Expert Systems with Applications 40 (2013) 3282–3298

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Local and global optimization for Takagi–Sugeno fuzzy system by memetic genetic programming

Athanasios Tsakonas *

Smart Technology Research Centre, Bournemouth University, Talbot Campus, Fern Barrow, Poole BH12 5BB, UK

A R T I C L E   I N F O

Keywords:
Memetic genetic programming
Neuro-fuzzy systems
Context-free grammars
Evolutionary computation
Recursive least squares

A B S T R A C T

This work presents a method to incorporate standard neuro-fuzzy learning for Takagi–Sugeno fuzzy systems that evolve under a grammar driven genetic programming (GP) framework. This is made possible by introducing heteroglossia in the functional GP nodes, enabling them to switch behavior according to the selected learning stage. A context-free grammar supports the expression of arbitrarily sized and composed fuzzy systems and guides the evolution. Recursive least squares and backpropagation gradient descent algorithms are used as local search methods. A second generation memetic approach combines the genetic programming with the local search procedures. Based on our experimental results, a discussion is included regarding the competitiveness of the proposed methodology and its properties. The contributions of the paper are: (i) introduction of an approach which enables the application of local search learning for intelligent systems evolved by genetic programming, (ii) presentation of a model for memetic learning of Takagi–Sugeno fuzzy systems, (iii) experimental results evaluating model variants and comparison with state-of-the-art models in benchmarking and real-world problems, (iv) application of the proposed model in control.

1. Introduction

Among successful approaches in system modeling, the fuzzy paradigm retains a privileged position for its ability to model the qualitative properties of human knowledge and reasoning, without the requirement of complex quantitative processing (Zadeh, 1973). While early models proposed by Mamdani were dealing with classification tasks (Cordon, 2012; Mamdani, 1977), additional research by Takagi and Sugeno led to fuzzy systems where the consequent part was defined as a function of the input space, commonly in the form of polynomials (Takagi & Sugeno, 1985). The use of fuzzy systems under a supervised learning concept by incorporating machine learning methods changed radically their applicability, allowing for automatic, data-driven construction or tuning of fuzzy rule bases. This advance effectively enabled the application of fuzzy systems in regression and system identification tasks. Among the popular training approaches, several *neuro-fuzzy* or heuristic techniques retain a privileged position (Angelov & Filev, 2004; Jang, 1992; Nauck & Kruse, 1999), mainly due to their speed and high generalization performance. Under the generic umbrella term of *fuzzy-evolutionary* methods on the other hand (Cordon, 2012; Gacto, Alcala, & Herrera, 2012; Hoffmann, 2001; Nawa, 1998), focus is often given on benefits arising from robustness and dimensionality reduction. Research has led to many successful paradigms that use global optimization for fuzzy systems. A multi-objective genetic fuzzy system is proposed in Alcalá, Gacto, and Herrera (2011) to deal with high dimensional regression problems and the approach is shown to be statistically superior to comparative models. In Wang, Kwong, Jin, Wei, and Man (2005), a multi-objective hierarchical genetic algorithm is used for the generation of fuzzy systems. A fuzzy clustering method and the recursive least squares method precede to determine an initial fuzzy system optimized thereafter by the genetic algorithm. The model is effectively applied to system identification tasks. A combination of local and global optimization for the generation of Takagi–Sugeno–Kang fuzzy systems augmented with the use of Taylor series as consequent part is demonstrated in Herrera, Pomares, Rojas, Guillen, and Valenzuela (2011). In that work, a clustering algorithm is combined with local search and a heuristic membership function merging approach takes place in order to improve the model transparency.

Genetic programming (GP) is a successful branch of evolutionary computation (Koza, 1992). The representation of a population individual in GP is in the form of a program tree with inner nodes representing functions and outer nodes representing arguments. The main advantage of using GP lies in its ability to evolve variable-length solutions and to represent programs or complex mathematical formulas, effectively enabling *symbolic regression*. Efficient guidance of the structure and evolution of GP by

* Tel.: +44 0 7867938402.
E-mail address: thanos.tsakonas@gmail.com

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.12.099

context-free grammars (Sipser, 1997) makes possible the expression of very complex structures or complete intelligent systems, such as electrical circuits (Koza et al., 1997), feedforward neural networks (Tsakonas, 2006), or hybrid combined systems (Tsakonas & Gabrys, 2012b). The use of grammar guided GP to generate fuzzy rule based systems has been investigated. In Alba, Cotta, and Troya (1999), the production of Mamdani-type output fuzzy rule bases for classification is proposed. In Tsakonas (2006) a similar approach is used in a comparison framework with other GP-derived structures. The production of Takagi–Sugeno fuzzy systems has also been an investigated research topic. In Delgado, Zuben, and Gomide (2004), a genetic algorithm is used for a co-evolutionary system that generates Takagi–Sugeno–Kang (TSK) fuzzy systems. In Chen, Yang, Abraham, and Peng (2007), the production of a specific class of TSK fuzzy systems is examined using evolutionary programming. In Hoffman and Nelles (2001), a GP system is applied to improve a greedy algorithm (LOLIMOT) in finding data clusters by means of partitioning the search space. The application of GP for the production of TSK fuzzy systems is also shown in Tsakonas and Gabrys (2011), where the generation of genetic higher-order Takagi–Sugeno–Kang fuzzy systems is proposed.

The term *meme*, originally coined by Dawkins (1976), describes a cultural element that can be passed on subsequent generations by no-genetic means. The meme concept was included in a unifying theory describing the evolution of complex systems under the term *Universal Darwinism* (Dawkins, 1983). In Moscato (1989) a *memetic algorithm* was defined as a hybrid genetic algorithm bearing additional local search learning. The first generation of memetic algorithms included simple combinations of evolution and local search learning, without however featuring any methods of memetic transmission. The second generation of memetic algorithms addressed memetic transmission, and it refers to systems where the meme is encoded in the genotype of the genetic algorithm. A third generation differs to the second generation in that the memes are not predetermined but they may arise during evolution, to represent frequently repeated patterns. Although most memetic approaches apply local search to all individuals in every generation, it has been shown that this is not necessary (Krasnogor & Smith, 2005). In Hart (1994), selective local search is applied and two adaptive methods for applying local search are examined: in *distribution-based adaptive* methods, redundancy in the population is used to avoid unnecessary local search; in *fitness-based adaptive* methods, information of the population fitness is used to evoke local search in more fitted individuals. In Land (1998), the application of local search to only optimal individuals is proposed and variable ratios between local and evolutionary search are tested. The use of local searchers that modify their behavior according to the convergence of the population is also proposed in Krasnogor and Pelta (2002). In Nguyen, Ong, and Lim (2009), a comprehensive empirical study was presented that models the memetic algorithm as a process involving the decision of embracing separate actions of evolution or individual learning.

There are several approaches that effectively incorporate local search in GP. A GP framework is used in Meuth (2010) to incorporate meme-based learning, or meta-learning, is proposed and examined in a problem of autonomous agents in changing environments. A memetic genetic programming approach that incorporates a decision-tree based operator in a GP framework is proposed in Wang et al. (2005). The model is shown to improve classification scores over three GP variants. Gradient descent learning was shown in Smart and Zhang (2004) to improve GP in object recognition tasks. In that work, the hybrid model outperformed the purely evolutionary model in all examined cases. Local search is also incorporated to GP in order to improve its fundamental learning properties. In Topchy and Punch (2001), a model of GP that incorporates gradient descent local search to tune the random constants of the tree is shown to converge in less generations. Linear regression is used in Topchy and Punch (2003) to improve the performance of GP using interval arithmetic and linear scaling aiming to overcome problems from the protected operators.

Although the aforementioned systems have demonstrated the synergy between GP and local search, at the best knowledge of the authors, there is not any methodology defined which would describe the principles for coupling the evolutionary learning for a GP-generated intelligent system with its local-search learning equivalent. The potential however of such a synergy is high: a model using both learning techniques could combine the desirable properties of local search, such as improved generalization and efficiency, on top to the ones of evolutionary search, such as unbiased, automatic design and configuration and avoidance of local minima. This gap is addressed in this work, by proposing an enriched functional set for GP, which supports *heteroglossia* (Dentith, 1994) and it allows both learning methods (i.e. evolutionary and local search) to co-operate obliging the functional nodes to switch behavior. The method is demonstrated in a novel model, named *MEMFIS* (MEMetic genetic programming Fuzzy Inference System). MEMFIS incorporates local-search training methods for fuzzy systems, namely backpropagation gradient descent for the linguistic variables and recursive least squares for the coefficients of the linear functions. In addition to local search, complete fuzzy rule bases evolve under a GP framework with a context-free grammar guidance. MEMFIS effectively generates and trains first-order Takagi–Sugeno fuzzy systems for regression and control. Since the incorporation of local search at all times and to all population individuals can be computationally expensive and it is not considered necessary (Krasnogor & Smith, 2005), we propose a simple parametric model for applying local search in a variable rate. This rate is inspired by the exponential growth seen in cultural outbreaks (Dawkins, 1993) and it is tested in terms of generalization performance and efficiency with constant rate counterparts.

The paper is organized as follows. Section 2 describes the architecture of MEMFIS, its grammar and functional elements and the training process together with an insightful example. In Section 3 we compare MEMFIS with several variants and other machine learning methods in regression and time-series forecasting tasks. In the same section we also generate a fuzzy controller for the inverted pendulum problem. Section 4 includes a discussion on our results. The paper concludes with Section 5 proposing further lines of research.

2. MEMFIS

2.1. Overview

An instance of local-search learning synergy with GP-generated intelligent structures, MEMFIS automatically generates and trains complete Takagi–Sugeno fuzzy rule bases. Coupling local search with evolutionary optimization in MEMFIS attempts to combine the following properties:

- Ability to represent the knowledge in a set of fuzzy constraints, inherited through the fuzzy system modeling.
- High generalization performance by local search methods. These methods for MEMFIS consist of backpropagation for the fuzzy partitions and least squares optimization for the coefficients of the linear functions in the consequent part of the fuzzy rules.
- Robustness of evolutionary search. Training fuzzy systems using an evolutionary approach is expected to reduce user-imposed design bias, prevent premature convergence to local minima solutions and to deal well with dimensionality issues.