



Lévy flights in neutral fitness landscapes



Marco Tomassini

Department of Information Systems, Complex Systems Group, University of Lausanne, Switzerland

HIGHLIGHTS

- Neutrality is common in biological and artificial evolutionary fitness landscapes.
- Hill-climbing techniques are not effective on regions of equal fitness.
- Neighbor-to-neighbor random drift is useful to escape neutral plateaus and to reach higher fitness ones.
- Lévy flights always outperform random drift on model neutral landscapes NK_p and NK_q .

ARTICLE INFO

Article history:

Received 18 June 2015

Received in revised form 6 November 2015

Available online 29 December 2015

Keywords:

Neutrality

Fitness landscapes

Lévy flights

ABSTRACT

Regions of equal or close fitness are common in biological and artificial evolutionary systems. Customary hill-climbing optimizing paradigms turn out to be unsuitable to walk and search such large neutral networks. Here we propose a new technique to quickly jump out of neutral networks and to reach better fitness regions. The algorithm, based on Lévy flights, is compared to an established nearest neighbors random drift technique on two families of constructive neutral landscapes called the NK_q and the NK_p ensembles. The results of our numerical simulations clearly show that the new algorithm performs better than the nearest neighbors random drift for all studied landscapes. We conclude with some explanations of the observed behavior and some suggestions for the use of Lévy flights in more general search and optimization heuristics.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The fitness landscape metaphor [1] has been a standard tool for visualizing biological evolution and speciation. Fitness landscapes are often depicted as “rugged” surfaces with many local peaks of different heights flanked by valleys of different depth [2,3]. This view is now acknowledged to be only part of the story. In fact, starting in the sixties, it has been cogently argued in the biological literature, especially by M. Kimura [4,5] and Ohta [6], that most genetic change in a population is neutral, i.e. it does not cause significant fitness variations. In this view, the population is seen as evolving by randomly drifting along connected neutral genotype networks of equal, or almost equal, fitness with occasional jumps to nearby neutral networks in the fitness landscape. This view is rather different from the ordinary one in which selective pressure pushes populations towards local maxima of fitness in the landscape. Many confirmations have been found of neutral evolution since its original formulation by Kimura [5]. For instance, Schuster, Stadler, and coworkers have investigated in detail the role of neutrality in RNA folding landscape evolution [7–13] finding that neutrality is an important factor in adaptation. It has also been observed that the huge dimensionality of biologically interesting fitness landscapes, considering the redundancy in the genotype-fitness map, brings naturally the existence of neutral and nearly neutral networks [14]. The relevance and benefits of neutrality for the robustness and evolvability in living systems has been recently discussed in Ref. [15].

E-mail address: marco.tomassini@unil.ch.

Growing evidence has been found that such large-scale neutrality is also present in artificial landscapes. Not only in combinatorial fitness landscapes such as randomly generated SAT instances [16], cellular automata rules [17], and Boolean genetic programming search spaces [18], but also in complex real-world design and engineering applications such as evolutionary robotics [19,20] and evolvable hardware [21–23] among others. The existence of extended neutral networks in these and other fitness landscapes suggests that the search processes on these landscapes are qualitatively different from the dynamics on rugged landscapes [10,21,24–29]. As a consequence, techniques for effective evolutionary search on landscapes with neutrality may be quite different from more traditional approaches to evolutionary search [27,30].

To avoid difficulties related to the chemistry of molecules in biological systems and to facilitate the study of artificial landscapes, it is convenient to use more abstract neutrality models. In this paper we study fitness evolution on synthetic neutral landscapes derived from Kauffman's NK landscapes [3], which are often used as model landscapes with tunable rugosity. We selected two extensions of the NK family as example landscapes with synthetic neutrality, namely the NK_p ('probabilistic' NK) [26], and NK_q ('quantized' NK) [31] families. Two dynamical processes on these landscapes are compared: one in which a walker drifts over neutral networks at constant fitness through one-bit, single locus mutations, and another in which mutation distances are drawn from a power-law distribution and can sometimes affect several loci. We shall show that such a Lévy walker performs better on these neutral landscapes allowing the search to attain higher fitness values in all conditions and providing more innovation capabilities.

In the next section we describe in some detail the neutral families of landscapes under study. Then, after defining the properties of the search processes to be compared, we present and discuss the results obtained by extensive numerical simulations on a large sample of landscapes. This is followed by our conclusions.

2. NK landscapes with neutrality

S. Kauffman [2,3] introduced the NK family of landscapes as a model for constructing multimodal landscapes that can gradually be tuned from smooth to rugged. In the model, N refers to the number of (binary) genes in the genotype (i.e. the string length) and K to the number of genes that influence a particular gene (the epistatic interactions). By increasing the value of K from 0 to $N - 1$, NK landscapes can be tuned from smooth to rugged. For $K = 0$ all contributions can be optimized independently and there is a single maximum. At the other extreme when $K = N - 1$ the landscape becomes completely random, the probability of any given configuration of being the optimum is $1/(N + 1)$, and the expected number of local optima is $2^N/(N + 1)$. Intermediate values of K interpolate between these two cases and have a variable degree of "epistasis", i.e. of gene interaction [3].

The fitness function of a NK -landscape $f_{NK} : \{0, 1\}^N \rightarrow [0, 1)$ is defined on binary strings with N bits. An "atom" with fixed epistasis level is represented by a fitness component $f_i : \{0, 1\}^{K+1} \rightarrow [0, 1)$ associated to each bit i . Its value depends on the allele at bit i and also on the alleles at the K other epistatic positions. (K must fall between 0 and $N - 1$). The fitness $f_{NK}(s)$ of $s \in \{0, 1\}^N$ is the average of the values of the N fitness components f_i :

$$f_{NK}(s) = \frac{1}{N} \sum_{i=1}^N f_i(s_i, s_{i_1}, \dots, s_{i_K}), \quad (1)$$

where $\{i_1, \dots, i_K\} \subset \{1, \dots, i - 1, i + 1, \dots, N\}$. There are two customary ways for choosing the K other bits from N bits in the bit string: adjacent and random neighborhoods. With an adjacent neighborhood, the K bits nearest to the bit i are chosen (the genotype is taken to have periodic boundaries). With a random neighborhood, the K bits are chosen randomly on the bit string. Here we shall use the more common and general random neighborhood. It may be interesting to note that the NK model is related to spin glasses, and more precisely to p -spin models [32,33], where p plays a role similar to K .

The two variants of NK landscapes that we introduce next are representative of the way to obtain neutrality in additive fitness landscapes. Indeed, for the two families, the fitness value of a solution is computed as a sum. Modifying a term in the sum would alter the probability to get the same fitness value.

The NK_p landscapes have been introduced by Barnett [26]. In this variant, one term of the sum is null with probability p . Formally, the fitness components are modified and tuned by the parameter $p \in [0, 1]$ which controls the neutrality of the landscape. The fitness component $f_{s_i, s_{i_1}, \dots, s_{i_K}}^i$ is null with probability p , i.e. $P(f_{s_i, s_{i_1}, \dots, s_{i_K}}^i = 0) = p$. Apart from this, the fitness computation is as in Eq. (1). The probability that two neighbor solutions have the same fitness value increases with the parameter p and thus neutrality increases with increasing p .

The NK_q landscapes have been introduced by Newman and Engelhardt [31]. For these landscapes, the terms of the sum are integer numbers between 0 and $q - 1$. Thus, when some terms are modified, it is possible to get the same sum. Formally, as for NK_p landscapes, the fitness components are defined with a parameter q which tunes the neutrality. Parameter q is an integer number ≥ 2 . Each $f_{s_i, s_{i_1}, \dots, s_{i_K}}^i$ is one of the fractions $\frac{k}{q}$ where k is an integer number randomly chosen in $[0, q - 1]$. In order to have the fitness between 0 and 1 the sum has to be normalized in the following way:

$$f_{NK_q}(s) = \frac{1}{N(q - 1)} \sum_{i=1}^N f_i(s_i, s_{i_1}, \dots, s_{i_K}).$$

Download English Version:

<https://daneshyari.com/en/article/974485>

Download Persian Version:

<https://daneshyari.com/article/974485>

[Daneshyari.com](https://daneshyari.com)