainly exists instances of the same size for which a given technique (for example 2-opt *HC*) feels more or less "at ease", *i.e.* local optima values are relatively farther from the global optimum value in one instance than in an other. Being able to differentiate between such instances is also important, and this cannot be achieved with a correlation length given as a function of the instance size.

Returning to the *TSP*, we think that if one is interested in finding a global optimum, using fitness values to guide oneself along a *HC* walk, then one has to assume that there exists some correlation between the fitness values of encountered solutions and their distance to the researched optimum. This follows Jones and Forrest's proposition [4] about the importance of fitness-distance correlation (*FDC*), which we also think of as being a determining correlation measure. Obviously, there cannot be a good *FDC* if the correlation between neighbors is not good. Nonetheless *FDC* seems to us more like a necessary condition for the local search to be effective than a simple correlation between neighbors. In the following, we study reference instances with known global optima, so we can compute their *FDC*. We exhibit differences in fitness-distance correlation between city-swap and 2-opt operators, that seems to us related to the behavior of *HC* walks in these landscapes.

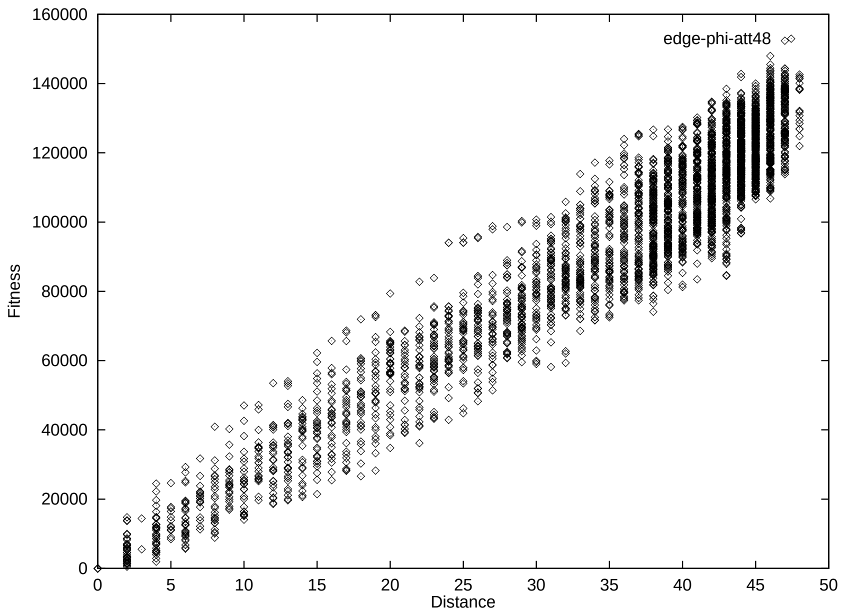


Figure 2: Correlation between Δ_e and ϕ for att48 instance.

4 Correlation measures in the *TSP* Landscape

From several *TSP* instances where an optimal solution is known (such as att48, kroA100, tsp225, pa561, pr1002 — the number at the end of the name is the size of the instance), we study the relationship between Δ_e and fitness. At least 10000 tours are sampled for each instance, with distance from global optimum ranging from 0 to $n-1$. These tours are generated by starting from the global optimum and applying a random mix of 2-opt and city swap, in order to obtain samples that are not biased towards one of these operators. Results are shown in Figure 2.

We can draw a lesson from these experiments: it is quite clear that there is a high degree of fitness-distance correlation in the 2-opt landscape (0.94). This means that the local search algorithm can rely on fitness values encountered in order to get closer to the global optimum. This correlation between Δ_e and ϕ has been observed in all experiments we realized. Furthermore, the more cities in the instance, the higher the correlation between Δ_e and ϕ (*e.g.* 0.99 *for tsp225*). This gives an affirmative answer to a question in [4] about the existence of non artifical problems exhibiting such a structure. We are confident that most symetric, euclidean, planar *TSP* instances have very similar correlation structures.

When plotting the same data with the city swap distance, Δ_c, one obtains clouds with a worse fitness-distance correlation (0.86 *for att48*), as can be seen in Figure 3. Moreover, this value decreases with the instance size : 0.61 *for tsp225.* We think this lack of *FDC* is the main reason that prevents a city-swap *HC* from having as good performance as a 2-opt *HC*.

The huge majority of randomly generated starting tours (*i.e.* random permutations) are situated in the right of Figure 3 pointed out by arrow 1. Starting

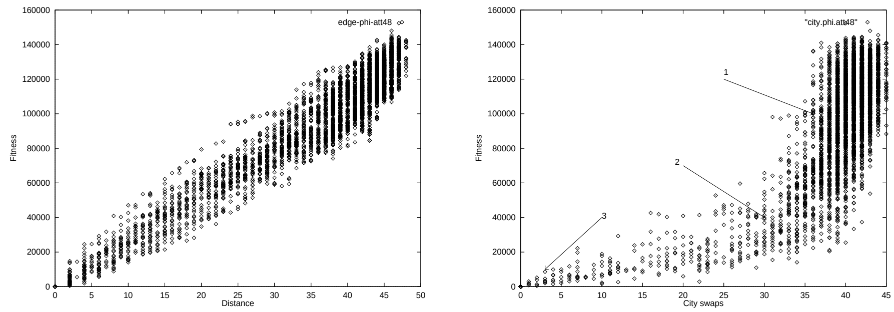
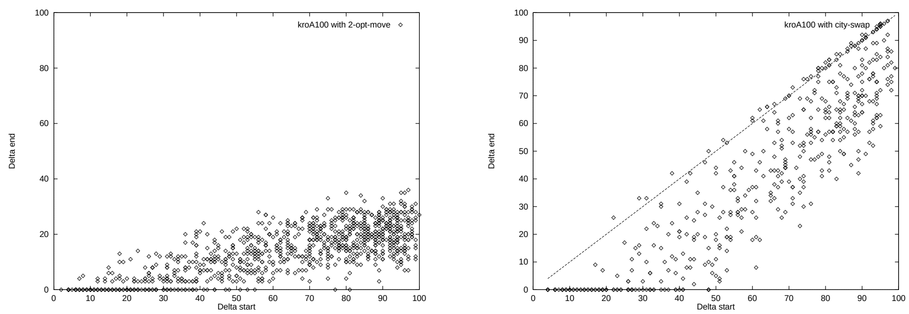


Figure 3: Comparison of the fitness landscape for 2-opt-move (on the left, distance is Δ_e) and city-swap (on the right, distance is Δ_c) for att48 instance.



(a) 2-opt move *HC* runs with distance Δ_e

(b) City-swap *HC* runs with distance Δ_c

Figure 4: Δ(start) against Δ(end)

out with random tours, the search begins at a big distance from the global optimum. The city-swap operator does improve the overall fitness, reaching the region pointed at by arrow 2 on Figure 3. But this fitness improvement does not decrease that much the distance from the global optimum relatively to the Δ_c distance. Furthermore regions pointed at by arrow 2 and 3 are separated by a gentle slope that seems, in our opinion, unable to lead the *HC* towards the optimum. In order to go deeper into this comparison, we have recorded the distance to the global optimum for a set of random start points and their associated local optima obtained after a steepest descent *HC* run. This gives the plots Δ(start) against Δ(end) for city-swap and 2-opt-move *HC* runs in Figure 4. It is clear that the city-swap operator only slightly improves the distance. Even worse, some end points are further in distance than the starting points, as can be seen by comparison with the reference line $y = x$. This behavior can also be observed in Figure 5, where complete walks of *HC* using steepest descent strategy are plotted. The trajectories recorded accentuate the overall shape of clouds, due to the steepest descent rule. Again, differences in final distance to the optimum are clearly visible.

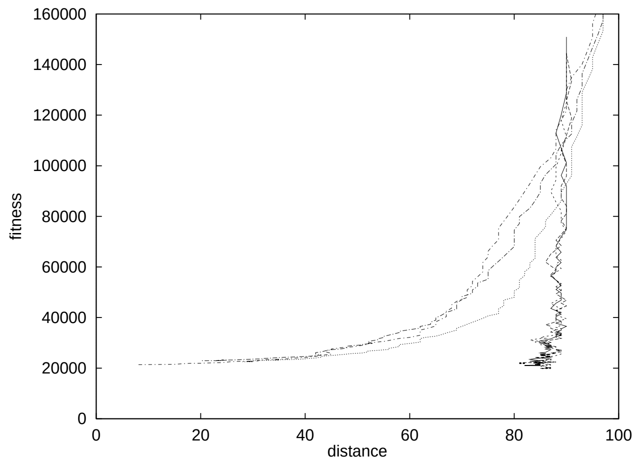


Figure 5: Examples of 2opt and city swap HC walks on tsp225 instance. 2opt HCs searches finish their trajectories on the left, city-swap HCs searches on the right. 2opt trajectories are plotted for Δ_e distance, city swap trajectories are plotted relatively to the Δ_c distance.

5 Conclusion and Future Works

In this paper, we investigate one aspect of the structure of the *TSP* problem. We focus on the behavior of a hill-climbing algorithm using two different operators. The fitness landscape, a metaphor taken from biology, helps us in figuring out the topology of the research space where our algorithms travel. We conduct benchmarks in order to gain a quantitative knowledge about these landscapes and about the trajectories of HC walks.

Studying the results allows us to compare the city-swap and 2opt-move operators on the basis of fitness-distance correlation relatively to known optima. We observe a better FDC for 2opt move than for city swap. We argue that this seems a powerful explanation for the difference in performance between these two well-known operators. We show that a city-swap HC even if it improves the starting points is not able to give tours near the global optimum relatively to its own metric. We observe that 2opt move HC are able to come much closer to the global optimum.

This last point is in accordance with Stadler's conjecture [11] about the existence of a "massif central" regrouping local optima in a area of restricted size. This property could be exploited, as an example by recombining local optima. On the opposite, local optima in the city-swap landscape are separated by long distances.

In future work we will be aiming at:

- implementing intensifying search techniques on the massif central of 2opt local optima;
- verifying if the same ideas are fruitfull on other problems, such as non-symetric *TSP* and job-shop scheduling. This includes the modification of operators for the job-shop scheduling in order to improve their FDC value, and figuring out whether searches benefit from this;

- investigating further the structure of the search space of different kinds of TSP (non-symetric, partial graphs).

Knowledge on fitness landscape should also provide us with clues to help in designing good operators and representation.

References

[1] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In D. F. Jones, editor, *Proceedings of the Sixth International Congress on Genetics*, volume I, 1932.

[2] E.D. Weinberger. Correlated and uncorrelated landscapes and how to tell the difference. *Biological Cybernetics*, 63:325–336, 1990.

[3] P. F. Stadler and W. Schnabl. The landscape of the traveling salesman problem. *Physics Letters A*, (161):337–344, 1992.

[4] T. Jones and S. Forrest. Fitness distance correlation as a measure of problem difficulty for genetic algorithm. In *6th International Conference on Genetic Algorithm*, pages 184–192, 1995.

[5] Bernard Manderick, Mark de Weger, and Piet Spiessens. The genetic algorithm and the structure of the fitness landscape. In *4th International Conference on Genetic Algorithms*, pages 143–149, 1991.

[6] G. Reinelt. Tsplib — a traveling salesman problem library. *ORSA J. Comput.*, 3:376–384, 1991.

[7] Marc Lipsitch. Adaptation on rugged landscapes generated by iterated interactions of neighboring genes. In *4th International Conference on Genetic Algorithms*, pages 128–135, 1991.

[8] W. Hordijk. A measure of landscapes. *Evolutionary Computation*, 4(4), 1996.

[9] B. Krakhofer. Local optima in landscapes of combinatorial optimization problems. Master's thesis, Univ. Vienna, Austria, 1995.

[10] G. E. P. Box and G. M. Jenkins. *Time Series Analysis, Forecasting and Control*. Holden Day, 1970.

[11] Peter F. Stadler. Towards a theory of landscapes. Technical Report SFI-95-03-030, Santa Fe Institute, 1995.