On the Mutual Information as a Fitness Landscape Measure

Rebeka Coric
Josip Juraj Strossmayer University of Osijek
Osijek 31000, Croatia
rcoric@mathos.hr

Stjepan Picek
KU Leuven, ESAT/COSIC and imec
Leuven-Heverlee B-3001, Belgium
stjepan@computer.org

Domagoj Jakobovic
University of Zagreb, Faculty of Elec. Eng. and Computing
Zagreb 10000, Croatia
domagoj.jakobovic@fer.hr

Carlos A. Coello Coello
CINVESTAV-IPN, Department of Computer Science
Mexico D.F. 07360, Mexico
ccoello@cs.cinvestav.mx

ABSTRACT

Fitness landscape analysis plays an important role in both theoretical and practical perspectives when using evolutionary algorithms. In this paper, we develop a new measure based on the mutual information paradigm and we show how it can help to deduce further information about the fitness landscape. In order to validate it as a valuable source of information when conducting fitness landscape analysis, we investigate its properties on a well-known benchmark suite. Moreover, we investigate the usefulness of the obtained information when choosing crossover operators. Finally, we show that when using our new measure, a number of classifiers can be constructed that offer an improved accuracy.

CCS CONCEPTS

•**Computing methodologies → Continuous space search;** *Randomized search;*

KEYWORDS

Fitness landscape, Single-objective, Mutual Information, Performance comparison

ACM Reference format:

Rebeka Coric, Stjepan Picek, Domagoj Jakobovic, and Carlos A. Coello Coello. 2017. On the Mutual Information as a Fitness Landscape Measure. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 2 pages.

DOI: http://dx.doi.org/10.1145/3067695.3076082

1 INTRODUCTION AND BACKGROUND

When dealing with difficult optimization problems, evolutionary algorithms (EAs) have shown their strength a plethora of times. Still, obtaining an optimal solution is a goal that is often difficult to reach which stems from the huge diversity of possible (difficult) optimization problems one can encounter. Furthermore, the "No Free Lunch" theorem states that, informally speaking, when averaged over all optimization problems, all algorithms behave the same [4]. Since we know that we cannot find a single best algorithm for all problems, we should rather turn our attention to finding the best algorithms for certain classes of problems. In this process, the concept of the fitness landscape can play a significant role. Indeed, by checking the properties of the landscape corresponding to the problem at hand, we can gain new knowledge that can be then used to understand and design better suited metaheuristics. However, even by using all existing measures, important characteristics of fitness landscapes are often left unnoticed. Therefore, designing new measures that have a sound theoretical basis, but are also clear on a more intuitive level is an important goal.

The main contribution of this paper is the design of a new fitness landscape measure based on the mutual information paradigm. With this measure, we are able to better distinguish from among some problem classes on the basis of their fitness landscapes. Additionally, we conduct an extensive analysis of the current fitness landscape measures for a multitude of standard benchmarks with a varying number of dimensions. Moreover, we analyze the amount of information obtained from fitness landscapes with respect to the size of the initial population. More precisely, in this paper we:

(1) design new fitness landscape measures based on mutual information,
(2) empirically assess the validity of our measure on a number of problems, problems' dimensions, and sampled fitness landscape sizes,
(3) show that our measures are relatively independent of the size of the sampled fitness landscape, which makes them easier to use in practice, and
(4) show that our measures can help in choosing crossover operators in genetic algorithms for solving different problems.

Regarding the test problems, we use those available from the COCO (COmparing Continuous Optimisers) platform [1], where we investigate the performance of 24 noise-free real-parameter single-objective problems in 30 and 100 variables.

2 MUTUAL INFORMATION LANDSCAPE ANALYSIS

Mutual information is a concept which comes from the probability theory and information theory. It is the measure of the mutual dependence between two variables. More precisely, we can use mutual information to quantify the amount of information (i.e., the entropy of a random variable) obtained from one random variable, by observing the other random variable. Mutual information can

GECCO '17 Companion, Berlin, Germany

DOI: http://dx.doi.org/10.1145/3067695.3076082

Table 1: Classification results (population size: 1 000, dimension: 30).

	Naive Bayes		C4.5		Random Forest		SVM (polykernel)	
	with NMI	without NMI	with NMI	without NMI	with NMI	without NMI	with NMI	without NMI
arithmetic	50.000%	37.500%	54.167%	54.167%	62.500%	45.833%	54.167%	50.000%
average	45.833%	50.000%	54.167%	54.167%	37.500%	62.500%	41.667%	45.833%
BGA	83.333%	45.833%	62.500%	50.000%	83.333%	75.000%	54.167%	54.167%
BLX α	37.500%	33.333%	45.833%	45.833%	45.833%	58.333%	50.000%	54.167%
discrete	25.000%	29.167%	29.167%	25.000%	29.167%	29.167%	37.500%	41.667%
flat	37.500%	33.333%	45.833%	45.833%	45.833%	58.333%	50.000%	54.167%
heuristic	37.500%	33.333%	45.833%	45.833%	45.833%	58.333%	50.000%	54.167%
local	33.333%	45.833%	70.833%	70.833%	41.667%	54.167%	66.667%	66.667%
one point	20.833%	20.833%	29.167%	33.333%	33.333%	33.333%	37.500%	33.333%
SBX	75.000%	62.500%	79.167%	91.667%	70.833%	87.500%	70.833%	75.000%

Table 2: Average Info Gain for fitness landscape measures.

Related measures			NMI based measures					
$d(P)$	$dmm(P)$	δ_{dmm}	AM	GM	HM	VAR	SD	MAD
0.026461538	0.068615385	0.039846154	0.067923077	0.100307692	0.100307692	0.033615385	0.033615385	0.033615385

be formally defined in the following way:

$$MI(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \tag{1}$$

where $p(x, y)$ is the joint probability density function of variables X and Y, and $p(x)$ and $p(y)$ are the marginal probability density functions of X and Y.

If $H(X)$ denotes entropy of variable X, and $H(X, Y)$ the joint entropy of variables X and Y, mutual information can be calculated by using the following expression:

$$MI(X;Y) = H(X) + H(Y) - H(X,Y).$$

Due to the marginal entropies, mutual information is not an invariant measure and, therefore, to increase the strength of the measure, we can also use the normalized mutual information:

$$NMI(X;Y) = \frac{H(X) + H(Y) - H(X,Y)}{H(X,Y)}. \tag{2}$$

With the mutual information measure, for every function we calculate the centroid $C = (c_1, \cdots, c_D)$ of all the vectors in the final population by using the expression $c_j = \frac{\sum_{i=1}^{n} x_i}{n}, j = 1, \cdots, D$. Then, for every vector, the mutual information between that vector and the centroid is calculated. After that the arithmetic mean (AM), geometric mean (GM), and harmonic mean (HM) for the normalized mutual information (NMI) are obtained.

From the obtained values we observe that the normalised mutual information does not significantly depend on the choice of the mean value. Also, conducted experiments show that mutual information doesn't depend on the size of the population. We consider these to be good characteristics, since it enables easier analysis without the need for further tuning with respect to the population sizes.

Estimating Crossover Operator Efficiency. We investigate to see if the new measures can help in classifying crossover operators into good, average or bad in terms of optimizing the selected benchmark functions as done in [2], where a separate classifier is developed for each of the crossover operators. We conducted two sets of experiments. First, a 4-fold cross-validation was used with Naive Bayes, C4.5, Random Forest, and SVM with polykernel classifiers without NMI measure. After that, six new measures are included: arithmetic mean (AM), geometric mean (GM), harmonic mean (HM), variance (VAR), standard deviation (SD), and mean absolute deviation (MAD) of NMI. Then, 4-fold cross-validation is repeated for each of the four mentioned classifiers and crossover operators. The results are given in Table 1. Each row gives a percentage of correctly classified examples for one crossover operator, with and without NMI measure. Additionally, we used InfoGain Ranker to determine whether the new measures are useful for classification [3], as shown in Table 2.

3 CONCLUSIONS

In this paper, we show how mutual information and normalized mutual information can be used as tools in the fitness landscape analysis. Our results show that both of those measures are consistent with respect to the size of the population and the dimensionality of the problem. On the basis of the presented results, we believe that mutual information (both standard and normalized) poses a viable option when investigating fitness landscapes. Still, our results represent only a starting point where more experiments are necessary to determine the role of mutual information in the fitness landscape analysis.

REFERENCES

[1] N. Hansen, S. Finck, R. Ros, and A. Auger. 2009. *Real-Parameter Black-Box Optimization Benchmarking: Noiseless Functions Definitions.* Technical Report RR-6829. INRIA. http://hal.inria.fr/inria-00362633/en/

[2] Stjepan Picek and Domagoj Jakobovic. 2014. From Fitness Landscape to Crossover Operator Choice. In *Proceedings of the 2014 Conference on Genetic and Evolutionary Computation (GECCO '14).* 815–822.

[3] Ian H. Witten and Eibe Frank. 2005. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems).* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[4] David H. Wolpert and William G. Macready. 1997. No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation* 1, 1 (April 1997), 67–82.