Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems ••• (••••) •••–•••

FUZZY sets and systems

www.elsevier.com/locate/fss

Some new qualitative insights into quality of fuzzy rule-based models

Jeremy Kerr-Wilson [a,*], Witold Pedrycz [a,b,c]

[a] *Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, T6R 2V4, Canada*
[b] *Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, 21589, Saudi Arabia*
[c] *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Received 19 February 2014; received in revised form 21 April 2016; accepted 2 May 2016

Abstract

Rules in fuzzy rule-based models convey essential knowledge about the system under discussion. As such, they capture the essence of relationships occurring among input and output variables. While the quality of such fuzzy models is predominantly related with the accuracy and eventual interpretability of rules (although to a limited extent), the quality of rules being regarded as generic pieces of knowledge has not been studied so far. In this study, we formulate and investigate this problem by looking at the quality of rules, including aspects of (a) stability, (b) generalizability, and (c) conflict. We identify a concept of rule multiplicity, conflict, and study an emergence of rule generalization. A number of pertinent performance indices are developed, and their usage is presented through a series of experimental studies.

Keywords: Fuzzy systems; Fuzzy models; Interpretability; Stability; Consistency; FCM; Fuzzy rule; Data-driven; Rule quality; Machine learning

1. Introduction

Fuzzy rule-based models have been around for a number of decades, almost from the inception of fuzzy sets. To a significant extent, they are viewed as being synonymous with fuzzy models. Over the course of time, we have witnessed many major developments and visible trends in their design and analysis. The Takagi–Sugeno fuzzy model [18] is one such architecture which has gained notable popularity, and remains commonly considered to this day, being used in diverse areas of system modeling [4,8,14].

At the beginning of the era of fuzzy modeling, the primary focus was on the development of different model designs; however, over time, the accuracy of fuzzy models became a key design criterion, while the interpretability, transparency, and comprehensibility of models started to play a less visible role. Recently, some very sophisticated methodologies have been proposed, with the aim of improving the interpretability of complex fuzzy systems through

* Corresponding author.
E-mail addresses: kerrwils@ualberta.ca (J. Kerr-Wilson), wpedrycz@ualberta.ca (W. Pedrycz).

http://dx.doi.org/10.1016/j.fss.2016.05.002
0165-0114/

rule reduction and optimization [10]. While such methods prove effective, the effort required is daunting for many tasks. Other studies have examined the tradeoffs between rule interpretability and accuracy [12,15], with some focusing on the human readability of rules and linguistic quantifiers [2].

Zhou et al. [22] extensively explore methodologies for the creation of fuzzy models, taking particular note of the differences between low-level versus high-level model interpretability, noting that many studies concentrate on just one of these aspects, while neglecting the other. Of particular interest, [22] considers fuzzy models created through rules generated by Fuzzy C-Means (FCM) to be the closest of current methods to constructing both low and high level interpretable models; however, FCM is still lacking with regards to linguistically interpretable rules, in that rule consistency and completeness are not guaranteed and rules may not be very distinguishable.

While the improvement of model accuracy and the development of new modeling algorithms is still an active area of study [4,8,13,14,19,23], there have been fewer recent developments with regards to the consistency or stability of rules. Some efforts have been made, such as those discussed in [6], which explores measures of rule consistency and completeness to be used in fuzzy rule evaluation. Pancho et al. [16] recently defined a quadrant of interpretability indices, stating that those indices dealing with rule consistency, co-firing of rules, and rule transparency contain the least number of works in current literature. Of those studies examining interpretability, many focus on the comprehensibility and linguistic representation of rules; these evaluation measures are quite subjective and focus on the human readability of rules, rather than on rule quality.

There is another interesting and fundamental facet of rule-based models that has not been investigated so far. Intuitively, the rules are supposed to be crucial to the understanding of the system. As such, they should convey essential knowledge about the system that does not fluctuate if the acquired data is slightly affected. We do not want to see a situation where minor fluctuations in collected data, and/or new data collected, suddenly produce a completely different set of rules. This volatility of rules is a highly detrimental feature; we may say that the rules become unstable, and the lack of stability may manifest itself in significantly different rules being generated. Our intent is to investigate this aspect in detail. First, we formulate the problem and offer its formal description. Second, as the rules change, the landscape of the pieces of knowledge (information granules) present in the condition and conclusion parts of the rules may also change. We consider three aspects of rule-based models, and offer the corresponding indices which quantify these aspects:

Multiplicity of rules – there are several identical rules having the same conditions and conclusions. This is an indicator that some rules repeat themselves and in this sense become more profoundly visible.

Conflicting rules – there are rules whose condition parts are the same but the conclusion parts are different. These rules are brittle in the sense that they are affected by some changes in data in the consequent resulting in lower quality rules.

Generalizable rules – there could be rules whose condition parts are slightly different; however, their conclusions are the same. These can be combined into a single rule whose condition part becomes more general; hence, the quality of the rule itself becomes higher.

Obviously, the intensity of the situations outlined above depends upon the data (and its variability), as well as the granularity of the condition and conclusion parts. It is needless to say that information granularity plays a pivotal role in these scenarios and deserves more thorough study. To the best of our knowledge, the proposed approach and the ensuing indices (metrics) are novel and not discussed in the existing literature.

In the experimental part of this study, we consider a given data set, split it randomly into a series of equally sized subsets, and for each subset we construct a rule-based model using identical modeling parameters. Next, the rules of the individual models are investigated and analyzed *en mass* for stability using the metrics described above.

A number of related methods have been proposed, which give some indication of rule quality. One of the metrics discussed in [6] is rule consistency, which bears a complementary (though inverted) definition to the conflict metric discussed in this study. Other measures of rule quality, including completeness, readability, simplicity, consistency, distinguishability, and transparency (among others), are discussed in detail in [22], which cites a number of prior studies detailing definitions for these terms [3,5,11,12]. While some of these concepts share similar goals with the metrics we introduce in this paper, no previous work has used the same methodology as applied here, and rule quality analysis exhibits a significant level of originality. We proposed this new method of fuzzy rule based model analysis with the express intent of defining a methodology for evaluating the stability of a rule base for a given data set.

Please cite this article in press as: J. Kerr-Wilson, W. Pedrycz, Some new qualitative insights into quality of fuzzy rule-based models, Fuzzy Sets Syst. (2016), http://dx.doi.org/10.1016/j.fss.2016.05.002