Information Sciences 257 (2014) 301–312

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

Adaptability, interpretability and rule weights in fuzzy rule-based systems

Andri Riid [a,*], Ennu Rüstern [b]

[a] Laboratory of Proactive Technologies, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia
[b] Department of Computer Control, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

ARTICLE INFO

Article history:
Available online 16 January 2013

Keywords:
Fuzzy modeling
Fuzzy control
Classification
Interpretability of fuzzy systems

ABSTRACT

This paper discusses interpretability in two main categories of fuzzy systems – fuzzy rule-based classifiers and interpolative fuzzy systems. Our goal is to show that the aspect of high level interpretability is more relevant to fuzzy classifiers, whereas fuzzy systems employed in modeling and control benefit more from low-level interpretability. We also discuss the interpretability–accuracy tradeoff and observe why various rule weighting schemes that have been brought into play to increase adaptability of fuzzy systems rather just increase computational overhead and seriously compromise interpretability of fuzzy systems.

1. Introduction

Fuzzy rule-based systems can be divided into two main categories. In the fields of modeling and control usually **interpolative fuzzy systems** are employed because what we need is a continuous output variable. The inference algorithm of interpolative fuzzy systems is based on rule interpolation, hence the name.[1] In classification, on the other hand, the principal task of a **fuzzy rule-based classifier** is just to assign a class label (the number of which is limited) to the sample presented to it. The inference algorithm relies on rule competition, rather than on cooperation.

Interpretability of fuzzy systems – an ability to explain the behavior of the system in an understandable way – has attracted many researchers in recent years (see e.g. [2–5,9–11,14,17,21,30,31,33,37,39–41,44,50,45,65]). While the ultimate definition of interpretability with all its implications (interpretability measures and requirements) is still underway [17], a taxonomy of fuzzy system interpretability into low-level and high-level interpretability has been firmly established [17,65]. Low-level interpretability issues can be tracked down to fuzzy partition properties such as normality, coverage, convexity, distinguishability and complementarity. High-level interpretability, on the other hand, is associated with rule base properties and in many studies, high-level interpretability improvement essentially boils down to complexity reduction (i.e. reducing the number of variables, rules and conditions per rule). Some recent studies have included further semantic issues to deal with [6,18,43].

Our previous research [51,55] has mostly focused on low-level interpretability aspects in interpolative fuzzy systems under the label of fuzzy system **transparency**.

In the present paper we take a step further from [57] and show that these two levels of interpretability are not equally relevant to the two categories of fuzzy rule-based systems. In fuzzy interpolative systems, the high level of rule interaction dictates that partition properties are a primary concern from interpretability viewpoint and complexity reduction is a

* Corresponding author.
E-mail addresses: andri@dcc.ttu.ee (A. Riid), ennu.rystern@dcc.ttu.ee (E. Rüstern).

[1] Not to be confused with an alternative concept of fuzzy interpolative systems [34].

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2012.12.048

background issue (Section 2). This is because fuzzy models and controllers cannot usually have a large number of variables without falling prey to the curse of dimensionality.

On the other hand, it can be shown that due to specific characteristics of the inference algorithm in classification, low-level interpretability requirements appear in a milder formulation. Interpretability improvement is therefore largely a matter of finding a small number of concise fuzzy rules with a limited number of conditions per rules while preserving a satisfying balance between interpretability and accuracy (interpretability–accuracy tradeoff). The roots of this approach date back to the 1990s [25,26] and we revisit its main aspects in Section 4.

We also discuss adaptability of fuzzy rule-based systems. Some interpolative fuzzy systems, such as Mamdani systems greatly benefit from product–sum inference as it provides an analytical expression of the inference function and permits us to apply computationally efficient methods for the identification of consequent parameters (Section 3). In classification, rule weights are often introduced to improve the classification rate [15,24,29,32,38,47,66] and they are sometimes considered as an improvement to the way in which the rules in fuzzy interpolative systems interact [12,48,63]. Our purpose, however, is to show that in both fuzzy system categories (see Sections 5 and 6, respectively), the rate at which rule weights contribute to adaptability is often overestimated and their true identity is usually overlooked.

2. Interpolative fuzzy systems

It is generally acknowledged that of the two prevailing types of interpolative fuzzy systems, Mamdani systems are inherently more interpretable than Takagi–Sugeno (TS) systems.[2] This is because Mamdani systems provide a better (more intuitive) mechanism for the integration of expert knowledge into the system as fuzzy rules in Mamdani systems closely follow the format of natural languages and deal with fuzzy sets exclusively. These rules are based on the disjunctive rule format

$$\begin{aligned}&\text{IF } x_1 \text{ is } A_{1r} \text{ AND } x_2 \text{ is } A_{2r} \text{ AND } \ldots\\ &\ldots \text{ AND } x_N \text{ is } A_{Nr} \text{ THEN } y \text{ is } B_r\\ &\qquad\qquad \text{OR } \ldots,\end{aligned} \tag{1}$$

where A_{ir} denote the linguistic labels of the ith input variable associated with the rth rule ($i = 1, \ldots, N$; $r = 1, \ldots, R$), and B_r is the linguistic label of the output variable, associated with the same rule.

Each A_{ir} has its representation in the numerical domain – the membership function μ_{ir} (the same applies to B_r represented by γ_r) and in a general case the inference function that computes the fuzzy output $F(y)$ of the system (1) has the following form

$$F(y) = \bigcup_{r=1}^{R}\left(\left(\bigcap_{i=1}^{N}\mu_{ir}(x_i)\right)\cap\gamma_r\right), \tag{2}$$

where $\cup_r^R$ denotes the aggregation operator (corresponds to OR in (1)), $\cap$ is the implication operator (THEN) and $\cap_i^N$ is the conjunction operator (AND). In order to obtain a numerical output, (2) is generally defuzzified with the center-of-gravity method

$$y = Y_{cog}(F(y)) = \frac{\int_Y yF(y)dy}{\int_Y F(y)dy}. \tag{3}$$

In the following, the activation degree of the rth rule – the result of the conjunction operation in (2) – is denoted as

$$\tau_r = \bigcap_{i=1}^{N}\mu_{ir}(x_i). \tag{4}$$

In a normal Mamdani system the number of membership functions (MFs) per ith variable (S_i) is relatively small – this number is rarely equal to R as the notation style in (1) implies, moreover, for the sake of coverage it is often desired that all possible unique combinations of input MFs are represented $\left(R = \prod_{i=1}^{N} S_i\right)$. MFs of the system are thus shared between the rules and a separate $R \times N$ dimensional matrix that accommodates the identifiers $m_{ri} \in \{1,2,\ldots,S_i\}$ maps the existing input MFs μ_i^s to the rule slots. Unless unique output MFs are assigned to all rules, they also need some external allocation mechanism.

According to [55], Mamdani systems are subject to transparency constraints to ensure low level interpretability.

The most convenient way to satisfy the transparency constraints for the input MFs μ_i^s $(s = 1,\ldots,S_i;\ i = 1,\ldots,N)$ is to use the following definition:

$$\mu_i^s(x_i) = \begin{cases} \frac{x_i - a_i^{s-1}}{a_i^s - a_i^{s-1}}, & a_i^{s-1} < x_i \leqslant a_i^s,\\ \frac{a_i^{s+1} - x_i}{a_i^{s+1} - a_i^s}, & a_i^s < x_i < a_i^{s+1},\\ 0, & \text{otherwise} \end{cases} \tag{5}$$

[2] First and higher order TS systems are not considered in this paper but reader may refer to a related study [53].