Neurocomputing 135 (2014) 203–217

Contents lists available at ScienceDirect

Neurocomputing

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

A new method for designing neuro-fuzzy systems for nonlinear modelling with interpretability aspects

K. Cpałka **, K. Łapa *, A. Przybył, M. Zalasiński

Częstochowa University of Technology, Institute of Computational Intelligence, Poland

ARTICLE INFO

Article history:
Received 7 August 2013
Received in revised form
12 December 2013
Accepted 13 December 2013
Communicated by R. Tadeusiewicz
Available online 21 January 2014

Keywords:
Nonlinear modelling
Neuro-fuzzy systems
Evolutionary algorithms
Hierarchical clustering
Interpretability

ABSTRACT

In this paper we propose a new approach to nonlinear modelling. It uses capabilities ofthe so-called flexible neuro-fuzzy systems and evolutionary algorithms. The aim of our method is not only to achieve appropriate accuracy of the model, but also to ensure the possibility of interpretability of the knowledge within it. The proposed approach was achieved by, among others, appropriate selection of operational criteria applied to evolutionary model creation. It allows to extract interpretable fuzzy rules in the cases which use the learning data e.g. from identification. The possibility of interpretation of knowledge accumulated in the model seems to be important in practice, because it guarantees operation predictability and facilitates production of efficient and accurate control methods. Our method was tested with the use of well-known simulation problems from the literature.

1. Introduction

The analysis of technical issues aims at finding and understanding the essence of the problem, it tries to create a model. The reason for this is the willingness to ensure predictability, which guarantees safety, decreases costs and ensures control. It aims not only to have an accurate model, which is able to work in the real-time, but also the model which is interpretable (data mining). The knowledge of the model facilitates designing efficient and accurate controllers for the modelling process.

In the literature the following approaches to modelling are considered:

- *White-box model*: This approach uses phenomenological (theoretical) description of physical phenomena. In the cases for dynamic modelling it is presented as differential equations. Most commonly algebraic form is implemented with the use of theory of state variables. It is well known issue in control theory [11]. It is worth mentioning that the phenomenological model is interpretable but not necessarily accurate enough. It stems from simplifying assumptions or insufficient knowledge of the modelled phenomena. An example of this are models of electromechanical drives [45]. In such implementations the simplifying assumptions apply mainly to (a) assume symmetry and linearity, (b) idealization of actors and sensors characteristics, and (c) neglecting saturation of the magnetic circuit. In order to improve the quality of such a model the compensation of influence of selected phenomena (e.g. nonuniform magnetic circuit) or extending mathematical description by the description of physical phenomena in the cooperating components are proposed [4].
- *Black-box model*: In this approach the behaviour of the object is recreated on the basis of observations of cause and effect dependencies. Parameters of standard (usually very complex) model are tuned to data derived from observation of the object. In such a case it is theoretically possible to obtain high accuracy of the model. It is worth mentioning that in this type of modelling the interpretation of the model is often impossible. It stems from the characteristics of methods which support this type of modelling. These methods include, among others, systems of computational intelligence such as neural networks [21,41,42,55,58]. Neural networks have the ability to learn based on data derived from observation of the object. Unfortunately, the knowledge accumulated in the neural networks is non-interpretable.
- *Grey-box model*: This approach is based on model structure derived from some laws and parameters tuned to the data defining behaviour of the object. This aims at ensuring that the mapping model creates possibility of interpretability of the knowledge accumulated within the model. These methods include, among others, hybrid solutions and systems of computational intelligence such as fuzzy systems and neuro-fuzzy

* Corresponding author.
** Principal corresponding author.
E-mail addresses: krzysztof.cpalka@iisi.pcz.pl (K. Cpałka), krystian.lapa@iisi.pcz.pl (K. Łapa), andrzej.przybyl@iisi.pcz.pl (A. Przybył), marcin.zalasinski@iisi.pcz.pl (M. Zalasiński).

0925-2312/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.12.031

systems. Special attention should be paid to neuro-fuzzy systems which will be considered in this work. These systems unite the advantages of both neural networks and fuzzy systems [15,30,49,52,54,60]. The knowledge accumulated in the fuzzy systems is described by if–then rules, formulated by an expert, which are easy to read. The neuro-fuzzy systems unite the ability to learn the neural networks with the possibility of easily read representation of the knowledge of the fuzzy systems. Thanks to this combination the neuro-fuzzy systems are perfect for nonlinear modelling. It is worth mentioning that only the use of neuro-fuzzy system for nonlinear modelling does not ensure interpretability of the knowledge accumulated within it. The reason for that is (a) use of a large number of fuzzy rules, (b) use of a large number of antecedents and consequents of the rules, (c) overlapping of the fuzzy sets, etc. As already mentioned the grey box model also includes the hybrid solutions. These solutions combine e.g. fuzzy sets theory and fuzzy systems theory with the state variables theory (for dynamic modelling). In such correlation the following approaches are known: (a) approaches based on implementation of sector-nonlinearity method and (b) approaches based on nonlinearity of selected parameters of linear model (described by fuzzy rules). The first approach identifies sectors which are the basis for local linear object approximation in different operating points. Then states variable technique describes (locally linear) behaviour for the whole object in the selected sector. This description is based on well known methods of the control theory which applies to linear models (it is their big advantage). It should also be mentioned that the interpretation of the knowledge accumulated in such a model is difficult. Even the use of fuzzy rules does not change this fact [25]. If–then fuzzy rules describe only the way of switching between linear models while the reason for switch of the sector is hard to substantiate. In the second approach the interpretation of the accumulated knowledge is a bit easier. It stems from the fact that the if–then fuzzy rules describe the change of the values of the selected coefficients of the linear model in reaction to change in the input data values, not the change of the whole model [3,34,44].

There is still a search for such nonlinear modelling methods which will be characterized by good accuracy and possibility to interpret the knowledge accumulated within it. The interpretability issue in the context of nonlinear modelling is much harder than in the case of classification (in the system the exact value of the output signal is important). Each limitation put upon the system structure (used to increase the interpretability) has a negative effect on the accuracy (and vice versa). On the other hand, the attempt to directly interpret the accumulated knowledge, for example in the case of fuzzy system without techniques to increase interpretability, is difficult. In the literature there are different approaches to increase the interpretability of fuzzy systems. It can be noted that these approaches are mainly based on a suitable structure of the fuzzy system (e.g. a hybrid structure) or on the use of specific training algorithm (e.g. methods in the field of multiobjective optimization or evolutionary optimization) (see e.g. [1,5,16,22–24,36,46,53,56,62]). It seems that the approach used to increase interpretability should be simple to implement and subsequent to modify, and should allow the use of dynamic programming techniques.

The accuracy versus interpretability trade-off is also a common topic in the literature [19,62]. An interesting approach for nonlinear modelling is the use of potential of neuro-fuzzy systems and extortion of interpretability of the knowledge accumulated within it. The interpretability of the neuro-fuzzy system is defined in many ways [5,35,37,62], in most cases it is defined as interpretability of fuzzy partitions, also known as integrity or similarity, and interpretability of rules, also known as complexity. In [14] a taxonomy based on a double axis 'complexity versus semantic interpretability′ considering the two main kinds of measures was presented; and 'rule base versus fuzzy partitions′ considering the different components of the knowledge base to which both kinds of measures can be applied. This systematics assumes four quadrants of the interpretability of fuzzy-rule based systems:

- Q_1: Complexity at the rule base quadrant. This quadrant includes, among others, number of rules, number of conditions.
- Q_2: Complexity at the fuzzy partition quadrant. This quadrant includes, among others, number of membership functions, number of features.
- Q_3: Semantics at the rule base quadrant. This quadrant includes, among others, consistency of rules, rules fired at the same time, transparency of rule structure.
- Q_4: Semantics at the fuzzy partition quadrant. This quadrant includes, among others, completeness or coverage, normalization, distinguish ability, complementarily, relative measures.

As mentioned before, considering the interpretability in the cases connected to nonlinear modelling is not easy. It is due to the fact that it requires specified assumptions (e.g. resulting from considered Q_1–Q_4 quadrants) where each can affect the accuracy in a negative way.

The population based algorithms are a convenient tool used for learning neuro-fuzzy systems in nonlinear modelling tasks. They exist in many variants based on behaviour of different populations, for example ants, bacterial, birds. It stems from their implementation flexibility and simplicity, and effectiveness. Population based algorithms are used for example for system parameters′ selection [27,32], system structure selection [17], system structure reduction [23,24,56]. These algorithms also allow us to implement mechanisms which ensure system′s interpretability [7,16,19].

Using the population based algorithms regarding the learning neuro-fuzzy systems in nonlinear modelling tasks, it is very important to pay attention to the initial population generation [2,46]. The initialization cannot be random, as in the case of other kind of problems solved with the use of population based algorithms. It stems from the fact that some groups of genes within chromosome are connected (for example, genes which code parameters of membership functions) and depend on each other (for example, genes which code fuzzy sets with singular input or output). Then the initialization is based most often on the capabilities of unsupervised learning and allows among others to initially select rules of the systems and their number. If it is conducted correctly, it shortens learning time of the system and enhances system′s quality in the meaning of selected criteria. The appropriate initialization of population based algorithm is not simple. It stems from the fact that the aim of the initialization algorithm is selection of the whole population, not a single system. The diversity of population needs to be ensured.

In this paper we propose a new approach to nonlinear modelling. It includes a new system initialization algorithm and the new method of neuro-fuzzy system structure and parameters′ selection, aimed at obtaining both high accuracy and high interpretability in accordance to the Q_1–Q_4 criteria presented in [14]. The method we propose

- *Uses the potential of the flexible neuro-fuzzy systems for nonlinear modelling*, proposed in our previous works [8,48,50]. These systems have a high accuracy, which is very important in the nonlinear modelling cases. These systems also allow us to enter a hierarchy of antecedents of rules and whole rules. It seems to