FuzzyR: An Extended Fuzzy Logic Toolbox for the R Programming Language

Chao Chen¹, Tajul Rosli Razak¹-², Jonathan M. Garibaldi¹

¹Lab for Uncertainty in Data and Decisions Making (LUCID), Intelligent Modelling & Analysis Group (IMA), School Of Computer Science, The University of Nottingham, UK.
²Faculty of Computer & Mathematical Science, Universiti Teknologi MARA, Perlis Branch, MALAYSIA.

Email: ¹{chao.chen, tajul.razak, jon.garibaldi}@nottingham.ac.uk, ²tajulrosli@uitm.edu.my

Abstract—This paper presents an R package FuzzyR which is an extended fuzzy logic toolbox for the R programming language. FuzzyR is a continuation of the previous Fuzzy R toolboxes such as FuzzyToolkitUoN. Whilst keeping existing functionalities of the previous toolboxes, the main extension in the FuzzyR toolbox is the capability of optimising type-1 and interval type-2 fuzzy inference systems based on an extended ANFIS architecture. An accuracy function is also added to provide performance indicators featuring eight alternative accuracy measures, including a new measure UMBRAE. In addition, graphical user interfaces have been provided so that the properties of a fuzzy inference system can be visualised and manipulated, which is particularly useful for teaching and learning. Note that this paper illustrates some of the new features of the FuzzyR toolbox, but does not provide a complete list of all functions available. More details about the new features of FuzzyR and a complete description of all functions can be found in the manual of the toolbox.

I. INTRODUCTION

Since its introduction by Zadeh in 1965 [1], fuzzy logic has become more and more popular, both as a specific field of research, as well as a tool for the modelling and computation with data applied in a myriad of disciplines [2, 3]. As mentioned by Wagner [4], while the research interest in fuzzy logic (in particular type-2 fuzzy logic) has steadily grown in the last 10 years, the number of practical applications in real-world use is still limited. There should be no doubt that the provision of free and open-source toolboxes is vital for the wider dissemination, accessibility and implementation of theoretic work and applications on fuzzy logic systems [5, 6].

In the literature, there have been many fuzzy logic toolkits available for type-1, interval type-2 and general type-2 fuzzy logic systems under different programming languages. For example, there are: MATLAB®, toolboxes [7, 8, 9]; GUAJE, FisPro, Juzzy and Juzzy Online for Java [10, 11, 12, 13], and Fuzzycreator for Python [6]. On the other hand, R is an ideal choice of language for such a toolbox as it is freely available and used by a wide variety of researchers from a variety of research fields: creating a tool in R will provide a generic infrastructure for a wide range of users and will hopefully lead to an increase in the uptake of fuzzy logic systems (of all flavours). In a previous paper, Wagner et al. [4] presented a fuzzy toolbox for the R programming language. The toolbox was later released as R package called FuzzyToolkitUoN [14].

As already described [4], the previous Fuzzy R toolbox can provide functions to create and perform fuzzy inference in Mamdani style. However, the Takagi–Sugeno–Kang (TSK) type fuzzy inference systems, which are also widely used, were not supported. Also, in many applications such as fuzzy logic controllers, effective methods such as ANFIS [15, 16] are widely used to tune the membership functions in order to gain better performance for a pre-defined fuzzy inference system. Such optimisation functions are missing in the previous R toolbox. This paper presents a newly released R package called FuzzyR [17], which is an extended version of the previous R fuzzy logic toolbox.

The remainder of this paper is structured as follows. Section II briefly summarises existing features of the previous toolbox. Section III presents the newly added features in FuzzyR. Section IV provides some code examples for the newly added features. Finally, conclusions are drawn in Section V.

II. SUMMARY OF PREVIOUS TOOLBOX FEATURES

The original fuzzy logic toolbox for the R programming language was presented by Wagner et al. [4]. It was later published in two forms. One is the R package FuzzyToolkitUoN [14], which is mainly for type-1 fuzzy logic. The other was an intermediate version in which the toolkit was extended to feature both type-1 and interval type-2. This version was released in source code format, available via a direct web link, but was never released as an official R package. This derived version (fuzzy-v1_7) is now superseded by the package described in this paper, and hence is now deprecated.

As mentioned by Wagner et al. [4], the R toolkit has been designed to maximise the transparency for people who are already familiar with the parameters and specification of fuzzy sets and systems while also making it easy to learn for people who are new to the area. Incidentally, the command line interface employed in the R language results in a high similarity compared to the command line interface operation of the MATLAB® Fuzzy Logic Toolbox. The main features of the previous fuzzy logic toolbox for the R programming language can be summarised as follows.

- Create, evaluate and plot type-1 and interval type-2 membership functions. Most of the commonly used types (e.g.
Triangular, Trapezoidal and Gaussian) of membership functions have been defined and ready-to-use. Users can also use customised membership functions.

- Create, export and import Mamdani style fuzzy inference systems with user-specified membership functions, rules and defuzzification methods etc.
- Produce a summary of the constructed fuzzy inference systems showing all the properties such as input and output membership functions, the rule antecedents and consequents.
- Print the rule base of the constructed fuzzy inference systems using IF-THEN sentences which are easy-to-understand.
- Evaluate the given fuzzy inference system with user-specified input values.
- Visualise the control surface of the given fuzzy inference system under the pre-defined domain of inputs.

### III. NEW TOOLKIT FEATURES

This sections presents the new toolkit features in FuzzyR.

#### A. Redesigned membership functions

In the previous R toolbox, specifying membership functions are actually generating discretised fuzzy sets. By calling a membership function with specified input values \( x \) and parameters \( p \), the output will be the membership grades \( y \) for the corresponding input \( x \). See Fig. 1 as an example.

```r
x <- 1:10
p <- c(2, 5, 8)
y <- trimf(x, p)
```

![Fig. 1. Specify a Triangular membership function with the previous R toolbox.](image)

With **FuzzyR**, a membership function can be specified with only the parameters \( p \). Internally, it generates a real function which can later be evaluated with input \( x \) values. See Fig. 2 as an example.

```r
p <- c(2, 5, 8)
trimf1 <- genmf("trimf", p)
class(trimf1)
```

![Fig. 2. Specify and evaluate a Triangular membership function with the **FuzzyR** toolbox.](image)

#### B. Fuzzy set composition

Based on the newly designed membership functions described above, **FuzzyR** provides methods for the composition of fuzzy sets based on t-norm and t-conorm operators. These methods can generate new membership functions as required (See Fig. 3). This would be quite useful for designing and implementing fuzzy inference systems.

```r
p1 <- c(2, 5, 8)
trimf1 <- genmf("trimf", p1)
p2 <- c(3, 6, 9)
trimf2 <- genmf("trimf", p2)
trimf3 <- fuzzy.tnorm("min", trimf1, trimf2)
class(trimf3)
```

![Fig. 3. An example of using the t-norm operator with the **FuzzyR** toolbox.](image)

#### C. TSK type fuzzy inference system

As mentioned above in the Introduction, the Takagi–Sugeno–Kang (TSK) type inference was not supported in the previous toolbox. It has now been added as an option (see below) in **FuzzyR** for both type-1 and interval type-2 fuzzy inference systems. A complete demo for creating a TSK model can be found in the source file of the function `tipper.tsk`.

```r
fis = newfis('tipper', fisType = 'tsk')
```

#### D. ANFIS - TSK type fuzzy system optimisation

The architecture of ANFIS was proposed by Jang [15] to serve as a basis for constructing a set of fuzzy rules with appropriate membership functions to generate the desired fuzzy inference system. A hybrid learning algorithm, based on the gradient method and the least-squares estimate, is applied to tune the parameters of membership functions. The original ANFIS was designed for type-1 TSK models. The implementation of ANFIS in **FuzzyR** is based on an extended ANFIS architecture which can be applied to both type-1 and interval type-2 fuzzy inference systems. There are many options, as shown below, available when using ANFIS to optimise a fuzzy inference system in **FuzzyR**. Details about the usage can found in the manual of the toolbox. Some demonstration code for illustrating the use of ANFIS in **FuzzyR** will also be presented in Section IV.

```r
anfis.optimise(anfis, data.trn, data.chk = NULL, epoch.total = 100, stepsize = 0.1, rate.inc = 1.1, rate.dec = 0.9, method = c("gradient", "lse"), err.log = F, online = 0, lambda = 1, opt.by = "err.opt", err.trn.fix = T)
```

#### E. Type-reduction algorithms

Interval type-2 fuzzy logic systems have been demonstrated to have better abilities to handle uncertainties than their type-1 counterparts in many applications [18, 19, 20, 21, 22, 23, 24]. However, the high computational cost of type-reduction algorithms makes it more expensive to deploy interval type-2 systems, especially for certain cost-sensitive real-world
applications. Improving the efficiency of type-reduction algorithms continues to attract research interest. In FuzzyR, the mostly commonly used type-reduction algorithms such as EKM [25] and EIASC [26] have been implemented (see Fig. 4). We have also implemented other efficient algorithms such as DA, DAND, COSTRWSR and SC presented in recent studies [27, 28, 29, 30, 31].

```r
N <- 2^5
wr <- runif(N, 0, 1)
w1 <- wr * runif(N, 0, 1)
x <- abs(rnorm(N, 0, 1))

ekm(w1, wr, x, maximum = T, w.which = F, sorted = F, k.which = F)
## [1] 0.809751
dand(w1, wr, x, maximum = T, w.which = F, sorted = F, k.which = F)
## [1] 0.809751
eiasc(w1, wr, x, maximum = T, w.which = F, sorted = F, k.which = F)
## [1] 0.809751
costrwsr(w1, wr, x, maximum = T, w.which = F)
## [1] 0.809751
sc(w1, wr, x, maximum = T, w.which = F)
## [1] 0.809751
```

Fig. 4. An example of the use of type-reduction functions in FuzzyR.

**F. Accuracy measures**

Without a suitable measure, it is not possible to carry out an objective evaluation of methodologies. Various accuracy measures have been proposed over the past decades as being the best to use in various situations [32]. However, many of these measures are not generally applicable due to issues such as sometimes generating infinite values, or being undefined under certain conditions, which may produce misleading results. In a recent study [33], a new accuracy measure called UMBRAE, which is based on bounded relative error, was proposed. Though the title of the paper may suggest an area of time series forecasting, UMBRAE is actually a good choice for general use as an alternative to other accuracy measures such as RMSE and GMRAE. It combines the best features of various alternative measures without suffering their common issues. Specifically, the new accuracy measure has the following properties:

(i) Informative: it can provide an informative result without the need to trim errors.
(ii) Resistant to outliers: it can hardly be dominated by a single forecasting outlier.
(iii) Symmetric: overestimates and underestimates are treated equally.
(iv) Scale-independent: it can be applied to datasets on different scales.
(v) Interpretability: it is easy to understand and can provide intuitive results.

In FuzzyR, there is an accuracy function which can provide performance indicators by using eight different accuracy measures including the new measure UMBRAE (see Fig. 5).

```r
y <- rnorm(100, sd = 0.1)
x <- rnorm(100, sd = 0.1)
tref <- 0

fuzzyr.accuracy(t, y, tref)
## MAE RMSE MASE MRAE GMRAE MAPE sMAPE uMbRAE
## 0.117 0.148 NA 2.389 1.298 238.904 138.207 1.263
```

Fig. 5. An example of the use of accuracy function with the FuzzyR toolbox.

**G. Graphical user interface**

In the previous toolkit, we use a command-line interface to build the fuzzy logic system which similar to MATLAB®. In order to allow users to construct FLSs easily, a graphic user interface (GUI) is presented, which is also similar to that of MATLAB® Fuzzy Logic Toolbox. Thus, the toolkit will be more user-friendly and more interactive, especially for a new user of FuzzyR. This GUI enables the user to view every element of FLS such as membership function (MFs) graph, rules and to evaluate the fuzzy inferences. As part of describing this GUI, we present the Waiter Tipping application, before briefly introducing some of the the visual graphical user interface features.

For example, the source code snippet below produces the initial main GUI for tipper example, as shown in Fig. 6. Note that the complete source code of the tipper function can be obtained from the toolbox.

```r
fis = tipper();
showGUI(fis);
```

As can be seen in Fig. 6, the main GUI consists of: (i) Menu option namely; Membership function, Rules and Defuzzier. (ii) Selection panes which show the input and output variables and function evaluate fuzzy inference; and (iii) Two buttons for exit and download.

In addition, the details of what can be done using this GUI are described in the following subsystem.

1) **Display membership functions plot:** This GUI enables the inputs and output of the membership function to be displayed, as shown in Fig. 7. For example, select input/output dropdown menus allow the user to choose which input or output of membership function to be displayed. In the case of Waiter-Tipping example, the user can make a selection based...
on service, food and tip, as shown in Fig. 7. Then, Fig. 8 shows the MFs plot that chosen by the user for the Waiter Tipping example. Additionally, download plot button is also provided for the user if they intend to download the plot of that particular MF.

2) View Rule Base of FIS: In the previous toolkit, the user has to write a source code snippet, namely showrule(), in order to view all the rules inside the FIS created. Then, as a result, the rules will be shown in the R programming console. However, by using this GUI, it enables the user to graphically view all the rules, by clicking on a button provided, as shown in Fig 9.

Fig 9 shows an example of the rule-based extracted from the FIS of the Waiter-Tipping example.

3) Evaluate Fuzzy Inference System (FIS): The main advantage of this GUI is that it provides the user with an easy to understand visual representation of the fuzzy inference system (FIS). Conversely, in the previous toolkit, users have to write a source code snippet, namely evalfis() and also need to provide the values in order to evaluate the FIS object. Then, the toolkit shows the result on the R programming console only. However, in this FuzzyR, we have transformed the evaluation FIS function to be more user-friendly and interactive by using the GUI, as shown in Fig. 10.

As an example, Fig. 10 shows the FIS of the Waiter Tipping FIS example, which is evaluated using the GUI. In this GUI, the user may insert the values using a slider provided on the left at the menus, as shown in Fig. 10. The number of the slider is generated based on how many input variables that are given into FIS. Then, the results are generated, which are displayed at the centre of this main interface, as shown in Fig. 10. Note that the results and graph automatically change to reflect the values that the user provides using the slider. Likewise, the user also can download the graph by using the download button provided, as shown in Fig. 10.

IV. CODE DEMO

In this section, we use demo code to illustrate how to use ANFIS with FuzzyR for both type-1 and interval type-2 models. The demo models are built based on a commonly used Iris flower data set, which has 150 instances. As shown in Table I, there are five columns in the data set. The first four columns are the features, and the fifth column is the species of Iris (setosa, virginica and versicolor).

Firstly, we need to load the FuzzyR package. Then we do the data preprocessing to setup the training, validation and testing
TABLE I
DATA SAMPLES OF THE IRIS FLOWER DATA SET.

<table>
<thead>
<tr>
<th>Sepal Length</th>
<th>Sepal width</th>
<th>Petal Length</th>
<th>Petal width</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>3.5</td>
<td>1.4</td>
<td>0.2</td>
<td>I. setosa</td>
</tr>
<tr>
<td>4.9</td>
<td>3.0</td>
<td>1.4</td>
<td>0.2</td>
<td>I. setosa</td>
</tr>
<tr>
<td>7.0</td>
<td>3.2</td>
<td>4.7</td>
<td>1.4</td>
<td>I. versicolor</td>
</tr>
<tr>
<td>6.3</td>
<td>3.3</td>
<td>6.0</td>
<td>2.5</td>
<td>I. virginica</td>
</tr>
</tbody>
</table>

sets. In this demo, only first two columns are used as inputs and the last column is used as the output.

```r
library(FuzzyR)
## First two columns are used
data.all <- iris[,3:5]
species <- as.factor(data.all[,3])
data.all[,3] <- as.numeric(species)
data.all <- as.matrix(data.all)
## Split the data set into training/validation/testing sets.
n.trn <- 120
n.chk <- 15
n.tst <- 15
data.trn <- data.all[1:n.trn,]
data.chk <- data.all[(n.trn+1):(n.trn+n.chk),]
data.tst <- data.all[-(1:(n.trn+n.chk)),]

To build the fuzzy inference models, we need to define the number of inputs and membership functions. In this demo, there are two inputs and we define three membership functions for each input.

```r
input.num <- 2
input.mf.num <- rep(3, input.num)
```

The rule base can be customised, or selected from the whole rule set. In this case, we use random four rules from the whole rule set.

```r
rule.num <- 4
rule.which <- sort(sample(1:prod(input.mf.num), rule.num))
```

To use relative accuracy measures such as MRAE, GMRAE and UMBRAE, we need to define the reference benchmark method. In this case, the reference benchmark method is to simply use the values of the second input as the output. A scaling factor needs to be defined for the accuracy measure MASE.

```r
naive.trn <- data.trn[,2]
native.chk <- data.chk[,2]
native.tst <- data.tst[,2]
scale.mase <- mean(abs(data.trn[,3] - naive.trn))
```

Before optimising the model by ANFIS, some settings need to be defined.

```r
## optimisation method: alg1 -- the hybrid algorithm; alg2 -- pure back propagation
alg1 <- c("gradient", "lse")
alg2 <- c("gradient", "gradient")
algorithm <- alg1
## epochs: the number of training epochs
## stepsize: the length of each gradient transition in the parameter space
epochs <- 100
stepsize <- 0.01
## online training flag: 0 -- batch; 1 -- online; 2 -- semi-online
online.flag <- 0
## err.log: T or F, the flag indicate whether to save the error log.
err.log <- T
## opt.by: the optimum model will be chosen based on smallest error obtained. "err.trn" -- training error; "err.chk" -- validation error.
opt.by <- "err.chk"
```

After building the ANFIS architecture, the fuzzy inference model can be optimised with the function `anfis.optimise`.

```r
fis.it2 <- fis.builder(x.range, input.num, input.mf.num, input.mf.type="IT2", rule.num=rule.num, rule.which=rule.which, defuzzMethod="KM")
anfis.it2 <- anfis.builder(fis.it2)
anfis.it2 <- anfis.optimise(anfis.it2, data.trn, data.chk, epoch.total=epochs, stepsize=stepsize, rate.inc=1.1, rate.dec=0.9, method=algorithm, err.log=T, online=online.flag, lambda=1, opt.by)
```

The training/validation errors during training epochs are saved in `err.trn` and `err.chk`, which can be used for further analysis such as plotting the error curves (example code provided below at the end of this section).

```r
if(err.log){
  err.trn.log.it2 <- err.trn
  err.chk.log.it2 <- err.chk
}
```

Finally, the optimised model can be used for further evaluations (e.g. to obtain the testing errors).

```r
f.tst.it2 <- anfis.eval(anfis.it2, data.tst[,1:input.num])
fuzzyr.accuracy(f.tst.it2, data.tst[,-(1:input.num)], naive.tst, scale.mase)
```

Code for the interval type-2 model is listed below.

```r
fis.it2 <- fis.builder(x.range, input.num, input.mf.num, input.mf.type="IT2", rule.num=rule.num, rule.which=rule.which, defuzzMethod="KM")
anfis.it2 <- anfis.builder(fis.it2)
anfis.it2 <- anfis.optimise(anfis.it2, data.trn, data.chk, epoch.total=epochs, stepsize=stepsize, rate.inc=1.1, rate.dec=0.9, method=algorithm, err.log=T, online=online.flag, lambda=1, opt.by)
if(err.log){
  err.trn.log.it2 <- err.trn
  err.chk.log.it2 <- err.chk
}
```

f.tst.it2 <- anfis.eval(anfis.it2, data.tst[,1:input.num])
fuzzyr.accuracy(f.tst.it2, data.tst[,-(1:input.num)], naive.tst, scale.mase)
Below is a code example for plotting the training and validation RMSE error curves, shown in Fig. 11, during the training epochs of type-1 and interval type-2 ANFIS models.

```r
require(ggplot2)
require(reshape2)

rmse.trn.t1 <- err.trn.log.t1[,2]
rmse.chk.t1 <- err.chk.log.t1[,2]
rmse.trn.it2 <- err.trn.log.it2[,2]
rmse.chk.it2 <- err.chk.log.it2[,2]

errors <- cbind(rmse.trn.t1, rmse.chk.t1, rmse.trn.it2, rmse.chk.it2)
colnames(errors) <- c("training.t1", "validation.t1", "training.it2", "validation.it2")
df <- data.frame(errors)
df$epochs <- factor(1:epochs)
df <- melt(df,id.vars="epochs")
cbbPalette <- c(rainbow(15), "#000000")
g <- ggplot(df, aes(x=epochs, value, group=variable, shape=variable, fill=variable, colour=variable, linetype=variable)) +
  scale_x_discrete(breaks=seq(epochs/10, epochs, epochs/10)) +
  geom_line(size=0.8) + geom_point(size=2) +
  scale_shape_manual(values=c(2,6,1,4,3)) +
  labs(x="Epochs", y="RMSE") +
  ggtitle("") +
  theme_bw() +
  theme(plot.title = element_text(lineheight=.8, face="bold"))
print(g)
```

![Fig. 11](image1.png)

Fig. 11. The training and validation RMSE error curves during the training epochs of type-1 and interval type-2 ANFIS models.

With the function `anfis.plotmf` which can be used to plot the antecedent membership functions of an ANFIS object, the user can visually check the changes of membership functions made by the optimisation. The below code is an example to show the membership functions (see Fig. 12) of the second input used in the interval type-2 ANFIS model demonstrated above.

```r
par(mfrow=c(1,2))
anfis.plotmf(anfis.builder(fis.it2), 'input', 2)
anfis.plotmf(anfis.it2, 'input', 2)
```

![Fig. 12](image2.png)

Fig. 12. An illustration of the antecedent membership functions for an interval type-2 ANFIS model before (left) and after (right) optimisation.

V. CONCLUSION

In this paper, we have presented a newly released fuzzy logic toolbox, called FuzzyR, for the R programming language. The toolbox is available as a standard R package on CRAN [17]. FuzzyR keeps the main functionalities of the previous R toolbox, as summarised in this paper, but many new features have been added in this updated toolbox. Specifically, we have redesigned membership functions and fuzzy operators. TSK type models have been supported in FuzzyR and ANFIS functions can be used to optimise the performance of type-1 and interval type-2 TSK models. Moreover, we have implemented the state-of-the-art type-reduction algorithms (e.g. EIASC, DAND and SC) presented in recent studies [30, 31]. In FuzzyR, there is an accuracy function which uses eight different accuracy measures including a new accuracy measure UMBRAE proposed in [33]. Additionally, graphical user interfaces have been provided, which can be very useful for the purposes of teaching and learning.

Though many new features have been included in FuzzyR compared to the previous R toolbox, there are still a lot of things that can be done in the future. For example, currently, the ANFIS models in FuzzyR can only be based on generalised bell-shaped membership functions. This could be expanded with the support of more types of membership functions. Also, other optimisation methods (derivative-free algorithms such as simulated annealing and particle swarm optimisation) could also be considered. Furthermore, since there have been more and more research interest in general type-2 fuzzy inference systems, the support of general type-2 functionalities in FuzzyR is also in the plan. Also, since the Computational Intelligence Society has endorsed a standard for Fuzzy Markup Language (FML) which enables the interoperability of fuzzy inference systems irrespective of the environment they are executed [34, 35, 36, 37], we are aiming to extend FuzzyR in a future version that complies with the FML.

REFERENCES


