
Online modelling based on Genetic Programming

Stephan Winkler*, Hajrudin Efendic
and Luigi Del Re

Institute for Design and Control of Mechatronical Systems,
Johannes Kepler University, Linz, Austria

E-mail: Stephan@HeuristicLab.com E-mail: Hajrudin.Efendic@jku.at

E-mail: Luigi.DelRe@jku.at

*Corresponding author

Michael Affenzeller and Stefan Wagner

Department of Software Engineering,
Upper Austrian University of Applied Sciences,
College of Information Technology at Hagenberg, Austria

E-mail: Michael@HeuristicLab.com

E-mail: Stefan@HeuristicLab.com

Abstract: Genetic Programming (GP), a heuristic optimisation technique based on the theory of Genetic Algorithms (GAs), is a method successfully used to identify non-linear model structures by analysing a system's measured signals. Mostly, it is used as an offline tool that means that structural analysis is done after collecting all available identification data. In this paper, we propose an enhanced on-line GP approach that is able to adapt its behaviour to new observations while the GP process is executed. Furthermore, an approach using GP for online Fault Diagnosis (FD) is described, and finally test results using measurement data of NO_x emissions of a BMW diesel engine are discussed.

Keywords: genetic programming; GP; data driven model identification; self-adaption; machine learning; online modelling; fault diagnosis; FD.

Reference to this paper should be made as follows: Winkler, S., Efendic, H., Del Re, L., Affenzeller, M. and Wagner, S. (2007) 'Online modelling based on Genetic Programming', *Int. J. Intelligent Systems Technologies and Applications*, Vol. 2, Nos. 2/3, pp.255–270.

Biographical notes: Stephan Winkler received his MSc (Dipl.-Ing.) from Johannes Kepler University, Linz, Austria, his research interests include genetic programming, non-linear model identification and machine learning. He is currently a Research Associate within the Translational Research project L284 funded by the Austrian Science Fund (FWF), the institutions involved in the execution of this research project are the Department of Software Engineering at Upper Austrian University of Applied Sciences, Hagenberg, Austria and the Institute for Design and Control of Mechatronical Systems at Johannes Kepler University, Linz, Austria.

Hajrudin Efendic received his MSc (Dipl.-Ing.) in automatic control and electronics and his MSc in telecommunications from the University of Sarajevo, Bosnia and Herzegovina, and he is a Research Associate at the Institute of Design and Control of Mechatronical Systems at Johannes Kepler University. His research interests include control oriented design and monitoring, data-based analysis and fault diagnosis of complex technical systems.

Luigi Del Re is Professor of the Institute of Design and Control of Mechatronical Systems at Johannes Kepler University, Linz, Austria. His main research interests are non-linear control, parameter and structure identification, automotive control, model predictive control and systematic design assessment.

Michael Affenzeller has published several papers and journal papers dealing with theoretical aspects of genetic algorithms and evolutionary computation in general. He is Professor at the Department of Software Engineering, Upper Austrian University of Applied Sciences, Hagenberg. He also holds the position of an Associate Professor at the Institute of Formal Models and Verification at Johannes Kepler University, Linz, Austria.

Stefan Wagner received his MSc (Dipl.-Ing.) from Johannes Kepler University, Linz, Austria, his research interests include applications and theory of evolutionary algorithms, software engineering, parallel and distributed computing, genetic programming, machine learning and data mining. He currently holds the position of an Associate Professor at the Department of Software Engineering, Upper Austrian University of Applied Sciences, Hagenberg.

1 Introduction

The problem of finding a model for a system's measured signal, i.e., discovering the mathematical relationship between empirically observed variables measuring a system, is an important problem in technical fields such as mechatronics, but also in economics and other areas of science (Langley et al., 1987). In practice, the observed data may be noisy and there may be no known way to express the relationships involved in a precise way. Problems of this type are of scientific interest in the context of systems control and data mining; usually they are called symbolic system identification problems, black box problems or modelling problems.

In this paper we present an online structure identification method based on a GP identification approach. This identification algorithm is able to adapt its behaviour during its execution to new data, i.e., to a changing of the algorithm's environment. By doing so, this method combines the advantages of enhanced, hybrid variants of GAs and heuristic optimisation techniques with real-time knowledge discovery and data mining.