GEOMETRIC SEMANTIC GENETIC PROGRAMMING AND ITS REAL-LIFE APPLICATIONS

Leonardo Vanneschi

For all supervised learning problems, where the quality of solutions is measured by a distance between target and output values (error), geometric semantic operators (GSOs) of genetic programming (GP) induce an error surface characterized by the absence of locally suboptimal solutions (unimodal error surface). So, GP that uses GSOs, called geometric semantic GP (GSGP), has a potential advantage in terms of evolvability compared to many other existing computational methods. This fosters GSGP as a possible new state-of-the-art machine learning method. Nevertheless, GSGP is still very young (it was first introduced in 2013) and research is still much in demand. This tutorial is oriented to researchers and students that are not familiar with GSGP, and are willing to deepen and contribute to this exciting and promising field. The first objective of this tutorial is to explain why the error surface induced by GSOs is unimodal, and why this fact is important. This will be done after a soft introduction to the important concept of fitness landscape, and its applications to GP. Then, the tutorial presents some known drawbacks of GSOs and discusses workarounds to limit those drawbacks. In particular, a very efficient implementation of these operators, that makes them usable in practice, is presented. After that, the tutorial focuses on the numerous interesting applicative results that have been obtained so far by GSGP. Among the different real-life problems that have been tackled with success using GSGP, covering several different applicative domains, the tutorial will discuss the prediction of the relative positions of computer tomography slices, the forecasting of energy consumption, the prediction of pharmacokinetic parameters in drug discovery and development, the prediction of the unified Parkinson's disease rating scale assessment and the prediction of high performance concrete strength. While discussing the results that have been obtained on these applications, the audience will also discover that some properties of GSOs that may help limiting overfitting, bestowing on GSGP a very interesting generalization ability. Finally, the tutorial suggests further readings and discusses open issues of GSGP.