Towards Designing and Measuring Interpretable Hierarchical Fuzzy Systems



University of Nottingham
UK | CHINA | MALAYSIA

Tajul Rosli Razak

School of Computer Science
University of Nottingham

A thesis submitted for the degree of
Doctor of Philosophy

December 2019

I would like to dedicate this thesis to my loving family - Hasila Jarimi and Mia Dahlia.

Declaration

The work in this thesis is based on research carried out at the Intelligent Modelling and Analysis (IMA), and the Lab for Uncertainty in Data and Decision Making (LUCID) of Research Groups, the School of Computer Science, the University of Nottingham, UK. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text.

Tajul Rosli Razak
December 2019

Acknowledgements

I would like to thank Allah for his blessings, inspirations, and the passion that He gave me for this enjoyable PhD journey.

I also would like to thank the following people without whom this thesis would not have happened. Foremost, my main supervisor and co-supervisor; Professor Jonathan M. Garibaldi and Dr Christian Wagner, for their dedicated help, advice, encouragement, and inspiration throughout my PhD. I also would like to thank Dr Daniele Soria from the University of Westminster, the UK; Dr Amir Pourabdollah from Nottingham Trent University, the UK; Ministry of Higher Education (MOHE) Malaysia and Universiti Teknologi MARA (UiTM) Perlis, Malaysia for the doctoral scholarship; the School of Computer Science , University of Nottingham for the facilities.

Special thanks to my mother Che Mah Harun, my lovely wife Hasila Jarimi, and my daughter Mia Dahlia for their love, patience and understanding. Another sincere thank you message to my family in Malaysia for all their support and prayers.

A sincere thanks also go to all participants of FUZZ-IEEE 2017 for their time participating in my survey. Their valuable answers and detailed comments have contributed enormously to this study. Also a special thanks to Dr Sharkey and all the people involved in the Neonatal Intensive Care Unit (NICU) at QMC Hospital, Nottingham in which I have collaborated for providing the medical dataset and also guidance and sharing of vast experience in the field of NICU.

Last but not least, to all IMA and LUCID research groups members, thank you for the friendly working environment in the group, especially friends in room B38.

Abstract

This thesis presents a detailed study of the interpretability of hierarchical fuzzy systems (HFSs). It focuses on the development of a design guidelines framework for interpretable HFSs. This thesis aims to fill some of the gaps in the body of knowledge. Several research questions are raised, including: "How can the interpretability of HFSs be measured with indices?", "How can complexity be comprehensively measured in HFSs?", "How can user perception on the interpretability and complexity of HFSs be captured?" and "How can interpretable HFSs be designed?" Thus, to study the interpretability of HFSs, this thesis includes the following methodology discussed in different chapters.

First, measures of interpretability are investigated, and an index measuring interpretability specifically in HFSs is introduced. The best way to know about interpretability is by learning how to measure it. Although many researchers have suggested indices to measure interpretability, none of them try to measure the interpretability of HFSs. Indeed, all of them only focus on measuring the interpretability of fuzzy logic systems (FLSs). This is due to the HFSs' architecture, i.e., multiple subsystems, layers, and topologies, and this presents a significant challenge to measure the interpretability of HFSs. Based on this investigation, this study successfully introduces an initial index for measuring the interpretability of HFSs. The initial index is built based on the challenges arising from the structure of HFSs mentioned.

Due to the subjective nature of the interpretability, the best way to validate the proposed measurements of interpretability of HFSs is by asking the users. However, it is not an easy task to get the user perception, particularly on the interpretability, and there is a lack of research on this issue. Therefore, the second focus of this thesis presents research on a new method of capturing user perceptions of the interpretability and also complexity of HFSs. This is the first time that user study has been used to obtain and assess both qualities in HFSs. Based on this, a new analysis of the relationship between interpretability and complexity of HFSs is presented, and this provides insights into the process of developing measures of interpretability of HFSs.

However, rather than just using the user study to evaluate the measurements directly, this study also uses input from the user study to 'guide' the measurement of interpretability of

HFSs in what is known as a participatory design approach. The participatory design approach enables the subjective views of a range of users to be taken into account in shaping the measurement of interpretability of HFSs. Thus, the use of the participatory design approach to configure the resulting measurement of interpretability of HFSs is also evaluated.

Complexity is seen as an essential component in determining interpretability. In FLSs, complexity is expressed by the number of rules, variables, and fuzzy terms, called rule-based complexity. Several studies have used indicators (for example, the number of rules) to measure the complexity of FLSs. However, none of the studies considered the structure of HFSs, i.e., multiple subsystems, layers and varied topologies, which may also affect the complexity of HFSs. In addition, the user study revealed different perceptions of complexity in HFSs. Therefore, research is then presented on improving the complexity measurement in HFSs. The measurement is based on combining rule-based complexity with the structural complexity.

Designing an interpretable HFS is a challenging task because of the need to define the interpretability of the architecture of the HFS (the subsystems, the input variables of each subsystem, and the interactions between subsystem), as well as the rules of each subsystem. To assist with this, a design guidelines framework for interpretable HFSs is produced. The framework is based on the measurement index of interpretability and complexity that is presented earlier. The framework consists of five guidelines for building interpretable HFSs.

Finally, to demonstrate a design guidelines framework for interpretable HFSs on the real-world example, a design for an interpretable HFS for a neonatal intensive care unit (NICU) is produced. It is aimed to provide understandable decision support model to clinicians. In the medical context, it is essential for people to understand the importance of the system features. The HFSs is then practically illustrated by using real physiological data at NICU and compared with a flat FLS system. The results show that the design guidelines framework can offer the ability to design an interpretable HFS in practice efficiently.

Table of contents

List of Figures **xii**

List of Tables **xvi**

1 Introduction **1**

1.1 Background and Motivation 1
1.2 Aims and Objectives 4
1.3 Contributions to Knowledge 5
1.4 Thesis Outline 6

2 Literature Review **9**

2.1 Introduction 9
2.2 Fuzzy Logic Systems 9
2.2.1 Fuzzy Sets and Membership Functions 10
2.2.2 Linguistic Variables, Values and Rules 13
2.2.3 Overview of Designing Fuzzy Logic Systems 15
2.2.4 Fuzzy Logic in Medical Domains 16
2.2.5 The Curse of Dimensionality 17
2.3 Hierarchical Fuzzy Systems 18
2.3.1 Introduction 18
2.3.2 Structure of Hierarchical Fuzzy Systems 21
2.3.3 Approaches to Building Hierarchical Fuzzy Systems 23
2.3.4 Advantages of Hierarchical Fuzzy Systems 29
2.3.5 Issues of the interpretability in Hierarchical Fuzzy Systems 31
2.4 Interpretability 33
2.4.1 State of The Art 33
2.4.2 Taxonomy to Analyse Interpretability of Fuzzy Systems 35

2.4.3 Measuring Interpretability using Interpretability Indices 42
2.4.4 Measuring Interpretability of Fuzzy Systems Using a User Study . . 47
2.4.5 Interpretability Issues in Fuzzy Systems 48
2.5 Complexity 49
2.5.1 State of the Art 49
2.5.2 Complexity in Fuzzy Systems 50
2.5.3 Complexity in Other Fields 52
2.6 Case Study: Neonatal Monitoring 54
2.6.1 Summary of Existing Neonatal Monitoring Systems using Fuzzy Logic Systems 55
2.6.2 Problems 58
2.7 Summary 58

3 An Initial Approach to Measuring Interpretability of Hierarchical Fuzzy Systems 60
3.1 Introduction 60
3.2 Motivation 61
3.3 The Proposed Initial Index for HFSs 62
3.4 Detail of Process for Measuring Interpretability of HFSs 63
3.5 Seesaw Control Application Experiment 64
3.5.1 Parallel Implementation of the Seesaw control 67
3.5.2 Serial Implementation of the Seesaw control 70
3.6 Result and Discussion 72
3.7 Summary 73

4 An Exploratory User Study on Perception of Interpretability in Hierarchical Fuzzy Systems 75
4.1 Introduction 76
4.2 Capturing Users' Perception on Interpretability and Complexity 77
4.2.1 Motivation 77
4.2.2 Application Background – Iris Classification Systems 77
4.2.3 The User Study Approach 78
4.3 Experiment and Results 82
4.3.1 Descriptive Analysis of the Survey Data 82
4.3.2 Participants' Feedback 92
4.3.3 Interpretability vs Complexity 93

4.4 Discussion . . . 97
4.5 Summary . . . 98

5 Participatory Design Approach to Guide the Measurement of Interpretability in Hierarchical Fuzzy Systems 100
5.1 Introduction . . . 100
5.2 Motivation . . . 101
5.2.1 Multiple Individual Subsystems . . . 101
5.2.2 Aggregation . . . 102
5.2.3 Topology and Layering . . . 102
5.3 A General Index for Interpretability of Hierarchical Fuzzy Systems . . . 103
5.3.1 HFS Interpretability . . . 103
5.3.2 Layer-weighting Strategy . . . 104
5.4 A General Index Demonstration in Principle . . . 105
5.4.1 Method . . . 106
5.4.2 Results . . . 107
5.5 Beyond the Mean – A Participatory Design Approach . . . 110
5.5.1 Participatory User Study . . . 111
5.5.2 Exploration of Alternatives Configurations of H . . . 115
5.5.3 Matching H to the Participatory Study . . . 119
5.6 Discussion . . . 121
5.7 Summary . . . 122

6 An Improved Complexity Measure in Hierarchical Fuzzy Systems 123
6.1 Introduction . . . 123
6.2 Motivation . . . 124
6.3 A Comprehensive Measurement of Complexity in Hierarchical Fuzzy Systems 125
6.3.1 Step 1: Calculating the Rule-based Complexity of HFSs . . . 125
6.3.2 Step 2: Calculating the Structural Complexity of HFSs . . . 126
6.3.3 Step 3: Overall Measurement of the Complexity of HFSs . . . 129
6.4 Experimental Results . . . 131
6.4.1 Seesaw Control Problem . . . 131
6.4.2 Iris classification Problem . . . 140
6.5 Results and Discussion . . . 143
6.6 Summary . . . 145

7 A Framework for Guiding the Design of Interpretable Hierarchical Fuzzy Systems 147
7.1 Introduction . . . 147
7.2 A Design Guidelines Framework for Interpretable HFSs . . . 148
7.2.1 Guideline 1 – Selection of Topology . . . 148
7.2.2 Guideline 2 – Simplicity and Readability of Subsystems . . . 150
7.2.3 Guideline 3 – The Meaning of Intermediate Variables . . . 151
7.2.4 Guideline 4 – Link-up Subsystems using Meaningful Intermediate Variables . . . 153
7.2.5 Guideline 5 – Constructing the Subsystems' Rules in HFSs . . . 153
7.3 Case Study: Neonatal Intensive Care Unit . . . 161
7.3.1 Medical Background . . . 161
7.3.2 Data Collection . . . 162
7.3.3 Designing an Interpretable Hierarchical Fuzzy System: Oxygenation Decision . . . 168
7.3.4 A Practical Illustration of an Interpretable HFS . . . 181
7.3.5 Measuring the Interpretability of Oxygenation Decision Systems . . . 188
7.4 Discussion . . . 188
7.5 Summary . . . 190

8 Conclusions 191
8.1 Introduction . . . 191
8.2 Thesis summary . . . 191
8.3 Contributions . . . 193
8.4 Future work . . . 194
8.5 Dissemination . . . 195
8.5.1 Journal Papers . . . 195
8.5.2 Conference Papers . . . 195
8.5.3 Presentations . . . 196

References 198

Appendix A Interpretability of Fuzzy Logic Systems Survey 212

Appendix B Participants' Feedback 228

Appendix C Interpretability on the various design of Fuzzy Logic Systems Survey 234

Appendix D Data Acquisition: OxyGUI System 259

List of Figures

2.1 The membership function for the crisp sets of tall men 11
2.2 The membership function for the fuzzy sets of *tall* men. 12
2.3 Triangular MF 13
2.4 Trapezoidal MF 13
2.5 Gaussian MF 13
2.6 Generalized Bell MF 13
2.7 Membership functions for the linguistic variable 'height' 14
2.8 Components of Fuzzy Logic Systems. 16
2.9 FLS: Iris classification 19
2.10 Serial HFS: Iris classification 19
2.11 Example of rules in FLS and HFS 21
2.12 Parallel HFS: : Iris classification 22
2.13 Structured Hierarchical Fuzzy System. 24
2.14 Conventional HFS. x_1, x_2, x_3 are the input while y_1 and y are the intermediate output and output respectively. 25
2.15 The sorting process in L-HFS. 26
2.16 Two-layer Hierarchical Fuzzy System. 28
2.17 HIFAM with n inputs x_n, one output y and n-1 rule base RB_{n-1}. 29
2.18 A taxonomy of interpretability of fuzzy systems. 36
2.19 Fuzzy sets without distinguishability. 37
2.20 Example of membership function of *service* that shows complete coverage. 38
2.21 Example of normal fuzzy set (a) and subnormal fuzzy set (b). 38
2.22 Fuzzy partition with a poor semantic interpretability. 41
2.23 Hierarchical fuzzy system for assessing interpretability for Fuzzy Index . . 46
2.24 A typical relationship between complexity and interpretability of a model. . 50
2.25 One main program M, two subroutines A and B. 53

3.1 The Seesaw System. 64
3.2 MFs for all inputs x_1, x_2, x_3, x_4 and also intermediate output in HFSs namely y_1 and y_2 65
3.3 MFs for output variable, y for Seesaw control 65
3.4 Interpretability of a Parallel HFS using Nauck index (N) and Fuzzy index (F). 67
3.5 Interpretability of a Serial HFS using Nauck index (N) and Fuzzy index (F). 70
3.6 Interpretability measurement between the FLS and HFSs of the seesaw control application using an initial index, H_I. 73

4.1 Membership functions of input *Sepal Length* with two MFs. 78
4.2 Membership functions of input *Sepal Length* with three MFs. 78
4.3 Four steps towards developing a user study 79
4.4 *Card system*: P-2 80
4.5 Participants rank F-2, P-2 and S-2 based on the interpretability and complexity of design in relation to creating the Iris classification systems 82
4.6 Participants ranked the topologies of the Iris systems based on simplicity. . 83
4.7 Participants ranked the Iris systems in Set MF-2 based on interpretability. . 85
4.8 Participants ranked the Iris classification systems in Set MF-2 based on the complexity of design. 86
4.9 Participants ranked the Iris systems in Set MF-3 based on interpretability. . 87
4.10 Participants ranked the Iris systems in Set MF-3 based on the complexity of design. 88
4.11 Participants chose HFS Iris systems in Set MF-3 as preferred to design. . . 89
4.12 Participants agree to choose the HFS rather than the FLS in Set MF-2 . . . 90
4.13 Participants agree to choose the HFS rather than the FLS in Set MF-3 . . . 91
4.14 Participants ranked HFS Iris systems in Set MF-3 based on interpretability . 92
4.15 Interpretability vs complexity for F-2 in Set MF-2 93
4.16 Interpretability vs complexity for P-2 in Set MF-2 94
4.17 Interpretability vs complexity for S-2 in Set MF-2 94
4.18 Interpretability vs complexity for F-3 in Set MF-3 95
4.19 Interpretability vs complexity for P-3 in Set MF-3 95
4.20 Interpretability vs complexity for S-3 in Set MF-3 95

5.1 Summary of the overall H_{mean} index 109
5.2 Rotary crane system: A 112
5.3 Rotary crane system: B 112

5.4 Rotary crane system: C . . . 112
5.5 Rotary crane system: D . . . 112
5.6 Rotary crane system: E . . . 112
5.7 Rotary crane system: F . . . 112
5.8 Rotary crane system: G . . . 112
5.9 Rotary crane system: H . . . 112
5.10 Rotary crane system: I . . . 112
5.11 Rotary crane system: J . . . 112
5.12 Rotary crane system: K . . . 112
5.13 Rotary crane system: L . . . 112
5.14 The figure shows a pairwise comparison of Rotary systems (System A and B) shown to participants. The participants select a system that is *more* interpretable for them. If both systems seem *equally* interpretable, they may select the '*Equal*' option as their answer. . . . 113

6.1 Complexity measurement in HFSs. . . . 125
6.2 Seesaw control Application: FLS . . . 131
6.3 K-S Test: FLS vs Parallel L-HFS . . . 136
6.4 K-S Test: FLS vs Serial L-HFS . . . 137
6.5 K-S Test: Parallel L-HFS vs Serial L-HFS . . . 137
6.6 Summary of the overall complexity (C_{HFS}) for FLS, Parallel L-HFS and Serial L-HFS, using different aggregation functions . . . 140
6.7 K-S Test: FLS vs Parallel HFS . . . 142
6.8 K-S Test: FLS vs Serial HFS . . . 142
6.9 K-S Test: Parallel HFS vs Serial HFS . . . 143

7.1 Topology: Parallel HFS . . . 149
7.2 Topology: Serial HFS . . . 149
7.3 Two inputs and a single output (2-ISO) . . . 150
7.4 Three inputs and a single output (3-ISO) . . . 150
7.5 Membership functions of inputs: *Sepal length*, *Sepal width*, *Petal length* and *Petal width* . . . 151
7.6 Mixed-mode method: step-wise procedure . . . 154
7.7 Example of a survey completed by an expert for determining Sepal classification. . . . 155
7.8 Transform the expert surveys into the table for Sepal classification. . . . 155

7.9 Parallel HFS rules – FLS_1 . . . 157
7.10 Parallel HFS rules – FLS_2 . . . 158
7.11 Parallel HFS rules – FLS_3 . . . 158
7.12 Serial HFS rules – FLS_1 . . . 159
7.13 Serial HFS rules – FLS_2 . . . 160
7.14 Serial HFS rules – FLS_3 . . . 160
7.15 Levels of care available at NICU. The movement of babies is depended on health progression. . . . 161
7.16 Data flow in the data acquisition process . . . 163
7.17 Sample of physiological data . . . 164
7.18 Histogram of each of physiological parameter for Baby001. . . . 166
7.19 Scatter plot of the parameter correlation . . . 167
7.20 Physiological parameter : Heart Rate (bpm) . . . 168
7.21 Physiological parameter : Oxygen Saturation (%) . . . 169
7.22 HFS topology: Oxygenation Decision . . . 170
7.23 MFs of input variable: Heart Rate (bpm) . . . 171
7.24 MFs of input variable: Oxygen Saturation (%) . . . 171
7.25 MFs of input variable: Baby's condition (Crying) . . . 171
7.26 MFs of input variable: Baby's condition (Sleeping) . . . 172
7.27 MFs of input variable: Baby's condition (Stress) . . . 172
7.28 MFs of intermediate variable: Heart Rate class . . . 172
7.29 MFs of intermediate variable: Oxygen Saturation class . . . 173
7.30 MFs of output variable: Oxygenation Decision . . . 173
7.31 The process involved in producing interpretable HFS rules from human experts 178
7.32 FLS topology: Oxygenation Decision . . . 181
7.33 The graphical user interface (GUI) for the Oxygenation Decision system: Main interface . . . 183
7.34 The Kolmogorov–Smirnov test comparing the FLS and HFS Oxygenation Decision systems over six days . . . 187

List of Tables

2.1 The rule by random access. P, Z, N are linguistic terms for each input variable. 26
2.2 The rules for F_1 27
2.3 The rules for F_2 27
2.4 Definitions of various terms related to interpretability from online dictionaries 34
2.5 A taxonomy to analyze the interpretability of FLSs. 40
2.6 Summary of the current interpretability indices 43
2.7 Fuzzy index of Waiter-Tipping Example 47
2.8 Summary of existing neonatal monitoring systems using Fuzzy Logic Systems 56

3.1 Description of the FLS, Parallel HFS and Serial HFS for Seesaw Control Applications 66
3.2 Interpretability of the Seesaw Control Application using the Nauck and Fuzzy Indices 67

4.1 Iris classification systems 78
4.2 Frequency of topology rankings in Question 1 extracted from survey (from simpler (1) to the least (3)). 83
4.3 Frequency of interpretability rankings in Question 2 extracted from survey (from most interpretable (1) to the least (3)). 84
4.4 Frequency of complexity rankings in Question 3 extracted from survey (from less complex (1) to more complex (3)). 85
4.5 Frequency of interpretability rankings in Question 4 extracted from survey (from more interpretable (1) to the least (3)). 86
4.6 Frequency of complexity rankings in Question 5 extracted from survey (from less complex (1) to the more complex (3)) 87
4.7 Frequency of participants' answers in Question 6 extracted from survey . . 88
4.8 Frequency of participants' answers in Question 7 extracted from survey . . 89

4.9 Frequency of participants' answers in Question 8 extracted from survey . . 90
4.10 Frequency of participants' answers in Question 9 extracted from survey . . 91
4.11 Pearson correlation coefficient (r) between Interpretability and Complexity of design for Iris classification systems in Set MF-2 and Set MF-3 96

5.1 Description of the six Iris classification systems 106
5.2 Interpretability of the Iris systems; F-2, P-2 and S-2 using the H_{mean} . . . 108
5.3 Interpretability of the Iris systems; F-3, P-3 and S-3 using the H_{mean} . . . 108
5.4 A summary of the rotary crane for various HFS designs 111
5.5 Frequency of the users interpretability rating for pairwise comparisons of rotary crane, as extracted from user study 114
5.6 The interpretability ratings for the pairwise comparisons of the Rotary crane system extracted from participatory user study 115
5.7 Base interpretability of each of the Rotary Crane Systems measured using the Fuzzy interpretability index 116
5.8 Interpretability of the Rotary Crane Systems using H with different aggregation strategies and layer weighting strategies 117
5.9 Pairwise comparison for the interpretability of the Rotary crane systems obtained from the H with different aggregation strategies at each HFS layer 118
5.10 Pairwise comparison for the interpretability of the Rotary Crane systems obtained from the H with different layer-weighting strategies 119
5.11 The Agreement Score between the preferences given by each of the users (in Table 5.5) and the preference indicated by H framework (using different aggregation and layer-weighted strategies as shown in Tables 5.9 and 5.10) . 120

6.1 The rules of flat FLS 132
6.2 Description of the FLS, Parallel L-HFS and Serial L-HFS for Seesaw Control Applications 133
6.3 The rules of Parallel L-HFS: FLS_1 133
6.4 The rules of Parallel L-HFS: FLS_2 134
6.5 The rules of Parallel L-HFS: FLS_3. 134
6.6 The rules of Serial L-HFS: FLS_1 134
6.7 The rules of Serial L-HFS: FLS_2 135
6.8 The rules of Serial L-HFS: FLS_3 135
6.9 The Kolmogorov-Smirnov Test between FLS, Parallel L-HFS and Serial L-HFS of Seesaw Control Application 136

6.10 The Rule-based Complexity between the FLS, Parallel L-HFS and the Serial L-HFS of Seesaw Control Application 138
6.11 The Structural Complexity between the FLS, Parallel L-HFS and the Serial L-HFS of Seesaw Control Application 139
6.12 Summary of overall measurement complexity for FLS, Parallel L-HFS and Serial L-HFS with various alternatives combination strategies 140
6.13 The Kolmogorov-Smirnov Test between FLS, Parallel HFS and Serial HFS of Iris classification 141

7.1 Rules of Sepal classificaion 156
7.2 Physiological data 164
7.3 The pre-process data 165
7.4 Mean (μ) and standard deviation (σ) of Baby001 data 165
7.5 Abbreviations and symbols used in the survey 174
7.6 Results extracted from the survey (Heart Rate vs Crying) 175
7.7 Results extracted from the survey (Heart Rate vs Sleeping) 175
7.8 Results extracted from the survey (Heart Rate vs Stress) 175
7.9 Results extracted from the survey (Oxygen Saturation vs Crying) 176
7.10 Results extracted from the survey (Oxygen Saturation vs Sleeping) 176
7.11 Results extracted from the survey (Oxygen Saturation vs Stress) 176
7.12 Fuzzy Rule Base (FLS_1): Heart rate classification 177
7.13 Fuzzy Rule Base (FLS_2): Oxygen saturation class 177
7.14 Fuzzy Rule Base (FLS_3): Oxygenation decision 178
7.15 HFS rules in subsystem FLS_1 using 'AND' logical connective 179
7.16 HFS rules in subsystem FLS_2 using 'AND' logical connective 180
7.17 HFS rules in subsystem FLS_3 180
7.18 A flat FLS rules using 'AND' logical connective 182
7.19 Features of both FLS and HFS implementations of the oxygenation decision system 183
7.20 The Sample Evaluation of FIS output for FLS and HFS of Oxygenation Decision 185
7.21 The Kolmogorov-Smirnov Test between FLS and HFS of Oxygenation Decision 186
7.22 A summary of interpretability of Oxygenation decision systems using $H_{\min}$ 188

B.1 Question 1 – Participants' feedback on the simplest topologies for Iris Systems 228

B.2 Question 2 – Participants' feedback on the more interpretable of Iris Systems in Set MF-2 229
B.3 Question 3 – Participants' feedback on the less complex of Iris Systems in Set MF-2 230
B.4 Question 4 – Participants' feedback on the more interpretable of Iris Systems in Set MF-3 230
B.5 Question 5 – Participants' feedback on the less complex interpretable of Iris Systems in Set MF-3 231
B.6 Question 6 – Participants' feedback on HFS Iris systems (P-3 or S-3) that favoured to design in Set MF-3 231
B.7 Question 7 – Participants' feedback on agreeing to choose HFS (P-2) than FLS (F-2) with rule reduction from 16 to 12 rules in Set MF-2 232
B.8 Question 8 – Participants' feedback on agreeing to choose HFS (P-3) than FLS (F-3) with rule reduction from 81 to 27 rules in Set MF-3 232
B.9 Question 9 – Participants' feedback on HFS Iris systems (P-3 or S-3) that is more interpretable in Set MF-3 233

Chapter 1

Introduction

1.1 Background and Motivation

Interpretability is often known as a synonym for comprehensibility, intelligibility, or transparency [1], which is a desirable property for all kinds of applications such as knowledge extraction and decision support [2, 3]. As reported in [4], a model can be defined as interpretable if its behaviour is *intelligible,* i.e. it can be easily perceived and understood by a user. Interpretability has a more cognitive aspect, and its study penetrates several disciplines, including machine learning [5]. In mathematical logic, interpretability is a relation between formal theories that expresses the possibility of interpreting or translating one into the other [6]. Thus, interpretability is closely connected with the ability of users to understand the model. The concept of interpretability is considered in various fields such as education, medicine, computer science, engineering, physics and social sciences, which show a universal interest in interpretability. For instance, the concept of interpretability in medical domains is essential, that is, allowing physicians to understand why predictions or decisions are made, giving reasons behind the decision to further accept or overcome model output [7].

Fuzzy logic systems (FLSs) are one of the currently used techniques for modelling nonlinear, uncertain and complex systems. An essential characteristic of FLSs is the partitioning of the space of system variables into fuzzy regions using fuzzy sets [8]. In each region, the characteristics of the system can be described merely using a rule. Generally, an FLS consists of a rule base with rules associated to particular regions, where the information available is transparent and easily readable. This characteristic of fuzzy systems has been employed in many fields including medical [9], engineering [10], decision support [11], pattern recognition [12], and others.

One of the strengths of FLSs is their interpretability [13]. In contrast, neural networks are viewed as black boxes because mathematical formulae set the mapping between inputs and outputs. Black-box modelling can simulate a real-world system reliably and precisely, but the model structure and parameters usually give no explicit explanation about the system behaviour [14]. Conversely, FLSs can be seen as grey boxes in the sense that every element of the whole system can be checked for plausibility [15], in which relationships between input and output variables are established in terms of a fuzzy logic-based descriptive language. In addition, FLSs consist of linguistic variables [16] and linguistic rules [17] which are easy to interpret by the user [18, 19] because they are quite close to expert natural language. Also, interpretable FLSs produce a model that not only performs good global prediction, but also generates a parsimonious and understandable rule base, and as according to the principle of Occam's razor: "The simplest hypothesis or theory agreeing with the facts is the best one" [20]. Hence, the "unique selling point" of FLSs is usually related to a capability in producing simple rule-based solutions using simple terms [21].

Interpretability indicates how easily an FLS can be understood by human beings [22]. In recent years, the interest of researchers in obtaining more interpretable fuzzy models has increased. It is usually assumed that the interpretability of FLSs is automatically related directly to the number of linguistic rules [14]. However, other factors should also be considered to measure the interpretability of FLSs, such as the number of conditions, number of membership functions (MFs), number of features, consistency of rules, transparency of rule structure, completeness or coverage of MFs and so on [23]. Due to the large number of factors involved and subjective nature of interpretability (e.g. it depends on the background of the person who makes the assessment), the choice of an appropriate interpretability measure is still an open discussion. Thus, substantial research on interpretability measures [14, 23, 21, 15, 24–28] proposes interpretability indices for FLSs. Additionally, complexity is an essential component to determine the interpretability of FLSs. Thus far, complexity has often been used as an indirect measurement of the interpretability of FLSs. Several researchers claim that reduction of complexity in a system will lead to better interpretability of the fuzzy systems [22, 29, 30].

At present, a fundamental issue in FLSs is the *curse of dimensionality*; that is, the number of fuzzy rules required increases exponentially with the dimensionality of the input space [31]. Suppose there are n input variables and m membership function for each input variables, then m^n rules are needed to construct a complete fuzzy system. This problem of rules explosion increases complexity and reduces transparency and interpretability, which is then a major disadvantage [32]. To address this problem, several methods have been proposed

for optimizing the size of the rule-base such as rule selection [33–37], feature selection [38], rule interpolation [39, 40], singular-value decomposition-QR [41], evolutionary algorithm [42], fuzzy similarity measure [22] and rule learning [14]. Most of these techniques involve the modification of the original input variables of fuzzy systems, which can cause the loss of the original meaning of variables [2] and also reduce the model interpretability. Nevertheless, one effective way to deal with this problem is through the use of a special type of FLS, namely hierarchical fuzzy systems (HFSs), which reduce the number of rules while retaining the original meaning of variables [43–45].

Raju et al. [43] first introduced HFSs as a valuable approach to overcome the curse of dimensionality. The idea of HFSs is to put the input variables into a collection of low-dimensional fuzzy logic systems connected hierarchically, instead of creating a single high-dimensional rule base for a fuzzy logic system. Implementing this approach provides a way to deal with this problem of rule explosion, thus minimising complexity, and improving model interpretability. HFSs have an essential property that the total number of rules increase linearly rather than exponentially as in conventional fuzzy systems [46, 43, 44, 47]. In HFSs, if we consider two input variables per low-dimensional fuzzy system and define m fuzzy sets for each variable, then each low-dimensional fuzzy system consists of m^2 rules; therefore, the total number of rules is $(n-1)m^2$, which is a linear function of the number of input variables n [47]. Thus, the total number of rules in HFSs will always be equal to or lower than that of the 'equivalent' FLS. However, the main challenge with HFSs is to address a problem of loss of physical meaning at the intermediate output variable which may make them difficult to design. This problem can also present a drawback concerning their interpretability. Therefore, several researchers [48, 46, 49–53] have proposed a new approach in designing HFSs in an attempt to tackle the issue of lack of physical meaning of intermediate output in HFSs.

According to Benítez and Casillas [54], an HFS may improve the interpretability as "*(i) the hierarchical structure generates a lower number of variables, (ii) the algorithm does not generate artificial linking variables, and so, all of the variables are interpretable because they belong to the system, and (iii) the rules are simpler because the number of variables per module is lower. At the same time, the accuracy is maintained or improved*". Despite the fact that several researchers have claimed that HFSs can reduce complexity and thus improve the interpretability in FLSs, the remaining issue is: "How can the interpretability of HFSs be measured?" To date, there has been no attempt to extend interpretability indices created for ordinary (non-hierarchical) FLSs, such that there is no formal measurement for the interpretability of HFSs which includes issues such as their topological structure, the number of layers, or the meaning of intermediate variables.

Intuitively, the fewer the number of rules in HFSs the more complexity is reduced. This can give an advantage to the interpretability of HFSs, as stated in [55], "*interpretability maximisation is often handled as complexity minimisation*". Also, Alcala et al. [56], stated that reducing the model complexity is a way to improve the system readability, i.e., a system with a minor number of rules requires a minor effort to be interpreted. Thus, complexity is often used as an indirect measurement of interpretability in FLSs. However, questions arise such as: "Does the reduction of the complexity in HFSs improve their interpretability?" and "How can the complexity of HFSs be measured, taking into account its structure?" These issues are still unclear since there is no formal measurement for complexity in HFSs that considers the structure of multiple subsystems, layers and different topologies.

Torra in [57] has claimed that building an HFS is a challenging task because of the need to define the architecture of the HFS (the subsystems, the input variables of each subsystem, and the interactions between each subsystem), as well as the rules of each subsystem. Also, the fundamental issue with HFSs is how to handle the loss of physical meaning of the intermediate output variables and consequently they are difficult to design. Several researchers [48, 46, 49–53] have proposed a new approach in designing HFSs as they tried to tackle the issue of the lack of physical meaning of intermediate output in HFS. However, it still not clear whether these design approaches can really improve the model interpretability and minimise the model complexity. This is due to the fact that to fully understand the concept of interpretability and complexity in HFSs is a challenge and remains uncertain.

This thesis aims to fill these research gaps.

1.2 Aims and Objectives

The main aim of this thesis is to develop a design guidelines framework for building interpretable hierarchical fuzzy systems.

Several research questions and hypotheses underlying the overall work were identified at the beginning of this thesis. This knowledge gap has led to the formulation of the following research questions:

- Is there a way to measure the interpretability of HFSs?
- What is the best way to obtain users' perception of the interpretability?
- How can a measure of interpretability be validated?

- How can the complexity of an HFSs' structure be measured?
- How can interpretable HFSs be designed?

In order to achieve the aims stated above and to answer the research questions, the following objectives were identified:

1. To investigate the assessment of interpretability of HFSs.
2. To establish standard methodologies for capturing perceptions of the interpretability and complexity of HFSs from users.
3. To guide the measurement of interpretability of HFSs.
4. To explore the measurement of complexity in HFSs.
5. To build a design guidelines framework for interpretable HFSs.
6. To demonstrate the design guidelines framework for interpretable HFSs in a real-world example.

The first objective is discussed in Chapter 3; the second is presented in Chapter 4; the third is given in Chapter 5; the fourth is addressed in Chapter 6. Finally, the last two objectives are described in Chapter 7.

1.3 Contributions to Knowledge

This thesis makes a number of contributions to the field of interpretability of fuzzy systems, particularly for HFSs, as detailed below:

1. An interpretability assessment of HFSs is introduced. This is the first work that undertakes a formal assessment to measure the interpretability of HFSs, which has been built based on the challenges arising from the structure of HFSs, including their multiple subsystems, layers and varied topologies. Therefore, this thesis makes a major contribution to the knowledge of interpretability assessment.
2. A novel study to capture users' perception on the interpretability and complexity of HFSs is presented. Due to the subjective nature of interpretability, it is not an easy task to obtain users' perceptions, particularly of interpretability, and there exists a lack of research addressing this issue in fuzzy systems. Thus, this method of designing

the user study is a significant contribution to the knowledge of interpretability in this context.

3. Novel analysis of interpretability and complexity of HFSs based on the user study is undertaken. This is the first time that such a user study has been used to assess the relationship between interpretability and complexity in HFSs. This analysis adds insight and new knowledge into this relationship, particularly in HFSs from the user's perspective.

4. The development of an interpretability measurement of HFSs using a participatory design approach is proposed. The participatory design approach is utilised to guide the measurement of interpretability. This approach enables the subjective views of a range of users to be taken into account in shaping the measurement of interpretability of HFSs.

5. This thesis contributes to the current debate of complexity measurement with the introduction of a combined rule-based and structural complexity index, in order to establish a comprehensive measurement of complexity in HFSs.

6. A novel design guidelines framework for interpretable HFSs is proposed. The framework is based on the measurement index of interpretability and complexity of HFSs consisting of five key design guidelines for building interpretable HFS.

7. The design of an interpretable HFS for supporting decision-making in oxygenation levels to be administered with a neonatal intensive care unit (NICU) is detailed, constituting a demonstration of the proposed design framework in creating an interpretable HFS for a real-world application.

This research work has led to several refereed papers —- one journal paper (under review) and three conference papers. A complete list of publications is reported in Section 8.5.

1.4 Thesis Outline

The thesis is presented in eight chapters, as follows:

Chapter 1. *Introduction*: This chapter presents an introduction to the scope of research in this thesis. This chapter outlines a discussion on the background, motivation, research questions, research objectives and also contributions of this research.

Chapter 2. *Literature Review*: In this chapter, some of the related studies conducted by previous researchers are reviewed. The key research areas include the state of the art of the fuzzy logic systems, hierarchical fuzzy systems, interpretability, complexity and the case study of neonatal monitoring.

Chapter 3. *An Initial Approach to Measuring Interpretability of Hierarchical Fuzzy Systems*: This chapter describes the interpretability measurement for fuzzy systems and consider two essential aspects, namely complexity-based interpretability and semantic-based interpretability. However, both aspects of measurement do not take into account the structure of fuzzy systems, i.e., when it is hierarchical (HFS). Therefore, this chapter investigates the measurement of interpretability of HFSs that have multiple subsystems, layers and varied topologies. At the end of this chapter, an initial index for measuring the interpretability of HFSs is presented.

Chapter 4. *An Exploratory User Study on Perception of Interpretability in Hierarchical Fuzzy Systems*: This chapter presents an exploration user study on perception of interpretability in HFSs. The use of user study as a means to obtain and assess the interpretability and also complexity in HFSs is the highlight of this chapter. This is due to the fact such study has never been researched previously. Later, this chapter elaborates on the analysis of the relationship between interpretability and complexity of HFSs, including the descriptive analysis of the survey data, analysis of individual differences in perceived interpretability and complexity, and statistical analysis of the interpretability and complexity.

Chapter 5. *Participatory Design Approach to Guide the Measurement of Interpretability in Hierarchical Fuzzy Systems*: In this chapter, a participatory design approach is utilised to guide the measurement of interpretability that presented in Chapter 3. This approach enables the subjective views of a range of users to be taken into account in shaping the measurement of interpretability of HFSs. At the end of the chapter, a demonstration and evaluation using the participatory design approach to configure the measurement interpretability for HFSs is presented.

Chapter 6. *An Improved Complexity Measure in Hierarchical Fuzzy Systems*: Complexity is an essential component in interpretability. Therefore, this thesis looks in detail at the complexity of fuzzy systems, particularly HFSs. In FLSs, complexity is expressed by the number of rules, variables, and fuzzy terms – rule-based complexity. Several studies have used indices, e.g., the number of rules, to measure the complexity of

FLSs. However, none of the studies consider the structure of HFSs, i.e., multiple subsystems, layers and varied topologies which may also affect the complexity of HFSs. Therefore, this chapter introduces an approach to obtain a comprehensive measurement of complexity in HFSs by combining the rule-base complexity with the structural complexity.

Chapter 7. *A Framework for Guiding the Design of Interpretable Hierarchical Fuzzy Systems*: This chapter presented a design guidelines framework for interpretable HFSs. The framework is built based on the measurement index of interpretability and complexity that presented in Chapter 5 and 6 respectively. The framework consists of five key design guidelines for building interpretable HFS. Finally, a demonstration of the framework to guide the designing of an interpretable HFS on real-world is presented. Specifically, the demonstration focuses on the design an interpretable HFS for a neonatal intensive care unit (NICU).

Chapter 8. *Conclusions*: This chapter discusses the contributions and the limitations of the thesis. The recommendations for future research are also discussed in this chapter. A list of publications and oral presentations derived from this work is also reported.

Chapter 2

Literature Review

2.1 Introduction

This chapter presents relevant background information on the research subjects in order to identify gaps in the body of knowledge. The key research areas include the state of the art of FLSs, HFSs, interpretability, complexity, and a case study of neonatal monitoring. As stated in Chapter 1, the aim of this thesis is to develop a design guidelines framework for interpretable hierarchical fuzzy systems. The literature review begins with Section 2.2 which provides a brief overview of the components in FLSs including fuzzy sets, membership functions, linguistic variables, rule bases and an overview of designing FLSs. The introduction of HFSs is discussed in Section 2.3. This section includes an overview of the definition and structure of HFSs, related work on current HFSs, and the advantages and disadvantages of HFSs. Section 2.4 presents the state of the art research on the interpretability of FLSs including definition, assessment and current issues. Section 2.5 presents a summary of the complexity of FLSs as a subcomponent of interpretability. This also includes complexity measurement in other fields such as psychology, science and circuit design. Section 2.6 presents a case study on neonatal monitoring including the existing work on neonatal monitoring in the literature using FLSs. Finally, a summary of this chapter is presented Section 2.7.

2.2 Fuzzy Logic Systems

FLSs were first introduced by Lotfi Zadeh [8] in 1965 who was a Professor at the University of California, Berkeley. He introduced the notion of fuzzy sets to handle such uncertain environments and vagueness in information, and which now forms the basis for modern

control theory. For more than fifty years, FLSs provided the basis for a successful method of modelling uncertainty, vagueness and imprecision particularly in consumer products and control applications [58].

The field of FLSs has been making rapid progress in recent years. There has been an increasing number of works in many areas such as science, manufacturing, business and also in the medical domain for decision making. Many researchers [9–11, 59–61] have used fuzzy systems as a tool for controlling and modeling in many fields, proving it to be a useful technology. This is mainly due to the flexibility and ease of which knowledge can be expressed by means of fuzzy rules as well as some theoretical results in these fields [8].

Additionally, FLSs are becoming an acceptable methodology for designing robust controllers that can overcome uncertainty and imprecision [42],[62]. FLSs have been used in many problems and have proven capable in generating more robust and cost-effective solutions in the face of uncertainties. Moreover, it is often said that fuzzy solutions are more humanistic, because they offer modelling and processing of linguistic variables and rules that are easy to understand [63]. This is why one of the strengths of FLSs is their interpretability [13], particularly in applications such as knowledge extraction and decision support [2, 3]. This thesis is concerned with the concept of interpretability of FLSs and will further discuss this topic in Chapters 3, 4 and 5.

The remainder of this section presents the background and properties of FLSs including fuzzy sets, membership functions, linguistic variables, rules and an overview of designing FLSs.

2.2.1 Fuzzy Sets and Membership Functions

Fuzzy sets can be considered as an extension of classical or 'crisp' set theory. In classical set theory, an element x is either a member or non-member of set A. Thus, the membership $\mu_A(x)$ of x into A is given by [64]:

$$\mu_A(x) = \begin{cases} 1, & \text{if} \quad x \in A \\ 0, & \text{if} \quad x \notin A \end{cases} \tag{2.1}$$

Consider *tall* men, as an example. One might say that "a height more than 190 cm is a *tall* man". This statement can be represented in the form of a classical set as *tall* = $\{x|x \geq 190\}$ and the membership function characterising this set is shown in Fig. 2.1.

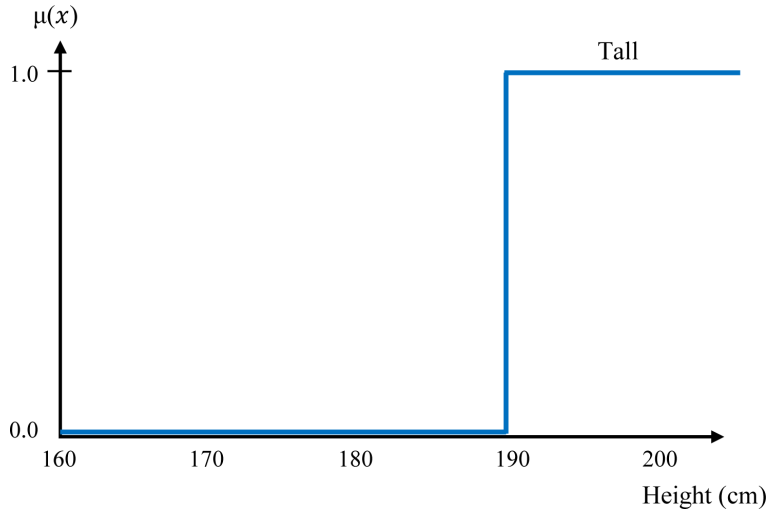


Fig. 2.1 The membership function for the crisp sets of tall men. Adapted from [65].

In contrast to classical set theory, the fuzzy set methodology introduced the concept of *degree* to the notion of membership. More formally, a fuzzy set A of a universe of discourse X (the range over which the variable spans) is characterised by a *membership function* $\mu_A(x) : X \rightarrow [0,1]$ which associates with each element x of X a number $\mu_A(x)$ in the interval $[0,1]$, with $\mu_A(x)$ representing the *grade of membership* of x in A. The precise meaning of the membership grade is not rigidly defined, but is supposed to capture the 'compatibility' of an element to the notion of the set.

By using the above example again, the statement "a height more than 190 cm is a *tall* man" can be represented in the form of the fuzzy set shown in Fig. 2.2. In comparison with a classical set in which only sharp boundaries are permitted, the concept of membership degree in fuzzy sets allows fuzzy or blurred boundaries to be defined. In Fig. 2.2, it can be seen that a height of 188 cm can also be considered as *tall* but with a lesser degree of membership than for 190 cm (i.e., $\mu_{tall}(x = 188) = 0.90$); whereas in a classical set the degree of membership is zero (i.e., a height of 188 cm does not belong to the set of *tall* at all). Fuzzy sets provide the tools to represent problems in everyday language, and it is this property that provides a problem solving technique that mimics the characteristics of human reasoning and decision making.

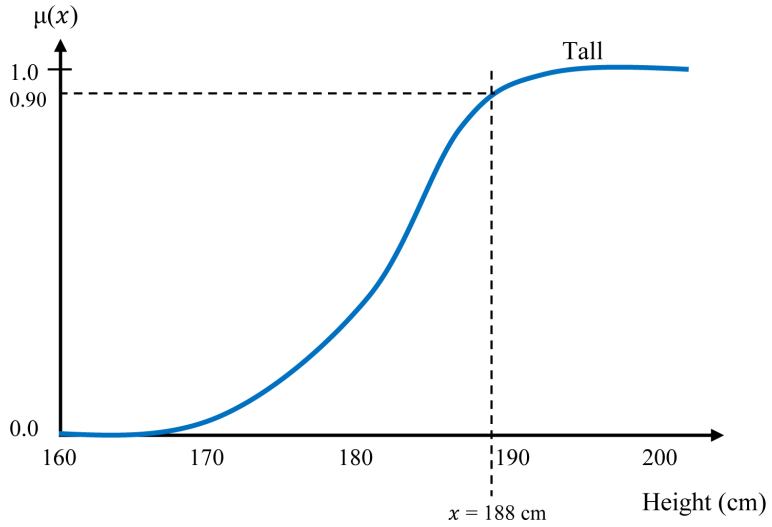


Fig. 2.2 The membership function for the fuzzy sets of *tall* men. Adapted from [65].

The membership function, usually given the designation of μ, is the principle of fuzzy sets. A membership function defines how each point in the universe of discourse is mapped to a degree of membership usually taken as a real number in the interval $[0, 1]$. For example, Figs. 2.3, 2.4, 2.5 and 2.6 show some of the membership functions that are commonly be used for the fuzzy sets. Note that, in this example, the membership functions were generated using the Fuzzy Toolbox found in the R programming language [66].

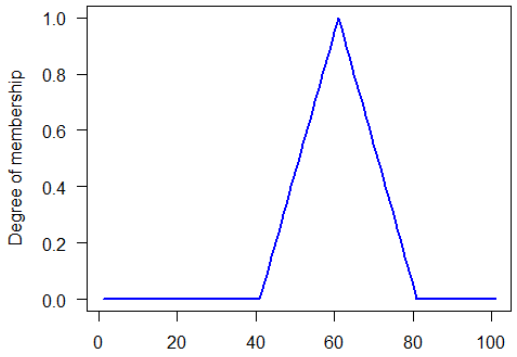


Fig. 2.3 Triangular MF

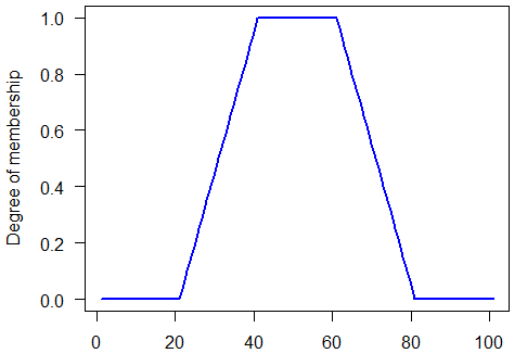


Fig. 2.4 Trapezoidal MF

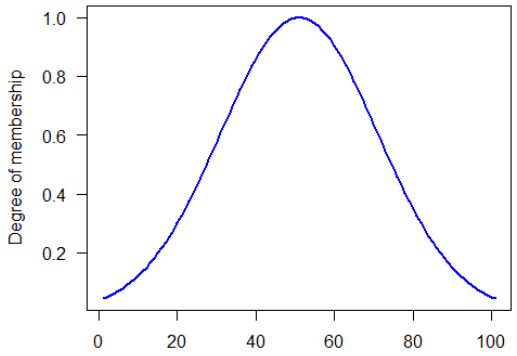


Fig. 2.5 Gaussian MF

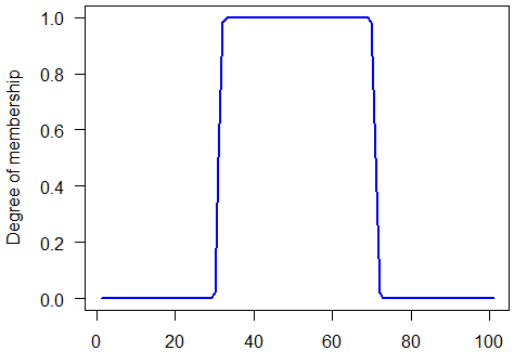


Fig. 2.6 Generalized Bell MF

2.2.2 Linguistic Variables, Values and Rules

The term 'linguistic variable' was introduced by Zadeh [16] to refer to a variable whose values are in the form of "linguistic expressions" rather than numerical values. In the example shown in Fig. 2.2, 'height' is a linguistic variable with a linguistic value '*tall*'.

Other possible linguistic values for the linguistic variable 'height' could include terms such as '*short*' and '*medium*'. Each linguistic value is represented by a fuzzy set (membership function) in which the characteristic of each fuzzy set is dependent on the context of the particular problem. Although these linguistic terms are very subjective, they might be interpreted as for example in Fig. 2.7.

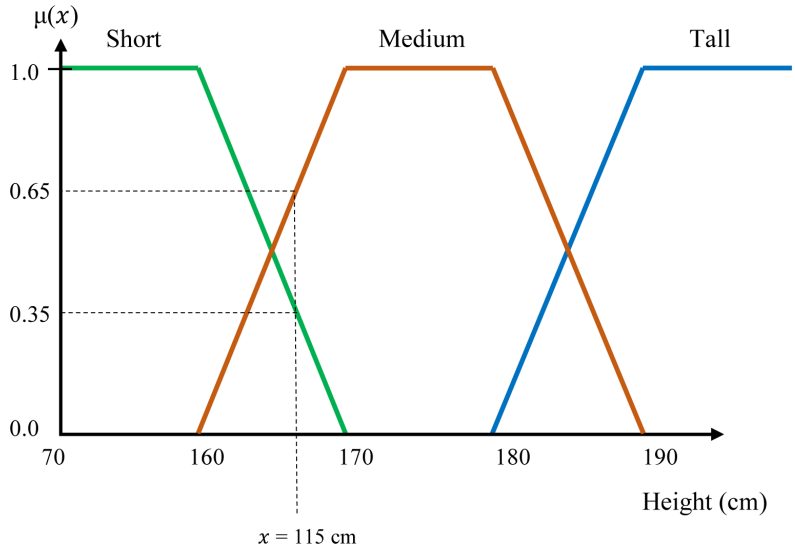


Fig. 2.7 Membership functions for the linguistic variable 'height'

As shown in Fig. 2.7, the 'height' is partitioned into three fuzzy sets – *short*, *medium* and *tall*. These fuzzy sets are partially overlapping. Hence, it can be seen that the height of 155 cm has partial membership in both the fuzzy set *short* and the fuzzy set *medium*, where

$$\mu_{short}(x = 155) = 0.35,$$
$$\mu_{medium}(x = 155) = 0.65$$

Fuzzy sets in the form of linguistic expressions can then be connected using IF-THEN statement(s), to form a series of fuzzy rules. Fuzzy rules are represented in the form:

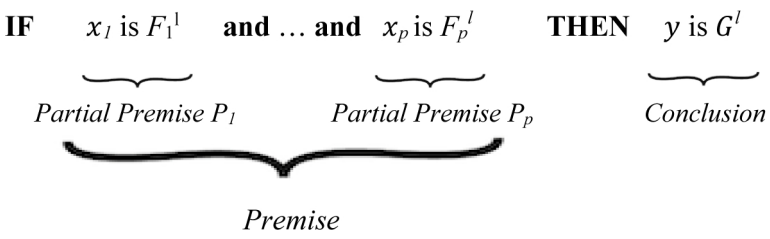
IF antecedent(s) THEN consequent(s)

Interpreting an if-then rule involves two distinct parts: first evaluating the antecedent, which involves fuzzifying the input, and second applying that result to the consequent known as implication. This is known as fuzzy reasoning [67]. Moreover, as discussed in Mendel et al. [68], a fuzzy rule has p inputs $x_1 \in X_1, \ldots, x_p \in X_p$, and one output $y \in Y$, and is characterized by M rules, where the lth rule has the form

$$\begin{aligned} R^l : \text{IF } x_1 \text{ is } F_1^l \text{ and} \ldots \text{and } x_p \text{ is } F_p^l, \\ \text{THEN } y \text{ is } G^l, \qquad l = 1, \ldots M. \end{aligned} \tag{2.2}$$

The rule represents a fuzzy relation between the input space $X_1 \times \cdots \times X_p$ and the output space, Y, of an FLS. On the subject of rule structure, Mamdani rules [69] (whose conclusion

is a fuzzy set) are universally acknowledged as the more interpretable kind of rules. As discussed in [15], from the interpretability point of view, Mamdani rules are preferred because they are linguistic rules of the form:



Rule premises are made up of tuples (input variable, fuzzy sets). For instance, x_p is the name of the input variable and F_p^l represents the fuzzy sets (or linguistic term) for that variable. Notice that the absence of an input in a rule means that variable is not considered when firing that rule. Besides Mamdani rules, there are many other rule formats. One of the commonly used rules is the well-known Takagi–Sugeno rules [70], where the conclusion is a linear combination of the input values. However, only the Mamdani rules structure is used in this thesis because of its increased interpretability, which is a goal of the design guidelines this thesis will produce.

2.2.3 Overview of Designing Fuzzy Logic Systems

Fig. 2.8 shows the five interconnected components of a fuzzy system. The fuzzification component computes the membership grade for each crisp input variable based on the membership functions defined. The inference engine then conducts the fuzzy reasoning process as defined by a set of fuzzy rules, by applying the appropriate fuzzy operators in order to obtain the fuzzy sets to be accumulated in the output variable. The defuzzifier transforms the output fuzzy set to a crisp output by applying a specific defuzzification method.

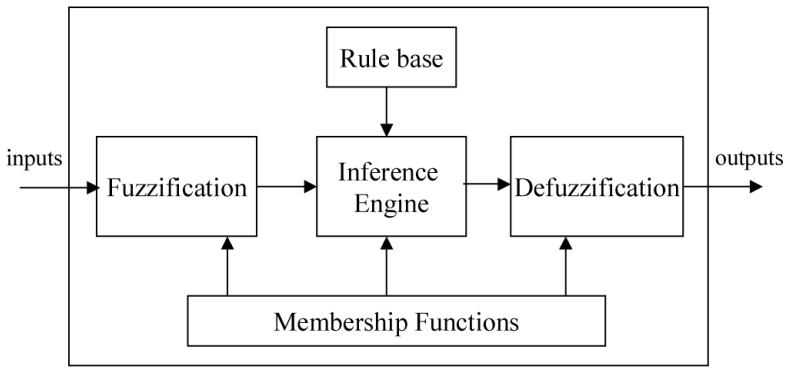


Fig. 2.8 Components of Fuzzy Logic Systems.

In general, the main steps as shown in Fig. 2.8 performed in the FLS are as follows:

- Analyse and understand the problem in consideration.
- Determine the linguistic variables (the inputs and outputs). For each linguistic variable, identify the linguistic values and define the fuzzy sets (membership functions).
- Identify and define the fuzzy rule set.
- Choose the appropriate methods for fuzzification, fuzzy inference and defuzzification.
- Evaluate the system.

Recently, as discussed earlier, the field of FLS research has been making rapid progress in many areas such as science, manufacturing, business and also the medical domain for decision making. This thesis is concerned with FLSs particularly in the medical domain, and therefore, this is reviewed in the following section.

2.2.4 Fuzzy Logic in Medical Domains

Real world knowledge is characterised by incompleteness, inaccuracy and inconsistency. Fuzzy set theory introduced by Zadeh [8] makes it possible to define inexact medical entities. It provides an excellent approach in approximating medical concepts. Furthermore, fuzzy logic also provides reasoning methods to approximate inference [71], easiness of understanding, low computational costs, and the ability to be incorporated into systems that the human experts use. These attributes make this approach an extremely interesting option to represent medical models [72].

Fuzzy logic has been successfully applied in many medical domains such as Breast Cancer [73–75], Heart Disease [76–78], Lung cancer [79–81], Liver [82–84], Diabetes [85–87], and Newborn infants [72, 88, 89]. From the published literature , it may be concluded that fuzzy logic has proven reliable and effective in many medical domains since it can address the specific problem of disease classification in the presence of uncertain or vague knowledge of a linguistic nature [90]. However, as discussed in [43, 47, 57], the conventional FLS is rather limited, suffering from so-called "rule explosion" (discussed below) which can lead to a long execution time in computing the output for a complex system. In addition, as reported in Gacto et al. [23], researchers have usually focused on the improvement of the accuracy of the FLS models obtained without paying special attention to the interpretability.

Therefore, it is worth emphasising that, a hierarchical fuzzy system for neonatal care is essential in order to overcome the limitation of conventional FLS whilst improving the interpretability of the model. One of the outputs of this thesis is such an FLS, specifically a novel HFS that addresses these issues.

2.2.5 The Curse of Dimensionality

As mentioned earlier, FLSs have been successfully applied in many areas, especially those that require the handling of uncertainty and imprecise information. However, a fundamental issue of FLSs is that as the number of variables increases, the number of rules increases exponentially [45]. Thus, the complexity of an FLS also increases exponentially with the number of variables involved. This phenomenon was referred to by Bellman [91] as the '*curse of dimensionality*'. Zeng and Keane in their study [92] viewed the curse of dimensionality from three perspectives.

- *Rule dimensionality*

  The total number of rules in the fuzzy rule base increases exponentially with the number of input variables.

- *Parameter dimensionality*

  The total number of parameters in the mathematical formulas of fuzzy systems increases exponentially with the number of input variables.

- *Data or information dimensionality*

  The number of data or knowledge sets required to identify fuzzy systems increases exponentially with the number of input variables.

To address this problem, several methods have been proposed for optimizing the size of the rule base such as rule selection [33–37], feature selection [38], rule interpolation [39, 40], singular-value decomposition-QR [41], evolutionary algorithms [42], fuzzy similarity measures [22] and rule learning [14]. Most of these techniques involve the modification of the original input variables of fuzzy systems that can cause the loss of the original meaning of variables [2] and also reduce the model interpretability. Nevertheless, one effective way to deal with this problem is through the use of a special type of FLS, namely hierarchical fuzzy systems (HFSs), that will reduce the number of rules while retaining the original meaning of variables [43–45].

2.3 Hierarchical Fuzzy Systems

2.3.1 Introduction

As introduced earlier, FLSs have successfully been used in various fields of study involving handling uncertainties and imprecise information. However, as problems have rapidly become more complex, requiring larger sets of input variables, the conventional FLS has encountered limitations with the increased number of rules concerning leading to increased times taken for fuzzy inference [43],[44].

At present, the important issues in the conventional FLS are reducing the number of rules involved and their corresponding computation requirement [48, 46, 93]. One effective way to deal with this problem is through the use of a special type of FLS, namely hierarchical fuzzy systems (HFSs) [43–45, 47, 94, 95].

HFSs were first introduced by Raju et al. [43],[44] and are an important approach to overcome this problem. The Iris dataset [96] is used as an example to provide a formal definition of HFSs. The Iris Dataset consists of four input parameters, which are *sepal length (SL), sepal width (SW), petal length (PL), petal width (PW)* and one output variable which is Flower; *Setosa*, *Versicolor*, and *Virginica*.

HFSs are characterized by composing the input variables into a collection of low-dimensional FLSs—fuzzy logic subsystems [43],[44]. HFSs can be illustrated as a cascade structure where the output of each layer, called intermediate output, is considered as an input to the following layer. For example, Figs. 2.9 and 2.10 show the topology of FLS and HFS for iris classification respectively. The HFS for iris classification is obtained by decomposing the input variables of the FLS as shown in Fig. 2.9 into a collection of subsystems, namely

FLS_1, FLS_2 and FLS_3 as shown in Fig. 2.10. In addition, the intermediate outputs *Sepal (S)* and *Sepal_Petal (S_P)* are produced in this HFS.

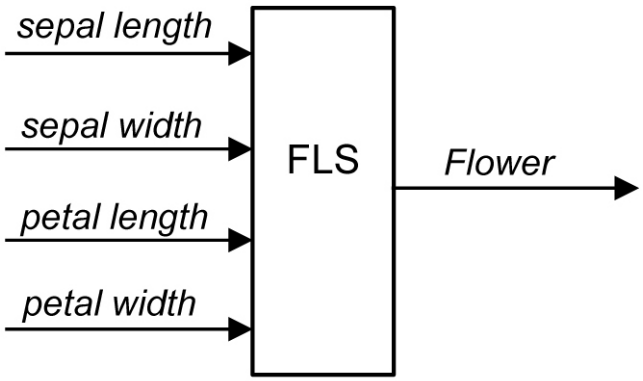


Fig. 2.9 FLS: Iris classification

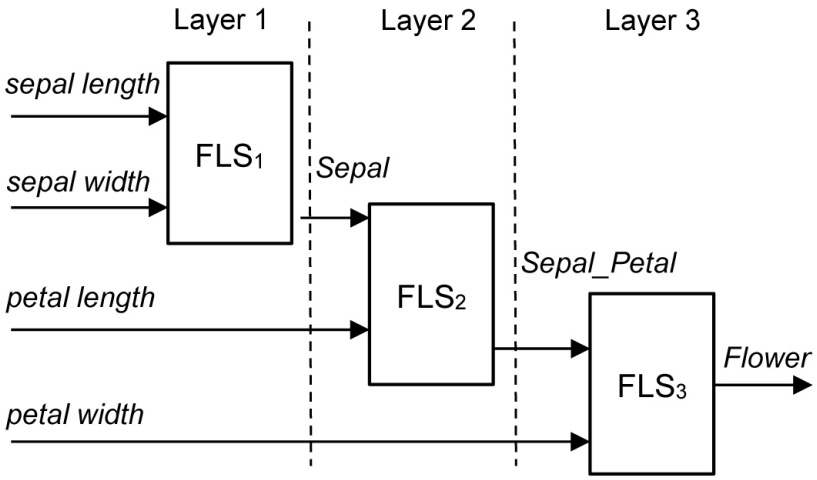


Fig. 2.10 Serial HFS: Iris classification

HFSs can be viewed as a functional decomposition of FLSs [97]. For instance, the FLS and HFS for iris classification as shown in Fig. 2.9 and Fig. 2.10 respectively can be presented, from a functional point of view as

$$Flower = F(SL, SW, PL, PW) \implies Flower = f_3(f_1(SL, SW), f_2(PL, PW)).$$

An FLS that goes from one layer as shown in Fig. 2.9 to two layers as in Fig. 2.10 has a smaller number of rules when considering a fully specified rule base. The most extreme reduction of rules will be if the structure of the HFS has two input variables for each low dimensional FLS and has $(n-1)$ layers [43], where n is the total number of input variables as can be seen in Fig. 2.10. If we consider two input variables per low-dimensional FLS and define m fuzzy sets for each input variable, including the intermediate output variables $y_1, ..., y_{n-2}$, the total number of rules (R_{HFS}) is a linear function [45] of the number of input variables n and can be expressed as:

$$R_{HFS} = (n-1)m^2. \tag{2.3}$$

Note that equation (2.3) only valid for HFS structures that consists of two input variable per FLS subsystem in a serial structural form, for example as in Fig. 2.10. Conversely, in conventional FLSs, the number of rules increases exponentially with the number of input variables [46, 43]. Suppose there are n input variables and m fuzzy sets for each input variable, then the number of rules (R_{FLS}) needed to construct a complete fuzzy system with fully specified rule base (using the “AND" logical connective) can be expressed as:

$$R_{FLS} = m^n \tag{2.4}$$

From these equations (2.3) and (2.4), it is clear that the total number of rules in the FLSs (R_{FLS}) is always greater than or equal to the equivalent HFSs (R_{HFS}). For example, Fig. 2.9 and Fig. 2.10 show an FLS and HFS with 4 input variables ($n = 4$) and, assuming that 3 fuzzy sets ($m = 3$) are defined for each input variable, the total number of rules for this FLS is $R_{FLS} = m^n = 3^4 = 81$ whereas for the HFS, the total number of rules is $R_{HFS} = (n-1)m^2 = (4-1)3^2 = 27$.

FLS	**HFS**
Rules: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa*	Rules: FLS_1: IF *sepal length* is *small* AND *sepal width* is *small* THEN *Sepal* is *small* FLS_2: IF *petal length* is *small* AND *petal width* is *small* THEN *Petal* is *small* FLS_3: IF *Sepal* is *small* and *Petal* is *small* THEN *Flower* is *Sentosa*

Fig. 2.11 Example of rules in FLS and HFS

Fig. 2.11 shows an example of the rule structure in FLS and HFS for iris classification. Rules in HFSs are decomposed from FLSs into small rules in multiple subsystems, namely FLS_1, FLS_2 and FLS_3. By doing this, the rules structure in HFSs are reducing the rule length since it creates a smaller rule for each subsystem. Thus, rules in HFSs are simpler than those in FLSs because the number of variables per subsystem is lower [54]. Consequently, this may improve the human readability of rule base in HFSs [4].

2.3.2 Structure of Hierarchical Fuzzy Systems

As presented earlier, an HFS is a special type of FLSs, which has the structure of multiple subsystems, layers and different topologies. In the following subsection, the structure of HFSs is reviewed.

2.3.2.1 Multiple Subsystems

HFSs are characterized by having several subsystems that contribute to the computation of the final solution. Each subsystem will have a small number of inputs and outputs, a smaller rulebase, and serve a single purpose [98]. Then, these subsystems are linked in such a way that the output of a subsystem is the input of subsequent ones. For example, in Fig. 2.12, subsystem FLS_1 has input variables *sepal length* and *sepal width*, and *Sepal* output. This system has its own task which is to determine the *Sepal* classification. Also, it contains a small rule base which can be illustrated as follows:

- IF *sepal length* is *small* AND *sepal width* is *small* THEN *Sepal* is *small*

- IF *sepal length* is *small* AND *sepal width* is *medium* THEN *Sepal* is *small*
- IF *sepal length* is *small* AND *sepal width* is *large* THEN *Sepal* is *medium*
- IF *sepal length* is *medium* AND *sepal width* is *small* THEN *Sepal* is *small*
- IF *sepal length* is *medium* AND *sepal width* is *medium* THEN *Sepal* is *medium*
- IF *sepal length* is *medium* AND *sepal width* is *large* THEN *Sepal* is *large*
- IF *sepal length* is *large* AND *sepal width* is *small* THEN *Sepal* is *medium*
- IF *sepal length* is *large* AND *sepal width* is *medium* THEN *Sepal* is *large*
- IF *sepal length* is *large* AND *sepal width* is *large* THEN *Sepal* is *large*

2.3.2.2 Layers and Topologies

Also, as mentioned earlier, HFSs are produced by decomposing the input variables in FLSs into multiple low-dimensional FLSs. By doing this, several *layers* are produced in HFSs. Based on the same input variables, HFSs may be produced using different topologies, e.g., serial and parallel [54]. The parallel HFS can have more than one low-dimensional FLS per layer, while serial HFSs use strictly one FLS per layer as shown in Fig. 2.12 and Fig. 2.10 respectively. Thus, these topologies commonly have a different number of layers. For example, Fig. 2.12 and Fig. 2.10 show two different HFS topologies using the same four input variables. Both HFS topologies use the same number of subsystems, but with different numbers of layers in their structure.

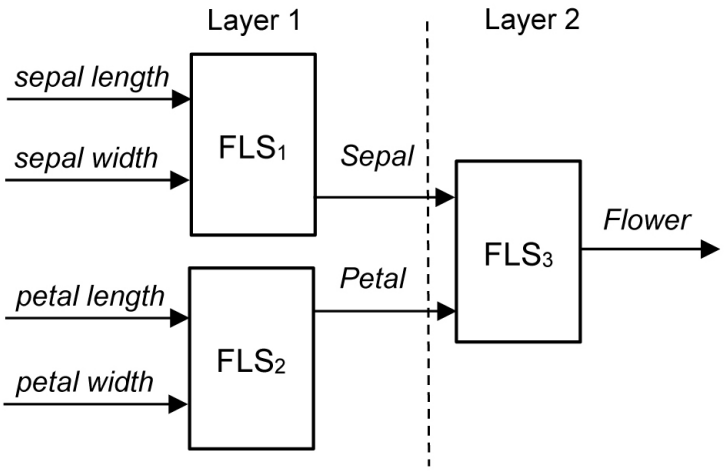


Fig. 2.12 Parallel HFS: : Iris classification

Note that, Campello and Caradori do Amaral have claimed that, for a topology with two input variables, when they are put into a fuzzy system of which the output is combined with another input variable into the second fuzzy system, and this procedure continues until all input variables are used, this topology will lead to the most parsimonious models [95]. Raju et al. [43, 44] have further claimed that the most influential input variables should be chosen as the system variables in the first layer, the next most important variables should be chosen in the second level, and so forth. However, there is an argument on this issue. Wang et al. [47] show in their study that there is no general conclusion about which inputs are more influential to the system output.

2.3.3 Approaches to Building Hierarchical Fuzzy Systems

Torra [57] has claimed that building an HFS is a challenging task because of the need to define the architecture of the HFS (the subsystems, the input variables of each subsystem, and the interactions between subsystems), as well as the rules of each subsystem. Also the fundamental issue with HFSs is that they must handle more abstract values in linguistic variables produced as intermediate outputs, and are consequently more difficult to design. Additionally, a study conducted by Cross and Sudkamp [53] concluded that decomposition of an HFS cannot reproduce the same result to the original FLS but rather generate approximations with fewer fuzzy rules.

Despite these issues, several researchers [48, 46, 49–53] have proposed a new approach in designing HFSs. Therefore, the approaches to building HFSs will be discussed in detail in the following subsections.

2.3.3.1 Structured-Hierarchical Fuzzy System (S-HFS)

In the conventional HFS, the outputs of FLSs in the first layer which are called intermediate outputs are used as input variables of the next layer. However, these intermediate outputs are more abstract and possess less meaning in the real-world model they are trying to represent. Thus, when the outputs act as input variables for the next layer, the next fuzzy rule will be even more abstract and consequently be more difficult to design. This problem becomes more crucial when the HFS grows larger.

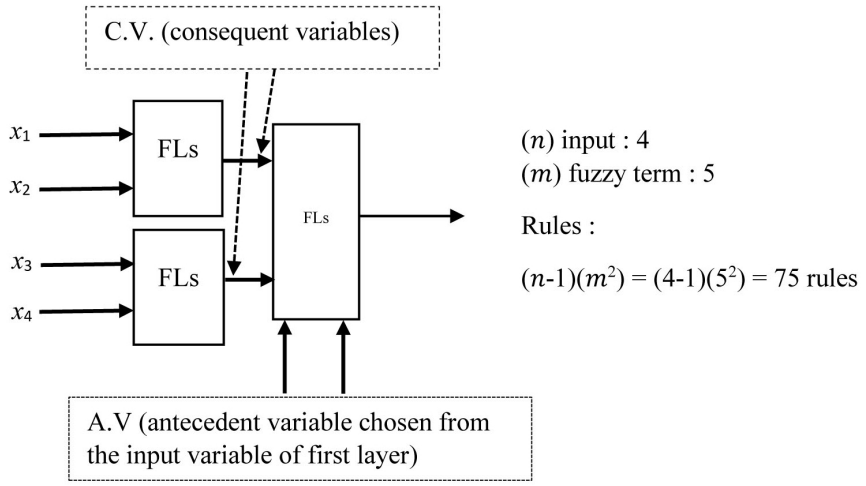


Fig. 2.13 Structured Hierarchical Fuzzy System. Adapted from [48].

Joo and Lee [48] proposed the structured-HFS (S-HFS) to overcome this problem as shown in Fig. 2.13. In the first layer of a S-HFS, all input variables (x_1, x_2, x_3, x_4) can be used as antecedent linguistic variables whereas in the second layer of a S-HFS, the outputs of previous layers are not used as antecedent linguistic variables because they have a more abstract meaning. However, they are used in the THEN-parts of fuzzy rules in this layer. The antecedent linguistic variables of the next layer may be chosen from the input variables of previous layers that have physical meaning. Therefore, only input variables with concrete meaning are used in IF-parts of fuzzy rule and consequently the fuzzy rules are easier to design.

2.3.3.2 Limpid-Hierarchical Fuzzy System (L-HFS)

Lee et al. [46] have also proposed a new method to overcome the issue of intermediate outputs which do not possess physical meaning in the middle layer. From the study of S-HFSs in Joo and Lee [48], they introduced a new mapping rule base called a Limpid-Heirarchical Fuzzy System (L-HFS) in order to obtain the HFS rule base. To illustrate this method, an HFS as in Fig. 2.14, is used that consists of three input variables: x_1, x_2 and x_3. F_1 and F_2 are subsystems for this HFS at layer 1 and layer 2, respectively.

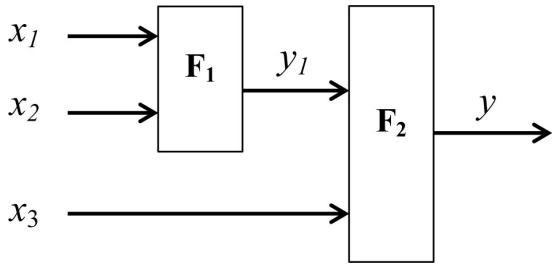


Fig. 2.14 Conventional HFS. x_1, x_2, x_3 are the input while y_1 and y are the intermediate output and output respectively.

By defining each variable using three linguistic terms, where N is negative, Z is zero and P is positive, the complete rules for a flat FLS consists of 27 $(3 \times 3 \times 3 = 27)$ rules, as shown in Table 2.1. From these complete rules, Algorithm 1 is then applied to the mapping rule for subsystems F_1 and F_2 of the HFS.

Algorithm 1 L-HFS

Step 1: Determine the form of HFS and label the subsystem namely; F_1 and F_2 as shown in Fig. 2.14.

Step 2: From Fig. 2.14, fix the input variables x_1,x_2 and use the sorting process to the subsystem F_1, in order to obtain the new output mapping variables.

Step 3: Repeat Step 2 up to and including the penultimate layer.

Step 4: In the final layer (in this example is layer 2), tabulate the final subsystem (F_2) rules from the involved linguistic outputs (y_1) of the previous layer's mapping variables.

Table 2.1 The rule by random access. P, Z, N are linguistic terms for each input variable.

	x_1								
x_3	P			Z			N		
	x_2			x_2			x_2		
	P	Z	N	P	Z	N	P	Z	N
P	P	N	Z	N	P	N	N	Z	N
Z	Z	N	P	N	P	N	N	P	N
N	P	P	Z	P	Z	N	N	Z	N
	↓	↓	↓	↓	↓	↓	↓	↓	↓
	A	B	C	B	D	E	E	C	E

By implementing Algorithm 1 (as in Step 2), five new mapping variables (A, B, C, D and E) are obtained for intermediate output $y1$; the process of sorting is shown in Fig. 2.15. Tables 2.2 and 2.3 show the rules obtained using Algorithm 1 and the total number of rules for this HFS is 24 rules (9 in F_1 and 15 in F_2). The detailed process of Algorithm 1 can be seen in [46].

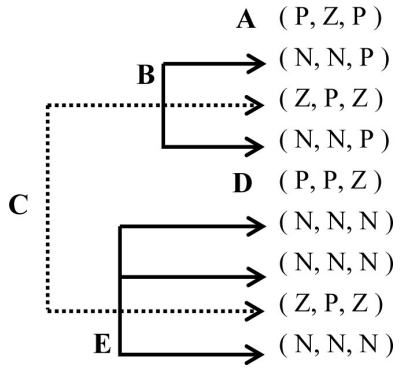


Fig. 2.15 The sorting process in L-HFS. Adapted from [46].

Table 2.2 The rules for F_1

x_2	x_1		
	P	Z	N
P	A	B	E
Z	B	D	C
N	C	E	E

Table 2.3 The rules for F_2

y_1	x_3		
	P	Z	N
A	P	Z	P
B	N	N	P
C	Z	P	Z
D	P	P	Z
E	N	N	N

Lee et al. [46] claimed that, by implementing a L-HFS, a smaller rule base is achievable with the same input-output model as in the original FLS. Furthermore, the design of the fuzzy rules involved in this approach is easier than in a conventional HFS. The method was further investigated by Masmoudi et al. [51]. In their study, an HFS was applied to compute the optimal control strategies of discrete large-scale nonlinear systems. They implemented an L-HFS algorithm to obtain the rule base of their system and proved that this approach is easy for designing the fuzzy rules involved .

2.3.3.3 Two-Layer Hierarchical Fuzzy Systems

Joo and Sudkamp [49] proposed a method for converting a multidimensional FLS to a two-layer hierarchical FLS that reduces the number of rules and also improves the run-time efficiency. For the first layer, they used linearly dependent rule (m^2) and for the second layer the applied the Takagi-Sugeno-Kang method to produce the rules ($m^{(n-2)}$) as shown in Fig. 2.16. In their research, they also proposed a solution such that optimal number of input variables should be selected for the first layer.

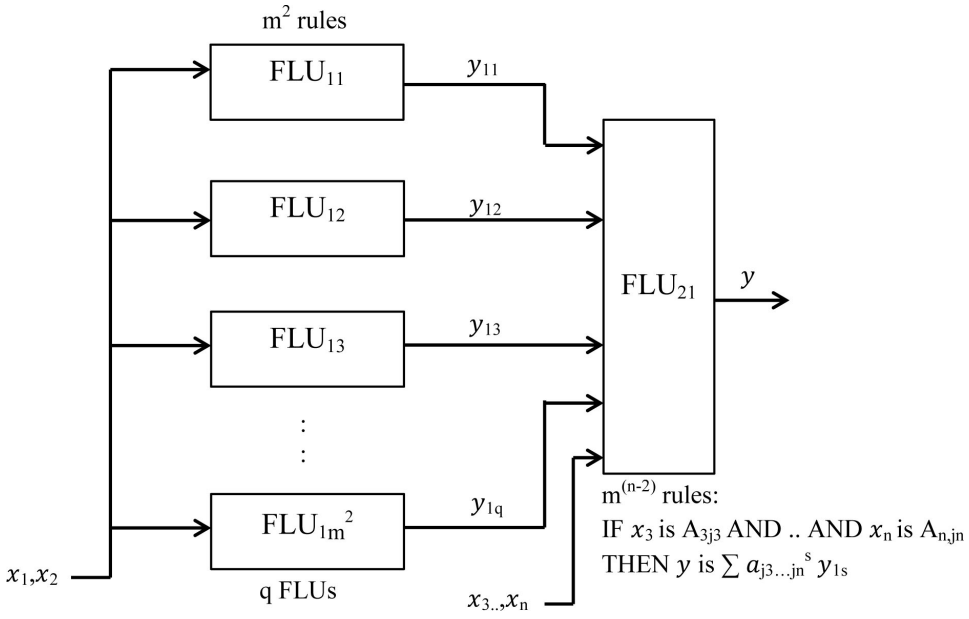


Fig. 2.16 Two-layer Hierarchical Fuzzy System. Adapted from [49].

For their experiment, the ball and beam example in [99] was applied. It was concluded that their method does not only give fewer fuzzy rules, but also gives equivalent results to the original system. However, the similarities between the result obtained from their proposed HFS and conventional FLS were not justified in their work.

2.3.3.4 Hierarchical Fuzzy Associative Memory (HIFAM)

Fuzzy Associative Memories (FAM) form a lattice partition of the input space which yields a very intuitive set of if-then-rules. Yet, this approach cannot be used for problems with many input dimensions because the number of possible rules grows exponentially with the number of inputs.

A new method of rule generation for hierarchical fuzzy systems (Hierarchical Fuzzy Associative Memory, (HIFAM)) was introduced by Holve [50] to overcome the problem with many input dimensions. A HIFAM is structured as a binary tree and overcomes the exponential growth of the rulebases when the number of inputs increases and is suited for approximation and classification problems as shown in Fig. 2.17.

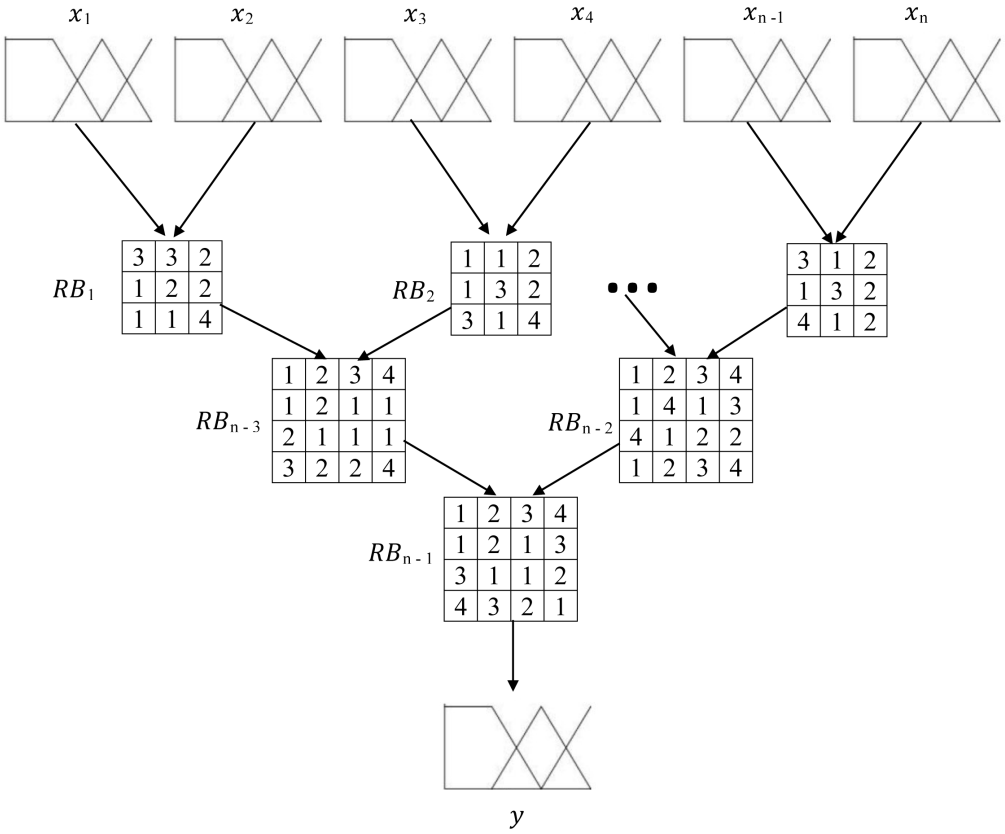


Fig. 2.17 HIFAM with n inputs x_n, one output y and n-1 rule base RB_{n-1} (adapted from [50]).

The presented approach is an efficient method for automatic generation of fuzzy rules for hierarchical fuzzy systems. It is well suited for classification problems as well as for regression tasks and competes well with existing classification and regression techniques like neural networks or decision trees.

However, among the proposed approaches, none of them has explained whether their approach can also improve the interpretability of HFSs. Several aspects should be considered in capturing interpretability of HFSs such as semantic interpretability, meaningful fuzzy sets, and intermediate variables of HFSs. For example, if the layers in HFSs have meaning, they can be useful for interpretability.

2.3.4 Advantages of Hierarchical Fuzzy Systems

In the most of the cases, the underlying idea behind the construction of HFSs is to cope with the complexity of problems. Here are some other advantages of the implementation of the HFS.

2.3.4.1 Rule reduction

A very undesirable problem may arise when the number of system inputs is large, causing the total number of rules to grow exponentially [50], [47]. This also enormously increases the computational complexity of the rule base as well as the difficulty in building a fuzzy system. This phenomenon is known as the *curse of dimensionality*, which may be overcome by hierarchically restructuring the fuzzy subsystems, thus resulting in the number of rules to linearly increase and this will reduce the computational complexity [100],[95]. Also, Brown and Harris suggest in [101] to use a hierarchical structure of fuzzy rule bases to cause the number of rules to grow linearly.

It can be concluded that one of the main purposes of using HFSs is to minimize the computational complexity and the size of rule base, which in turn may reduce the need of large system memory and speed up the processing time [100].

2.3.4.2 Improving the Interpretability

Recently, the interpretability of HFSs has been the subject of numerous researchers. As mentioned earlier, the main objective of HFSs is to minimize the computational complexity and the size of rule base. Thus, reducing the model complexity is also a way to improve the system interpretability, i.e., a system with only a few rules requires a minor effort to be interpreted [56]. Also, Salgado and Cunha in [102] claimed that the small number of rules of each fuzzy sub-system can help to avoid this problem, without degrading global model accuracy. In some cases, it is assumed that the linguistic terms and their membership functions of such input variables are the same in each fuzzy sub-system. The purpose of this assumption is to keep interpretation consistency and for simplification [92].

According to Benítez and Casillas [54], the interpretability of HFSs is good due to several reasons: "*(1) the hierarchical structure generates a lower number of variables in each subsystem, (2) the algorithm does not generate artificial linking variables, and so, all of the variables are interpretable because they belong to the system, (3) the rules are simpler because the number of variables per subsystem is lower. At the same time, the accuracy is maintained or improved*". However, it should be noted that this statement is only valid in the context of their application, and it cannot be generalised.

2.3.4.3 Trade-off between Accuracy and Interpretability

Shukla and Tripathi in their survey [103] claimed that HFSs are one of the methods to maintain a good trade-off in the design of complex fuzzy systems related to rules, rule bases,

membership functions, and fuzzy partitions. Also, Delgado et al. in [104] have shown that HFSs can be adjusted not only to improve the model's performance (accuracy), but also to guarantee the interpretability of the resulting fuzzy models. In addition, linguistic modeling using a hierarchical knowledge base can be regarded as a search for a decomposition of a non-linear system that gives a desired balance between the interpretability and the accuracy of the model [105].

However, the choice between how interpretable and how accurate the model must be usually depends on the user's needs for a specific problem and will condition the kind of FLSs selected to model it [106].

2.3.4.4 Universal approximation

Zeng and Keane [92] investigated the approximation capabilities of hierarchical fuzzy systems. This analysis shows that, compared to standard fuzzy approximation, hierarchical fuzzy approximation can significantly reduce both the number of rules and parameters required to achieve a desired degree of accuracy. Wang proved that a certain class of HFS can serve as a universal approximator to any real continuous function on a compact set [45],[47], and this was again proven by Joon and Lee [93].

2.3.5 Issues of the interpretability in Hierarchical Fuzzy Systems

Despite its success in handling the explosion of the number of rule bases, there are some issues in relation to HFSs that must be addressed in order to investigate the interpretability of HFSs.

2.3.5.1 Interpretability assessment

As mentioned earlier, none of the researchers have investigated how interpretability can similarly be measured using indices in HFSs. Up to now, no specific interpretability assessment has been introduced to measure interpretability of HFSs. The measurement of interpretability is important for HFSs in order to:

(i) understand the concept of interpretability in HFSs, i.e., the effect of HFS architecture on interpretability.

(ii) measure interpretability of HFSs, i.e., to select the most interpretable HFS topologies, e.g. parallel or serial.

(iii) compare the interpretability between FLSs and HFSs, i.e., to show that HFSs are more interpretable than FLSs.

In this thesis, the issue of measuring the interpretability of HFSs is explored. Thus, a measure of assessing interpretability of HFSs is proposed as presented in Chapter 3 and further developed in Chapter 5.

2.3.5.2 Intermediate Outputs do not Possess Physical Meaning

In the conventional HFS, the outputs of FLSs in the first layer, which are called intermediate outputs, are used as the input linguistic variables of the next layer. However, these intermediate outputs do not possess physical meaning, i.e. are abstract and bear little meaning to the real world. Thus, when the outputs act as input variables for the next layer, the subsequent fuzzy rule will have less physical meaning and consequently is difficult to design. Furthermore, this problem will become more crucial when the HFS grows larger [46],[93],[95]. Thus, the problem of intermediate outputs with little to no physical meaning will reduce the interpretability of HFSs [93].

Many researchers have proposed different types of HFS [48], [46],[51], [52] to tackle these issues. However, none of these methods has acknowledged or described the advantages of interpretability.

Note that this issue is very important to HFSs. However, at the moment, this thesis is not focussing on solving this issue as it is a challenging concept and this will be a suitable recommendation for future work. Consequently, only the HFS examples that already have a clear semantic meaning for all input, intermediate output and output variables are used in this thesis. For instance, the meaningful intermediate output in HFSs are obtained by involving human experts such as doctors and nurses, as discussed in Chapter 7.

2.3.5.3 The Structure of HFSs

As discussed earlier, in Section 2.3.2, the questions arise on the following issues:

(i) Which topology of HFSs can offer better interpretability?

(ii) What is the optimum number of layers when designing an interpretable HFS?

(iii) Which aggregation function should be used to aggregate multiple interpretability values of subsystems in HFSs?

(iv) How can the rule bases for each subsystem be designed?

The primary challenge to build interpretable HFSs is due to their structures. To tackle this issue, in this thesis, a design guidelines framework to build interpretable HFSs is proposed and discussed in Chapter 7.

2.3.5.4 Input Variables Selection in Each Subsystem

As discussed in [43, 44], it is claimed that in a hierarchical structure, typically, the most influential input variables are chosen as the system variables in the first layer, the next most important variables are chosen in the second layer, and so forth. However, a question arose:

How are the most influential variables selected?

There is no general conclusion as to which inputs to the hierarchical fuzzy system are more influential to the output, although rough observations were drawn for some particular fuzzy systems [47].

In summary, HFSs has been claimed as a useful method to overcome the problem of the curse of dimensionality in FLSs, thus reducing the model complexity and improving model interpretability. However, the interpretability is very problematic to understand since it is of a subjective nature. This thesis intends to understand the concept of interpretability of HFSs. In doing so, in the next section, the state of the art of research into the interpretability of FLSs is presented.

2.4 Interpretability

2.4.1 State of The Art

In the literature, interpretability is often known as a synonymous to comprehensibility, intelligibility, transparency, understandability, readability, etc. But it is still difficult for us to answer the questions “What does interpretability actually mean?” and “What are the definitions of intelligibility, transparency, understandability, readability?" Indeed, it is difficult to find a formal definition in the literature about interpretability and others synonyms due to inherent subjectivity of the topics [14],[107].

For an informal approach to obtain the definition of interpretability and synonyms, the definitions were retrieved from online dictionaries, namely *Dictionary.com*[1] and *Merriam-Webster*[2], that can be seen in Table 2.4.

[1] http://www.dictionary.com/
[2] https://www.merriam-webster.com/

Table 2.4 Definitions of various terms related to interpretability from online dictionaries

Term	**Dictionary.com**	**Merriam-Webster**
Interpretability	to give or provide the meaning of; explain; explicate; elucidate.	to explain or tell the meaning of: present in understandable terms.
Comprehensibility	capable of being comprehended or understood; intelligible.	capable of being comprehended: intelligible.
Understandability	capable of being understood; comprehensible.	to have understanding: have the power of comprehension.
Intelligibility	the quality or condition of being intelligible; capability of being understood.	capable of being understood or comprehended.
Transparency	easily seen through, recognized, or detected.	readily understood.
Readability	capable of being read; legible.	able to be read easily.

Table 2.4 shows the relationship of the meaning between various terms related to interpretability obtained from *Dictionary.com* and *Meriam-Webster*. From these, it is apparent that interpretabilty is closely related to the ability to understanding something, e.g. a system. This view is confirmed by a study by [108] that also defined interpretability as referring to the capability to express the behaviour of the real system understandably.

The concept of interpretability is considered in various fields such as education, medicine, computer science, engineering, physics, social sciences, etc which shows a universal interest for interpretability. In Mathematical Logic, interpretability is defined as a relation between formal theories that expresses the possibility of interpreting or translating one into the other [109]. For example, in [110], they applied the concept of interpretability to examine a statistical result returned by root-mean-square error (RMSE), the R-square, average absolute error (AAE) and relative average absolute error (RAAE). Two classifiers were introduced for interpretability of statistical results in terms of prediction accuracy. The results were either: (i) interpretable; or (ii) not interpretable without context.

In Psychology, it has been shown that there are many factors influencing people's interpretation due to their subjective views of the world around them. Cultural values, needs, beliefs, experiences, expectations, involvement, self-concept, and other personal influences all have tremendous bearing on how we interpret stimuli in our environment [111]. A study in [112] attempted to assess the effect of limited experience with the road signs used in

United States on their interpretability. They measured the interpretability score by evaluating the percentage of correct matching. That is, an interpretation was scored as correct if the three experts independently assessed five responses and conveyed the same meaning as the European sign meanings. The results showed that signs which are difficult to interpret generally use abstract, unfamiliar symbols or include ambiguous cues.

In Machine Learning, interpretability is also taken as meaning there is some degree of capability to be understood. This implies that a model constructed to map the feature space (predictors) into the output space (dependent variables) has more chance to be understandable if it contains some human-accessible explicit representation of the mapping, as in decision trees [113]. Also, James et al. [114] claimed that decision trees are easier to explain to people than linear regression since they mirror more closely the human decision-making process than other predictive models. Nevertheless, they may suffer from fundamental issues, such as overfitting and variance.

Interpretability is acknowledged as one of the most appreciated advantages of fuzzy systems [115] in many applications, particularly in those with high human interaction where it actually becomes a strong requirement [107]. Fuzzy systems use fuzzy sets to describe domains of values of certain variables. Similarly to human thinking, linguistic terms can be used for this purpose. This property makes fuzzy systems rather unique among modeling systems because while maintaining some intuitive conditions about the collection of fuzzy sets, they possess the ability to be easily interpreted, i.e. easily understandable even for laymen users [116]. However, the interpretability of FLSs is a subjective property that depends on several factors, mainly the model structure, the number of input variables, the number of fuzzy rules, the number of linguistic terms, the shape of the fuzzy sets and so on [108]. Also, it depends on the person who makes the assessment [15]. To date, the study of interpretability in FLS is still an open discussion.

This thesis is concerned with how interpretability is affected in FLSs when it is hierarchical, featuring various subsystems, layers and topologies. As far as the literature is concerned, no attempt has been made in this area. The next section provides the discussion on the interpretability assessment of FLSs.

2.4.2 Taxonomy to Analyse Interpretability of Fuzzy Systems

In recent years, the enthusiasm of researchers in obtaining more interpretable fuzzy models has increased. Assessing the interpretability of fuzzy systems still remains an open and challenging problem. Defining a good index is extremely difficult mainly due to the inherently

subjective nature of interpretability [117]. It strongly depends on the background of the person who makes the assessment according to their own knowledge, but also takes into account their previous experience and preferences [117]. Furthermore, interpretability is difficult to quantify because it is very subjective [118],[119]. Due to the subjectivity of the concept the choice of appropriate interpretability measures is still an open problem [23].

Some researchers [14, 23] have suggested taxonomy as a way of exploring the interpretability of FLSs. That is, it could help give a better understanding of which aspects that should be considered and be taken into account when measuring the interpretability of FLSs.

Zhou and Gan [14] proposed a taxonomy of fuzzy model interpretability and categorise it into *low-level* interpretability and *high-level* interpretability as can be seen in Fig. 2.18. The low-level interpretability of fuzzy models is achieved on the fuzzy set level by optimising specific criteria, such as optimising MFs, and the high-level interpretability is obtained on the fuzzy rule level by conducting overall complexity reduction in relation to other criteria, such as a moderate number of variables and rules and consistency of rules.

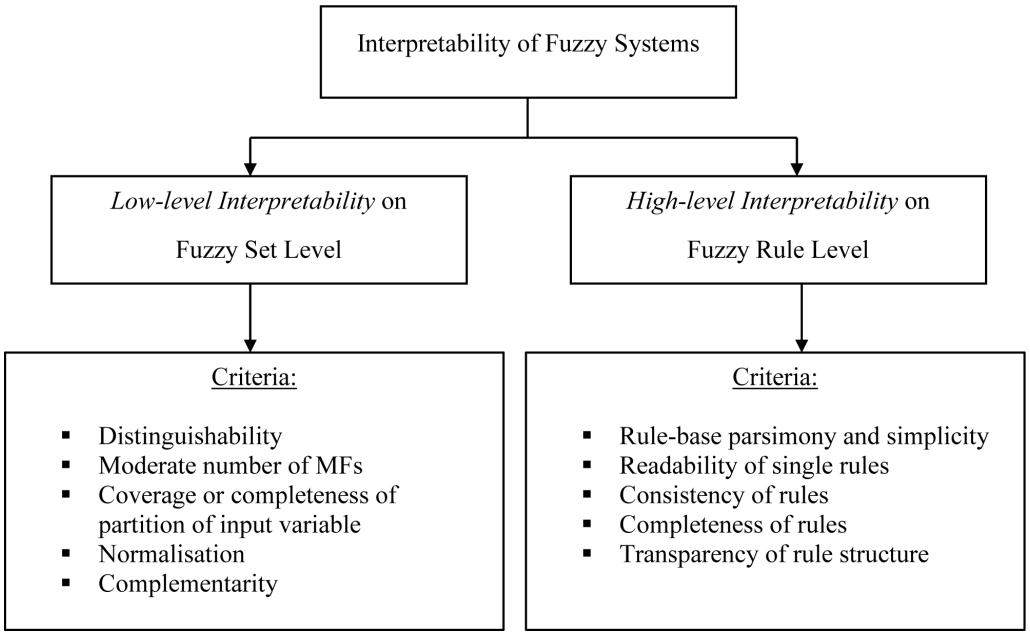


Fig. 2.18 A taxonomy of interpretability of fuzzy systems (adapted from [14]).

The criteria of low-level interpretability and high-level interpretability are then explained as follow:

1. Criteria for low-level interpretability

   - *Distinguishability*

     Distinguishability is an important criterion, because, in the input partition space

to achieve an interpretable FLS, the fuzzy sets should clearly define the unique ranges in the universe of discourse of a variable, and each MF must be different from each other. Representing a linguistic term with a clear semantic meaning is essential. Fig. 2.19 shows that certain distinguishability of fuzzy sets may be lost, and will result in difficulties in assigning different linguistic labels and semantic meaning to these fuzzy sets.

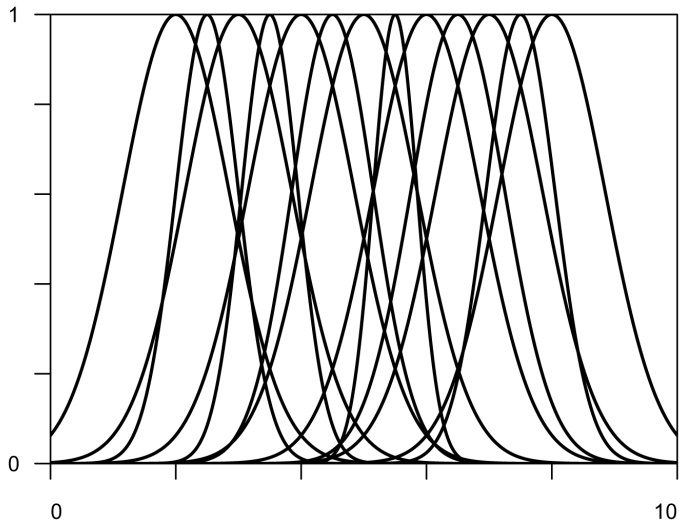


Fig. 2.19 Fuzzy sets without distinguishability. Adapted from [14].

- *Moderate number of MFs*
  Based on a study in cognitive psychology [120], the number of different entities efficiently stored in the short-term memory should not exceed the limit of 7 ± 2. Therefore, the number of variable MFs should not be arbitrary but must be compatible with the number of human conceptual entities that can be efficiently handled and occur during inferential activities.
- *Coverage or completeness of partition of input variable*
  The entire universe of discourse of a variable must be covered by MFs generated, and each data point should belong to at least one of the fuzzy sets and have a linguistic representation as shown in Fig. 2.20.

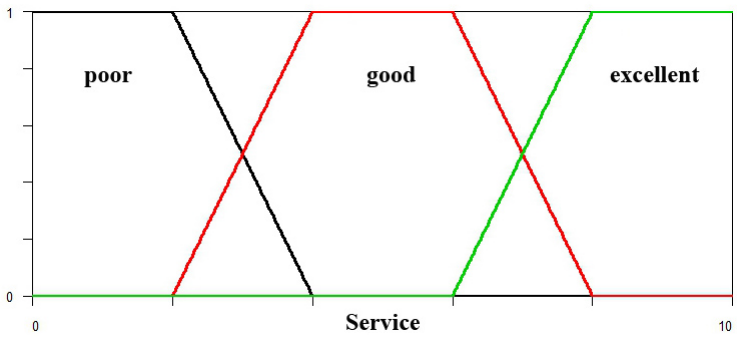


Fig. 2.20 Example of membership function of *service* that shows complete coverage.

- *Normalisation*
  Each MF of a variable is expected to represent a linguistic label with clear semantic meaning. Thus, at least one data point in the universe of discourse should have a membership value equal to one, that is, MFs of a variable should be normal. An example of this normalization can be seen in Fig. 2.21.

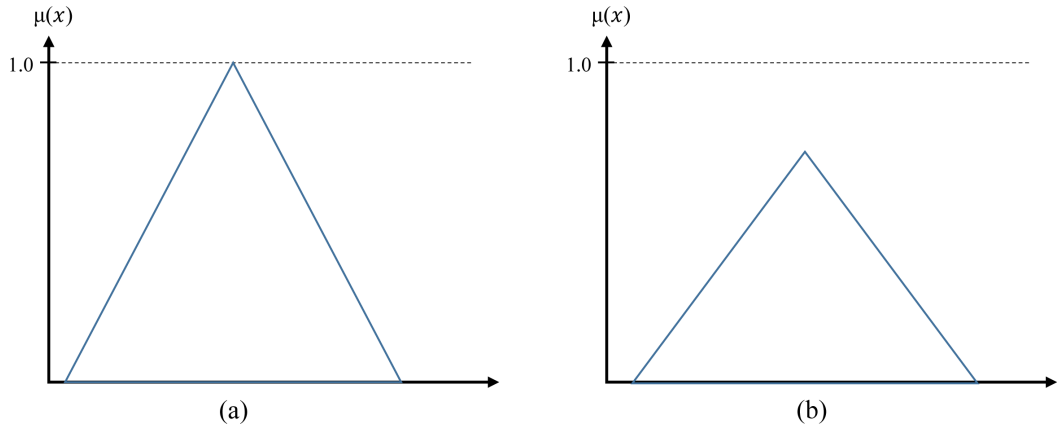


Fig. 2.21 Example of normal fuzzy set (a) and subnormal fuzzy set (b).

- *Complementarity*
  The sum of all its membership values should be equal to one for each element of the universe of discourse. This ensures a uniform distribution of meaning among the elements.

2. Criteria for high-level interpretability

- *Rule base parsimony and simplicity*
  According to the principle of Occam's razor,

"The best model is the simplest one fitting the system behaviours well".

Hence, in order to preserve the model performance at a satisfied level, the set of fuzzy rules must be as small as possible. This is because a large rule base would lead to a lack of global understanding of the system.

- *Readability of single rule*
  As discussed earlier, the number of conceptual entities a human being can efficiently handle should not exceed the limit of 7 ± 2. Hence, to improve readability, the number of conditions in the premise part of the rule should not exceed the limit of 7 ± 2 distinct conditions.
- *Consistency of rules*
  This criterion indicates uniformity in the rule base in the sense that rules with similar premise parts should have similar consequent parts.
- *Completeness of rules*
  To prevent the fuzzy system from breaking inference, at least one rule should be fired for any possible input vector.
- *Transparency of rule structure*
  A fuzzy rule should characterise human knowledge or system behaviours in a clear way.

Another taxonomy by Gacto et al. [23], introduced the combination of; (i) complexity versus semantic interpretability; and (ii) rule base (RB) versus fuzzy partition; as the guideline that should be taken into account to measure the interpretability of FLSs. Complexity-based interpretability is devoted to decreasing the complexity of the obtained model (usually measured as the number of rules, variables, labels per rule, and other factors as shown in Table 2.5). Meanwhile, semantics-based interpretability is dedicated to preserving the semantics associated with the MFs. Approaches trying to ensure semantic integrity may impose constraints on the MFs or consider factors such as distinguishability, or coverage. Table 2.5 presents a matrix between complexity versus semantic interpretability and rule base versus fuzzy partition.

Table 2.5 A taxonomy to analyze the interpretability of FLSs. Adapted from [23].

	Rule base level	Fuzzy partition level
Complexity-based interpretability	G1 number of rules number of conditions	G2 number of membership functions number of features
Semantic-based interpretability	G3 consistency of rules rules fired at the same time transparency of rules structure cointension	G4 completeness or coverage normalisation distinguishability complementarity relative measure

As shown in Table 2.5, the taxonomy has four groups which focus on specific criteria to measure the interpretability of FLSs. The description of each group can be explained as follow:

G1 Complexity at the rule base level

This group is concerned with the criteria to reduce or to control the complexity of the rule base. The first criterion is the *number of rules*. To improve the model interpretability, the set of fuzzy rules must be as small as possible under conditions in which the model performance is preserved. The second criterion is the *number of conditions*. The number of conditions in the antecedent of a rule must not exceed the limit of 7 ± 2 distinct conditions. This is due to the number of conceptual entities a human being can handle [120]. Therefore, the number of conditions should be as small as possible in order to ease the readability of the rules.

G2 Complexity at the level of fuzzy partitions

The criteria for this group emphasise controlling the complexity at the level of fuzzy partitions. The first criterion is the *number of features or variables*, which need to be reduced to improve the readability of the knowledge base. This can reduce the dimensionality in high dimensional problems. The second criterion is the *number of MFs*. Similarly with the number of conditions in G1, the number of MFs should not exceed the limit of 7 ± 2 distinct MFs [120].

G3 Semantics at the rule base level

This group contains criteria for controlling the semantic interpretability at the rule base (RB) level. The first criterion is the *consistency of the RB* which indicates that the rules

with similar premise parts should have similar consequent parts. The second criterion is the *number of rules fired at the same time*. This criterion favours of minimising the number of rules firing that are activated for a given input.

G4 Semantics at the fuzzy partition level

According to Gacto et al. [23], a complex fuzzy partition that has high overlap among MFs reduces the global semantic interpretability as shown in Fig. 2.22. Thus, the criteria for this group (G4) involve maintaining semantic interpretability at the fuzzy partition level. The first criterion is *completeness or coverage*, indicates that the universe of discourse of a variable should be covered by the MFs, and every data point should belong to at least one of the fuzzy sets and have a linguistic representation. The second criterion is *normalisation*, which indicates that each MF of a variable is supposed to describe a linguistic label with clear semantic meaning; consequently, at least one data point in the universe of discourse should have a membership value equivalent to one, that is, MFs of a variable should be normal, although this principle has been contested by Garibaldi et al. [121]. The third criterion, *distinguishability*, indicates that the MFs should represent a linguistic term with clear semantic meaning and should be easily distinguished from the rest of the MFs of the same variable. The final criterion is the *complementarity*, which indicates for each element of the universe of discourse, the sum of all membership values should approach one. This is to ensure a uniform distribution of the meaning among the elements.

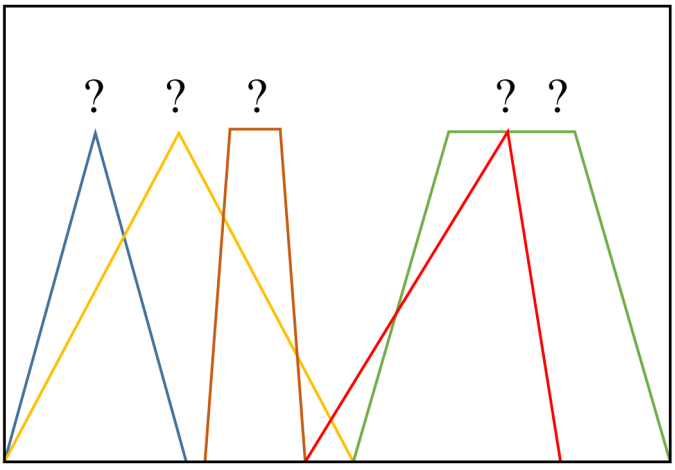


Fig. 2.22 Fuzzy partition with a poor semantic interpretability. Adapted from [23].

2.4.3 Measuring Interpretability using Interpretability Indices

The introduction of these taxonomies provides pathways for researchers to understand the criteria that should be considered for measuring the interpretability of FLSs. Consequently, substantial research on interpretability measures has proposed a range of alternative interpretability indices for FLSs.

2.4.3.1 Summary of the Current Interpretability Indices

Table 2.6 shows a summary of the current works that propose indices for measuring the interpretability of FLSs, grouped by authors, year, Gacto's taxonomy group and description indices.

Table 2.6 Summary of the current interpretability indices

Authors	Year	Indices	Gacto's taxonomy	Description Indices
Nauck [21]	2003	Nauck Index	G1, G2	An index is proposed to measure the interpretability of fuzzy linguistic models, particularly for classification problems.
Guillaume & Charnomordic [28]	2003	Guillaume Index	G1, G2	A method is proposed for generating interpretable fuzzy rules from data.
Alonso et al. [24]	2006	Fuzzy Index	G1, G3	A Fuzzy index is proposed to measure the interpretability of FLSs, particularly for fuzzy rule-based classification systems. Six variables are taken as the input of an HFS and combined into a single index.
Ishibuchi & Nojima [27]	2017	Simple indices	G1	An index is proposed to measure the interpretability of multi-objective fuzzy genetics-based machine learning, concerning the rule base readability.
Alonso et al. [1]	2008	HILK	G1, G3	A methodology for designing Highly Interpretable Linguistic KBs (HILK) is proposed for classification problems, considering both expert knowledge and knowledge extracted from data.
Alonso et al. [15]	2009	An experimental approach and Web Poll	G1, G3	A methodology is proposed to analyse the different indices measures by using a web poll aimed at determining how different people assess interpretability by giving priority to different criteria.
Botta et al. [118]	2009	Novel index	G2	An index is proposed using the mean square error which focus on the readability of fuzzy partitions.
Marquez et al. [122]	2010	Adaptive defuzzification	G1	An index is proposed for regression problems with a mechanism to improve the interpretability in the sense of complexity for linguistic fuzzy rule-based systems with adaptive defuzzification.
Alonso & Magdalena [117]	2010	Flexible Index : Combine user's preference	G1, G3	Several indices are combined to get a measurement that incorporates the preferences of the user flexibly and efficiently.
Mencar et al. [25]	2011	A cointension based approach	G4	An approach for automatically evaluating interpretability of rule-based fuzzy classifiers is proposed that exploits the propositional view of rules as a means to define cointension.
Pancho et al. [123]	2016	Fingrams	G3	Interaction between rules is demonstrated at the inference level graphically regarding co-fired rules, i.e., rules fired at the same time by a given input.
Pota et al. [124]	2016	Pota Index	G1, G3	An index is proposed for quantitatively assessing either readability or semantic interpretability particularly for fuzzy classification problem.

Among the indices suggested in Table 2.6, none of the indices propose to measure the interpretability of FLSs when it is a hierarchical system, featuring multiple subsystems, layers and different topologies. To date, no one has investigated how interpretability can similarly be measured using indices when the FLS concerned is hierarchical.

Additionally, the fourth column in Table 2.6 represents Gacto's taxonomy indicating which group (G1, G2, G3, G4) each proposed index is included in. In this thesis, two interpretability indices in Table 2.6, namely the Nauck and Fuzzy indices, are used to investigate how interpretability can similarly be measured in Heirarchical FLSs, as further discussed in Chapter 3. The Nauck and Fuzzy indices are chosen due to the fact that both are

common indices which cover several interpretability criteria in the aforementioned Gacto's taxonomy, namely (G1 and G2) and (G1 and G3), respectively, as shown in Table 2.6. Also, both indices produce a numerical value in $[0,1]$ for measuring the interpretability of FLSs. These indices are discussed in the next subsections.

2.4.3.2 Nauck Index

The Nauck index is a numerical index introduced by Nauck [21] in order to measure fuzzy rule-based classification systems. Hence, it will be used in this thesis as a useful index for assessing the interpretability of FLSs. It is computed as the product of three terms:

$$Nauck\ index = comp \times \overline{cov} \times \overline{part}. \tag{2.5}$$

- *comp* represents the complexity of FLSs measured as the number membership functions (MFs) of output variables divided by the number of input variables in the FLS's rules. It is computed as

$$comp = m / \sum_{i=1}^{r} n_i, \tag{2.6}$$

  where m is the number MFs of output variables, r is the number of rules and n_i is the number of input variables used in the ith rule.

- *cov* is the coverage degree of the fuzzy partition. If X_i is the domain of ith input variable partitioned by p_i MFs $\{\mu_i^{(1)},\ldots,\mu_i^{(p_i)}\}$, then *cov* is computed as

$$\begin{aligned} cov_i &= \frac{\int_{X_i} \hat{h}_i(x)\,\mathrm{d}x}{N_i} \\ \hat{h}_i(x) &= \begin{cases} h_i(x) & \text{if} \quad 0 < h_i(x) < 1 \\ \frac{p_i - h_i(x)}{p_i - 1}, & \text{otherwise} \end{cases} \\ h_i(x) &= \sum_{k=1}^{p_i} \mu_i^{(k)}(x), \end{aligned} \tag{2.7}$$

  where $h_i(x)$ is the total MFs of ith input variable with $N_i = \int_{X_i}$ for continuous domains. The integral in (2.7) is replaced by a sum for discrete finite domains with $N_i = |X|$. Then, $\overline{cov} = \sum_{i=1}^{r} cov_i / n_i$, denotes the average normalized coverage for all input variables.

- *part* stands for the partition index which is computed as the inverse of the number of MFs minus one for each input variable;

$$part_i = \frac{1}{p_i - 1}, \tag{2.8}$$

where p_i is the number of MFs in the ith input variable. Then, $\overline{part} = \sum_{i=1}^{r} part_i/n_i$, denotes the average normalized partition index for all input variables.

An FLS model is said to be less interpretable when its Nauck index is closer to 0 and more interpretable when Nauck index is closer to 1. To illustrate how the interpretability index is generated by Nauck index, the Waiter-Tipping example [125] which is frequently used as a benchmark is considered. The rule base of the example is as follows:

IF *service* is *poor* THEN *tip* is *cheap*,

IF *service* is *good* THEN *tip* is *average*,

IF *service* is *excellent* THEN *tip* is *generous*.

The *Nauck index* can be generated as:

$$\begin{aligned} comp &= \frac{3}{1+1+1} = 1 \\ \overline{cov} &= 1 (\text{i.e., as shown in Fig. 2.20}) \\ \overline{part} &= \frac{1}{3-1} (\text{only one input variable is used}) \\ Nauck\ index &= comp \times \overline{cov} \times \overline{part} = 0.5. \end{aligned}$$

Hence, the Nauck index of 0.5, indicates the interpretability value for the Waiter-Tipping example.

2.4.3.3 Fuzzy Index

As discussed in [24], the Fuzzy index, which is inspired by Nauck's index, is a proposed interpretability assessment. Six variables are taken as the input of a hierarchical fuzzy system to combine these into a single index. The six variables are: (i) total number of rules (NR); (ii) total number of premises in all the rules (NP) — in a complete rule-set, this equals the number of rules multiplied by the number of input variables; (iii) number of rules which use one input variable ($NR_{i=1}$); (iv) number of rules which use two input variables ($NR_{i=2}$); (v) number of rules which use three or more input variables ($NR_{i\geq 3}$); and (vi) average number of linguistic terms defined for each input variable ($\overline{terms}$). The index also depends on the

number of classes (NC), also referred to as number of output terms. It should be noted that although the Fuzzy Index is generated using an HFS, it only focuses on measuring the interpretability of standard FLSs, and has not previously been applied to HFSs.

The Fuzzy index itself is computed as the result of an aforementioned hierarchical FLS as shown in Fig. 2.23 which consists of four linked knowledge bases, namely: RB1, RB2, RB3 and RB4. RB1 gives an estimation of the rule base dimension based on the total number of rules and input variables. At the same time, the rule base complexity is evaluated at RB2 according to the number of input variables used by the rules. Then, RB3 combines rule base dimension and complexity, and as a result it yields a rule base interpretability index. Lastly, RB4 integrates the rule base interpretability index and the average number of linguistic terms defined by then input variables.

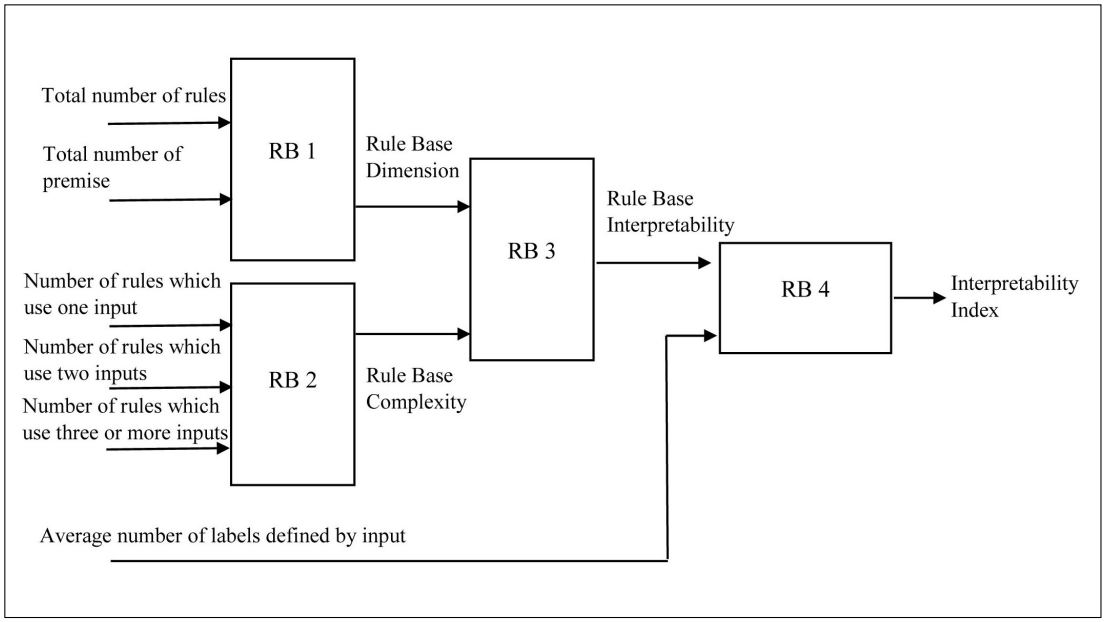


Fig. 2.23 Hierarchical fuzzy system for assessing interpretability via Fuzzy index. Adapted from [24].

Similarly to the Nauck index, a Fuzzy index closer to 0 implies that the FLS model is less interpretable. Meanwhile, a Fuzzy index closer to 1 implies higher interpretability. To show how this index is generated, we use the same Waiter-Tipping example rule base:

IF *service* is *poor* THEN *tip* is *cheap*,

IF *service* is *good* THEN *tip* is *average*,

IF *service* is *excellent* THEN *tip* is *generous*.

The Fuzzy index calculation for Waiter-Tipping example is summarized in Table 2.7.

Table 2.7 Fuzzy index of Waiter-Tipping Example

NR	:	3
NP	:	3
$NR_{i=1}$	:	3
$NR_{i=2}$	:	0
$NR_{i\geq 3}$	:	0
$\overline{terms}$	:	3
Fuzzy index	:	0.647

Thus, the Fuzzy index of 0.647, indicates the interpretability value for the Waiter-Tipping example.

2.4.4 Measuring Interpretability of Fuzzy Systems Using a User Study

Since interpretability is of a subjective nature, it is essential to assess how people perceive interpretability in fuzzy systems. This is because assessment of interpretability is strongly dependent on the background of the person who assesses according to their knowledge, but also taking into account their previous experience and preferences [117]. A user study is an excellent way to assess this.

A user study is one of the methods that can be used to convey design research. A user study allows researchers to identify specific variables that are interesting and observe the impact on the result of varying those factors [126]. Examples of user studies include that of Balazs and Koczy [116] in which they conducted interviews to ask users to define fuzzy sets, i.e., to get to know what a user meant by '*hot*'. Based on the user-defined linguistic terms, fuzzy rules and rule bases can be constructed easily. This was claimed to lead to a complexity reduction and improvement of interpretability.

Mencar and Fanelli [4] conducted a survey which attempted to provide a complete presentation of interpretability constraints in the literature. The survey aimed to: (i) give a homogeneous description of all interpretability constraints; (ii) to provide for a critical review of such constraints; and (iii) to identify potentially different meanings of interpretability. Alonso et al. in [15] evaluated the most used interpretability indices with a user study (in the form of a web poll) to extract useful information regarding interpretability assessment. The results showed that a fuzzy index was more easily adapted to the context of each problem as well as the quality criteria of the users.

A user study is proposed in this thesis to assess how users perceive the interpretability and complexity of FLSs. The user study is implemented at Fuzz-IEEE 2017 conference which was held in Naples, Italy. The sample of participants was selected from a range of academics (from doctoral students to full Professors), with expertise in FLSs. This user study will be discussed in detail in Chapter 4. As explained earlier, interpretability is a problematic concept to understand as it is of a subjective nature. In the next section, some of the issues regarding interpretability that are still open for discussion are presented.

2.4.5 Interpretability Issues in Fuzzy Systems

Interpretability is an ill-posed property, and there are several issues that have been addressed by researchers over the past 20 years. In other words, those researchers came up with a number of questions discussing interpretability which have led to the development of a new flourishing research direction. In general, the issues and challenges of interpretability can be grouped into four categories.

2.4.5.1 Interpretability Definition

Although interpretability issues in fuzzy system and statistical system modelling have received much attention in recent years, there is no well-established definition about model interpretability [14],[25],[115],[119],[127]. For that reason, there are still open questions on this issue; “What does interpretability actually mean?" and “Why is interpretability worth considering?"

This thesis will not intend to propose a new definition to interpretability but will utilise some of the definitions that have already been proposed in the literature [108].

2.4.5.2 Interpretability Assessment

The choice of an appropriate interpretability measure is still an open discussion due to its subjective nature and the large number of factors involved. Also, according to Alonso et al. in [15], interpretability is of a subjective nature because it depends on the talent and background of the end-user. Substantial research on interpretability measures [14, 23, 1, 15, 21, 24, 25, 27, 28, 117, 118, 122, 128–130] proposed interpretability indices for FLSs. However, questions arise including; “What is the best index to represent interpretability?", “Do we need to evaluate all criteria when measuring interpretability?" and “ What is the universal interpretability index that can represent a different types of FLSs i.e., including HFSs? ". This thesis is concerned with taking interpretability indicies for FLSs and applying

them to HFSs, which consist of multiple layers, subsystems and varied topologies. To date, no one has investigated how interpretability in HFSs be measured using indices.

2.4.5.3 Design Interpretable Fuzzy Systems

Model accuracy and interpretability are two conflicting objectives, so preserving interpretation during adaptation in data-driven fuzzy system modelling is a challenging task, which has received much attention in the fuzzy system modelling community [14]. Questions that the literature raises include; “How can interpretable fuzzy systems be designed?” and “ How can interpretability and accuracy be balanced in fuzzy systems?". As reported in [103], several methods have been introduced to maintain a good trade-off in the design of complex fuzzy systems. They include Evolutionary Multiobjective Optimization [104],[131], Context Adaptation (CA) [118], Hierarchical Fuzzy Modeling [93],[106] and many other algorithms and approaches related to rule bases, membership functions, and fuzzy partitions. Although HFSs are claimed to be a good design for interpretable fuzzy models, it not an easy task to design HFSs that having multiple subsystems, layers and meaning of intermediate outputs. This issue is explored and further discussed in Chapter 7.

2.4.5.4 Representation of Fuzzy Systems

As discussed in [119], it is claimed that for very complex problems the use of novel forms of representation (different from the classical rule-based form) may help in representing complex relationships in comprehensible ways, thus yielding a valid aid in designing interpretable fuzzy systems. However, the literature has not yet provided a complete answer as to how interpretable fuzzy systems should be represented.

In the following subsection, in order to have more understanding of the interpretability of HFSs, the complexity will be discussed, which is a key component of interpretability.

2.5 Complexity

2.5.1 State of the Art

The complexity of a system is a rather universal concept. The concept of complexity has recently become a serious challenge for scientific research within a multi-disciplinary context. For example, it is quite common to find complex systems in biology, cosmology, engineering, computing, finance and other areas [132]. Apparently, there is no consensus

about just what “complexity” means, but there is a cluster of associated ideas which has established some common intellectual ground amongst researchers: “system”, “interactions”, “emergence”, “self-organisation”, “learning and adaptation”, “evolution and coevolution”, “positive feedbacks”, “networks” and “distributed control” [133].

Johnson in [134] described complexity as characterising the behaviour of a system or model whose components interact in multiple ways and follow local rules, meaning that there is no reasonable higher instruction to define the various possible interactions. Also, complexity has been defined by Gegov [132] to be associated with some attributes such as nonlinearity, uncertainty, dimensionality and structure, leading to a more challenging management of systems with these attributes.

In the following subsections, the concept of complexity in relation to FLSs and also other fields will be discussed.

2.5.2 Complexity in Fuzzy Systems

As discussed earlier in Table 2.5, complexity is an essential component to determine the interpretability of FLSs. Thus far, complexity has often been used as an indirect measurement of the interpretability of FLSs. Several researchers claim that the reduction of complexity in a fuzzy system can lead to better interpretability [22, 29, 30]. This relationship between complexity and interpretability can be illustrated, as in Fig. 2.24.

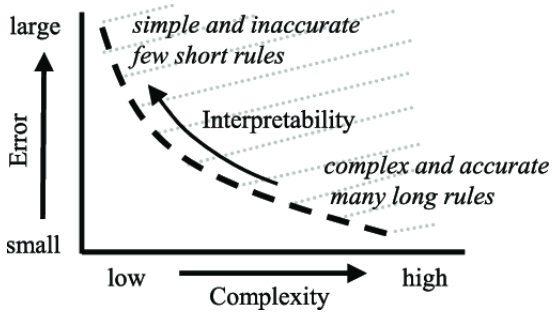


Fig. 2.24 A typical relationship between complexity and interpretability of a model. Taken from [135].

In FLSs, complexity could be related to the specific problem described by the fuzzy model. In other words, from the structural analysis of a knowledge base, we should expect to gain information concerning the complexity of the underlying problem [25]. The complexity in FLSs grows exponentially with their input variables [22], which is also known as *complexity explosion*. Model complexity can have a significant impact on the accuracy, efficiency and

interpretability of FLSs [136]. Complexity is usually evaluated by a simple measure, mainly looking at the rule-based complexity.

2.5.2.1 Rule-based Complexity

Rule-based complexity is primarily related to the readability of the knowledge base [137]; that is, the complexity of the rule base can be minimised to improve its readability [25]. Standards which are used to measure rule-based complexity include the number of rules, variables, and labels per rule, amongst others [55],[138]. Furthermore, a study by [139] has shown the relationship between the number of inputs, the number of rules and the complexity in FLSs to be as follows:

$$\uparrow \textit{Number of inputs} \text{ (FLSs)}, \uparrow \textit{Number of rules} \text{ (FLSs)} \tag{2.9}$$

$$\uparrow \textit{Number of rules} \text{ (FLSs)}, \uparrow \textit{Complexity} \text{ (FLSs)} \tag{2.10}$$

where $\uparrow$ denotes increasing. The relationships of (2.9) and (2.10), indicate that increasing the number of inputs in FLSs will increase the number of rules in FLSs and consequently, this will also increase the complexity of FLSs.

In the related literature, there are several ways to measure the rule-based complexity in FLSs. For example, Nauck in [21] has defined the complexity of an FLS as measurable by the number of classes (NC) divided by the total number of premises (NP). This complexity measurement is a part of his proposed interpretability index – Nauck's index. It is computed as:

$$Comp = \frac{\text{NC}}{\text{NP}} \tag{2.11}$$

Another study, by Alonso et al. [15], measured the complexity by one minus the *Comp* value in (2.11), which can be expressed as follows:

$$\begin{aligned} C &= 1 - Comp \\ &= 1 - \left(\frac{\text{NC}}{\text{NP}}\right) \end{aligned} \tag{2.12}$$

According to Alonso et al. [15], an FLS model is more complex when the rule-based complexity C is close to 1 and less complex when the rule-based complexity C is close to 0.

2.5.3 Complexity in Other Fields

As mentioned earlier, the complexity of systems is a universal concept. Thus, it is relevant to consider how complexity in other fields is measured or assessed, specifically, on how structural complexity is measured. The measure of structural complexity from the other fields will be utilised to the HFS design, as further examined and discussed in Chapter 6. Therefore, some literature related to the complexity is presented in order to give some necessary background.

Simon in [140] claimed that complexity takes the form of hierarchy — a complex system is composed of subsystems that, in turn, have their subsystems, and so on. He said that the term "*hierarchy*" has usually been used to refer to a complex system in which each of the subsystems is subordinated by an authority relating to the system it belongs to. More exactly, in a formal heirarchical organisation, each system consists of a "boss" and a set of subordinate subsystems. Each of the subsystems has a "boss" who is the immediate subordinate of the "boss" and this implies that the heirarchy is exactly three layers deep in the system. He also claimed that *the relations among subsystems are more complex* than in the formal organisational hierarchy.

Commons et al. [141] introduced the notion of the order of hierarchical complexity of tasks. Hierarchical complexity represents a form of information that is orthogonal to the traditional information theory form, in which information is coded as bits that rise quantitatively with more information. Information in tasks and their subtasks can be characterized by both hierarchical (vertical) and non-hierarchical (horizontal) complexity. This is due to the fact that every task has an order of complexity associated with it.

Furthey study by Commons et al. [142] presented the model of hierarchical complexity which included a framework for scoring reasoning stages in any domain as well as in any cross cultural setting. The scoring is based not upon the content or the participant material, but instead on the mathematical complexity of the hierarchical organization of information. Specifically, Commons et al. [142] claimed that hierarchical complexity refers to the mathematical complexity of the task presented to the participant, but not directly to the complexity of the participant's performance that will successfully complete the given task.

Most complex systems have an inherently hierarchical organisation and, correspondingly, the networks behind them also exhibit hierarchical features. Mones et al. [143] developed an approach and proposed a quantity (measure) which is simple enough to be widely applicable, reveals a number of universal features of the organization of real-world networks and is capable of capturing the essential features of the structure and the degree of hierarchy in a complex network. Complexity arises from either the structure of the interactions between

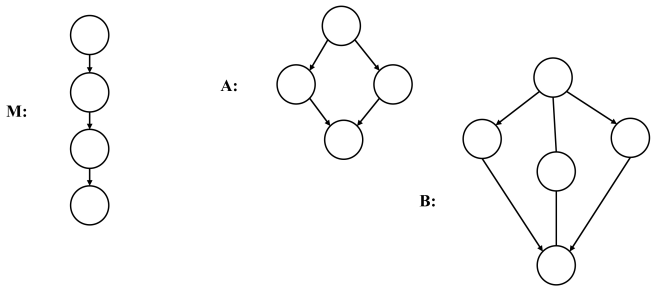


Fig. 2.25 One main program M, two subroutines A and B. Adapted from McCabe [145].

very similar units or other specific characteristics of those interactions. In both cases, the abstract representation of a complex system can be achieved by a collection of nodes (units) and edges (representing interactions between the units) forming a network (or graph).

Overall, from the published literature, one common view of complexity measure is that it is dependent on the number of structural features contained within an organisation rather than simply on the number of its basic elements [144]; this idea is also known as structural complexity, and is examined further in the next section.

2.5.3.1 Structural Complexity

Structural complexity is an attribute of any general type of system. This attribute can be assessed by different measures, and it is often linked to interaction among systems' properties such as nodes, edges and networks [132, 143].

McCabe in [145] proposed a graph-theoretic structural complexity measure that measures and controls the number of paths through a computer program. The complexity measure developed here is defined in terms of basic paths that when taken in combination will generate every possible path. The cyclomatic number $v(G)$ of graph G with n vertices, e edges, and p connected components is

$$v(G) = e - n + p \tag{2.13}$$

The overall strategy involves measuring the complexity of a programme by computing the number of linearly independent paths $v(G)$. However, calculating the complexity of a collection of programmes ($p \neq 1$), particularly a hierarchical nest of subroutines, (2.13) becomes as follows:

$$v(G) = e - n + 2p \tag{2.14}$$

For instance, assume a main program M calls two subroutines A and B and has a control structure as shown in Fig. 2.25. Let us denote the total graph above with 3 connected components as $M \cup A \cup B$. Now, since $p = 3$, the complexity can be computed using (2.14) as follow

$$\begin{aligned} v(M \cup A \cup B) &= e - n + 2p \\ &= 13 - 13 + 2 \times 3 \\ &= 6 \end{aligned}$$

A graph is more complex when the structural complexity $v(G)$ value is large and less complex when the $v(G)$ value is small.

In this thesis, the structural complexity in HFSs relating to the interaction amongst the structure of HFSs, namely multiple subsystems, layers and different topologies will be assessed. That is examined in detail in Chapter 6.

2.6 Case Study: Neonatal Monitoring

A complex medical problem was selected as a case study to explore the use of HFSs in the real world, and so the medical background to this case study will now be presented. The Neonatal Intensive Care Unit (NICU) is where the sickest newborn babies in hospitals are managed and treated. The NICU has one of the most information sensitive environments where efforts are made to continuously deliver the optimal healthcare for fragile, critically ill patients. "These babies that were born prematurely need special medical attention in the NICU. They need hospital supervision until their systems and organs can function with no external assistance" [146]. As a result of evolution in clinical practice and equipment in the NICU, "the gestational age at which 50% of neonates survive decreased from 29 weeks to 24 weeks over the last few decades" [147].

Prematurity is defined as babies born before completion of 37 weeks of pregnancy. "Premature delivery is divided into 3 subcategories based on gestational age: extremely preterm (less than 28 weeks), very preterm (28-32 weeks), and moderate to late preterm (32 to 37 weeks)" [148]. Infants who are ≥ 37 weeks gestation are classified as term. According to the World Health Organization, approximately 15 million premature babies are born each year globally and this number continues to increase. Prematurity accounted for one million deaths per year and is the leading cause of mortality in children under age of five [149]. Data from the UK Office for National Statistics highlights there are around 60,000 premature deliveries every year in the UK. In Western Europe, the UK has the second worst neonatal

mortality rate of 2.1 per 1000 live-births at the early neonatal phase of 0-6 days, 0.7 deaths per 1000 live-births at the late neonatal phase of 7-28 days, and 4.9 deaths per 1000 live-births at the age of less than five years [150].

In the neonatal setting, "the current practice for oxygenation management aims to target oxygen saturations (SpO_2) between 91% and 95%. Increased time outside of this SpO_2 range is associated with increased morbidity including retinopathy of prematurity (ROP) and chronic lung disease (CLD)" [151]. Two clinical trials, BOOST II and SUPPORT, have provided evidence that SpO_2 should be targeted in the 91-95% for preterm infants. "Both trials conducted their studies in infants born before 28 weeks of gestational age and found that targeting SpO_2 of more than 95% is associated with more ROP and CLD, while targeting SpO_2 less than 91% is associated with increased in mortality rate. In conclusion, keeping the SpO_2 range in 91-95% minimized the risks of significant morbidity and mortality" [152], [153].The need for prolonged oxygen administration is associated with poor pulmonary outcomes in childhood life [154].

Better management of SpO_2 can be beneficial in decreasing the mortality and disability rates among preterm babies. This can be improved with better respiratory practice, avoiding inappropriate oxygenation and timely intervention before respiratory deterioration. Therefore, there are opportunities to further improve the neonatal care we deliver that can improve the overall respiratory outcome of premature infants and potentially reduce the long-term complications.

Blood pressure, heart rate, respiratory rate, and temperature are also routinely collected vital signs on the NICU. They are valuable in order to determine a patient's progress during hospitalization [155]. The pulse oximeter is used as the fifth vital sign to determine oxygen saturation in paediatrics [156]. Healthcare professionals use vital signs as an important component of their clinical evaluation when trying to predict deterioration and a number of studies have identified the vital signs trend is beneficial as an early warning indicator of clinical status [157]. In the preterm population, additional events such as the pattern of desaturations and bradycardias can also help guide the impending clinical deterioration.

2.6.1 Summary of Existing Neonatal Monitoring Systems using Fuzzy Logic Systems

The literature concerning the use of fuzzy systems in neonatal monitoring is summarised in Table 2.8.

Table 2.8 Summary of existing neonatal monitoring systems using Fuzzy Logic Systems

Neonatal Monitoring system	Author	Results
A knowledge-based system using fuzzy logic for classifying plethysmogram pulses into two categories either valid or artefact.	Belal et al. [158]	Able to classify 679 (82%) valid segments and 543 (93%) distorted segments correctly. The calculations of the system's performance showed 82% sensitivity, 86% accuracy and 93% specificity.
An intelligent system that provides a tool enabling the user to make the final decision to accept or reject the advice given for neonatal intensive care management.	Belal et. al [159]	The overall percentage of the agreement between the system and the clinician was 93%.
A fuzzy expert system to predict the need for advanced neonatal resuscitation efforts in the delivery room.	Reis et al. [160]	Result shows sensitivity of 76.5% and specificity of 94.8% in the identification of the need for advanced neonatal resuscitation measures, considering a cut-off value of 5 on a scale ranging from 0 to 10. The area under the receiver operating characteristic curve was 93%.
Theoretical fuzzy linguistic model to estimate the risk of neonatal death based on birth weight and gestational age.	Nascimento and Ortega [72]	The results were compared with experts' opinions and the fuzzy model was able to capture the expert knowledge with a strong correlation ($r = 0.96$).

A fuzzy controller for the adjustment of inspired oxygen (FIO_2) concentration in ventilated newborns.	Sun et al. [161]	This system was able to maintain a target oxygen saturation (SpO_2) better than routine manual control and demonstrated the ability to produce a safe and clinically efficacious control system.
An Automated Fuzzy logic based Skull Stripping (AFSS) in Neonatal and Infantile Magnetic Resonance (MR) Images.	Yamaguchi et al. [162]	The results showed that the proposed method stripped skull well from any neonatal and infantile MR images.
A physiology-based fuzzy logic algorithm to assign a confidence level to heart rate and respiratory rate time-series data as they are collected, with the assumption that neither electrocardiogram (ECG) nor respiratory waveforms are available.	Liu et al. [163]	Performance in detecting four types of faults that result in low-confidence data points (receiver operating characteristic areas under the curve ranged from 0.67 (SD 0.04) to 0.83 (SD 0.03), where mean and SD are measured over all faults).
Investigation into whether fuzzy logic could offer an improvement in cardiotocography (CTG) analysis over the crisp expert system; and to investigate whether retrospective analysis of complete CTG traces could be automated.	Skinner et al. [89]	The fuzzy score produced by the system has good correlation with both the assessment of clinical experts and existing crisp system.
Theory of fuzzy sets in determining the severity of acute respiratory distress of a patient in an intensive care unit by implementing fuzzy multicriteria decision-making.	Velasevic et al. [164]	The preliminary evaluation of the approach regarding clinical practice, on the small sample of data, has produced encouraging results: more than 90% of the diagnoses obtained through the system confirmed a physician's diagnosis.

From the literature review in Table 2.8, it can be concluded that FLSs have been proven reliable when implemented in neonatal care with high accuracies of the prediction models. FLSs may also be concluded as an effective method since they can reason with imprecise information and also uncertainties [58]. Moreover, interpretability is acknowledged as the main advantage of the FLSs which is an essential for applications that require high interaction with humans [15] which is true of neonatal monitoring.

2.6.2 Problems

At the NICU, babies develop their own normal individual behavioural patterns, which clinicians and parents can identify. When a baby deviates from this pattern, it is often a sign that the baby is unwell, typically from infection. Identification and prompt treatment of deterioration are the keys to effective monitoring of babies at the NICU. However, the time taken to identify deviations from the usual pattern is often experience dependent. That is, a senior clinician or consultant is likely to determine this before more junior doctors do.

As mentioned earlier, FLSs have been proven reliable when implemented in neonatal care with high accuracies of the prediction models. However, key challenges remain in the design of FLSs, such as the fact that the number of rules required commonly increases exponentially with the number of input variables [31], which is usually high in this field. Therefore, it is worth emphasising that, the introduction of advanced fuzzy systems, namely heirarchical fuzzy systems (HFSs), for neonatal care is essential to overcome the limitation of conventional FLSs and improve model interpretability. A design of interpretable HFSs for NICU which can be easily understood by clinicians is proposed and discussed in Chapter 7.

2.7 Summary

In this chapter, the necessary background of fundamental concepts in the literature for the presentation and definitions of the methodology used in this thesis was provided. The chapter briefly discussed a brief overview of the components in FLSs including fuzzy sets, membership functions, linguistic variables, rule bases and an overview of designing FLSs. In the next section, the concept of HFSs was introduced. This section included an overview of the definition and structure of HFSs, related work on current HFSs, the advantages of HFSs and issues in HFSs. This chapter then described the state of the art on the interpretability, particularly in FLSs, including definition, assessment and the current issues. A summary was then presented of the complexity of FLSs as a subcomponent in the interpretability.

This also included complexity measurement in other fields such as psychology, science and circuit design. In the final section, this chapter presented a case study on neonatal monitoring including the existing works on neonatal monitoring in the literature using FLSs.

In the next chapter, the first objective of this thesis, which is to investigate the assessment of interpretability of HFSs, will be addressed.

Chapter 3

An Initial Approach to Measuring Interpretability of Hierarchical Fuzzy Systems

HFSs have been shown to have the potential to improve the interpretability of FLSs. In recent years, a variety of indices have been proposed to measure the interpretability of FLSs such as the Nauck index [21] and Fuzzy index [24]. However, interpretability indices associated with HFSs have not so far been discussed. The structure of HFSs, with multiple layers, subsystems, and varied topologies, is the main challenge in constructing interpretability indices for HFSs. As discussed in Chapter 1, the first objective of this thesis is to investigate the assessment of interpretability of HFSs. In doing so, this chapter begins to address these challenges by introducing extensions to the Nauck and Fuzzy indices to create an initial index for measuring the interpretability of HFSs. Using the proposed initial index, this chapter explores the concept of interpretability in relation to the different structures in FLSs and HFSs. Initial experiments on benchmark datasets show that based on the proposed initial index, HFSs with equivalent function to FLSs produce higher interpretability values, i.e. are more interpretable than their corresponding FLSs.

3.1 Introduction

One of the strengths of FLSs is their interpretability [13], particularly in applications such as knowledge extraction and decision support [2, 3]. However, key challenges remain around FLS interpretability, including the *curse of dimensionality*: the number of required rules

commonly increases exponentially with the number of input variables [31]. This challenge is also known as rule explosion and may reduce the transparency and interpretability of FLSs [32]. One effective way to deal with this problem is through the use of a special type of FLS, namely HFSs [43–45, 47, 94, 95].

HFSs were introduced by Raju et al. [43] as an approach to overcome the *curse of dimensionality* which arises in conventional FLSs. In HFSs, the original FLSs are decomposed into a series of low-dimensional FLSs — fuzzy logic subsystems (see Chapter 2.3). Moreover, the rules in HFSs commonly have antecedents with fewer variables than the rules in FLSs with equivalent function, since the number of input variables of each subsystem is lower [54, 165]. Thus, HFSs tend to reduce rule explosion, thus minimizing complexity, and improving model interpretability. So far, the potential of HFSs to support good interpretability in FLSs has not been explored in detail, with only a small number of works considering it [48, 50, 32, 106, 166].

In this chapter, an extension of the two most common FLS interpretability indices, namely the Nauck index and Fuzzy index is proposed, to create an initial index for measuring the interpretability of HFSs, with a specific focus on the complex structure of HFSs such as having multiple layers, subsystems and varied topologies.

The rest of this chapter is organised as follows; Section 3.2 presents a motivation of this chapter. Section 3.3 introduces an initial index to measuring interpretability of HFSs. The step-by-step calculation of this initial index to measuring interpretability of HFSs is presented in Section 3.4. Section 3.5 performs an experiment to this initial index using a real-world application, namely Seesaw Control Application. Then, the result and discussion is presented in Section 3.6. Finally, a summary of this chapter is discussed in Section 3.7.

3.2 Motivation

HFSs have been shown to have the potential to improve interpretability of FLSs. However, interpretability indices associated with HFSs have so far not been discussed. The architecture of HFSs, with multiple layers, subsystems, and varied topologies, is the main challenge in constructing interpretability indices for HFSs. This raises questions, such as: “What is the best aggregation technique across and within layers?", and “How should new interpretability indices deal with different topologies?". Taking these questions into consideration, an initial index for measuring the interpretability of HFSs has been proposed.

3.3 The Proposed Initial Index for HFSs

This chapter begins to extend the Nauck and Fuzzy indices by proposing an initial index for such indices from FLSs to HFSs and referred as H_I. However, a question may arise; "Why do we need a new aggregation strategy for this initial index, H_I?" This is due to the fact that subsystems in HFSs are commonly located across and within layers. Therefore, by considering these challenges that arise from HFSs' architecture, an aggregation strategy capturing the HFS topology with *layer-weight* has been proposed.

In this context, the H_I is computed as follows:

$$H_I = \sum_{j=1}^{q} \left(l_j \sum_{k=1}^{s_j} E_{jk}/s_j \right) \tag{3.1}$$

where,

- E_{jk} is the Nauck (N) or Fuzzy (F) index of a subsystem k at layer j,
- l_j is the associated weight to the layer j of the HFSs,
- s_j is the number of subsystems located at the layer j,
- q is the number of layers.

Note that (3.1) returns the original FLS index when applied to a standard FLS because it has only one subsystem. Also, it should be noted that only one output variable and no cyclic connections between internal FLSs is focused on HFSs.

Layer-weights, l_j are associated with each subsystem according to their layer. A *layer-weighting* is introduced which has the following properties:

(i)
$$\sum_{j=1}^{q} l_j = 1, \quad l_j \in [0,1]$$

The summation value of all *layer-weights* l_j should be equal to 1 regardless of the number of layers q.

(ii)
$$l_1 > l_2 > ... > l_q, \quad j = 1,...,q.$$

The *layer-weights*, l_j are arranged in descending order. This is because, as discussed in [43, 44], most HFSs' structure is formed by having the most influential input variables

in the first layer of the hierarchy, the next most important inputs in the second layer, and so on.

In order to achieve the above properties, l_j is proposed as:

$$l_j = \frac{2\Big(q-j+1\Big)}{q\Big(q+1\Big)}, \quad j = 1,...,q. \tag{3.2}$$

An HFS model is less interpretable when the H_I is close to 0 and more interpretable when H_I is close to 1.

3.4 Detail of Process for Measuring Interpretability of HFSs

There are three key steps in order to calculate the interpretability of HFSs using the H_I.

Step 1: Calculate Interpretability for each Subsystem

First, the interpretability of each subsystem is calculated using both the Nauck and the Fuzzy indices. For example, assume the values of the Nauck index calculated for the three subsystems in a Parallel HFS are $\mathrm{N}_1 = 0.3$, $\mathrm{N}_2 = 0.2$ and $\mathrm{N}_3 = 0.5$.

Step 2: Identify Layer-weights

Next, the values of the layer-weights are computed using layer-weight as in (3.2). For example, our hypothetical Parallel HFS could have two layers, and the layer weights computer at each layer could be $l_1 = 0.667$ and $l_2 = 0.333$.

Step 3: Calculate the Overall Interpretability

Finally, the overall interpretability can be calculated using the H_I index as given in (3.1). For instance, the interpretability of our hypothetical Parallel HFS is computed as follows:

$$\begin{aligned}
\mathrm{H}_I &= \sum_{j=1}^{q}(l_j \sum_{k=1}^{s_j} E_{jk}/s_j) \\
&= l_1(\mathrm{N}_1+\mathrm{N}_2)/2 + l_2(\mathrm{N}_3/1) \\
&= 0.667(0.3+0.2)/2 + 0.333(0.5) \\
&= 0.333
\end{aligned}$$

3.5 Seesaw Control Application Experiment

As a demonstration, a seesaw control application to solve the problem of balancing a seesaw has used as an example. Note that the source of this example was initially taken from [46]. They used the example to stimulate and compare the result of balancing a seesaw system using both conventional single layer FLS and an HFS. Moreover, this example of seesaw control application was also adopted in other studies such as [167–170], particularly in the context of control systems.

In this section, to balance the seesaw, the involved parameters are the distance of the cart from the origin ($x = x_1$), the angle that the wedge makes with the vertical line ($\theta = x_2$), the height of the wedge ($r_1 = x_3$) and center of mass of the wedge ($r_2 = x_4$) as shown in Fig. 3.1.

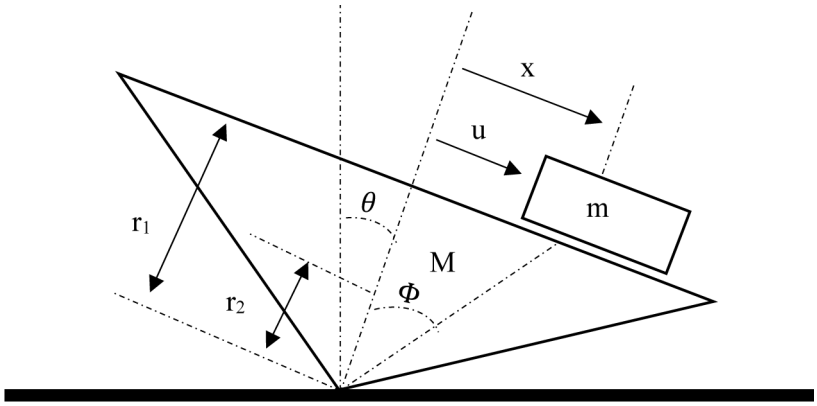


Fig. 3.1 The Seesaw System. Adapted from [46].

As reported in [46], the the conventional FLS consist of 4 input variables ($n = 4$) namely x_1, x_2, x_3 and x_4 are modelled with 3 fuzzy sets ($m = 3$): *pb*, *ze* and *nb*. Figs. 3.2 and 3.3 show the MFs of all input, intermediate output and output variable for seesaw control. A summary of FLS and both HFSs is presented in Table 3.1. The total number of rules for this standard FLS of a seesaw control is obtained by $m^n = 3^4 = 81$ rules and can be illustrated as follows:

- IF x_1 is *nb* AND x_2 is *nb* AND x_3 is *nb* AND x_4 is *nb* THEN y is *nb*,
- IF x_1 is *nb* AND x_2 is *nb* AND x_3 is *nb* AND x_4 is *ze* THEN y is *nb*,
- IF x_1 is *nb* AND x_2 is *nb* AND x_3 is *nb* AND x_4 is *pb* THEN y is *nm*,
- IF x_1 is *nb* AND x_2 is *nb* AND x_3 is *ze* AND x_4 is *nb* THEN y is *nb*,
- IF x_1 is *nb* AND x_2 is *nb* AND x_3 is *ze* AND x_4 is *ze* THEN y is *nm*,

- IF x_1 is *nb* AND x_2 is *nb* AND x_3 is *ze* AND x_4 is *ze* THEN y is *ns*,
  ⋮
- IF x_1 is *pb* AND x_2 is *pb* AND x_3 is *pb* AND x_4 is *pb* THEN y is *pb*.

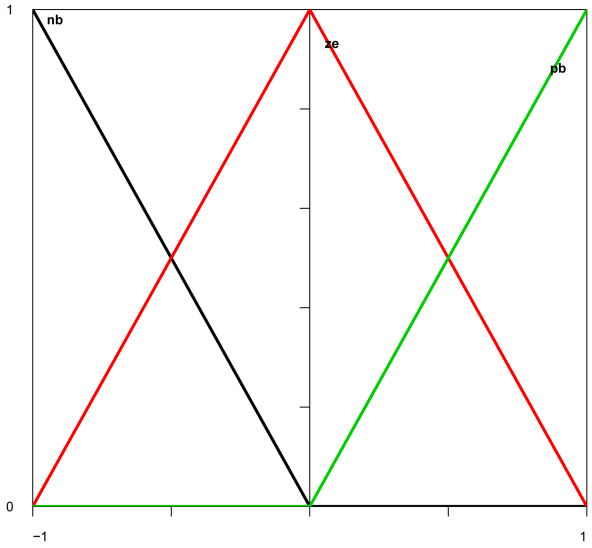


Fig. 3.2 MFs for all inputs x_1, x_2, x_3, x_4 and also intermediate output in HFSs namely y_1 and y_2

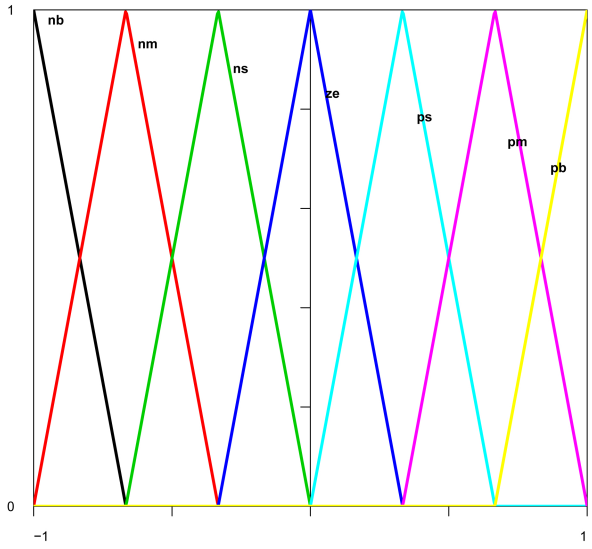


Fig. 3.3 MFs for output variable, y for Seesaw control

Table 3.1 Description of the FLS, Parallel HFS and Serial HFS for Seesaw Control Applications

Descriptions	FLS	Parallel HFS	Serial HFS
Number of inputs	4	4	4
Number of layers	1	2	3
Number of subsystems	1	3	3
Number of rules in FLS_1	81	9	9
Number of rules in FLS_2	-	9	9
Number of rules in FLS_3	-	9	9
Total number of rules	81	27	27

In this experiment, two types of HFS topologies (parallel and serial) were chosen and developed using the Fuzzy Toolbox in the R programming language [66]. The HIFAM [50] approach is implemented to building these HFSs for Seesaw control.

Table 3.2 shows the interpretability measurements of the standard FLS and all subsystems of the HFSs for the seesaw example for both the Nauck and Fuzzy indices. The aim of this experiment is to explore how the proposed initial index of interpretability behaves for several HFS toplogies. Then, the overall results of the interpretability index are compared with those of the conventional FLS.

Table 3.2 Interpretability of the Seesaw Control Application using the Nauck and Fuzzy Indices

Seesaw systems	Nauck Index				Fuzzy Index							
	comp	$\overline{cov}$	$\overline{part}$	Index	NR	NP	$NR_{i=1}$	$NR_{i=2}$	$NR_{i\geq 3}$	$\overline{terms}$	NC	Index
FLS	0.222	1	0.5	**0.011**	81	324	0	0	81	3	7	**0.241**
Parallel HFS:												
FLS_1	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_2	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_3	0.389	1	0.5	0.194	9	18	0	9	0	3	7	0.605
H_I				**0.120**								**0.530**
Serial HFS:												
FLS_1	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_2	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_3	0.389	1	0.5	0.194	9	18	0	9	0	3	7	0.605
H_I				**0.102**								**0.512**

3.5.1 Parallel Implementation of the Seesaw control

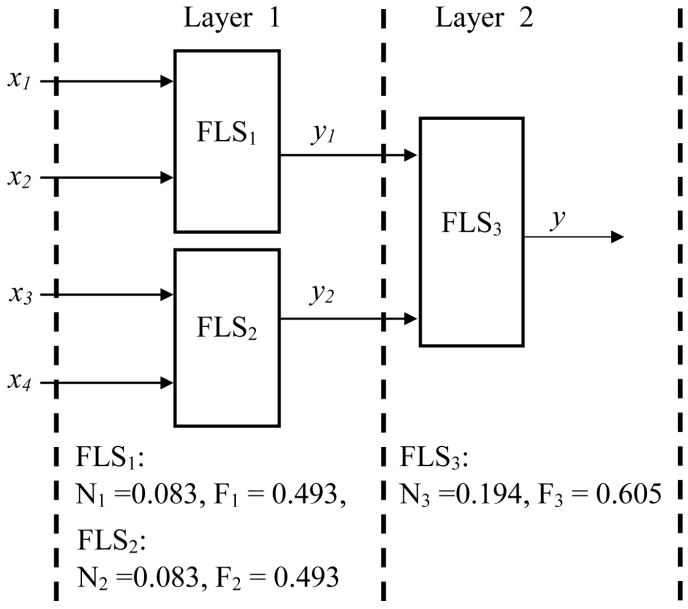


Fig. 3.4 Interpretability of a Parallel HFS using Nauck index (N) and Fuzzy index (F).

Figure 3.4 shows the interpretability indices measured for each subsystem in the parallel HFS implementation of the seesaw control by using the Nauck and Fuzzy indices respectively. The rules for subsystems FLS_1, FLS_2, and FLS_3 can be illustrated as follows:

(i). Rules for FLS_1 :

		x_2		
		nb	ze	pb
x_1	nb	*nb*	*nb*	*ze*
	ze	*nb*	*ze*	*pb*
	pb	*ze*	*pb*	*pb*

(1). IF x_1 is *nb* AND x_2 is *nb* THEN y_1 is *nb*,
(2). IF x_1 is *nb* AND x_2 is *ze* THEN y_1 is *nb*,
(3). IF x_1 is *nb* AND x_2 is *pb* THEN y_1 is *ze*,
(4). IF x_1 is *ze* AND x_2 is *nb* THEN y_1 is *nb*,
(5). IF x_1 is *ze* AND x_2 is *ze* THEN y_1 is *ze*,
(6). IF x_1 is *ze* AND x_2 is *pb* THEN y_1 is *pb*,
(7). IF x_1 is *pb* AND x_2 is *nb* THEN y_1 is *ze*
(8). IF x_1 is *pb* AND x_2 is *ze* THEN y_1 is *pb*,
(9). IF x_1 is *pb* AND x_2 is *pb* THEN y_1 is *pb*.

(ii). Rules for FLS_2 :

		x_4		
		nb	ze	pb
x_3	nb	*nb*	*nb*	*ze*
	ze	*nb*	*ze*	*pb*
	pb	*ze*	*pb*	*pb*

(1). IF x_3 is *nb* AND x_4 is *nb* THEN y_2 is *nb*,
(2). IF x_3 is *nb* AND x_4 is *ze* THEN y_2 is *nb*,
(3). IF x_3 is *nb* AND x_4 is *pb* THEN y_2 is *ze*,
(4). IF x_3 is *ze* AND x_4 is *nb* THEN y_2 is *nb*,

(5). IF x_3 is *ze* AND x_4 is *ze* THEN y_2 is *ze*,
(6). IF x_3 is *ze* AND x_4 is *pb* THEN y_2 is *pb*,
(7). IF x_3 is *pb* AND x_4 is *nb* THEN y_2 is *ze*
(8). IF x_3 is *pb* AND x_4 is *ze* THEN y_2 is *pb*,
(9). IF x_3 is *pb* AND x_4 is *pb* THEN y_2 is *pb*.

(iii). Rules for FLS_3 :

		y_2		
		nb	ze	pb
y_1	nb	*nb*	*nm*	*ns*
	ze	*ns*	*ze*	*ps*
	pb	*ps*	*pm*	*pb*

(1). IF y_1 is *nb* AND y_2 is *nb* THEN y is *nb*,
(2). IF y_1 is *nb* AND y_2 is *ze* THEN y is *nm*,
(3). IF y_1 is *nb* AND y_2 is *pb* THEN y is *ns*,
(4). IF y_1 is *ze* AND y_2 is *nb* THEN y is *ns*,
(5). IF y_1 is *ze* AND y_2 is *ze* THEN y is *ze*,
(6). IF y_1 is *ze* AND y_2 is *pb* THEN y is *ps*,
(7). IF y_1 is *pb* AND y_2 is *nb* THEN y is *ps*
(8). IF y_1 is *pb* AND y_2 is *ze* THEN y is *pm*,
(9). IF y_1 is *pb* AND y_2 is *pb* THEN y is *pb*.

The values of Nauck and Fuzzy indices for the three subsystems are $N_1 = 0.083$, $N_2 = 0.083$, $N_3 = 0.194$, $F_1 = 0.493$, $F_2 = 0.493$ and $F_3 = 0.605$ (the details are shown in Table 3.2). There are two layers in this HFS ($q = 2$). The values of the layer-weights using equation (3.2) are $l_1 = 0.667$ and $l_2 = 0.333$. According to equation (3.1), the suggested interpretability index H_I of this HFS are:

$$\begin{aligned} H_I &= \sum_{j=1}^{q} (l_j \sum_{k=1}^{s_j} E_{jk}/s_j) \\ &= l_1(N_1 + N_2)/2 + l_2(N_3/1) \\ &= 0.667(0.083 + 0.083)/2 + 0.333(0.194) \\ &= \mathbf{0.120} \end{aligned}$$

$$\begin{aligned} \mathrm{H}_I &= \sum_{j=1}^{q} (l_j \sum_{k=1}^{s_j} E_{jk}/s_j) \\ &= l_1(F_1 + F_2)/2 + l_2(F_3/1) \\ &= 0.667(0.493 + 0.493)/2 + 0.333(0.605) \\ &= \mathbf{0.530} \end{aligned}$$

3.5.2 Serial Implementation of the Seesaw control

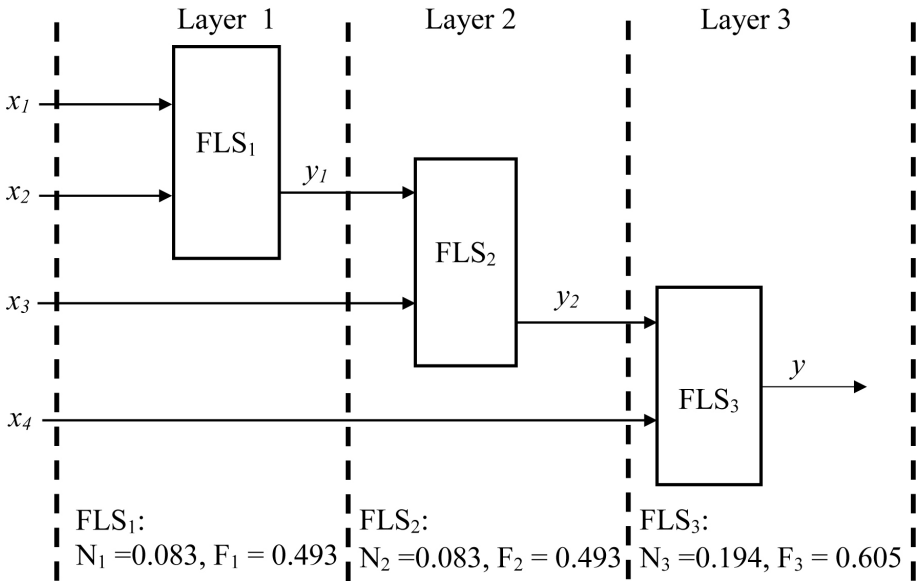


Fig. 3.5 Interpretability of a Serial HFS using Nauck index (N) and Fuzzy index (F).

Figure 3.5 shows the interpretability indices measured for each subsystem in the serial HFS implementation of the seesaw control by using the Nauck and Fuzzy indices respectively. The rules for subsystems FLS$_1$, FLS$_2$, and FLS$_3$ can be illustrated as follows:

(i). Rules for FLS$_1$:

		x_2		
		nb	ze	pb
x_1	nb	*nb*	*nb*	*ze*
	ze	*nb*	*ze*	*pb*
	pb	*ze*	*pb*	*pb*

(1). IF x_1 is *nb* AND x_2 is *nb* THEN y_1 is *nb*,
(2). IF x_1 is *nb* AND x_2 is *ze* THEN y_1 is *nb*,
(3). IF x_1 is *nb* AND x_2 is *pb* THEN y_1 is *ze*,
(4). IF x_1 is *ze* AND x_2 is *nb* THEN y_1 is *nb*,
(5). IF x_1 is *ze* AND x_2 is *ze* THEN y_1 is *ze*,
(6). IF x_1 is *ze* AND x_2 is *pb* THEN y_1 is *pb*,
(7). IF x_1 is *pb* AND x_2 is *nb* THEN y_1 is *ze*
(8). IF x_1 is *pb* AND x_2 is *ze* THEN y_1 is *pb*,
(9). IF x_1 is *pb* AND x_2 is *pb* THEN y_1 is *pb*.

(ii). Rules for FLS_2 :

		x_3		
		nb	ze	pb
y_1	nb	*nb*	*nb*	*ze*
	ze	*nb*	*ze*	*pb*
	pb	*ze*	*pb*	*pb*

(1). IF y_1 is *nb* AND x_3 is *nb* THEN y_2 is *nb*,
(2). IF y_1 is *nb* AND x_3 is *ze* THEN y_2 is *nb*,
(3). IF y_1 is *nb* AND x_3 is *pb* THEN y_2 is *ze*,
(4). IF y_1 is *ze* AND x_3 is *nb* THEN y_2 is *nb*,
(5). IF y_1 is *ze* AND x_3 is *ze* THEN y_2 is *ze*,
(6). IF y_1 is *ze* AND x_3 is *pb* THEN y_2 is *pb*,
(7). IF y_1 is *pb* AND x_3 is *nb* THEN y_2 is *ze*
(8). IF y_1 is *pb* AND x_3 is *ze* THEN y_2 is *pb*,
(9). IF y_1 is *pb* AND x_3 is *pb* THEN y_2 is *pb*.

(iii). Rules for FLS_3 :

		x_4		
		nb	ze	pb
y_2	nb	*nb*	*nm*	*ns*
	ze	*ns*	*ze*	*ps*
	pb	*ps*	*pm*	*pb*

(1). IF y_2 is *nb* AND x_4 is *nb* THEN y is *nb*,
(2). IF y_2 is *nb* AND x_4 is *ze* THEN y is *nm*,
(3). IF y_2 is *nb* AND x_4 is *pb* THEN y is *ns*,
(4). IF y_2 is *ze* AND x_4 is *nb* THEN y is *ns*,
(5). IF y_2 is *ze* AND x_4 is *ze* THEN y is *ze*,
(6). IF y_2 is *ze* AND x_4 is *pb* THEN y is *ps*,
(7). IF y_2 is *pb* AND x_4 is *nb* THEN y is *ps*
(8). IF y_2 is *pb* AND x_4 is *ze* THEN y is *pm*,
(9). IF y_2 is *pb* AND x_4 is *pb* THEN y is *pb*.

Similar to the parallel case, the index values obtained for subsystems 1, 2 and 3 are $\mathrm{N}_1 = 0.083$, $\mathrm{N}_2 = 0.083$, $\mathrm{N}_3 = 0.194$, $\mathrm{F}_1 = 0.493$, $\mathrm{F}_2 = 0.493$ and $\mathrm{F}_3 = 0.605$ (the details are shown in Table 3.2). There are three layers in this HFS ($q = 3$). The layer-weights using equation (3.2) are $l_1 = 0.5$, $l_2 = 0.333$ and $l_3 = 0.167$. According to equation (3.1), the suggested interpretability index H_I of this HFS are:

$$\begin{aligned}
\mathrm{H}_I &= \sum_{j=1}^{q}(l_j \sum_{k=1}^{s_j} E_{jk}/s_j) \\
&= l_1((N_1/1) + l_2(N_2/1) + l_3(N_3/1) \\
&= 0.5(0.083) + 0.333(0.083) + 0.167(0.194) \\
&= \mathbf{0.102}
\end{aligned}$$

$$\begin{aligned}
\mathrm{H}_I &= \sum_{j=1}^{q}(l_j \sum_{k=1}^{s_j} E_{jk}/s_j) \\
&= l_1((F_1/1) + l_2(F_2/1) + l_3(F_3/1) \\
&= 0.5(0.493) + 0.333(0.493) + 0.167(0.605) \\
&= \mathbf{0.512}
\end{aligned}$$

3.6 Result and Discussion

The proposed initial index, H_I was applied to the example of a seesaw control application for both the Nauck and Fuzzy indices. The overall interpretability measurements of the FLS and HFSs were summarised and are shown in Fig. 3.6.

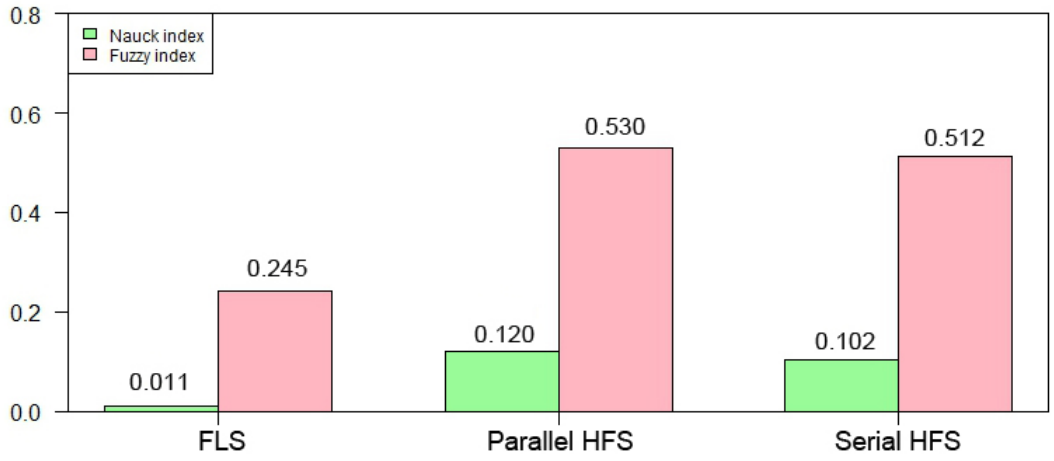


Fig. 3.6 Interpretability measurement between the FLS and HFSs of the seesaw control application using an initial index, H_I.

For the discussed examples that have similar performance (see Section 6.4.1.2), the computed values of interpretability index H_I in the HFSs are larger compared to the FLS regardless of its type (i.e. parallel or serial HFS). As captured in Fig. 3.6, the initial index using Nauck indices for both parallel and serial HFSs are larger than those of the FLS i.e. 0.120 and 0.102 respectively. Figure 3.6 also shows the same pattern with the initial index using Fuzzy indices with values of 0.530 and 0.512 respectively. Overall, in this example of HFSs , the values of the interpretability indices of parallel HFSs are larger than for serial HFSs as shown in both the Nauck and Fuzzy indices of H_I.

The results generated for the initial index H_I follow intuition in the sense that one expects *better* interpretability for HFSs than for their FLS counterparts. Clearly, the aggregation strategy adopted in this chapter is not unique and a number of different aggregation strategies may be appropriate. In the same context, based on the initial results, it is clear that further work is required to establish ground truth on what the appropriate interpretability index should be, so that it reflects or approximates the interpretability of a system as perceived by an actual human user.

3.7 Summary

Intuitively, HFSs have the potential to improve interpretability in FLSs. However, interpretability indices associated with HFSs have not so far been discussed. In this chapter, as a first step to study the interpretability of HFSs, an initial index has been proposed and applied

to the popular FLS interpretability indices, namely the Nauck and Fuzzy indices. Note that this initial index is not perfect as this was the initial proposal to measure interpretability of HFSs. Thus, further exploration of the index will be discussed in Chapter 5.

In initial experiments, the initial index using both Nauck and Fuzzy indices produced higher results for HFSs than for their functionally equivalent FLS counterparts. This follows intuition as HFSs have been credited with better interpretability compared to FLSs. In the future, a key part of work in this thesis will be to conduct experiments with real users to establish the *perceived* interpretability and to use it to validate and inform appropriate interpretability indices (e.g. : "When is an index based on the Nauck and Fuzzy index most appropriate?"). At the moment, beyond the initial motivation of HFSs, it for example is not clear whether a larger number of layers is necessarily easier to interpret than a larger number of single-layer rules. It should be noted that the seesaw control application used in this Chapter does not have clear semantic meaning on the labels, and hence does not have an 'intuitive' interpretation. This example was also missing the essential semantic properties of the partition. Therefore, the Iris classification problem which has a clear semantic meaning of input, intermediate and output variables was introduced as an example, and used in Chapters 4, 5 and 7.

Chapter 4 focuses on leveraging the information gained from studying interpretability in its human context to inform index aggregation in HFSs. Further, other aspects of interpretability will be explored, including the linguistic meaning of fuzzy sets, the relationship between interpretability and complexity, and the logical complexity of the rules, i.e. the impact on interpretability of for example using AND, rather than OR.

Chapter 4

An Exploratory User Study on Perception of Interpretability in Hierarchical Fuzzy Systems

As discussed in Chapter 2, one of the fundamental issues in FLSs is rule explosion, which may reduce interpretability of FLSs. As is known, HFSs can address rule explosion and thus provide a potentially valuable pathway to improve interpretability and minimise complexity in FLSs. However, nothing is known regarding how interpretability can be measured when a system is hierarchical, that is, with multiple subsystems, layers and topologies. Also, basic interpretability indices that have been proposed to measure the interpretability only focus on FLSs. Therefore, in Chapter 3, an initial index to measure the interpretability of HFSs was proposed. However, since interpretability is of a subjective nature, it is essential to see how people perceive interpretability in FLSs and also HFSs.

In this chapter, a user study for capturing the interpretability of FLSs and HFSs is presented. The user study attempts to provide a complete presentation of the interpretability from users' point of view in the Iris flower classification example with the following objectives: (i) to examine how people perceive interpretability in FLSs and HFSs; (ii) to investigate the influence of HFSs' architecture, namely the impact of multiple subsystems, layers and topologies towards interpretability; and (iii) to explore the relationship between interpretability and complexity of FLSs and HFSs. Hence, this chapter addresses the second objective of this thesis; that is to establish standard methodologies for capturing the users' perceptions on the interpretability and complexity of HFSs.

4.1 Introduction

Interpretability is acknowledged as one of the most appreciated advantages of FLSs [115], particularly in relation to applications such as knowledge extraction and decision support [2, 3]. Specifically, one of the most important motivations for using FLSs for system modelling is that a fuzzy system uses linguistic variables and rules [63] that are easy to understand. However, this rule based structure also poses significant challenges, including the *curse of dimensionality*, where the number of required rules and model complexity commonly increase exponentially with the number of input variables [22, 31], thus potentially reducing the transparency and interpretability of FLSs. From these challenges, we can summarise the relationship between the number of rules, complexity and interpretability in FLSs as follows:

$$\uparrow \textit{Number of rules} \text{ (FLSs)}, \uparrow \textit{Complexity} \text{ (FLSs)} \tag{4.1}$$

$$\uparrow \textit{Complexity} \text{ (FLSs)}, \downarrow \textit{Interpretability} \text{ (FLSs)} \tag{4.2}$$

where $\uparrow$ denotes increasing and $\downarrow$ indicates decreasing.

Many of the approaches [14, 22, 33, 38, 40–43] as mentioned earlier (See Section 2.2.5) are designed to overcome the *curse of dimensionality* by reducing the model complexity of FLSs [92]. Moreover, by reducing model complexity, a way may be found to improve model interpretability, i.e., a system with a smaller number of rules requires a smaller effort to be interpreted [56]. However, the following questions arise: (i) "Is it always true that reducing the complexity of FLSs improves the interpretability of FLSs?"; and (ii) "Does an FLS with fewer rules always have better interpretability compared to an FLS with more rules?".

In [55] Ishibuchi et al. demonstrated, using simple test problems, that the minimisation of complexity does not always lead to maximisation of interpretability, particularly in the design of fuzzy rule-based systems. However, the problem is that interpretability of FLSs or HFSs is a very difficult concept, because of its subjective nature. In Chapter 3, an initial index to measure interpretability of HFSs has been introduced. Still, evaluating how different people perceive interpretability is not a trivial matter. Therefore, it is important to analyse how people perceive interpretability and also the complexity of design when it comes to the creation of FLSs. In other words, from a human perspective, "does the minimisation of FLS complexity of design lead to the maximisation of FLS interpretability?". From the literature, this intriguing question has never been addressed through the use of a user study.

This chapter introduces a user study to assess users' perception on interpretability and also complexity of FLSs. While the number of participants in the study is still limited, the

chapter puts forward a valuable first step both in terms of using user studies in respect to exploring the properties of FLSs, and in addressing the key question of complexity verses interpretability.

The rest of this chapter is organised as follows: Section 4.2 introduces a novel method for capturing users' perception of interpretability and complexity. The experiments to showcase the features of the proposed method using a real-world example are shown in Section 4.3. Section 4.4 presents a general discussion on the produced result. Finally, a summary of this chapter is presented in Section 4.5.

4.2 Capturing Users' Perception on Interpretability and Complexity

4.2.1 Motivation

Due to the subjective nature of the interpretability, conducting a user study on the interpretability is not an easy task. The challenges include "How can users be asked about the interpretability and complexity of FLSs?", "What are the questions?" and "How can this user study be conducted?". Thus, a new approach in conducting the user study on assessing how users perceive the interpretability and complexity of FLSs and HFSs has been proposed.

4.2.2 Application Background – Iris Classification Systems

This chapter uses the well known Iris flower classification as an example in this user study. Note that this Iris flower classification example has been introduced earlier in Chapter 2.3. The Iris classification problem consists of four attributes as input features: *sepal length*, *sepal width*, *petal length* and *petal width*. Three classes of Iris flowers are output features: *Setosa*, *Versicolor* and *Virginica*.

In this chapter, three Iris classification systems are introduced (flat FLS, Parallel HFS and Serial HFS) as can be seen in topologies Fig. 2.9, Fig. 2.12 and Fig. 2.10 (see Chapter 2), respectively. They are classified into two configurations: three systems with two membership functions (MFs), referred to as Set MF-2, termed F-2, P-2 and S-2; and three systems with three MFs, referred to as Set MF-3, termed F-3, P-3 and S-3. The examples of MFs for Set MF-2 and MF-3 are shown in Figs. 4.1 and 4.2, and the complete rules for these Iris classification systems can be seen in Appendix A. The description of these Iris classification systems can be seen in Table 4.1.

Table 4.1 Iris classification systems

Sets	Systems	Number of inputs	Number of MFs	Number of rules
	F-2	4	2	16
Set MF-2	P-2	4	2	12
	S-2	4	2	12
	F-3	4	3	81
Set MF-3	P-3	4	3	27
	S-3	4	3	27

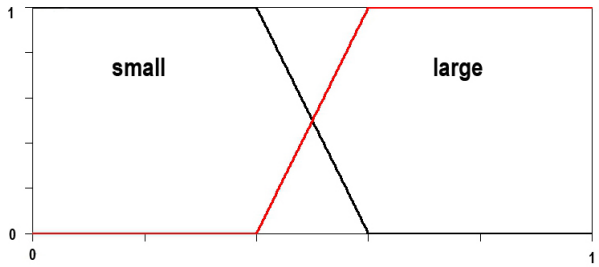


Fig. 4.1 Membership functions of input *Sepal Length* with two MFs.

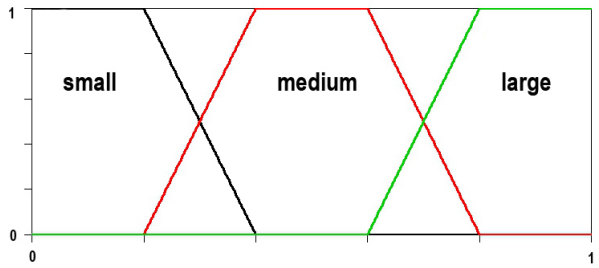


Fig. 4.2 Membership functions of input *Sepal Length* with three MFs.

4.2.3 The User Study Approach

Setting it apart from the existing research on interpretability, this chapter looks at the interpretability and complexity of FLSs and HFSs using a user study. Specifically, a user study to assess the interpretability and complexity of the designs in relation to creating FLSs and HFSs is proposed. Fig. 4.3 shows the steps involved in developing this user study:

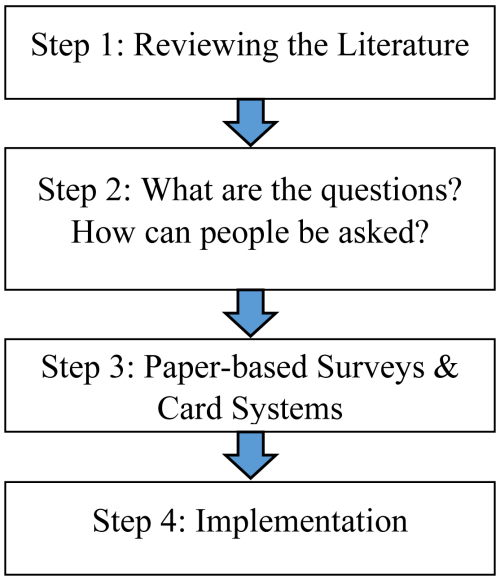


Fig. 4.3 Four steps towards developing a user study

4.2.3.1 Step 1 - Reviewing the Literature

This chapter begins by carefully reviewing the literature that focuses on the problems and issues in relation to interpretability and complexity.

4.2.3.2 Step 2 - What are the questions? How to ask people?

It is not a straightforward task to represent a system in a paper-based (survey), together with a set of questions of interpretability, to involve users in a study. For example, how can we transform these systems into a paper-based format that can illustrate all of the descriptions as in Table 4.1, while achieving the aim of assessing how interpretability and complexity are perceived by users when evaluating these systems on paper?

4.2.3.3 Step 3 - Paper-based Surveys and *Card Systems*

As it is challenging to present Set MF-2 and Set MF-3 in a paper-based survey, it was decided that each of the Iris classification systems in Set MF-2 and Set MF-3 should be printed on an A4 card – called a *card system*. Within this context of survey research, the card system is a visual aid used predominantly during in-person surveys [171]. The topology, membership functions and rule set of each system in Set MF-2 and Set MF-3 were printed on the *card system* (as can be seen in Fig. 4.4).

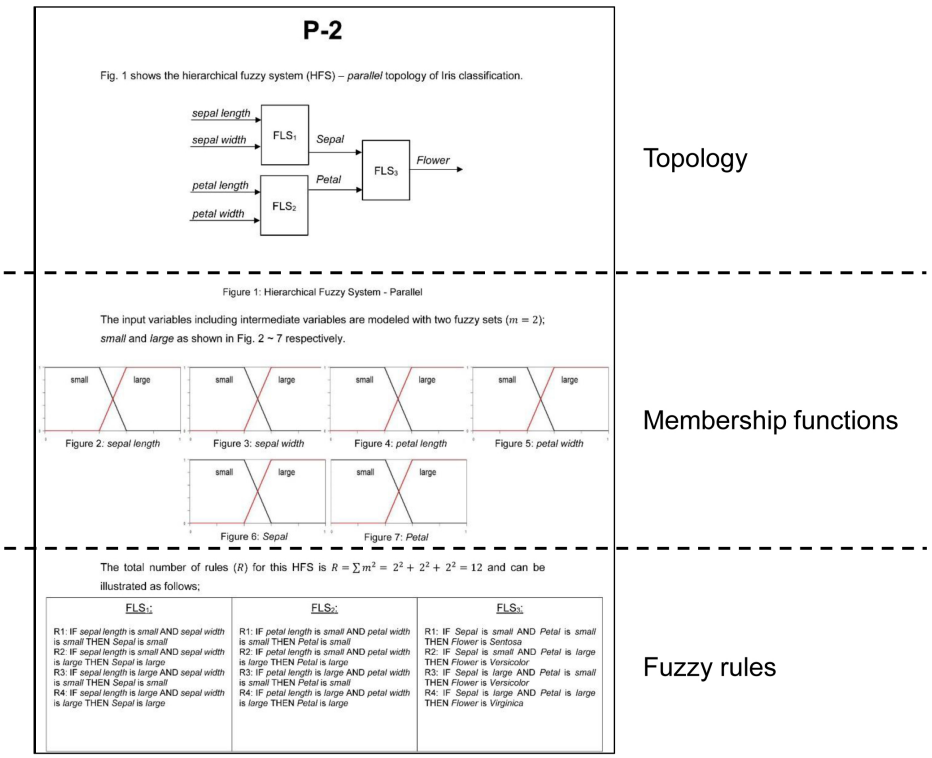


Fig. 4.4 *Card system*: P-2

In their study, Alonso et al. [15] found that, when it comes to interpretability, participants felt much more confident in giving rankings rather than giving numerical values. By following their method in the design of questions, it was decided to ask the participants to provide a rank for the interpretability and complexity of the design of each system in Set MF-2 and Set MF-3. The questionnaire consists of nine main questions and can be illustrated as follows:

i. *Please rank the following topologies; FLS, Parallel HFS and Serial HFS from simpler (1) to the least (3). You may provide an equal rank if you feel there is no difference between options.*

ii. *Please rank the following systems; F-2, P-2 and S-2 in order of interpretability from 1 to 3, where 1 is more interpretable, and 3 is less interpretable. You may provide an equal rank if you feel there is no difference between options.*

iii. *Please rank the following systems; F-2, P-2 and S-2 in term of design complexity from 1 to 3, where 1 is less complex, and 3 is more complex. You may provide an equal rank if you feel there is no difference between options.*

iv. *Please rank the following systems; F-3, P-3 and S-3 in order of interpretability from 1 to 3, where 1 is more interpretable, and 3 is less interpretable. You may provide an equal rank if you feel there is no difference between options.*

v. *Please rank the following systems; F-3, P-3 and S-3 in term of design complexity from 1 to 3, where 1 is less complex, and 3 is more complex. You may provide an equal rank if you feel there is no difference between options.*

vi. *Consider P-3 and S-3. Which of these HFS systems would you prefer to design?*

vii. *Consider F-2 and P-2. Does the reduction of 16 to 12 rules in system F-2 to P-2 make it worth to choose HFS?*

viii. *Consider F-3 and P-3. Does the reduction of 81 to 27 rules in system F-3 to P-3 make it worth to choose HFS?*

ix. *Consider P-3 and S-3. Which of HFS system is more interpretable to you?*

To view all the questions in this questionnaire completed with all *card systems*, participation information sheet and consent form, please refer to the Appendix A.

4.2.3.4 Step 4 - Implementation

The paper-based survey has been approved by the School of Computer Science Ethics committee, University of Nottingham. The survey was carried out at the IEEE International Conference on Fuzzy Systems 2017, Naples, Italy. For example, as shown in Fig. 4.5, the participants were asked to rank the three Iris systems on the *card systems* (namely F-2, P-2 and S-2 cards) based on their view of the interpretability and complexity of the designs. In terms of the ranking of interpretability, the participants may provide rank 1 to indicate the most interpretable of the Iris systems, and rank 3 to indicate the least interpretable of the Iris systems. In contrast, in terms of ranking complexity of design, the participants may provide rank 1 to indicate the least complex to design, and rank 3 for the most complex.

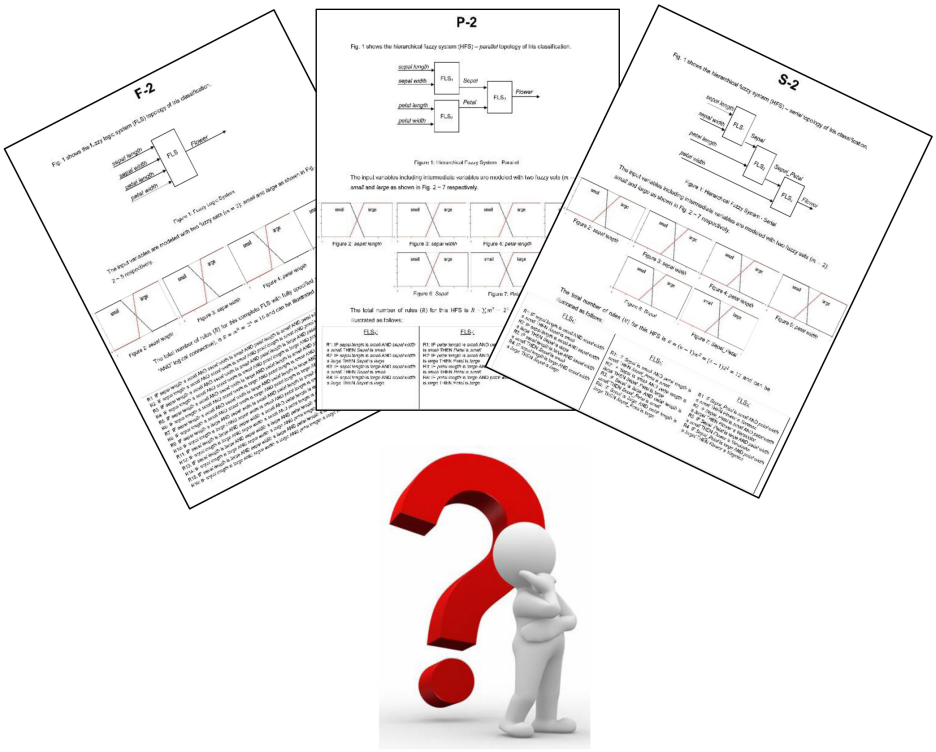


Fig. 4.5 Participants rank F-2, P-2 and S-2 based on the interpretability and complexity of design in relation to creating the Iris classification systems

4.3 Experiment and Results

This section presents the experiment and results of participant perceptions on the interpretability and complexity of the implemented paper-based survey. The aim of this section is to explore the relationship between interpretability and complexity of different FLS designs when creating various Iris classification systems. The analysis process consists of three main steps: (i) descriptive analysis of the survey data; (ii) participants' feedback; and (iii) interpretability vs complexity.

4.3.1 Descriptive Analysis of the Survey Data

This paper-based survey received responses from 25 participants at Fuzz-IEEE 2017. From this data, the survey data are summarised and described in a meaningful way using basic frequency tables.

i. Question 1 – Which topology is simpler?

In this survey, Question 1 aims to examine how participants perceived topologies of the FLS, Parallel HFS and Serial HFS based on Iris classification systems.

Table 4.2 Frequency of topology rankings in Question 1 extracted from survey (from simpler (1) to the least (3)).

Topologies	Rank						Total		Average Rank
	1		2		3				
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	
FLS	17	68	1	4	7	28	25	100	1.60
Parallel HFS	7	28	11	44	7	28	25	100	2.00
Serial HFS	4	16	5	20	16	64	25	100	2.48

Table 4.2 shows participants' perception by providing the topology rankings of the FLS, Parallel HFS and Serial HFS of Iris classification systems. The last column summarises the entire table, providing the average rank. In general, the result showed that participants feel that the topology FLS is simpler by looking at it rather than the Parallel HFS and Serial HFS with the average rank value is close to 1. Also, the result of which topology is simpler that chosen by participants is illustrated in Fig 4.6. Consequently, Fig 4.6 provides a clear picture of the topology rankings from participants' point of view.

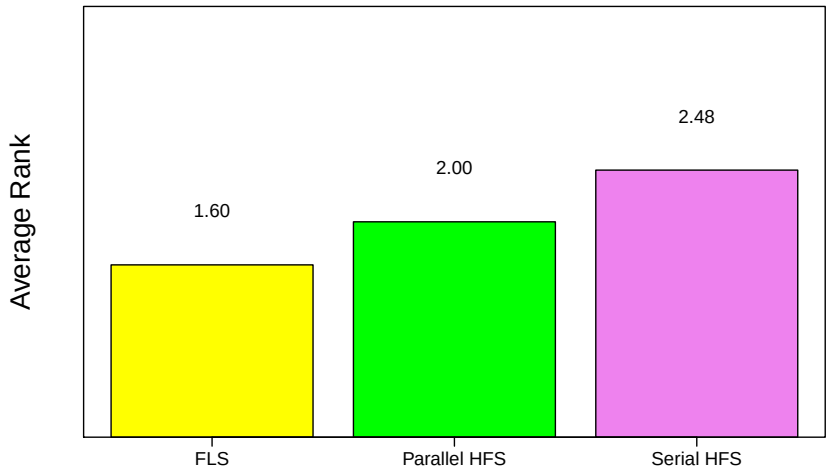


Fig. 4.6 Participants ranked the topologies of the Iris systems based on simplicity.

ii. Question 2 – Which systems in Set MF-2 is more interpretable?

In this survey, Question 2 intends to analyse how participants perceived interpretability of F-2, P-2 and S-2 in Set MF-2 of Iris classification systems.

Table 4.3 Frequency of interpretability rankings in Question 2 extracted from survey (from most interpretable (1) to the least (3)).

Iris systems	Rank						Total		Average Rank
	1		2		3				
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	
F-2	5	20	8	32	12	48	25	100	2.28
P-2	19	76	4	16	2	8	25	100	1.32
S-2	2	8	9	36	14	56	25	100	2.48

Table 4.3 shows the result of interpretability rankings for F-2, P-2 and S-2 in Set MF-2 of the Iris classification systems. The last column summarises the entire table, providing the average rank. In general, most of the participants found P-2 to be more interpretable than F-2 and S-2 in Set MF-2, as seen from the average rank that is close to 1. Additionally, the average rank reveals that the participants also ranked F-2 as more interpretable than S-2, with the average rank of 2.28 being less than 2.48, as shown in Table 4.3. To summarise this result, Fig. 4.7 gives a clear overview of which systems in Set MF-2 that are more interpretable from the participants' point of view.

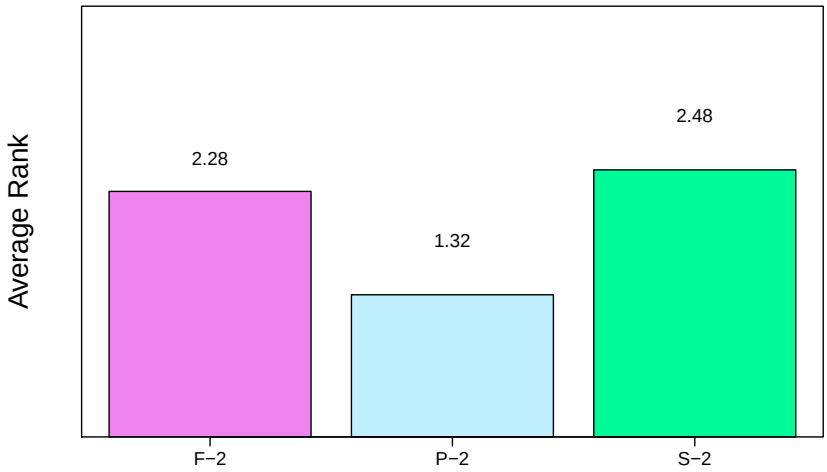


Fig. 4.7 Participants ranked the Iris systems in Set MF-2 based on interpretability.

iii. Question 3 – Which systems in Set MF-2 is less complex to design?

In contrast to Question 2, Question 3 aims to investigate how participants perceived the complexity of the previous systems, namely F-2, P-2 and S-2 in Set MF-2.

Table 4.4 Frequency of complexity rankings in Question 3 extracted from survey (from less complex (1) to more complex (3)).

Iris systems	Rank						Total		Average Rank
	1		2		3				
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	
F-2	12	48	1	4	12	48	25	100	2.00
P-2	10	40	10	40	5	20	25	100	1.80
S-2	6	24	7	28	12	48	25	100	2.24

Table 4.4 shows the result of complexity rankings for F-2, P-2 and S-2 in Set MF-2 of the Iris classification systems. The last column summarises the entire table, providing the average rank. From the observation, most of the participants found P-2 to be less complex to design than F-2 and S-2 in Set MF-2, with the average rank being close to 1. Additionally, the result shows that the participants also rated F-2 as less complex to design than S-2, with the average rank of 2.00 being less than 2.24, as shown in Table 4.3.

To summarise this result, Fig. 4.8 presents a clear picture of which Iris classification systems in Set MF-2 are less complex to design from the participants' point of view.

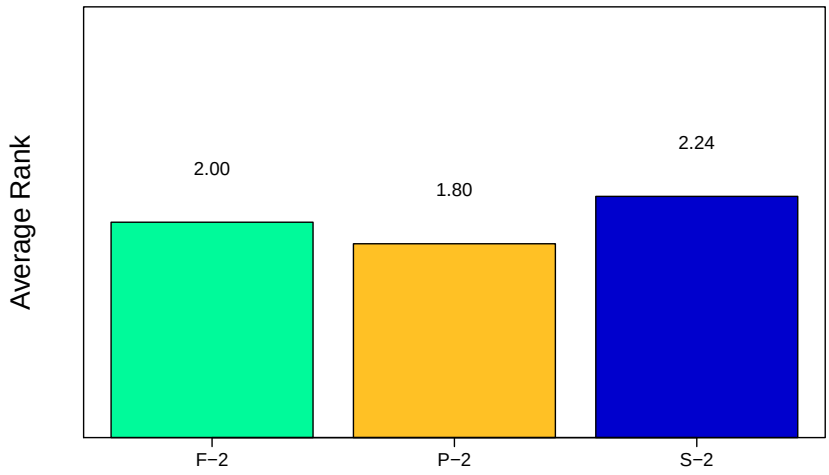


Fig. 4.8 Participants ranked the Iris classification systems in Set MF-2 based on the complexity of design.

iv. Question 4 – Which systems in Set MF-3 is more interpretable?

Similar to Question 2, Question 4 also aims to analyse how participants perceived interpretability but now in Set MF-3, namely F-3, P-3 and S-3.

Table 4.5 Frequency of interpretability rankings in Question 4 extracted from survey (from more interpretable (1) to the least (3)).

Iris systems	Rank						Total		Average Rank
	1		2		3				
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	
F-3	5	20	5	20	15	60	25	100	2.40
P-3	18	72	4	16	3	12	25	100	1.40
S-3	4	16	11	44	10	40	25	100	2.24

In Set MF-3, by increasing the number of membership functions from two to three, all rule sets in these Iris classification systems also increase as can be seen in Table 4.1. The last column summarises the entire table, providing the average rank. Participants

still perceived similar interpretability for these Iris classification systems as shown in Table 4.5. The participants still found P-3 to be more interpretable than F-3 and S-3 in Set MF-3, with the average rank being close to 1. However, for the case of Set MF-3, the result shows that the participants rated S-3 as *more* interpretable than F-3, with the average rank of 2.24 is less than 2.4, as shown in Table 4.5. To review this result, Fig 4.9 provides a clear overview of Iris classification systems in Set MF-3 that are more interpretable from the participants' point of view.

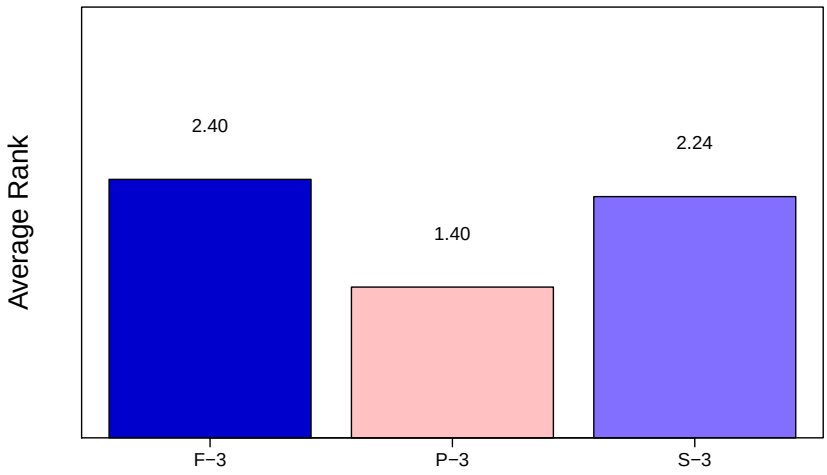


Fig. 4.9 Participants ranked the Iris systems in Set MF-3 based on interpretability.

v. Question 5 – Which systems in Set MF-3 is less complex to design?

Similar to Question 3, Question 5 also aims to examine how participants perceived the complexity of these systems, but now in Set MF-3 namely F-3, P-3 and S-3.

Table 4.6 Frequency of complexity rankings in Question 5 extracted from survey (from less complex (1) to the more complex (3))

Iris systems	Rank						Total		Average Rank
	1		2		3				
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	
F-3	10	40	2	8	13	52	25	100	2.12
P-3	10	40	10	40	5	20	25	100	1.20
S-3	10	40	6	24	9	36	25	100	1.96

Table 4.6 shows the result of complexity rankings for F-3, P-3 and S-3 in Set MF-3 of the Iris classification systems. The last column summarises the entire table, providing the average rank. Likewise, for the case of Set MF-3, most of the participants also chose P-3 to be less complex to design than S-3 and F-3, with the average rank being close to 1. The results also show that the participants chose S-3 is less complex to design than F-3, with the average rank of 1.96 being less than 2.12, as shown in Table 4.6. To summarise this result, Fig 4.10 presents a clear overview of Iris classification systems in Set MF-3 that are less complex to design from the participants' point of view.

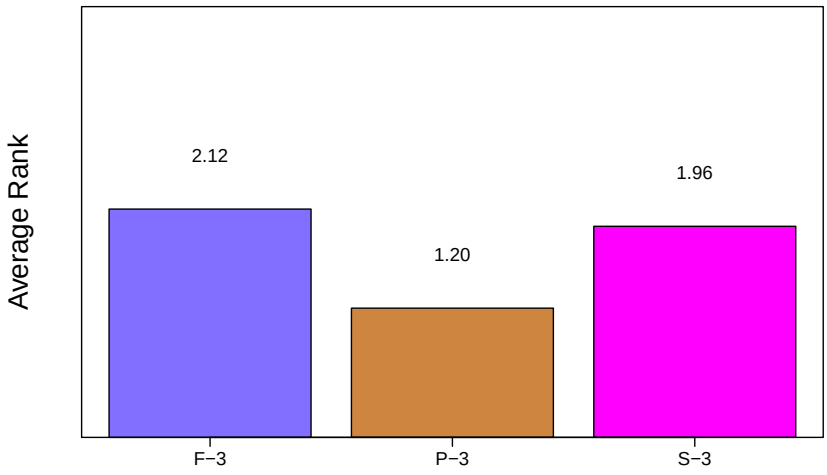


Fig. 4.10 Participants ranked the Iris systems in Set MF-3 based on the complexity of design.

vi. Question 6 – Which HFS systems in Set MF-3 would you prefer to design?

In this survey, Question 6 aims to examine the participants' perception of specific HFS Iris classification systems in Set MF-3 that they would prefer to design.

Table 4.7 Frequency of participants' answers in Question 6 extracted from survey

Iris systems	Frequency	
	Count	(%)
P-3	18	72
S-3	7	28
Total	25	100

Table 4.7 shows the result of participants' answers on the HFS Iris classification systems that they would prefer to design. As can be seen in Table 4.7, most of the participants favoured designing for P-3 rather than S-3. The result reveals that 72% of the participants chose P-3 and only 28% of the participants chose S-3 as also shown in Fig 4.11. Note that this result is associated with the result in Table 4.6 and Fig. 4.10 that showed the participants also perceived P-3 as less complex than S-3 to be designed.

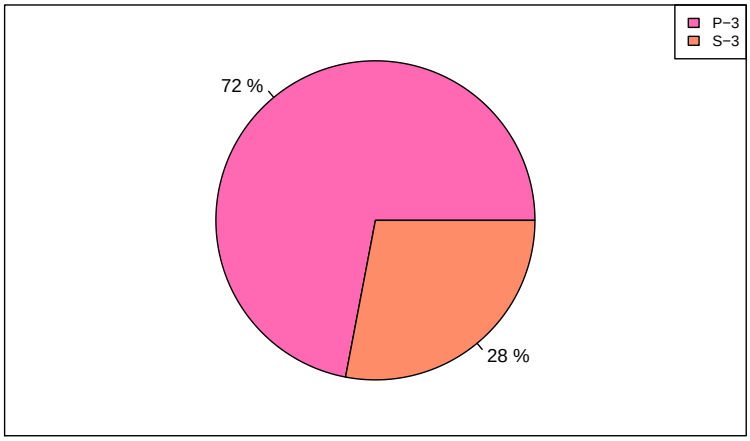


Fig. 4.11 Participants chose HFS Iris systems in Set MF-3 as preferred to design.

vii. Question 7 – Does the reduction from 16 to 12 rules (F-2 to P-2) in Set MF-2 make the HFS worth choosing?

In this survey, Question 7 aims to assess whether the participants prefer the HFS (P-2) rather than the FLS (F-2) when designing Iris classification systems in Set MF-2 with the rule reduction from 16 to 12 rules.

Table 4.8 Frequency of participants' answers in Question 7 extracted from survey

Participants' answer	Frequency	
	Count	(%)
YES	12	48
NO	13	52
Total	25	100

Table 4.8 shows the results received from participants on whether it is worth choosing the HFS (P-2) over the FLS (F-2) in Set MF-2 for the rule reduction from 16 to 12 rules. The results show that a near-equal number of participants *agree* and *disagree* to chose P-2 over F-2. That is, 52% of the participants did not agree to chose P-2 over F-2, whereas 48% of the participants agreed as shown in Fig 4.12.

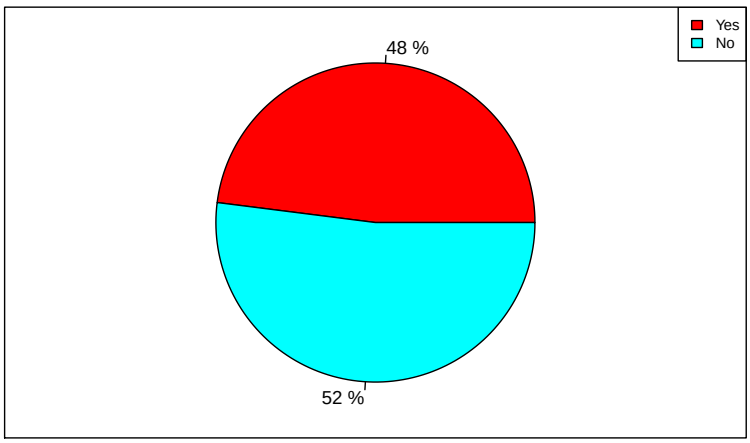


Fig. 4.12 Participants agree to choose the HFS rather than the FLS in Set MF-2

viii. Question 8 – Does the reduction of 81 to 27 rules (F-3 to P-3) in Set MF-3 make the HFS worth choosing?

Similar to Question 7, Question 8 aims to assess whether the participants agree to prefer the HFS (P-3) rather than the FLS (F-3) when designing Iris classification systems in Set MF-3 with the rule reduction from 81 to 27 rules.

Table 4.9 Frequency of participants' answers in Question 8 extracted from survey

Participants' answer	Frequency	
	Count	(%)
YES	19	76
NO	6	24
Total	25	100

Table 4.9 shows the results received from participants on whether that is worths to chose the HFS (P-3) over the FLS (F-3) in Set MF-3 for the rule reduction from 81 to 27 rules. In contrast to Set MF-2, considering the substantial rule reduction (54 rules), most of the participants agreed to choose P-3 over F-3 for Iris classification systems as can be seen in Table 4.9. The results reveal that 76% of the participants agreed to choose P-3 over F-3, whereas 24% of the participants did not agree as shown in Fig 4.13.

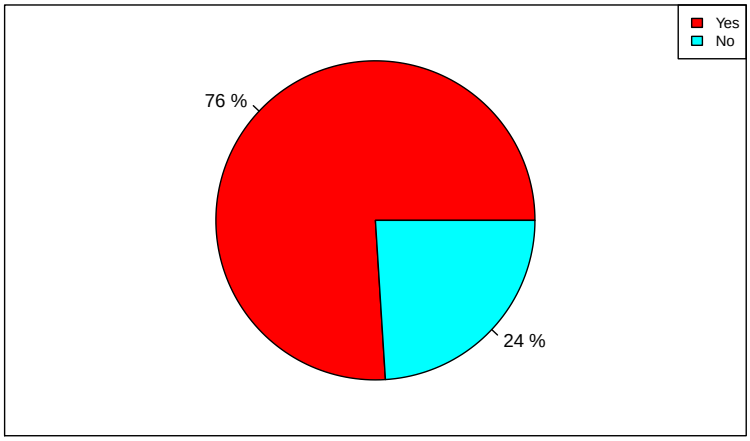


Fig. 4.13 Participants agree to choose the HFS rather than the FLS in Set MF-3

ix. Question 9 – Which HFS systems in Set MF-3 is more interpretable to you?

In this survey, Question 9 aims to observe participants' perception of interpretability of the HFS Iris classification systems in Set MF-3.

Table 4.10 Frequency of participants' answers in Question 9 extracted from survey

Iris systems	Frequency	
	Count	(%)
P-3	20	80
S-3	4	16
Equal	1	4
Total	25	100

Table 4.10 indicates the result of participants' answers on which HFS Iris classification systems are more interpretable from their point of view. As can be seen in Table 4.10, most of the participants chose P-3 rather than S-3 as more interpretable. The result reveals that 80% of the participants chose P-3 and only 28% of the participants chose S-3 as more interpretable as shown in Fig 4.14. However, there was a participant who claimed that both of these HFSs (P-3 and S-3) had equal interpretability.

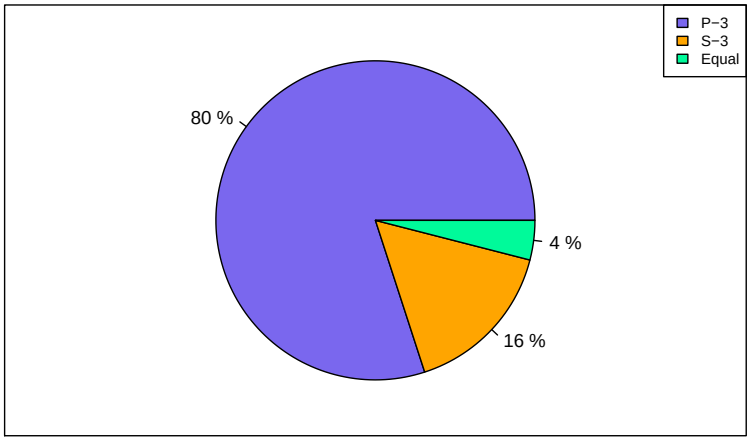


Fig. 4.14 Participants ranked HFS Iris systems in Set MF-3 based on interpretability

4.3.2 Participants' Feedback

In this section, the most important participants' feedback is summarised from the survey data. Whilst, for the illustrative example, there is an absence of a clear relationship between the numerical results obtained for Parallel and Serial HFS systems, the comments of participants (which can be seen in Appendix B) indicate that the Parallel form is the most suited to the example of the Iris system. According to the participants, it is more intuitive when *sepal* and *petal* are classified separately with the resulting outputs driving the classification of the Iris flowers. Similarly, the participants expressed that fewer rules in each subsystem improved their readability. This preference is not intrinsic to the Parallel or Serial form of decomposition but is related to the natural structure (petals and sepals of flowers) inherent in this particular example. The details of comments can be seen in the Appendix B.

4.3.3 Interpretability vs Complexity

In this section, the relationship between interpretability and complexity is investigated from the survey data. Note that, for this investigation, only the survey data extracted from Questions 2 - 5 is covered.

4.3.3.1 Individual Differences in Interpretability and Complexity of Design

First, in this section, the following questions, "Are the FLSs that participants ranked as *more interpretable* those that they ranked as the *less complex* to design?" or indeed "Are the FLSs that participants ranked as *more interpretable* those that they ranked as the *more complex* to design?" are explored. To explore this, Figs. 4.15 - 4.20 present the distribution of the participants' answers in relation to the interpretability and complexity of the design of each Iris classification system as extracted from Tables 4.3 - 4.6.

Figs. 4.15, 4.16 and 4.17 present the distribution of participants' answers in relation to the interpretability and complexity of the designs for creating F-2, P-2 and S-2 respectively in Set MF-2. It is apparent from these figures that there is substantial diversity in terms of the answers with respect to the interpretability and complexity of the design/creation of F-2, P-2 and S-2. Overall, in the case of F-2, the most salient relationship from Fig. 4.15 shows that nine participants (36%) ranked F-2 as *less interpretable* and *more complex* to design. Similar to F-2, the most salient relationship from Fig. 4.17 shows that nine participants (36%) ranked S-2 as *less interpretable* and *more complex* to design. However, for the case of P-2, the most salient relationship from Fig. 4.16 shows that eight participants (32%) ranked P-2 as *more interpretable* and *less complex* to design.

			Complexity of design		
			Less → More		
		Rank	*1*	*2*	*3*
Interpretability	More ↕ Less	*1*	3	0	2
		2	7	0	1
		3	2	1	9

Fig. 4.15 Interpretability vs complexity for F-2 in Set MF-2

			Complexity of design		
			Less ⟷ More		
		Rank	*1*	*2*	*3*
Interpretability	More ↕ Less	*1*	8	7	4
		2	1	3	0
		3	1	0	1

Fig. 4.16 Interpretability vs complexity for P-2 in Set MF-2

			Complexity of design		
			Less ⟷ More		
		Rank	*1*	*2*	*3*
Interpretability	More ↕ Less	*1*	1	0	1
		2	3	4	2
		3	2	3	9

Fig. 4.17 Interpretability vs complexity for S-2 in Set MF-2

Meanwhile, for the case of Set MF-3, Figs. 4.18, 4.19 and 4.20 present a distribution of participant answers ranking the interpretability and complexity of the designs for creating F-3, P-3 and S-3 respectively. As shown in these figures, there is also substantial difference in the answers in relation to the interpretability and complexity of design for creating F-3, P-3 and S-3. Likewise, the relationship in Set MF-3 shows the same pattern as in Set MF-2. The pattern indicates that the most meaningful relationship from Fig. 4.18 identified is that eleven participants (44%) ranked F-3 as *less interpretable* and *more complex* to design. Similar to F-3, the most important relationship from Fig. 4.20 shows that six participants (24%) ranked S-3 as *less interpretable* and *more complex* to design. However, in the case of P-3, the most important relationship from Fig. 4.19 shows that eight participants (32%) ranked P-3 as *more interpretable* and *less complex* to design.

			Complexity of design		
			Less ⟷ More		
		Rank	*1*	*2*	*3*
Interpretability	More ↕ Less	*1*	3	1	1
		2	3	1	1
		3	4	0	11

Fig. 4.18 Interpretability vs complexity for F-3 in Set MF-3

			Complexity of design		
			Less ⟷ More		
		Rank	*1*	*2*	*3*
Interpretability	More ↕ Less	*1*	8	7	3
		2	1	3	0
		3	1	0	2

Fig. 4.19 Interpretability vs complexity for P-3 in Set MF-3

			Complexity of design		
			Less ⟷ More		
		Rank	*1*	*2*	*3*
Interpretability	More ↕ Less	*1*	2	1	1
		2	5	4	2
		3	3	1	6

Fig. 4.20 Interpretability vs complexity for S-3 in Set MF-3

Although only the most important relationships (based on the highest number of participant answers) were discussed for each system, there were also other relationships between interpretability and complexity of design in relation to the creation of F-2, P-2 and S-2 for Set MF-2, and F-3, P-3 and S-3 for Set MF-3. For instance, some participants ranked a

less interpretable system as *less complex* to design. Also, some participants ranked a *less interpretable* system as *more complex* to design as can be seen in Figs. 4.15 - 4.20. Thus, while weaker, these relationships also should be taken into account in order to investigate the significance of the relationship between interpretability and complexity of designs when creating FLSs.

4.3.3.2 Pearson Correlation Coefficient between Interpretability and Complexity

Next, the significance of the relationships between interpretability and complexity of design for creating Iris classification systems that have been captured in Subsection 4.3.3.1 are investigated, using Pearson's correlation coefficient (r) [172]. This statistically establishes how influential a relationship is between two variables where $+1$ indicates a strong positive relationship and -1 indicates a strong negative relationship [173]. Then, to conclude this correlation, the p-value is used to measure the significant level. In this experiment, a significant level of alpha (α) $= 0.05$ is used an indicator as to whether the correlation is statistically significant (p-value < 0.05) or not (p-value > 0.05).

Table 4.11 Pearson correlation coefficient (r) between Interpretability and Complexity of design for Iris classification systems in Set MF-2 and Set MF-3

Iris systems	r	p-value
Set MF-2:		
F-2	0.421	0.036
P-2	0.052	0.804
S-2	0.316	0.123
Set MF-3:		
F-3	0.410	0.042
P-3	0.231	0.266
S-3	0.275	0.184

Table 4.11 provides the summary of the statistics for interpretability and complexity of design for Iris classification systems in Set MF-2 and Set MF-3 that were computed using Pearson's correlation coefficient. First, results indicate that a positive correlation was found between interpretability and complexity of design in relation to creating F-2, P-2 and S-2 for Set MF-2, and F-3, P-3 and S-3 for Set MF-3.

However, as shown in Table 4.11, although all Iris classification systems in Set MF-2 and Set MF-3 gave positive correlations (r) between interpretability and complexity of design, only F-2 and F-3 show statistical significance of this positive correlation (r) in Set MF-2

and Set MF-3 respectively. The computed p-value for F-2 and F-3 are *less than* 0.05 in both Set MF-2 and Set MF-3. The positive correlations (r) in P-2 and P-3 ($r = 0.052$, $r = 0.231$), and also S-2 and S-3 ($r = 0.316$, $r = 0.275$) in Set MF-2 and Set MF-3 respectively, are not statistically significant.

4.4 Discussion

This chapter has presented an initial exploration of the relationship between FLS interpretability and complexity. It was focused specifically on FLSs designed to implement the Iris classification system and consisted of three main steps.

For the first step, the survey data particularly in Questions 2 to 9 were presented, which consists of the interpretability and complexity of designs ranked by participants, as shown in Tables 4.3 to 4.6. It has been found that the majority of the participants chose Parallel HFSs (P-2 and P-3) as more interpretable than the FLS and Serial HFSs (S-2 and S-3) for both Set MF-2 and Set MF-3. Likewise, a similar pattern was also found for the case of complexity of design. It has been found that Parallel HFSs (P-2 and P-3) was chosen by most of the participants as less complex to design when compared to Serial HFSs (S-2 and S-3) and FLSs, in both Set MF-2 and Set MF-3. These findings also correspond with the results presented in Tables 4.7 to 4.10, in that participants perceived that the Parallel HFS was more interpretable and less complex to design than the Serial HFS for both Set MF-2 and Set MF-3. It seems possible that these results are due to different topologies, the number of membership functions, and also the number of rules in each system for both Set MF-2 and Set MF-3, which influences the participants in their ranking of the interpretability and complexity of the designs. From participants' feedback (as described in Section 4.3.2), they found it is more intuitive when sepal and petal are classified separately (as in Parallel HFS) with the resulting outputs driving the classification of the Iris flowers. Clearly, this is probably related to the natural structure (petals and sepals of flowers) inherent in this particular example.

For the second step, an analysis of the individual differences in the interpretability and design complexity rankings for each system was performed as captured in Figs. 4.15 - 4.20. It has been found considerable diversity of relationships between interpretability and complexity of design as given by the participants. For example, some participants stated that the *less interpretable* a system was, the *more complex* it was to design. However, some participants also stated that the *less interpretable* a system was, the *less complex* it was to design. Some possible explanations of this finding include: (i) participants ranked the system as *more interpretable* because they perceived it as being more intuitive and easy to understand

for both Set MF-2 and Set MF-3; (ii) the difference in the number of rules between FLS and HFS in Set MF-3 is larger than Set MF-2– therefore, participants feel that it will be challenging to design an FLS model; and (iii) some participants find it challenging to design parallel and serial HFS as they are composed of multiple subsystems and layers.

In order to examine the significance of the relationship between interpretability and complexity of design, the Pearson correlation coefficient was calculated for the third step. A key outcome from the evaluation is that a positive correlation (r) between interpretability and complexity of design was found for F-2, P-2 and S-2 in Set MF-2 and F-3, P-3 and S-3 in Set MF-3, as shown in Table 4.11. That is, it indicates that the *more interpretable* a system is, the *less complex* it is to design the system, and, also, the *less interpretable* a system is, the *more complex* it is to design the system – at least in the context of the Iris classification systems considered. However, this positive relationship is statistically significant only for FLSs (F-2 and F-3) and not for Parallel HFSs (P-2 and P-3) and Serial HFSs (S-2 and S-3) for both Set MF-2, and Set MF-3. It is hard to interpret the cases for Parallel HFSs and Serial HFSs because the statistical power is low due to the small sample size. Nevertheless, it suggests that the effect is weak for the positive correlation for Parallel HFSs and Serial HFSs.

4.5 Summary

In this chapter, an initial user study has been used to explore the relationship between interpretability and complexity of design when creating FLSs. This has never been addressed by researchers in this field. Although the current study is based on a small sample of participants and focuses on the particular case of FLSs designed for Iris classification, the findings suggest that, in creating fuzzy systems, it does not always hold true that the *less complex* the design of a system is, the *more interpretable* the system is. This seems especially true when an FLS is hierarchical. It has also been found that the topologies of the FLSs (e.g. flat, parallel and serial) have a significant influence on the interpretability and complexity of the design of FLSs.

In Chapter 5, the information gained from this user study will be leveraged to guide the measurement of interpretability that was presented in Chapter 3. This approach, called a participatory design approach, enables the subjective views of a range of users to be taken into account in shaping the measurement of interpretability of HFSs.

Importantly, this chapter found that complexity is an essential component in interpretability. It shows that from the users' perspective, reducing complexity can improve the interpretability. However, it also shows that from the users' perspective, there is diversity in the

complexity, particularly for HFSs. They perceived that it is not only the rules that affect complexity, but also the structure of HFSs. Therefore, in Chapter 6, the complexity of HFSs will be investigated, with the aims of improving the complexity measurement in HFSs.

Chapter 5

Participatory Design Approach to Guide the Measurement of Interpretability in Hierarchical Fuzzy Systems

As discussed in Chapter 3, the challenges of measuring the interpretability of HFSs include issues such as their topological structure, the number of layers, and the meaning of intermediate variables. Thus, a question arises: “How can we make an informed assessment of how HFSs should be designed to enhance interpretability?". In this chapter, a general index to measure the interpretability of HFSs is proposed, combined with a participatory user design process to create a specific instance of the index for an application context; thus addressing the third objective of this thesis such that to guide the interpretability measurement of HFSs. The participatory design approach is built based on the study presented in Chapter 3, that produced an initial index. This approach enables the subjective views of a range of practitioners and experts in the design and creation of FLSs (see Chapter 4) to be taken into account in shaping the design of a general index for measuring interpretability in HFSs. This design process and general index is demonstrated through a classification application, showing the ability of the resulting index to appropriately capture interpretability as perceived by system design experts.

5.1 Introduction

Participatory design is an approach that involves the participation of users in the design development process to help ensure that the result meets their needs and is usable in practice

[174]. Participatory design has been used to develop solutions to complex problems, especially when dealing with people, such as in educational [175] and medical [176] fields. It provides a methodology towards making the design process co-operative and efficient. As discussed earlier, interpretability is subjective in nature in that it depends on people who judge it. Hence, this approach may be a useful method to assist the development of a measure for the interpretability of HFSs.

This chapter introduces a general index to measure the interpretability of HFSs. A participatory design approach is then used to guide the development of a general index for capturing the interpretability of HFSs, building on the initial index introduced in Chapter 3. The index addresses key challenges arising from the structure of HFSs, incorporates an elementary index for assessing the interpretability of each subsystem, an aggregation strategy for combining the indices of the various subsystems within a single layer, and a layer-weighting strategy that combines layers while capturing the topology of the HFS. Initial demonstration and evaluation using the participatory design approach is presented to configure the general index so as to allow its implementation in practice.

The rest of this chapter is organised as follows: Section 5.2 presents the motivation of this chapter on the introduction of the general index of interpretability for HFSs. The general index for interpretability of HFSs is discussed in Section 5.3, followed by an outline of how the framework is demonstrated in principle in Section 5.4. Section 5.5 introduces the participatory design approach process, consisting of: (i) a participatory user study on peoples' perception of interpretability in which users are provided with various configurations of HFSs; (ii) an experimental study on alternative aggregation and layer-weighting strategies on the same systems; and (iii) a statistical comparison to explore the level of agreement between users' perceptions and the experimental data. The results are discussed in Section 5.6, and finally a summary is presented in Section 5.7.

5.2 Motivation

This chapter will introduce a more general index to capturing interpretability of HFSs, in order to address the following three challenges.

5.2.1 Multiple Individual Subsystems

HFSs are produced by structuring the input variables in FLSs into multiple subsystems (see Chapter 2.3). Each subsystem commonly has a small number of inputs and outputs and a

small rule base, and serves commonly a single purpose [98]. The first challenge may be expressed as “How can the interpretability of each subsystem in an HFS be measured using an index?". This challenge is akin to the principal challenge of capturing standard FLS interpretability using an index.

5.2.2 Aggregation

The second challenge is the choice of aggregation strategy to combine the indices of the various subsystems in an HFS. Several aggregation strategies may be suitable, such as *mean*, *min*, *max* and Order Weighted Average (*OWA*) [177, 178]. An *OWA* is calculated by reordered the associated indices of subsystems in descending order before multiplying them by the weights. In [179], Yager introduces the linguistic quantifier to calculate weights (w) in which he defines certain values of alpha (α) to capture labels such as “At least one" ($\alpha = 0.0$), “At least a few" ($\alpha = 0.1$), “A few" ($\alpha = 0.5$), “Half" ($\alpha = 1.0$), “Most" ($\alpha = 2.0$), “Almost all" ($\alpha = 10.0$) and “All" ($\alpha = \infty$). This is done by assigning weights according to:

$$w_i = \left(\frac{i}{s}\right)^{\alpha} - \left(\frac{i-1}{s}\right)^{\alpha}, \quad i = 1, ..., s, \tag{5.1}$$

where s is the total number of subsystems.

The specific attractiveness of the general operator is that it enables dynamic weighting of the individual interpretability of subsystems (based on the individual interpretability of subsystems such as established by the traditional FLS indices). For example, choosing $\alpha = 0.1$ results in a weighting strategy closely resembling the *max*, in which the most interpretable subsystem in the layer is given the highest weight, the second-most interpretable a substantially lower weight, and so on.

5.2.3 Topology and Layering

Based on the same input variables, HFSs with different topologies may be produced, such as the serial and parallel HFSs shown in [54]. A Parallel HFS can have more than one subsystem per layer (e.g. Fig. 2.12 – see Chapter 2), while Serial HFSs use strictly one subsystem per layer (e.g. Fig. 2.10 – see Chapter 2). Thus, these topologies commonly have a different number of layers. For example, Fig. 2.12 and Fig. 2.10 show two different topologies of HFSs using the same four input variables. Both topologies use the same number of subsystems, but with different numbers of layers in their structure. This leads us to the question: “How

can the interpretability of HFSs with different topologies and number of layers be measured systematically?".

5.3 A General Index for Interpretability of Hierarchical Fuzzy Systems

A general index for interpretability of HFSs needs to assess the interpretability of each subsystem, which occur in layers (as illustrated in Figs. 2.12 and 2.10), and then aggregate these together into a single overall measure of interpretability of the whole system.

5.3.1 HFS Interpretability

This chapter proposes a generic index, referred to as H (for HFS *interpretability*), that allows for alternative aggregation strategies as follows:

$$H = \sum_{j=1}^{q} \left(l_j \bigoplus_{k=1}^{s_j} E_{jk} \right), \tag{5.2}$$

where:

- E_{jk} is the underlying (standard) FLS index associated with the subsystem k at layer j, for example the Nauck (N) or the Fuzzy (F) index;
- $\bigoplus$ represents a general aggregation operator;
- l_j is the weight associated with layer j of the HFS (see below);
- s_j is the number of subsystems located in layer j;
- q is the number of layers of the HFS.

Note that (5.2) returns the original FLS index when applied to a standard FLS because it has only one subsystem and one layer. Further, in this chapter, (5.2) is purposely designed as a linear aggregation across layers, i.e. the *OWA* is only applied within each layer, but not for the aggregation across layers.

Layer-weights, l_j, are associated with each subsystem according to their layer, such that the sum of all *layer-weights* l_j is equal to one regardless of the number of layers q, i.e.:

$$\sum_{j=1}^{q} l_j = 1, \quad l_j \in [0,1]. \tag{5.3}$$

Based on the above, an HFS model is less interpretable when the H is close to 0 and more interpretable when the H is close to 1.

5.3.2 Layer-weighting Strategy

A variety of weighting strategies for the individual layers within HFSs is possible. Here, a key set of alternatives is introduced.

5.3.2.1 Layer Weights Decreasing with Layer

The l_j are arranged in descending order. This is intended to reflect the fact that the structure of most HFSs is formed by having the most influential input variables in the first layer of the hierarchy, the next most important inputs in the second layer, as for example in [43, 44]. Hence:

$$l_1 > l_2 > ... > l_q, \quad j = 1,...,q.$$

In order to achieve this and satisfy (5.3), l_j can be given by:

$$l_j = \frac{2\Big(q-j+1\Big)}{q\Big(q+1\Big)}, \quad j = 1,...,q. \tag{5.4}$$

5.3.2.2 Increasing with Layer

The same principle as above, but with the layer-weights *increasing* with layer depth. This is indicated that the input variables in the last layer of the hierarchy are most important, as given by:

$$l_j = \frac{2j}{q(q+1)}, \quad j = 1,..,q \tag{5.5}$$

5.3.2.3 Equal Weighting

Assigning an equal weight for all layers, as given by:

$$l_j = \frac{1}{q} \tag{5.6}$$

5.4 A General Index Demonstration in Principle

It is intuitive to initially explore the *mean* as an aggregation operator, to both demonstrate the functionality of the H index generally and to explore the behavior of the resulting 'mean-based' H in principle. This approach has initially been explored in Chapter 3 (as in (3.1)) and is summarised here.

Considering the *mean* as aggregation operator, thus it is reffered to as H_{mean}. Now, (5.2) becomes:

$$H_{\text{mean}} = \sum_{j=1}^{q} \left(l_j \sum_{k=1}^{s_j} E_{jk}/s_j \right) \tag{5.7}$$

To demonstrate the behaviour of the resulting H_{mean}, both the Nauck and the Fuzzy indices (as the underlying indices applied to each subsystem) using the well-known Iris flower classification problem [96] have been considered. Again, note that this Iris example is the same as used in Section 2.3 and Subsection 4.2.2. The Iris dataset has four attributes as input features, namely: *sepal length*, *sepal width*, *petal length* and *petal width*; and three classes of iris flowers as output, namely: *Setosa*, *Versicolor* and *Virginica*.

Three individual systems are defined to capture the variety of HFSs' architectures, namely a (standard) FLS (F), a Parallel HFS (P) and a Serial HFS (S). The three systems were each designed in two configurations, where all variables have either two or three membership functions. The various systems are characterised by seven attributes as follows:

1. **Model Type**: Type of fuzzy model, namely F, P and S as shown in Figs. 2.9, 2.12 and 2.10 (see Chapter 2), respectively.

2. **NI**: Number of input variables.

3. **NMF**: Number of membership functions used for all the variables of each model, as shown in Figs. 4.1 and 4.2 (see Chapter 4).

4. **NR**: Total number of rules.

5. **NS**: Number of subsystems in the model.

6. **NL**: Number of layers in the model.

7. **NRS**$_k$: Number of rules in subsystem k.

The attributes for the six systems are summarised in Table 5.1. The complete rule set for each of the six variations of the systems are given in Appendix A.

Table 5.1 Description of the six Iris classification systems

Iris systems	Model Type	NI	NMF	NR	NS	NL	NRS_1	NRS_2	NRS_3
F-2	FLS	4	2	16	1	1	16	-	-
P-2	Parallel HFS	4	2	12	3	2	4	4	4
S-2	Serial HFS	4	2	12	3	3	4	4	4
F-3	FLS	4	3	81	1	1	81	-	-
P-3	Parallel HFS	4	3	27	3	2	9	9	9
S-3	Serial HFS	4	3	27	3	3	9	9	9

5.4.1 Method

In this section, the application of the H_{mean} index to the six variations of the classification systems described above is shown in detail. Both the Nauck and Fuzzy indices are used within the H_{mean} index to enable their comparison. The six systems are then also used in the participatory design experiments described later.

The application of the H_{mean} index to measure the interpretability of each of the six systems is carried out in the following steps:

5.4.1.1 Calculate interpretability for each subsystem

First, the interpretability of each subsystem is calculated using both the Nauck and the Fuzzy indices. For example, the values of the Nauck index for the three subsystems in P-2 (Parallel HFS with 2 MFs) are $N_1 = 0.250$, $N_2 = 0.250$ and $N_3 = 0.375$ (the details of the calculations are shown in Table 5.2).

5.4.1.2 Identify the layer-weights

Next, the values of the layer-weights are computed using (5.4). For instance, for P-2 and P-3 which consist of two layers, the values of the layer weights at each layer are $l_1 = 0.667$ and

$l_2 = 0.333$; where for S-2 and S-3 which consist of three layers, the values of layer weights at each layer are $l_1 = 0.500$, $l_2 = 0.333$ and $l_3 = 0.167$.

5.4.1.3 Calculate the overall interpretability

Then, the overall interpretability can be calculated using the H_{mean} as given in (5.7). For example, the interpretability of model P-2 is computed as follows:

$$\begin{aligned} H\text{mean} &= \sum_{j=1}^{q} (l_j \sum_{k=1}^{s_j} E_{jk}/s_j) \\ &= l_1(\text{N}_1 + \text{N}_2)/2 + l_2(\text{N}_3/1) \\ &= 0.667(0.250 + 0.250)/2 + 0.333(0.375) \\ &= 0.292 \end{aligned}$$

5.4.2 Results

The overall interpretability measurements of the six Iris classification systems calculated using the H_{mean} are shown in Tables 5.2 and 5.3 and summarised in Fig. 5.1. In general, it can be seen that the computed H_{mean} interpretability indices in the various hierarchical models are always larger (i.e. more interpretable) as compared to the interpretability of the standard FLSs, regardless of whether the hierarchical topology is parallel or serial, and regardless of the number of membership functions.

Table 5.2 Interpretability of the Iris systems; F-2, P-2 and S-2 using the H_{mean}

Iris models	Nauck Index				Fuzzy Index							
	comp	$\overline{cov}$	$\overline{part}$	Index	NR	NP	$NR_{i=1}$	$NR_{i=2}$	$NR_{i\geq 3}$	$\overline{terms}$	NC	Index
F-2	0.047	1	1	**0.047**	16	64	0	0	16	2	3	**0.500**
P-2:												
FLS_1	0.250	1	1	0.250	4	8	0	4	0	2	2	0.611
FLS_2	0.250	1	1	0.250	4	8	0	4	0	2	2	0.611
FLS_3	0.375	1	1	0.375	4	8	0	4	0	2	3	0.663
H_{mean}				**0.292**								**0.628**
S-2:												
FLS_1	0.250	1	1	0.250	4	8	0	4	0	2	2	0.611
FLS_2	0.250	1	1	0.250	4	8	0	4	0	2	2	0.611
FLS_3	0.375	1	1	0.375	4	8	0	4	0	2	3	0.663
H_{mean}				**0.271**								**0.620**

Table 5.3 Interpretability of the Iris systems; F-3, P-3 and S-3 using the H_{mean}

Iris models	Nauck Index				Fuzzy Index							
	comp	$\overline{cov}$	$\overline{part}$	Index	NR	NP	$NR_{i=1}$	$NR_{i=2}$	$NR_{i\geq 3}$	$\overline{terms}$	NC	Index
F-3	0.009	1	0.5	**0.005**	81	324	0	0	81	3	3	**0.194**
P-3:												
FLS_1	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_2	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_3	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
H_{mean}				**0.083**								**0.493**
S-3:												
FLS_1	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_2	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
FLS_3	0.167	1	0.5	0.083	9	18	0	9	0	3	3	0.493
H_{mean}				**0.083**								**0.493**

As shown in Fig. 5.1, considering the Nauck index for the two membership function case, the resultant H_{mean} value (i.e. the calculated overall interpretability) is greatest for the parallel HFS model (P-2 $= 0.292$), followed by the serial HFS model (S-2 $= 0.271$), and finally the plain FLS (F-2 $= 0.047$). The same pattern is observed for the Fuzzy index, although the absolute values of interpretability obtained are higher.

Further, as seen in Fig. 5.1, considering the Nauck index for the three membership function case, the computed interpretabilities are higher for both the hierarchical models (P-3 $=$ S-3 $= 0.083$) compared to the plain FLS (F-3 $= 0.005$). However, in this case, the interpretability of both the hierachical models are the same — i.e. the interpretability of the parallel and serial models featuring three membership functions are the same. The same pattern is obtained with the Fuzzy index, albeit with higher absolute values of interpretability.

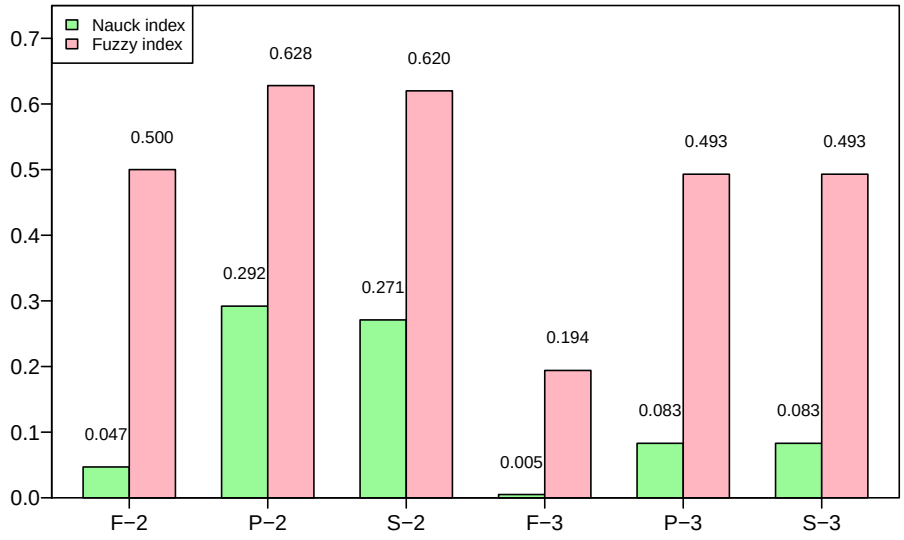


Fig. 5.1 Summary of the overall H_{mean} index

The results generated for the H_{mean} follow intuition in the sense that the HFSs do have better interpretability than the FLSs for all systems. Further, the parallel topology, P-2, is seen to have a better interpretability than the serial topology, S-2. This feature is actually due to a combination of three factors: (i) both the Nauck and the Fuzzy indices rate the interpretability of a (sub-)system consisting of two inputs each with two MFs and an output with three MFs ($2 \times 2 \rightarrow 3$) *higher than* that of a (sub-)system consisting of two inputs each with 2 MFs and an output with two MFs ($2 \times 2 \rightarrow 2$); (ii) the proposed H_{mean} gives higher interpretability to sub-systems in earlier layers; and (iii) P-2 features the ($2 \times 2 \rightarrow 3$) sub-system in layer 2,

whereas S-2 features it in layer 3. This is not repeated in the case of P-3 and S-3, as in these cases all sub-systems are of form $(3 \times 3 \rightarrow 3)$ and so all have equal interpretability; hence, the parallel and serial HFS topologies result in the same interpretability.

The Fuzzy index is designed to provide a measurement of interpretability which is closer to the user's point of view than the Nauck or other indices [15]. Given this and this finding that both produce similar results in this H_{mean} experiments, only the Fuzzy index will be used for the remainder of this chapter in refining the H index. It is clear that the identical values of interpretability of subsystems in Parallel and Serial HFSs in this context will produce the same overall interpretability result regardless of any aggregation operator, as shown in Table 5.3. Therefore, due to this issue, in the next section, another case study will be introduced as an example.

5.5 Beyond the Mean – A Participatory Design Approach

While aggregating the subsystems using the *mean* can be used as an ad-hoc strategy, in order to capture the interpretability of HFSs as perceived by actual users, a participatory design approach to derive H parameters within the general index is proposed. As mentioned earlier, participatory design is an approach that involves users in the design development process to ensure the result satisfies their needs [174].

Following the problem mentioned above (identical interpretability values of subsystems in the Iris classification system), a new case study is introduced. In this experiment, a more complex set of FLS and alternative HFSs is presented, based on the rotary crane system as described in [51]. The rotary crane system is commonly used to resolve the problem of load rotation and hoisting systems such as in the construction, factories, and harbours. A summary of the rotary crane for various HFS designs is presented in Table 5.4.

Table 5.4 A summary of the rotary crane for various HFS designs

Rotary Crane Systems	Number of inputs	Number of MFs	Number of subsystems	Number of layers	Total number of rules
Rotary crane system - A	6	3	1	1	729
Rotary crane system - B	6	3	2	2	252
Rotary crane system - C	6	3	2	2	252
Rotary crane system - D	6	3	2	2	108
Rotary crane system - E	6	3	2	2	108
Rotary crane system - F	6	3	3	2	99
Rotary crane system - G	6	3	3	2	63
Rotary crane system - H	6	3	3	2	63
Rotary crane system - I	6	3	4	3	54
Rotary crane system - J	6	3	4	3	54
Rotary crane system - K	6	3	4	4	54
Rotary crane system - L	6	3	5	5	45

The experiment aims to guide the refinement of the H index using the participatory design approach, in particular around: (i) the aggregation strategy for combining the sub-system indices within a single layer; and (ii) the strategy for assigning weights to the layers.

The H design process consists of three main steps:

1. a participatory user study on peoples' perception of interpretability in which users are provided with various configurations of HFSs, in this case variations of the rotary crane systems as in Table 5.4;

2. an experimental study of alternative aggregation and layer-weighting strategies on the same systems; and

3. the level of agreement between users' perceptions and the experimental data obtained to find the aggregation strategy and layer-weighting strategy to best fit with the perceptions of users.

These processes are now described in detail for the Rotary crane systems.

5.5.1 Participatory User Study

Twelve Rotary crane systems were constructed, termed A ... L. Each system has a different configuration such as the number of rules, number of subsystems and number of layers, as shown in Table 5.4. Illustrations of the topology of each can be seen in Figs. 5.2 to 5.13.

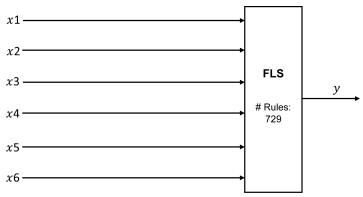


Fig. 5.2 Rotary crane system: A

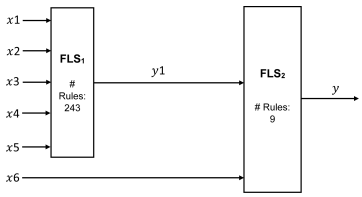


Fig. 5.3 Rotary crane system: B

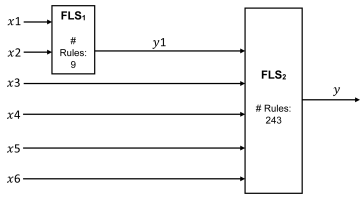


Fig. 5.4 Rotary crane system: C

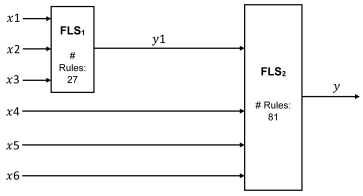


Fig. 5.5 Rotary crane system: D

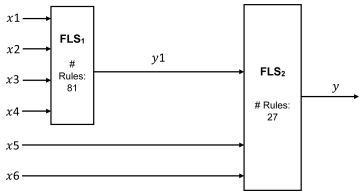


Fig. 5.6 Rotary crane system: E

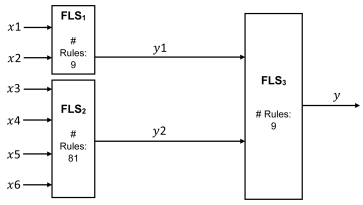


Fig. 5.7 Rotary crane system: F

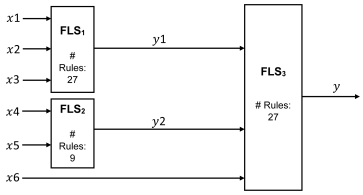


Fig. 5.8 Rotary crane system: G

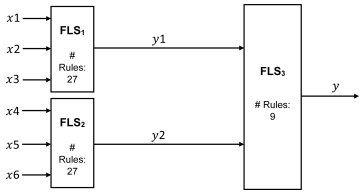


Fig. 5.9 Rotary crane system: H

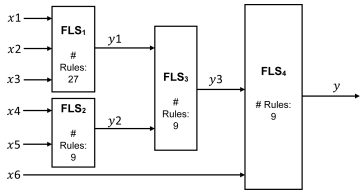


Fig. 5.10 Rotary crane system: I

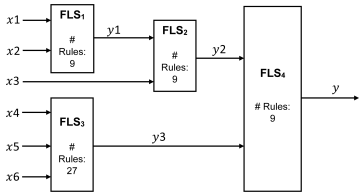


Fig. 5.11 Rotary crane system: J

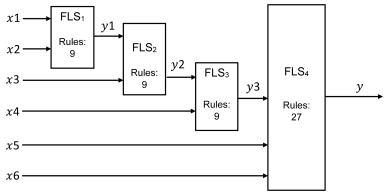


Fig. 5.12 Rotary crane system: K

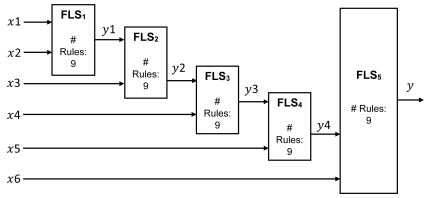


Fig. 5.13 Rotary crane system: L

Similar to the user study in Chapter 4 (see Section 4.2.3), each system was represented on an *A4* card – *card systems*. However, this time, only the topology and rule structures were presented. Participants were asked to choose which system they favoured in terms of interpretability in a set of pairwise comparisons drawn from the total set of possible pairs. The combination of the pairwise comparisons (labelled PW-1 to PW-20) are as follows: (PW-1: A, B), (PW-2: A,C), (PW-3: B,C), (PW-4: B,D), (PW-5: B,E), (PW-6: C,D), (PW-7: C,E), (PW-8: D,E), (PW-9: D,F), (PW-10: E,F), (PW-11: F,G), (PW-12: F,H), (PW-13: G,H),

(PW-14: H,I), (PW-15: I,J), (PW-16: I,J), (PW-17: I,L), (PW-18: J,K), (PW-19: J,L) and (PW-20: K,L).

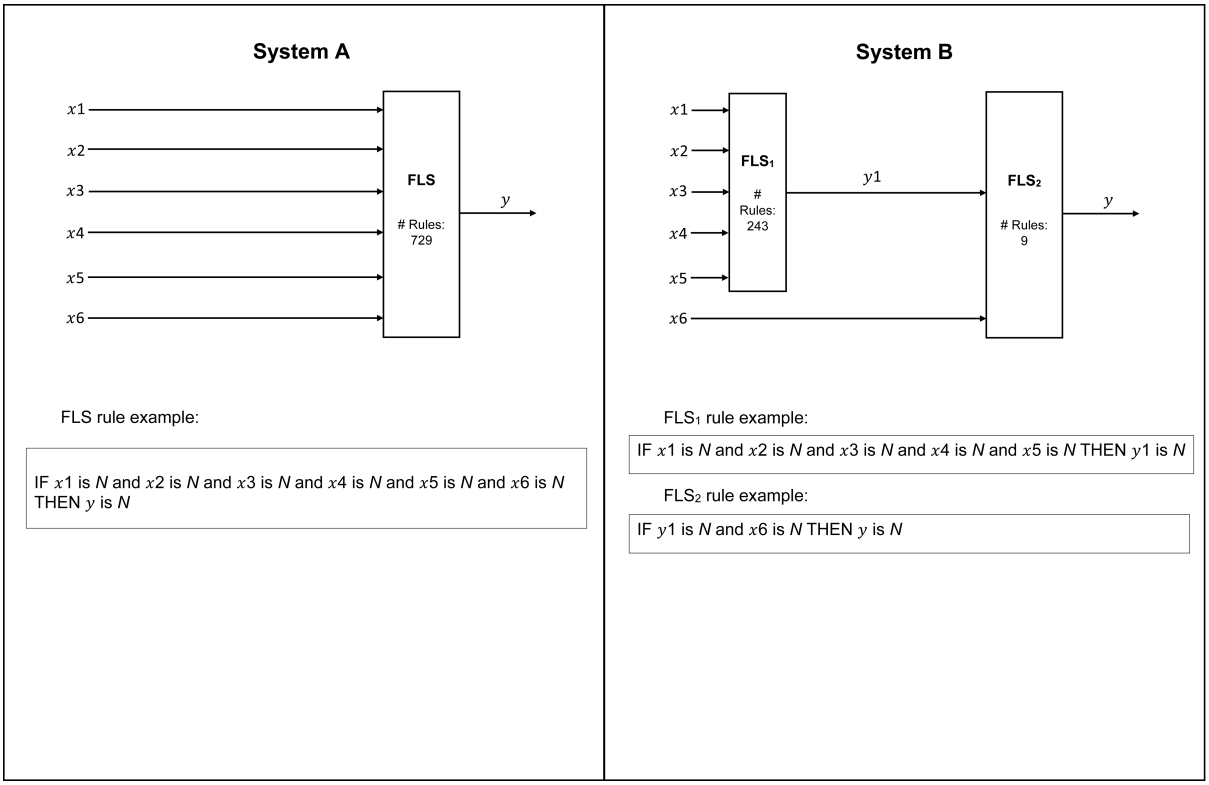


Fig. 5.14 The figure shows a pairwise comparison of Rotary systems (System A and B) shown to participants. The participants select a system that is *more* interpretable for them. If both systems seem *equally* interpretable, they may select the '*Equal*' option as their answer.

For instance, for PW-1, the participants were asked to choose between system A and B, based on their perceived interpretability preference (see in Fig. 5.14 for a mock-up of PW-1). If both systems seem equally interpretable, they could indicate '*Equal*' (EQ) as their answer. This experiment was carried out through an online-survey with 40 participants (termed as U1 to U40) from a wide range of expertise. The complete pairwise comparison questions on these Rotary crane system can be seen in Appendix C.

Table 5.5 Frequency of the users interpretability rating for pairwise comparisons of rotary crane, as extracted from user study

Pairwise comparisons	Users Interpretability Rating												
	A	B	C	D	E	F	G	H	I	J	K	L	EQ
PW-1	21	11	-	-	-	-	-	-	-	-	-	-	8
PW-2	20	-	12	-	-	-	-	-	-	-	-	-	8
PW-3	-	17	8	-	-	-	-	-	-	-	-	-	15
PW-4	-	16	-	14	-	-	-	-	-	-	-	-	10
PW-5	-	16	-	-	13	-	-	-	-	-	-	-	11
PW-6	-	-	8	21	-	-	-	-	-	-	-	-	11
PW-7	-	-	10	-	19	-	-	-	-	-	-	-	11
PW-8	-	-	-	9	13	-	-	-	-	-	-	-	18
PW-9	-	-	-	13	-	25	-	-	-	-	-	-	2
PW-10	-	-	-	-	13	22	-	-	-	-	-	-	5
PW-11	-	-	-	-	-	26	11	-	-	-	-	-	3
PW-12	-	-	-	-	-	3	-	17	-	-	-	-	20
PW-13	-	-	-	-	-	-	6	26	-	-	-	-	8
PW-14	-	-	-	-	-	-	-	25	11	-	-	-	4
PW-15	-	-	-	-	-	-	-	-	18	5	-	-	17
PW-16	-	-	-	-	-	-	-	-	17	-	16	-	7
PW-17	-	-	-	-	-	-	-	-	20	-	-	14	6
PW-18	-	-	-	-	-	-	-	-	-	16	14	-	10
PW-19	-	-	-	-	-	-	-	-	-	15	-	19	6
PW-20	-	-	-	-	-	-	-	-	-	-	7	17	16

Table 5.5 presents the frequency of the participants interpretability ratings for 20 pairwise comparisons. The detail of answers given by each participant to each of the pairwise comparisons are shown in Table 5.6.

Table 5.6 The interpretability ratings for the pairwise comparisons of the Rotary crane system extracted from participatory user study

Users	Pairwise comparisons (PW) of the Rotary crane system																			
	PW-1	PW-2	PW-3	PW-4	PW-5	PW-6	PW-7	PW-8	PW-9	PW-10	PW-11	PW-12	PW-13	PW-14	PW-15	PW-16	PW-17	PW-18	PW-19	PW-20
U-1	A	A	EQ	EQ	EQ	EQ	EQ	EQ	F	F	F	H	H	H	EQ	K	L	EQ	L	L
U-2	B	C	C	D	E	D	EQ	D	D	E	G	H	G	H	EQ	K	L	K	L	L
U-3	B	C	EQ	EQ	EQ	EQ	EQ	EQ	F	F	G	H	H	I	J	I	I	J	J	EQ
U-4	A	A	EQ	B	B	C	C	EQ	F	F	F	F	H	I	I	I	I	J	J	EQ
U-5	B	C	B	B	E	D	E	EQ	F	EQ	G	EQ	EQ	H	EQ	K	I	K	L	L
U-6	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	F	E	F	EQ	H	H	I	K	I	K	J	K
U-7	A	C	EQ	D	E	D	E	EQ	D	E	G	H	H	H	I	K	L	K	L	L
U-8	B	C	EQ	D	E	D	E	EQ	F	F	G	EQ	G	I	EQ	K	L	K	L	L
U-9	A	A	EQ	EQ	EQ	EQ	EQ	EQ	F	F	G	EQ	H	H	EQ	I	I	J	J	EQ
U-10	EQ	EQ	EQ	D	E	EQ	E	EQ	F	F	EQ	EQ	EQ	H	EQ	I	I	J	J	EQ
U-11	B	C	C	D	E	D	E	D	F	E	F	H	EQ	H	EQ	EQ	L	EQ	L	L
U-12	B	C	EQ	D	E	D	E	EQ	D	E	G	H	EQ	H	EQ	EQ	L	EQ	L	L
U-13	A	A	B	B	B	C	C	D	D	E	F	EQ	H	H	EQ	EQ	EQ	EQ	EQ	EQ
U-14	B	C	B	D	E	D	E	E	F	F	G	H	G	I	I	I	L	J	L	L
U-15	A	A	EQ	EQ	EQ	C	E	EQ	D	E	F	EQ	H	H	EQ	I	I	J	J	K
U-16	A	A	B	B	B	EQ	EQ	EQ	D	EQ	F	EQ	H	H	I	I	I	EQ	J	EQ
U-17	A	A	EQ	D	E	D	E	E	F	F	F	H	H	H	J	I	I	K	L	L
U-18	A	A	B	B	E	D	E	E	D	E	F	H	H	I	J	K	I	J	J	L
U-19	B	C	B	B	B	D	E	D	D	F	F	F	H	H	I	K	L	K	L	L
U-20	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ
U-21	A	A	C	B	B	C	C	D	F	F	F	EQ	H	H	I	I	I	J	J	K
U-22	EQ	EQ	C	D	E	D	E	E	F	F	G	EQ	H	I	I	I	I	J	J	EQ
U-23	A	EQ	C	D	E	C	C	D	D	E	F	EQ	G	I	I	I	I	J	J	K
U-24	A	A	C	D	B	D	C	E	F	F	F	F	G	I	I	I	L	K	L	L
U-25	EQ	EQ	B	B	EQ	C	C	D	D	E	G	EQ	EQ	EQ	EQ	I	I	J	J	EQ
U-26	A	A	C	D	B	D	C	D	F	F	F	H	H	H	I	K	I	J	L	K
U-27	EQ	A	B	B	EQ	D	C	EQ	D	E	EQ	EQ	EQ	H	J	EQ	I	K	L	EQ
U-28	A	A	B	EQ	B	D	E	E	F	F	F	H	H	H	I	K	I	EQ	L	K
U-29	EQ	EQ	EQ	B	EQ	EQ	E	EQ	EQ	EQ	F	EQ	G	EQ	EQ	EQ	EQ	EQ	EQ	EQ
U-30	A	A	B	B	B	EQ	E	E	F	F	F	EQ	H	H	I	K	L	K	L	EQ
U-31	A	A	EQ	B	B	C	C	E	F	F	F	H	H	I	I	K	L	K	L	L
U-32	A	A	B	EQ	B	C	EQ	EQ	F	F	F	H	H	H	I	I	I	J	J	EQ
U-33	A	A	EQ	EQ	B	EQ	EQ	EQ	F	EQ	F	EQ	H	H	EQ	K	EQ	EQ	EQ	L
U-34	A	A	B	EQ	EQ	D	EQ	EQ	D	E	F	H	H	H	I	K	L	K	L	K
U-35	B	C	B	B	B	D	E	E	F	F	F	H	H	H	EQ	EQ	EQ	EQ	EQ	EQ
U-36	EQ	EQ	C	D	EQ	EQ	EQ	E	F	F	G	H	H	I	J	K	L	K	L	L
U-37	A	A	B	D	B	D	E	E	F	F	F	EQ	H	H	I	I	I	J	J	EQ
U-38	B	C	B	B	B	D	E	E	F	F	F	EQ	H	EQ	EQ	I	EQ	J	EQ	L
U-39	A	A	B	B	E	D	C	D	D	E	F	H	H	H	I	K	L	K	L	L
U-40	B	C	B	B	B	D	E	E	F	F	F	EQ	EQ	I	EQ	I	I	J	J	EQ

From initial observation of Table 5.6, there is appreciable diversity of opinion in the 40 participants as to the interpretability of the various systems. This also shows that interpretability is very subjective because each participant may perceive the interpretability differently.

5.5.2 Exploration of Alternatives Configurations of *H*

The next step was conducted to explore various alternative aggregations and layer-weighting strategies as described in Section 5.3. Firstly, the Fuzzy index for each of the subsystems present in the twelve different Rotary crane system configurations was calculated, as shown in Table 5.7.

Table 5.7 Base interpretability of each of the Rotary Crane Systems measured using the Fuzzy interpretability index

Rotary Crane Systems	Fuzzy Index for each HFS subsystem				
	FLS_1	FLS_2	FLS_3	FLS_4	FLS_5
System A	0.0216	-	-	-	-
System B	0.0647	0.4932	-	-	-
System C	0.4932	0.0647	-	-	-
System D	0.2408	0.1941	-	-	-
System E	0.1941	0.2408	-	-	-
System F	0.4932	0.1941	0.4932	-	-
System G	0.2408	0.4932	0.2408	-	-
System H	0.2408	0.2408	0.4932	-	-
System I	0.2408	0.4932	0.4932	0.4932	-
System J	0.4932	0.4932	0.2408	0.4932	-
System K	0.4932	0.4932	0.4932	0.2408	-
System L	0.4932	0.4932	0.4932	0.4932	0.4932

Five different aggregation strategies, *mean*, *min*, *max*, and two linguistic *OWAs* (using alpha (α) of 0.1 and 2) were explored. Each was used as a general aggregation operator $\bigoplus$ in the H framework presented in (5.2) — note that only the *decreasing weight* layer-weighting strategy was used in conjunction with the various aggregation strategies. For example, for the case of Linguistic $OWA_{\alpha=0.1}$, the values of $\alpha = 0.1$ will be used in (5.1) to obtain its weights (w) before multiplying them by the reordered subsystems in descending order. Assuming that three Fuzzy index values for each sub-system in System F are $F_1 = 0.4932$, $F_2 = 0.1941$ and $F_3 = 0.4932$, the overall interpretability of Rotary crane system F was computed using H with $OWA_{\alpha=0.1}$ and a *decreasing weight* layer-weighting, as follows:

$$
\begin{aligned}
H_{\text{owa}\alpha=0.1} &= \sum_{j=1}^{q} \left(l_j \sum_{k=1}^{s_j} E_{jk} w_k \right) \\
&= l_1(F_1 w_1 + F_2 w_2) + l_2(F_3) \\
&= 0.667(0.4932)(0.933) + (0.1941)(0.067)) + 0.333(0.4931) \\
&= 0.4798
\end{aligned}
$$

Meanwhile, in the layer-weighting experiment, the aforementioned three layer-weighting strategies, *decreasing weight*, *increasing weight* and *equal weight* as described in Section 5.3.2 were investigated. All these strategies were used as the layer-weight l_j in the H framework presented in (5.2) — note that only the *mean* aggregation strategy was used in conjunction with these layer-weighting strategies. For instance, for the case of *increasing weight*, the values of layer-weight can be computed using (5.5). For the Parallel models which consists of two layers, the values of layer weights at each layer are $l_1 = 0.333$ and $l_2 = 0.667$. Meanwhile, for the Serial models which consists of three layers, the values of layer weights at each layer are $l_1 = 0.167$, $l_2 = 0.333$ and $l_3 = 0.500$. Assuming that three Fuzzy index values for each sub-system in System F are $\mathrm{F}_1 = 0.4932$, $\mathrm{F}_2 = 0.1941$ and $\mathrm{F}_3 = 0.4932$, the overall interpretability of system F computed using H with *increasing weight* layer-weighting and *mean* aggregation strategy can be expressed as follows:

$$\begin{aligned} H\text{mean} &= \sum_{j=1}^{q} \left(l_j \sum_{k=1}^{s_j} E_{jk}/s_j \right) \\ &= l_1(\mathrm{F}_1 + \mathrm{F}_2)/2 + l_2(\mathrm{F}_3/1) \\ &= 0.333((0.4932 + 0.0.1941)/2) + 0.667(0.4932/1) \\ &= 0.4434 \end{aligned}$$

Table 5.8 Interpretability of the Rotary Crane Systems using H with different aggregation strategies and layer weighting strategies

Rotary Crane Systems	H (using *decreasing weight* layer strategy)					H (using H_{mean} aggregation)		
	H_{mean}	\Hmin	$H_{\max}$	$H_{\text{owa}\alpha=0.1}$	\HowaD	*decreasing weight*	*increasing weight*	*equal weight*
System A	0.0216	0.0216	0.0216	0.0216	0.0216	0.0216	0.0216	0.0216
System B	0.2075	0.2075	0.2075	0.2075	0.2075	0.2075	0.3504	0.2789
System C	0.3504	0.3504	0.3504	0.3504	0.3504	0.3504	0.2075	0.2789
System D	0.2253	0.2253	0.2253	0.2253	0.2253	0.2252	0.2097	0.2175
System E	0.2097	0.2097	0.2097	0.2097	0.2097	0.2097	0.2252	0.2175
System F	0.3935	0.2938	0.4932	0.4798	0.3437	0.3935	0.4434	0.4184
System G	0.3249	0.2408	0.4091	0.2521	0.3670	0.3249	0.2829	0.3039
System H	0.3249	0.3249	0.3249	0.3249	0.3249	0.3249	0.4091	0.3670
System I	0.4301	0.3670	0.4932	0.3755	0.4617	0.4301	0.4722	0.4511
System J	0.4301	0.3670	0.4932	0.4847	0.3986	0.4301	0.4722	0.4511
System K	0.4679	0.4679	0.4679	0.4679	0.4679	0.4680	0.3922	0.4301
System L	0.4932	0.4932	0.4932	0.4932	0.4932	0.4932	0.4932	0.4932

The results obtained are shown in Table 5.8. From these results, the H index produced a diversity of answers for various systems, aggregations and layer-weighting strategies. These

results were then transformed to obtain the H scores for the 20 pairwise comparisons. For the case of the H_{mean} example in aggregation strategies, the first pairwise comparison is between System A and B. In this example, System B was chosen as it scores higher than System A based on the overall interpretability, indicating that the H index suggests that System B is more interpretable than System A. The complete results of pairwise comparison for the interpretability of the Rotary crane systems obtained from the H with different aggregation and layer weighted strategies can be seen in Tables 5.9 and 5.10 respectively.

Table 5.9 Pairwise comparison for the interpretability of the Rotary crane systems obtained from the H with different aggregation strategies at each HFS layer

Pairwise comparisons	H_{mean}	H_{min}	H_{max}	$H_{\text{owa}\alpha=0.1}$	$H_{\text{owa}\alpha=2}$
PW-1	B	B	B	B	B
PW-2	C	C	C	C	C
PW-3	C	C	C	C	C
PW-4	D	D	D	D	D
PW-5	E	E	E	E	E
PW-6	C	C	C	C	C
PW-7	C	C	C	C	C
PW-8	D	D	D	D	D
PW-9	F	F	F	F	F
PW-10	F	F	F	F	F
PW-11	F	F	F	F	G
PW-12	F	H	F	F	F
PW-13	EQ	H	G	H	G
PW-14	I	I	I	I	I
PW-15	EQ	EQ	EQ	J	I
PW-16	K	K	I	K	K
PW-17	L	L	L	L	L
PW-18	K	K	J	J	K
PW-19	L	L	L	L	L
PW-20	L	L	L	L	L

Table 5.10 Pairwise comparison for the interpretability of the Rotary Crane systems obtained from the H with different layer-weighting strategies

Pairwise comparisons	*decreasing weight*	*increasing weight*	*equal weight*
PW-1	B	B	B
PW-2	C	C	C
PW-3	C	B	EQ
PW-4	D	B	B
PW-5	E	B	B
PW-6	C	D	C
PW-7	C	E	C
PW-8	D	E	EQ
PW-9	F	F	F
PW-10	F	F	F
PW-11	F	F	F
PW-12	F	F	F
PW-13	EQ	H	H
PW-14	I	I	I
PW-15	EQ	EQ	EQ
PW-16	K	I	I
PW-17	L	L	L
PW-18	K	J	J
PW-19	L	L	L
PW-20	L	L	L

5.5.3 Matching H to the Participatory Study

This section was conducted to determine the level of agreement between the interpretability ratings provided by the participatory user study (as in Subsection 5.5.1) and various alternative configurations of the H framework (as shown in Subsection 5.5.2).

Specifically, the agreement scores between the results in Tables 5.9 and 5.10 with those in Table 5.5 were computed. For example, for pairwise comparison PW-1, the participants preferences are A= 20, B= 11 and EQ= 8, as shown in Table 5.5. Accordingly, from the preference indicated by H index in Table 5.9, the H_{mean} prefers B, and consequently the

level agreement score obtained is 11 agreements (B was preferred by 11 users). Full details of the agreement score are provided in Table 5.11.

Table 5.11 The Agreement Score between the preferences given by each of the users (in Table 5.5) and the preference indicated by H framework (using different aggregation and layer-weighted strategies as shown in Tables 5.9 and 5.10)

	H aggregation strategies					H layer-weighted strategies		
	H_{mean}	H_{min}	H_{max}	$H_{\text{owa}\alpha=0.1}$	$H_{\text{owa}\alpha=2}$	*decreasing weight*	*increasing weight*	*equal weight*
PW-1	11	11	11	11	11	11	11	11
PW-2	12	12	12	12	12	12	12	12
PW-3	8	8	8	8	8	8	17	15
PW-4	14	14	14	14	14	14	16	16
PW-5	13	13	13	13	13	13	16	16
PW-6	8	8	8	8	8	8	21	8
PW-7	10	10	10	10	10	10	19	10
PW-8	9	9	9	9	9	9	13	18
PW-9	25	25	25	25	25	25	25	25
PW-10	22	22	22	22	22	22	22	22
PW-11	26	26	26	26	11	26	26	26
PW-12	3	17	3	3	3	3	3	3
PW-13	8	26	6	26	6	8	26	26
PW-14	11	11	11	11	11	11	11	11
PW-15	17	17	17	5	18	17	17	17
PW-16	16	16	17	16	16	16	17	17
PW-17	14	14	14	14	14	14	14	14
PW-18	14	14	16	16	14	14	16	16
PW-19	19	19	19	19	19	19	19	19
PW-20	17	17	17	17	17	17	17	17
Mean	14	**15**	14	14	13	14	**17**	16
SD	6	**6**	6	7	5	6	**6**	6

Table 5.11 presents the level of agreement score between the pairwise comparisons of interpretability from a participatory user study and the alternative configurations of the H, respectively. The last two rows summarise the entire table, providing the mean and standard deviation (SD) for each column. From Table 5.11, the H_{min} aggregation strategy and *increasing weight* layer weight strategy achieve the higher agreement score with means of 15 and 17, respectively. Thus, it can be concluded that most of the answers given by users are closer to the ratings obtained using H with the H_{min} aggregation strategy and *increasing weight* layer weight strategy.

5.6 Discussion

This chapter studied the newly proposed generic H index through a participatory design process consisting of three steps to explore the effects of utilising alternative strategies in two key components, namely the aggregation and layer-weighting strategies. The experiment was carried out to derive the configuration of H using a more complex system, the Rotary crane example.

In the first step, another user study was carried out to assess how people perceive the interpretability of twelve different configurations of the system through 20 pairwise comparisons. Based on the opinions of 40 participants with a range of expertise, a diversity of perception regarding interpretability was found. The results imply that interpretability is very subjective and challenging to understand, as views may vary greatly as to the interpretability of different system topologies. In the second step, alternative configurations of the H with various aggregation and layer-weighting strategies were used to measure interpretability. It can be seen from Table 5.8 that this more complex system produces different interpretability scores for almost all the different configurations of the system.

The final step was to examine the level of agreement in terms of interpretability between the pairwise comparison produced from aggregation and layer-weighting strategies as in Step 2, with the pairwise comparison obtained from participatory user study as in Step 1. The number of agreements between the users' views in Step 1 and H results in Step 2 show that the $H_{\min}$ aggregation strategy and *increasing weight* layer-weighted strategy produced the highest agreement score when compared with the others. While the differences are relatively small, these results *suggest* that the H with configuration $H_{\min}$ aggregation strategy and *increasing weight* layer-weighted strategy is to be preferred, as it produced the highest agreement with the users.

5.7 Summary

In summary, this chapter has contributed a new generic index for the measurement of the interpretability of hierarchical fuzzy systems, namely the H index. This index allows the use of any index for measuring the interpretability of a plain fuzzy system to be combined in any configuration of hierarchical systems with different numbers of subsystems, organised in differing topologies. This chapter has then presented a participatory design process, consisting of three main steps, to determine the selection of the best strategies for combining subsystems into an overall index of interpretability. Based on the current evidence, the use of the *min* operator to aggregate subsystems within a layer, together with the weighted mean operator using a *increasing weight* strategy to combine layers, are suggested within the generic H index for capturing the interpretability of HFSs.

Finally, the H index can also contribute to addressing the interpretability of HFSs, and thus can help to build a design guidelines framework for interpretable HFSs. This will be discussed further in Chapter 7.

Chapter 6

An Improved Complexity Measure in Hierarchical Fuzzy Systems

Interpretability is an important and necessary topic that needs to be discussed in relation to the fields of Artificial Intelligence and Machine Learning. Within FLSs, HFSs have been suggested as a key component to help improve the interpretability of FLSs. Indeed, as discussed in Chapter 2, one of the advantages in HFSs is to cope with the complexity which arises in real-world FLSs. In FLSs, the complexity is commonly measured using in a rule-based complexity, considering the number of rules, variables, and fuzzy terms. Several studies have used indicators (for example, the number of rules) to measure the complexity of FLSs. However, this is not a perfect way of assessing complexity in HFSs that have the structure of multiple subsystems, layers and different topologies. Thus far, complexity assessment associated with the structure of HFSs has not been discussed. This chapter aims to put forward a new approach in assessing the complexity of HFSs which will combine rule-based complexity and structural complexity. A detailed measurement of complexity for different HFSs' topologies, namely parallel and serial, will be presented to showcase the features of the new approach. This chapter contributes to the current debate with the introduction of a combined rule-based and structural complexities-based approach in order to establish a comprehensive measurement of complexity in HFSs.

6.1 Introduction

HFSs have been shown to be an effective way to reduce the number of rules in FLSs, which in turn reduces the model complexity and improves its interpretability. Indeed, in literature, it has been claimed that reducing the complexity is a way of improving the model

interpretability in FLSs [22, 29, 30]. Thus, complexity appears to be a key component in the interpretability of FLSs. However, to fully understand the concept of interpretability and complexity in HFSs is a challenge and still remains uncertain.

In Chapter 3, as an introductory step to understanding the interpretability of HFSs, an initial index has been proposed as a way to measure their interpretability; this considered the structure of multiple subsystems, layers and different topologies. However, since interpretability is of a subjective nature, it is not clear as to how different people may perceive interpretability in HFSs. Hence, in Chapter 4, the user perception of interpretability and complexity in HFSs has been investigated with the aid of a user study. It has been found that users had different perceptions regarding interpretability and complexity. For instance, the study found that the rules and structure affected the users' perceptions of interpretability and complexity. Although the parallel and serial HFS in the Iris classification example have the same number of rules (see Section 4.2.2), users felt the structure of the serial system (which has more layers) is more complex to design than the parallel system. Therefore, this chapter will focus on understanding the complexity of HFSs, and to identify the requirement of measuring the complexity in HFSs.

Despite the fact that people have usually measured complexity in FLSs by looking principally at the rule-based complexity level [22, 27, 104, 180], but this is only one quality that can be examined especially for HFSs that have a structure of multiple subsystems, layers and different topologies. This chapter proposes a more comprehensive way of assessing the complexity in HFSs. Here, a new approach to complexity measurement in HFSs is set out, with a specific focus on the combination of rule-based complexity and structural complexity.

The rest of this chapter is organised as follows. Section 6.3 introduces a new approach to complexity measurement in HFSs that considers both rule-based complexity and structural complexity. Section 6.4 demonstrates the new approach of complexity measurement with a real-world example, specifically seesaw control systems. Section 6.5 presents results and discussion. Finally, a summary of this chapter is presented in Section 6.6.

6.2 Motivation

In Chapter 4, the complexity has been found as a key component in determining interpretability of both FLSs and specifically HFSs. The users perceived that the complexity of HFSs is not only dependent on the number of rules, but is also affected by its structure that has multiple subsystems, layers and different topologies. For example, the study found that the number of layers in HFSs affects users' perception of the complexity of HFSs. However, in

FLSs, complexity is usually evaluated by a simple measure, often looking at the rule-based complexity such as the number of input variables, number of rules and fuzzy terms. It is still unclear on how can we measure the complexity of an HFS in a way that takes into account its structure. Thus, the motivation of this chapter is to address these issues by introducing an approach to improve the complexity measurement of HFSs.

6.3 A Comprehensive Measurement of Complexity in Hierarchical Fuzzy Systems

A comprehensive measurement of complexity is needed to evaluate complexity in HFSs that have multiple subsystems, layers and different topologies. A new approach to complexity measurement referred to as C_{HFS} (for HFS *complexity*) is proposed that combines rule-based complexity and structural complexity. The C_{HFS} process for measuring complexity in HFSs is illustrated in Fig. 6.1 and explained in the following steps.

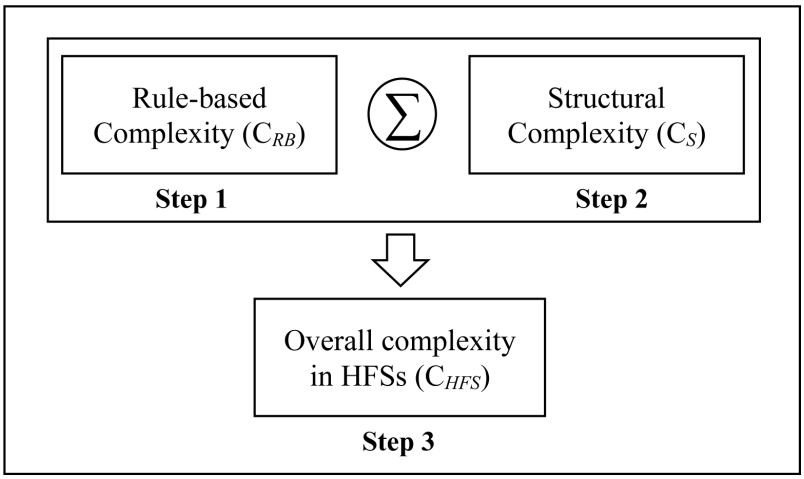


Fig. 6.1 Complexity measurement in HFSs, as explained in detail in Section 6.3.3.2.

6.3.1 Step 1: Calculating the Rule-based Complexity of HFSs

Although there are several ways of measuring rule-based complexity in FLSs, it is challenging to measure the complexity of HFSs that have multiple subsystems. It should be noted that each subsystem will have a small number of inputs and outputs, and a small rule base that will serve a single purpose [98].

For an initial solution, in this step, equation (2.12) as introduced in Chapter 2 is used to calculate the rule-based complexity of each subsystem in a given HFS. However, another challenge is to aggregate all the rule-based complexity values at each subsystem across and within layers, in order to obtain the overall rule-based complexity of HFSs. For this challenge, an aggregation strategy namely a layer-weighted average strategy has been proposed in Chapter 3 which considers all subsystems across and within layers. Therefore, in this step, the layer-weighted average strategy is used to obtain the final rule-based complexity of a given HFS. This strategy now becomes as follows;

$$\mathrm{C}_{RB} = \sum_{j=1}^{q} \left(l_j \sum_{k=1}^{s_j} C_{jk}/s_j \right), \tag{6.1}$$

where:

- C_{jk} is the complexity value associated with the subsystem k at layer j, for example as shown in (2.12);
- l_j is the weight associated with layer j of the HFS;
- s_j is the number of subsystems located in layer j;
- s is the total number of subsystems;
- q is the number of layers of the HFS;
- C_{RB} is the rule-based complexity of the HFS.

An HFS model is less complex when the C_{RB} is close to 0 and more complex when C_{RB} is close to 1.

6.3.2 Step 2: Calculating the Structural Complexity of HFSs

As discussed in Chapter 2, structural complexity in the case of HFSs is related to the interaction amongst its elements; that is, the multiple subsystems, layers and different topologies. However, to evaluate structural complexity in HFSs is a challenging task, particularly when quantifying the interaction between the HFSs' elements. So far, however, assessment on the structural complexity in HFSs has not been investigated.

Therefore, in this step, as an initial solution, a well-established measure, namely the McCabe measure [145] (as discussed in Chapter 2) is used to measure the complexity in the structure of HFSs. The McCabe measure is chosen based on the following reasons.

Firstly, the control graph theory used in McCabe has a similar structure to fuzzy cognitive maps (FCMs). FCMs are fuzzy-graph structures in FLSs that were used for representing causal reasoning [181]. However, none of the indices or measures have been proposed in measuring FCMs' structure of complexity. Secondly, the McCabe measure takes into account the structural complexity in their measures. Thirdly, the McCabe measure can also be applied to assess the complexity of a collection of control graphs known as a hierarchical nest (see Section 2.5.3.1). Finally, the decomposition in the McCabe measure is very similar to the concept with the structure of HFSs that may feature an arbitrary number of multiple subsystems.

In order to transfer McCabe's approach into the setting of HFSs, an approach of mapping the McCabe measure to the HFSs' design is introduced and explained in the following subsections.

(i) Identify elements in HFSs' Design

First, all the elements which are contained in HFSs' design are denoted as follows: (i) x indicates the number of input variables in HFSs, (ii) l indicates the number of layers in HFSs, (iii) a indicates the number of input-output (IO) connections in HFSs, and (iv) s indicates the number of subsystems in HFSs.

(ii) Mapping Edges (e) to the HFSs' Design

The first element of the McCabe measure is e, that is the number of edges in the control graph that need to be mapped to the HFSs' design. The mapping possibilities may be viewed as follows:

$$e \rightarrow (x \quad \text{or} \quad l \quad \text{or} \quad a \quad \text{or} \quad s).$$

One may say that the element of a (number of the IO connections in an HFS) could be the best option for mapping to the element e (number of edges). However, one might also argue that the element l (number of layers in HFSs' design) could be a better mapping to e. Thus, delaying the best choice, two alternative mappings are presented which can now be viewed as follows:

$$e \rightarrow a \quad \text{OR} \quad e \rightarrow l.$$

(iii) Mapping Nodes (n) to the HFSs' Design

The next element of the McCabe measure is n, that is the number of nodes in the control graph that will be mapped. Taking into account the mapping in (ii), there are two

alternative mapping possibilities:

$$\text{EITHER} \quad n \rightarrow (x \text{ or } l \text{ or } s) \quad \text{OR} \quad n \rightarrow (x \text{ or } a \text{ or } s).$$

For the first alternative, it seems most reasonable to choose s (number of subsystems) in HFSs' design to map to n (number of nodes) in the McCabe measure. For the second alternative, x (number of inputs) is chosen to map to n. Consequently, both alternative mappings can now be represented as follows:

$$n \rightarrow s \quad \text{OR} \quad n \rightarrow x.$$

(iv) Mapping Components (p) to the HFSs' Structure

The final element of the McCabe measure is p, that is the number of connected components in the control graph. Now, given the assignments in (ii) and (iii), the remaining possibilities are:

$$\text{EITHER} \quad p \rightarrow (x \text{ or } l) \quad \text{OR} \quad p \rightarrow (a \text{ or } s).$$

The number of connected components in the control graph (p) is a very similar concept as the number of layers (l) in HFSs. In HFSs, all subsystems are placed and attached through the HFS layers. However, one may argue that the number of connected components in the control graph can also be related to the number of subsystems (s) in HFSs. Therefore, for the first alternative, l is mapped to p; whilst in the second alternative, s is mapped to p. Hence, both alternative mappings now are presented as follows:

$$p \rightarrow l \quad \text{OR} \quad p \rightarrow s.$$

(v) Construct complexity measure of HFSs – (C_S)

Following the mapping selection in (ii), (iii), and (iv), two alternative mappings are obtained, namely the first alternative (from now on referred to as A), which constitutes the mapping $e \rightarrow a$, $n \rightarrow s$ and $p \rightarrow l$, whereas the second alternative (termed B) which constitutes $e \rightarrow l$, $n \rightarrow x$ and $p \rightarrow s$. Finally, a structural complexity measure for HFSs is assembled using both alternative mappings, and now (2.14) appears as follows:

$$\mathrm{C}_S\mathrm{A}\Big(\bigcup_{i=1}^{s}\Big) = a - s + 2l \tag{6.2}$$

$$C_S B\Big(\bigcup_{i=1}^{s}\Big) = l - x + 2s \tag{6.3}$$

where $C_S A$ and $C_S B$ indicate the non-normalised structural complexity in HFSs using alternative mappings A and B respectively.

6.3.3 Step 3: Overall Measurement of the Complexity of HFSs

The final challenge is to combine the rule-based complexity and structural complexity calculation from Step 1 and Step 2 respectively in order to obtain the overall complexity of the HFS. However, to combine both rule-based complexity and structural complexity is not a straightforward task. In this case, two important questions (i) "What is a way to make both rule-based complexity and structural complexity comparable in that both should have the same scale of output?" (ii) "What is the best operator to combine rule-based complexity and structural complexity?" are examined. Therefore, as a result of this step, a method to address the challenges has been proposed and explained in the following subsection.

6.3.3.1 Normalise the structural complexity values to be in the same range as rule-based complexity values

McCabe's measure ranges from $[1,\infty]$ whereas $C_S A$ and $C_S B$ range from $[1,\infty]$ and $[-\infty,\infty]$ respectively. Whilst there is no inherent problem with this, other measures of interpretability for FLSs (such a Nauck's index [21] or Alonso et al. [24]) have traditionally been defined over $[0,1]$. Given this, it will be easier to subsequently combine a HFS complexity measure with other components of interpretability if it lies over the same range. Thus, in this chapter, the complexity values of $C_S A$ and $C_S B$ will be normalised to the range $[0,1]$. There are several alternative functions that can be used but for this chapter, one of the principal functions in mathematics, which is an inverse trigonometric function is chosen. This function can be expressed as follows:

$$\tilde{C}_S = \frac{4\arctan(C_S) - \pi}{2\pi} \tag{6.4}$$

For the case of the $C_S B$ that has a different range to $C_S A$, the function is slightly modified as follows:

$$\tilde{C}_S = \frac{2\arctan(C_S) + \pi}{2\pi} \tag{6.5}$$

where $\arctan()$ is the inverse of the tangent function, C_S represents the values of $C_S A$ or $C_S B$, and $\tilde{C}_S$ indicates the normalised structural complexity in HFSs. An HFS model is less

complex when the $\tilde{C}_S$ is close to 0 and more complex when $\tilde{C}_S$ is close to 1, in terms of structural complexity.

6.3.3.2 Combining the structural and rule-based complexity using a generic operator

In order to combine structural complexity and rule-based complexity, it may be sufficient to use an aggregation operator that selects something between min and max values. Alternatively, it also could be something else such as a t-conorm operator. Hence, in the second challenge, several options to combine the aforementioned rule-based complexity and structural complexity are employed, which is shown below.

$$C_{HFS} = C_{RB} \textcircled{\Sigma} \tilde{C}_S \tag{6.6}$$

where $\textcircled{\Sigma}$ indicates the generic operator. In this case, the generic operator will represent basic aggregation operators namely *min*, *max* and *mean*, and also another operator strategy that is a t-conorm operator, specifically one of *probabilistic sum*, *bounded sum* and *einstein sum*. These aggregation and t-conorm operators can be demonstrated as follows:

(i). *Aggregation operator: min*

$$C_{HFS} = \text{MIN}(C_{RB}, \tilde{C}_S)$$

(ii). *Aggregation operator: max*

$$C_{HFS} = \text{MAX}(C_{RB}, \tilde{C}_S)$$

(iii). *Aggregation operator: mean*

$$C_{HFS} = \text{MEAN}(C_{RB}, \tilde{C}_S)$$

(iv). *T-conorm operator: probabilistic sum*

$$C_{HFS} = C_{RB} + \tilde{C}_S - C_{RB} \cdot \tilde{C}_S$$

(v). *T-conorm operator: bounded sum*

$$C_{HFS} = \text{MIN}(C_{RB} + \tilde{C}_S, 1)$$

(vi). *T-conorm operator: einstein sum*

$$C_{HFS} = \frac{C_{RB} + \tilde{C}_S}{1 + C_{RB} \cdot \tilde{C}_S}$$

An HFS model is expected to be less complex when the C_{HFS} is close to 0 and expected to be more complex when the C_{HFS} is close to 1.

6.4 Experimental Results

In this experiment, the examples of the seesaw control application (as used in Chapter 3) and the iris classification problem (as used in Chapters 4 and 5) were used to demonstrate the features of proposed complexity measure.

6.4.1 Seesaw Control Problem

As defined earlier (see Section 3.5), the involved parameters of the seesaw are the distance of the cart ($x = x_1$), the angle that the wedge makes with the vertical line ($\theta = x_2$), the height of the wedge ($r_1 = x_3$), and the centre of mass of the wedge ($r_2 = x_4$), as shown in Fig. 3.1. The topology and complete rules set for this FLS (81 rules) can be seen in Fig. 6.2 and Table 6.1, respectively.

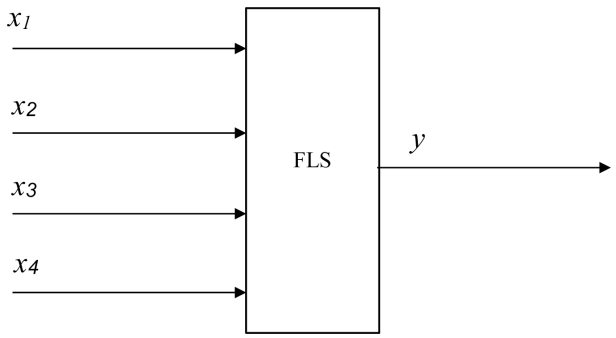


Fig. 6.2 Seesaw control Application: FLS

Table 6.1 The rules of flat FLS

x_3	x_4	x_1								
		nb			ze			pb		
		x_2			x_2			x_2		
		nb	ze	pb	nb	ze	pb	nb	ze	pb
	nb	nb	nb	nm	nb	nm	ns	nm	ns	ze
nb	ze	nb	nm	ns	nm	ns	ze	ns	ze	ps
	pb	nm	ns	ze	ns	ze	ps	ze	ps	pm
	nb	nb	nm	ns	nm	ns	ze	ns	ze	ps
ze	ze	nm	ns	ze	ns	ze	ps	ze	ps	pm
	pb	ns	ze	ps	ze	ps	pm	ps	pm	pb
	nb	nm	ns	ze	ns	ze	ps	ze	ps	pm
pb	ze	ns	ze	ps	ze	ps	pm	ps	pm	pb
	pb	ze	ps	pm	ps	pm	pb	pm	pb	pb

This experiment aims to explore the features of the proposed complexity measurement by using a seesaw control as an example. The experiment process consists of three main steps: (i) produce a Parallel L-HFS and a Serial L-HFS for the Seesaw Control; (ii) examine the equivalence between the Parallel L-HFS and the Serial L-HFS; and (iii) measure the complexity of the Parallel L-HFS and the Serial L-HFS by using the proposed measure.

6.4.1.1 Produce a Parallel L-HFS and a Serial L-HFS for the Seesaw Control

First, this experiment begins by producing two types of HFSs, namely parallel and serial. In contrast to Chapter 3, in this chapter, the L-HFS algorithm is used to decompose the FLS from the seesaw system to reproduce two types of HFS models, also known as Parallel L-HFS and Serial L-HFS. However, they utilise similar topologies for parallel and serial L-HFSs, as in Chapter 3 (as in Figs. 3.4 and 3.5 respectively). The complete rules for these HFSs are shown in Tables 6.3–6.8. A summary of both HFSs is presented in Table 6.2.

Table 6.2 Description of the FLS, Parallel L-HFS and Serial L-HFS for Seesaw Control Applications

Descriptions	FLS	Parallel L-HFS	Serial L-HFS
Number of inputs	4	4	4
Number of layers	1	2	3
Number of subsystems	1	3	3
Number of rules in FLS_1	81	9	9
Number of rules in FLS_2	-	9	15
Number of rules in FLS_3	-	25	21
Total number of rules	81	43	45

As can be seen in Table 6.2, the decomposition process from FLS to HFSs (Parallel L-HFS and Serial L-HFS) has effectively reduced the number of rules by almost 50% from the original FLS rules. Intuitively, one may say that both the Parallel L-HFS and Serial L-HFS have reduced approximately 50% of the complexity from the original FLS if only considering the total number of rules as an indicator to measure HFS complexity. However, the uncertainty of the structure of HFSs, namely due to multiple subsystems, layers, and different topologies, may also affect the complexity of HFSs. Therefore, in the following step, the proposed measure, C_{HFS}, is used to measure complexity for both the Parallel L-HFS and Serial L-HFS.

Table 6.3 The rules of Parallel L-HFS: FLS_1

x_2	x_1		
	nb	ze	pb
nb	A	B	C
ze	B	C	D
pb	C	D	E

Table 6.4 The rules of Parallel L-HFS: FLS_2

x_4	x_3		
	nb	ze	pb
nb	F	G	H
ze	G	H	I
pb	H	I	J

Table 6.5 The rules of Parallel L-HFS: FLS_3. y_1 and y_2 are the intermediate outputs generated from subsystem FLS_1 and FLS_2, respectively.

y_2	y_1				
	A	B	C	D	E
F	nb	nb	nm	ns	ze
G	nb	nm	ns	ze	ps
H	nm	ns	ze	ps	pm
I	ns	ze	ps	pm	pb
J	ze	ps	pm	pb	pb

Table 6.6 The rules of Serial L-HFS: FLS_1

x_2	x_1		
	nb	ze	pb
nb	A	B	C
ze	B	C	D
pb	C	D	E

Table 6.7 The rules of Serial L-HFS: FLS_2

x_3	y_1				
	A	B	C	D	E
nb	F	G	H	I	J
ze	G	H	I	J	K
pb	H	I	J	K	L

Table 6.8 The rules of Serial L-HFS: FLS_3

x_4	y_2						
	F	G	H	I	J	K	L
nb	nb	nb	nm	ns	ze	ps	pm
ze	nb	nm	ns	ze	ps	pm	pb
pb	nm	ns	ze	ps	pm	pb	pb

6.4.1.2 Examine the Equivalence between the Parallel L-HFS and the Serial L-HFS

Before the complexity of Parallel L-HFS and Serial L-HFS are measured, the systems must be shown to be equivalent. For example, if the result showed that they were not equivalent, there would be no reason for conducting a detailed measurement and comparison regarding their complexity because it would be measuring two different systems.

Therefore, in this step, a statistical test was conducted using the Kolmogorov-Smirnov (K-S) test [182] to compare the equivalence of the Parallel L-HFS and the Serial L-HFS. The K-S test statistic measures the maximum distance (D) between the empirical cumulative distribution function (ECDF) of the Parallel L-HFS and the Serial L-HFS, measured in a vertical direction. It is computed as:

$$D = \sup_x \left| F_{P-LHFS}(x) - F_{S-LHFS}(x) \right| \tag{6.7}$$

where, $F_{P-LHFS}(x)$ indicates cumulative distribution of output mapping of Parallel L-HFS, $F_{S-LHFS}(x)$ indicates cumulative distribution of output mapping of Serial L-HFS.

The D statistic is then compared with the critical value (D_α) to find the equivalence between the two HFSs namely Parallel L-HFS and Serial L-HFS. If the D is greater than

(D_α), then both output mappings are distributed differently. However, if the D is less than (D_α), then both distributions are equal. The critical values can be computed as follows [182]:

$$D_\alpha = \frac{c(\alpha)}{\sqrt{n}} \tag{6.8}$$

where, n is sample size, α is level of significance that will indicate value of $c(\alpha)$ and both values can be referred at K-S Test Critical values in [182]. In this case, the value of $\alpha = 0.05$ and $c(\alpha) = 1.36$ are applied.

Table 6.9 The Kolmogorov-Smirnov Test between FLS, Parallel L-HFS and Serial L-HFS of Seesaw Control Application

Input samples (n)	FLS vs Parallel L-HFS		FLS vs Serial L-HFS		Parallel L-HFS vs Serial L-HFS		Critical Value
	D	p-value	D	p-value	D	p-value	$D_\alpha = \frac{c(\alpha)}{\sqrt{n}}$
500	0.042	0.7699	0.040	0.8186	0.048	0.6121	0.061

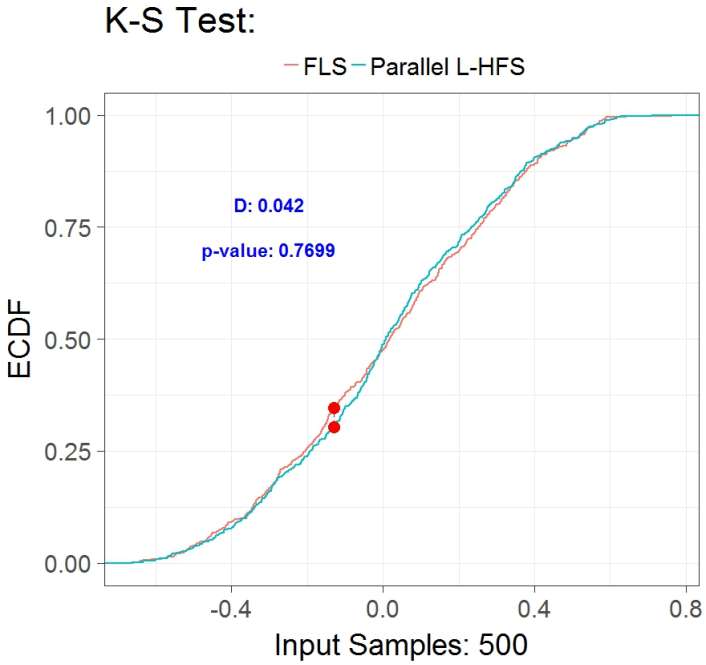


Fig. 6.3 K-S Test: FLS vs Parallel L-HFS

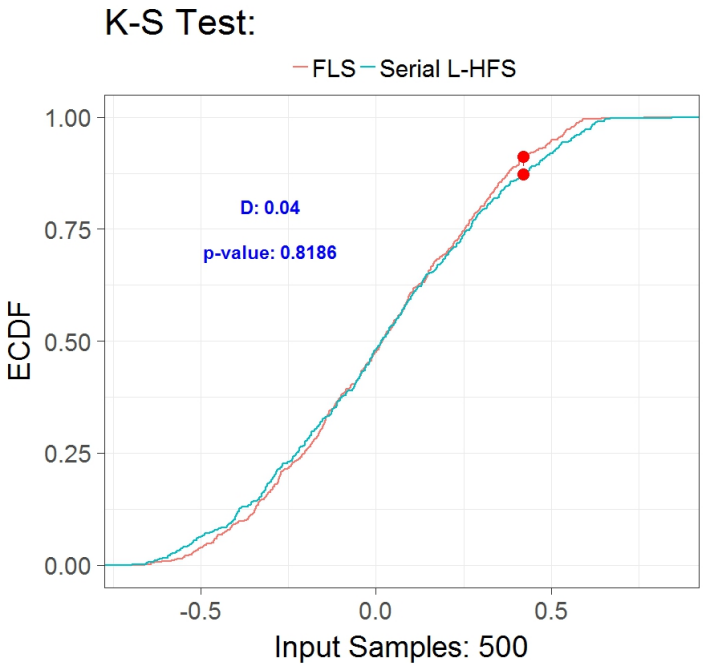


Fig. 6.4 K-S Test: FLS vs Serial L-HFS

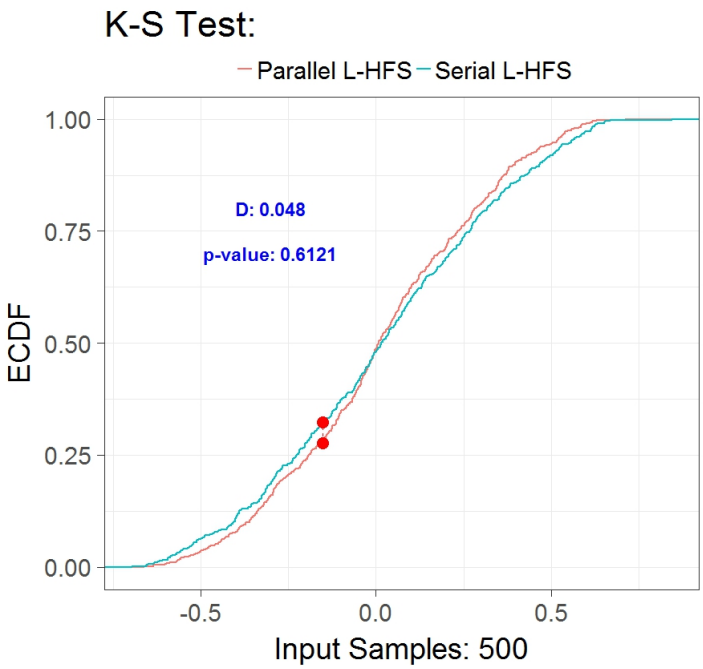


Fig. 6.5 K-S Test: Parallel L-HFS vs Serial L-HFS

Table 6.9 presents the result of the K-S test that examined the equivalence between the FLS, Parallel L-HFS and Serial L-HFS of the seesaw system. The result can also be viewed graphically in Figs. 6.3, 6.4 and 6.5. From Table 6.9 and Fig. 6.5, the K-S test showed that both HFS seesaw systems – namely, Parallel L-HFS and Serial L-HFS were equivalent when tested with the 500 input samples. The K-S test statistic, which measured the maximum distance (D) between two HFS seesaw systems, is less than (D_α), indicating they have the same distribution. Additionally, the results revealed that both HFSs were also equivalent to FLS when tested with 500 samples, as shown in Figs. 6.3 and 6.4.

6.4.1.3 Measuring Complexity of the Parallel and Serial L-HFS

Finally, in this step, the proposed measure, C_{HFS} is utilised to measure the complexity in the Parallel L-HFS and the Serial L-HFS. The measurement is described in detail in the following subsections.

Step 1 *Calculating the rule-based complexity of the Parallel and Serial L-HFS*:

To calculate the rule-based complexity, equation (6.1) is applied to measure the complexity in Parallel L-HFS and Serial L-HFS. Table 6.10 shows a summary of the two systems' rule-based complexity.

Table 6.10 The Rule-based Complexity between the FLS, Parallel L-HFS and the Serial L-HFS of Seesaw Control Application

Seesaw Systems	Rule-based Complexity			
	NC	NP	Comp	C (1- Comp)
FLS	7	324	0.022	0.978
Parallel L-HFS				
- FLS_1	5	18	0.278	0.722
- FLS_2	5	18	0.278	0.722
- FLS_3	7	50	0.140	0.860
C_{RB}				0.768
Serial L-HFS				
- FLS_1	5	18	0.278	0.722
- FLS_2	7	30	0.233	0.767
- FLS_3	7	42	0.167	0.833
C_{RB}				0.756

As seen in Table 6.10, the computed rule-based complexity, C_{RB}, produces a higher complexity value of 0.768 for the Parallel L-HFS, compared to 0.756 for the Serial

L-HFS. From this it may be concluded that the Parallel L-HFS is more complex than the Serial L-HFS in terms of rule-based complexity.

Step 2 *Calculating the structural complexity of the Parallel and Serial L-HFS*:

For the structural complexity, only alternative mapping B as in (6.3) is used to measure the structural complexity, $C_S B$, in the two HFSs. This is due to the fact that $C_S B$ incorporates inputs in the range $[-\infty, \infty]$. Table 6.11 presents an overview of the summary of this structural complexity, $C_S B$, measured for the Parallel L-HFS and Serial L-HFS. The equation in (6.5) is then used to normalise the computed structural complexity of Parallel L-HFS and Serial L-HFS to be within the range [0,1], as shown in the last column of Table 6.11.

Table 6.11 The Structural Complexity between the FLS, Parallel L-HFS and the Serial L-HFS of Seesaw Control Application

Seesaw Systems	Structural Complexity				
	l	n	s	$C_S B$	$\tilde{C}_S$
FLS	1	4	1	-2	0.250
Parallel L-HFS	2	4	3	4	0.844
Serial L-HFS	3	4	3	5	0.874

Table 6.11 shows the pattern of computed structural complexity that produces a higher value for the Serial L-HFS ($C_S B = 5$) in comparison to the Parallel L-HFS ($C_S B = 4$). From this it may be concluded that the Serial L-HFS is more complex than the Parallel L-HFS in terms of structural complexity.

Step 3 *Overall measurement of the complexity of the Parallel and Serial L-HFS*:

The structural complexity has been normalised to be the same scale as the rule-based complexity in the previous step as shown in Table 6.11. Therefore, both rule-based complexity and structural complexity for Parallel L-HFS and Serial L-HFS are comparable. Next, to obtain the overall complexity of Parallel L-HFS and Serial L-HFS, the proposed C_{HFS} as shown in (6.6) is used to combine the rule-based complexity and structural complexity by using several combination options, namely aggregation and t-conorm operators. Table 6.12 provides a summary of the overall measurement complexity in FLS, Parallel L-HFS and Serial L-HFS using various combination strategies.

In addition, the results of the overall measurement complexity in FLS, Parallel L-HFS and Serial L-HFS can also be viewed graphically in Fig. 6.6.

Table 6.12 Summary of overall measurement complexity for FLS, Parallel L-HFS and Serial L-HFS with various alternatives combination strategies

Alternative combination strategies	FLS			Parallel L-HFS			Serial L-HFS		
	C_{RB}	$\tilde{C}_S$	C_{HFS}	C_{RB}	$\tilde{C}_S$	C_{HFS}	C_{RB}	$\tilde{C}_S$	C_{HFS}
Aggregation - min	0.978	0.250	0.250	0.768	0.844	0.768	0.756	0.874	0.756
Aggregation - max	0.978	0.250	0.978	0.768	0.844	0.844	0.756	0.874	0.874
Aggregation - mean	0.978	0.250	0.614	0.768	0.844	0.806	0.756	0.874	0.815
T-conorm - probabilistic sum	0.978	0.250	0.984	0.768	0.844	0.964	0.756	0.874	0.969
T-conorm - bounded sum	0.978	0.250	1.000	0.768	0.844	1.000	0.756	0.874	1.000
T-conorm - einstein sum	0.978	0.250	0.987	0.768	0.844	0.978	0.756	0.874	0.981

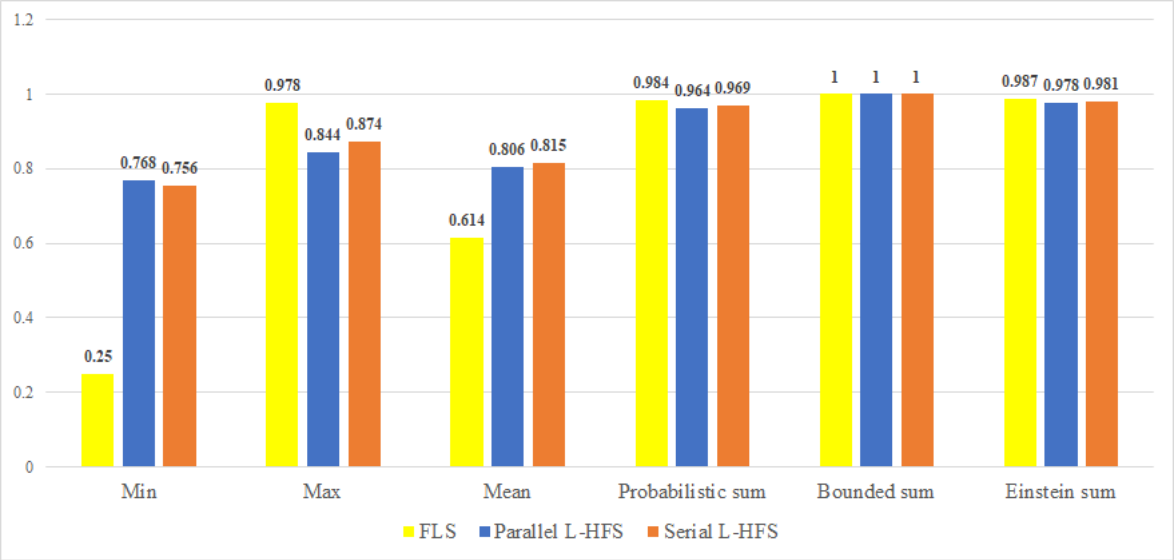


Fig. 6.6 Summary of the overall complexity (C_{HFS}) for FLS, Parallel L-HFS and Serial L-HFS, using different aggregation functions

6.4.2 Iris classification Problem

The Iris flower classification example that has been introduced earlier in Chapters 2, 4 and 5, was used again in this section. Specifically, three Iris classification systems with the 3-MFs

configuration, namely FLS, Parallel HFS and Serial HFS as shown in Figs. 2.9, 2.12 and 2.10 (see Chapter 2), respectively. In contrast to the previous section, this experiment aims to only examine the equivalence between the FLS, Parallel HFS and Serial HFS. This is important because previously the iris classification problem has only been used to focus on interpretability without being concerned at the performance of the various configurations. A similar method as in Section 6.4.1.2, was used to reveal their equivalent agreement, as described in the following subsection.

6.4.2.1 Equivalence Between Systems for Iris Classification

The Kolmogorov-Smirnov statistical test, as in (6.7), was again used to find the level of equivalence between the various system configurations.

Table 6.13 The Kolmogorov-Smirnov Test between FLS, Parallel HFS and Serial HFS of Iris classification

Input samples (n)	FLS vs Parallel HFS		FLS vs Serial HFS		Parallel HFS vs Serial HFS		Critical Value
	D	p-value	D	p-value	D	p-value	$D_\alpha = \frac{c(\alpha)}{\sqrt{n}}$
150	0.1233	0.0209	0.2833	$6.947e^{-11}$	0.3133	$3.233e^{-13}$	0.1333

Table 6.13 presents the results of the K-S test on the various Iris classification systems. The results show that the FLS and Parallel HFS are equivalent, due to the fact that (D) between them is less than (D_α), which indicates that they have the same distribution. However, the results between FLS and Serial, and between Parallel and Serial systems, indicated that these were not equivalent, as shown in Table 6.13.

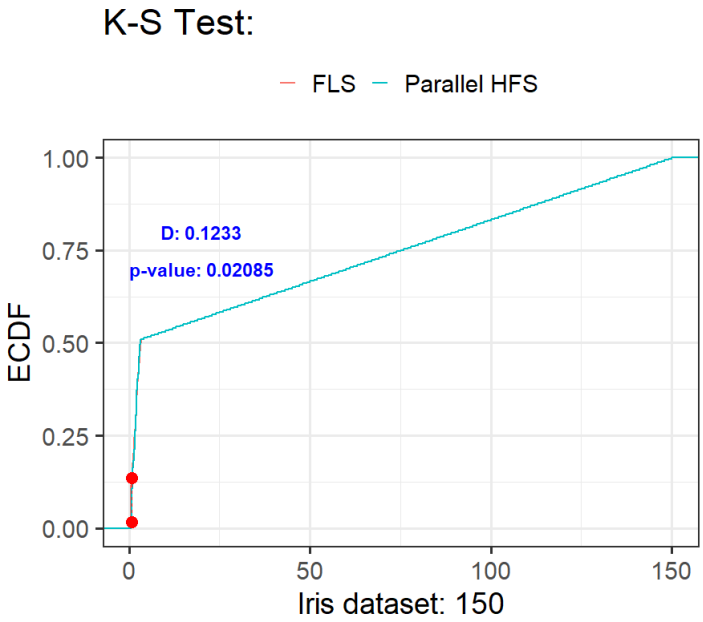


Fig. 6.7 K-S Test: FLS vs Parallel HFS

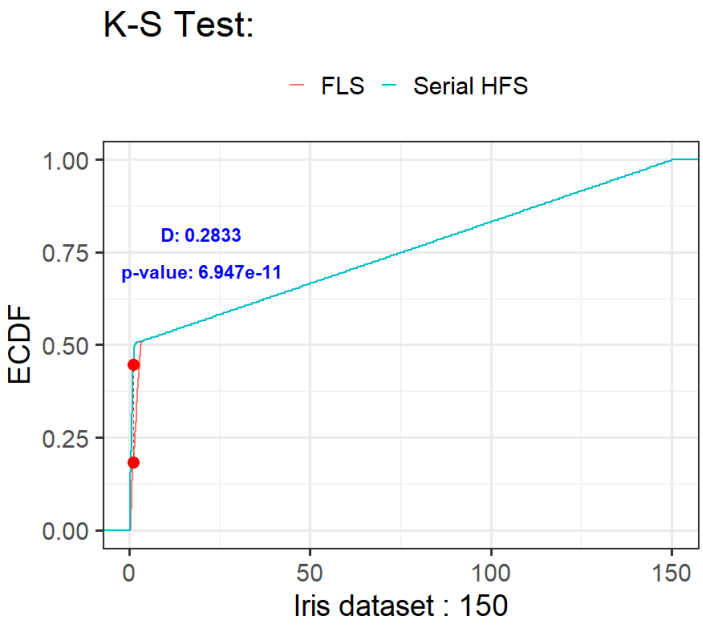


Fig. 6.8 K-S Test: FLS vs Serial HFS

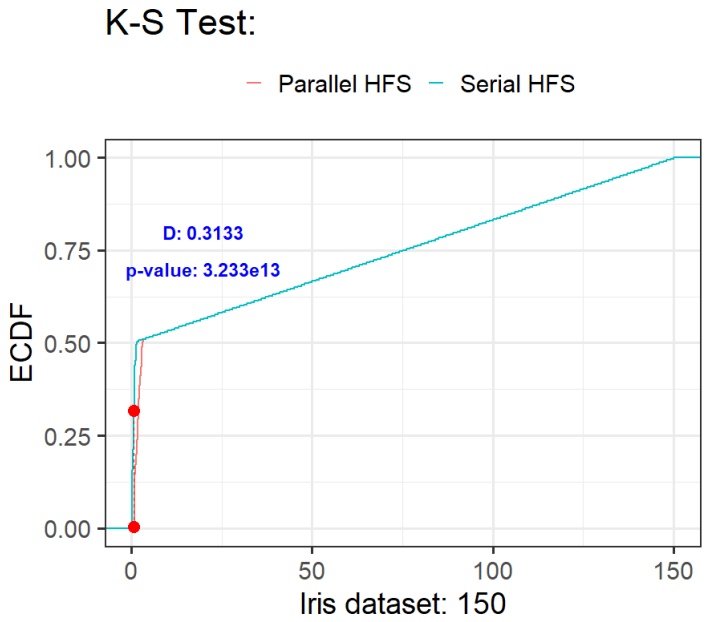


Fig. 6.9 K-S Test: Parallel HFS vs Serial HFS

The result can also be viewed graphically in Figs. 6.7, 6.8 and 6.9. From Fig. 6.7, the K-S test showed that the FLS and Parallel HFS were equivalent. Despite the K-S test revealing that the results between the FLS and the Serial HFS, and the Parallel and Serial HFS, were not equivalent, nevertheless Figs. 6.8 and 6.9 illustrate that they only produce *slightly* different results in terms of the observed sdistribution.

6.5 Results and Discussion

This chapter has presented a new approach to measuring complexity in HFSs, namely the C_{HFS} approach, which focused on combining rule-based complexity and structural complexity. An experiment was conducted and observed the proposed measure using a real-world example; in this case, the Seesaw systems and Iris classification problem.

First, for the case of the Seesaw system, an FLS was decomposed into two HFS models, namely Parallel L-HFS and Serial L-HFS. By doing so, both Parallel L-HFS and Serial L-HFS produced approximately 50% fewer rules than the original FLS of the Seesaw System. Also, the rule-based complexity reveals that the FLS is more complex than the Parallel L-HFS and Serial L-HFS, as shown in Table 6.10. Intuitively, this may assume that the reduction of the number of rules reduces the complexity of both models. However, it is not appropriate

to consider only the number of rules as an indicator when measuring complexity in HFSs, without taking into account its structure, namely the existence of multiple subsystems, layers and different topologies.

Next, both the Parallel L-HFS and the Serial L-HFS together with FLS, were examined to make sure that they were equivalent systems. This was achieved by using a statistical analysis method, namely a Kolmogorov-Smirnov (K-S) test. Although both the Parallel L-HFS and the Serial L-HFS had differed slightly on the total number of rules, the K-S test result proved that they were equivalent systems. The K-S test also reveals that both HFSs were equivalent to FLS.

Finally, measuring the complexity of both the Parallel L-HFS and the Serial L-HFS was implemented using C_{HFS}, consisting of three main steps. For the first step, the rule-based complexity for the FLS, Parallel L-HFS and Serial L-HFS was measured using the proposed C_{RB} as shown in (6.1). In summary, the results show that FLS is *more complex* than both HFSs in term or rule-based complexity, i.e. the number of rules in FLS is higher than both HFSs. Moreover, the findings also reveal that the Parallel L-HFS produced a higher value of C_{RB} than the Serial L-HFS. This would indicate that the Parallel L-HFS was more complex than the Serial L-HFS in terms of rule-based complexity. Interestingly, the total number of rules in the Parallel L-HFS was slightly less than in the Serial L-HFS. This result of C_{RB} gives insight into the complexity of HFSs, nevertheless, it is not appropriate to consider only the number of rules when measuring complexity in HFSs. Next, for the second step, the structural complexity for the FLS, Parallel L-HFS and Serial L-HFS was measured using the proposed C_SB as shown in (6.3). In contrast, for this case, the results showed that the Serial L-HFS produce a higher value of C_SB than the Parallel L-HFS, which would indicate that the Serial L-HFS is more complex than the Parallel L-HFS. This finding was expected and suggests that in HFSs the complexity increases exponentially with the number of HFS layers in terms of the structural complexity. Further, the results confirm the expectation that the structural complexity of FLS is *less complex* than both HFSs as its structure is simpler than the HFSs, possessing just a single layer and subsystem.

For the third step, the overall complexity of the FLS, Parallel L-HFS and Serial L-HFS was calculated using the proposed C_{HFS} as shown in (6.6). This aimed to combine the result of the rule-based complexity and structural complexity calculations from Steps 1 and 2 respectively, using several combination alternatives. The results from this step indicate that the combination from operators *max*, *mean*, *probabilistic sum* and *einstein sum* presented a similar pattern, namely that the overall complexity value of C_{HFS} for the Serial L-HFS was higher than the Parallel L-HFS, indicating that the Serial L-HFS is more complex than

the Parallel L-HFS. However, the operator *min* showed a different pattern, namely that the overall complexity value of C_{HFS} for the Parallel L-HFS is higher than that of the Serial L-HFS, indicating that the Parallel L-HFS is more complex than the Serial L-HFS. However, the operator *bounded sum* produced an equal complexity value for both Parallel L-HFS and Serial L-HFS, which could indicate that they are equally complex. Overall, the results generated for the proposed measure using most operators follow intuition in the sense that one expects the Serial L-HFS to be more complex than the Parallel L-HFS. For the case of FLS, both aggregation and t-conorm operators produced a variety of complexity results. For instance, the *max, probabilistic sum* and *Einstein sum* operators all show the FLS is more complex than both HFSs. In contrast, the *min* and *mean* operators indicated that the FLS is less complex than both HFSs.

There are several possible explanations for this result, of which the most likely are as follows: (i) the overall complexity of HFSs is more affected by the structural complexity than rule-based complexity; (ii) complexity in HFSs intuitively increases exponentially with the number of layers; and (iii) rule-based complexity in HFSs does not entirely depend on the total number of rules, but it also depends on the aggregation process of individual rule-based complexities in each HFS subsystem.

Second, for the case of Iris classification problem, the three Iris systems were evaluated and examined to determine the equivalence between them. Likewise to the previous experiment, a statistical K-S test was used to reveal their equality. The results showed that there was an equivalent between the FLS and Parallel HFS. This indicates that, despite the Parallel HFS having fewer (27) rules, it still offers the same performance as the FLS having 81 rules. However, the results also suggested that the FLS was *not* equivalent to the Serial HFS; and that the HFSs (Parallel and Serial) were *not* equivalent. Nevertheless, Figs. 6.8 and 6.9 show that the K-S test showed only slight difference in distribution between these systems.

6.6 Summary

In summary, this chapter has contributed to the field a new approach to measuring complexity in HFSs, known as C_{HFS}. The new approach allows the combination of rule-based complexity and structural complexity to be a way to obtain a more comprehensive measurement of complexity, particularly in HFSs.

Although the combination strategies explored in this chapter are not exhaustive and some different approaches may be appropriate, based on the current evidence, the proposed measure (C_{HFS}) shows promise for measuring the complexity in HFSs and it behaves in

an intuitive manner. Also, the proposed measure (C_{HFS}) appears to be a better approach for measuring complexity in HFSs because it considers both rule-based complexity and structural complexity.

Finally, a comprehensive way to measure the complexity in HFSs was obtained. The introduction of this measure is hoped to give insight into a deeper understanding of the complexity in HFSs, and thus also contribute to improving the interpretability of HFSs. Therefore, in Chapter 7, a design guidelines framework towards an interpretable HFSs will be discussed.

Chapter 7

A Framework for Guiding the Design of Interpretable Hierarchical Fuzzy Systems

HFSs are claimed to be an excellent approach to reduce the number of rules in FLSs, as discussed in Chapter 2. Moreover, HFSs have also been shown to have the potential to reduce complexity and improve interpretability for FLSs. However, to design an interpretable HFS is a challenging task. This is due to the HFSs' structures, that have multiple subsystems, layers and different topologies, as explained in Chapters 3, 5, and 6. This chapter uses the proposed measures of interpretability and complexity in Chapters 5 and 6 respectively, to present a design guidelines framework to build interpretable HFSs. Then, the proposed framework is demonstrated using a real-world example, specifically interpretable HFS design for the neonatal intensive care unit (NICU) at the QMC hospital in Nottingham.

7.1 Introduction

As discussed in Chapter 2, Torra [57] claimed that building an HFS is a challenging task because of the need to define the architecture of the HFS (the subsystems, the input variables of each subsystem, and the interactions between subsystem), as well as the rules of each subsystem. Furthermore, a fundamental issue with HFSs is how to handle the abstract meaning of intermediate output variables, which consequently makes HFSs difficult to design [53]. Although several researchers have proposed approaches to handle these issues [48, 46, 49–53], their proposed approaches do not consider the interpretability of HFSs,

particularly semantic interpretability of intermediate variables. Several aspects should be considered when measuring the interpretability of HFSs, such as semantic interpretability, meaningful fuzzy sets and intermediate variables. For example, if a layer in an HFS has a meaning, it can be useful for interpretability. This shows that the interpretability of HFSs is very problematic to understand since it is of a subjective nature. Consequently, it is not easy to design or build an interpretable HFS.

In Chapters 5 and 6 of interpretability and complexity for HFSs has been proposed. This chapter uses that measure to introduce a design guidelines framework to build an interpretable HFS.

7.2 A Design Guidelines Framework for Interpretable HFSs

In this section, a design guidelines framework for building interpretable HFS is proposed. The proposed design guidelines framework consists of five key guidelines that are explained in the following subsections. Note that, the Iris classification example as discussed in previous Chapters 2, 4 and 5 is used in order to illustrate the proposed design guidelines framework.

7.2.1 Guideline 1 – Selection of Topology

The first design guideline to build an interpretable HFS is the selection of HFSs' topologies such as parallel, serial and hybrid. This chapter only discusses parallel and serial topologies, and HFSs using each topology to solve the Iris classification problem can be seen in Figs. 7.1 and 7.2 respectively.

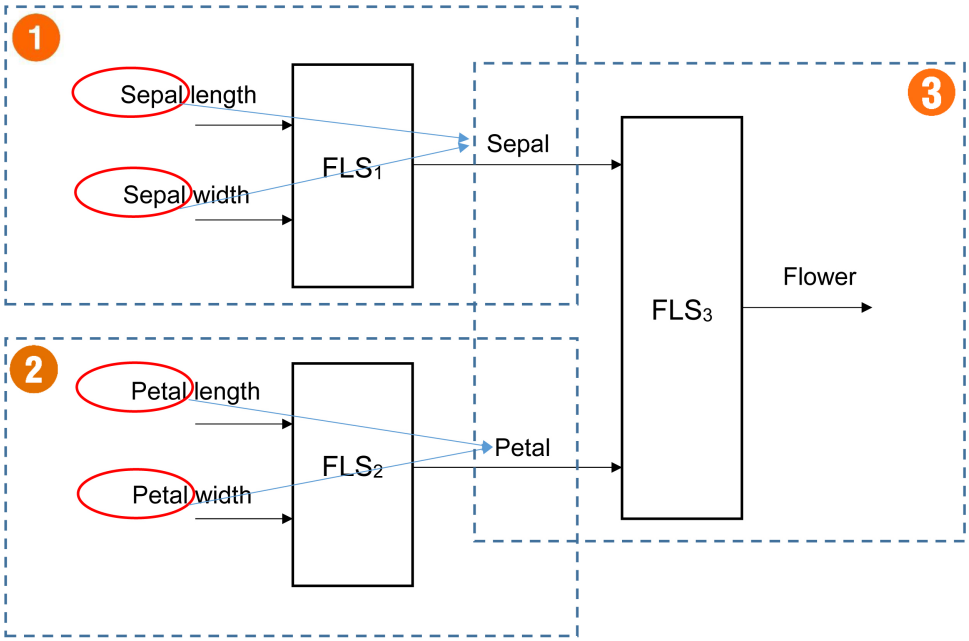


Fig. 7.1 Topology: Parallel HFS

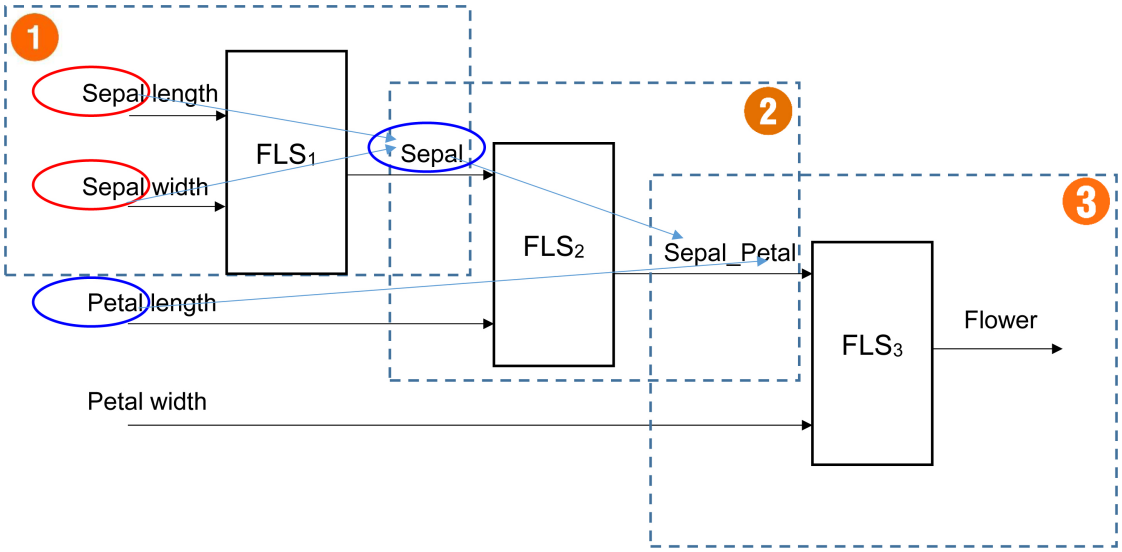


Fig. 7.2 Topology: Serial HFS

According to the measure of interpretability and complexity in Chapters 5 and 6, based on current evidence in their experiments, the parallel topology is more interpretable than the serial topology. Also, the measure found that parallel topology is less complex to design than serial. Moreover, this finding is also equivalent with user perceptions obtained in Chapter 4.

However, this finding cannot be generalised for the other HFS examples, which implies that it is essential to choose the correct HFS topology in order to design an interpretable HFS.

7.2.2 Guideline 2 – Simplicity and Readability of Subsystems

The second guideline is on simplicity and readability in relation to subsystems in HFSs. Several considerations are provided when designing simple and readable subsystems and are defined in the following subsections.

(i) Two inputs and a single output – "2-ISO"

The use of two inputs and a single output for all the subsystems in HFSs is suggested, referred to as 2-ISO. Figs. 7.3 and 7.4 illustrate the examples of 2-ISO and three inputs and a single output (3-ISO) respectively. For the case of the Iris classification examples as in Figs. 7.1 and 7.2, all the subsystems in both topologies are designed using the 2-ISO architecture. In doing this, the subsystems will have fewer rules. As a result, the readability of rules in the subsystems of HFS is improved and will lead to increasing the interpretability of the HFS.

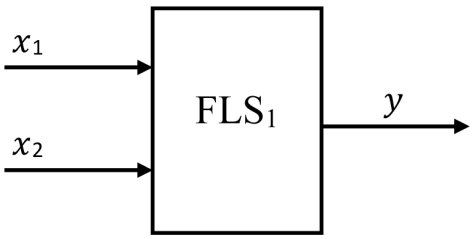


Fig. 7.3 Two inputs and a single output (2-ISO)

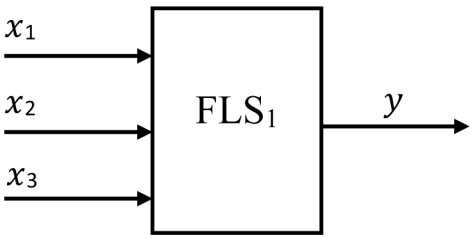


Fig. 7.4 Three inputs and a single output (3-ISO)

(ii) Pair all input variables in each subsystem

Following the design guideline of 2-ISO for each subsystem, all input variables are then paired for each subsystem according to their roles in the system. That indicates that each subsystem will get two input variables which have the same role or character. For instance, for the case of Parallel HFS as shown Fig. 7.1, input variables of *sepal length* and *sepal width* are paired together in subsystem FLS$_1$ because they both concern the *sepal*. Likewise, *petal length* and *petal width* are paired together in FLS$_2$ because they both represent the *petal*.

(iii) Small number of linguistic terms (7 ± 2).

As discussed in Chapter 2, according to a study in cognitive psychology [120], the number of different entities efficiently stored in the short-term memory should not exceed the limit of 7 ± 2. Therefore, in this case, the number of linguistic terms should not be arbitrary but must be compatible with the number conceptual entities that humans can efficiently handle and reason about during inferential activities.

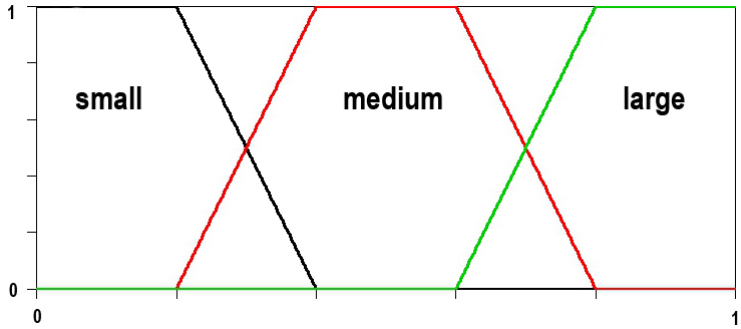


Fig. 7.5 Membership functions of inputs: *Sepal length*, *Sepal width*, *Petal length* and *Petal width*

The number of linguistic terms for HFSs should not be greater than 7 is advised in order to improve readability and simplicity. In Iris classification as in Figs. 7.1 and 7.2, the number of linguistic terms is three for each input variables: *small* (S), *medium* (M) and *large* (L), as shown in Fig. 7.5.

(iv) Set the constant linguistic terms

For simplicity and readability, all input variables including the intermediate output variables are suggested should be set with constant linguistic terms. This suggestion will also help to ease the process of designing HFSs. For example of Iris classification in Fig. 7.1, all input variables (*sepal length*, *sepal width*, *petal length* and *petal width*), including the intermediate outputs (*sepal* and *petal*), will have the same linguistic terms as shown in Fig. 7.5. However, it should be noted that this suggestion may be utilised in some instances, depending on the specific problem under consideration.

7.2.3 Guideline 3 – The Meaning of Intermediate Variables

The next design guideline focuses on the meaning of intermediate variables in HFSs. This is a fundamental issue, particularly in designing an interpretable HFS [183]. As discussed in Chapter 2, having more concrete meaning of intermediate variables will ease the process of

designing an HFS. Whereas, if the meaning of intermediate variables is more abstract, this will result in challenges in designing an HFS.

Two key points to overcome this issue are recommended in order to facilitate the design development of an interpretable HFS. The two key points are explained in the following subsection.

(i). Extract the meaningful intermediate variable from the paired input variables

The first point is to extract the meaningful intermediate variable from the paired input variables of a particular subsystem. Following on from the design guideline suggested earlier, the input variables are paired based on the same characteristic. Hence, the particular subsystem should now have a clear semantic interpretability of input and intermediate variables.

For example, in the parallel HFS in Fig. 7.1, the intermediate variable of *sepal* is extracted from the paired input variables *sepal length* and *sepal width* in subsystem FLS_1. Likewise, in the case of subsystem FLS_2, the intermediate variable of *petal* is also extracted from the paired input variables *petal length* and *petal width*. Consequently, the produced intermediate variables *sepal* and *petal* have proper meaning and thus increased semantic interpretability.

However, it is not always as easy to extract meaningful intermediate variables from the paired input variables. This is because it depends on the context of the HFSs' examples and topologies. For example, as can be seen in the serial HFS in Fig. 7.2, the intermediate variable of *sepal_petal* is extracted from the paired intermediate and input variables *sepal* and *petal length* in subsystem FLS_2. The meaning of this variable is more abstract and may reduce its semantic interpretability.

(ii). Identify from the experts

The second option is to obtain meaningful intermediate variables from human experts – human interpretability. Several researchers claim that an FLS would have the advantage regarding interpretability when involving human experts in its design [117], [184], [185].

Therefore, it is suggested to engage with experts in order to acquire meaningful intermediate variables in order to ease designing an interpretable HFS. Additionally, this especially helps for designing interpretable HFS in fields requiring human interaction such as Medicine.

7.2.4 Guideline 4 – Link-up Subsystems using Meaningful Intermediate Variables

Another design guideline is to link up all subsystems in HFSs. In this case, using an intermediate variable to connect all the subsystems is proposed. As stated in Chapter 2, an intermediate variable from the subsystem in the previous layer will serve as an input variable to another subsystem in the next layer. Therefore, the intermediate variable should have a well-defined meaning because it will be an input variable to another subsystem and thus increase the overall interpretability of HFSs in term of semantic interpretability.

For example, in the parallel HFS as in Fig. 7.1, all subsystems are connected using the intermediate variables, namely *sepal* and *petal*. Likewise, in the serial HFS as in Fig. 7.2, all subsystems are also connected using the intermediate variables, namely *sepal* and *sepal_petal*. Intuitively, parallel HFS appears to have better interpretability than serial HFS in terms of clear semantic interpretability from input, intermediate and output variables for the Iris classification.

7.2.5 Guideline 5 – Constructing the Subsystems' Rules in HFSs

The final design guideline is constructing the rules for each subsystem in HFSs. Even though all the subsystems developed using the 2-ISO model will have a smaller number of rules, it is challenging to construct the rules for subsystems, especially when intermediate variables are involved.

Two key approaches in constructing the rules for each subsystem in HFSs are recommended:

(i) Generating rules from human experts

The most interpretable rules for FLSs are rules that are acquired from human experts [127]. However, one of the crucial challenges in FLSs is to construct their rule base [186], especially when it comes to acquiring input from human experts. This is because there is no general methodology for constructing fuzzy rules from human experts. Moreover, converting the experts' know-how into IF-THEN rules is difficult, and often results are incomplete, unnecessary and involve different knowledge since experts are not capable of specific details or cannot express all their knowledge including intuition and inspiration [187].

There are several methods available to acquire the knowledge from human experts in order to generate rules. However, in this case, the use of a conventional method to

obtain the knowledge from human experts is proposed, that is to use a mixed-mode method; which combines interviews and surveys. The mixed-mode method is used to obtain knowledge which can be transformed into the rules for particular subsystems of HFSs as defined below, and illustrated in Fig. 7.6.

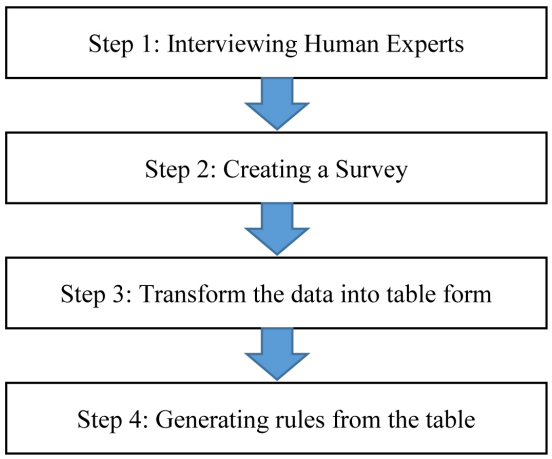


Fig. 7.6 Mixed-mode method: step-wise procedure

Step 1. *Interviewing human experts*

In the first step, an interview is conducted with human experts. At this stage, the interviewer aims to get an overview of the problem in order to extract knowledge from human experts [188–191]. Questions may include "What is the problem to be solved for the system?", "What are the input variables that should be fed into the system?" and "What possible solutions may be produced by combining inputs into the system?".

Step 2. *Creating a survey*

In this step, a survey is constructed based on the information obtained in the previous step. However, it is not a straightforward task to create this survey for capturing rules from human experts. Wagner et al. [186] introduced a novel survey-centric methodology which enables the capture of individual, subjective input from domain (not fuzzy logic) experts with minimal prior training and provides mechanisms to aggregate the resulting data into a working and interpretable fuzzy system. Given the capability of capturing the subjective views of human experts using survey in Wagner's method, this method is used to help create the survey and capture rules from experts. The survey is structured as a *survey-questionnaire* where each of N experts is presented with

a set of M matrices, determining the relationship between two variables which will respectively form the antecedent and the consequent of a rule.

		Sepal length								
		S			M			L		
		S	M	L	S	M	L	S	M	L
Sepal width	S	√			√				√	
	M		√			√				√
	L			√		√			√	

Fig. 7.7 Example of a survey completed by an expert for determining Sepal classification.

Following the suggestion of 2-ISO for subsystems of HFSs, a survey-questionnaire form is designed based on the paired inputs of each particular subsystem. For example, for Iris classification, Fig. 7.7 shows a sample of a survey-questionnaire forms that have been filled out by an expert to identify the *Sepal* classification based on the paired inputs *sepal length* and *sepal width*. An expert in particular is instructed to put one tick (✓) per column in the row which best captures the relationship between the *sepal length* and *sepal width* in order to determine the *Sepal* class.

Step 3. *Transform the data into the table (matrix) form*

Next, multiple answers of survey-questionnaires from several experts are combined and transformed into the meaningful table. For example, for *Sepal* class, two experts, namely Expert 1 and Expert 2 have answered this survey-questionnaire as shown in Fig. 7.8.

Expert 1:

		Sepal length								
		S			M			L		
		S	M	L	S	M	L	S	M	L
Sepal width	S	√			√				√	
	M		√			√				√
	L			√		√			√	

Expert 2:

		Sepal length								
		S			M			L		
		S	M	L	S	M	L	S	M	L
Sepal width	S		√				√		√	
	M			√			√			√
	L			√		√			√	

		Sepal length		
		S	M	L
Sepal width	S	*S*	*S*	*M*
	M	*M*	*L*	*L*
	L	*L*	*M*	*M*

Sepal

Fig. 7.8 Transform the expert surveys into the table for Sepal classification.

Then, both survey-questionnaire forms are aggregated into one summary table. In order to generate both survey-questionnaires, the ticks in each field across are all counted, providing an overall view of the meaning of the given relationships. For the *Sepal* class example as in Fig. 7.8, the meaning of all the ticks has represented the particular meaning of each of the linguistic terms of *Sepal* namely *small* (S), *medium* (M) and *large* (L).

Step 4. *Generating rules from the table*

In this step, all data in the matrix from the previous step, are translated into IF-THEN rules. For *Sepal* class, the information in the table taken from experts as in Fig. 7.8, is translated into IF-THEN rules format. All these rules are summarised in Table 7.1.

Table 7.1 Rules of Sepal classificaion

Rules	IF		THEN
	sepal length is	*sepal width* is	*Sepal* is
1	small	small	small
2	small	medium	medium
3	small	large	large
4	medium	small	small
5	medium	medium	large
6	medium	large	medium
7	large	small	medium
8	large	medium	large
9	large	large	medium

(ii) Decomposition from standard FLSs

The second approach to constructing the rules for each subsystem in HFSs is by using the method of decomposition from standard FLSs rules. The proposed approaches that have been presented in Chapter 2 are presented in order to decompose FLS rules into the HFS subsystems. Specifically, the approaches are S-HFS, L-HFS, two-layer HFS, and HIFAM. Note that, the proposed approaches also consider the issue of the less concrete meaning of intermediate variables for HFSs. Note that the proposed approaches can only be implemented if the FLS rules are already predefined.

For example, in the case of Iris classification as in Figs. 7.1 and 7.2, the HIFAM approach was used to construct rules for all subsystems for both the parallel and serial HFSs. The complete rules are presented in the following subsections.

- Iris classification: Parallel HFS

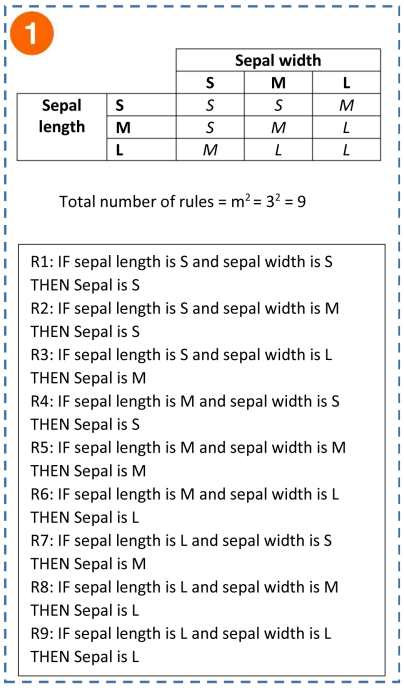


Fig. 7.9 Parallel HFS rules – FLS_1

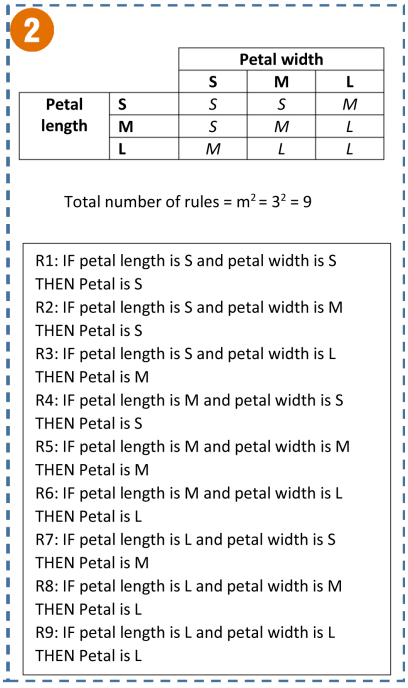


Fig. 7.10 Parallel HFS rules – FLS_2

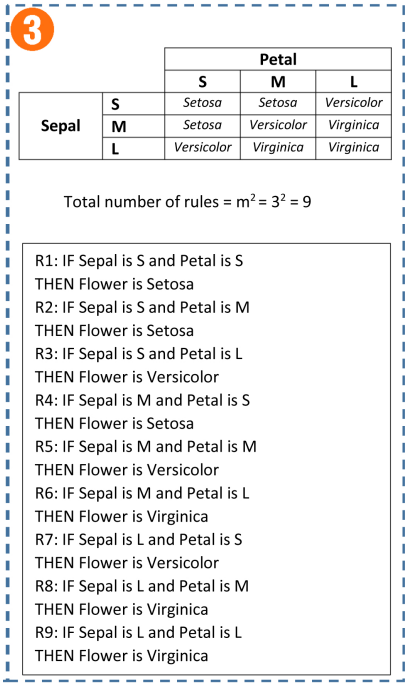


Fig. 7.11 Parallel HFS rules – FLS_3

Figs. 7.9, 7.10 and 7.11 reveal the rules created using the HIFAM approach for subsystems FLS_1, FLS_2 and FLS_3 respectively, for parallel HFS. Each subsystem contains 9 rules (3×3), and thus the total number of rules for this parallel HFS is equal to 27 rules.

- Iris classification: Serial HFS

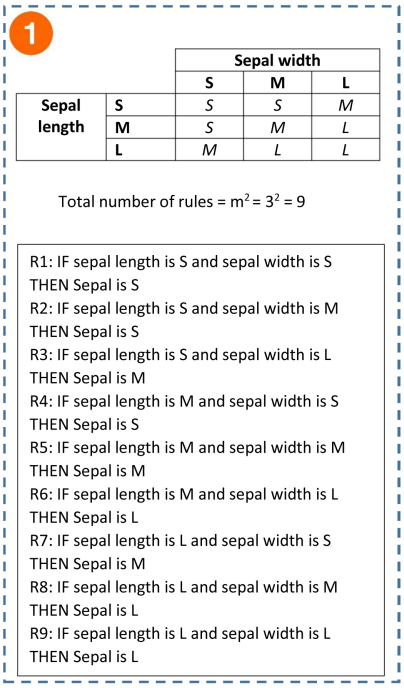


Fig. 7.12 Serial HFS rules – FLS_1

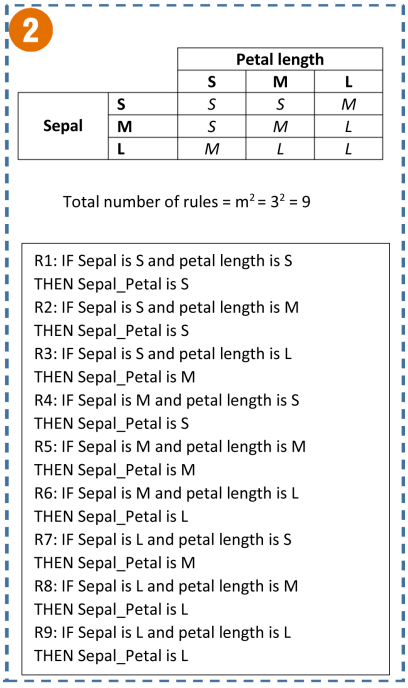


Fig. 7.13 Serial HFS rules – FLS_2

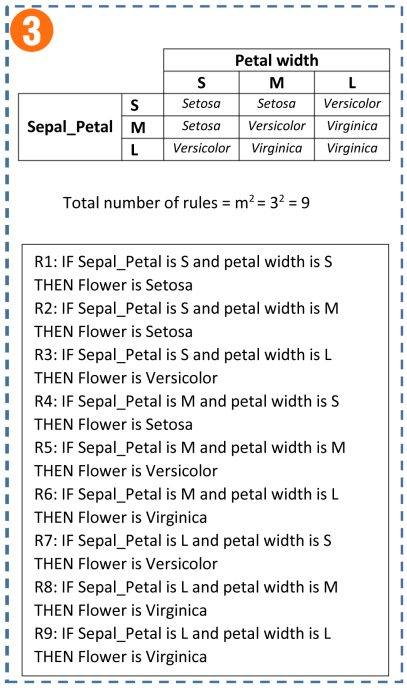


Fig. 7.14 Serial HFS rules – FLS_3

In addition, for the serial HFS, Figs. 7.12, 7.13 and 7.14 show the rules generated using the HIFAM approach for subsystems FLS_1, FLS_2 and FLS_3 respectively. Again, each subsystem contains 9 rules (3×3), and thus the total number of rules for this serial HFS is equal to 27 rules.

Although both parallel HFS and serial HFS have the same number of rules, according to the measure of interpretability and complexity in Chapters 5 and 6, the differing topologies result in different levels of both complexity and interpretability.

In the next section, the proposed design guidelines framework is demonstrated within the context of a real-world medical application.

7.3 Case Study: Neonatal Intensive Care Unit

7.3.1 Medical Background

The Neonatal Intensive Care Unit (NICU) is where the sickest newborn babies in hospitals are managed and treated. These babies are born prematurely (*less* than 32 weeks) and need special medical attention. They need hospital supervision until their systems and organs can function with no external assistance [146].

According to Larcher and Hird [192], they describe intensive care as

"*the use of invasive treatment intended to save or extend the life of a neonate who might otherwise die from organ failure. It includes artificial ventilation, other forms of organ support and artificial nutrition delivered by invasive means*".

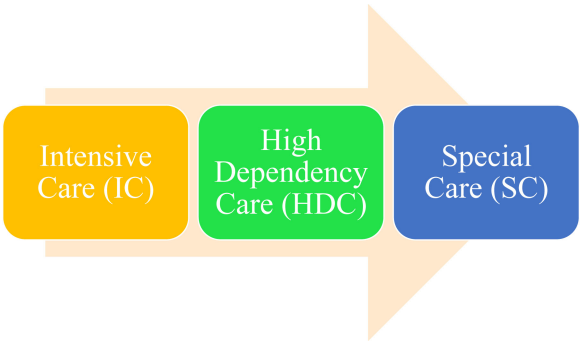


Fig. 7.15 Levels of care available at NICU. The movement of babies is depended on health progression.

Fig. 7.15 shows different levels of care available to neonatal unit [193] namely; *Neonatal intensive care* for the most seriously ill babies, *High-dependency care*, for babies who do not need to be in the NICU but still require complex care, and *Special Care* for babies who catching up on growth and development after premature birth.

7.3.2 Data Collection

This thesis collaborated with Dr Sharkey at NICU, Queen's Medical Centre (QMC) Hospital, Nottingham. He is a neonatologist and acts as the domain expert for this study. All the collected information is used to design an interpretable HFS for *oxygenation decisions* of premature babies, particularly to lead the process of creating rules and validating components of the resulting HFS.

7.3.2.1 Physiological Data Acquisition

Currently, in NICU, the patient monitor is programmed to set off an alarm sound to alert nurses and doctors when bradycardia and desaturation events occur. However, according to Dr Sharkey, the limitation of the patient monitor is that the baby's physiological data cannot be recorded and stored during such events. Thus, the events cannot be further evaluated by the NICU team members for analysis since they do not have sufficient or required information.

To tackle this issue, a system called OxyGUI was developed which would allow the data associated to the two above-mentioned events to be recorded. This will help the NICU team members to establish standard methodologies for the acquisition, storage, and extraction of physiological data from a cot-side monitor.

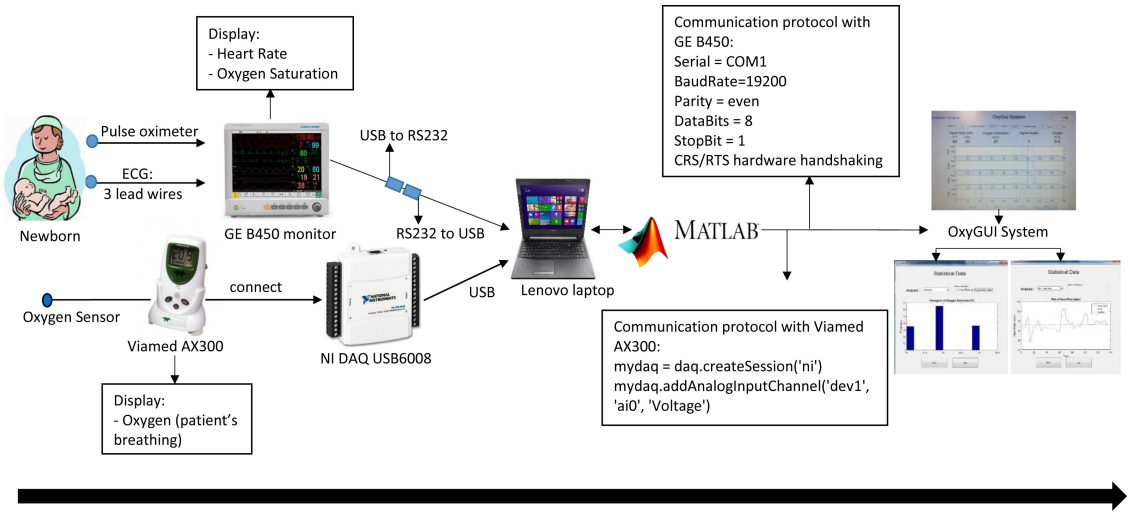


Fig. 7.16 Data flow in the data acquisition process

Fig. 7.16 shows the whole process involved in obtaining physiological data of premature babies at NICU. This OxyGUI system will record data for every one second starting when the baby is admitted into the NICU including the baby ID, admission date, date of birth, gestation age, HR, SpO_2, signal quality and inspired oxygen (FiO_2). Note that 'signal quality' intends to assess whether reliable HR can be obtained from the three-lead Electrocardiogram (ECG), as shown in Fig. 7.16. The descriptions and sample of these data are presented in Table 7.2 and Fig. 7.17 respectively. Also, the user manual on the OxyGUI system is provided in Appendix D. The OxyGUI provides the convenience of viewing the statistical analysis of these recorded data such as the frequency of these events, i.e., bradycardia, tachycardia and desaturation that arise during the day or night. This data is very important in establishing the validity of the proposed interpretable HFS.

Table 7.2 Physiological data

Attributes	**Description**
Baby_ID	The ID is assigned for each baby at NICU
Admission_Date	Date of start admission to NICU
Gestation_age	Gestation age (in week) of baby when admission to NICU
Date_Time	A timestamp is the time at which an event is recorded
HR	Heart rate, beat per minutes (bpm)
SpO_2	Oxygen Saturation (%)
FiO_2	Inspired oxygen (%)- Percentage of oxygen in the space being measured
signal_quality	Signal quality

Baby_ID	Admission_date	Gestation_age	Date_Time	SpO2	HR	signal_quality	FiO2
Baby 001	09/08/2018	29	09/08/2018 10:51:43	88	160	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:44	94	158	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:45	94	158	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:46	94	158	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:47	94	158	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:48	94	158	130	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:49	94	158	133	29.6
Baby 001	09/08/2018	29	09/08/2018 10:51:50	94	158	130	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:51	94	158	130	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:52	94	158	123	30.1
Baby 001	09/08/2018	29	09/08/2018 10:51:53	94	158	127	30
Baby 001	09/08/2018	29	09/08/2018 10:51:54	94	158	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:55	95	159	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:56	95	159	120	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:57	95	159	120	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:58	95	159	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:51:59	95	160	127	29.6
Baby 001	09/08/2018	29	09/08/2018 10:52:00	95	160	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:01	95	160	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:02	95	160	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:03	94	160	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:04	95	161	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:05	94	161	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:06	94	161	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:07	95	161	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:08	94	161	123	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:09	94	161	120	29.6
Baby 001	09/08/2018	29	09/08/2018 10:52:10	94	161	127	29.5
Baby 001	09/08/2018	29	09/08/2018 10:52:11	94	161	123	30.2
Baby 001	09/08/2018	29	09/08/2018 10:52:12	94	161	123	30.2

Fig. 7.17 Sample of physiological data

7.3.2.2 An Initial Data Analysis on Physiological Data

For the initial data analysis, the physiological data from one premature baby (referred to here as Baby001) over the six days of period at NICU is analysed. It is a large dataset consisting of 448208 rows, since it was recorded for every single second. Therefore, the data was first pre-processed for each day as shown in Table 7.3a. The data was then processed second time and 106 records of outliers were removed (heart rate = 1/bpm) from the original data as shown in Table 7.3b. This data was then used for the analysis phase.

Table 7.3 The pre-process data

(a) Data with outlier

Day	**Hours**	**Minutes**	**Seconds**
Day 1	24	1440	86400
Day 2	24	1440	86400
Day 3	24	1440	86400
Day 4	24	1440	86400
Day 5	24	1440	86400
Day 6	4.5	270.13	16208

(b) Clean data

Day	**Hours**	**Minutes**	**Seconds**
Day 1	24	1439.93	86396
Day 2	24	1439.65	86379
Day 3	24	1439.52	86371
Day 4	24	1439.48	86369
Day 5	24	1439.65	86379
Day 6	4.5	270.13	16208

Mean (μ) and standard deviation (σ) are used to show the distribution of this physiological of Baby001 data, as shown in Table 7.4 over the six days duration. Fig. 7.18 shows the histogram of each parameter which are oxygen saturation, heart rate, signal quality and inspired oxygen illustrating the distribution of data collected over six days.

Table 7.4 Mean (μ) and standard deviation (σ) of Baby001 data

	Physiological parameter			
	SpO_2 (%)	HR (bpm)	Signal Quality	FiO_2 (%)
μ	90.1	155.5	144.6	41.2
σ	5.1	9.4	78.0	5.9

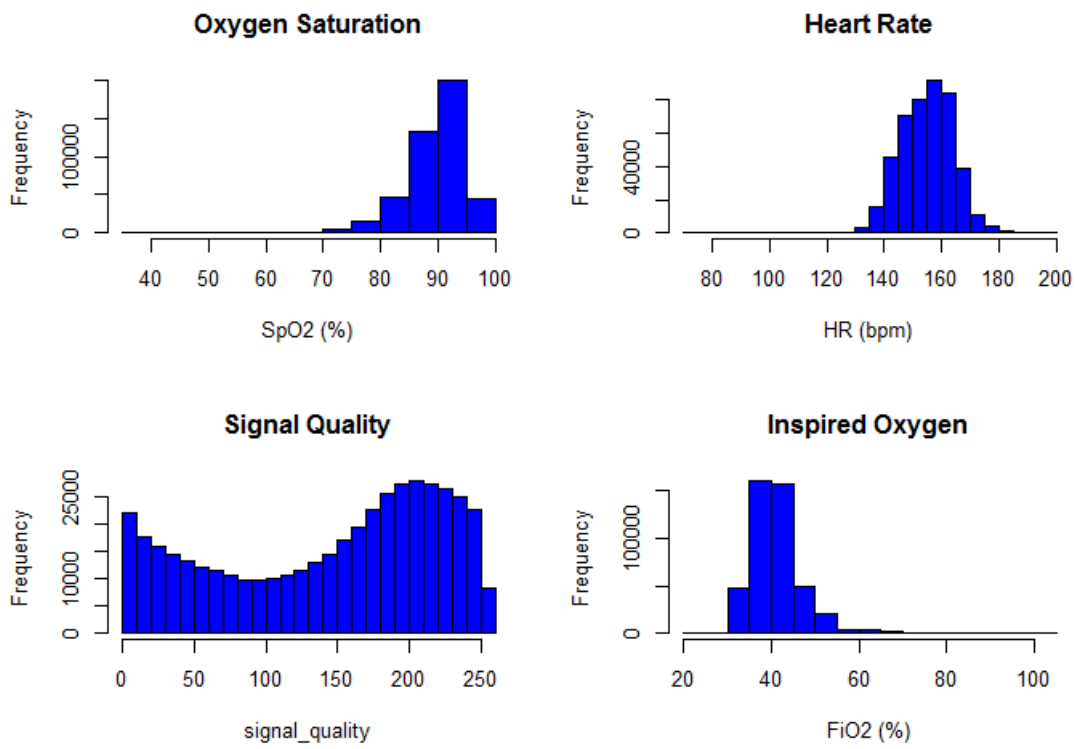


Fig. 7.18 Histogram of each of physiological parameter for Baby001.

Scatter plots can be used to examine the relationship between parameters in this data. Six scatter plots (in Fig. 7.19) show the correlation between four parameters namely; heart rate (HR), oxygen saturation (SpO_2), signal quality and inspired oxygen (FiO_2).

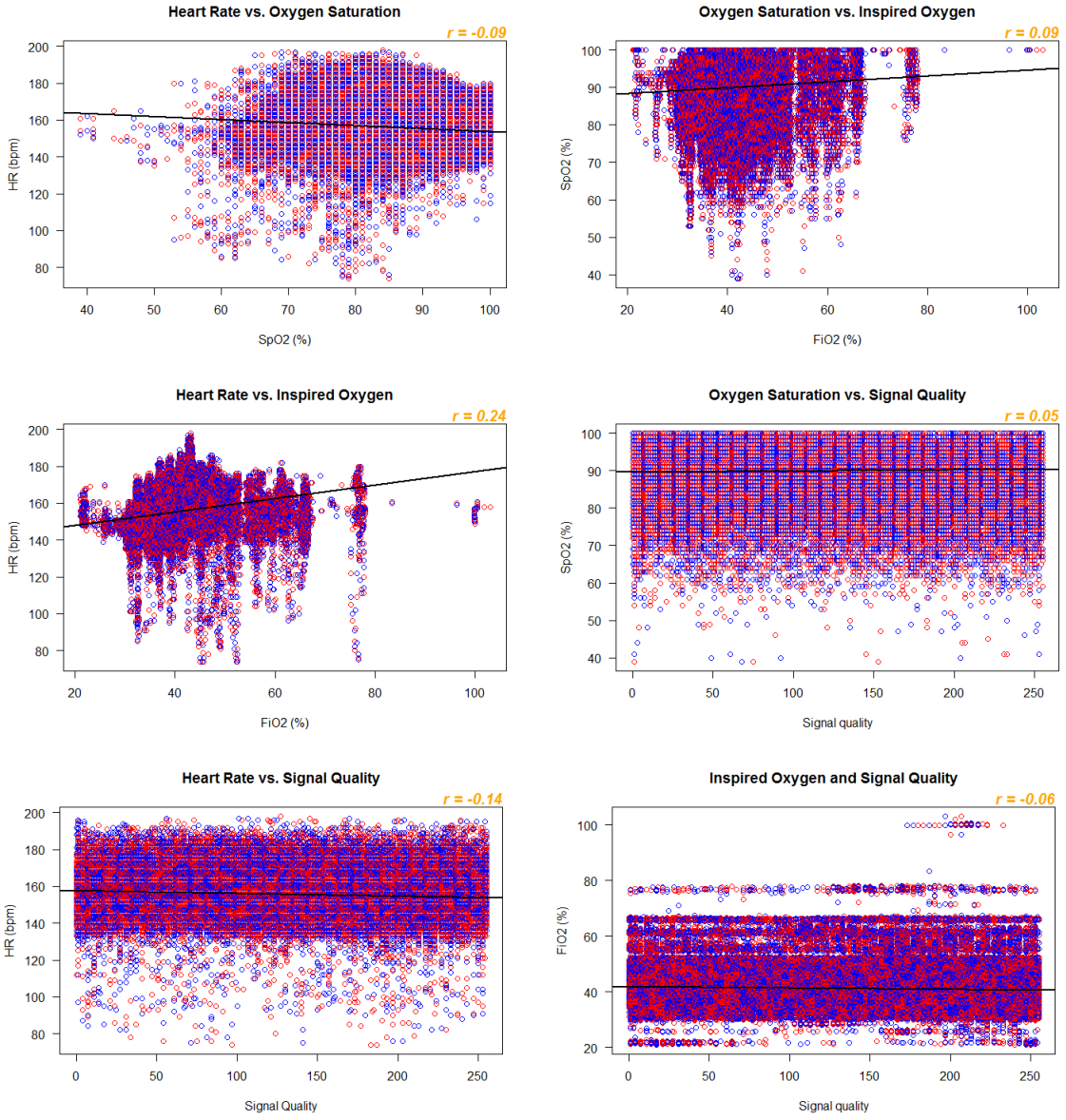


Fig. 7.19 Scatter plot of the parameter correlation

Fig. 7.19 shows that the plots do not seem to produce a significant relationship between the parameters' combination. In fact, some of the correlations such as heart rate and oxygen saturation, heart rate and signal quality, inspired oxygen and signal quality, show negative relationships. However, it is worth emphasising that, the correlation between oxygen saturation and inspired oxygen, and heart rate and inspired oxygen, are considered as significant since the computed Pearson's correlation coefficient for the aforementioned relationship is 24% and 9% respectively. Therefore, these type of relationships will be further

investigated and will be as the main focus in the process of designing an interpretable HFS for NICU, illustrated in the following subsection.

7.3.3 Designing an Interpretable Hierarchical Fuzzy System: Oxygenation Decision

In this section, the proposed design guidelines framework is demonstrated in building an interpretable HFS for the *Oxygenation decision* and is described in the following subsection. Note that the *oxygenation decision* at NICU is essential to ensure adequate oxygenation for the premature baby based on their vital signs such as heart rate and oxygen saturation.

7.3.3.1 Interviewing Human Experts

Following the findings in the initial analysis in Subsection 7.3.2.2, first, the interview process to obtain the physiological parameters, symptoms and vital signs with the domain experts at NICU is discussed. The information from this interview is essential as it will ease the process of constructing HFS topology and the survey-questionnaires.

Dr Sharkey was interviewed to obtain information for each of the parameters, i.e., HR and SpO_2 and capture the primary rule base taken into account when any decision is made at the NICU. The results from the interview for HR and SpO_2 are illustrated below.

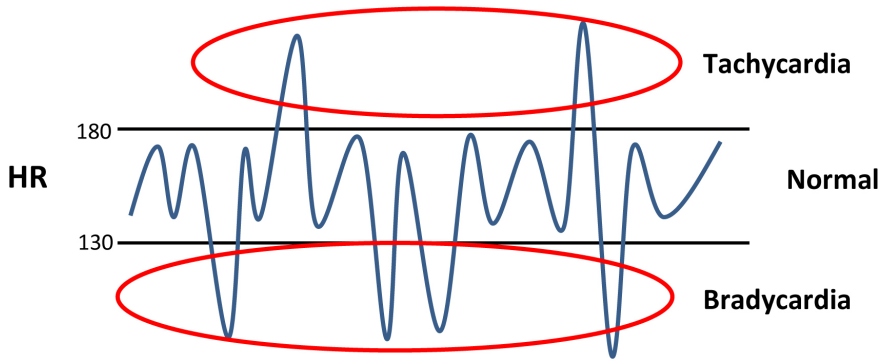


Fig. 7.20 Physiological parameter : Heart Rate (bpm)

Preterm babies are very sensitive and need special care. Their hearts are usually not fully developed and require continuous monitoring [194]. Fig. 7.20 shows the HR classification of a baby's heartrate at NICU indicating bradycardia, normal heart rate and tachycardia. The HR is measured in a beat per minutes (bpm) which is shown on the bedside monitor. Fig. 7.20 indicates that the HR for a normal baby is in the range of 130 to 180 bpm. If the baby's

heart rate drops below 130 bpm, that implies the baby is suffering from bradycardia and needs attention. Meanwhile, if the baby's heart rate rises above 180 bpm that means the baby has tachycardia and also needs attention. However, these are all dependent on the baby's condition, i.e., weight, crying or sleeping. For example, if the baby is crying or touched then the heart rate will rise. Note that the nurse on duty will record the baby's condition information into the nursing chart for every hour.

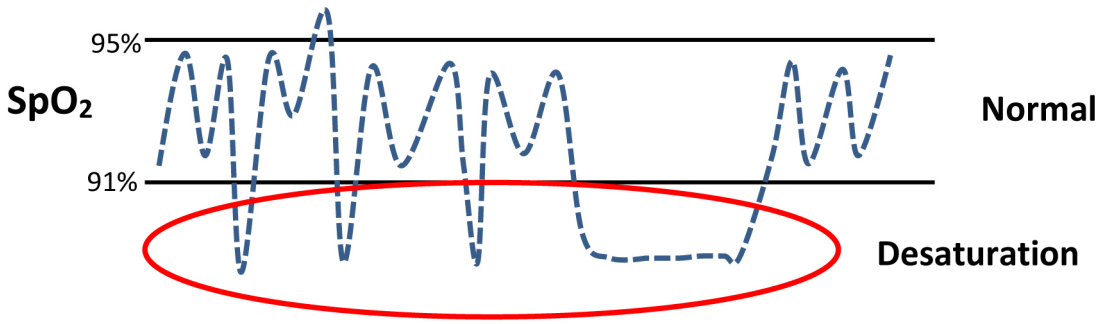


Fig. 7.21 Physiological parameter : Oxygen Saturation (%)

Pulse oximetry measures oxygen saturation (SpO_2) and is a well-established method of monitoring patient oxygenation status [161]. The SpO_2 is the amount of oxygen bound to hemoglobin in the blood expressed as a percentage of the maximal binding capacity which is also shown on the bedside monitor. Fig. 7.21 shows the SpO_2 classification of a baby at NICU and includes normal and desaturation. A normal baby will have a SpO_2 reading in between 91% to 95% whereas if the SpO_2 reading decreases to below 91% then this indicates that the baby is experiencing desaturation. Usually, when desaturation occurs, the nurse will try to increase the oxygen in the atmosphere (FiO_2) to correct this.

7.3.3.2 Guideline 1: Selection of Topology

Following the information obtained during the interview session discussed in Subsection 7.3.3.1, an HFS topology for *Oxygenation decision* at NICU is constructed as shown in Fig. 7.22.

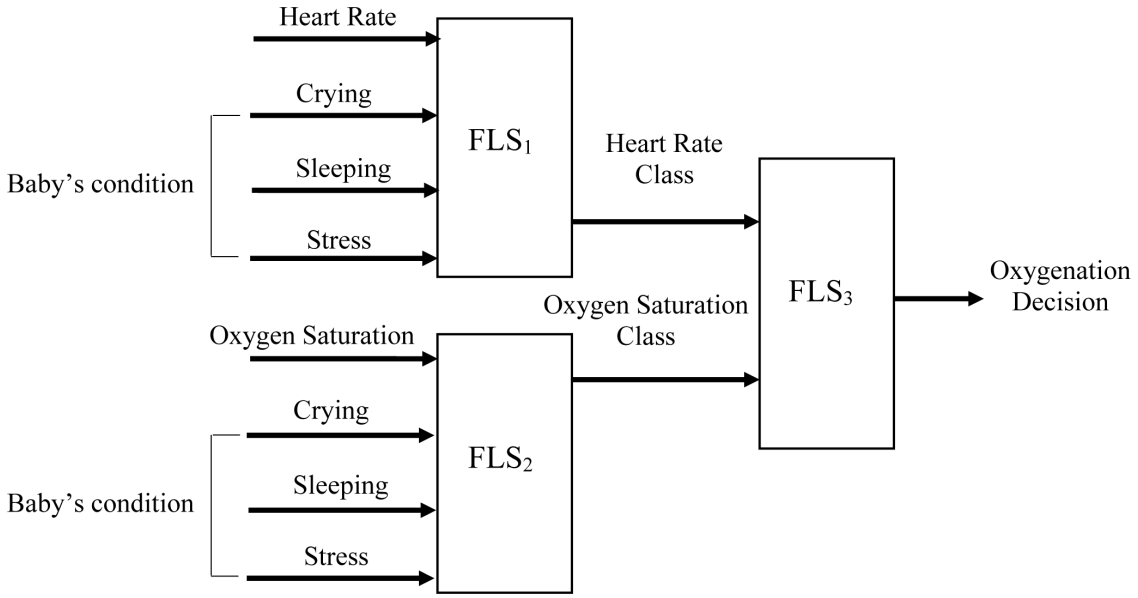


Fig. 7.22 HFS topology: Oxygenation Decision

The HFS topology as shown in Fig. 7.22 consists of three subsystems namely FLS_1, FLS_2 and FLS_3. These subsystems have paired inputs such as (HR, Baby's condition) $\rightarrow$ Heart Rate class, (SpO_2, Baby's condition) $\rightarrow$ Oxygen saturation class, and (Heart Rate class, Oxygen saturation class) $\rightarrow$ *Oxygenation decision*.

Note that the input called baby's condition that appears in subsystems FLS_1 and FLS_2 is the same. From the interview session with experts, it was found that both inputs HR and SpO_2 are very sensitive to the baby's condition input and influence the decision in the heart rate class and oxygen saturation class respectively. Thus, the topology of HFS for *Oxygenation decision* is designed as shown in Fig. 7.22. It should also be noted that the baby's condition input comprises several conditions, namely *crying*, *sleeping* and *stress*.

7.3.3.3 Guideline 2: Simplicity and Readability of Subsystems

For simplicity and readability of the subsystem, it is suggested that the number of linguistic terms for HFS *oxygenation decision* must not be greater than seven. Figs. 7.23 to 7.30 show the membership functions for the inputs, intermediate and output variables that are formed for an HFS of the *oxygenation decision* respectively.

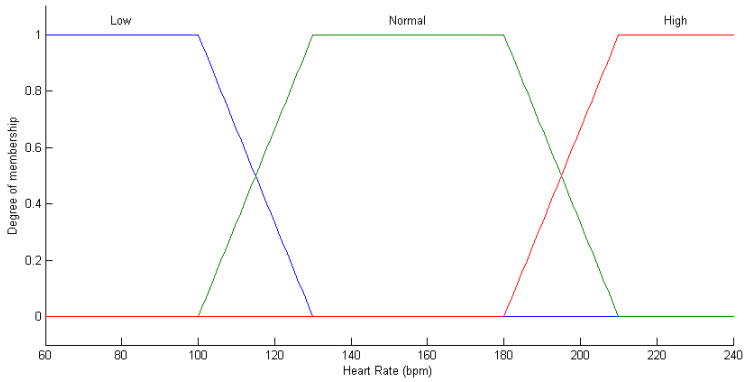


Fig. 7.23 MFs of input variable: Heart Rate (bpm)

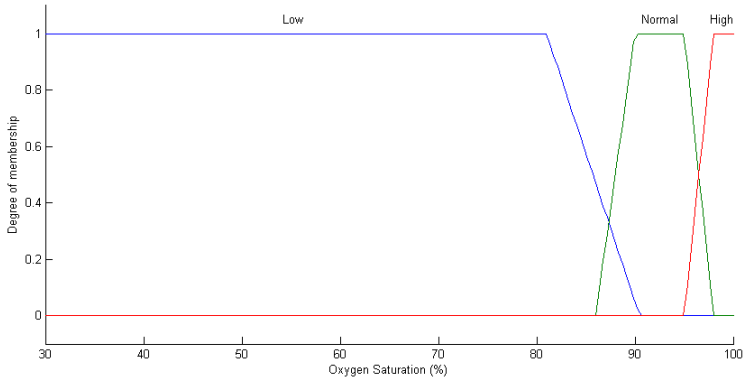


Fig. 7.24 MFs of input variable: Oxygen Saturation (%)

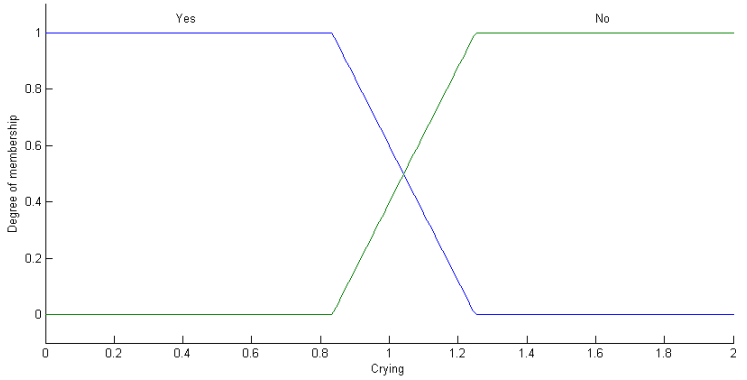


Fig. 7.25 MFs of input variable: Baby's condition (Crying)

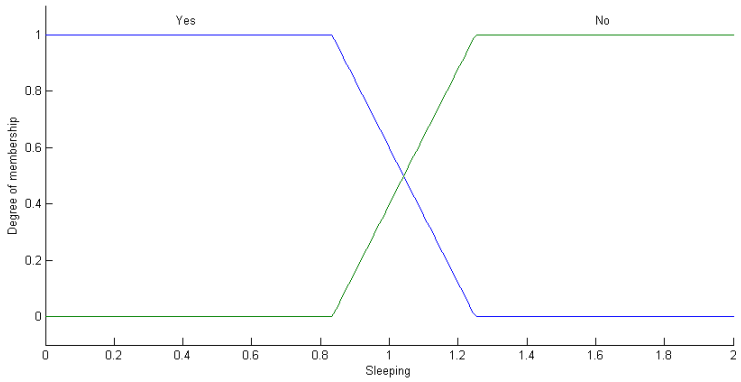


Fig. 7.26 MFs of input variable: Baby's condition (Sleeping)

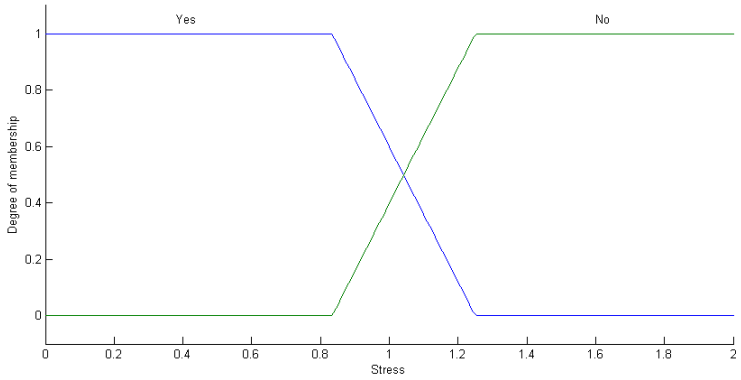


Fig. 7.27 MFs of input variable: Baby's condition (Stress)

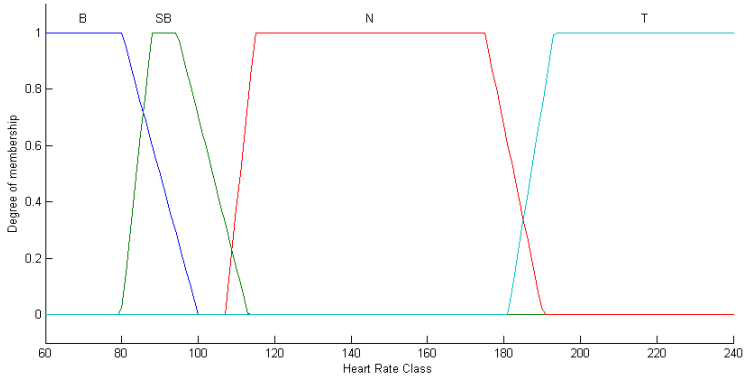


Fig. 7.28 MFs of intermediate variable: Heart Rate class

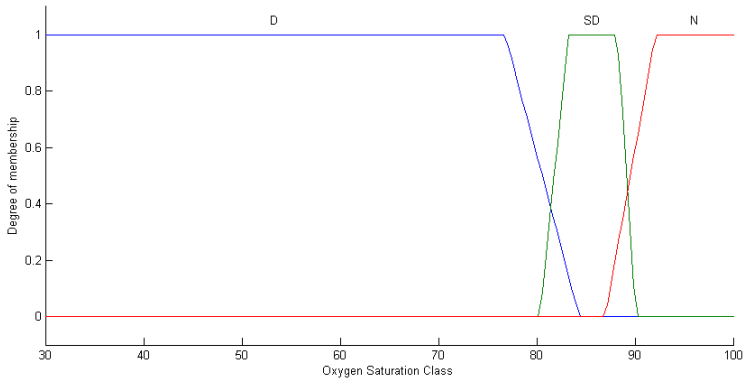


Fig. 7.29 MFs of intermediate variable: Oxygen Saturation class

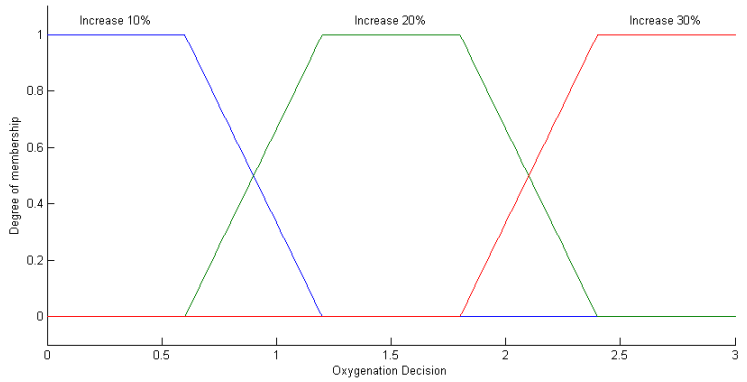


Fig. 7.30 MFs of output variable: Oxygenation Decision

7.3.3.4 Guideline 3 & 4: Meaning of Intermediate Variables and Linked Subsystems

In this subsection, the meaningful intermediate variables of the proposed HFS *oxygenation decision* are identified from a human expert, in interview session with Dr Sharkey as described in Subsection 7.3.3.1. Therefore, in Fig. 7.22, the intermediate variables *heart rate class* and *oxygen saturation class* are labelled as the output of subsystems FLS_1 and FLS_2, respectively. Note that these meaningful intermediate variables play an important role in increasing the semantic interpretability of HFSs. Consequently, these intermediate variables are used as a connection between subsystems FLS_1 and FLS_3, and subsystems FLS_2 and FLS_3, as shown in Fig. 7.22.

7.3.3.5 Guideline 5: Constructing the Rules of Subsystems

(i) *Creating a Survey*

In order to obtain HFS rules from domain experts, a survey-questionnaire is constructed. Following the proposed design guidelines framework, several survey-questionnaires based on the information collected earlier in the interview session and the HFS topology have been designed. These survey-questionnaires are intended to simplify the process of extracting the rules used by experts when making *oxygenation decisions*. Note that the same domain expert, Dr Sharkey, was asked to answer these survey-questionnaires. The samples of these survey-questionnaires that has been filled out by domain experts can be viewed in Tables 7.6 to 7.11. Also, for clarity, the abbreviations and symbols used in the survey-questionnaires can be seen in Table 7.5.

Table 7.5 Abbreviations and symbols used in the survey

Abbreviations / Symbols	Descriptions
Y	Yes
N	No
L	Low
N	Normal
H	High
B	Bradycardia
SB	Slight Bradycardia
N	Normal
T	Tachycardia
D	Desaturation
SD	Slight Desaturation
N	Normal
$\uparrow$	Increase
O_2	Oxygen

Table 7.6 Results extracted from the survey (Heart Rate vs Crying)

		Heart Rate (bpm)											
		HR <130				130 <HR <180				HR >180			
		B	SB	N	T	B	SB	N	T	B	SB	N	T
Crying	Y	✓						✓					✓
	N		✓					✓					✓

Table 7.7 Results extracted from the survey (Heart Rate vs Sleeping)

		Heart Rate (bpm)											
		HR <130				130 <HR <180				HR >180			
		B	SB	N	T	B	SB	N	T	B	SB	N	T
Sleeping	Y	✓						✓					✓
	N		✓					✓					✓

Table 7.8 Results extracted from the survey (Heart Rate vs Stress)

		Heart Rate (bpm)											
		HR <130				130 <HR <180				HR >180			
		B	SB	N	T	B	SB	N	T	B	SB	N	T
Stress	Y	✓						✓					✓
	N		✓					✓					✓

Table 7.9 Results extracted from the survey (Oxygen Saturation vs Crying)

		Oxygen Saturation (%)								
		SpO_2 <91			91 <SpO_2 <95			SpO_2 >95		
		D	SD	N	D	SD	N	D	SD	N
Crying	Y	✓					✓			✓
	N		✓				✓			✓

Table 7.10 Results extracted from the survey (Oxygen Saturation vs Sleeping)

		Oxygen Saturation (%)								
		SpO_2 <91			91 <SpO_2 <95			SpO_2 >95		
		D	SD	N	D	SD	N	D	SD	N
Sleeping	Y	✓					✓			✓
	N		✓				✓			✓

Table 7.11 Results extracted from the survey (Oxygen Saturation vs Stress)

		Oxygen Saturation (%)								
		SpO_2 <91			91 <SpO_2 <95			SpO_2 >95		
		D	SD	N	D	SD	N	D	SD	N
Stress	Y	✓					✓			✓
	N		✓				✓			✓

(ii) *Transforming Survey Data into Rule Matrices*

Following the results of the survey-questionnaires obtained earlier, in this subsection, the information is transformed into a meaningful table of rules based on the proposed

design guidelines framework. That is, the ticks (✓) in the survey-questionnaires are then defined as the specific linguistic terms as shown in Tables 7.12, 7.13 and 7.14.

Table 7.12 Fuzzy Rule Base (FLS_1): Heart rate classification

		Heart Rate (bpm)		
		Low	Normal	High
Crying	Y	B	N	T
	N	SB	N	T
Sleeping	Y	B	N	T
	N	SB	N	T
Stress	Y	B	N	T
	N	SB	N	T

Table 7.13 Fuzzy Rule Base (FLS_2): Oxygen saturation class

		Oxygen Saturation (%)		
		Low	Normal	High
Crying	Y	D	N	N
	N	SD	N	N
Sleeping	Y	D	N	N
	N	SD	N	N
Stress	Y	D	N	N
	N	SD	N	N

Table 7.14 Fuzzy Rule Base (FLS_3): Oxygenation decision

		HR Class			
		B	SB	N	T
SpO_2 Class	N	$\uparrow O_2$ - 20%	-	-	-
	SD	$\uparrow O_2$ - 30%	$\uparrow O_2$ - 10%	-	-
	D	$\uparrow O_2$ - 30%	$\uparrow O_2$ - 20%	$\uparrow O_2$ - 10%	$\uparrow O_2$ - 20%

(iii) *Generating HFS Rules from Tables*

In this subsection, all the rules for the subsystems of the HFS are generated. Fig. 7.31 shows the process involved from the beginning in order to acquire the rules from the domain experts.

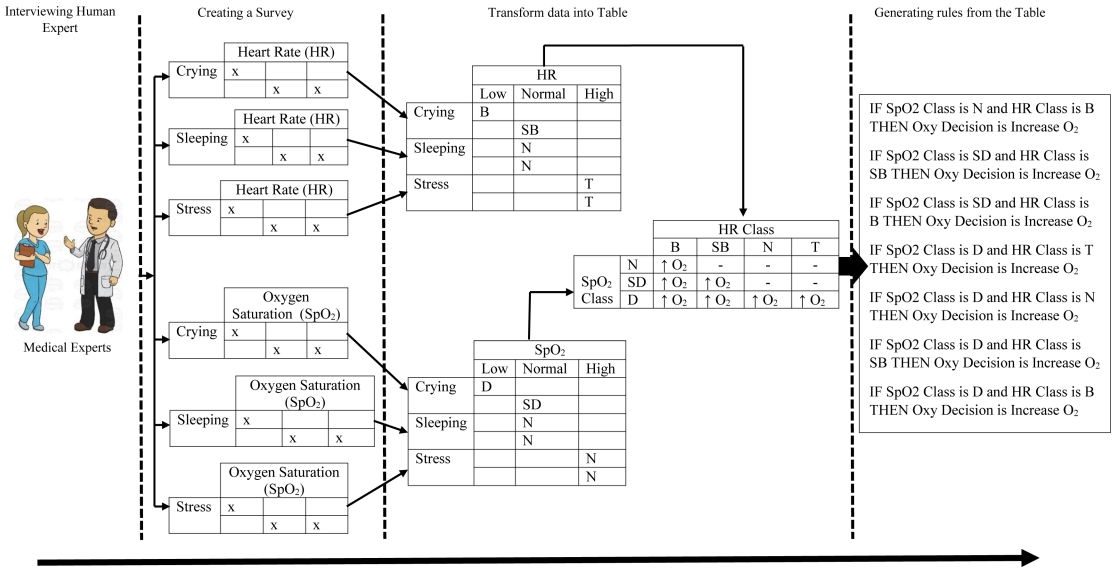


Fig. 7.31 The process involved in producing interpretable HFS rules from human experts

Following the process involved in Fig. 7.31, the final IF-THEN rules in the subsystems FLS_1, FLS_2, and FLS_3 are produced as shown in Tables 7.15, 7.16, and 7.17 respectively.

Table 7.15 HFS rules in subsystem FLS_1 using 'AND' logical connective

Rules No.	IF				THEN
	HR is	Crying is	Sleeping is	Stress is	HR Class is
1	L	Y	N	Y	B
2	L	Y	N	N	B
3	L	N	Y	Y	B
4	L	N	Y	N	SB
5	L	N	N	Y	SB
6	L	N	N	N	SB
7	N	Y	N	Y	B
8	N	Y	N	N	N
9	N	N	Y	Y	N
10	N	N	Y	N	N
11	N	N	N	Y	N
12	N	N	N	N	N
13	H	Y	N	Y	T
14	H	Y	N	N	T
15	H	N	Y	Y	T
16	H	N	Y	N	T
17	H	N	N	Y	T
18	H	N	N	N	T

Table 7.16 HFS rules in subsystem FLS_2 using 'AND' logical connective

Rules No.	IF				THEN
	SpO_2 is	Crying is	Sleeping is	Stress is	SpO_2 Class is
1	L	Y	N	Y	D
2	L	Y	N	N	D
3	L	N	Y	Y	D
4	L	N	Y	N	SD
5	L	N	N	Y	SD
6	L	N	N	N	SD
7	N	Y	N	Y	D
8	N	Y	N	N	N
9	N	N	Y	Y	N
10	N	N	Y	N	N
11	N	N	N	Y	N
12	N	N	N	N	N
13	H	Y	N	Y	SD
14	H	Y	N	N	N
15	H	N	Y	Y	N
16	H	N	Y	N	N
17	H	N	N	Y	N
18	H	N	N	N	N

Table 7.17 HFS rules in subsystem FLS_3

Rules No.	IF		THEN
	SpO_2 Class is	HR Class is	Oxygenation Decision is
1	D	B	$\uparrow O_2$ - 30%
2	D	SB	$\uparrow O_2$ - 20%
3	D	N	$\uparrow O_2$ - 10%
4	D	T	$\uparrow O_2$ - 20%
5	SD	B	$\uparrow O_2$ - 30%
6	SD	SB	$\uparrow O_2$ - 10%
7	N	B	$\uparrow O_2$ - 20%

Finally, an interpretable HFS for making *oxygenation decision* at the NICU is designed using the proposed design guidelines framework. In the next subsection, the *oxygenation*

decision system is developed using the proposed interpretable HFS design and is illustrated with the physiological data from one premature baby at the NICU.

7.3.4 A Practical Illustration of an Interpretable HFS

In this section, both a flat and heirarchical FLS is created for the oxygenation decision making system. In the case of the flat system, its topology is shown in Fig. 7.32, which consists of 54 rules. Table 7.18 shows a sample of the rules used in the system.

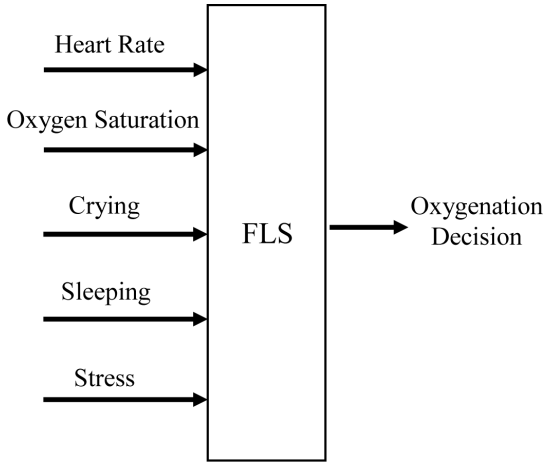


Fig. 7.32 FLS topology: Oxygenation Decision

Table 7.18 A flat FLS rules using 'AND' logical connective

Rules No.	IF					THEN
	HR is	SpO_2 is	Crying is	Sleeping is	Stress is	Oxygenation Decision
1	L	L	Y	N	Y	$\uparrow O_2$ - 30%
2	L	L	Y	N	N	$\uparrow O_2$ - 30%
3	L	L	N	Y	Y	$\uparrow O_2$ - 30%
4	L	L	N	N	Y	$\uparrow O_2$ - 30%
⋮						
16	L	H	N	N	Y	$\uparrow O_2$ - 20%
17	L	H	N	Y	N	$\uparrow O_2$ - 20%
18	L	H	N	N	N	$\uparrow O_2$ - 10%
19	N	L	Y	N	Y	$\uparrow O_2$ - 30%
⋮						
52	H	H	N	N	Y	$\uparrow O_2$ - 10%
53	H	H	N	Y	N	$\uparrow O_2$ - 20%
54	H	H	N	N	N	$\uparrow O_2$ - 20%

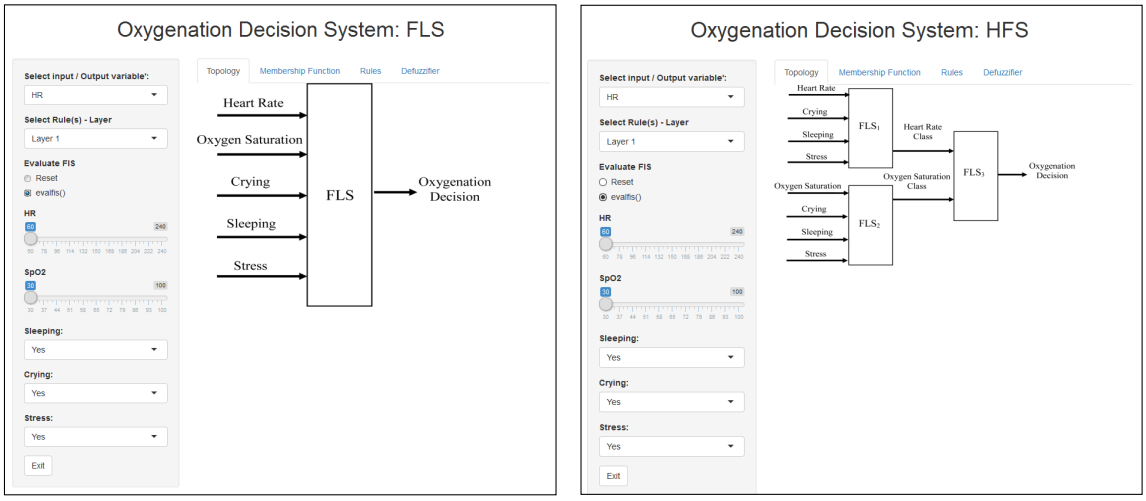
This section illustrates how the oxygenation decision systems (HFS and FLS) behave in practice using a real-world dataset based on data from one of the premature babies (Baby 001). This is done in three main steps, which are defined in the following subsection.

7.3.4.1 Generating Fuzzy Systems using the FuzzyToolkitUoN

R programming is becoming popular among researchers across disciplines. Wagner et al. [66] had developed a fuzzy logic package (FuzzyToolkitUoN) in R programming for Type-1 and Type-2 fuzzy logic systems. In the first step, both FLS and HFS designs are used as a guideline in developing both fuzzy systems using the FuzzyToolkitUoN. That is, all the information about membership functions and rules that was previously discussed in Sections 7.3.3.3, 7.3.3.5 and 7.3.4 are used to develop both fuzzy systems. The key features of both the FLS and HFS for this oxygenation decision system are described in Table 7.19.

Table 7.19 Features of both FLS and HFS implementations of the oxygenation decision system

Features	Oxygenation decision system	
	FLS	HFS
Number of input variables	5	5
Number of intermediate variables	0	2
Number of output variable	1	1
Average of number of membership functions	2.5	2.75
Number of rules	54	43
Number of subsystems	1	3
Number of layers	1	2



(a) FLS (b) HFS

Fig. 7.33 The graphical user interface (GUI) for the Oxygenation Decision system: Main interface

This step also provides a graphical user interface (GUI) in the R programming for both the FLS and HFS as shown in Fig. 7.33. The GUI of these systems consists of the number of membership functions for input, intermediate and output variables, topology, rules, defuzzification view and the *evalfis()* function. The *evalfis()* function used in this chapter to evaluate the fuzzy inference system is described below.

7.3.4.2 Evaluating Fuzzy Inference Systems (FIS) of the FLS and HFS

Then, both fuzzy systems are used to simulate the oxygenation decision-making problem by using real data of one premature baby (Baby001) over a six days period. It should be noted that the data of Baby 001 is same as discussed earlier in Section 7.3.2.2. In order to that, in this step, the function *evalfis()* in FuzzyToolkitUoN is adopted to evaluate the output of the fuzzy system for each given input in the dataset. Finally, the results obtained from the evaluation were then compared, as shown in Table 7.20. The comparison of the results is discussed in the next step.

Table 7.20 The Sample Evaluation of FIS output for FLS and HFS of Oxygenation Decision

HR	SpO2	Crying	Sleeping	Stress	FLS output:	HFS output:		
					Oxygenation decision	HR Class	SpO2 Class	Oxygenation decision
150	93	N	Y	N	**0.4589**	146.9188	94.8293	**1.5000**
149	92	N	N	Y	**0.4589**	94.4470	85.4288	**0.9672**
149	92	N	Y	Y	**0.5102**	147.6064	94.3597	**1.5000**
149	91	N	N	Y	**0.5141**	82.9117	58.9035	**2.0890**
149	91	N	N	N	**0.4589**	146.9188	94.8293	**1.5000**
149	90	Y	N	N	**0.5768**	94.4470	77.7779	**1.6730**
149	90	Y	N	N	**0.5768**	94.4470	77.7779	**1.6730**
149	90	Y	N	N	**0.5892**	94.5198	77.4740	**1.6671**
149	90	N	Y	Y	**0.5788**	146.7159	87.2695	**1.5000**
149	90	N	N	N	**0.6940**	133.6602	79.9252	**0.5094**
149	90	Y	N	N	**0.6222**	94.7543	76.6634	**1.6546**
149	89	N	Y	N	**0.9669**	147.6541	92.9283	**1.5000**
149	89	N	Y	Y	**0.8685**	142.0905	76.4446	**0.4589**
148	90	N	Y	Y	**0.5768**	146.9188	87.3354	**1.5000**
148	90	Y	N	Y	**0.5768**	75.0950	55.1613	**2.5411**
148	90	N	Y	Y	**0.6665**	132.1867	71.0655	**0.4589**
147	90	N	N	N	**0.5942**	145.1435	94.1103	**1.5000**
147	91	Y	Y	N	**1.5000**	150.0000	65.0000	**0.4589**
147	91	Y	N	Y	**0.4589**	75.0950	55.1613	**2.5411**
147	91	N	N	N	**0.4589**	146.9188	94.8293	**1.5000**
147	91	N	N	Y	**0.4709**	111.3399	86.7614	**0.5780**
147	91	N	N	Y	**0.4589**	94.4470	85.4288	**0.9672**
147	91	Y	Y	Y	**1.5000**	150.0000	65.0000	**0.4589**
147	91	N	N	N	**0.4780**	147.1717	94.6512	**1.5000**
147	91	Y	Y	Y	**0.5642**	78.6337	56.6428	**2.5411**
147	90	N	N	N	**0.6195**	142.5988	85.8297	**1.5000**
147	90	Y	Y	N	**1.5000**	150.0000	65.0000	**0.4589**
147	90	N	Y	N	**0.5768**	146.9188	94.4967	**1.5000**
147	90	Y	N	Y	**0.6582**	76.5334	55.8016	**2.5411**
148	90	Y	Y	Y	**0.7784**	100.5482	58.5183	**1.5000**
148	90	Y	N	Y	**0.6768**	76.7829	55.9090	**2.5411**
148	90	Y	N	Y	**0.6223**	78.8592	56.5957	**2.5411**
148	91	N	Y	Y	**0.4589**	146.9188	94.8293	**1.5000**

7.3.4.3 Examining equivalence between FLS and HFS for oxygenation decision

In this step, a statistical test was conducted using the K-S test [182] (as used in Chapter 6) to compare the equivalence of the output produced from the FLS and HFS as in Section 7.3.4.2.

The K-S test statistic measures the maximum distance (D) between the empirical cumulative distribution function (ECDF) of an FLS and HFS, measured in the vertical direction.

Table 7.21 The Kolmogorov-Smirnov Test between FLS and HFS of Oxygenation Decision

Input samples (n)	FLS vs HFS		Critical Value
	D	p-values	$D_\alpha = \frac{1.36}{\sqrt{n}}$
Day 1 ($n = 84600$)	0.29954	$2.2e^{-16}$	0.00468
Day 2 ($n = 84600$)	0.23806	$2.2e^{-16}$	0.00468
Day 3 ($n = 84600$)	0.16498	$2.2e^{-16}$	0.00468
Day 4 ($n = 84600$)	0.23157	$2.2e^{-16}$	0.00468
Day 5 ($n = 84600$)	0.22406	$2.2e^{-16}$	0.00468
Day 6 ($n = 16208$)	0.28059	$2.2e^{-16}$	0.01072

Table 7.21 presents the result of the K-S test that examined the equivalence between the FLS and the HFS of the oxygenation decision system. Additionally, Fig. 7.34 provides a graphical view with the K-S tests for both systems over the six day period. From Table 7.21 and Fig. 7.34, the K-S test results showed that the FLS and HFS were not equivalent when tested with the Baby 001 dataset. That is, the K-S test statistic, which measured the maximum distance (D) between the FLS and HFS, is greater than (D_α), indicating the outputs produced from FLS and HFS have different distributions.

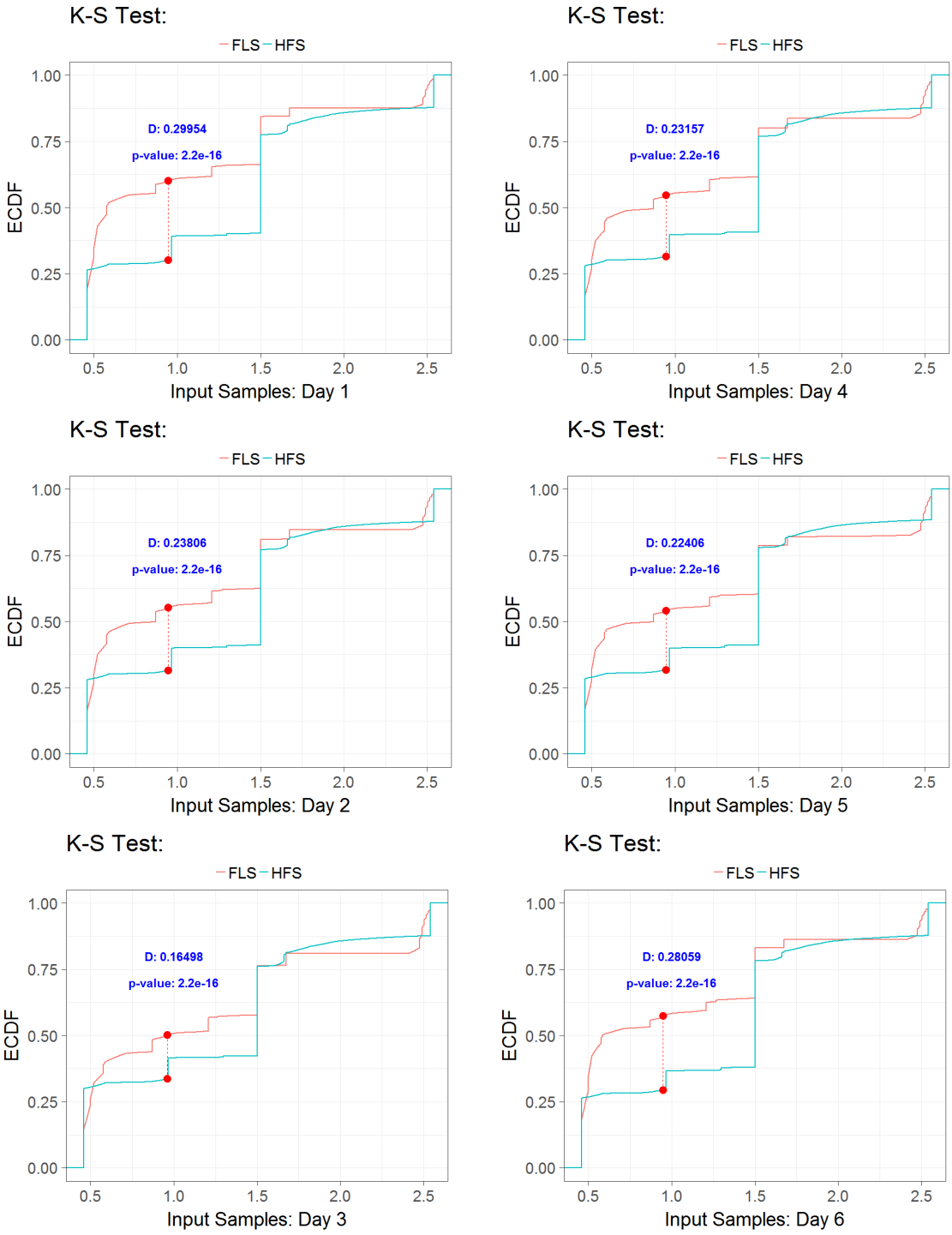


Fig. 7.34 The Kolmogorov–Smirnov test comparing the FLS and HFS Oxygenation Decision systems over six days

7.3.5 Measuring the Interpretability of Oxygenation Decision Systems

In this section, the general index as proposed in (5.2) (see Chapter 5) was used to measure and compare the interpretability of the FLS and HFS for NICU Oxygenation Decision. Following the results obtained in Chapter 5, the *min* aggregation operator together with the *increasing weight* layer-weighted strategy, were used in the general index. Now, (5.2) becomes:

$$H_{\min} = \sum_{j=1}^{q} \left(l_j \sum_{k=1}^{s_j} \min(E_{jk}) \right) \tag{7.1}$$

Table 7.22 A summary of interpretability of Oxygenation decision systems using $H_{\min}$

Oxygenation decision systems	Fuzzy Index							
	NR	NP	a	a	a	term	NC	Index
FLS	54	270	0	0	54	2.4	3	0.1941
$H_{\min}$								**0.1941**
HFS:								
FLS_1	18	72	0	0	18	2.25	4	0.4962
FLS_2	18	72	0	0	18	2.25	3	0.4472
FLS_3	7	14	0	7	0	3.5	3	0.5129
$H_{\min}$								**0.4910**

Table 7.22 shows the interpretability values for the FLS and HFS systems, which was calculated using the $H_{\min}$ as in (7.1). As can be seen in Table 7.22, the computed interpretability value ($H_{\min}$) produce a higher value for the HFS than the FLS. This indicates that, as hoped, the HFS is *more interpretable* than the FLS for the NICU Oxygenation Decision system.

7.4 Discussion

In this chapter, an interpretable HFS for an oxygenation decision-making system to be used within a NICU is designed using the proposed design guidelines framework. Then the systems for the HFS together with a flat system (FLS) were created using the FuzzyToolkitUoN

which utilises the R Programming language. Next, both systems were practically illustrated using a real premature baby dataset, and the results were observed.

The findings do not show that an HFS is better than an FLS. Conversely, it cannot be concluded that the FLS is better than the HFS, as both systems produce different results when applied to a real dataset of a premature baby. A possible explanation for this finding is that the HFS and FLS have a different topologies, including a different number of layers, subsystems and rules. This case study has also shown that it is challenging to develop a system for this complex medical application. Based on these observations, it also cannot be decided whether the FLS or HFS system is the best to represent this system.

However, the HFS design is a promising choice to represent this complex medical application, providing a clear and understandable model. This is because the HFS was designed with the involvement of human experts. To further explore this, the interpretability of FLS and HFS for oxygenation decision system were measured and compared using the H index ($H_{\min}$). The interpretability index showed that the HFS produced *higher* interpretability value than FLS, indicating the HFS is *more interpretable* than the FLS for this application. Also, the HFS can be claimed to be a more interpretable model for the oxygenation decision due to: (i) the HFS has a more intuitive structure when HR Class and SpO_2 class were classified and then fed to the final subsystem, producing meaningful intermediate variables that offer additional information to the clinician; (ii) the HFS has a clear meaning and high semantic interpretability for all inputs, intermediate and output variables, aided by including the domain expert in the design process; (iii) the HFS has fewer total rules (43) in the subsystems ($FLS_1 = 18$ rules, $FLS_2 = 18$, $FLS_3 = 7$) compared to the number of rules (54) in the flat FLS; and (iv) the use of the proposed design guidelines framework made it easy to build.

There are some limitations in the interpretable HFS design for oxygenation decision using the proposed design guidelines framework that need to be addressed in the future. The limitations include: (i) the study should involve a larger number of domain experts in order to get better information for the complex medical problem; (ii) the proposed design guidelines framework does not capture all aspects of interpretability and complexity, e.g. the semantic meaning of fuzzy sets and intermediate variables in HFSs; and (iii) although the interpretable HFS involved human experts in the process of designing it, the model still should be validated by the other human experts who are not associated with the design process.

7.5 Summary

This chapter presented a design guidelines framework for interpretable HFSs. The first section of this chapter provided a brief discussion as to the motivations and background of related work and theories. Then, a design guidelines framework that has five key guidelines to build an interpretable HFS was introduced. These key design guidelines were composed based on the measure of interpretability and complexity in Chapters 5 and 6. Further, a real case study was used to demonstrate the proposed design guidelines framework, specifically to design an interpretable HFS to determine oxygenation decisions for the NICU at QMC Hospital. Later, a practical illustration of this interpretable HFSs was then applied with a real premature baby dataset, and the results were observed. A practical illustration of a flat system (FLS) was also made with the same dataset.

Although it was found that the NICU Oxygenation Decision FLS and HFS produced different results, no clear indication is given as whether the FLS or HFS is better for the oxygenation decision system; however, it can be said that the HFS shows promise in representing this complex medical problem. Indeed, as reported in Section 7.3.5, the computed $H_{\min}$ gave a *higher* interpretability value for the HFS than FLS, indicating that the HFS is indeed *more interpretable* than the FLS. Furthermore, informally, the HFS design was found to be more interpretable to the clinicians, due to providing meaningful information within the system itself (including its topology structure, fewer rules and meaningful intermediate variables), thereby intuitively offering a more understandable decision support model. Hence, this also indicates that the design guidelines framework provides the ability to efficiently design an interpretable HFS. Thus, the fifth and sixth objectives of this thesis are accomplished.

Chapter 8

Conclusions

8.1 Introduction

The research begins with the investigation of interpretability assessment in HFSs. Since interpretability is very subjective because it depends on how humans view it, then the research focuses on capturing the subjective view of interpretability from the human perspective. Next, the research incorporates the subjective view from humans to shape the development of a interpretability measurement of HFSs using a participatory design approach. Later, since it is found that complexity is a key component of interpretability of HFSs, the research next focused on obtaining a comprehensive assessment to measure complexity in HFSs. Finally, a design guidelines framework for interpretable HFSs is developed. The framework is built based on the proposed measurement of interpretability and complexity of HFSs, demonstrated using a real-world case study in the medical domain.

8.2 Thesis summary

As discussed in Chapter 1, the main aim of this thesis is to develop a design guidelines framework for interpretable hierarchical fuzzy systems. To reach this aim, the following objectives were identified:

(i). To investigate the assessment of interpretability of HFSs.

(ii). To establish standard methodologies for capturing perceptions of the interpretability and complexity of HFSs from users.

(iii). To guide the measurement of interpretability of HFSs.

(iv). To explore the measurement of complexity in HFSs.

(v). To build a design guidelines framework for interpretable HFSs.

(vi). To demonstrate the design guidelines framework for interpretable HFSs in a real-world example.

This thesis began with the introduction chapter (Chapter 1) which provided the background and motivations of the work in this study. The aim and objectives of this thesis were highlighted clearly. Next, the contribution of the thesis were outlined. The first chapter ended with the organisation of the thesis.

In Chapter 2, some of the related studies conducted by previous researchers are reviewed. The key research areas include the state of the art of the fuzzy logic systems, hierarchical fuzzy systems, interpretability, complexity and the case study of neonatal monitoring.

In Chapter 3, the research in this thesis began with an exploratory study into the measurement of interpretabilty of fuzzy systems and considered two essential aspects, namely complexity-based interpretability and semantic-based interpretability. However, both aspects of measurement do not take into account the structure of fuzzy systems, i.e., when it is hierarchical (HFS). Therefore, this study investigated the measurement of interpretability of HFSs that have multiple subsystems, layers and varied topologies. In the end, an initial index for measuring the interpretability of HFSs was presented and thus addressed the first objective of this thesis.

Chapter 4 presented an exploratory user study on the perception of interpretability in HFSs. The use of a user study as a means to obtain and assess the interpretability and also complexity in HFSs is the highlight of this chapter. This is due to the fact that such a study has never been used in this field previously. Later, this chapter elaborated on the analysis of the relationship between interpretability and complexity of HFSs, including the descriptive analysis of the survey data, analysis of individual differences in perceived interpretability and complexity, and statistical analysis of the interpretability and complexity. Thus, the second objective of this thesis was achieved.

In order to address the third objective of this thesis, Chapter 5 presented a participatory design approach to guide the initial index of interpretability of HFSs (from Chapter 3). This approach enabled the subjective views of a range of users, to be taken into account in shaping the measurement of interpretability of HFSs. At the end of this study, a demonstration and evaluation using the participatory design approach to configure the measurement of interpretability for HFSs was presented. The finding suggested the use of the *min* aggregation

strategy and *increasing weight* layer-weighting strategy, within the general H index for capturing the interpretability of HFSs.

Chapter 6 studied in detail the complexity of fuzzy systems, particularly HFSs. In FLSs, complexity is expressed by the number of rules, variables, and fuzzy terms – rule-based complexity. Several studies have used indices, e.g., the number of rules, to measure the complexity of FLSs. However, none of the studies considered the structure of HFSs, i.e., multiple subsystems, layers and varied topologies which may also affect the complexity of HFSs. Therefore, this chapter presented an approach to obtain a comprehensive measurement of complexity in HFSs by combining the rule-base complexity with the structural complexity. Therefore, the fourth objective of this thesis was achieved.

Chapter 7 addressed both the fifth and sixth objectives of this thesis. It presented a design guidelines framework for interpretable HFSs. The framework was built based on the measurement index of interpretability and complexity that was presented in Chapters 5 and 6 respectively. The framework consists of five key design guidelines for building interpretable HFS. Subsequently, a demonstration of the framework to guide the designing of an interpretable HFS in the real-world was presented. Specifically, the demonstration focuses on the design an interpretable HFS for supporting decision-making for oxygenation administration in a neonatal intensive care unit (NICU).

In the next section, the main contributions of this thesis are reported.

8.3 Contributions

The study of interpretability, particularly in the context of hierarchical fuzzy systems is an important area, which is likely to gain interest as it has clear relevance to explainable AI (XAI). This thesis has resulted in the following contributions:

1. **An interpretability assessment of HFSs.**
   This is the first work that undertake a formal assessment to measure the interpretability of HFSs which has been built based on the challenges arising from the structure of HFSs, i.e., multiple subsystems, layers and varied topologies. Therefore, this thesis makes a major contribution to the knowledge of interpretability assessment.

2. **A novel method in capturing users' perception on the interpretability and complexity of HFSs.**
   Due to the subjective nature of the interpretability, it is not an easy task to get the users' perception, particularly of interpretability, and there exists a lack of research

addressing the issue. Thus, this method of designing the user study is a significant contribution to the knowledge of interpretability in fuzzy systems.

3. **A novel analysis of interpretability and complexity of HFSs using a user study.**
   This is the first time that the user study has been used to assess the relation between interpretability and complexity in HFSs. This analysis adds insight into the new knowledge of the relation between interpretability and complexity particularly in HFSs from the user's perspective.

4. **The development of an interpretability measurement of HFSs using a participatory design approach.**
   A participatory design approach is utilised to guide the measurement of interpretability. This approach enables the subjective views of a range of users to be taken into account in shaping the measurement of interpretability of HFSs.

5. **A complexity assessment of HFSs.**
   This thesis contributes to the current debate of complexity measurement with the introduction of a combined rule-based and structural complexity index in order to establish a comprehensive measurement of complexity in HFSs.

6. **A novel design guidelines framework for interpretable HFSs.**
   The framework is based on the measurement index of interpretability and complexity of HFSs consisting of five key design guidelines for building interpretable HFS.

7. **A design for an interpretable HFS to Oxygenation decision at NICU.**
   A demonstration of the framework to guide the designing of an interpretable HFS on real-world is presented. Specifically, the demonstration focuses on the design of interpretable HFS for a neonatal intensive care unit (NICU).

In the next section, the future works of this thesis are explained.

8.4 Future work

Several areas can be further developed in future, including as follows:

1. While the proposed interpretability assessment is promising, it does not capture all aspects of the interpretability of HFSs. Thus, future work should explore other aspects of the interpretability of HFSs such as the semantic interpretability of fuzzy sets and intermediate variables of HFSs.

2. While the user study provides intriguing insight into the relationship between interpretability and complexity as perceived by people (users, participants, etc.), more research on this topic needs to be undertaken, with a larger sample size, in order to develop a fine-grained understanding of the relationship between interpretability and complexity of design when creating HFSs.

3. Future work should focus on conducting more experiments involving different case studies with more complex and varied hierarchical systems, and a wider set of people from both within the fuzzy community and from other communities in the participatory design approach. In doing so, it may be possible to gain further insight into different configurations of the proposed index, in order to ultimately gain a deeper understanding of the interpretability of HFSs.

4. Future work could focus on exploring other aspects of complexity, including the semantic complexity of fuzzy sets, intermediate output and the logical complexity of the rules. This may gain further insight into the complexity of HFSs.

8.5 Dissemination

The research work reported in this thesis has been used in various conference and journal papers as well as several internal and international talks. What follows is a list of publications and presentations derived from this work, together with a reference to the chapter in which the topic is covered.

8.5.1 Journal Papers

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D., *Towards a Framework for Capturing Interpretability of Hierarchical Fuzzy Systems — A Participatory Design Approach*, submitted to IEEE Transactions on Fuzzy Systems (Under Review). (Chapter 5)

8.5.2 Conference Papers

- Razak, T. R., Garibaldi, J. M., & Wagner, C.(2019). *A Measure of Structural Complexity of Hierarchical Fuzzy Systems Adapted from Sofware Engineering*, In 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE. (Chapter 6)

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D. (2018). *Interpretability and Complexity of Design in the Creation of Fuzzy Logic Systems — A User Study* In 2018 IEEE Symposium Series on Computational Intelligence (IEEE SSCI) (pp. 420–426). IEEE. (Chapter 4)

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D. (2017). *Interpretability indices for hierarchical fuzzy systems.* In 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1–6). IEEE. (Chapter 3)

8.5.3 Presentations

- Razak, T. R., Garibaldi, J. M., & Soria, D., *Analysis on Preliminary Neonatal Data,* (oral presentation) at a meeting with Medical Expert; Dr Don Sharkey, C32, School of Computer Science, University of Nottingham, UK, 25th May 2016.

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D., *Interpretability indices for hierarchical fuzzy systems,* (oral presentation) at Intelligent Modelling and Analysis Seminar – Fuzz-IEEE Practice, School of Computer Science, University of Nottingham, UK, 27th June 2017.

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D., *Interpretability indices for hierarchical fuzzy systems,* (oral presentation) at 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, Italy, 12th July 2017.

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D., *Interpretability of Fuzzy Logic System Survey,* (oral presentation) at Weekly Research Catchups, Lab for Uncertainty in Data and Decision Making Research Group, School of Computer Science, University of Nottingham, UK, 11th August 2017.

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D., *OxyGUI System: Data Acquisition,* (oral presentation) at Weekly Research Catchups, Lab for Uncertainty in Data and Decision Making Research Group, School of Computer Science, University of Nottingham, UK, 20th October 2017.

- Razak, T. R., Garibaldi, J. M., Wagner, C., Pourabdollah, A., & Soria, D., *Interpretability and Complexity of Design in the Creation of Fuzzy Logic Systems – A User Study,* (oral presentation) at 2018 IEEE Symposium Series on Computational Intelligence (IEEE-SSCI), Bengaluru, India, 18th November 2018.

- Razak, T. R., Garibaldi, J. M., & Wagner, C., *A Measure of Structural Complexity of Hierarchical Fuzzy Systems Adapted from Software Engineering,*(oral presentation) at 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, Louisiana, USA, 23rd June 2019.

References

[1] J.M. Alonso, L. Magdalena, and S. Guillaume. HILK: A new methodology for designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. *Int. J. Intell. Syst.*, 23(7):761–794, jul 2008.

[2] D. Nauck and R. Kruse. Obtaining interpretable fuzzy classification rules from medical data. *Artif. Intell. Med*, 16(2):149–169, 1999.

[3] R. Mikut, J. Jäkel, and L. Gröll. Interpretability issues in data-based learning of fuzzy systems. *Fuzzy Sets Syst.*, 150:179–197, 2005.

[4] C. Mencar and A.M. Fanelli. Interpretability constraints for fuzzy information granulation. *Inf. Sci.*, 178(24):4585–4618, 2008.

[5] A Giboin. ML comprehensibility and KBS explanation: Stating the problem collaboratively. In *Proceedings of IJCAI-95 Workshop on Machine Learning and Comprehensibility*, pages 1–12, 1995.

[6] R. Nisbet, G. Miner, and J. Elder IV. *Handbook of statistical analysis and data mining applications*. Academic Press, 2009.

[7] G. Valdes, J. M. Luna, E. Eaton, C. B. Simone, L. H. Ungar, and T. D. Solberg. MediBoost: a Patient Stratification Tool for Interpretable Decision Making in the Era of Precision Medicine. *Scientific reports*, 6(August):37854, 2016.

[8] L.A. Zadeh. Fuzzy sets. *Inf. and Control*, 8(3):338–353, jun 1965.

[9] A. Bárdossy, A. Blinowska, W. Kuzmicz, J. Ollitrault, M. Lewandowski, A. Przybylski, and Z. Jaworski. Fuzzy logic-based diagnostic algorithm for implantable cardioverter defibrillators. *Artif. Intell. Med*, 60(2):113–21, feb 2014.

[10] A. Gad and M. Farooq. Application of Fuzzy Logic in Engineering problems. In *Proc. IECON*, volume 3, pages 2044–2049. IEEE, 2001.

[11] O.W. Samuel, M.O. Omisore, and B.A. Ojokoh. A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever. *Expert Systems with Applications*, 40(10):4164–4171, aug 2013.

[12] W. Pedrycz. Fuzzy sets in pattern recognition: Methodology and methods. *Pattern Recognition*, 23(1-2):121–146, jan 1990.

[13] T. Furuhashi and T. Suzuki. On interpretability of fuzzy models based on conciseness measure. In *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*, volume 1, pages 284–287. IEEE, 2001.

[14] S.-M. Zhou and J.Q. Gan. Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets Syst.*, 159(23):3091–3131, dec 2008.

[15] J. M. Alonso, L. Magdalena, and G. González-Rodríguez. Looking for a good fuzzy system interpretability index: An experimental approach. *Int. J. Approx. Reas.*, 51(1):115–134, dec 2009.

[16] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning—I. *Inf. Sci.*, 8(3):199–249, jan 1975.

[17] E. H. Mamdani. Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis. *IEEE Trans. on Computers*, 12(12):1182–1191, 1977.

[18] M. Setnes, R. Babuska, and H.B. Verbruggen. Rule-based modeling: precision and transparency. *IEEE Trans. Syst., Man, and Cybern. - Part C*, 28(1):165–169, 1998.

[19] J. M. Alonso, O. Cordón, A. Quirin, and L. Magdalena. Analyzing interpretability of fuzzy rule - based systems by means of fuzzy inference - grams. In *World Congress on Soft Computing*, 2011.

[20] S.-M. Zhou and J.Q. Gan. Constructing accurate and parsimonious fuzzy models with distinguishable fuzzy sets based on an entropy measure. *Fuzzy Sets and Systems*, 157:1057–1074, 2006.

[21] D.D. Nauck. Measuring interpretability in rule-based classification systems. In *Proc. 12th IEEE Int. Conf. Fuzzy Syst.*, pages 196–201. IEEE, 2003.

[22] Y. Jin. Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement. *IEEE Trans. Fuzzy Syst.*, 8(2):212–221, apr 2000.

[23] M. J. Gacto, R. Alcalá, and F. Herrera. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. *Inf. Sci.*, 181(20):4340–4360, 2011.

[24] J. M. Alonso, S. Guillaume, and L. Magdalena. A hierarchical fuzzy system for assessing interpretability of linguistic knowledge bases in classification problems. In *Proc. PIMU*, pages 348–355, 2006.

[25] C. Mencar, C. Castiello, R. Cannone, and A. M. Fanelli. Interpretability assessment of fuzzy knowledge bases: A cointension based approach. *Int. J. Approx. Reas.*, 52:501–518, 2011.

[26] D. P. Pancho, J. M. Alonso, and J. Alcala-Fdez. A new fingram-based software tool for visual representation and analysis of fuzzy association rules. In *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–7. IEEE, jul 2013.

[27] H. Ishibuchi and Y. Nojima. Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning. *Int. J. Approx. Reas.*, 44(1):4–31, jan 2007.

[28] S. Guillaume and B. Charnomordic. A new method for inducing a set of interpretable fuzzy partitions and fuzzy inference systems from data. In *Interpretability Issues in Fuzzy Modeling*, pages 148–175. Springer Berlin Heidelberg, 2003.

[29] M. Setnes and R. Babuska. Rule base reduction: some comments on the use of orthogonal transforms. *IEEE Trans. Syst., Man, and Cybern. - Part C*, 31(2):199–206, may 2001.

[30] J. Yen and L. Wang. Simplifying fuzzy rule-based models using orthogonal transformation methods. *IEEE Trans. Syst., Man, and Cybern. - Part B*, 29(1):13–24, 1999.

[31] S.-M. Zhou, J.M. Garibaldi, R.I. John, and F. Chiclana. On Constructing Parsimonious Type-2 Fuzzy Logic Systems via Influential Rule Selection. *IEEE Trans. Fuzzy Syst.*, 17(3):654–667, jun 2009.

[32] S. Jin and J. Peng. Towards hierarchical fuzzy rule interpolation. In *Proc. IEEE 14th Int. Conf. Cognitive Inform. & Cognitive Comput. (ICCI*CC)*, pages 267–274. IEEE, jul 2015.

[33] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka. Selecting fuzzy if-then rules for classification problems using genetic algorithms. *IEEE Trans. Fuzzy Syst.*, 3(3):260–270, 1995.

[34] H. Ishibuchi, T. Murata, and I.B. Turksen. Selecting linguistic classification rules by two-objective genetic algorithms. In *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, volume 2, pages 1410–1415. IEEE, 1995.

[35] H. Ishibuchi, T. Murata, and I.B Türkşen. Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. *Fuzzy Sets Syst.*, 89(2):135–150, jul 1997.

[36] H. Ishibuchi and T. Murata. Multi-objective genetic local search algorithm. In *Proceedings of IEEE International Conference on Evolutionary Computation*, pages 119–124. IEEE, 1996.

[37] H. Ishibuchi, T. Nakashima, and T. Murata. Three-objective genetics-based machine learning for linguistic rule extraction. *Information Sciences*, 136(1):109–133, 2001.

[38] O. Cordón, M. J. Del Jesus, F. Herrera, L. Magdalena, and P. Villar. A Multiobjective Genetic Learning Process for joint Feature Selection and Granularity and Contexts Learning in Fuzzy Rule-Based Classification Systems. In *Interpretability issues in fuzzy modeling*, pages 79–99. Springer Berlin Heidelberg, 2003.

[39] L.T. Kóczy and K. Hirota. Approximate reasoning by linear rule interpolation and general approximation. *Int. J. Approx. Reas.*, 9(3):197–225, 1993.

[40] L.T. Koczy and K. Hirota. Size reduction by interpolation in fuzzy rule bases. *IEEE Trans. Syst. Man Cybern. - Part B*, 27(1):14–25, 1997.

[41] Q. Liang and J. M. Mendel. Designing interval type-2 fuzzy logic systems using an SVD-QR method: Rule reduction. *Int. J. Intell. Syst.*, 15(10):939–957, oct 2000.

[42] Y. Jin. Decentralized adaptive fuzzy control of robot manipulators. *IEEE Trans. Syst. Man Cybern. - Part B*, 28(1):47–57, 1998.

[43] G. V. S. Raju, J. Zhou, and R.A. Kisner. Hierarchical fuzzy control. *Int. J. Contr.*, 54(5):1201–1216, nov 1991.

[44] G.V.S. Raju and J. Zhou. Adaptive Hierarchical Fuzzy Controller. *IEEE Trans. Syst., Man, and Cybern.*, 23(4):973–980, 1993.

[45] L-X. Wang. Universal approximation by hierarchical fuzzy systems. *Fuzzy Sets Syst.*, 93(2):223–230, jan 1998.

[46] M-L. Lee, H-Y. Chung, and F-M. Yu. Modeling of hierarchical fuzzy systems. *Fuzzy Sets Syst.*, 138(2):343–361, sep 2003.

[47] L-X. Wang. Analysis and Design of Hierarchical Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*, 7(5):617–624, 1999.

[48] M.G. Joo and J.S. Lee. Hierarchical fuzzy control scheme using structured Takagi-Sugeno type fuzzy inference. In *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*, volume 1, pages 78–83. IEEE, 1999.

[49] M.G. Joo and T. Sudkamp. A Method of Converting a Fuzzy System to a Two-Layered Hierarchical Fuzzy System and Its Run-Time Efficiency. *IEEE Trans. Fuzzy Syst.*, 17(1):93–103, feb 2009.

[50] R. Holve. Rule generation for hierarchical fuzzy systems. In *Proc. Annu.Conf. North American Fuzzy Information Processing*, pages 444–449. IEEE, 1997.

[51] N.K. Masmoudi, C. Rekik, M. Djemel, and N. Derbel. Optimal Control for Discrete Large Scale Nonlinear Systems Using Hierarchical Fuzzy Systems. In *2010 Second Int. Conf. on Mach. Learning and Comput.*, pages 91–95. IEEE, 2010.

[52] A. Ge, Y. Wang, Z. Liu, and P. Jiang. Hierarchical fuzzy control of multivariable systems via semi-tensor product method. In *Proc. of the 33rd Chinese Control Conf.*, pages 4551–4556. IEEE, jul 2014.

[53] V. Cross and T. Sudkamp. Patterns of fuzzy rule-based inference. *Int. J. Approx. Reas.*, 11(3):235–255, oct 1994.

[54] A. D. Benítez and J. Casillas. Multi-objective genetic learning of serial hierarchical fuzzy systems for large-scale problems. *Soft Comput.*, 17(1):165–194, 2013.

[55] H. Ishibuchi, Y. Kaisho, and Y. Nojima. Complexity, interpretability and explanation capability of fuzzy rule-based classifiers. In *Proc. Int. Conf. Fuzzy Systems*, pages 1730–1735. IEEE, aug 2009.

[56] R. Alcalá, J. Ramón Cano, O. Cordón, F. Herrera, P. Villar, and I. Zwir. Linguistic modeling with hierarchical systems of weighted linguistic rules. *Int. J. Approx. Reas.*, 32(2-3):187–215, feb 2003.

[57] V. Torra. A Review of the Construction of Hierarchical Fuzzy Systems. *International Journal of Intelligent Systems*, 17(5):531–543, may 2002.

[58] L.A. Zadeh. The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy Sets and Systems*, 11(1-3):199–227, 1983.

[59] J. Harris. *Fuzzy logic applications in engineering science*, volume 29. Springer Science & Business Media, 2005.

[60] Yie-Tone Chen, Yi-Cheng Jhang, and Ruey-Hsun Liang. A fuzzy-logic based auto-scaling variable step-size MPPT method for PV systems. *Solar Energy*, 126:53–63, mar 2016.

[61] Salama A. Mostafa, Aida Mustapha, Mazin Abed Mohammed, Mohd Sharifuddin Ahmad, and Moamin A. Mahmoud. A fuzzy logic control in adjustable autonomy of a multi-agent system for an automated elderly movement monitoring application. *International Journal of Medical Informatics*, 112:173–184, apr 2018.

[62] A. Waldock and B. Carse. Learning a robot controller using an adaptive hierarchical fuzzy rule-based system. *Soft Computing*, 20(7):2855–2881, jul 2016.

[63] Lotfi A. Zadeh. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Trans. Syst., Man, and Cybern.*, SMC-3(1):28–44, 1973.

[64] J.M. Mendel. *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Prentice Hall PTR, 2001.

[65] O.H. Choon and S.L. Tilahun. Integrating fuzzy preference in genetic algorithm to solve multiobjective optimization problems. *Far East Journal of Mathematical Sciences*, 55(2):165—-179, 2011.

[66] C. Wagner, S. Miller, and J. M. Garibaldi. A Fuzzy Toolbox for the R Programming Language. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, pages 1185–1192. IEEE, jun 2011.

[67] M. Negnevitsky. *Artificial intelligence: A Guide to Intelligent systems*. Pearson Education, 2005.

[68] J.M. Mendel, R. I. John, and F. Liu. Interval Type-2 Fuzzy Logic Systems Made Simple. *IEEE Transactions on Fuzzy Systems*, 14(6):808–821, dec 2006.

[69] E. H. Mamdani and S. Assilian. An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. *Int. J. Man-Machine Studies*, 7:1–13, 1975.

[70] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15(1):116–132, jan 1985.

[71] M. Mahfouf, M.F Abbod, and D.A Linkens. A survey of fuzzy logic monitoring and control utilisation in medicine. *Artificial Intelligence in Medicine*, 21(1-3):27–42, jan 2001.

[72] L.F.C. Nascimento and N.R.S. Ortega. Fuzzy linguistic model for evaluating the risk of neonatal death. *Revista de Saúde Pública*, 36(6):686–692, dec 2002.

[73] C. A. Peña-Reyes and M. Sipper. A fuzzy-genetic approach to breast cancer diagnosis. *Artificial Intelligence in Medicine*, 17(2):131–155, oct 1999.

[74] M. Phillips, R.N. Cataneo, B.A. Ditkoff, P. Fisher, J. Greenberg, R. Gunawardena, C.S. Kwon, O. Tietje, and C. Wong. Prediction of breast cancer using volatile biomarkers in the breath. *Breast cancer research and treatment*, 99(1):19–21, sep 2006.

[75] M.U. Khan, J. P. Choi, H. Shin, and M. Kim. Predicting breast cancer survivability using fuzzy decision trees for personalized healthcare. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5148–5151, jan 2008.

[76] P.F. Stetson. Determining heart rate from noisy pulse oximeter signals using fuzzy logic. In *The 12th IEEE International Conference on Fuzzy Systems, 2003. FUZZ '03.*, volume 2, pages 1053–1058. IEEE, 2003.

[77] N. Bhatla and K. Jyoti. A Novel Approach for Heart Disease Diagnosis using Data Mining and Fuzzy Logic. *International Journal of Computer Applications*, 54(17):16–21, 2012.

[78] P.K. Anooj. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University - Computer and Information Sciences*, 24(1):27–40, jan 2012.

[79] J. Schneider, N. Bitterlich, H-G. Velcovsky, H. Morr, N. Katz, and E. Eigenbrodt. Fuzzy logic-based tumor-marker profiles improved sensitivity in the diagnosis of lung cancer. *International journal of clinical oncology*, 7(3):145–51, jun 2002.

[80] M. Phillips, N. Altorki, J.H.M. Austin, R.B. Cameron, R.N. Cataneo, J. Greenberg, R. Kloss, R.A. Maxfield, M.I. Munawar, H.I. Pass, A. Rashid, W.N. Rom, and P. Schmitt. Prediction of lung cancer using volatile biomarkers in breath. *Cancer Biomarkers*, 3(2):95–109, jan 2007.

[81] C.C. Samuel, V. Saravanan, and M.R.V. Devi. Lung Nodule Diagnosis from CT Images Using Fuzzy Logic. In *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, volume 3, pages 159–163. IEEE, dec 2007.

[82] A.M. Badawi, A.S. Derbala, and A-B.M. Youssef. Fuzzy logic algorithm for quantitative tissue characterization of diffuse liver diseases from ultrasound images. *International Journal of Medical Informatics*, 55(2):135–147, aug 1999.

[83] K. Polat, S. Şahan, H. Kodaz, and S. Güneş. Breast cancer and liver disorders classification using artificial immune recognition system (AIRS) with performance evaluation by fuzzy resource allocation mechanism. *Expert Systems with Applications*, 32(1):172–183, jan 2007.

[84] S. Shiomi, T. Kuroki, M. Kuriyama, H. Morikawa, and K. Masaki. Diagnosis of chronic liver disease from liver scintiscans by artificial neural networks. *Annals of Nuclear Medicine*, 11(2):75–80, 1997.

[85] M.F. Ganji and M.S. Abadeh. Using fuzzy ant colony optimization for diagnosis of diabetes disease. In *2010 18th Iranian Conference on Electrical Engineering*, pages 501–505. IEEE, may 2010.

[86] M.F. Ganji and M.S. Abadeh. A fuzzy classification system based on Ant Colony Optimization for diabetes disease diagnosis. *Expert Systems with Applications*, 38(12):14650–14659, nov 2011.

[87] C-S. Lee and M-H. Wang. A fuzzy expert system for diabetes decision support application. *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, 41(1):139–53, feb 2011.

[88] T. Ozen and J.M. Garibaldi. Investigating Adaptation in Type-2 Fuzzy Logic Systems Applied to Umbilical Acid-Base Assessment. In *Proceedings of 2003 European Symposium on Intelligent Technologies*, pages 289–294, 2003.

[89] J.F. Skinner, J.M. Garibaldi, J. Curnow, and E.C. Ifeachor. Intelligent fetal heart rate analysis. In *First International Conference on Advances in Medical Signal and Information Processing*, pages 14–21. IEEE, 2000.

[90] R.I. John and P.R. Innocent. Modeling Uncertainty in Clinical Diagnosis Using Fuzzy Logic. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 35(6):1340–1350, dec 2005.

[91] R.E. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961.

[92] X.-J. Zeng and J.A. Keane. Approximation Capabilities of Hierarchical Fuzzy Systems. *IEEE Trans. Fuzzy Syst.*, 13(5):659–672, oct 2005.

[93] M.G. Joo and J.S. Lee. A class of hierarchical fuzzy systems with constraints on the fuzzy rules. *IEEE Transactions on Fuzzy Systems*, 13(2):194–203, apr 2005.

[94] R.R. Yager. On the Construction of Hierarchical Fuzzy Systems Models. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 28(1):55–66, 1998.

[95] R.J.G.B. Campello and W. Caradori do Amaral. Hierarchical fuzzy relational models: linguistic interpretation and universal approximation. *IEEE Trans. Fuzzy Syst.*, 14(3):446–453, jun 2006.

[96] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179–188, sep 1936.

[97] L. Magdalena, F. Monasterio, and C. Rivero. Hierarchical decomposition of fuzzy controllers based on meta-knowledge. *Mathware & soft comput.*, 5:291–304, 1998.

[98] H. Hagras. A Hierarchical Type-2 Fuzzy Logic Control Architecture for Autonomous Mobile Robots. *IEEE Trans. Fuzzy Syst.*, 12(4):524–539, aug 2004.

[99] J. Hauser, S. Sastry, and P. Kokotovic. Nonlinear control via approximate input-output linearization: the ball and beam example. *IEEE Transactions on Automatic Control*, 37(3):392–398, mar 1992.

[100] W. Rattasiri and S.K. Halgamuge. Computational complexity of hierarchical fuzzy systems. In *PeachFuzz 2000. 19th International Conference of the North American Fuzzy Information Processing Society - NAFIPS (Cat. No.00TH8500)*, pages 383–387. IEEE, 2000.

[101] M. Brown, K.M. Bossley, D.J. Mills, and C.J. Harris. High dimensional neurofuzzy systems: overcoming the curse of dimensionality. In *Proceedings of 1995 IEEE International Conference on Fuzzy Systems. The International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium*, volume 4, pages 2139–2146. IEEE, 1995.

[102] P. Salgado and J.B. Cunha. Greenhouse climate hierarchical fuzzy modelling. *Control Engineering Practice*, 13(5):613–628, 2005.

[103] P. K. Shukla and S.P. P. Tripathi. A Survey on Interpretability-Accuracy (I-A) Trade-Off in Evolutionary Fuzzy Systems. In *2011 Fifth International Conference on Genetic and Evolutionary Computing*, pages 97–101. IEEE, aug 2011.

[104] M.R. Delgado, F. Von Zuben, and F. Gomide. Multi-objective decision making: towards improvement of accuracy, interpretability and design autonomy in hierarchical genetic fuzzy systems. In *Proc. Int. Conf. FUZZ-IEEE*, volume 2, pages 1222–1227. IEEE, 2002.

[105] O. Cordón, F. Herrera, and I. Zwir. A hierarchical knowledge-based environment for linguistic modeling: models and iterative methodology. *Fuzzy Sets and Systems*, 138(2):307–341, sep 2003.

[106] O. Cordon, F. Herrera, and I. Zwir. Linguistic modeling by hierarchical systems of linguistic rules. *IEEE Trans. Fuzzy Syst.*, 10(1):2–20, 2002.

[107] J. M. Alonso and L. Magdalena. Special issue on interpretable fuzzy systems. *Information Sciences*, 181:4331–4339, 2011.

[108] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. Interpretability Improvements to Find the Balance Interpretability-Accuracy in Fuzzy Modeling: An Overview. In *Interpretability issues in fuzzy modeling*, pages 3–22. Springer Berlin Heidelberg, 2003.

[109] Interpretability - Wikipedia.

[110] H. Hamad, A. Al-Zaben, and R. Owies. Interpretability and variability of metamodel validation statistics in engineering system design optimization: a practical study. *International Journal for Simulation and Multidisciplinary Design Optimization*, 5:A05, feb 2014.

[111] Wang Xiao. Interpretation. *Boundless Psychology*, 2016.

[112] R. W. Brainard, Ri. J. Campbell, and E. H. Elkin. Design and interpretability of road signs. *Journal of Applied Psychology*, 45(2):130–136, 1961.

[113] P. Geurts, Al. Irrthum, and L. Wehenkel. Supervised learning with decision tree-based methods in computational and systems biology. *Molecular BioSystems*, 5(12):1593, 2009.

[114] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning*, volume 103 of *Springer Texts in Statistics*. Springer New York, New York, NY, 2013.

[115] U. Bodenhofer and P. Bauer. A Formal Model of Interpretability of Linguistic Variables. In *Interpretability issues in fuzzy modeling*, pages 524–545. Springer Berlin Heidelberg, 2003.

[116] K. Balazs and L. T. Koczy. New parameterizable search space narrowing technique for adjusting between accuracy and interpretability in fuzzy systems. In *Proc. IEEE 13th Int. Symp. Comput. Intell. and Inform. (CINTI)*, pages 323–328. IEEE, nov 2012.

[117] J. M. Alonso and L. Magdalena. Combining user's preferences and quality criteria into a new index for guiding the design of fuzzy systems with a good interpretability-accuracy trade-off. In *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*, pages 1–8. IEEE, jul 2010.

[118] A. Botta, B. Lazzerini, F. Marcelloni, and D.C. Stefanescu. Context adaptation of fuzzy systems through a multi-objective evolutionary approach based on a novel interpretability index. *Soft Comput.*, 13(5):437–449, mar 2009.

[119] J. M. Alonso, C. Castiello, and C. Mencar. Interpretability of Fuzzy Systems: Current Research Trends and Prospects. In *Springer Handbook of Computational Intelligence*, pages 219–237. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.

[120] George A Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[121] J.M. Garibaldi, S. Musikasuwan, T. Ozen, and R.I. John. A case study to illustrate the use of non-convex membership functions for linguistic terms. In *2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542)*, volume 3, pages 1403–1408. IEEE, 2004.

[122] A. A. Marquez, F. A. Marquez, and A. Peregrin. A multi-objective evolutionary algorithm with an interpretability improvement mechanism for linguistic fuzzy systems with adaptive defuzzification. In *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*, pages 1–7. IEEE, jul 2010.

[123] D. P. Pancho, J. M. Alonso, and L. Magdalena. Enhancing Fingrams to deal with precise fuzzy systems. *Fuzzy Sets Syst.*, 297:1–25, 2016.

[124] M. Pota, M. Esposito, and G.D. Pietro. Interpretability indexes for Fuzzy classification in cognitive systems. In *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*, pages 24–31. IEEE, jul 2016.

[125] I. MathWorks and W-C. Wang. *Fuzzy Logic Toolbox: for Use with MATLAB: User's Guide*. Mathworks, Incorporated, 1998.

[126] K. J. Ostergaard, W. Wetmore, and J. D. Summers. A Methodology for the Study of the Effects of Communication Method on Design Review Effectiveness. In *Proc. 29th Design Automation Conf., Parts A and B*, volume 2003, pages 383–390. ASME, jan 2003.

[127] C. Mencar, G. Castellano, and A. M. Fanelli. Some fundamental interpretability issues in fuzzy modeling. In *Proceedings of Joint EUSFLAT-LFA*, pages 100–105, 2005.

[128] D.P. Pancho, J.M. Alonso, O. Cordon, A. Quirin, and L. Magdalena. FINGRAMS: Visual Representations of Fuzzy Rule-Based Inference for Expert Analysis of Comprehensibility. *IEEE Trans. Fuzzy Syst.*, 21(6):1133–1149, dec 2013.

[129] A. Riid and E. Rustern. Interpretability improvement of fuzzy systems: Reducing the number of unique singletons in zeroth order Takagi-Sugeno systems. In *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*, pages 1–6. IEEE, jul 2010.

[130] A. R. De Soto. A hierarchical model of a linguistic variable. *Inf. Sci.*, 181:4394–4408, 2011.

[131] H. Ishibuchi. Multiobjective Genetic Fuzzy Systems: Review and Future Research Directions. In *2007 IEEE International Fuzzy Systems Conference*, pages 1–6. IEEE, jun 2007.

[132] Alexander Gegov. *Fuzzy Networks for Complex Systems*, volume 259 of *Studies in Fuzziness and Soft Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[133] David Lane. Hierarchy, Complexity, Society. In *Hierarchy in Natural and Social Sciences*, pages 81–119. Springer-Verlag, Berlin/Heidelberg, 2006.

[134] Steven Johnson. *Emergence: The connected lives of ants, brains, cities, and software*. Simon and Schuster, 2002.

[135] A. Kroll and H. Schulte. Benchmark problems for nonlinear system identification and control using Soft Computing methods: Need and overview. *Applied Soft Comput.*, 25:496–513, dec 2014.

[136] H. Liu, A. Gegov, and M. Cocea. Complexity Control in Rule Based Models for Classification in Machine Learning Context. In *Advan. in Computat. Intell. Syst.*, pages 125–143. Springer, Cham, 2017.

[137] O. Cordón. A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems. *Int. J. Approx. Reas.*, 52(6):894–913, sep 2011.

[138] M. Galende-Hernández, G. I. Sainz-Palmero, and M. J. Fuente-Aparicio. Complexity reduction and interpretability improvement for fuzzy rule systems based on simple interpretability measures and indices by bi-objective evolutionary rule selection. *Soft Comput.*, 16(3):451–470, mar 2012.

[139] T. R. Razak, J. M. Garibaldi, C. Wagner, A. Pourabdollah, and D. Soria. Interpretability and Complexity of Design in the Creation of Fuzzy Logic Systems — A User Study. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 420–426. IEEE, 2018.

[140] Herbert A. Simon. The Architecture of Complexity, 1962.

[141] M.L. Commons, E.J. Trudeau, S.A. Stein, F.A. Richards, and S.R. Krause. Hierarchical Complexity of Tasks Shows the Existence of Developmental Stages. *Developmental Review*, 18(3):237–278, sep 1998.

[142] M.L. Commons. Introduction to the model of hierarchical complexity. *Behavioral Development Bulletin*, 13(1):1–6, 2007.

[143] Enys Mones, Lilla Vicsek, and Tamás Vicsek. Hierarchy Measure for Complex Networks. *PLoS ONE*, 7(3):e33799, mar 2012.

[144] W. G. Bail and M. V. Zelkowitz. Program complexity using Hierarchical Abstract Computers. *Computer Languages*, 13(3-4):109–123, jan 1988.

[145] T.J. McCabe. A Complexity Measure. *IEEE Trans. on Software Engineering*, SE-2(4):308–320, dec 1976.

[146] B. Lopez, F. I. Gamero, J. Coll, N. Mordvaniuk, J. Bassols, E. Bargallo, B. Guarch, and A. Lopez-Bermejo. Intelligent System for Premature Babies Healthcare at Home based on Case-based Reasoning. In *Proceedings Iwbbio 2014: International Work-Conference on Bioinformatics*, pages 1278–1289, 2014.

[147] A.G.S Philip. The evolution of neonatology. *Pediatric research*, 58(4):799–815, oct 2005.

[148] Julie-Anne Quinn, Flor M Munoz, Bernard Gonik, Lourdes Frau, Clare Cutland, Tamala Mallett-Moore, Aimee Kissou, Frederick Wittke, Manoj Das, Tony Nunes, and Others. Preterm birth: Case definition & guidelines for data collection, analysis, and presentation of immunisation safety data. *Vaccine*, 34(49):6047–6056, 2016.

[149] N.M. Nour. Premature delivery and the millennium development goal. *Reviews in obstetrics & gynecology*, 5(2):100–5, 2012.

[150] H. Wang et al. Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 384(9947):957–979, sep 2014.

[151] L. M. Askie, P. Brocklehurst, B. A. Darlow, N. Finer, B. Schmidt, and W. Tarnow-Mordi. NeOProM: Neonatal Oxygenation Prospective Meta-analysis Collaboration study protocol. *BMC Pediatrics*, 11(1):6, dec 2011.

[152] W. Tarnow-Mordi et al. Outcomes of Two Trials of Oxygen-Saturation Targets in Preterm Infants. *New England Journal of Medicine*, 374(8):749–760, feb 2016.

[153] W.A. Carlo et al. Target Ranges of Oxygen Saturation in Extremely Preterm Infants. *New England Journal of Medicine*, 362(21):1959–1969, may 2010.

[154] C. M. Majaesic, R. Jones, I. A. Dinu, M. D. Montgomery, R. S. Sauve, and C. M.T. Robertson. Clinical correlations and pulmonary function at 8 years of age after severe neonatal respiratory failure. *Pediatric Pulmonology*, 42(9):829–837, sep 2007.

[155] C. Lockwood, T. Conroy-Hiller, and T. Page. Vital signs. *JBI Reports*, 2(6):207–230, jul 2004.

[156] W R Mower, C Sachs, E L Nicklin, and L J Baraff. Pulse oximetry as a fifth pediatric vital sign. *Pediatrics*, 99(5):681–6, may 1997.

[157] M. M. Churpek, R. Adhikari, and D. P. Edelson. The value of vital sign trends for detecting clinical deterioration on the wards. *Resuscitation*, 102:1–5, may 2016.

[158] S. Y. Belal, A. F. G. Taktak, A. J. Nevill, and S. A. Spencer. A fuzzy system for detecting distorted plethysmogram pulses in neonates and paediatric patients. *Physiological Measurement*, 22(2):397–412, may 2001.

[159] S. Y. Belal, A. F. G Taktak, A. Nevill, and A. Spencer. An intelligent ventilation and oxygenation management system in neonatal intensive care using fuzzy trend template fitting. *Physiological Measurement*, 26(4):555–570, aug 2005.

[160] M.A.M. Reis, N.R.S. Ortega, and P.S.P. Silveira. Fuzzy expert system in the prediction of neonatal resuscitation. *Brazilian Journal of Medical and Biological Research*, 37(5):755–764, may 2004.

[161] Y. Sun, I. Kohane, and A.R. Stark. Fuzzy logic assisted control of inspired oxygen in ventilated newborn infants. In *Proceedings of the Annual Symposium on Computer Application in Medical Care*, pages 757–761, jan 1994.

[162] K. Yamaguchi, Y. Fujimoto, S. Kobashi, Y. Wakata, R. Ishikura, K. Kuramoto, S. Imawaki, S. Hirota, and Y. Hata. Automated fuzzy logic based skull stripping in neonatal and infantile MR images. In *International Conference on Fuzzy Systems*, pages 1–7. IEEE, jul 2010.

[163] J. Liu, T.M. McKenna, A. Gribok, B. A. Beidleman, W. J. Tharion, and J. Reifman. A fuzzy logic algorithm to assign confidence levels to heart and respiratory rate time series. *Physiological measurement*, 29(1):81–94, jan 2008.

[164] D.M. Velasevic, D.Z. Saletic, and S.Z. Saletic. A fuzzy sets theory application in determining the severity of respiratory failure. *International Journal of Medical Informatics*, 63(1-2):101–107, sep 2001.

[165] P. Salgado. Rule generation for hierarchical collaborative fuzzy system. *Appl. Math. Model.*, 32(7):1159–1178, jul 2008.

[166] A. Waldock, B. Carse, and C. Melhuish. Hierarchical fuzzy rule based systems using an information theoretic approach. *Soft Comput.-A Fusion of Foundations, Methodologies and Applications*, 10(10):867–879, 2006.

[167] A. Pedro Aguiar, J. P. Hespanha, and A. M. Pascoal. Switched seesaw control for the stabilization of underactuated vehicles. *Automatica*, 43(12):1997–2008, dec 2007.

[168] J. Lin, S.-Y. Guo, and J. Chang. Fuzzy coordinator compensation for balancing control of cart–seesaw system. *Journal of Sound and Vibration*, 330(26):6296–6310, dec 2011.

[169] Jonqlan Lin, You Sheng Ding, and Julian Chang. Balancing and swinging-up control for cart-pendulum-seesaw system by decomposed fuzzy coordination control. *Journal of Vibration and Control*, 20(6):925–942, apr 2014.

[170] J. Lin, H. Y. Lai, and J. Chang. Balancing control with neuro-fuzzy approach for electrical cart-seesaw system. In *2016 American Control Conference (ACC)*, pages 3958–3963. IEEE, jul 2016.

[171] Paul Lavrakas. Show Card. In *Encyclopedia of Survey Research Methods*. Sage Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States of America, 2013.

[172] P. Ahlgren, B. Jarneving, and R. Rousseau. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54(6):550–560, apr 2003.

[173] S. Siegel. *Nonparametric Statistics for the Behavioral Sciences*. 1956.

[174] F. Kensing and J. Blomberg. Participatory Design: Issues and Concerns. *Computer Supported Cooperative Work (CSCW)*, 7(3-4):167–185, sep 1998.

[175] F. Olga, H. Ioannis, K. Dimitris, and L. Spires. Implementing participatory design for developing a constructivist e-learning activity. In *2013 24th EAEEIE Annual Conference (EAEEIE 2013)*, pages 157–162. IEEE, may 2013.

[176] C. Sjöberg and T. Timpka. Participatory design of information systems in health care. *Journal of the American Medical Informatics Association : JAMIA*, 5(2):177–83, 1998.

[177] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Syst., Man, and Cybern.*, 18(1):183–190, 1988.

[178] R. Yager. Quantifier guided aggregation using OWA operators. *Int. J. Intell. Syst.*, 11(1):49–73, dec 1996.

[179] R.R. Yager. Families of OWA operators. *Fuzzy Sets Syst.*, 59(2):125–148, oct 1993.

[180] Xiaoqian Zhu, Jianping Li, Dengsheng Wu, Haiyan Wang, and Changzhi Liang. Balancing accuracy, complexity and interpretability in consumer credit decision making: A C-TOPSIS classification approach. *Knowledge-Based Syst.*, 52:258–267, nov 2013.

[181] B. Kosko. Fuzzy cognitive maps. *International Journal of Man-Machine Studies*, 24(1):65–75, jan 1986.

[182] Frank J. Massey and Jr. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253):68–78, mar 1951.

[183] L Magdalena. Designing interpretable Hierarchical Fuzzy Systems. In *2018 IEEE International Conference on Fuzzy Systems*, 2018.

[184] J.M. Alonso and L. Magdalena. A Conceptual Framework for Understanding a Fuzzy System. In *IFSA/EUSFLAT Conf.*, pages 119–124, 2009.

[185] C. Wagner and H. Hagras. Fuzzy Composite Concepts based on human reasoning. In *2010 IEEE International Conference on Information Reuse & Integration*, pages 308–313. IEEE, aug 2010.

[186] C. Wagner, M. Smith, K. Wallace, and A. Pourabdollah. Generating Uncertain Fuzzy Logic Rules from Surveys: Capturing Subjective Relationships between Variables from Human Experts. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 2033–2038. IEEE, oct 2015.

[187] E. W. T. Ngai and F. K. T. Wat. Design and development of a fuzzy expert system for hotel selection. *Omega*, 31(4):275–286, aug 2003.

[188] Ritu Agarwal and Mohan R. Tanniru. Knowledge Acquisition Using Structured Interviewing: An Empirical Investigation. *Journal of Management Information Systems*, 7(1):123–140, jun 1990.

[189] Andy Griffiths, Asbjørn Rachlew, and Asbjørn Rachlew. From interrogation to investigative interviewing. In *The Psychology of Criminal Investigation*, pages 154–178. Routledge, may 2018.

[190] Ioannis Mammas and Demetrios Spandidos. Practicing humility and medical education: Lessons learnt interviewing experts on Paediatric Virology. *Experimental and Therapeutic Medicine*, 18(4):3254–3256, aug 2019.

[191] Arnaud Gagneur, Virginie Gosselin, Julie Bergeron, Anne Farrands, and Geneviève Baron. Development of motivational interviewing skills in immunization (MISI): a questionnaire to assess MI learning, knowledge and skills for vaccination promotion. *Human Vaccines & Immunotherapeutics*, pages 1–7, jun 2019.

[192] V. Larcher and M. F. Hird. Withholding and withdrawing neonatal intensive care. *Current Paediatrics*, 12(6):470–475, dec 2002.

[193] BabyCentre Medical Advisory Board. The neonatal unit, 2012.

[194] Y. Abdel-Rahman, A. Jeremic, and K. Tan. Neonatal heart rate prediction. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society Conference*, pages 4695–4698, 2009.

Appendix A

Interpretability of Fuzzy Logic Systems Survey

Interpretability of Fuzzy Logic Systems Survey

Participant Information Sheet

Introduction

Interpretability is related to the capability to express something in an understandable way i.e. people may say that something is interpretable if they can understand it easily. One of the strengths of fuzzy logic systems (FLSs) is their interpretability which indicates how easily an FLS can be understood by a user.

This survey is designed to assess your opinion about the concept of interpretability of FLSs and hierarchical fuzzy systems (HFSs). The data collected will be anonymously used examine how people perceive interpretability in FLSs and HFSs using Iris flower classification application as an example.

All students, researchers, and academicians are very welcome to participate. The survey should take no longer than 15 minutes to complete.

Data collection consent and storage

Consent will be sought before research takes place. You will be guaranteed complete anonymity, with all recorded data from your participation identified by a number only. You are free to withdraw from the research at any time and have any data related to you destroyed – to withdraw, please contact researcher (either in person or via researcher's contact details below). All of the data collected will be stored in accordance with the Data Protection Act 1998 (DPA) and remains confidential. Only academic researcher directly related to the study will be able to access it.

Please do not hesitate to ask any questions. You may contact the researcher at any time about the research or in relation to your consent, including withdrawal from the study:

Researcher	Tajul Razak
Email	tajul.razak@nottingham.ac.uk
Address	B38, School of Computer Science, Jubilee Campus, The University of Nottingham, UK.

Participant Consent Form

Participant name: _____

Title of study: Interpretability Indices of Hierarchical Fuzzy Systems (Interpretability of Fuzzy Logic Systems Survey)

Name of researcher: Tajul Razak (PhD Student)
Address: B38, School of Computer Science, Jubilee Campus, University of Nottingham, UK.
Email: tajul.razak@nottingham.ac.uk

Name of supervisors:
Prof. Jon Garibaldi (jon.garibaldi@nottingham.ac.uk),
Dr Christian Wagner (christian.wagner@nottingham.ac.uk),
Dr Amir Pourabdollah (amir.pourabdollah@nottingham.ac.uk)

Please read and tick (/) each statement below:

	By completing this form, I confirm that I agree to take part in this research and I have read and understood the information sheet provided
	I understand that the survey will gather my opinion and I agree for this data to be used in anonymised form for research purposes.
	I understand that I can withdraw at any time, by contacting the researcher (either in person or via the researcher's contact details) and the data I have provided will be erased.
	I understand that any information that has already been anonymised and/or used in publications cannot be withdrawn.
	I confirm that I am over the age of 18.

If you have read the information sheet and agree to participate in the study, please complete this consent form.

I give permission for the data collected to be used in subsequent associated research. Yes / No Initial:

Please complete the details below to consent to this study:

Signature: _____ Date: _____

Interpretability of Fuzzy Logic Systems Survey

Introduction

This survey is designed to assess your opinion about the concept of interpretability of FLSs and hierarchical fuzzy systems (HFSs). The purpose of the survey is to examine how people perceive interpretability in FLSs and HFSs using Iris flower classification application as an example.

For this example, the four input variables ($n = 4$), i.e., sepal length, sepal width, petal length and petal width are modeled with two fuzzy sets ($m = 2$) and three fuzzy sets ($m = 3$). The output is a Flower; s*entosa, versicolor* and *virginica*.

Instructions

This survey consists of 9 questions: question 1 to 5 are the ranking-type questions, and question 6 to 9 are multiple choice questions. A set of cards that include an Iris flower classification application will be referred in answering the questions.

For the ranking-type questions, please rank the options from the most (1) to the least preferable (3). You may provide an equal rank if you feel there is no difference between options.

For example,

Please rank the ice cream from the most (1) to the least preferable (3). You may provide an equal rank if you feel there is no difference between options.

Strawberry	**1**		*Strawberry*	**3**
Chocolate	**3**	*OR*	*Chocolate*	**1**
Vanilla	**2**		*Vanilla*	**1**

Interpretability of Fuzzy Logic Systems Survey

Please provide your **opinion** by answering the following **questions**.

1. Please rank the topologies from simpler (1) to the least (3). You may provide an equal rank if you feel there is no difference between options.

   FLS ☐

   Parallel HFS ☐

   Serial HFS ☐

   Please explain your answer.

2. Please rank the following systems in order of **interpretability** from 1 to 3, where 1 is more interpretable, and 3 is less interpretable. You may provide an equal rank if you feel there is no difference between options.

   F-2 ☐

   P-2 ☐

   S-2 ☐

   Please explain your answer.

3. Please rank the following systems in term of **design complexity** from 1 to 3, where 1 is less complex, and 3 is more complex. You may provide an equal rank if you feel there is no difference between options.

F-2 ☐

P-2 ☐

S-2 ☐

Please explain your answer.

4. Please rank the following systems in order of **interpretability** from 1 to 3, where 1 is more interpretable, and 3 is less interpretable. You may provide an equal rank if you feel there is no difference between options.

F-3 ☐

P-3 ☐

S-3 ☐

Please explain your answer.

5. Please rank the following systems in term of **design complexity** from 1 to 3, where 1 is less complex, and 3 is more complex. You may provide an equal rank if you feel there is no difference between options.

F-3 ☐

P-3 ☐

S-3 ☐

Please explain your answer.

6. Consider P-3 and S-3. Which of these HFS systems would you prefer to design?

☐ P-3 ☐ S-3

Please explain your answer.

7. Consider F-2 and P-2. Does the reduction of 16 to 12 rules in system F-2 to P-2 make it worth to choose HFS?

☐ Yes ☐ No

Please explain your answer.

8. Consider F-3 and P-3. Does the reduction of 81 to 27 rules in system F-3 to P-3 make it worth to choose HFS?

☐ Yes ☐ No

Please explain your answer.

9. Consider P-3 and S-3. Which of HFS system is more interpretable to you?

☐ P-3 ☐ S-3

Please explain your answer.

Thank you for your participation, it is very much appreciated.

FLS

Fig. 1-A shows the fuzzy logic system (FLS) topology for Iris classification.

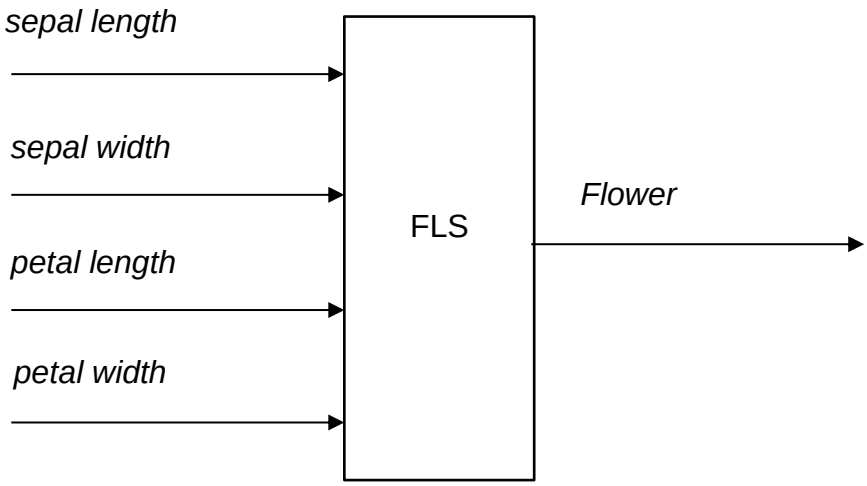


Figure 1-A: Fuzzy Logic System

F-2

Fig. 1 shows the fuzzy logic system (FLS) topology of Iris classification.

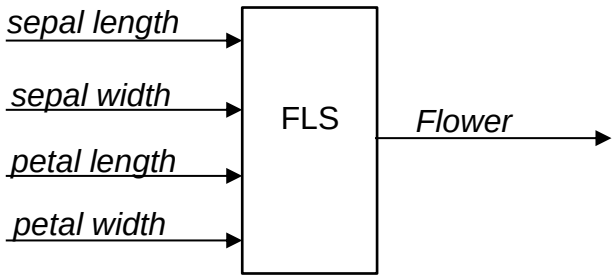


Figure 1: Fuzzy Logic System

The input variables are modeled with two fuzzy sets ($m = 2$); *small* and *large* as shown in Fig. 2 ~ 5 respectively.

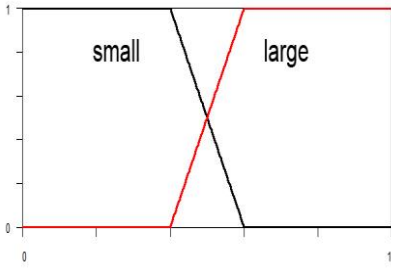


Figure 2*: sepal length*

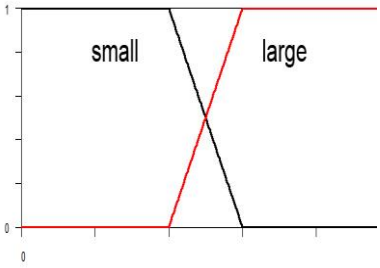


Figure 3: *sepal width*

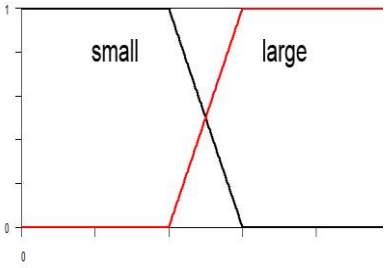


Figure 4: *petal length*

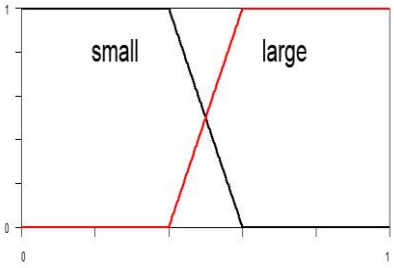


Figure 5: *petal width*

The total number of rules (R) for this complete FLS with fully specified rule base (using the "AND" logical connective), is $R = m^n = 2^4 = 16$ and can be illustrated as follows;

R1: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa*
R2: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *large* THEN *Flower* is *Versicolor*
R3: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *large* AND *petal width* is *small* THEN *Flower* is *Versicolor*
R4: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*
R5: IF *sepal length* is *small* AND *sepal width* is *large* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa*
R6: IF *sepal length* is *small* AND *sepal width* is *large* AND *petal length* is *small* AND *petal width* is *large* THEN *Flower* is *Versicolor*
R7: IF *sepal length* is *small* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *small* THEN *Flower* is *Virginica*
R8: IF *sepal length* is *small* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*
R9: IF *sepal length* is *large* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa*
R10: IF *sepal length* is *large l* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *large* THEN *Flower* is *Versicolor*
R11: IF *sepal length* is *large* AND *sepal width* is *small* AND *petal length* is *large* AND *petal width* is *small* THEN *Flower* is *Versicolor*
R12: IF *sepal length* is *large* AND *sepal width* is *small* AND *petal length* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*
R13: IF *sepal length* is *large* AND *sepal width* is *large* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Versicolor*
R14: IF *sepal length* is *large* AND *sepal width* is *large* AND *petal length* is *small* AND *petal width* is *large* THEN *Flower* is *Versicolor*
R15: IF *sepal length* is *large* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *small* THEN *Flower* is *Virginica*
R16: IF *sepal length* is *large* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*

F-3

Fig. 1 shows the fuzzy logic system (FLS) topology of Iris classification.

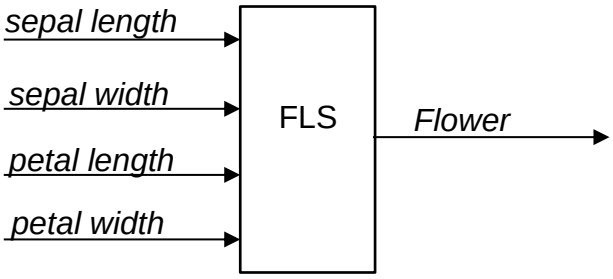


Figure 1: Fuzzy Logic System

The input variables are modeled with three fuzzy sets ($m = 3$); *small, medium* and *large* as shown in Fig. 2 ~ 5 respectively.

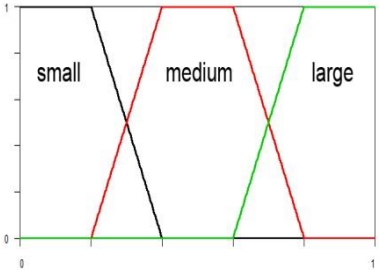


Figure 2*: sepal length*

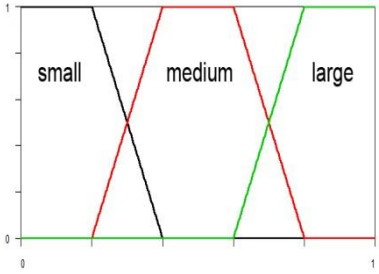


Figure 3: *sepal width*

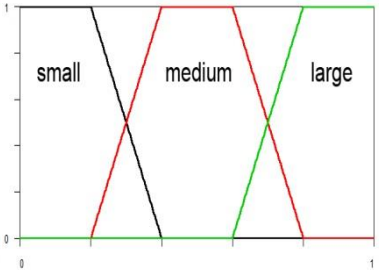


Figure 4: *petal length*

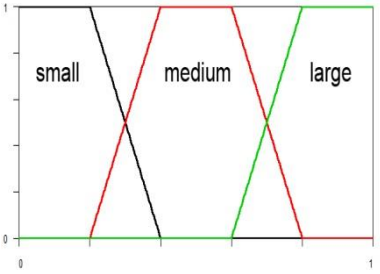


Figure 5: *petal width*

The total number of rules (R) for this complete FLS with fully specified rule base (using the "AND" logical connective), is $R = m^n = 3^4 = 81$ and can be illustrated as follows;

R1: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa*
:
:
R15: IF *sepal length* is *medium* AND *sepal width* is *small* AND *petal length* is *small* AND *petal width* is *large* THEN *Flower* is *Sentosa*
R16: IF *sepal length* is *large* AND *sepal width* is *medium* AND *petal length* is *small* AND *petal width* is *small* THEN *Flower* is *Versicolor*
:
:
R30: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *medium* AND *petal width* is *medium* THEN *Flower* is *Virginica*
R31: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *medium* AND *petal width* is *large* THEN *Flower* is *Virginica*
R32: IF *sepal length* is *small* AND *sepal width* is *small* AND *petal length* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*
:
:
R77: IF *sepal length* is *large* AND *sepal width* is *large* AND *petal length* is *medium* AND *petal width* is *large* THEN *Flower* is *Sentosa*
R78: IF *sepal length* is *medium* AND *sepal width* is *medium* AND *petal length* is *large* AND *petal width* is *small* THEN *Flower* is *Versicolor*
R79: IF *sepal length* is *medium* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *small* THEN *Flower* is *Virginica*
R80: IF *sepal length* is *medium* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *medium* THEN *Flower* is *Sentosa*
R81: IF *sepal length* is *medium* AND *sepal width* is *large* AND *petal length* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*

Parallel HFS

Fig. 2-A shows the hierarchical fuzzy system (HFS) – *parallel* topology for Iris classification.

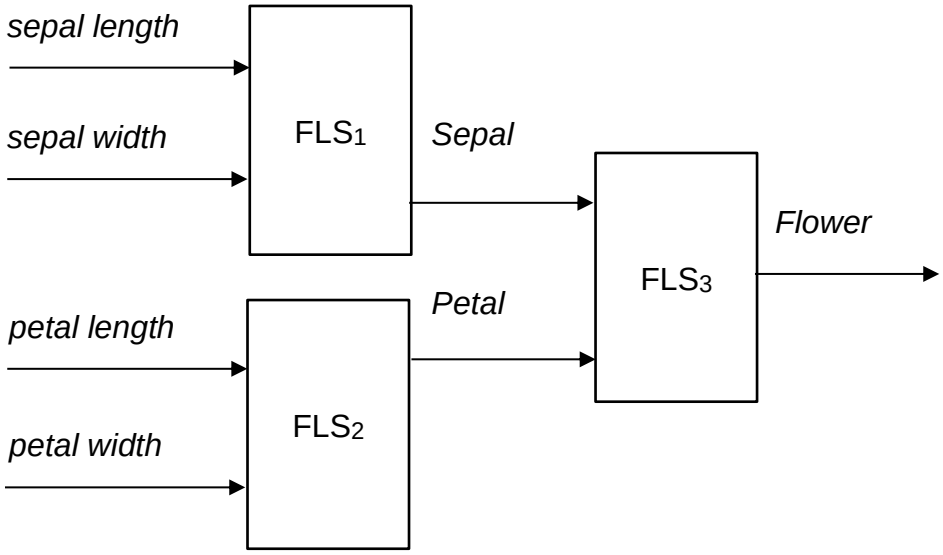


Figure 2-A: Hierarchical Fuzzy System - Parallel

P-2

Fig. 1 shows the hierarchical fuzzy system (HFS) – *parallel* topology of Iris classification.

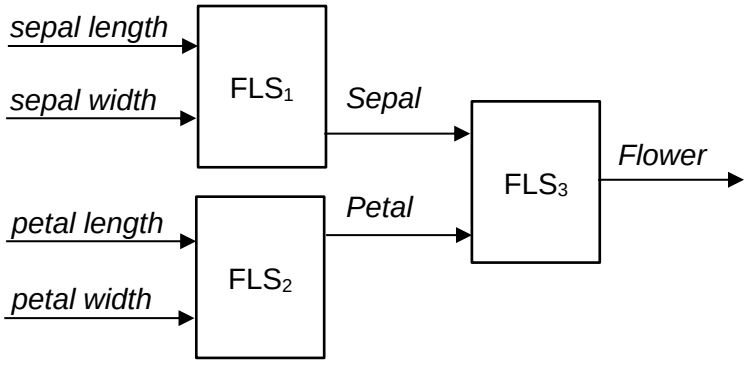


Figure 1: Hierarchical Fuzzy System - Parallel

The input variables including intermediate variables are modeled with two fuzzy sets ($m = 2$); *small* and *large* as shown in Fig. 2 ~ 7 respectively.

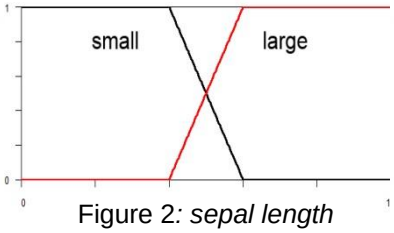


Figure 2*: sepal length*

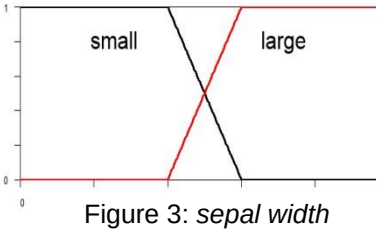


Figure 3: *sepal width*

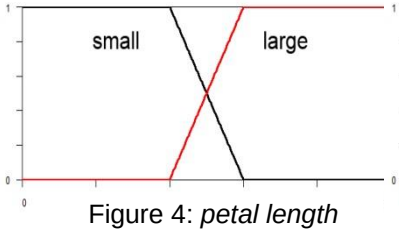


Figure 4: *petal length*

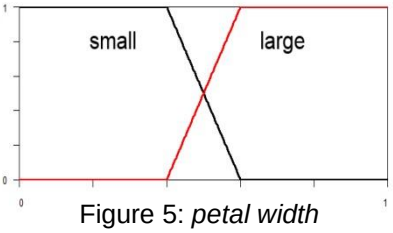


Figure 5: *petal width*

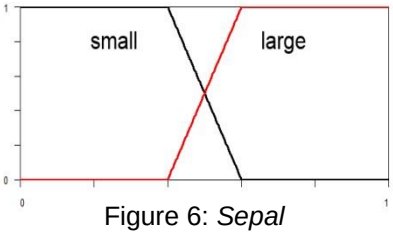


Figure 6: *Sepal*

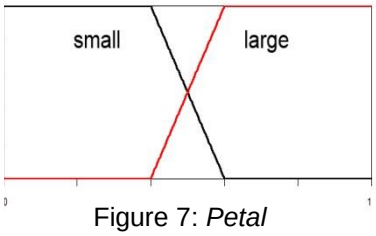


Figure 7: *Petal*

The total number of rules (R) for this HFS is $R = \sum m^2 = 2^2 + 2^2 + 2^2 = 12$ and can be illustrated as follows;

FLS$_1$:	FLS$_2$:	FLS$_3$:
R1: IF *sepal length* is *small* AND *sepal width* is *small* THEN *Sepal* is *small* R2: IF *sepal length* is *small* AND *sepal width* is *large* THEN *Sepal* is *large* R3: IF *sepal length* is *large* AND *sepal width* is *small* THEN *Sepal* is *small* R4: IF *sepal length* is *large* AND *sepal width* is *large* THEN *Sepal* is *large*	R1: IF *petal length* is *small* AND *petal width* is *small* THEN *Petal* is *small* R2: IF *petal length* is *small* AND *petal width* is *large* THEN *Petal* is *large* R3: IF *petal length* is *large* AND *petal width* is *small* THEN *Petal* is *small* R4: IF *petal length* is *large* AND *petal width* is *large* THEN *Petal* is *large*	R1: IF *Sepal* is *small* AND *Petal* is *small* THEN *Flower* is *Sentosa* R2: IF *Sepal* is *small* AND *Petal* is *large* THEN *Flower* is *Versicolor* R3: IF *Sepal* is *large* AND *Petal* is *small* THEN *Flower* is *Versicolor* R4: IF *Sepal* is *large* AND *Petal* is *large* THEN *Flower* is *Virginica*

P-3

Fig. 1 shows the hierarchical fuzzy system (HFS) – *parallel* topology of Iris classification.

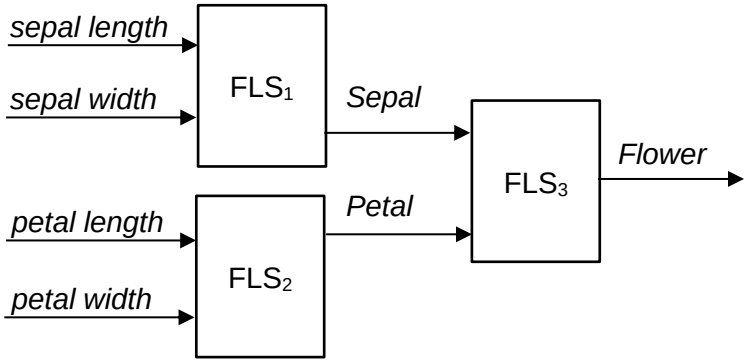


Figure 1: Hierarchical Fuzzy System - Parallel

The input variables including intermediate variables are modeled with three fuzzy sets ($m = 2$); *small, medium* and *large* as shown in Fig. 2 ~ 7 respectively.

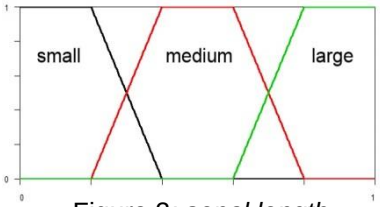


Figure 2*: sepal length*

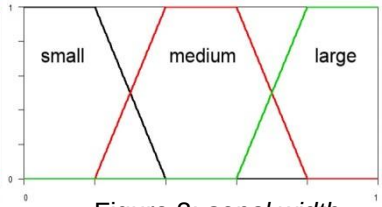


Figure 3: *sepal width*

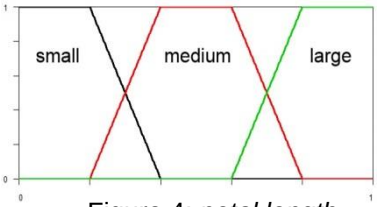


Figure 4: *petal length*

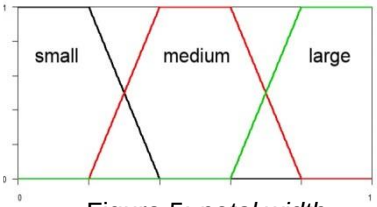


Figure 5: *petal width*

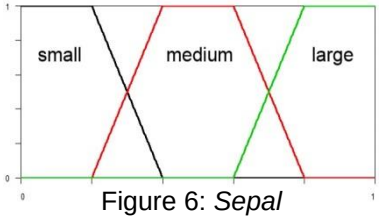


Figure 6: *Sepal*

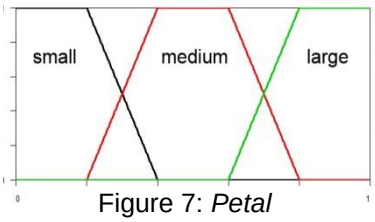


Figure 7: *Petal*

The total number of rules (R) for this HFS is $R = \sum m^2 = 3^2 + 3^2 + 3^2 = 27$ and can be illustrated as follows;

FLS$_1$:	FLS$_2$:	FLS$_3$:
R1: IF *sepal length* is *small* AND *sepal width* is *small* THEN *Sepal* is *small* R2: IF *sepal length* is *small* AND *sepal width* is *medium* THEN *Sepal* is *small* R3: IF *sepal length* is *small* AND *sepal width* is *large* THEN *Sepal* is *medium* R4: IF *sepal length* is *medium* AND *sepal width* is *small* THEN *Sepal* is *small* R5: IF *sepal length* is *medium* AND *sepal width* is *medium* THEN *Sepal* is *medium* R6: IF *sepal length* is *medium* AND *sepal width* is *large* THEN *Sepal* is *medium* R7: IF *sepal length* is *large* AND *sepal width* is *small* THEN *Sepal* is *medium* R8: IF *sepal length* is *large* AND *sepal width* is *medium* THEN *Sepal* is *large* R9: IF *sepal length* is *large* AND *sepal width* is *large* THEN *Sepal* is *large*	R1: IF *petal length* is *small* AND *petal width* is *small* THEN *Petal* is *small* R2: IF *petal length* is *small* AND *petal width* is *medium* THEN *Petal* is *small* R3: IF *petal length* is *small* AND *petal width* is *large* THEN *Petal* is *medium* R4: IF *petal length* is *medium* AND *petal width* is *small* THEN *Petal* is *small* R5: IF *petal length* is *medium* AND *petal width* is *medium* THEN *Petal* is *medium* R6: IF *petal length* is *medium* AND *petal width* is *large* THEN *Petal* is *medium* R7: IF *petal length* is *large* AND *petal width* is *small* THEN *Petal* is *medium* R8: IF *petal length* is *large* AND *petal width* is *medium* THEN *Petal* is *large* R9: IF *petal length* is *large* AND *petal width* is *large* THEN *Petal* is *large*	R1: IF *Sepal* is *small* AND *Petal* is *small* THEN *Flower* is *Sentosa* R2: IF *Sepal* is *small* AND *Petal* is *medium* THEN *Flower* is *Sentosa* R3: IF *Sepal* is *small* AND *Petal* is *large* THEN *Flower* is *Versicolor* R4: IF *Sepal* is *medium* AND *Petal* is *small* THEN *Flower* is *Sentosa* R5: IF *Sepal* is *medium* AND *Petal* is *medium* THEN *Flower* is *Versicolor* R6: IF *Sepal* is *medium* AND *Petal* is *large* THEN *Flower* is *Versicolor* R7: IF *Sepal* is *large* AND *Petal* is *small* THEN *Flower* is *Versicolor* R8: IF *Sepal* is *large* AND *Petal* is *medium* THEN *Flower* is *Virginica* R9: IF *Sepal* is *large* AND *Petal* is *large* THEN *Flower* is *Virginica*

Serial HFS

Fig. 3-A shows the hierarchical fuzzy system (HFS) – *serial* topology for Iris classification.

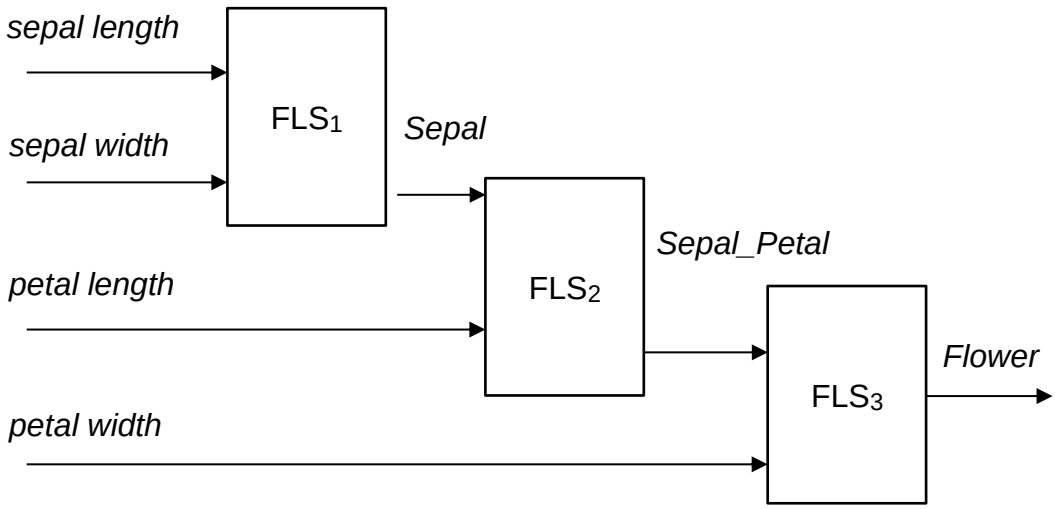


Figure 3-A: Hierarchical Fuzzy System - Serial

S-2

Fig. 1 shows the hierarchical fuzzy system (HFS) – *serial* topology of Iris classification.

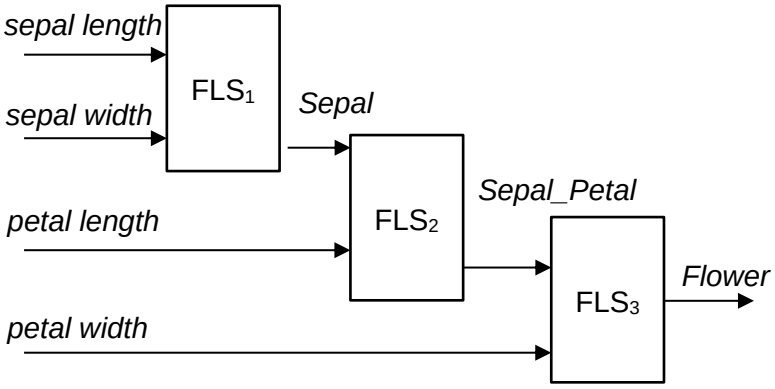


Figure 1: Hierarchical Fuzzy System - Serial

The input variables including intermediate variables are modeled with two fuzzy sets ($m = 2$); *small* and *large* as shown in Fig. 2 ~ 7 respectively.

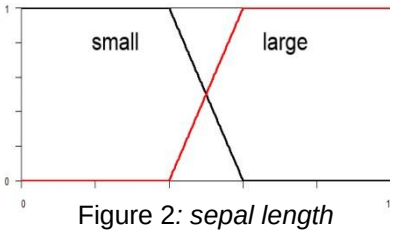


Figure 2*: sepal length*

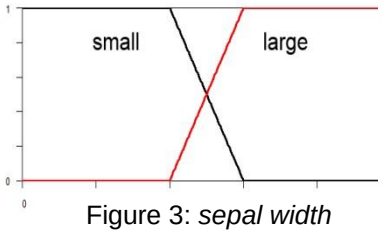


Figure 3: *sepal width*

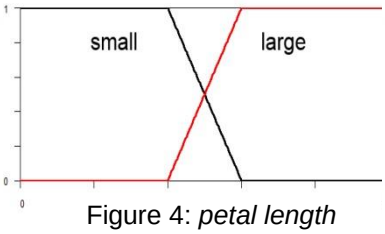


Figure 4: *petal length*

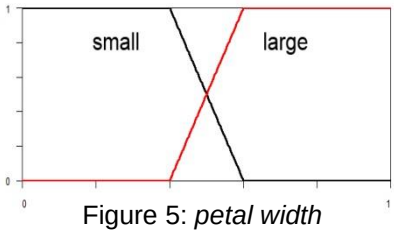


Figure 5: *petal width*

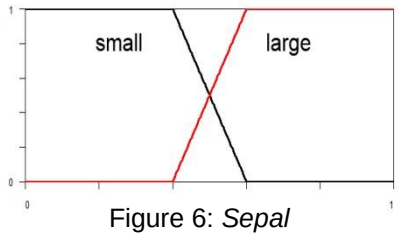


Figure 6: *Sepal*

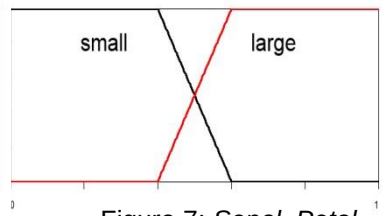


Figure 7: *Sepal_Petal*

The total number of rules (R) for this HFS is $R = (n-1)m^2 = (4-1)2^2 = 12$ and can be illustrated as follows;

FLS$_1$:	FLS$_2$:	FLS$_3$:
R1: IF *sepal length* is *small* AND *sepal width* is *small* THEN *Sepal* is *small* R2: IF *sepal length* is *small* AND *sepal width* is *large* THEN *Sepal* is *large* R3: IF *sepal length* is *large* AND *sepal width* is *small* THEN *Sepal* is *small* R4: IF *sepal length* is *large* AND *sepal width* is *large* THEN *Sepal* is *large*	R1: IF *Sepal* is *small* AND *petal length* is *small* THEN *Sepal_Petal* is *small* R2: IF *Sepal* is *small* AND *petal length* is *large* THEN *Sepal_Petal* is *large* R3: IF *Sepal* is *large* AND *petal length* is *small* THEN *Sepal_Petal* is *small* R4: IF *Sepal* is *large* AND *petal length* is *large* THEN *Sepal_Petal* is *large*	R1: IF *Sepal_Petal* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa* R2: IF *Sepal_Petal* is *small* AND *petal width* is *largel* THEN *Flower* is *Versicolor* R3: IF *Sepal_Petal* is *large* AND *petal width* is *small* THEN *Flower* is *Versicolor* R4: IF *Sepal_Petal* is *large* AND *petal width* is *largel* THEN *Flower* is *Virginica*

S-3

Fig. 1 shows the hierarchical fuzzy system (HFS) – *serial* topology of Iris classification.

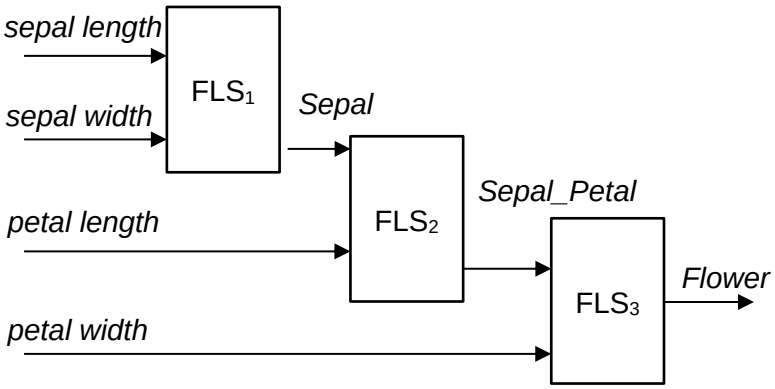


Figure 1: Hierarchical Fuzzy System - Serial

The input variables including intermediate variables are modeled with three fuzzy sets ($m = 3$); *small, medium* and *large* as shown in Fig. 2 ~ 7 respectively.

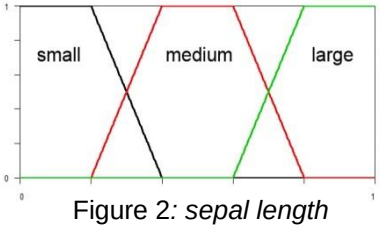


Figure 2*: sepal length*

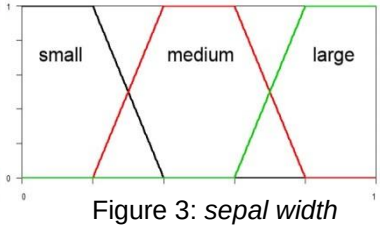


Figure 3: *sepal width*

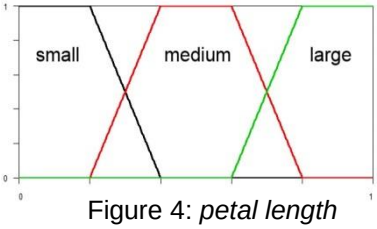


Figure 4: *petal length*

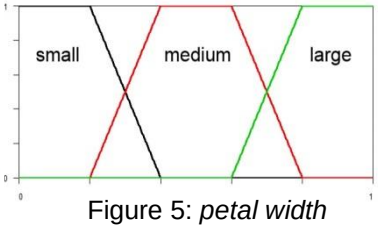


Figure 5: *petal width*

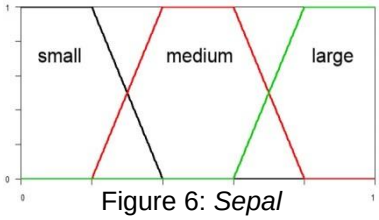


Figure 6: *Sepal*

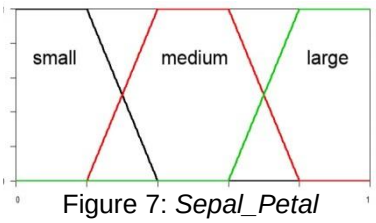


Figure 7: *Sepal_Petal*

The total number of rules (R) for this HFS is $R = (n-1)m^2 = (4-1)3^2 = 27$ and can be illustrated as follows;

FLS$_1$:	FLS$_2$:	FLS$_3$:
R1: IF *sepal length* is *small* AND *sepal width* is *small* THEN *Sepal* is *small* R2: IF *sepal length* is *small* AND *sepal width* is *medium* THEN *Sepal* is *small* R3: IF *sepal length* is *small* AND *sepal width* is *large* THEN *Sepal* is *medium* R4: IF *sepal length* is *medium* AND *sepal width* is *small* THEN *Sepal* is *small* R5: IF *sepal length* is *medium* AND *sepal width* is *medium* THEN *Sepal* is *medium* R6: IF *sepal length* is *medium* AND *sepal width* is *large* THEN *Sepal* is *medium* R7: IF *sepal length* is *large* AND *sepal width* is *small* THEN *Sepal* is *medium* R8: IF *sepal length* is *large* AND *sepal width* is *medium* THEN *Sepal* is *large* R9: IF *sepal length* is *large* AND *sepal width* is *large* THEN *Sepal* is *large*	R1: IF *Sepal* is *small* AND *petal length* is *small* THEN *Petal_Sepal* is *small* R2: IF *Sepal* is *small* AND *petal length* is *medium* THEN *Petal_Sepal* is *small* R3: IF *Sepal* is *small* AND *petal length* is *large* THEN *Petal_Sepal* is *medium* R4: IF *Sepal* is *medium AND petal length* is *small* THEN *Petal_Sepal* is *small* R5: IF *Sepal* is *medium* AND *petal length* is *medium* THEN *Petal_Sepal* is *medium* R6: IF *Sepal* is *medium* AND *petal length* is *large* THEN *Petal_Sepal* is *medium* R7: IF *Sepal* is *large* AND *petal length* is *small* THEN *Petal_Sepal* is *medium* R8: IF *Sepal* is *large* AND *petal length* is *medium* THEN *Petal_Sepal* is *large* R9: IF *Sepal* is *large* AND *petal length* is *large* THEN *Petal_Sepal* is *large*	R1: IF *Sepal_Petal* is *small* AND *petal width* is *small* THEN *Flower* is *Sentosa* R2: IF *Sepal_Petal* is *small* AND *petal width* is *medium* THEN *Flower* is *Sentosa* R3: IF *Sepal_Petal* is *small* AND *petal width* is *large* THEN *Flower* is *Versicolor* R4: IF *Sepal_Petal* is *medium* AND *petal width* is *small* THEN *Flower* is *Sentosa* R5: IF *Sepal_Petal* is *medium* AND *petal width* is *medium* THEN *Flower* is *Versicolor* R6: IF *Sepal_Petal* is *medium* AND *petal width* is *large* THEN *Flower* is *Versicolor* R7: IF *Sepal_Petal* is *large* AND *petal width* is *small* THEN *Flower* is *Versicolor* R8: IF *Sepal_Petal* is *large* AND *petal width* is *medium* THEN *Flower* is *Virginica* R9: IF *Sepal_Petal* is *large* AND *petal width* is *large* THEN *Flower* is *Virginica*

Appendix B

Participants' Feedback

Table B.1 Question 1 – Participants' feedback on the simplest topologies for Iris Systems

- *FLS is easy to see the input and output.*
- *FLS simple and direct. Parallel HFS and Serial HFS both ensemble not too much difference.*
- *Serialization is always simple as compared to parallelization.*
- *Considering all the variables equal (FLS) in one system where we only need 1 step looks the simpler for me. Although Parallel HFS and Serial HFS have both 3 steps, Serial HFS makes an unlogical association of variables that is why I consider it worse.*
- *From the implementation point of view, it is easier to implement one FLS in FLS than three FLSs in Parallel HFS or Serial HFS.*
- *The larger the number of FLS the harder to understand the architecture. Serial is harder than parallel because at least in this example it is somehow natural to group inputs two each side. However, this is problem dependent, it cannot be generalized.*
- *Flat FLSs are obviously the most interpretable; for Parallel HFS and Serial HFS, they both have to create additional hidden variables that are hard to interpret (or collect in real life) to human users.*
- *FLS requires fewest calculations. Parallel HFS requires more calculations, but clearly classifies sepal and petal separately before aggregation. Serial HFS aggregation of sepal and petal results is likely separate, but it's not clear.*
- *FLS is computationally efficient.*
- *FLS is has one rule base whereas Serial HFS has three rule bases within an asymmetric structure.*

Table B.2 Question 2 – Participants' feedback on the more interpretable of Iris Systems in Set MF-2

- *For P-2, the rule based system consists of simpler rules.*
- *F-2 although more rules, but more naturally and intuitionally. P-2 and S-2 less rules but difficult to connect the groups of rules.*
- *S-2 is less interpretable because of the way the rules feed forward.*
- *I think it is a good idea to have two sub-FLSs in terms of Sepal and Petal before classify the flower. But as I mentioned in question1, I think FLS_2 and FLS_3 are less interpretable in S-2.*
- *It is more intuitive to classify firstly sepals and petals separately and then the flowers.*
- *In this case, due to the nature of the problem, architecture P-2 fits very well the problem and accordingly P-2 is for me the most interpretable. S-2 has less rules than F-2 but the rules are not natural. For me, it is easier understanding the larger number of rules in F-2 than in S-2.*
- *It makes more sense to group the features of the same type together.*
- *P-2 has clear separate classification of sepal and petal. Easy to follow and understand because the rules are grouped. In F-2, the rules are simple to follow but because they are not grouped like P-2 it's slightly harder to read. In S-2, its unclear why sepal length and width are processed together, but petal length and width are separate. It's hard to follow the rules as a result.*
- *P-2 organises inputs in a semantically significant way. S-2 generates incomprehensible features*
- *More layers make rise to the complexity (time)*

Table B.3 Question 3 – Participants' feedback on the less complex of Iris Systems in Set MF-2

- *F-2 is less complex to design.*
- *P-2 and S-2 would allow me to decompose a complex problem into smaller one (divide and conquer)*
- *We assume F-2 as the base component. P-2 and S-2 are just combination.*
- *Less complex at every layer in term of combining the output rather than combining all layers output together at a time. Design is simple as compared to parallel approach.*
- *For me they are equal complex. With the rules, you have one system with 16 rules or 3 steps of 4 rules, but there is no difference in my opinion. (Neither of them are complex).*
- *I think F-2 has lower design complexity, since it is composed by only one FLS.*
- *F-2 has less fuzzy parameters and more rules to be optimised. P-2 and S-2 have more fuzzy parameters and less rules to be optimised.*
- *In F-2 and P-2, the rules are simply and equally easy to design. In S-2, the rules are intuitive.*
- *P-2 requires a semantically sound selection of inputs for the first layer.*
- *This is according to the number of FLSs – subsystems.*

Table B.4 Question 4 – Participants' feedback on the more interpretable of Iris Systems in Set MF-3

- *P-3 is more interpretable due to the simple rules.*
- *Less rules and organized in group make the system more interpretable.*
- *P-3 is easy because it has two layers and less number of rules.*
- *F-3 although more rules, but more naturally and intuitionally. P-3 and S-3 less rules but difficult to connect the groups of rules.*
- *Properties of same category (sepal, petal) are combined together make the system simple. This can be modelled using parallel approach.*
- *P-3 just seems more 'logical'.*
- *It is more intuitive to classify firstly sepals and petals separately and then the flowers.*
- *F-3 has many rules whereas P-3 has fewer rules within a symmetric structure.*
- *Neither is interpretable for obvious flaws. F-3 comes with too many rules; P-3/S-3 include hidden variables that are hard to interpret to human users plus a fairly large amount of rules.*
- *It makes more sense to group the features of the same type together.*

Table B.5 Question 5 – Participants' feedback on the less complex interpretable of Iris Systems in Set MF-3

- *F-3 is less complex to design.*
- *As in Question 3, decomposing the problem helps and I would also have a system with less rules.*
- *1 layer is easier to design but too many rules.*
- *Less complex at every layer in term of combining the output rather than combining all layers output together at a time. Design is simple as compared to parallel approach.*
- *The same as F-2, P-2 and S-2. There is no difference with three linguistic terms.*
- *I think F-3 has lower design complexity, since it is composed by only one FLS.*
- *F-3 has many rules whereas P-3 has fewer rules within a symmetric structure.*
- *Less rules in P-3 and S-3 to be optimised.*
- *The rule base cardinality for P-3 is log(n) m^2 smaller than others.*
- *P-3 requires a semantically sound selection of inputs for the first layer (same reason as in Question 3)*

Table B.6 Question 6 – Participants' feedback on HFS Iris systems (P-3 or S-3) that favoured to design in Set MF-3

- *There is in P-3, we can do a parallel processing (time computation).*
- *For this case, I prefer P-3 because the same type of input are together with same subsystems.*
- *Properties of same category (sepal, petal) are combined together make the system simple. This can be modelled using parallel approach.*
- *I have no idea what does this mean in S-3 by integrating the Sepal and petal length in FLS2 and further integrating the Sepal_Petal and petal width in FLS3.*
- *It may hold that S-3 is more complicated to implement but having 3 linguistics terms instead of 2 is better from a psychological point of view, due to it has an odd number of linguistic terms. If not, between small and large it is no middle point.*
- *P-3 is more intuitive for me.*
- *P-3 fits much naturally better the problem setting.*
- *Easier to understand. Each layer's purpose is clear in P-3, less so in S-3.*
- *P-3 organises inputs in a semantically significant way. S-3 generates incomprehensible features.*
- *P-3 is less complicated.*

Table B.7 Question 7 – Participants' feedback on agreeing to choose HFS (P-2) than FLS (F-2) with rule reduction from 16 to 12 rules in Set MF-2

- *Rules are simple and interpretable.*
- *The reduction is pretty small.*
- *Less number of rules - less mistake and less complexity.*
- *Assuming the results are just as accurate, it is worth it. P-2 is easier to interpret.*
- *I do not know. I believe that it depends on the specific problem. If the value of the second variable depends on the first one, then is better HFS. If not, the system with all the rules is more accurate. In this case, I will say no!*
- *I prefer F-2. It has four rules more but two subsystems (FLSs) less.*
- *The time to select the trigger rule is minimized.*
- *It is not the rule reduction the design criteria. HFS with less rules is less flexible than a flat FLS.*
- *It doesn't affect much. Difference small (number of rules).*
- *Depends on application based on computational time.*

Table B.8 Question 8 – Participants' feedback on agreeing to choose HFS (P-3) than FLS (F-3) with rule reduction from 81 to 27 rules in Set MF-3

- *The reduction is significant, P-3 looks much easier to design and understand.*
- *Scale of problem is reduced.*
- *Large reduction.*
- *Again, it depends on the problem and how the variables depend in each other. However, in this case the difference is so big that it may worth.*
- *In this case, I prefer P-3, because it has 54 rules less.*
- *Same answer as given in Question 7. Also, because of the huge reduction in the number of rules*
- *I would go for HFS given finer portioning and/or more features. However, I would still go for flat FLS if interpretability is of a big concern.*
- *Assuming the results are just as accurate, it is worth it. P-3 is easier to interpret.*
- *Smaller rule bases.*
- *Rules are simple and interpretable.*

Table B.9 Question 9 – Participants' feedback on HFS Iris systems (P-3 or S-3) that is more interpretable in Set MF-3

- *Is more interpretable because user can visualize the effect of each feature.*
- *The treatment of sepal and petal is separated.*
- *Topology is less complex.*
- *For this case, I prefer P-3 because the same type of input are together with same subsystems.*
- *P-3 can be constructed with parallel system, taking just constant time in the 1st layer of the FLS.*
- *Again, I have no idea what does this mean in S-3 by integrating the Sepal and petal length in FLS_2 and further integrating the Sepal_Petal and petal width in FLS_3.*
- *As explained better, for me is more logical to make independent systems of petal and sepal and later on joining them than adding a variable at each step.*
- *It is more intuitive to classify firstly sepals and petals separately and then the flowers.*
- *S-3 seems more straightforward, as you probably can merge one more attribute as per some significance measure. For P-3, it seems you have to separate all features into a number of pair-wise combinations.*
- *Both are the same.*

Appendix C

Interpretability on the various design of Fuzzy Logic Systems Survey

Interpretability on the various design of Fuzzy Logic System Survey

Introduction

'Interpretability' is related to the capability to express something in an understandable way i.e. people may say that something is interpretable if they can understand it easily. One of the strengths of fuzzy logic systems (FLSs) is their interpretability which indicates how easily an FLS can be understood by a user.

This survey is designed to assess your opinion about the concept of interpretability of FLSs and hierarchical fuzzy systems (HFSs). The data collected will be anonymously used in order to examine how people perceive interpretability in FLSs and HFSs using a *Rotary crane system* designs as an example.

All students, researchers, and academics are very welcome to participate.

Data Protection

Consent will be sought before research takes place. You will be guaranteed complete anonymity, with all recorded data from your participation identified by a number only. You are free to withdraw from the research at any time and have any data related to you destroyed – to withdraw, please contact the researcher (either in person or via the researcher's contact details below). All of the data collected will be stored in accordance with the General Data Protection Regulation (GDPR) and the Data Protection Act 2018 (DPA) and remains confidential.

Please do not hesitate to ask any questions. You may contact the researcher at any time about the research or in relation to your consent, including withdrawal from the study:

Researcher:	Tajul Razak
Email:	tajul.razak@nottingham.ac.uk
Address:	B38, School of Computer Science, Jubilee Campus, The University of Nottingham, UK.

Participant Consent Form

Taking part in the study (Please read and tick (/) each statement below): *Required*

☐ I have read and understood the project information sheet dated or it has been read to me. I have been able to ask questions about the study and my questions have been answered satisfactorily.

☐ I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.

Use of my data in the study (Please read and tick (/) each statement below): *Required*

☐ I understand that data which can identify me will not be shared beyond the project team.

☐ I understand that the survey will gather my opinion and I agree for these data to be used in an anonymised form for research purposes e.g. discussions, presentations, conferences, publications.

Reuse of my data (Please read and tick (/) each statement below): *Required*

☐ I give permission for the data that I provide to be reused for the sole purposes of future research and learning.

☐ I understand and agree that this may involve depositing my data in a data repository, which may be accessed by other researchers.

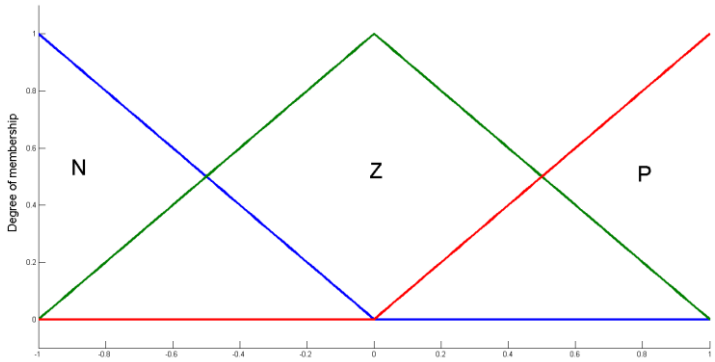
Copyright (Please read and tick (/) statement below): *Required*

☐ I give permission for data gathered during this project to be used, copied, excerpted, annotated, displayed and distributed for the purposes to which I have consented.

Instructions

This survey is designed to assess your opinion about the concept of interpretability of FLSs and hierarchical fuzzy systems (HFSs). The purpose of the study is to examine how people perceive interpretability in FLSs and HFSs using a rotary crane system[1] as an example.

For this example, the six input variables (n=6), i.e., x_1,x_2,x_3,x_4,x_5 and x_6, and the output, y are modelled with three fuzzy sets (m=3) as shown in the figure below.

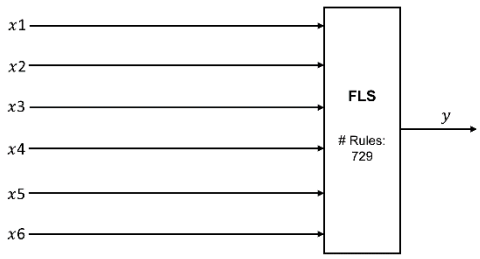


The survey consists of 20 questions for pairwise comparison. Please select **ONE** of the options which are **more interpretable** to you. If both systems seem equally interpretable, you may select '**Equal**' as your responds.

The survey should take no longer than 15 minutes to complete. Please click 'Next' to begin.

Pairwise comparison - 1

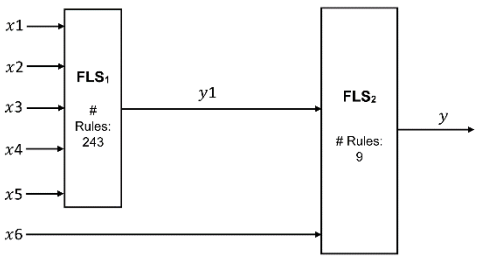
System A



FLS rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System B



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

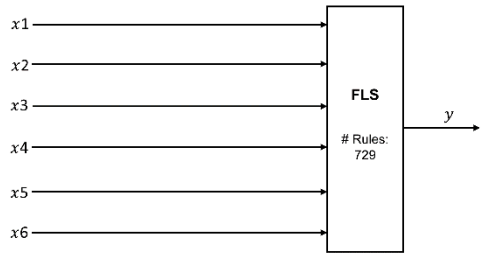
IF $y1$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System A
- System B
- Equal

Pairwise comparison - 2

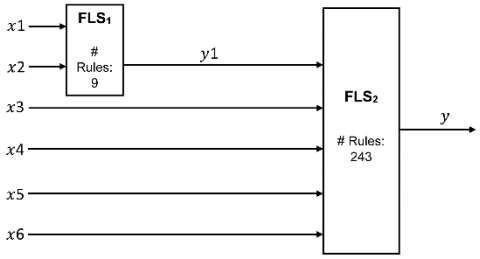
System A



FLS rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System C



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

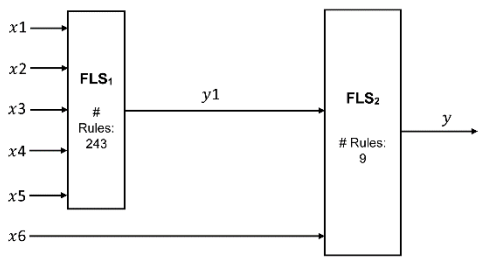
IF $y1$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System A
- System C
- Equal

Pairwise comparison - 3

System B



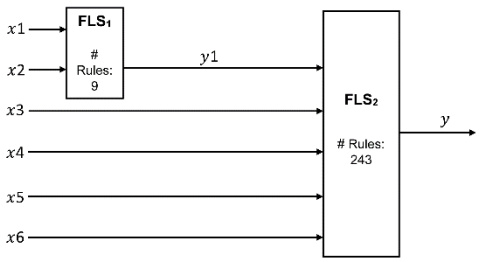
FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x6$ is *N* THEN y is *N*

System C



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

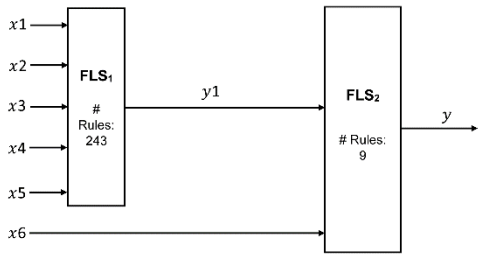
IF $y1$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System B
- System C
- Equal

Pairwise comparison - 4

System B



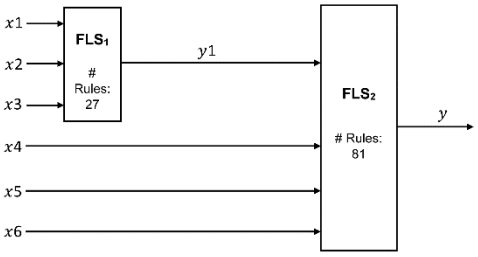
FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x6$ is *N* THEN y is *N*

System D



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

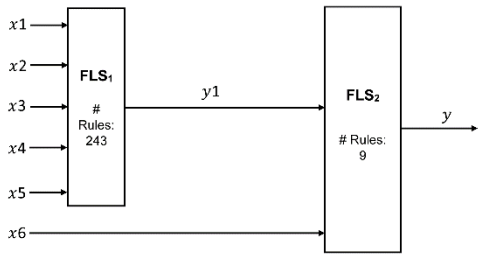
IF $y1$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System B
- System D
- Equal

Pairwise comparison - 5

System B



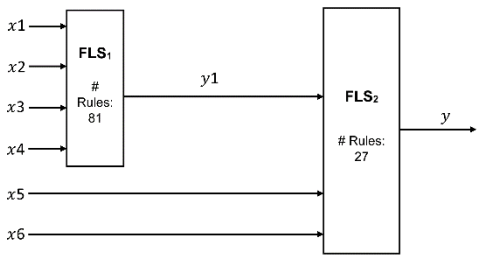
FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x6$ is *N* THEN y is *N*

System E



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

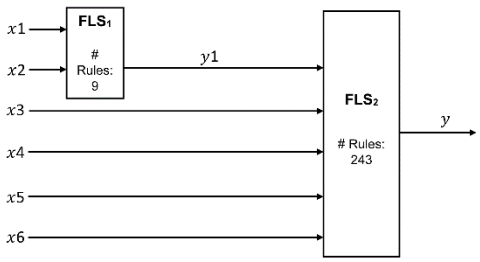
IF $y1$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System B
- System E
- Equal

Pairwise comparison - 6

System C



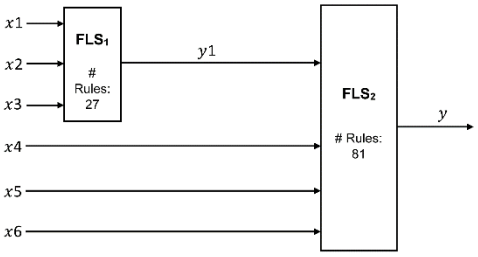
FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System D



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

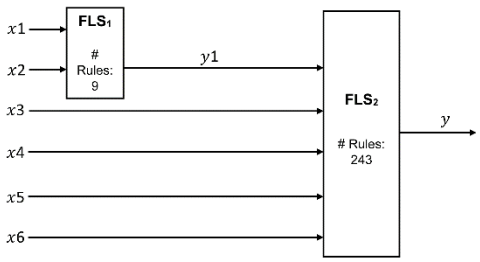
IF $y1$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System C
- System D
- Equal

Pairwise comparison - 7

System C



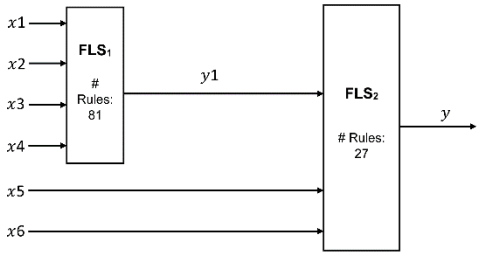
FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System E



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

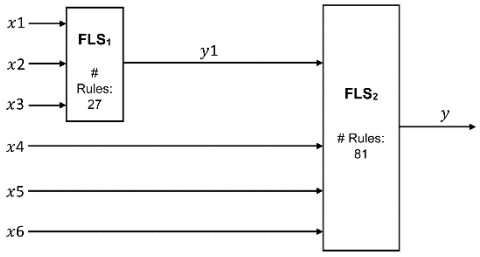
IF $y1$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System C
- System E
- Equal

Pairwise comparison - 8

System D



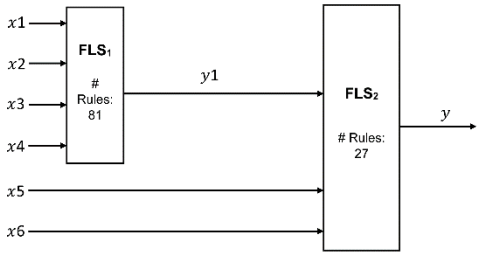
FLS₁ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS₂ rule example:

IF $y1$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System E



FLS₁ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* THEN $y1$ is *N*

FLS₂ rule example:

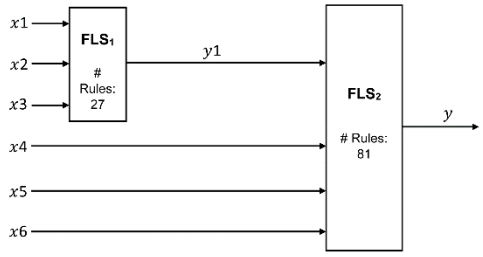
IF $y1$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System D
- System E
- Equal

Pairwise comparison - 9

System D



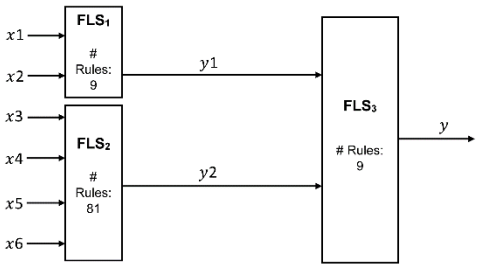
FLS₁ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS₂ rule example:

IF $y1$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System F



FLS₁ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS₂ rule example:

IF $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS₃ rule example:

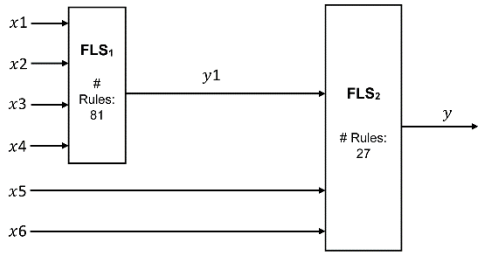
IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System D
- System F
- Equal

Pairwise comparison - 10

System E



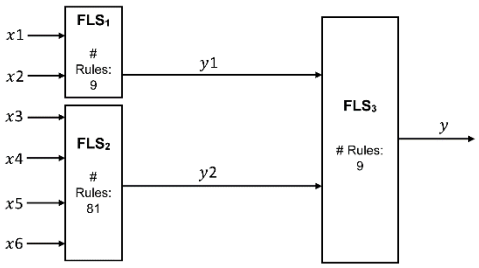
FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* and $x4$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System F



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

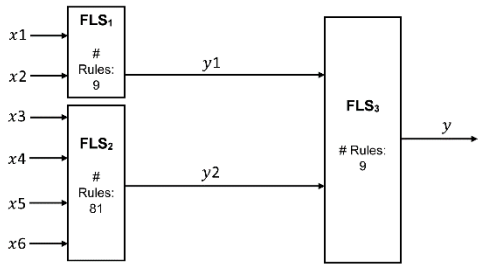
IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System E
- System F
- Equal

Pairwise comparison - 11

System F



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

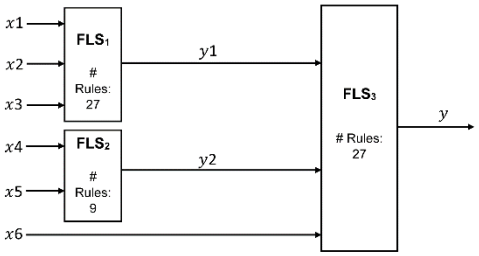
FLS$_2$ rule example:

IF $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

System G



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

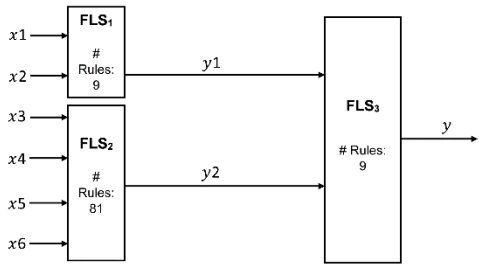
IF $y1$ is *N* and $y2$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System F
- System G
- Equal

Pairwise comparison - 12

System F



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

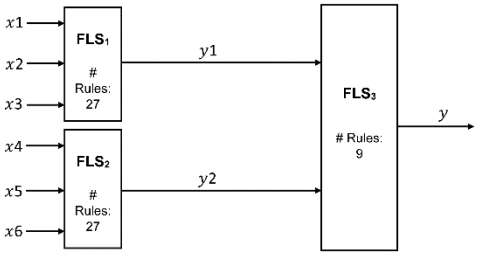
FLS$_2$ rule example:

IF $x3$ is *N* and $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

System H



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

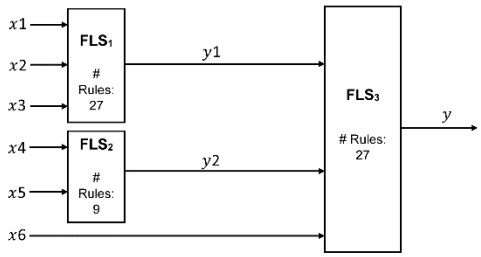
IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System F
- System H
- Equal

Pairwise comparison - 13

System G



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

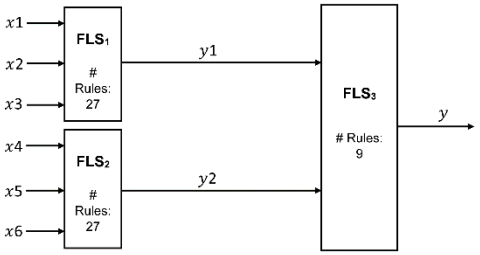
FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* and $x6$ is *N* THEN y is *N*

System H



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

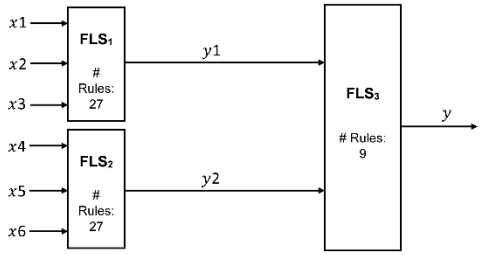
IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System G
- System H
- Equal

Pairwise comparison - 14

System H



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

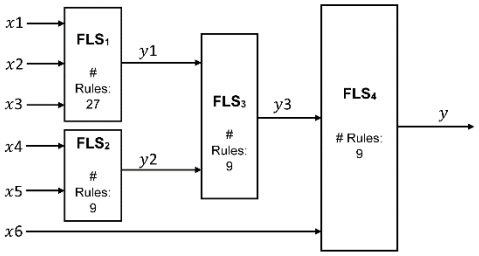
FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN y is *N*

System I



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

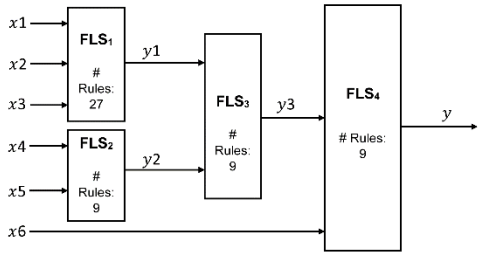
IF $y3$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System H
- System I
- Equal

Pairwise comparison - 15

System I



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* THEN $y2$ is *N*

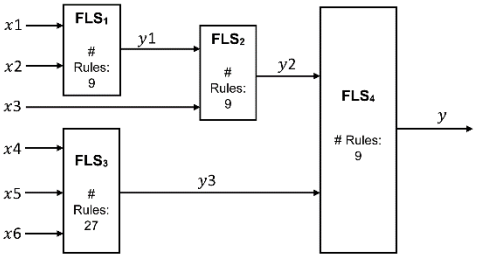
FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x6$ is *N* THEN y is *N*

System J



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

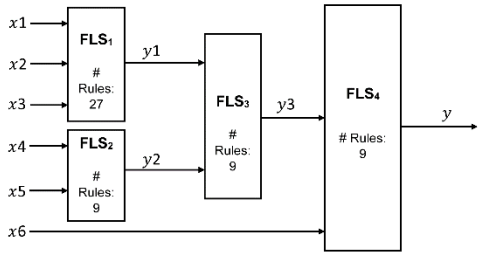
IF $y2$ is *N* and $y3$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System I
- System J
- Equal

Pairwise comparison - 16

System I



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* THEN $y2$ is *N*

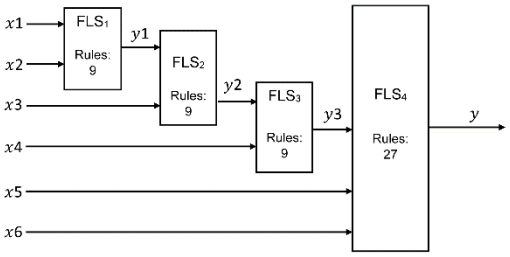
FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x6$ is *N* THEN y is *N*

System K



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y2$ is *N* and $x4$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

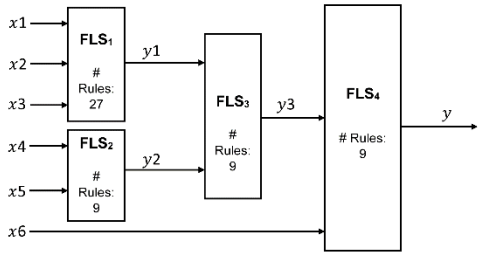
IF $y3$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System I
- System K
- Equal

Pairwise comparison - 17

System I



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* and $x3$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $x4$ is *N* and $x5$ is *N* THEN $y2$ is *N*

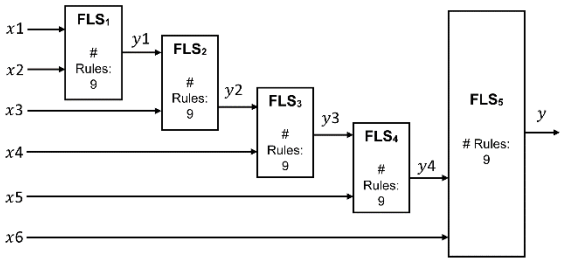
FLS$_3$ rule example:

IF $y1$ is *N* and $y2$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x6$ is *N* THEN y is *N*

System L



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y2$ is *N* and $x4$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x5$ is *N* THEN $y4$ is *N*

FLS$_5$ rule example:

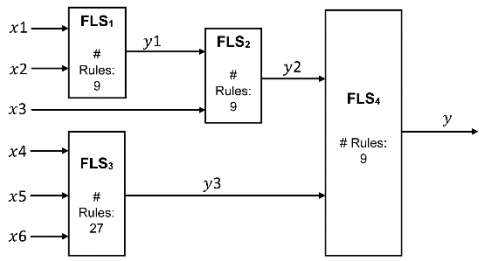
IF $y4$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System I
- System L
- Equal

Pairwise comparison - 18

System J



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

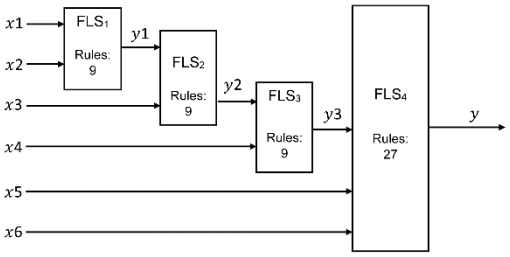
FLS$_3$ rule example:

IF $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y2$ is *N* and $y3$ is *N* THEN y is *N*

System K



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y2$ is *N* and $x4$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

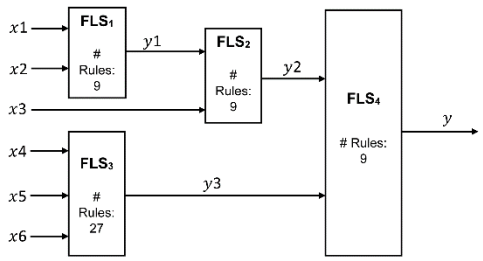
IF $y3$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System J
- System K
- Equal

Pairwise comparison - 19

System J



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

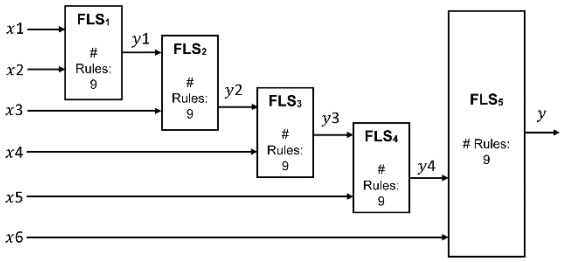
FLS$_3$ rule example:

IF $x4$ is *N* and $x5$ is *N* and $x6$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y2$ is *N* and $y3$ is *N* THEN y is *N*

System L



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y2$ is *N* and $x4$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x5$ is *N* THEN $y4$ is *N*

FLS$_5$ rule example:

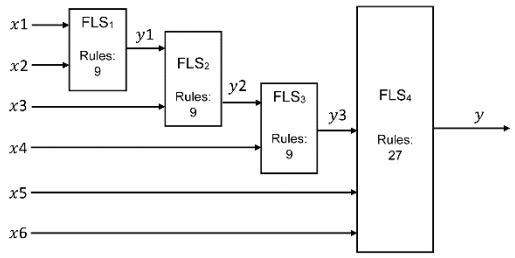
IF $y4$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System J
- System L
- Equal

Pairwise comparison - 20

System K



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

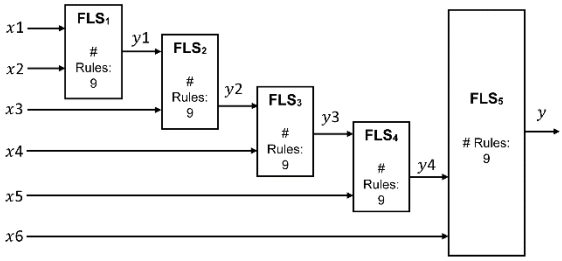
FLS$_3$ rule example:

IF $y2$ is *N* and $x4$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x5$ is *N* and $x6$ is *N* THEN y is *N*

System L



FLS$_1$ rule example:

IF $x1$ is *N* and $x2$ is *N* THEN $y1$ is *N*

FLS$_2$ rule example:

IF $y1$ is *N* and $x3$ is *N* THEN $y2$ is *N*

FLS$_3$ rule example:

IF $y2$ is *N* and $x4$ is *N* THEN $y3$ is *N*

FLS$_4$ rule example:

IF $y3$ is *N* and $x5$ is *N* THEN $y4$ is *N*

FLS$_5$ rule example:

IF $y4$ is *N* and $x6$ is *N* THEN y is *N*

Which of these two FLS designs would you prefer? *Required*

- System K
- System L
- Equal

Thank You

Thank you for your participation, it is very much appreciated.

Please do not hesitate to ask any question. You may contact us at any time about this survey.

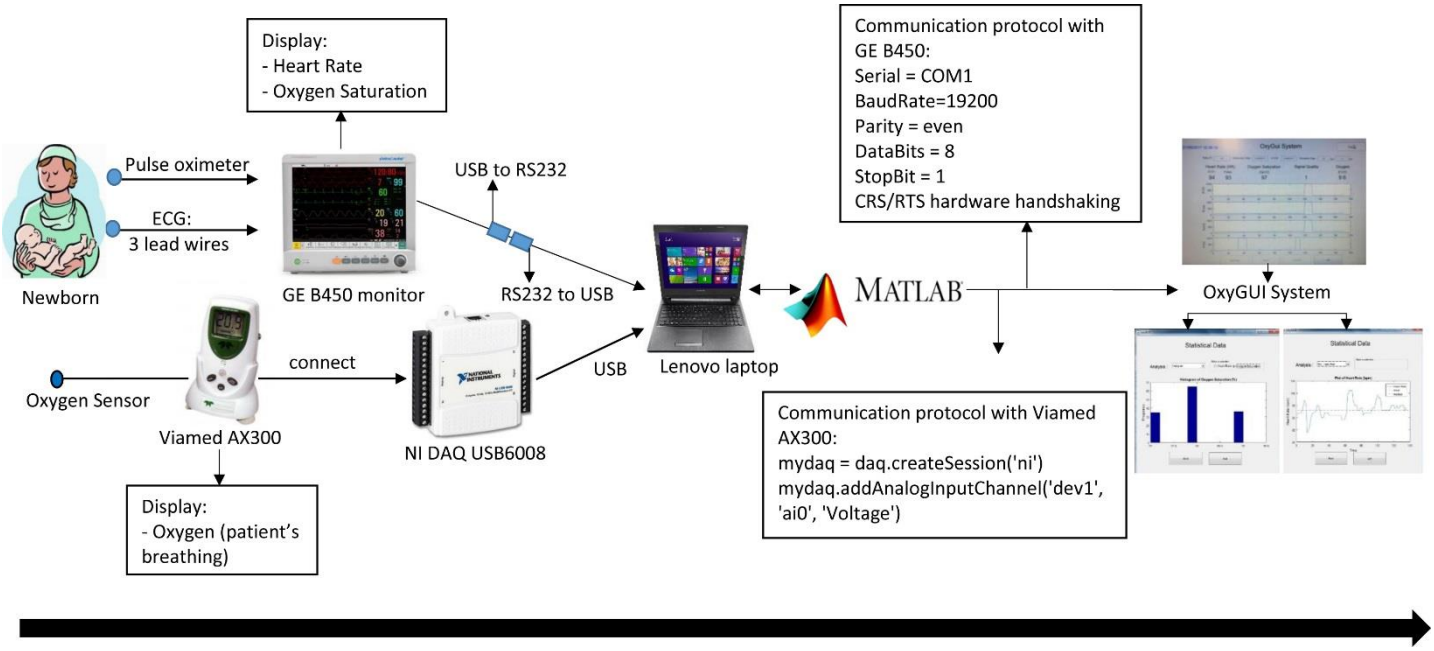
Email: tajul.razak@nottingham.ac.uk
Intelligent Modelling & Analysis (IMA) and Lab for Uncertainty in Data and Decision Making (LUCID) Research Groups,
School of Computer Science, The University of Nottingham,
United Kingdom

Appendix D

Data Acquisition: OxyGUI System

OxyGUI System

OxyGUI System – A flow process

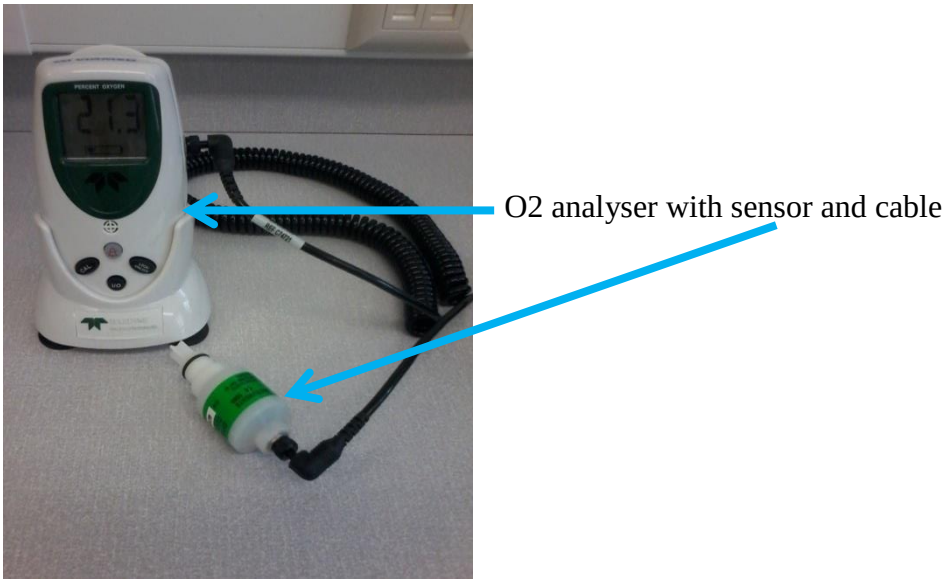


Setting up the OxyGUI System

The OxyGUI system is designed to monitor and record data in the background with minimal input or observation required. The system has been assessed, tested and passed for clinical use by the Clinical Medical Engineering team. The equipment should only be used by those who have are competent to use it and have received training and observation in this respect. If in doubt, do speak with one of the project team – Dr Don Sharkey.

Below you will find step-by-step setup instructions. For individual items, see the appendix for more information.

1. 1. Calibrate the O_2 analyser.



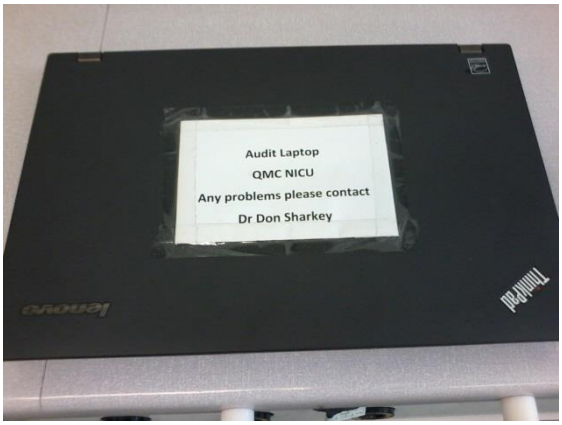




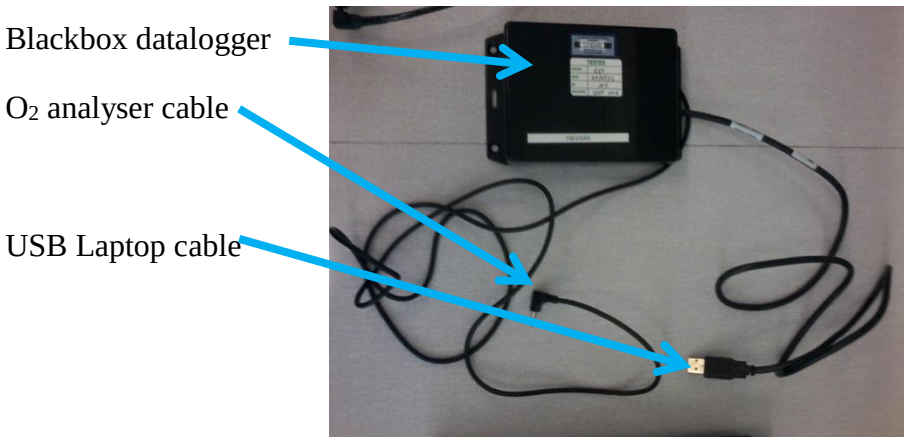


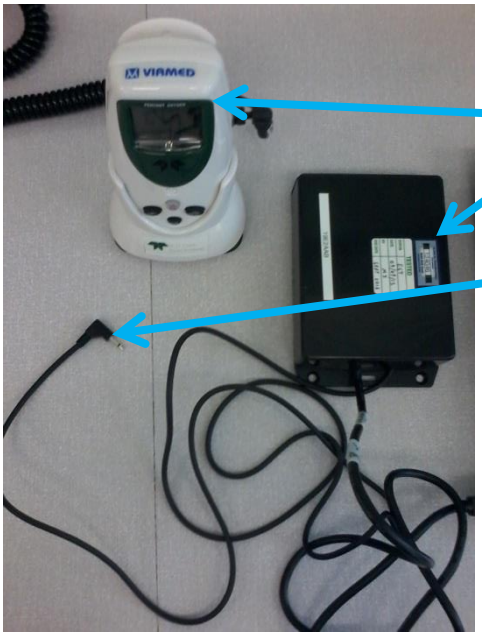
Switch the analyser on (bottom centre button). If the keys are locked press the lock button (key on right). The Cal button (on left) is pressed so the analyser then calibrates the sensor which must be in air at the calibration stage. The analyser is then ready to use.

2. Locate sufficient clear space close to the baby's monitor and respiratory support. Place the laptop in the chosen space. Plug the laptop in and place the power pack nearby.



3. Plug the blackbox, which holds the datalogger, into the oxygen analyser and laptop.





O_2 analyser

Blackbox datalogger

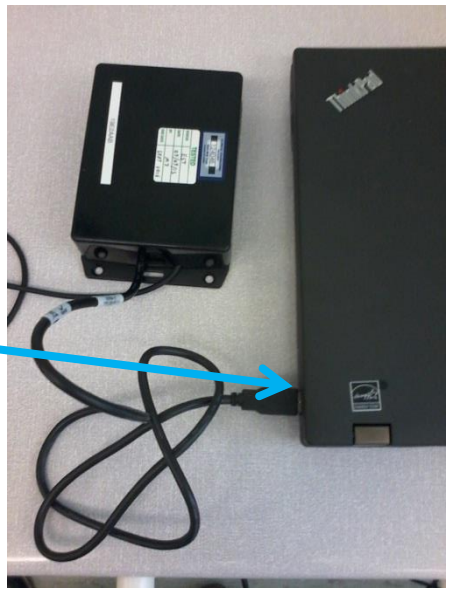
O_2 analyser cable connector inserts left side of analyser



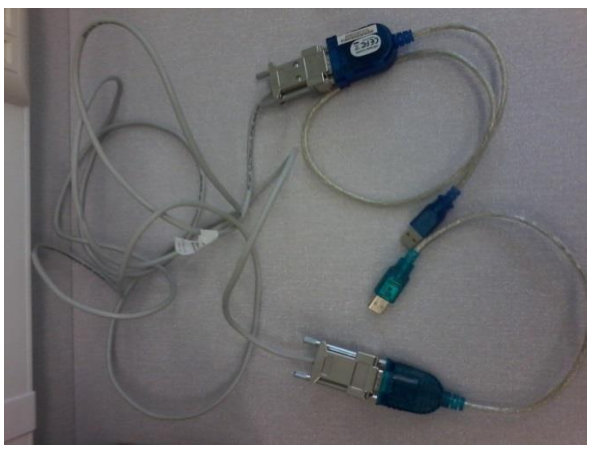


O_2 analyser cable connected

Blackbox datalogger USB cable is connected to the laptop USB port as shown



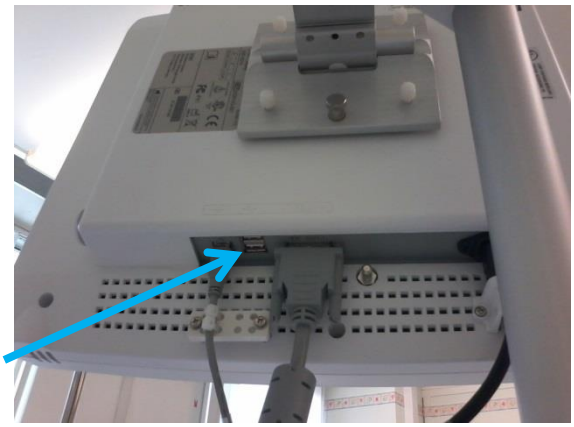
4. Next connect the laptop to the patient monitor. This is done using the long data cable with USB connectors at each end.

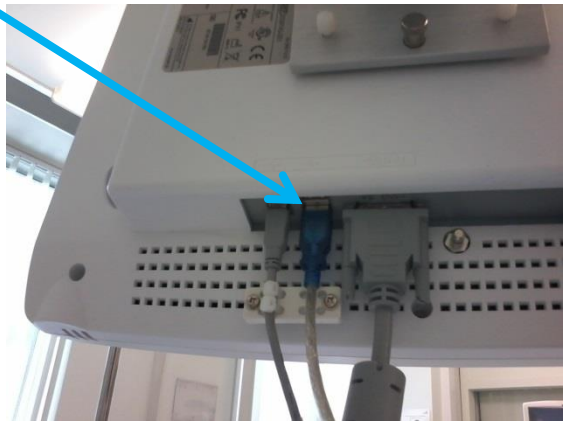


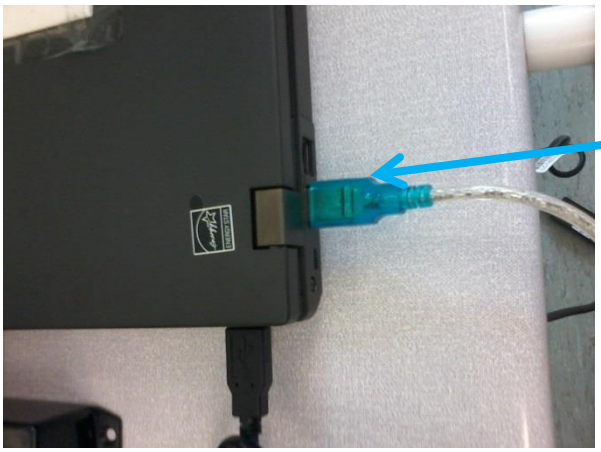


USB cable for patient monitor found at back on right hand side as you look at screen

USB port (actually 2 so use any) on back of patient monitor and with USB cable in place







USB cable from patient monitor connected to USB on laptop

5. The O2 analyser sensor is then connected to the patient circuit. This is done at the level of the humidifier. Identify the inspiratory arm from the ventilator/CPAP to the humidifier. This is where the T-piece that holds the sensor is placed. The O2 analyser sensor can then be placed in the T-piece.



T-piece for O_2 analyser sensor (single patient use only)



T-piece with O_2 analyser sensor in place



Inspiratory limb on humidifier where T-piece is connected

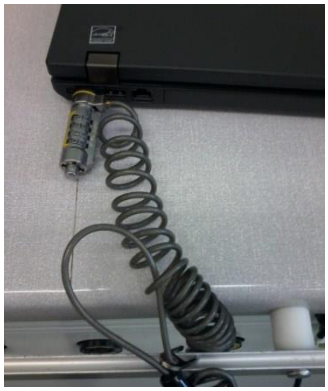
T-piece and sensor in place, here on a ventilator circuit



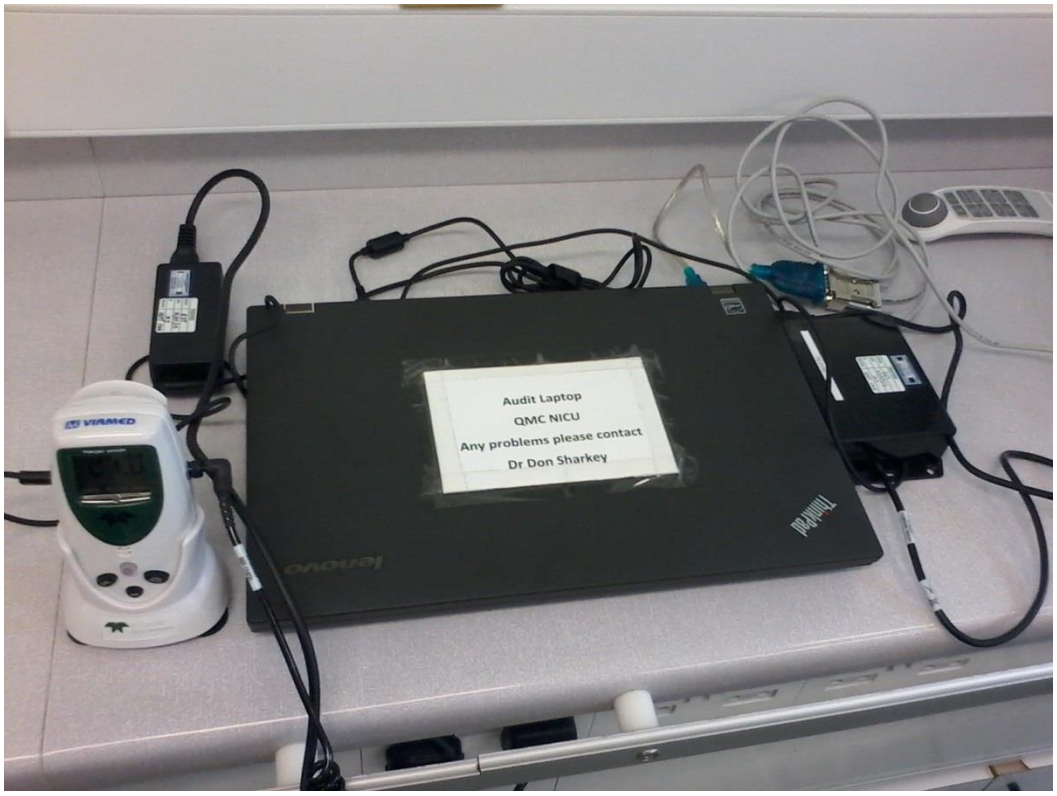
6. Secure the laptop with the coded cable to avoid it being removed. The cable is looped around a static fixture and then locked into the rear of the laptop using the security hole.







Rear of laptop showing security cable hole. Combination cable lock in place and secured around nearby fixture.



The final setup (without security lock).

Cleaning

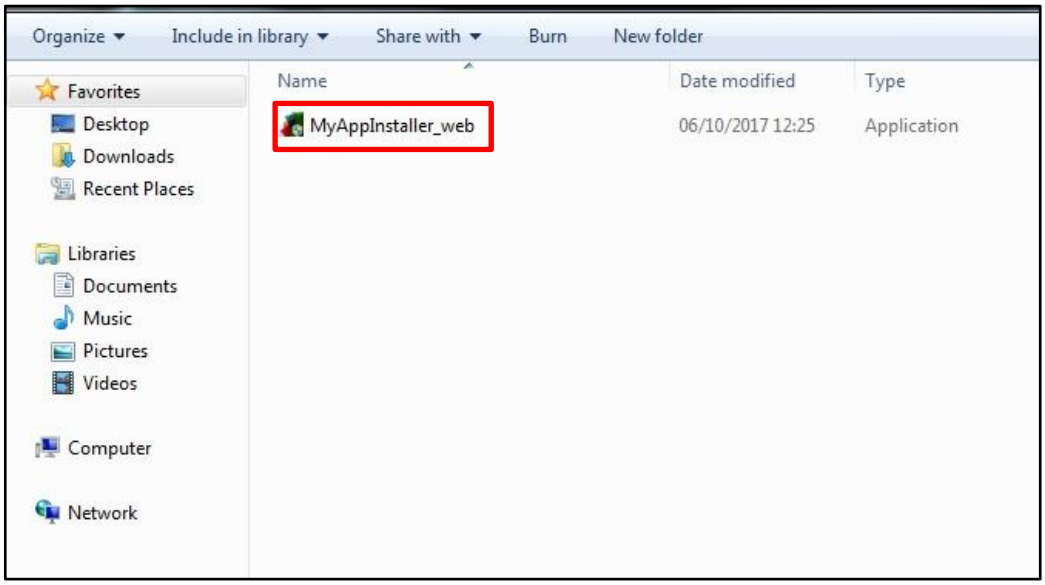
All equipment should be kept clean and periodically wiped using standard clinical wipes. Between patients this must be repeated including leads and surfaces. The T-piece O_2 analyser sensor holder should be disposed of as these single patient use.

The Software

The software on the laptop is written using the MatLab program which is installed on the computer. It uses a Graphical User Interface (GUI) to choose which equipment you want to use and the recording details. The password for the laptop is **Br@inSc@n52** precisely as written including the correct case.

i. Installation – OxyGUI System

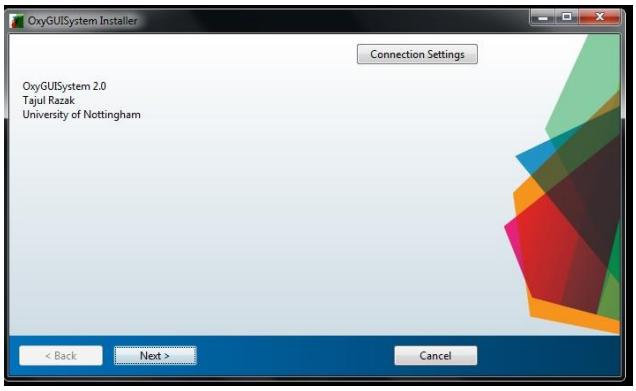
To install this OxyGUI System software, please open the OxyGUISystem folder on the main desktop and click on ‘ for_distribution’ folder. Then, please click MyAppInstaller_web as shown in the picture below.



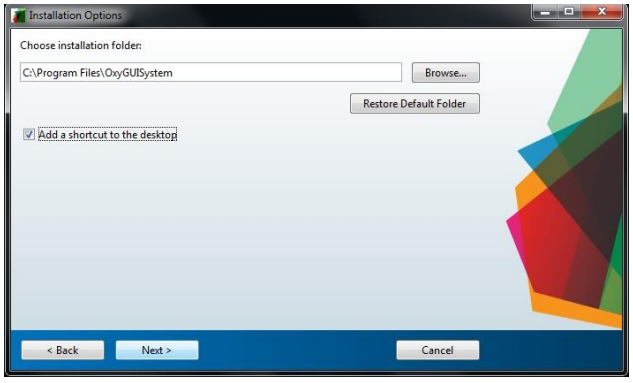
Then, a screen box will appear to indicate that it is preparing to be installed as shown below.



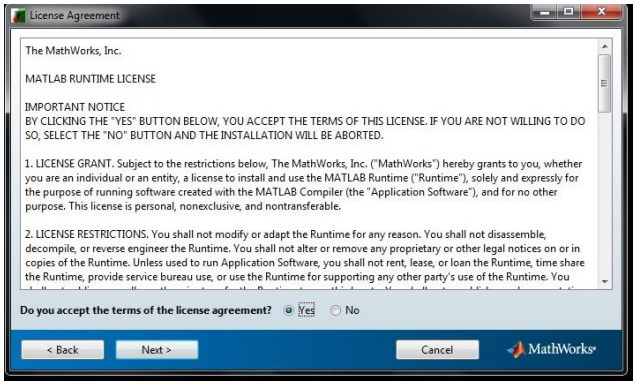
Please click ‘Next’ to proceed with the installation.



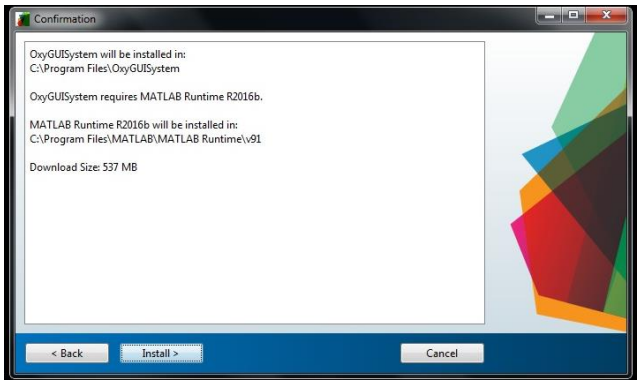
Please click ‘Next’ to proceed with the installation.



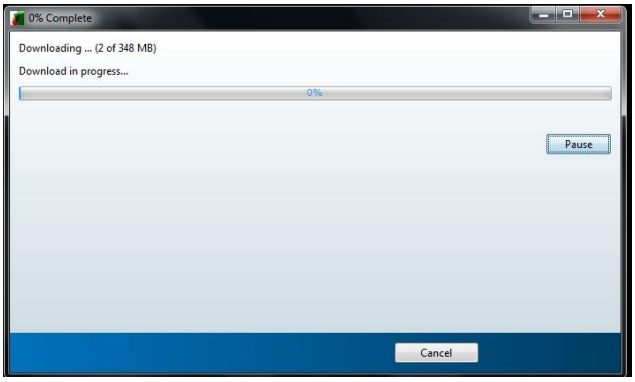
Please click 'Yes' to accept the license agreement for this software. Then, please click ‘Next’ to proceed with the installation.



Please click ‘Install’ to proceed with the installation.

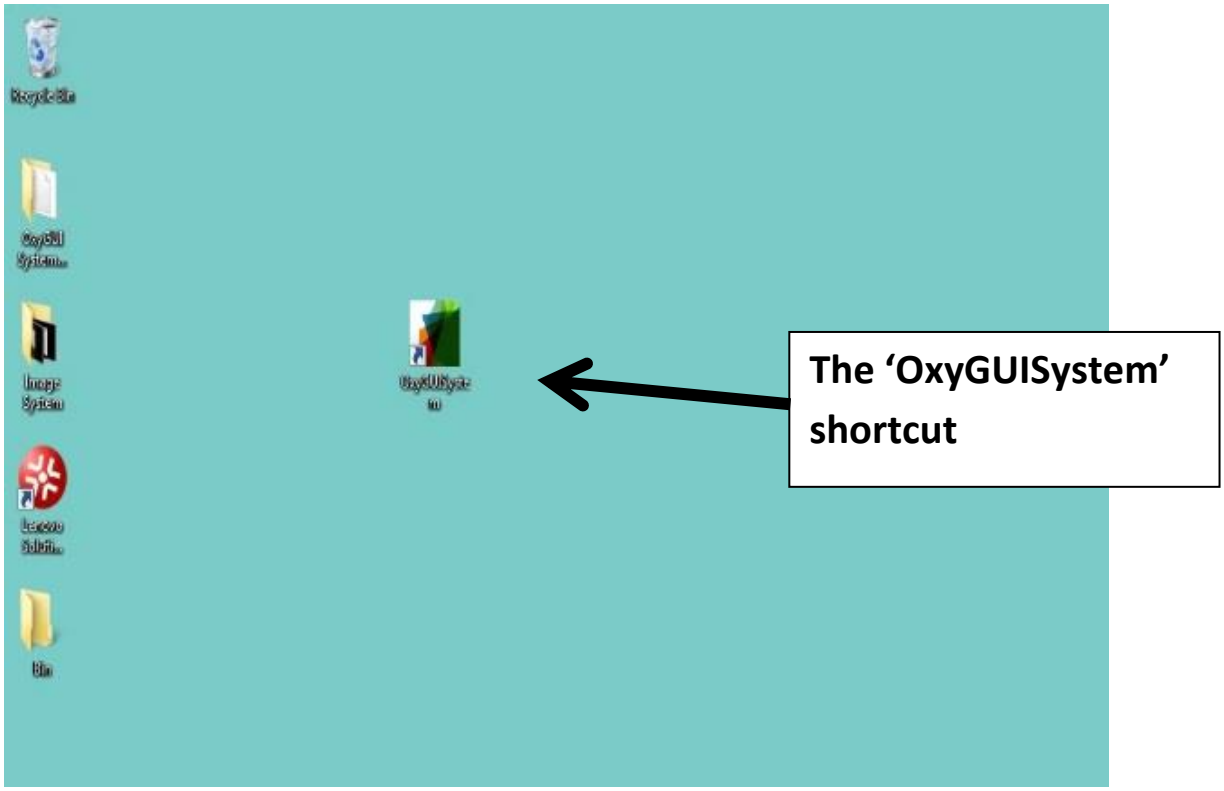


The ‘OxyGUI System’ is now being installed on this laptop.

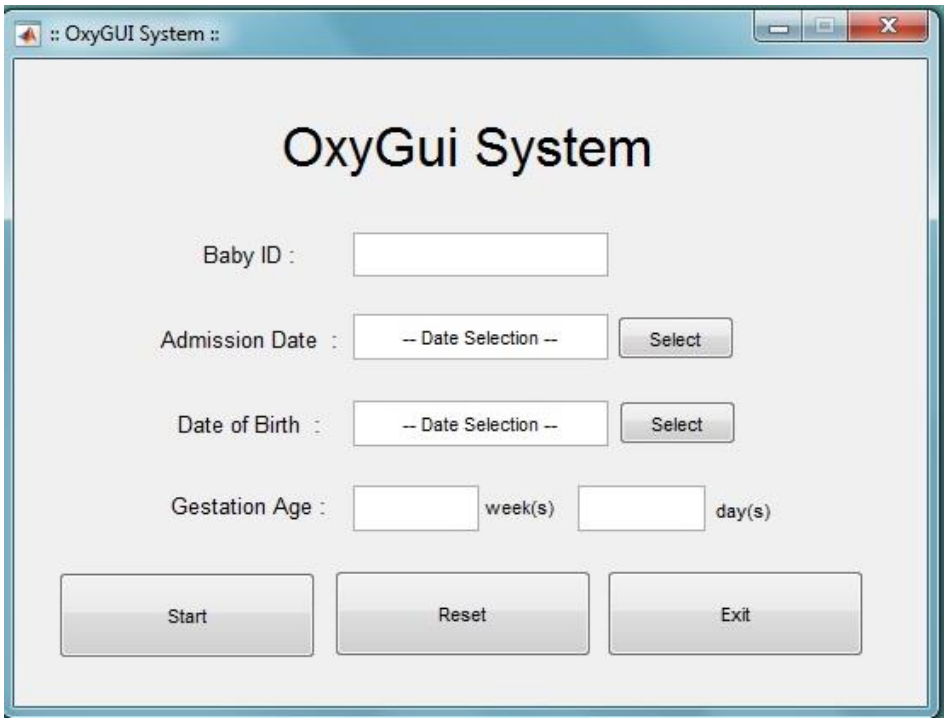


ii. Start OxyGUI System

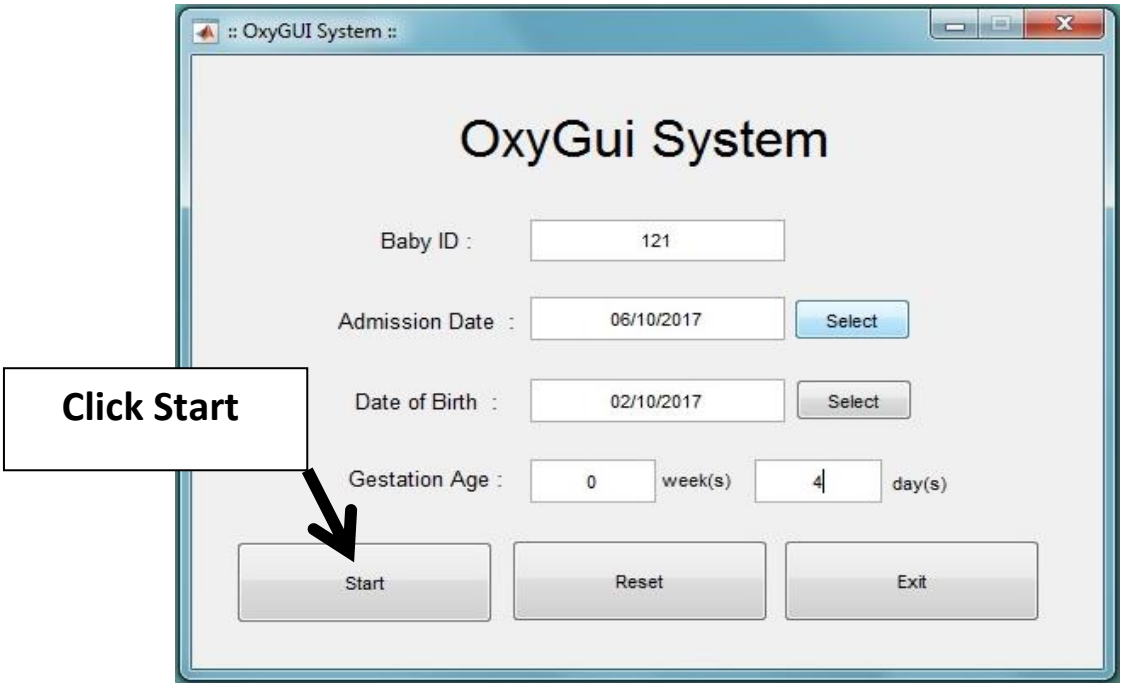
On the main desktop, there is a shortcut link to the **OxyGUISystem** software. Please click on this shortcut as in a picture below.



The main interface of the OxyGUI System appears as below.

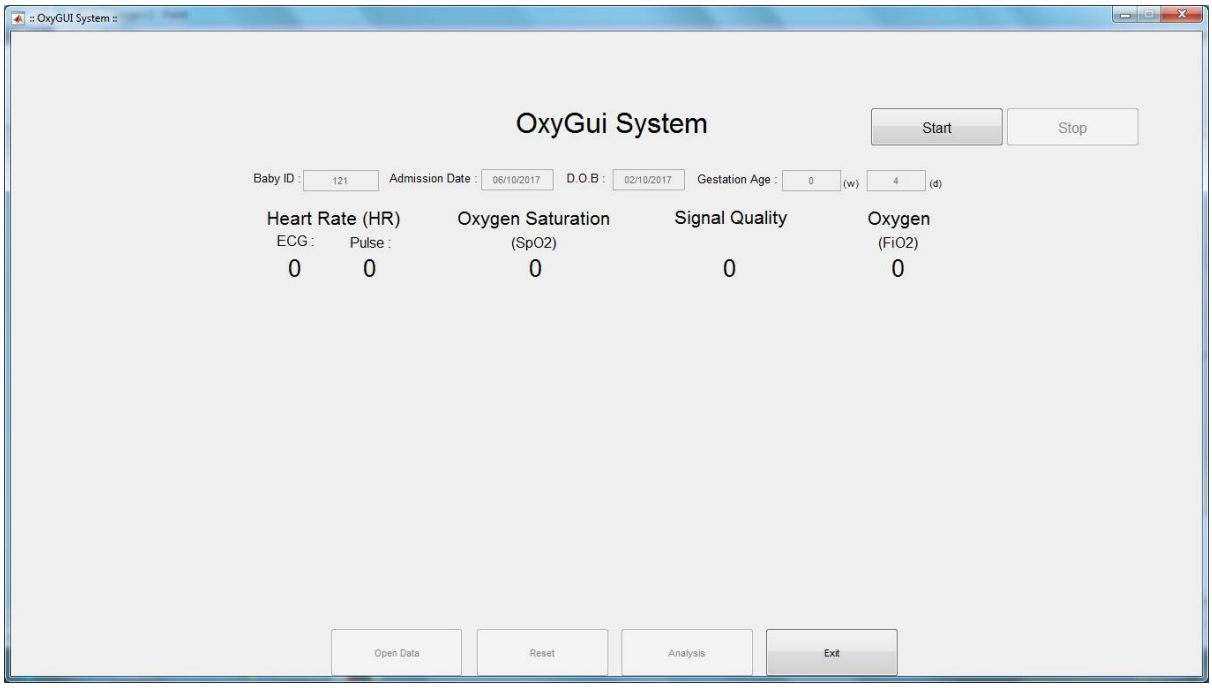


Please key in all the information in this interface. Once it complete, please click 'Start' to proceed with recording data.

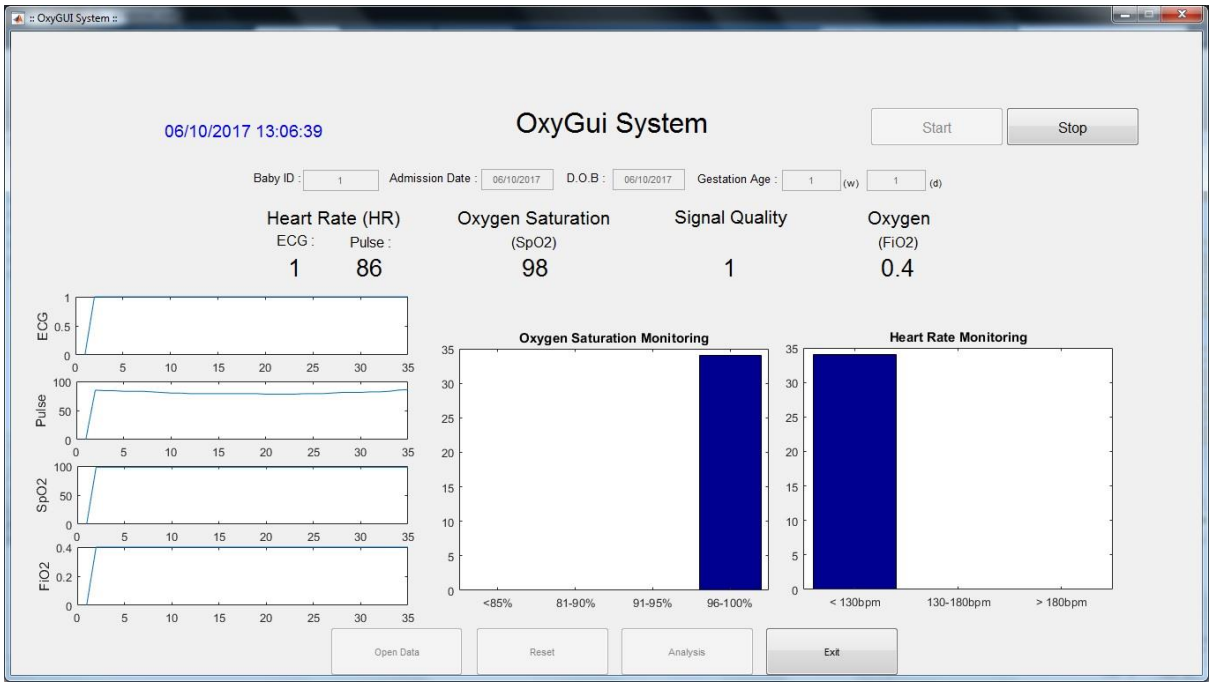


iii. OxyGUI System – Record data

Please click 'Start' to begin with the data recording process.

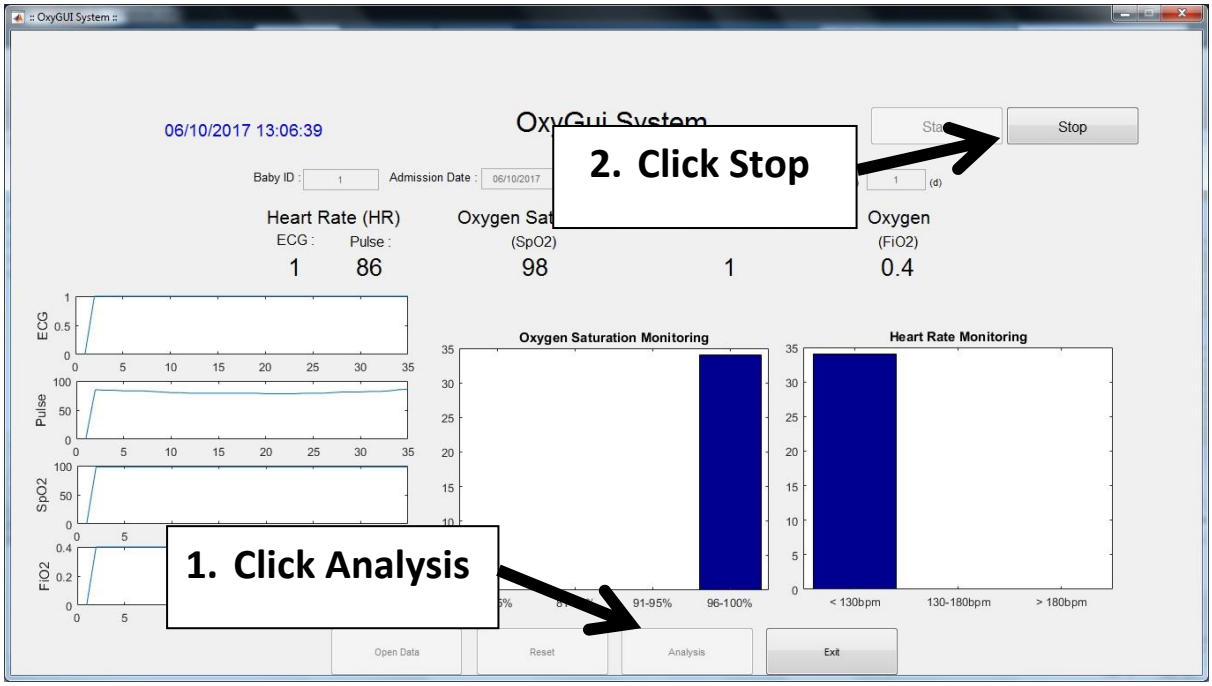


The interface of OxyGUI System below shows the physiological data; ECG-Heart Rate, Pulse, Oxygen Saturation and FiO2 are being recorded. The data will automatically be saved into a text file and named '**baby.txt**'. This file can be seen on the main desktop of this laptop.

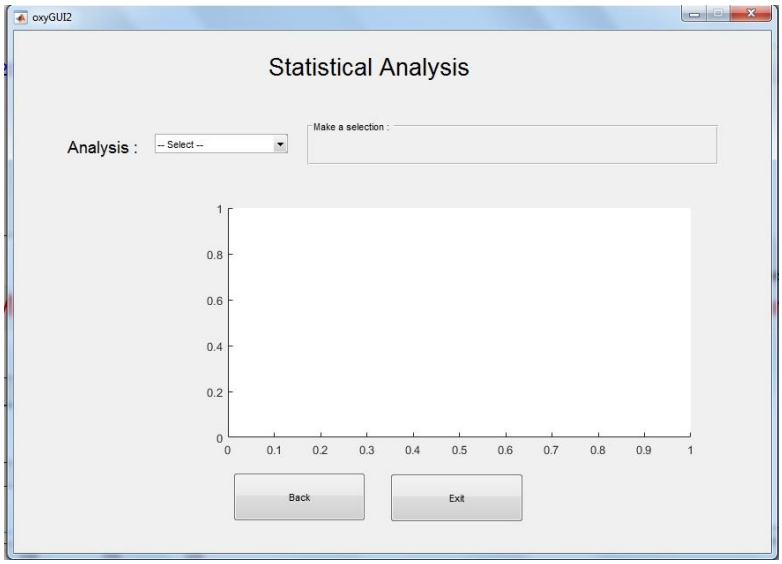


iv. OxyGUI System – Analysis

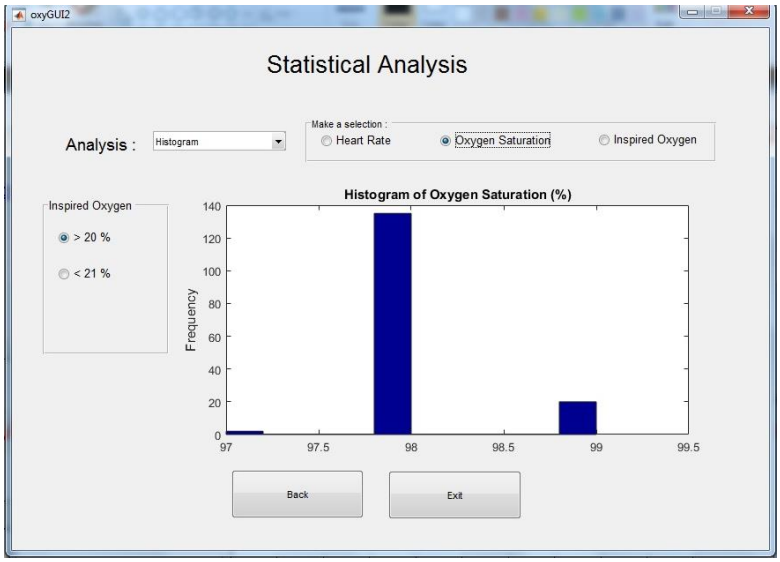
To see the analysis of current data, please click on 'Stop' and then click on 'Analysis' as shown in the picture below. This is because the analysis cannot be seen when the data is being recorded.



The interface of 'Statistical Analysis' appears as below. In this interface, you can view a simple statistical analysis such as Histogram, Scatterplot and Plot of current data.

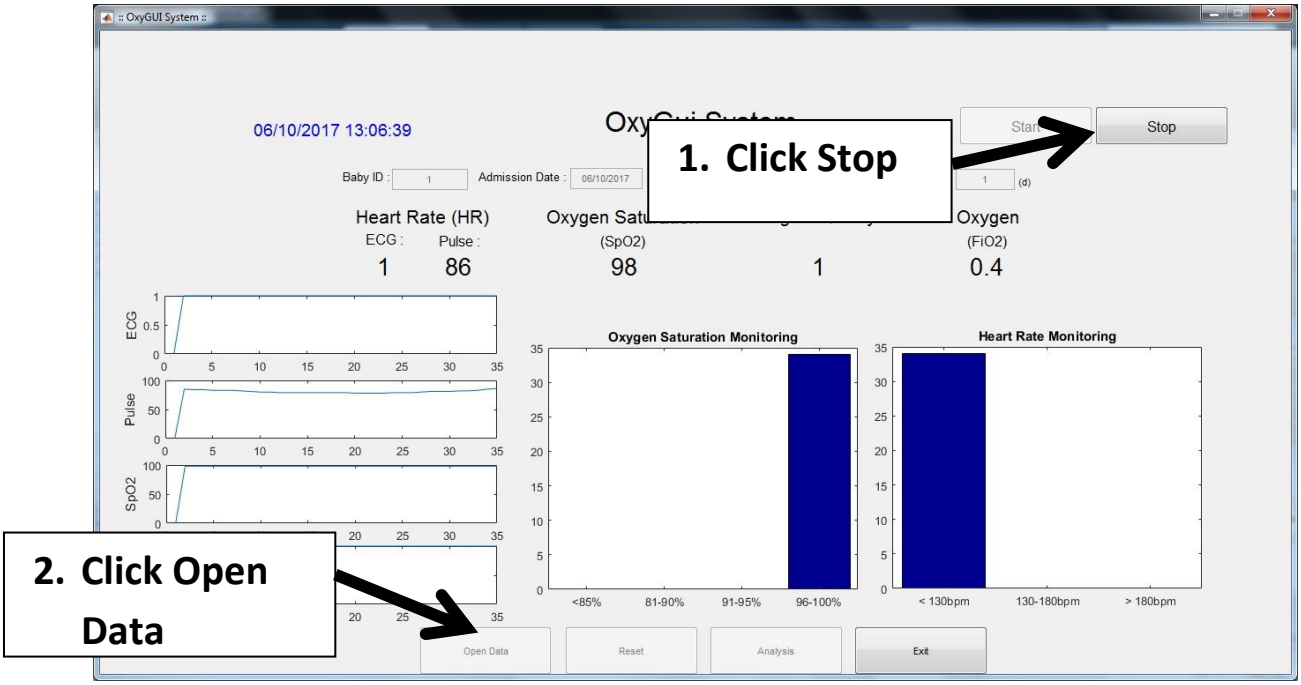


For example, if you choose to view a Histogram analysis, the interface will appear below. Then, you also can select to see the specific Histogram of Heart rate, SpO2 or FiO2 of current data.

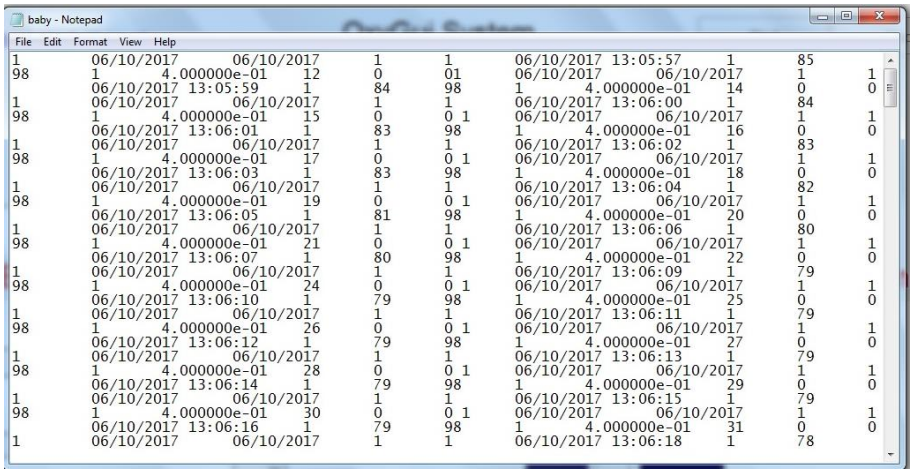


v. OxyGUI System – Save data

To save the current data, please click on '06/10/2017 13:06:39' and then click on 'Open Data' as shown in the picture below.

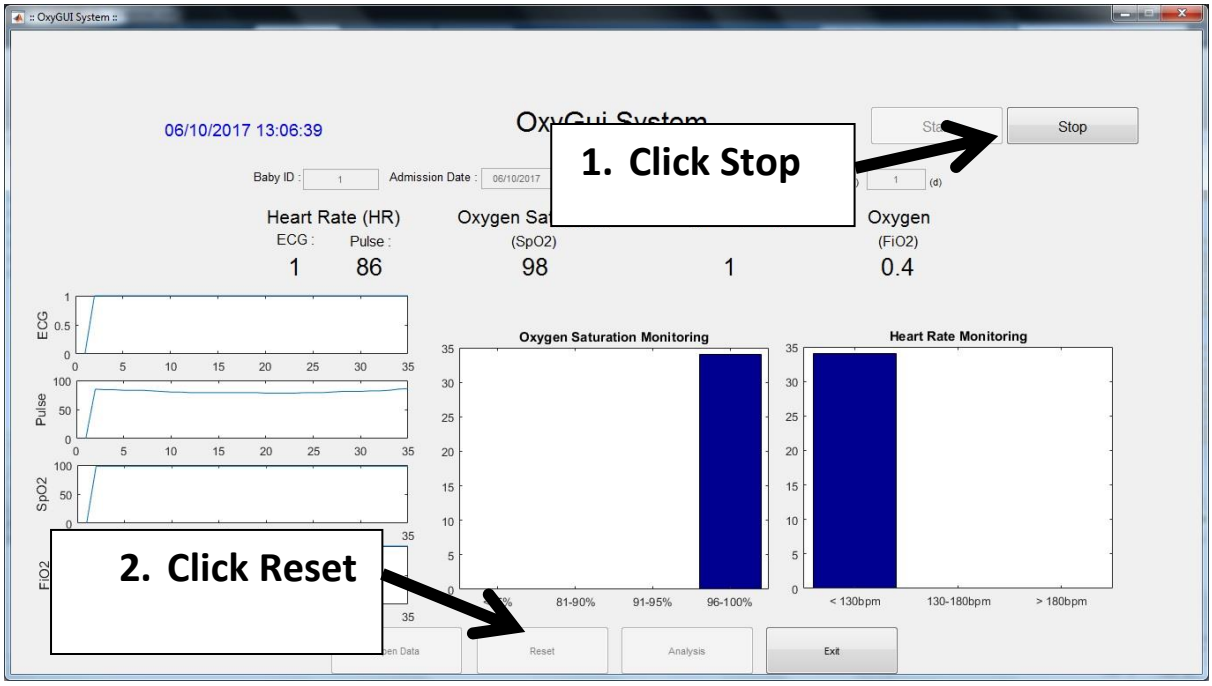


The text file 'baby.txt' appears as below. Then, you can save this file to any folder and rename it as you wish.

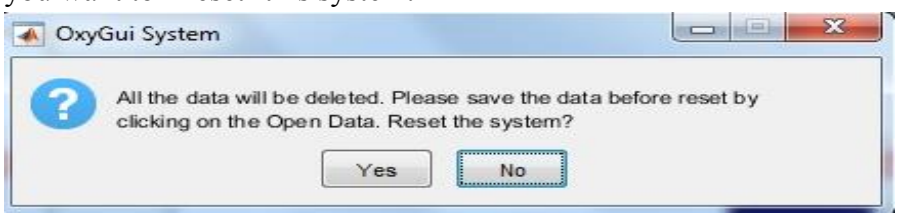


vi. OxyGUI System – Reset system

To reset the OxyGui system, please click on 'Stop' and then click on 'Reset' as shown in the picture below.

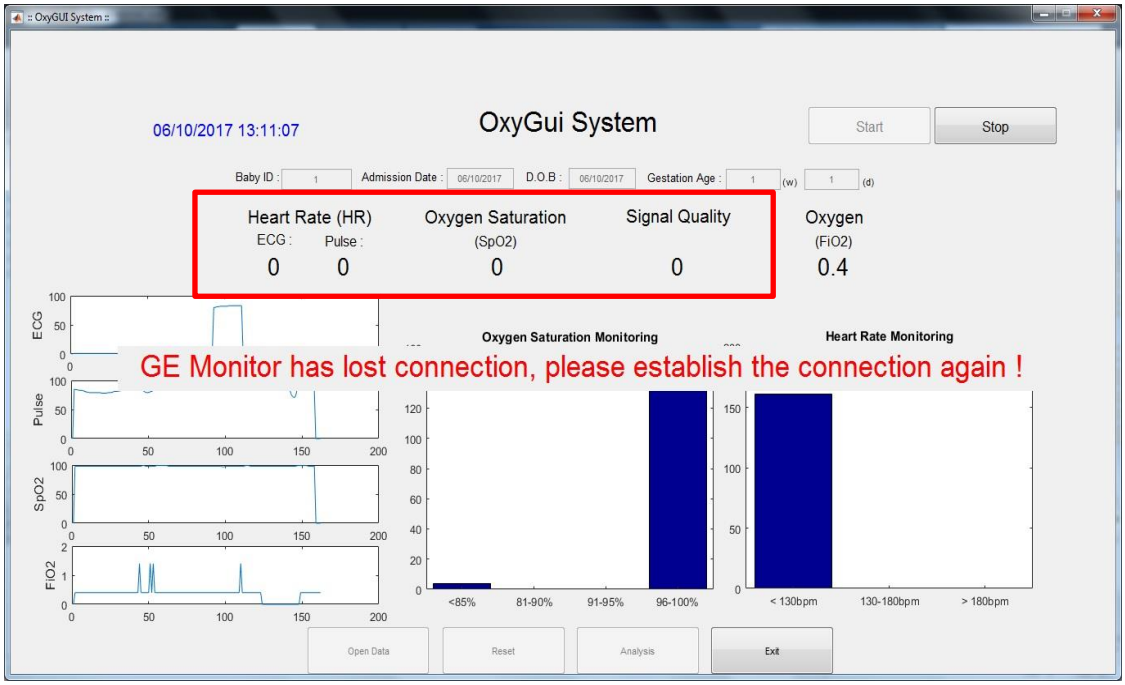


It should be noted, if you click 'Yes' as in the interface below, all the current data will be deleted from this laptop. Therefore, it is recommended that you save the data first before you want to ‘Reset’ this system.



vii. OxyGUI System – Reading errors

If the USB connector between GE monitor and Laptop is not connected, the OxyGUI system will display a warning message as shown in the interface below. Also, all the values of ECG, Pulse, SpO2 and Signal quality will change to be 0. However, the OxyGUI system still records these values together with FiO2. The warning message will disappear when the connection has been fixed again.



If the USB connector between O2 Analyser and Laptop is not connected, the OxyGUI system will display a warning message as shown in the interface below. Also, the value of FiO2 will change to be 0. However, the OxyGUI system still records this value together with ECG, Pulse, SpO2 and Signal quality. The warning message will disappear when the connection has been fixed again.

