



Chapter XVI

On Measuring the Attributes of Evolutionary Algorithms: A Comparison of Algorithms Used for Information Retrieval

J. L. Fernández-Villacañas Martín
Universidad Carlos III, Spain

P. Marrow and M. Shackleton
Intelligent Systems Laboratory, BTextreat Technologies, UK

ABSTRACT

In this chapter we compare the performance of two contrasting evolutionary algorithms addressing a similar problem, of information retrieval. The first, BTGP, is based upon genetic programming, while the second, MGA, is a genetic algorithm. We analyze the performance of these evolutionary algorithms through aspects of the evolutionary process they undergo while filtering information. We measure aspects of the variation existing in the population undergoing evolution, as well as properties of the selection process. We also measure properties of the adaptive landscape in each algorithm, and quantify the importance of neutral evolution for each algorithm. We choose measures of these properties because they appear generally important in evolution. Our

results indicate why each algorithm is effective at information retrieval, however they do not provide a means of quantifying the relative effectiveness of each algorithm. We attribute this difficulty to the lack of appropriate measures available to measure properties of evolutionary algorithms, and suggest some criteria for useful evolutionary measures to be developed in the future.

INTRODUCTION

Evolutionary methods have been the focus of much attention in computer science, principally because of their potential for performing a partially directed search in very large combinatorial spaces (Sloman, 1998). Evolutionary algorithms (EAs) have the potential to balance exploration of the search space with exploitation of useful features of that search space. However the correct balance is difficult to achieve and places limits on what can be predicted about the algorithm's behaviour. In addition, EAs are often implemented in system-specific ways, making it very difficult to predict and evaluate performance on different implementations. This makes the need for measures to evaluate and compare different algorithms all the more urgent.

In this chapter we focus upon the comparison between algorithms for information retrieval. This is one of the tasks at which evolutionary algorithms have been found particularly effective. Such algorithms deal with the situation where a relevant sub-set of documents or records must be isolated from a larger pool. This chapter considers two such algorithms which were developed for the task of information filtering in a telecommunications context. The BTGP is a genetic programming system where the programs produced execute Boolean searches through keywords (Fernández-Villacañas and Exell, 1996). The MGA is a genetic algorithm which also uses a Boolean tree representation, through a relatively complicated mapping between genotype and phenotype.

We compare the performance of these algorithms using a collection of measures chosen from consideration of evolutionary processes. Such measures have been developed within an evolutionary computation context and also within evolutionary biology. To understand why such measures might be useful, we first consider the evolutionary process itself.

Evolution can be described as “...any net directional change or any cumulative change in the characteristics of organisms or populations over many generations ...” (Endler, 1986).

But this evolutionary change may occur as the consequence of a number of different processes, acting to differing extent. Comparison of biological and

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/measuring-attributes-evolutionary-algorithms/6843

Related Content

A Computer-Assisted Diagnostic (CAD) of Screening Mammography to Detect Breast Cancer Without a Surgical Biopsy

Hadj Ahmed Bouarara (2019). *International Journal of Software Science and Computational Intelligence* (pp. 31-49).

www.irma-international.org/article/a-computer-assisted-diagnostic-cad-of-screening-mammography-to-detect-breast-cancer-without-a-surgical-biopsy/247134

Low-Frequency Data Embedding for DFT-Based Image Steganography

Petar Branislav Jeluši, Ante Poljiak, Davor Donevskiand Tomislav Cigula (2022). *International Journal of Software Science and Computational Intelligence* (pp. 1-11).

www.irma-international.org/article/low-frequency-data-embedding-for-dft-based-image-steganography/312558

Twitter Data Analysis Using Apache Streaming

Lavanya Sendhilvel, Kush Diwakar Desai, Simran Adake, Rachit Bisariaand Hemang Ghanshyambhai Vekariya (2023). *Convergence of Big Data Technologies and Computational Intelligent Techniques* (pp. 61-75).

www.irma-international.org/chapter/twitter-data-analysis-using-apache-streaming/314335

Adding Context into an Access Control Model for Computer Security Policy

Shangping Ren, Jeffrey J.P. Tsaiand Ophir Frieder (2007). *Advances in Machine Learning Applications in Software Engineering* (pp. 439-456).

www.irma-international.org/chapter/adding-context-into-access-control/4869

A Value-Based Framework for Software Evolutionary Testing

Du Zhang (2013). *Advances in Abstract Intelligence and Soft Computing* (pp. 355-373).

www.irma-international.org/chapter/value-based-framework-software-evolutionary/72792