Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Simulation Modelling Practice and Theory 15 (2007) 1128–1145
SIMULATION MODELLING PRACTICE AND THEORY
www.elsevier.com/locate/simpat

A genetic approach to modeling fuzzy systems based on information granulation and successive generation-based evolution method

Keon-Jun Park [a], Witold Pedrycz [b,c,*], Sung-Kwun Oh [a]

[a] *Department of Electrical Engineering, The University of Suwon, San 2-2, Wau-ri, Bongdam-eup, Hwaseong-si, Gyeonggi-do 445-743, South Korea*
[b] *Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G 2G6*
[c] *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Received 3 March 2007; received in revised form 23 June 2007; accepted 4 July 2007
Available online 13 July 2007

Abstract

In this paper, we introduce and investigate a new category of fuzzy inference systems based on information granulation and genetic optimization used to system identification. We show the applications of such systems to identification of nonlinear systems. The formal framework of information granulation and resulting information granules themselves become an important design facet of the fuzzy models. By embracing fuzzy sets, the model is geared towards capturing essential relationship between information granules rather than concentrating on plain numeric data. Information granulation realized with the use of the commonly exploited C-Means clustering helps determine the initial values of the parameters of the fuzzy models. This in particular concerns such essential components of the rules as the initial apexes of the membership functions standing in the premise part of the fuzzy rules and the points of the polynomial functions standing in the consequence part. The initial apexes (center points) of the membership functions based on C-Means algorithm are tuned with the aid of the genetic algorithm (GA), while the tuned apexes are also used to adjust the points of the consequent polynomials (conclusions) of the rules. In particular, the initial apexes of the membership functions and the initial points of the consequent polynomials are adjusted and updated every time through successive evolution process. The overall design methodology involves a hybrid structural and parametric optimization. Genetic algorithms and C-Means clustering are used to optimize the model with respect to its structure and parameters. To determine the structure and estimate the values of the parameters of the fuzzy model we consider the successive tuning method with generation-based evolution by means of genetic algorithms. The model is evaluated with the use of numerical experimentation and its quality is compared with respect to some other fuzzy models already encountered in the literature.

Keywords: Information granules (IG); Fuzzy inference system (FIS); C-Means clustering; Successive tuning method; Generation-based evolution; Genetic algorithms (GAs)

* Corresponding author. Address: Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G 2G6. Tel.: +1 780 492 3332; fax: +1 780 492 1811.
E-mail address: pedrycz@ee.ualberta.ca (W. Pedrycz).

1569-190X/$ - see front matter
doi:10.1016/j.simpat.2007.07.001

1. Introductory remarks

Fuzzy sets and fuzzy modeling proposed by Zadeh [1] have been widely investigated. In recent years, fuzzy systems have demonstrated their usefulness as a vehicle of system modeling. There has been a diversity of approaches to fuzzy modeling. To enumerate a few representative trends, we refer to some developments that have occured over time. In the early 1980s, linguistic modeling [2,3] and fuzzy relation equation-based approach [4,5] were proposed as dominant identification methods of fuzzy models. In the linguistic approach, Tong identified a gas furnace process by means of a logical examination of data [6]. Xu reported interesting results obtained through the use of modified Tong's method [7] and proposed an algorithm for a construction of an adaptive model based on decision tables. The main drawback of the method became apparent when dealing with high-order multivariable systems [8] where various issues of memory requirements and computing time started to become serious stumbling blocks. Pedrycz introduced an idea of identification of fuzzy systems realized within a formal framework of fuzzy relation equations. The proposed methodology dwelled on a concept of referential fuzzy sets regarded as generic modeling landmarks [3]. Xu and others presented a self-learning algorithm for a simple SISO fuzzy model [8]. Considering a fuzzy relation equation-based approach, Pedrycz identified fuzzy systems, using the referential fuzzy sets and conditional possibility distributions [4]. Xu constructed and identified the fuzzy relations of a model using referential fuzzy sets [7,8]. The general class of Sugeno–Takagi models [10] gave rise to more sophisticated rule-based systems where the rules come with conclusions forming local regression models. While being appealing with respect to the underlying topology (by supporting a modular fuzzy model being composed of a series of rules) [9,11], these models still await formal solutions as far as their structure optimization is concerned, say a construction of the underlying fuzzy sets – information granules being viewed as basic building blocks of any fuzzy model. Some enhancements to the models have been proposed by Oh and Pedrycz [12,23,24,27,28], however the problem of determining a suite of "suitable" initial parameters of the fuzzy sets in the rules remained open.

This study concentrates on a fundamental problem of fuzzy modeling that is a development of information granules-fuzzy sets and a formation of a general methodology aimed at the identification of structure and estimation of parameters of the fuzzy model. Taking into consideration the essence of the granulation process, we propose to cast the problem in the setting of clustering techniques and genetic algorithms. The design methodology emerges as a hybrid structural optimization (based on the C-Means clustering and genetic optimization) and parametric optimization (based on the least square error method, LSM, as well as the clustering mechanisms and genetic optimization). Information granulation realized with the aid of the C-Means clustering helps determine the initial parameters of fuzzy model such as the initial apexes of the membership functions forming the premise part of the fuzzy rules and the initial values of polynomial function forming the consequence part. The initial parameters are tuned (adjusted) in an effective manner with the aid of the genetic algorithms and the standard least square method. To optimize the structure and parameters of fuzzy model, we exploit the successive tuning method of generation-based evolution by means of genetic algorithms. The proposed model is evaluated through intensive numeric experimentation.

2. The mechanism of information granulation

Roughly speaking, information granules [13–16] are sought as collections of objects (data, in particular) brought together by the criteria of proximity, similarity, or functionality. Granulation of information is an inherent and omnipresent activity of human beings carried out with intent of gaining a better insight into a problem under consideration and arriving at its efficient solution. In particular, granulation of information is aimed at transforming the problem at hand into several smaller and therefore more manageable tasks. In this way, we partition the task into a series of well-defined subproblems (modules) of a far lower computational complexity than the original one. The form of information granulation (IG) itself becomes an important design feature of the fuzzy model given the fact that fuzzy models are geared toward capturing relationships between information granules. The key characteristics of experimental data being used in the construction of the fuzzy model are carefully reflected by fuzzy rules used therein.

Clustering is often regarded as a synonym of information granulation. The intent of clustering is to find a structure in the data and reveal clusters – information granules in the data set. The clustering algorithms have