The Quantification of Perception Based Uncertainty Using R-fuzzy Sets and Grey Analysis

Arjab Singh Khuman

B.Sc.(Hons), M.Sc.

A thesis submitted in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy
awarded
by

De Montfort University

School of Computer Science and Informatics
Centre for Computational Intelligence
De Montfort University
England

October 2016

Acknowledgments

I would first like to extend a monumental thank you and acknowledge the unwavering support and guidance of my supervisory team, Prof. Yingjie Yang and Prof. Robert John, whom I owe a great deal of gratitude and adulation. Their involvement throughout has seen the acceptance of several peer reviewed works, providing additional gravitas to the scope of my research. I would also like to extend a sincere thank you to Prof. Sifeng Liu, Dr. Jenny Carter and Dr. Simon Coupland, for their injection of much needed enthusiasm and inspiration when it was otherwise lacking. In addition to also allowing me to further enhance my academic abilities through the various appointed positions bestowed upon me. I feel it fitting to thank those members, past and present, of whom have been associated to the Centre for Computational Intelligence. Their support and words of wisdom have proved invaluable, providing not only technical advice, but equally important, advice on well-being and maintaining one's sanity. An acknowledgment section would not be complete if I did not thank my loved ones, family and friends. To them, a profound thank you for putting up with my highs and lows, of which there were many. The tirelessness of your patience and endurance has been inspirational. To you all, I am eternally grateful.

Archie

Abstract

The nature of uncertainty cannot be generically defined as it is domain and context specific. With that being the case, there have been several proposed models, all of which have their own associated benefits and shortcomings. From these models, it was decided that an R-fuzzy approach would provide for the most ideal foundation from which to enhance and expand upon. An R-fuzzy set can be seen as a relatively new model, one which itself is an extension to fuzzy set theory. It makes use of a lower and upper approximation bounding from rough set theory, which allows for the membership function of an R-fuzzy set to be that of a rough set. An R-fuzzy approach provides the means for one to encapsulate uncertain fuzzy membership values, based on a given abstract concept. If using the voting method, any fuzzy membership value contained within the lower approximation can be treated as an absolute truth. The fuzzy membership values which are contained within the upper approximation, may be the result of a singleton, or the vast majority, but absolutely not all. This thesis has brought about the creation of a significance measure, based on a variation of Bayes' theorem. One which enables the quantification of any contained fuzzy membership value within an R-fuzzy set. Such is the pairing of the significance measure and an R-fuzzy set, an intermediary bridge linking to that of a generalised type-2 fuzzy set can be achieved. Simply by inferencing from the returned degrees of significance, one is able to ascertain the true significance of any uncertain fuzzy membership value, relative to other encapsulated uncertain values. As an extension to this enhancement, the thesis has also brought about the novel introduction of grey analysis. By utilising the absolute degree of grey incidence, it provides one with the means to measure and quantify the metric spaces between sequences, generated based on the returned degrees of significance for any given R-fuzzy set. As it will be shown, this framework is ideally suited to domains where perceptions are being modelled, which may also contain several varying clusters of cohorts based on any number of correlations. These clusters can then be compared and contrasted to allow for a more detailed understanding of the abstractions being modelled.

List of Publications

Throughout the duration of this research, the following works have been published, accepted, in press, or otherwise being prepared. The content of this thesis borrows heavily from these published works.

International Journals:

[1] A. S. Khuman, Y. Yang, and R. John. (2016), 'Quantification of R-Fuzzy Sets', *Expert Systems with Applications*, vol. 55, 374-387.

[2] A. S. Khuman, Y. Yang, and S. Liu. (2016), 'Grey Relational Analysis and Natural Language Processing to: Grey Language Processing', *Journal of Grey System*, 28(1), 374-387.

Conference Papers:

[3] A. S. Khuman, Y. Yang, and S. Liu. 'Grey Relational Analysis and Natural Language Processing', *2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS)*, Leicester, 2015, pp. 107-112.

[4] A. S. Khuman, Yingjie Yang and R. John, 'A Significance Measure for R-Fuzzy Sets', *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, Istanbul, 2015, pp. 1-6.

[5] A. S. Khuman, Y. Yang and R. John, 'A Commentary on Some of the Intrinsic Differences Between Grey Systems and Fuzzy Systems', *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, San Diego, CA, 2014, pp. 2032-2037.

[6] A. S. Khuman, Y. Yang and R. John, 'A New Approach to Improve the Overall Accuracy and the Filter Value Accuracy of the GM (1,1) New-Information and GM (1,1) Metabolic Models', *2013 IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, 2013, pp. 1282-1287.

In Press:

[7] A. S. Khuman, Y. Yang, R. John, S. Liu, 'Quantification of Perception Clusters Using R-Fuzzy Sets and Grey Analysis', *2016 International Conference on Grey Systems and Uncertainty Analysis (GSUA2016)*, Leicester, 2016. Aug 8-11. pp. In Press.

[8] A. S. Khuman, Y. Yang, R. John, S. Liu, 'R-fuzzy Sets and Grey System Theory', *2016 IEEE International Conference on Systems, Man, and Cybernetics*, Budapest, 2016. Oct 9-12. pp. In Press.

In Preparation:

[9] A. S. Khuman, Y. Yang, R. John, 'Divergences Between Grey Theory and Fuzzy Theory', *Journal Paper.*

[10] A. S. Khuman, Y. Yang, R. John, 'Perception Based Uncertainty Modelling Using R-fuzzy Sets and Grey Relational Analysis', *Journal Paper*

Contents

Acknowledgments iii

Abstract iv

List of Publications v

List of Figures xiv

List of Tables xvi

1 Thesis Background 2

1.1 Introduction 2

1.2 Motivation 7

1.3 Hypothesis 9

1.4 Attributed Thesis Contributions 10

1.5 Experimentation 11

1.6 Thesis Structure 11

2 Literature Review 14

2.1 Introduction . . . 14

2.2 Fuzzy Theory . . . 15

2.2.1 A Crisp Set . . . 16

2.2.2 A Fuzzy Set . . . 17

2.2.3 Membership Functions . . . 20

2.3 Rough Theory . . . 22

2.3.1 Rough Approximations . . . 25

2.4 Uncertainty Variations and Extensions . . . 28

2.4.1 An Interval-Valued Fuzzy Set . . . 28

2.4.2 An Atanassov Intuitionistic Fuzzy Set . . . 29

2.4.3 A Type-2 Fuzzy Set . . . 29

2.4.4 A Shadowed Set . . . 31

2.4.5 Fuzzy-Rough & Rough-Fuzzy . . . 32

2.4.6 A Fuzzy-Rough Set . . . 33

2.4.7 A Rough-Fuzzy Set . . . 33

2.4.8 An Overview . . . 34

2.5 R-fuzzy Theory . . . 36

2.5.1 R-fuzzy Approximations . . . 37

2.5.2 An R-fuzzy Example . . . 38

2.6 Grey Theory . . . 42

2.6.1 Grey Numbers . . . 45

2.6.2 Grey Sets . . . 49

2.6.3 Grey Analysis . . . 50

2.6.4 Absolute Degree of Grey Incidence . . . 52

2.6.5 Relative Degree of Grey Incidence . . . 53

2.6.6 Synthetic Degree of Grey Incidence . . . 53

2.7 Closing Remarks . . . 54

3 Novel Contributions **57**

3.1 Introduction . . . 57

3.2 The Significance Measure . . . 59

3.2.1 Validation . . . 60

3.2.2 Intended Use . . . 61

3.2.3 A Significance Measure Example . . . 62

3.3 Streamlined Encapsulation . . . 74

3.3.1 Whitenisation . . . 75

3.3.2 A Whitenisation Example . . . 75

3.4 The Enhanced R-fuzzy Framework . . . 82

3.4.1 An Enhanced R-fuzzy Example 84

3.5 Similarity and Distance Measures 93

3.6 An R-fuzzy α-Cut 97

3.6.1 An R-fuzzy Shadowed Set 100

3.7 Closing Remarks 101

4 Empirical Observations **104**

4.1 Introduction 104

4.2 R-fuzzy Restrictions 105

4.3 Confliction 106

4.4 Descriptors 109

4.5 Thresholds 111

4.6 Disjoint Distribution 112

4.6.1 A Disjoint Enhanced R-fuzzy Example 114

4.7 Relationship to Fuzzy 123

4.7.1 Relationship to a Type-1 Fuzzy Set 124

4.7.2 Relationship to a Type-2 Fuzzy Set 124

4.8 Closing Remarks 125

5 Conclusion **129**

5.1 Introduction 129

5.2 The Significance Measure 130

5.3 Streamlined Encapsulation 132

5.4 The Enhanced R-fuzzy Framework 133

5.5 Future Work 134

List of Figures

1.1 The Research Pathway . . . 6

2.1 A Crisp Set . . . 17
2.2 A Fuzzy Set . . . 18
2.3 Multiple Fuzzy Sets . . . 19
2.4 A Triangular Membership Function . . . 20
2.5 A Trapezoidal Membership Function . . . 21
2.6 Rough Approximations for the Set *Walk* . . . 28
2.7 A Type-2 Fuzzy Set . . . 30
2.8 A Shadowed Representation of a Fuzzy Set . . . 31
2.9 A White, Grey & Black Number . . . 45

3.1 A Discrete Visualisation for *AC* . . . 63
3.2 A Possible Continuous Representation for *AC* . . . 64
3.3 A Discrete Visualisation for *NN* . . . 69
3.4 A Possible Continuous Visualisation for *NN* . . . 69
3.5 A Discrete Visualisation for *BN* . . . 70

3.6 A Possible Continuous Visualisation for *BN* 70

3.7 A Discrete Visualisation for *VN* 71

3.8 A Possible Continuous Visualisation for *VN* 71

3.9 A Continuous Visualisation for *NN*, *AC*, *BN* & *VN* 72

3.10 A Continuous Streamlined Representation for *AC* 76

3.11 A Continuous Streamlined Representation for *NN* 78

3.12 A Continuous Streamlined Representation for *BN* 80

3.13 A Continuous Streamlined Representation for *VN* 81

3.14 A Continuous Streamlined Visualisation for *NN*, *AC*, *BN* & *VN* 82

3.15 A Continuous Visualisation for *DR*, *R* & *LR* 86

3.16 The Comparability Between 20 and 25 Year Olds for *LR* 89

3.17 The Comparability Between 20 and 30 Year Olds for *LR* 90

3.18 The Comparability Between 25 and 30 Year Olds for *LR* 91

3.19 The Comparability Between 20 and 25 Year Olds for *R* 92

3.20 The Comparability Between 20 and 30 Year Olds for *R* 93

3.21 The Comparability Between 25 and 30 Year Olds for *R* 93

3.22 The Comparability Between 20 and 25 Year Olds for *DR* 94

3.23 The Comparability Between 20 and 30 Year Olds for *DR* 94

3.24 The Comparability Between 25 and 30 Year Olds for *DR* 95

3.25 α-Cut Representation 99

3.26 A Possible Readjusted Visualisation for *AC* & *BN* 99

4.1 A Strict Interval Visualisation for *M* 107

4.2 Another Possible Visualisation for *M* 107

4.3 A Possible Visualisation for *M* With Included Confliction 108

4.4 A Possible Continuous Visualisation for *VDR*, *DR*, *R*, *LR* & *VLR* 110

4.5 A Derivable Continuous Visualisation for *VDR*, *DR*, *R*, *LR* & *VLR* 111

4.6 A Possible Continuous Visualisation for *DB*, *V* & *LB* 117

4.7 A Derivable Continuous Visualisation for *DB*, *V* & *LB* 118

4.8 The Comparability Between Males and Females for *LB* 119

4.9 The Comparability Between Males and Females for *B* 121

4.10 The Comparability Between Males and Females for *DB* 121

List of Tables

2.1 An Information System 24

2.2 A Decision System 26

2.3 Recorded Noise of Flights 39

2.4 The Collected Subjective Perceptions of Individuals 39

2.5 A Digestive Summary Between Grey, Fuzzy & Prob. Statistics 44

3.1 The Collected Degrees of Significance for Noise Perception 67

3.2 The Whitenisation of *AC* 76

3.3 The Whitenisation of *NN* 78

3.4 The Whitenisation of *BN* 80

3.5 The Whitenisation of *VN* 81

3.6 Human Perception Based on the Variations for the Colour Red 85

3.7 The Degrees of Significance for Each of the Generated R-fuzzy Sets 86

3.8 The Degrees of Significance for Each of the Generated R-fuzzy Sets for the Age Cluster 20 Year Olds 87

3.9 The Degrees of Significance for Each of the Generated R-fuzzy Sets for the Age Cluster 25 Year Olds 88

3.10 The Degrees of Significance for Each of the Generated R-fuzzy Sets for the Age Cluster 30 Year Olds 89

3.11 A Comparable Summary of the Returned Absolute Degree of Grey Incidence for *LR*, *R* & *DR* 92

3.12 The Returned Jaccard Similarity Index For *LR*, *R* & *DR* 96

3.13 The Absolute Difference Between the Grey and the Jaccard Measure 96

4.1 Human Perception Based on the Variations for the Colour Blue 115

4.2 The Degrees of Significance for Each of the Generated R-fuzzy Sets 116

4.3 The Degrees of Significance for Males 119

4.4 The Degrees of Significance for Females 120

4.5 A Comparable Summary of the Returned Absolute Degree of Grey Incidence for *LB*, *B* & *DB* 120

1

Thesis Background

"Uncertainty is a quality to be cherished, therefore - if not for it, who would dare to undertake anything?"

– Villiers de L'Isle-Adam

1.1 Introduction

Uncertainty by its very nature is uncertain in its description, the notion of a singular form of *general* uncertainty is ill defined as it lacks contextualisation. The General Problem Solver, a computer program created by Herbert A. Simon, J.C. Shaw and Allen Newell in 1959, was the first attempt to create a universal problem solver (Newell et al., 1959). While it could handle relatively simple problems such as the Towers of Hanoi, it found great difficulty in providing solutions to problems that were difficult to formalise, and in turn, it would often find itself lost in the exponential possible walkthroughs generated. Even for simple problems, the exhaustive computational burden became problematic. However, the context of *knowing* the domain space in which one would be investigating, allows for a better suited approach. Problem spaces involving the proofs in the predicate logic and Euclidean geometry problem spaces, were ideal domains for the applicability of the General Problem Solver. In much the same way, understanding the environmental and contextual domain of uncertainty; identifying the key facets will allow one to employ a model which is better equipped. Uncertainty in its many guises may invoke certain features, optimisation is one such particularisation, involving the use of specialised tools and techniques. There is also probability and its distribution, where a considerable amount of research has been conducted on the focus on whether a clear event is likely to happen, rather than a vague event's existence (Feller, 1968). The commonly associated statistics is yet another example of a specific specialisation, albeit from a more analytical point of view, to describe, quantify and interpret the various forms of uncertainty that may be prevalent (Chance and Rossman, 2006).

Understanding is key, however, the availability of information can itself lead to different interpretations of a singular observation[1]. If the knowledge itself is subjective, finding an indicative, crisp representation can prove challenging, especially if there are multiple parties involved with varying degrees of confliction. The differences in subjective perception for a given observation can contain both a general consensus, and also individualised specific interpretations. This is not an issue of who should one side with when presented with differing perspectives, but rather, it is one with regards to understanding the uncertainty as to why there is a difference to begin with. It is the quantification of this perception based uncertainty that this thesis is concerned with. Making numerical sense of uncertainty is challenging, but it provides a basis from which one can infer; decisions can be made; classifications can be undertaken; trends discovered; events predicted.

From the initial beginnings of this thesis to its current incarnation, there have been several adaptations and inclusions. The original thesis title and proposal was given as:

'Emerging Uncertainty Models and Their Applications'

The main project objectives were to investigate the typical features of various uncertainties, from which a survey would be completed and the models analysed, investigating both the strengths and weaknesses. This would generate the candidates from which the most prospective model(s) would be considered for further enhancement. It was relatively early on at the beginning of the literature review, that the decision to make use of R-fuzzy sets was decided upon. The notion of an R-fuzzy set was first proposed by Prof. Yingjie Yang and Prof. Chris Hinde (Yang and Hinde, 2010). This new extension to fuzzy sets provided one with a means to represent uncertain fuzzy membership values within a rough

[1] An observation in the context of this thesis is to be understood as being an observation of an abstract concept.

approximation bounding. Therefore, the membership set of an R-fuzzy set is itself a set, from which a wider scope of uncertainty can be captured and inferenced.

One major problem that still continues to exist with relation to type-1 fuzzy sets, is that of deriving a crisp membership function. This is problematic as the membership function itself can be associated with varying degrees of vagueness and ambiguity. There have been several extensions put forward to overcome this pitfall (Deschrijver and Kerre, 2003). Atanassov intuitionistic fuzzy sets (Atanassov, 1986), where a degree of membership and degree of non-membership are presented. Shadowed sets (Pedrycz, 1998), where the evaluation of a membership is scored as either (1), (0) or belonging to the shadowed region $[0, 1]$. Interval-valued fuzzy sets (Sambuc, 1975), where the membership of an individual element is characterised as an interval itself. Type-2 fuzzy sets (Zadeh, 1975a,b,c), where the membership function itself is a type-1 fuzzy set. These *new* approaches involve the use of intervals, multiple parameters and additional fuzzy sets to describe the uncertain membership function values of fuzzy sets. However, there is still a problem, these new extensions are not able to recognise or distinguish between different object values within their shadow areas or intervals (Khuman et al., 2016a). The main reason for choosing an R-fuzzy approach to provide the foundation for this thesis, is that it allows for the possibility of multiple values to represent a crisp membership function.

The ongoing interest in type-2 fuzzy logic as a higher order form of fuzzy logic, has received a lot of attention. The use of interval type-2 fuzzy logic and the generalised approach of type-2 fuzzy logic has garnered much interest, particularly for its ability to handle higher degrees of uncertainty. As a result, its application areas are varied, but considerable work has been undertaken in clustering, classification and pattern recognition. A thorough and systematic review of type-2 fuzzy logic applications was undertaken by Melin and Castillo (2014). A vast majority of current type-2 fuzzy applications are concerned with interval type-2 fuzzy logic, which when compared to a generalised type-2 fuzzy approach has considerably less computational overhead. Interval type-2 fuzzy logic implies that every value in the secondary grade of membership be given a degree of membership equal to 1, therefore, as no new aspects of uncertainty are captured, only the *footprint-of-uncertainty* is used. It is widely agreed that a generalised approach will indeed allow for better management and the handling of uncertainty, compared to that of the interval type-2 approach. However, the problems associated to the complexities that one has to endure to make use of a generalised approach, is the reason why an interval approach is often ultimately chosen.

R-fuzzy sets encapsulates uncertainty via the use of rough sets to approximate the uncertain fuzzy membership values of a fuzzy set. By making use of the approximation bounding from rough set theory, an R-fuzzy set provides the functionality to encapsulate membership values from an entire populous, all within a single set, no matter how conflicting. The fundamental differences in what a rough set and fuzzy set capture, are also reasons as to why the hybridised concept of an R-fuzzy set works so well. A rough perspective is more concerned with ambiguity, a lack of information, whereas a fuzzy perspective is more concerned with vagueness, a lack of clear, sharp definable boundaries. As a result there have been several hybridisations between fuzzy sets and rough sets to allow for greater versatility in encapsulating uncertainty; Bodjanova (2007), Deng et al. (2007), Dubois (1980), Dubois and Prade (1990), Huynh and Nakamori (2005), Jensen and Shen (2008, 2009), Nanda and Majumdar (1992), Pawlak and Skowron (2007), Radzikowska and Kerre (2002), Sun et al. (2014), Wu et al. (2003), Xu et al. (2012), Zeng et al. (2015), all of which mainly incorporate the use of equivalence and similarity relations. The concept of an R-fuzzy set, was the first to use rough approximations of uncertain fuzzy membership values.

After reviewing some of the alternative approaches to uncertainty modelling, it became apparent that an R-fuzzy set was indeed the best choice. Not only for its greater breadth of encapsulation when compared to others, but also for its capacity for enhancement. This coupled with the fact that Prof. Yang constituted one half of the supervisory team, it made logical sense to extend his work of R-fuzzy sets. The contribution of the significance measure and the adage of grey analysis, for further inspection of the returned degrees of significance, provides for a framework which allows for the encapsulation of more complex uncertainty, while also offering more insight and greater levels of inference.

This thesis makes several novel contributions and enhancements to the R-fuzzy concept, in order to create a framework more capable of capturing a greater detail of uncertainty, and equally to provide one the means of inferencing in greater resolution. While the framework has been shown to work optimally within domains associated to perception modelling, it is believed that the versatility of the framework will allow for its deployment

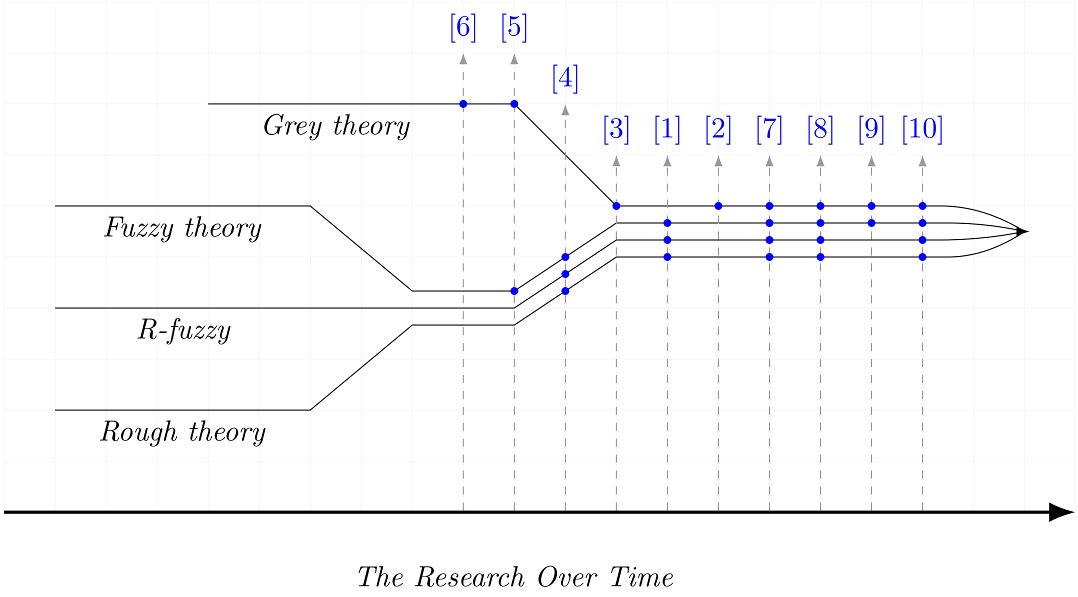


Figure 1.1: The Research Pathway

in other areas. A visualisation, with an indicative timeline over the course of the research is presented in Figure 1.1. The use of grey theory should not be seen as an afterthought, as the investigation into this paradigm was encouraged from the early stages of the research. The adage of grey methodologies has enhanced the overall resolution of detail one can expect, both from the returned results and the provided metrics.

The numeric references imposed onto the figure are directly correlated to the academically recognised contributions of this thesis, as presented in the List of Publications. Each publication used either a single paradigm or a combination of multiple concepts, as indicated by the blue circles and intersecting vertical slices. For example, the work associated to [4], is given as:

[4] A. S. Khuman, Yingjie Yang and R. John, 'A Significance Measure for R-fuzzy Sets', *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, Istanbul, 2015, pp. 1-6.

A publication that proposed the notion of the significance measure, hence why the blue circles are only concerned with *Fuzzy theory*, *R-fuzzy* and *Rough theory* time lines.

1.2 Motivation

The motivation for the enhancement of R-fuzzy sets comes from the *want* to extend an already impressive model for uncertainty encapsulation. However, it is rather surprising that there has been a lack of implementation of R-fuzzy in real world applications. This was another contributing factor to making it the basis of this thesis. If it can be shown and demonstrated that the enhancements that this thesis puts forward, are not only feasible, but intuitive and effective, then there should be no reason that the improved R-fuzzy framework be utilised more applicably, with regards to handling the uncertainty associated to perception based domains.

The newly proposed significance measure is able to quantify the importance of each and every membership value contained within any generated R-fuzzy set. Not only can it quantify, but it can also act as a validator for every contained uncertain fuzzy membership value, contained within the fuzzy membership set J_x. If the membership value is contained in the lower approximation, the returned degree of significance will always be a 1. If the returned degree of significance is a 0, then that specific fuzzy membership value has been disregarded as a viable representation. Any returned value within the interval [0,1] signifies that the membership value has some importance, to some degree, relative to its descriptor.

According to Klir et al. (1997), there exists three kinds of general uncertainty. From an empirical level, uncertainty is often associated with any type of measurement. Resolution can be a cause for concern when involving exactness, 0.1 is different from 0.01 as it to 0.001, and so on. There are an infinite number of variations that exist simply between the interval $[0, 1]$. From the cognitive level, uncertainty exists in the vagueness and ambiguity associated with natural language. This was the basis for the published work when using grey theory for the analysis of natural language processing, from which it was deemed to be an affective adaptation to employ within the R-fuzzy framework (Khuman et al., 2015b, 2016d). The meaning of a word from one individual to another may not be exact, however, an overlap between understandings can act as common ground; an agreement to the sentiment of the meaning of the word. At a social level, uncertainty can be used to benefit an agenda, it is at this level the interaction between privacy, secrecy and propriety occurs.

There could be several root causes for the existence of uncertainty, the information associated to the problem may be inherently noisy or incomplete, riddled with contradictions, vague and ambiguous. These deficiencies may result in sub-faceted aspects of uncertainty; uncertainty within uncertainty. From this, three states of uncertainty are given: *Vagueness* - associated to fuzzy with respect to the imprecise, vague boundaries of fuzzy sets. *Imprecision* - with regards to non-specificity of the cardinalities of sets and their alternatives. *Discord* - with regards to strife which expresses conflicts and contradictions of the various sets of alternatives (Klir and Wierman, 1999; Klir et al., 1997).

The need for models to encapsulate higher complex uncertainties, is the same need that makes a generalised type-2 fuzzy approach very enticing. Lessening the burden associated to a generalised approach would allow for its global applicability to reach new heights. The works of Wagner and Hagras (2010, 2013), Melin et al. (2014) and Mendel et al. (2009), are several attempts to alleviate the burden of computational complexity. The R-fuzzy and significance measure adaptation is yet another attempt to allow for a more efficient means to cater for higher orders of complexities. As it will be shown, the approach allows for the discrete values to be projected onto a continuous representation, allowing for robustness when inquiring about values not necessarily encompassed within the original R-fuzzy set. As such, there is a link to that of a general type-2 fuzzy set, one which will be later described in this thesis. As an additional contribution of this thesis, is the creation of a *streamlined* set for encapsulating the results of the significance measure. The original set was created using the amplitude attributed to each triggered uncertain fuzzy membership value, using a streamlined set allows for the same level of encapsulation with the minimal use of parameters. These streamlined sets themselves are based on the triangular and trapezoidal membership functions.

The ability to distinguish a membership value from one another allows for discernibility to be ensured. There is no need to concern oneself with the loss of information, as would be the case if an interval approach was adopted. An R-fuzzy and significance measure pairing allows for greater detailed information to be inferred from, which allows for better uncertainty management. This again is extended when making use of grey theory techniques, with the specific incorporation of the absolute degree of grey incidence. This allows for the metric spaces between perception clusters to be quantified, measuring the shift in perception from one collection of observations to the next.

1.3 Hypothesis

The following research hypothesis can be seen as the focus of this thesis:

'As a traditional R-fuzzy approach has already been shown to be able to encompass a greater breadth of uncertainty as compared to other uncertainty models; each and every membership value encapsulated can be distinguished from other encapsulated values. However, the importance of the values of the upper approximation lacks any type of inspection to determine how close they were deemed to be universally accepted. Additionally, R-fuzzy sets generated from the same initial data cannot be compared against one another as it does not cater for this functionality. With these short comings, the applicability of R-fuzzy is somewhat limited, and would partially answer the question as to why it is not as utilised as it should be. The creation of the significance measure provides a means to quantify and establish the conditional probability of any membership value relative to its descriptor. The thesis also explores the added benefit of employing grey analysis, via the notion of sequence generation. The improvements to the standard R-fuzzy framework that this thesis puts forward, allows for one to garner a greater amount of detail; detail which is of more resolution. By extending the current R-fuzzy framework by the novel contributions of this thesis such as the significance measure, the streamlined encapsulation concept, and the likes of the R-fuzzy α-cut for additional post analysis, ultimately allows for the framework to be enhanced. By providing one with a model that is more robust, versatile and resolute, the R-fuzzy approach for uncertainty modelling may become more applicable and favourable, when compared to other more established concepts.'

The short coming of the R-fuzzy concept was the motivation to provide it the means to be more robust and versatile, all the while, allowing for an additional dimension of complexity to be captured. Allowing for greater levels of detail to be contained, comes the need for a more detailed form of inspection. The divergences between R-fuzzy sets based on perception clusters can now be quantified, based on the metrics of returned sequences generated using the significance measure. The use of grey methodologies to provide for the analysis of sequences, as it will be shown, allows for a great deal of resolution.

1.4 Attributed Thesis Contributions

The novel contributions attributed to this thesis allows for an approach that encapsulates more detail, from which a greater level of analysis can be undertaken. This provides additional functionality to the concept of R-fuzzy and its application with regards to perception based uncertainty, which is inherently associated to subjective uncertainty. Based on this additional dimension of uncertainty, came the need for a higher degree of analysis. This was the reason of incorporating grey system theory, more specifically, the use of the absolute degree of grey incidence. The literature review will provide a more in-depth look at why grey theory was ultimately chosen for the analytical component of the new R-fuzzy framework. It has the ability to quantify the metric deviation between sequences of isolated clusters, allowing for the difference in perception between clusters to be measured and quantified. If it can be seen that there is a trend, linear or not, between perception as one propagates through clusters of cohorts, this in turn can aid policy and decision making. Knowing how an observation may be perceived before it is actually perceived, increases effectiveness, and streamlines efficiency.

The main contribution of this thesis is undoubtedly the significance measure, this allows for the R-fuzzy methodology to be paired with grey analysis, without such a component the introduction of grey analysis would not be possible. It is this enhanced R-fuzzy framework that this thesis is primarily concerned with. The significance measure allows for the detail and subjectivity encapsulated when using an R-fuzzy approach to be maintained and translated to a sequence, so that it may then be analysed via the use of grey techniques. The additional contributions of this thesis are with regards to the analysis and post-analysis after the use of the enhanced R-fuzzy framework. The likes of the streamlined encapsulation approach for providing the minimal configuration needed to correctly contain all relative and triggered uncertain fuzzy membership values. In addition, the thesis also describes the notion of an R-fuzzy α-cut that is essentially an α-cut but applied in a R-fuzzy setting. This is closely linked to the idea of an R-fuzzy shadowed set, using threshold values indicative of the most extreme precipices of inclusion and non-inclusion to the R-fuzzy set. These additional contributions allow for one to employ less restrictive means for analysis and understanding of the results.

1.5 Experimentation

The various examples contained within this thesis used real subjects, of which the majority were students belonging to taught modules that I myself am associated with. EXAMPLE 1 was taken from the original paper by Prof. Yingjie Yang and Prof. Chris Hinde (Yang and Hinde, 2010) that first proposed the notion of an R-fuzzy set. This would provide the foundation from which to extend and allow for conciseness throughout. EXAMPLE 2 uses EXAMPLE 1 as the basis from which to expand upon by demonstrating the use of the significance measure. EXAMPLE 3 also expands upon EXAMPLE 1, by demonstrating the use of *whitenisation* on the same initial data.

EXAMPLE 4 used a collection of students, friends and family members. The students belonged to the taught module of Fuzzy Logic and Knowledge Based Systems. The friends and family members were of varying ages and demographics, as too were the students. This included a mix of both female and male participants. There was no strict selection criteria, if they were willing to participate, they were allowed to. In much the same way, EXAMPLE 5 also made use of students, friends and family members, of which the majority from EXAMPLE 4 participated.

As it will be shown, the experiments reveal a wealth of additional detail, from which further experiments can be undertaken with specific requirements on inclusion and expectation. As this thesis is concerned with perception based uncertainty, one would feel more inclined to not be restrictive in terms of participation. The conclusion reached from the experiments contained in this thesis can be used to provide the background knowledge in the configuration of future experiments.

1.6 Thesis Structure

CHAPTER 2 provides one with a literature review and the background information needed for understanding the successive content of this thesis. A considerable amount of detail is given to the component parts that make up an R-fuzzy set; rough set approximations and fuzzy sets. Grey system theory and its foundational underpinnings, along with the absolute degree of grey incidence are also presented, highlighting the main ideology and facets used to enhance the R-fuzzy framework.

Chapter 3 describes the novel contributions attributed to this thesis, it proposes the notion of the significance measure, providing a full overview. The R-fuzzy and significance measure pairing is also investigated, demonstrating the enhanced R-fuzzy framework. Also demonstrated is the notion of a streamlined encapsulating set, one which is created using the minimal number of parameter values. One could make use of this if simply considering only the R-fuzzy and significance measure combination, without the addition of grey analysis. Such is the importance of the significance measure, it provides one with the component needed so that grey analysis can be undertaken. The significance measure acts as the facilitator, allowing for the output of an R-fuzzy set to be translated into a sequence, such that it can be given as the input for the absolute degree of grey incidence. Similarity and distance measures are also described, as too are the notions of an R-fuzzy α-cut and an R-fuzzy shadowed set, to aide post result analysis.

Chapter 4 presents the empirical observations of the proposed framework, linking together the use of an R-fuzzy foundation, the secondary grade membership that the significance measure allows for, and the use of grey analysis for the quantification of the metric spaces between sequences. The adage of worked examples better demonstrates the applicability of the proposed framework. In addition, the link of an R-fuzzy and significance measure pairing to that of fuzzy set theory is also discussed.

Chapter 5 puts forward a discussion and concluding summary of the research contained within this thesis. A breakdown of what has been created and the contribution to the knowledge base are also discussed, as too are the future enhancements that could be adopted.

2

Literature Review

> *"In these times I don't, in a manner of speaking, know what I want; perhaps I don't want what I know and want what I don't know."*
>
> – Marsilio Ficino

2.1 Introduction

This thesis ultimately combines the concepts of R-fuzzy and grey theory, creating a framework which is fundamentally R-fuzzy, while utilising grey analysis. The R-fuzzy approach itself is a hybridisation of fuzzy and rough set theory, the foundational literature on both of these fields is put forward to allow one to understand the core component parts of this *new* framework. As the R-fuzzy approach was decided upon relatively early on in the research, this brought about the need to understand its constituent parts, hence the ordering of this chapter. Section 2.2 introduces the notion of fuzzy, an established paradigm with regards to uncertainty modelling, one which has garnered much attention. The concept of fuzzy membership values and fuzzy sets are also described and also demonstrated. Section 2.3 puts forward the preliminaries associated to rough set theory. As R-fuzzy makes use of rough approximations, it is a vital construct of R-fuzzy and the new proposed framework. Section 2.4 describes some of the variations and extensions of fuzzy and rough which were also investigated in the initial stages of this research. Understanding the shortcomings that these alternative models have, allows for more credence to be given to the choice of why an R-fuzzy approach was ultimately chosen. Section 2.5

presents the concept of an R-fuzzy set, as was first proposed by Yang and Hinde (2010). Section 2.6 presents the paradigm of grey theory, originally proposed by Ju-Long (1982). Much like that of the fuzzy section, the concepts of grey numbers and sets are also described. It is also in this section that the absolute degree of grey incidence is introduced; the analysis component of the enhanced R-fuzzy framework. Section 2.7 remarks upon the chapter, and the justification of deciding upon a predominately R-fuzzy direction. The aforementioned sections set about providing one with a concrete understanding of the framework that this thesis is underpinned by.

2.2 Fuzzy Theory

With the pursuit of precision and exactness, comes the need to cater for increased levels of imprecision. This also implies additional costs, whether it be computational overhead, a time critical aspect, or the like. The amount and the quality of information available correlates directly with the amount of uncertainty involved. The more known regarding the environment, the less uncertainty there is as the formalisation of a possible solution, or solutions will be more certain. The amount of uncertainty will increase in relation to the amount of, or lack of knowledge available. This highlights the nature of a knowledge-gap, which can be understood as the level of understanding one may have, given the amount of information available, contrasted against an *ideal* understanding. The more information that is known regarding the abstraction, the smaller the scope of the uncertainty involved.

Conventional set theory makes use of Boolean logic (Boole, 1847, 1854), whereby an object is categorised as absolutely belonging, or absolutely not belonging to an inspected set (Cantor, 1895). The use of crisp boundaries applies an inherent level of strictness to what the set can model, instances where only two outcomes are allowed, such as an integer being either odd or even; such instances are easily handled using a classical approach. However, there is an underlying need to encapsulate uncertainty that is associated to vagueness when considering human based perception. Human nature and inferencing does not work in such a precise and crisp manner, a humanistic approach needs to cater for the existence of imprecision based uncertainty, along with vagueness. The understanding of a set from a classical perspective is not a fitting synthesis for human intuition. The notion of *mereology* described by Lesniewski (1929), considered the idea of an object being partially included in a set, this was the basis for the formulation of Max Black's vague set (Black, 1937), created in the 1930s.

The building blocks of any fuzzy implementation involves the use of fuzzy sets, first proposed by Zadeh (1965). A fuzzy set can be seen as an extension of the ideology of a vague set. From its inception fuzzy logic has been further expanded upon to establish itself as a powerful paradigm for modelling uncertainty (Zadeh, 1973). As logic is associated to propositions, fuzzy logic can be seen as the calculus of fuzzy propositions. Mathematical applications for precise reasoning will often need crisp understandings, however, this becomes problematic when concepts such as natural language are involved. Linguistic vernacular can be inherently vague, with a prevalent amount of ambiguity. Our daily existence will often be littered with varying degrees of uncertainty, further invoking various aspects of specific uncertainties (Zadeh, 1999a).

> *"... the remarkable human capability to perform a wide variety of physical and mental tasks without any measurements and any computations. Familiar examples of such tasks are parking a car; driving in heavy traffic; playing golf; understanding speech, and summarizing a story. Underlying this remarkable ability is the brain's crucial ability to manipulate perceptions, perceptions of size, distance, weight, speed, time, direction, smell, color, shape, force, likelihood, truth and intent, amongst others."*
>
> – L. A. Zadeh (1999)

2.2.1 A Crisp Set

The use of a classical set for the modelling of unclassical behaviour will often fall foul when considering the vagueness of uncertainty. For example, the abstract concept of *tall* cannot be universally defined, a single crisp value cannot be put forward as an indicative representation that is agreed upon by all. What is tall to some may not be as tall to others. Figure 2.1 provides a visualisation of what a typical crisp bounded set may look like. The plot in the figure describes any person being $6'$ or taller as a validated member

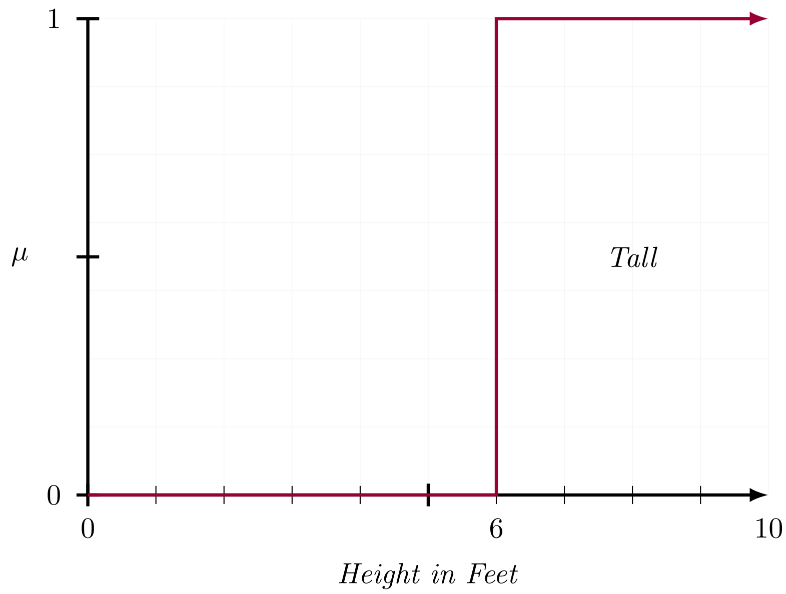


Figure 2.1: A Crisp Set

of the set *Tall*. However, using this precise and strict definition, it would neglect any instance of anything less than $6'$. It can be generally assumed that $6'$ is indeed tall, but so too is $5'11''$, at least to some extent. The only association one can attribute to a value is if it is included in the set *Tall* or not. The problem now becomes one of determining the bounding of the set, to realistically encapsulate all common held assumptions of what satisfies the notion of being tall. This echoes the sentiment of Sorites paradox, arising from a vague predicate. A fuzzy perspective will allow for a more forgiving approach, one which enables an object to have partial belongingness.

2.2.2 A Fuzzy Set

The most fundamental aspect of fuzzy set theory is its understanding of numbers. A fuzzy number is ideal for describing linguistic phenomena, where an exact description of its state is unknown. Fuzzy numbers were first introduced by Zadeh (1975a), for the purpose of approximating real numbers which deal with uncertainty and imprecision associated to quantities. It has great scope when approximating height, or weight and other such uncertain abstractions. The apex of a fuzzy number will generally be the only point where an object can be given a maximum degree of inclusion equal to 1. The varying degrees of

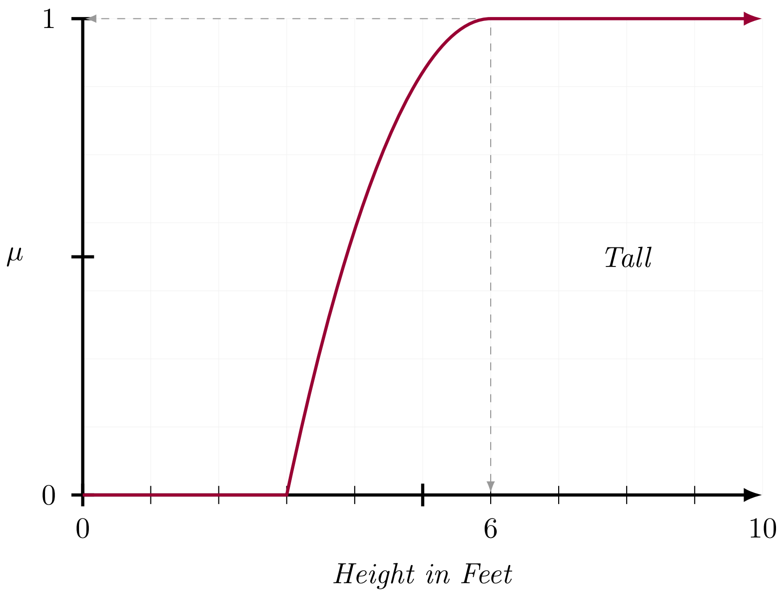


Figure 2.2: A Fuzzy Set

membership for other objects will be indicative to their proximity to the apex. Fuzzy sets extend the notion of fuzzy numbers to allow for more versatility. If one was to describe the set *Tall* as seen in Figure 2.1, using a fuzzy approach, a possible visualisation may look like the plot contained in Figure 2.2. In this plot, the inclusion of other possible values that could be deemed as tall, values such as $5'11''$ would also be included, but to a lesser degree than $6'$. Following this understanding, $5'10''$ would also be a viable candidate for inclusion, but to a lesser degree than $5'11''$, and so on. Using a fuzzy perspective for encapsulation, one is able to relax the expected strictness one would associate with a crisp set. Not only does a fuzzy set allow for this more harmonic understanding of uncertainty, but it also is able to fall back to a classical interpretation if need be. The degree of belongingness may be that of absolution, or absolutely not, in which case, a fuzzy set can replicate a crisp set (Klir and Folger, 1988). In essence, the process of mapping a membership value to an object is known as *fuzzification*. It is only when considering that an object may have partial belongingness, does the strength and applicability of a fuzzy set become apparent.

Definition 1 (Fuzzy Set): Let $\mathbb{U}$ represent the universe and let A be a set in $\mathbb{U}$ $\left(A \subseteq \mathbb{U}\right)$. The fuzzy set A is a set of ordered pairs given by the following expression:

$$A = \left\{\left\langle x, \mu_A(x)\right\rangle \,\middle|\, x \in \mathbb{U}\right\} \tag{2.2.1}$$

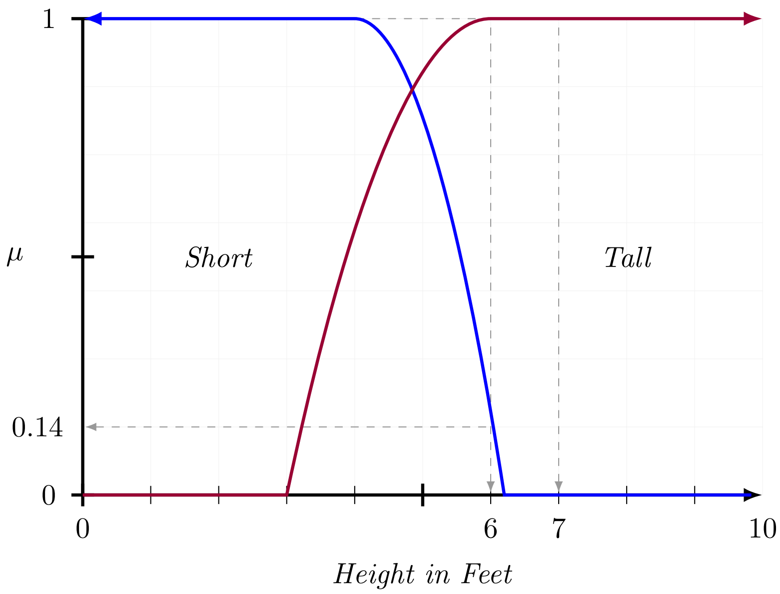


Figure 2.3: Multiple Fuzzy Sets

A fuzzy set on its own can only allow for a certain amount of functionality, the combination of multiple fuzzy sets allows one to extensively model an abstract concept, that would otherwise be very difficult to represent using a crisp understanding. As it is a set of ordered pairs, the object x is associated to a degree of inclusion $\mu_A(x)$, the same x may belong to more than one set, and as such may be attributed to multiple degrees of inclusion. A fuzzy set goes against the traditional approach of classical set theory, by allowing an object to belong to different sets by varying degrees of membership. Such is the methodology of fuzzy sets, the law of the excluded middle and the law of contradiction are ignored, these two prevalent laws would stop an object from belonging to more than one set if a crisp perspective was used. Continuing with the notion of tall, Figure 2.3 demonstrates how using an additional fuzzy set, one can allow for a more humanistic approach in understanding the significance of any given value. This plot contains an additional set labelled *Short*. The value $6'$ in this instance can be seen to have an absolute degree of association to the set *Tall* with a returned degree of membership of 1, and a partial degree of inclusion to the set *Short* with a returned degree of membership of 0.14. If one inspects the object value $7'$, it can be seen to have absolute inclusion to the set *Tall*, and complete non-inclusion to the set *Short*. This logical assumption that being $7'$ would never be regarded as being *Short* can be easily catered for, as too can a plethora of other abstract notions.

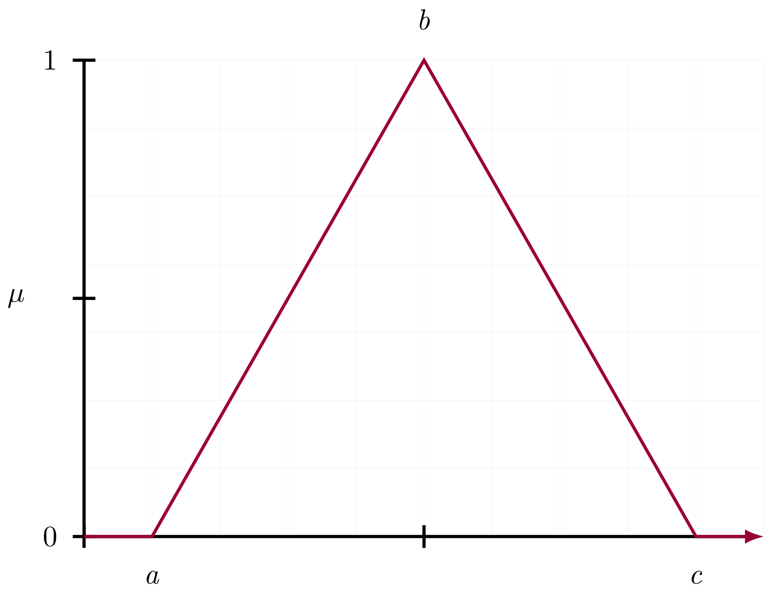


Figure 2.4: A Triangular Membership Function

2.2.3 Membership Functions

A fuzzy set is defined by its membership function, it is what gives a fuzzy set its shape. For example, the fuzzy set A will have a corresponding fuzzy membership function of μ_A, this will determine the degree of membership equal to, or within the range of the interval $[0, 1]$, for each value x in the universe of discourse X. A fuzzy membership is defined as follows:

$$\mu_A(x) : X \rightarrow [0, 1] \tag{2.2.2}$$

There are several varying membership function types, but this thesis is concerned with that of the triangular and trapezoidal membership functions, as seen in Figure 2.4 and Figure 2.5, respectively.

Inspecting the plots, one can see that a certain amount of similarities exist between the two. The only significant difference is that the trapezoidal function employs the use of a plateau region, which essentially acts as an interval, as any object to fall within this region will have an associated membership degree of absolute 1. The triangular membership as the one presented in Figure 2.4 is symmetric, however, they do not need to be, they can be skewed to better model abstraction.

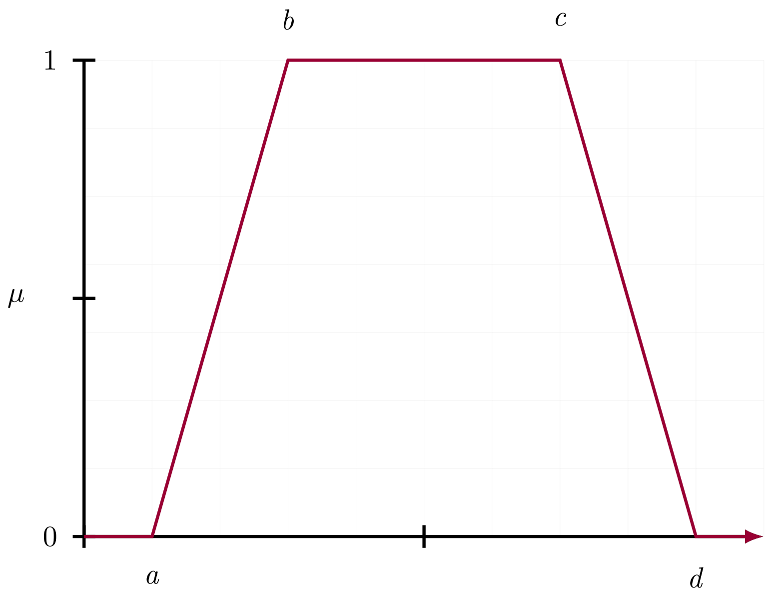


Figure 2.5: A Trapezoidal Membership Function

This thesis also puts forward a novel way of streamlining the creation of membership functions for the encapsulation of the returned degrees of significance, for each generated R-fuzzy set using the minimal number of parameters, as described in SECTION 3.3. This used a combination of both triangular and trapezoidal membership functions as the basis for encapsulating an entire abstract concept. The mathematical notation for the triangular membership can be defined as follows:

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ \dfrac{c-x}{c-b} & b \leq x \leq c \\ 0 & c \leq x \end{cases} \tag{2.2.3}$$

The mathematical notation for the trapezoidal membership can be defined in much the same way, with the addition of the interval region, and is given as follows:

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & b \leq x \leq c \\ \dfrac{d-x}{d-c} & c \leq x \leq d \\ 0 & d \leq x \end{cases} \tag{2.2.4}$$

The understanding of fuzzy sets and the fuzzy membership functions that derive them, is vital for the understanding of R-fuzzy sets. It is the uncertain, *possible* fuzzy membership values that are of concern, as there can be more than one which holds the same amount of sentiment indicative of a populous. Modelling a single perspective with no confliction is easily done as it is the only perspective to encapsulate. However, when more than one subjective perspective is being modelled, there can be a multitude of varying membership values that are agreed upon. The significance of these viable memberships need to be quantified, which in part provides the motivation of this thesis. Understanding the underlying significance of a membership value allows for one to better understand the perception of the populous. Fuzzy sets and the understanding of how a fuzzy membership value is obtained, is the main aspect of fuzzy theory that R-fuzzy is concerned with. This thesis will now go onto introduce the notion of rough theory, it is the rough approximations that allow for a rough set to be created, it is this rough set that provides the membership function of an R-fuzzy set.

2.3 Rough Theory

Rough set theory was first proposed by Pawlak (1982), in essence, the concept of a rough set deals with the synthesis of approximation with regards to the classificatory analysis of a data set, by framing a given concept, via the use of a lower and upper approximation. From the approximations, one is then able to construct a decision relative discernibility matrix, from which the discernibility functions can then be used to create a set of minimal reducts. These reducts describe the data in its most minimal form, without the loss of information, detail or sentiment. A conventional rough approach is used not only for data pruning, but also for the formulation of rules, to functionally map occurrences to the associated relative outcomes. These rules, can then be used to allow for the classification of novel, unseen data. Given the anticipation of contradiction, where a rule applies to cases which have the same attributable inputs, but different outcomes. These instances can be inferred from the *strength*, *coverage* and *certainty* values, to decide upon which rule is better suited.

The mathematical model used by rough theory was designed to be a new approach in handling imprecise, imperfect knowledge. The traditional mathematical remit of set theory stipulated that a set be clearly and exactly defined (Cantor, 1895). Philosophers with mathematical backgrounds, such as Bertrand Russell, were intrigued by the notion of a *vague* concept. A concept which allows for a set to partially include objects, or similarly, to allow for non-explicitly inclusive membership of objects, belonging to a certain degree. The classical ideology of set theory can lead to, in some instances contractions, as was described by Russell (1996)[1].

The standard rough model incorporates this notion of vagueness by using what is called the boundary set. If a boundary set is empty, it can be understood that the set being approximated is crisp. If the boundary set is non-empty, it can be implied that the set is *rough*. The boundary set contains objects that cannot be categorised as belonging unequivocally to any particular set. The theory assumes that there is a sufficient amount of knowledge known about the universe; the objects contained within the data set, so as to provide enough relative accuracy for categorisation. These objects are grouped together with other like objects that have the same conditional attribute values. If an object has the exact same properties as another object, the objects are indiscernible from one another, this gives rise to the indiscernibility and equivalence class relations. The property of indiscernibility plays heavily in the concept of rough theory. Indiscernibility and the equivalence relations allow for the reduction of the data longitudinally, via the rows of the data set.

Rough theory makes use of yet another data reduction technique, known as reducts. The reducts allow for the reduction in dimensionality, latitudinally, by reducing the number of conditional attributes. This is a means to discover the minimal number of conditional attributes that still hold true for the equivalence relation classes. The reduction in attributes correlates directly with a reduction of computational overhead, which is always beneficial regardless of the context of deployment. By removing superfluous objects and attributes, it allows for the approximation of the decision classes. Reducing the original data set such that the equivalence relation classes are maintained by the *new* reduced data set, will allow for greater effectiveness and overall efficiency (Lin and Cercone, 1996). Using

[1]The Principles of Mathematics republished via Norton

Table 2.1: An Information System

	Age	*LEMS*
x_1	16-30	50
x_2	16-30	0
x_3	31-45	1-25
x_4	31-45	1-25
x_5	46-60	26-49
x_6	16-30	26-49
x_7	46-60	26-49

the discernibility approach when finding the available subset of reducts, is acceptable for relatively small data sets. However, as the data set becomes larger, the discernibility approach in finding the minimal subset of attribute reducts becomes an *NP*-hard problem as stated by Komorowski et al. (1999). This in turn increases the associated computational burden with regards to the extra processing of information.

Assuming that $\Lambda = \left(U, A\right)$ is an information system, and that for any $B \subseteq A$ there is an associated equivalence relation of the form:

$$\text{IND}_{\Lambda}\left(B\right) = \left\{(x, x^{'}) \in U \times U \;\middle|\; \forall a \in B \quad a(x) = a(x^{'})\right\} \tag{2.3.5}$$

$\text{IND}_{\Lambda}\left(B\right)$ is called the B-indiscernibility relation. If it can be shown that $(x, x^{'}) \in \text{IND}_{\Lambda}\left(B\right)$, then the objects x and $x^{'}$ are indiscernible from one another with regards to the attributes from B. From the data contained in Table 2.1, which is with regards to the *Lower Extremity Motor Score* (*LEMS*) for varying age intervals, the indiscernibility relations for the conditional attributes, for both singular and combinatorial, would be of the following form:

$$\begin{aligned}
\text{IND}\left(\{\mathit{Age}\}\right) &= \left\{\{x_1, x_2, x_6\}, \{x_3, x_4\}, \{x_5, x_7\}\right\} \\
\text{IND}\left(\{\mathit{LEMS}\}\right) &= \left\{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7\}\right\} \\
\text{IND}\left(\{\mathit{Age},\ \mathit{LEMS}\}\right) &= \left\{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5, x_7\}, \{x_6\}\right\}
\end{aligned}$$

If one inspects the indiscernibility relation for *Age*, one can see that there are three partitions created on the universe. The set of objects $\{x_1, x_2, x_6\}$ creates one partition, the set of objects $\{x_3, x_4\}$ creates another partition, and the final set of objects, $\{x_5, x_7\}$ creates the third partition. Using both conditional attributes, *Age* and *LEMS* for the discernibility relation, ultimately creates five partitions of the universe. By simply inferring from any indiscernibility relation, one is able to ascertain the objects which satisfy its conditional belongingness.

2.3.1 Rough Approximations

Rough theory makes use of several approximations, of which provide the bounding utilised by an R-fuzzy approach. Assuming that $\Lambda = \Big(\mathbb{U}, A\Big)$ is an information system, and that $B \subseteq A$ and $X \subseteq \mathbb{U}$. Therefore B is a subset of the attributes and X is a subset of the universe, and also the set to be approximated. X can be approximated with only the information contained in B, by creating the lower and upper approximations of X.

Definition 2 (Approximations): Assume that $\Lambda = \Big(\mathbb{U}, A\Big)$ is an information system and that $B \subseteq A$ and $X \subseteq \mathbb{U}$. One can approximate set X with the information contained in B via a lower and upper approximation set.

The lower approximation is the set of all objects that absolutely belong to set X with respect to B. It is the union of all equivalence classes in $[x]_B$ which are contained within the target set X, and is given by the formal expression:

$$\underline{B}X = \Big\{x \;\Big|\; [x]_B \subseteq X\Big\} \tag{2.3.6}$$

$$\underline{B}(x) = \bigcup_{x \in \mathbb{U}} \Big\{B(x) : B(x) \subseteq X\Big\}$$

The upper approximation is the set of all objects which can be classified as being possible members of set X with respect to B. It is the union of all equivalence classes that have a non-empty intersection with the target set X, and is given by the formal expression:

$$\overline{B}X = \Big\{x \;\Big|\; [x]_B \cap X \neq \emptyset\Big\} \tag{2.3.7}$$

$$\overline{B}(x) = \bigcup_{x \in \mathbb{U}} \Big\{B(x) : B(x) \cap X \neq \emptyset\Big\}$$

Table 2.2: A Decision System

	Age	*LEMS*	*Walk*
x_1	16-30	50	*Yes*
x_2	16-30	0	*No*
x_3	31-45	1-25	*No*
x_4	31-45	1-25	*Yes*
x_5	46-60	26-49	*No*
x_6	16-30	26-49	*Yes*
x_7	46-60	26-49	*No*

The boundary region is the set that contains all objects that cannot be decisively categorised as belonging to X with respect to B. It is defined by the difference between the upper approximation Eq. (2.3.7) and the lower approximation Eq. (2.3.6), and is given by the formal expression:

$$BN(X) = \overline{B}(X) - \underline{B}(X) \tag{2.3.8}$$

Definition 3 (Rough Set): Assume that the pair, $apr = \left(\mathbb{U}, B\right)$ is an approximation space on $\mathbb{U}$ and assume that $\mathbb{U}/B$ denotes the set of all equivalence classes over B. The family of all definable sets in approximation space apr is denoted by $\mathrm{def}(apr)$. Given two subsets $\underline{A}$, $\overline{A} \in \mathrm{def}(apr)$ with $\underline{A} \subseteq \overline{A}$, the pair $\left(\underline{A}, \overline{A}\right)$ is called a rough set. If $x \in \underline{A}$ then $x \in \left(\underline{A}, \overline{A}\right)$. If $x \in \mathbb{U} - \overline{A}$ then $x \notin A$. If $x \in \overline{A}$ and $x \notin \underline{A}$ then x has an unknown relation to $\left(\underline{A}, \overline{A}\right)$. If the approximated set is crisp the boundary region of the set is empty, else if the approximated set is rough the boundary region of the set is non-empty.

To better understand the notion of rough set approximations, consider the data in Table 2.2, and notice the addition of the decision attribute *Walk*. A decision system, similar to that of an information system, is any information system of the following form $\Lambda = \left(\mathbb{U}, A \cup \{d\}\right)$, where $d \notin A$ is the decision attribute. If one was to approximate the decision attribute *Walk*, it can be clearly seen that it cannot be crisply defined in a classical sense. The objects x_3 and x_4 are contradictory, they have identical conditional values, but have conflicting decision outcomes. As the concept of *Walk* cannot be defined precisely, it is still possible to delineate the objects that definitely do belong to a particular classification, whether that is *Walk* = *Yes*, or *Walk* = *No*. The boundary set will include the objects that are inbetween the two approximations, in this case, that would include the objects x_3 and x_4. The very fact that the boundary set is non-empty, stipulates that the set is indeed *rough*.

Using the notation given for rough set approximations in DEFINITION 2, let $X = \{x \mid Walk(x) = Yes\}$. It can be seen that to satisfy this set, X would have to contain the objects x_1, x_4 and x_6. However, according to the established indiscernibility relations, this desired set cannot be created on the current known information. Using the conditional attributes of both *Age* and *LEMS*, the following approximations can be garnered:

Using Eq. (2.3.6) to obtain the lower approximation, one is presented with the set $\underline{B}(X) = \{x_1, x_6\}$. The objects $\{x_1, x_6\}$ can be understood as absolutely categorically belonging to the set *Walk* with absolute certainty. Using Eq. (2.3.7) to obtain the upper approximation, one is presented with the set $\overline{B}(X) = \{x_1, x_3, x_4, x_6\}$. Notice that the objects from the lower approximation are also included, this is due to the fact that the upper approximation includes all non-empty intersections of the target set. Therefore, it is implied that every object from a lower approximation, will also be included in the upper approximation. Also notice that even though object x_4 has a classification of *Yes*, and object x_3 has a classification of *No*, they are both contained within the upper approximation. This is due to the equivalence class relations. No matter which indiscernibility relation is used, the objects x_3 and x_4 are always members of the same partition, hence the inclusion of x_3, it cannot be distinguished from x_4 using only the conditional attributes. One cannot pick one object whilst ignoring the other, all objects of the set have to be included.

Using Eq. (2.3.8) to obtain the boundary region, one is presented with the set $BN(X) = \{x_3, x_4\}$. The boundary region will contain all objects that cannot be classified as belonging, or not belonging. With respect to the established equivalence relations, they can be described as being in *limbo*. The outside region, which includes all objects that do not belong to the set X is presented as $\mathbb{U} \Leftrightarrow \overline{B}(X) = \{x_2, x_5, x_7\}$. A graphical representation of these approximations can be seen in Figure 2.6, along with their associated contained objects.

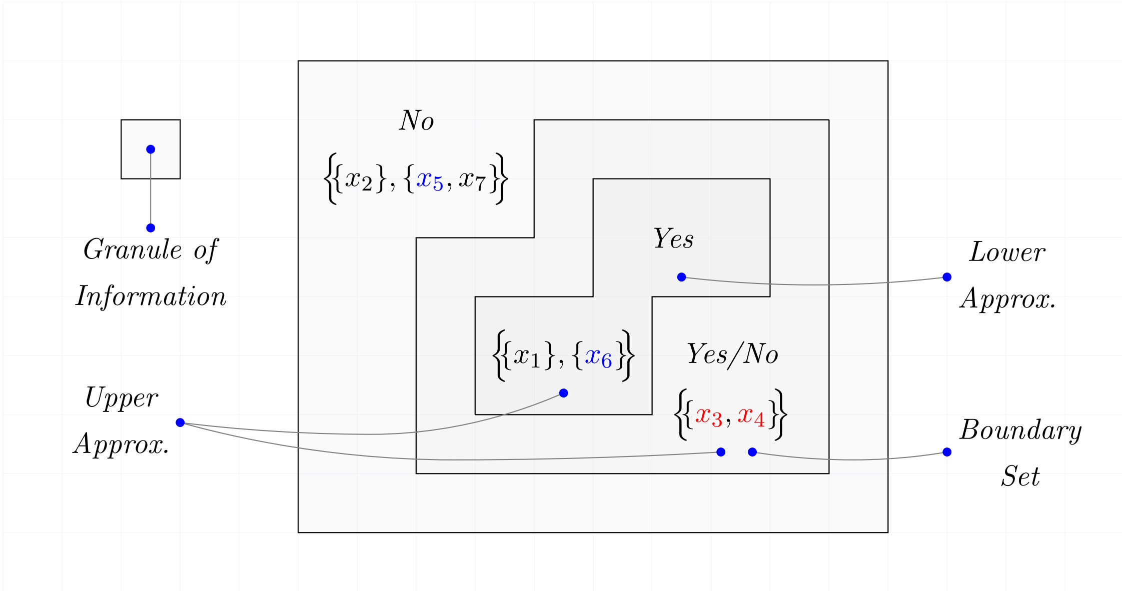


Figure 2.6: Rough Approximations for the Set *Walk*

2.4 Uncertainty Variations and Extensions

While researching the constituent components of an R-fuzzy set, other uncertainty models were also investigated, many of which were extensions or variations of the typical type-1 fuzzy set. Also described in this section is the notion of fuzzy-rough and rough-fuzzy models, whereby the use of equivalence and similarity relations are used. These alternative models are presented, concluding with a discussion as to why they were not ultimately chosen to be the basis of the new proposed framework.

2.4.1 An Interval-Valued Fuzzy Set

One such variation of a fuzzy set is the concept of an interval-valued fuzzy set. The membership of an individual object is characterised as belonging to an interval, rather than a single value as one would expect. An interval approach does indeed allow for the encapsulation of uncertain fuzzy membership values, by relatively simple means. However, as the approach adopts the use of intervals, information can be lost, it is not able to distinguish between values contained within its interval as it assumes unified distribution.

Definition 4 (Interval-Valued Fuzzy Set): Let $D[0,1]$ be the set of all closed sub-intervals from the interval $[0,1]$. Where $\mathbb{U}$ is the universe of discourse, and x is an object belonging to the universe $x \in X$. An interval-valued fuzzy set in $\mathbb{U}$ is given by set A, by the following expression:

$$A = \Big\{\Big\langle x, M_A(x)\Big\rangle \Bigm| x \in \mathbb{U}\Big\} \tag{2.4.9}$$

With $M_A : \mathbb{U} \rightarrow D[0,1]$

2.4.2 An Atanassov Intuitionistic Fuzzy Set

Similar to an interval-valued fuzzy approach, is that of an Atanassov intuitionistic fuzzy set, whereby, a degree of membership and non-membership are given. The sum of both memberships are generally less than or equal to 1. There have been extensions to this standard interpretation, Despi et al. (2013) allowed for the sum of memberships to be greater than 1, and for their differences to be either negative or positive. Much like that of interval-valued fuzzy sets, Atanassov intuitionistic fuzzy sets even with use of additional parameter values, will not be able to distinguish between values contained within its set.

Definition 5 (Atanassov Intuitionistic Fuzzy Set): Let $\mathbb{U}$ be a non-empty universe, where A in $\mathbb{U}$ is given by the following expression:

$$A = \Big\{\Big\langle x, \mu_A(x), v_A(x)\Big\rangle \Bigm| x \in \mathbb{U}\Big\} \tag{2.4.10}$$

Whereby μ_A stipulates the degree of membership of x with regards to A, and v_A is the degree of non-membership of x with regards to A.

$$\mu_A : X \rightarrow \Big[0,1\Big] \qquad v_A : X \rightarrow \Big[0,1\Big] \tag{2.4.11}$$

2.4.3 A Type-2 Fuzzy Set

A type-2 fuzzy set is a logical extension to that of type-1, whereby the addition of a secondary grade of membership is used. The secondary grade itself is a type-1 fuzzy mem-

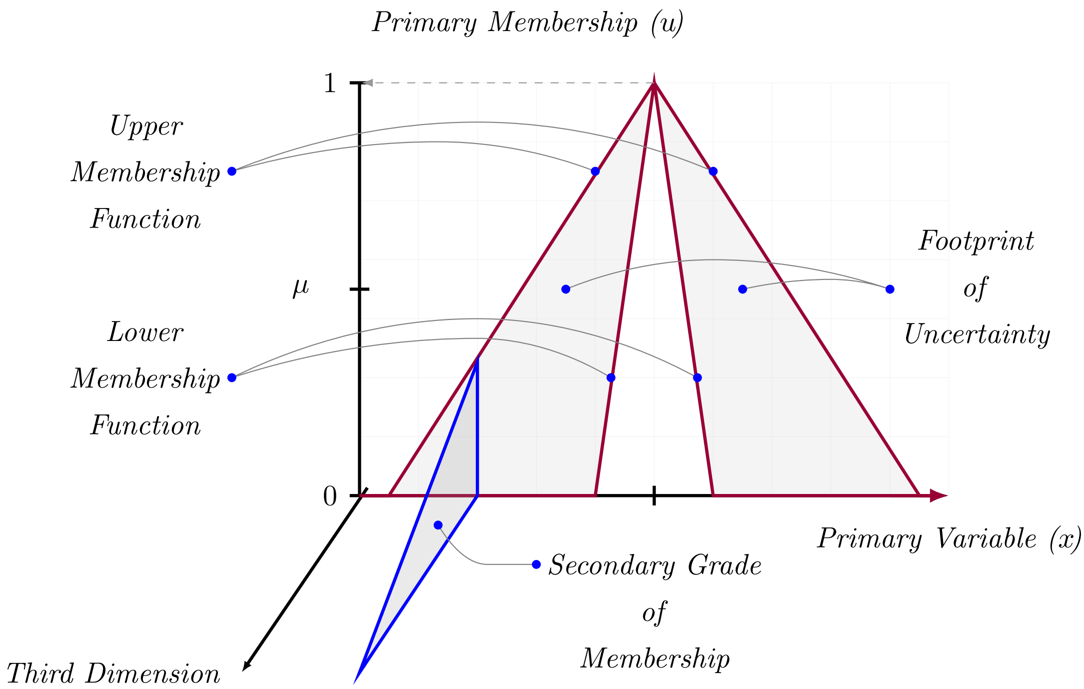


Figure 2.7: A Type-2 Fuzzy Set

bership, and provides a three-dimensional perspective, allowing for greater encapsulation of uncertainty. A visualisation of a type-2 fuzzy set is given in Figure 2.7.

DEFINITION 6 (Type-2 Fuzzy Set): A type-2 fuzzy set $\tilde{A}$ is characterised by a type-2 membership function $\mu_{\tilde{A}}(x,u)$, where $x \in \mathbb{U}$ and $u \in J_x \subseteq [0,1]$. A type-2 fuzzy set is given by the formal expression:

$$\tilde{A} = \left\{ \left\langle (x,u), \mu_{\tilde{A}}(x,u) \right\rangle \;\middle|\; \forall x \in \mathbb{U}, \forall u \in J_x \subseteq [0,1] \right\} \tag{2.4.12}$$

In which $\mu_{\tilde{A}} : \mathbb{U} \times J_x \rightarrow [0,1]$. $\tilde{A}$ can also be expressed as:

$$\tilde{A} = \int_{\forall x \in \mathbb{U}} \int_{\forall u \in J_x \subseteq [0,1]} \mu_{\tilde{A}}(x,u)/(x,u) \tag{2.4.13}$$

Where $\int \int$ denotes a union over all admissible x and u values. For discrete universes of discourse, $\int$ is replaced by that of $\sum$. The above definition is that of a generalised type-2 fuzzy set, as this is often too computationally exhaustive, an interval type-2 fuzzy approach may instead be utilised. In which case, $\mu_{\tilde{A}}(x,u)$ in Eq. (2.4.12) is replaced with $\mu_{\tilde{A}}(x,u) = 1$. This implies that every object to be contained within the secondary grade of membership is given an absolute degree of inclusion of 1. As no new information is contained in the secondary grade of membership, the *footprint-of-uncertainty* is instead used which is described by its upper and lower membership functions, thus reducing the computation needed to calculate each and every contained membership value of the secondary grade.

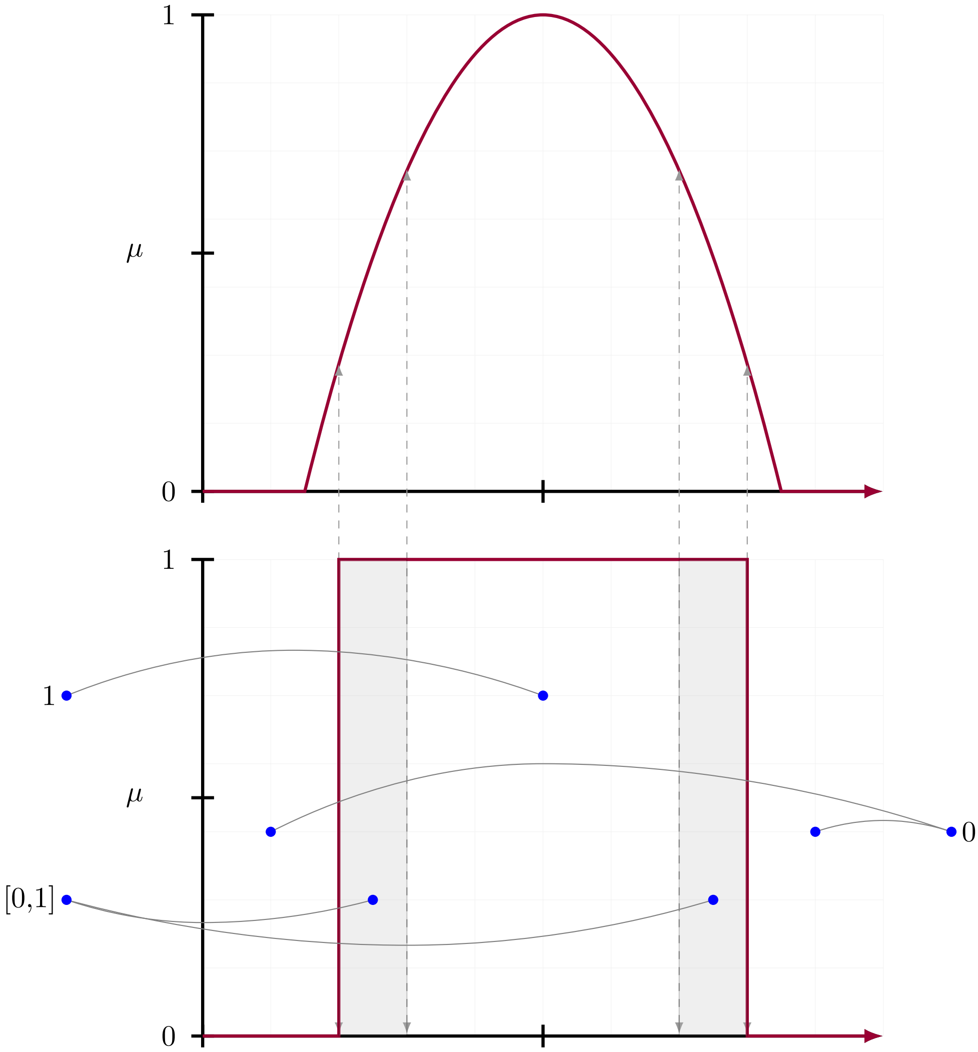


Figure 2.8: A Shadowed Representation of a Fuzzy Set

2.4.4 A Shadowed Set

A shadowed set is a special case fuzzy set, it employs the use of unit intervals to replicate areas of vagueness. As these regions are allocated to some regions of X, rather than the entire space, it contrasts with that of a type-1 fuzzy set. In much the same way an α-cut can create a crisp nesting of sets, a shadowed approach allows for an object to be associated to a clearly defined crisp area of the set. Therefore, a shadowed approach allows a type-1 fuzzy set to map an object x to 0, 1 or $[0, 1]$. A visualisation of a shadowed set is given in Figure 2.8.

DEFINITION 7 (Shadowed Set): Let $\mathbb{U}$ represent the universe and let A be a set in $\mathbb{U}$ $\big(A \subseteq \mathbb{U}\big)$. The fuzzy set A is a set of ordered pairs given by the following expression:

$$A = \Big\{\Big\langle x, \mu_A(x) \Big\rangle \,\Big|\, x \in \mathbb{U}\Big\} \tag{2.4.14}$$

The membership function of a shadowed set allows for an object to be declared as absolutely belonging to the set, absolutely not belonging to the set, or belonging to the set to some degree. The membership function is presented as follows:

$$\mu_A(x) : X \to \Big\{0, 1, [0, 1]\Big\} \tag{2.4.15}$$

The shadowed set was created as a means to not be constricted by precise numerical membership values. By allowing an object to be given a truth value of *Yes*, *No*, or belonging to the interval $[0, 1]$ which is indicative of the perceived uncertainty, allows for less computational overhead. However, once an object is given inclusion to the uncertainty interval, its significance is lost, an object cannot be distinguished once it is in the interval.

2.4.5 Fuzzy-Rough & Rough-Fuzzy

There have been several proposed hybridisations involving the use of fuzzy and rough components. The concepts contained in; Bodjanova (2007), Deng et al. (2007), Dubois (1980), Dubois and Prade (1990), Huynh and Nakamori (2005), Jensen and Shen (2008, 2009), Nanda and Majumdar (1992), Pawlak and Skowron (2007), Radzikowska and Kerre (2002), Sun et al. (2014), Wu et al. (2003), Xu et al. (2012), Zeng et al. (2015), mainly incorporate the use of equivalence and similarity relations. In much the same way that equivalence classes as given in Eq. (2.3.5), are vital for rough set theory, fuzzy-equivalence classes are vital to the concept of fuzzy-rough sets. A crisp equivalence class can be extended by the introduction of a similarity relation, which allows for the quantification of how similar two contained objects are to one another. For example, if S is a similarity relation on the universe and $\mu_S(x, y) = 0.1$, then it can be understood that the objects x and y are drastically different from one another. Equally, if the result was $\mu_S(x, y) = 0.8$, then both objects are indeed very similar. Expectedly, a value of 1 would indicate that they are identical. By using an equivalence relation on $\mathbb{U}$, one can incorporate the functionality of lower and upper approximations from rough set theory to create the concept of a rough-fuzzy set, extending the notion of classical set theory as stated by Dubois and Prade (2012, 1990). Alternatively, a fuzzy similarity relation can be used to replace the equivalence relation, the resulting concept is a deviation of rough set theory, referred to as fuzzy-rough sets. In a rough-fuzzy configuration, only the decision attribute values are fuzzy, whereas in a fuzzy-rough configuration, both conditional and decision values are allowed to be fuzzy. Using a fuzzy similarity relation, the fuzzy equivalence class $[x]_S$ for the objects that are classed as close to x, is defined as follows:

$$\mu_{[x]S}(y) = \mu_S(x, y) \tag{2.4.16}$$

2.4.6 A Fuzzy-Rough Set

The fuzzy B-lower and B-upper approximations as devised by Dubois and Prade (1990), are defined as follows:

$$\mu_{\underline{B}X}(F_i) = inf_x \max\Big\{1 - \mu_{F_i}(x), \mu_X(x)\Big\}\ \forall i \tag{2.4.17}$$

$$\mu_{\overline{B}X}(F_i) = sup_x \min\Big\{\mu_{F_i}(x), \mu_X(x)\Big\}\ \forall i \tag{2.4.18}$$

The tuple $\Big\langle \underline{B}X, \overline{B}X \Big\rangle$ can be used to represent the fuzzy-rough set.

2.4.7 A Rough-Fuzzy Set

A Rough-fuzzy set is another possible configuration, and can be seen as a generalisation of a rough set, derived from the approximation of a fuzzy set in a crisp approximation space. Under this configuration only the decision attribute values are fuzzy and the conditional values remain crisp. This is in contrast to the fuzzy-rough approach, where both conditional and decision attribute values may indeed be fuzzy. The lower and upper approximations represent the extent to which the objects belong to the set, and are defined as follows:

$$\mu_{\underline{R}X}\Big([x]_R\Big) = inf\Big\{\mu_X(x) \;\Big|\; x \in [x]_R\Big\} \tag{2.4.19}$$

$$\mu_{\overline{R}X}\Big([x]_R\Big) = sup\Big\{\mu_X(x) \;\Big|\; x \in [x]_R\Big\} \tag{2.4.20}$$

Where $\mu_X(x)$ is the degree of membership to which the object x belongs to the fuzzy equivalence class X, and each $[x]_R$ is a crisp equivalence class relation. The tuple $\Big\langle \underline{R}X, \overline{R}X \Big\rangle$ is called a rough-fuzzy set. This extends the classical rough set approach for both lower and upper approximations as given in Eq. (2.3.6) and Eq. (2.3.7), and can be seen if the $\mu_X(x)$ membership value returned is either 1 or 0. According to Dubois and Prade (2012, 1990), rough-fuzzy sets can be generalised to fuzzy-rough sets, where all the equivalence classes may be fuzzy. This will then allow for both the conditional and decision attribute values, to be represented as fuzzy, crisp, or a combination of the two.

The deployment of the two configurations is with regards to data sets that contain decision attributes, so that attribute reduction can be undertaken. An R-fuzzy approach is employed on perception based data, where each datum is a perception of a concept, two very different domains, and as such, different concepts.

2.4.8 An Overview

The problem of having to use excessive precision to describe increasingly imprecise phenomena has seen the creation of several variations, extensions and alternative approaches to try to overcome this dilemma. To some extent these alternatives such as; interval-valued fuzzy sets, Atanassov intuitionistic fuzzy sets, shadowed sets and type-2 fuzzy sets, do offer viable solutions but not wholeheartedly, as several questions still remain. The likes of interval-valued fuzzy sets, Atanassov intuitionistic fuzzy sets and shadowed sets, all allow for the means of encapsulating the uncertainty involved concerning the membership values of a fuzzy set. However, an interval-valued fuzzy set implies that the values contained within its interval are equally distributed. This is an unrealistic assumption for perception-based domains, as a specific subjective interpretation may not be shared with all concerned. As it will be seen, a unified distribution is not always the case, conflicting observations may exist. Other drawbacks may involve a value losing its inherent meaning if placed within an interval or shadowed region, as once placed in such a container, its uniqueness is dismissed, as it can no longer be distinguished from.

Perception based perspectives may not follow a universal interpretation, individuals may give varying results based on the same initial observations. These differences and similarities in their perceptions should all be taken into consideration. With this being the case, a single fuzzy membership value may not always be an ideal choice when representing a descriptive object, doing so would skew the underlying intent of the perceptions involved. The general consensus and the individual perceptions need to be taken into account. A type-2 fuzzy approach extends into the third dimension by using a type-1 fuzzy set to replace the use of crisp membership values. Nevertheless, the secondary membership function itself would still be using crisp values, as a result we are faced with the same initial problem. However, the higher levels of type-n one could implement, the closer one gets to precision and an agreed upon model, but not without consequence, the burden of complexity and computation becomes too costly.

The type-2 fuzzy set as presented in Definition 6 is the general interpretation, the simplified version is that of the interval type-2 fuzzy set. The use of the *footprint-of-uncertainty* reduces the computation to that of levels more forgiving. An interval type-2 fuzzy set makes use of interval mathematics, therefore making it easier to understand and compute compared to that of the generalised version. The review conducted by Melin and Castillo (2014), concisely inspected type-2 fuzzy logic applications in the areas mainly involving clustering, classification and pattern recognition, with the vast majority of them involving interval type-2 fuzzy logic.

There are indeed some interesting methods that exist for constructing interval-valued fuzzy sets. The work undertaken by Bustince Sola et al. (2015), shows that an interval-valued fuzzy set is a particular case of an interval type-2 fuzzy set. As such, both concepts should be treated differently from one another. The sentiment of Sola's paper is echoed by Mendel et al. (2016), reinforcing the perspective that they should indeed be treated differently.

As the membership of an R-fuzzy set itself is a set which contains discrete data, there is no loss of detail, unlike that of an interval valued fuzzy set approach (Yang and Hinde, 2010; Khuman et al., 2015a, 2016a). Once an object has entered the interval, there is no sense of how close to the bound of that interval it is; extremely pessimistic or overtly optimistic, the interval assumes generality and uniformed distribution. As the membership set of an R-fuzzy set is a rough set, the contents of which are crisply defined possible fuzzy membership values, that have an affinity to the descriptor it is being modelled for, no loss of information is suffered. Therefore, one can then quantify the distribution of that R-fuzzy set, ergo the proposed significance measure, which will be presented in Chapter 3.

An R-fuzzy approach offers a new perspective on uncertainty modelling, one which maintains all granules of information concerning its objects of uncertain fuzzy membership values. The substance of the following section was adapted from the research publications attributed to this thesis.

2.5 R-fuzzy Theory

The R-fuzzy concept is yet another proposal for encapsulating uncertainty, one which frames its fuzzy membership values via the approximations defined in Definition 2. The work undertaken by Yang and Hinde (2010) first put forward the notion of an R-fuzzy set, the capital 'R' distinguishes it from r-fuzzy, which was proposed by Li et al. (1996), yet another approach to encapsulate uncertainty. The membership value of an element of an R-fuzzy set is represented as a rough set. R-fuzzy sets are an extension of fuzzy set theory that allows for the uncertain fuzzy membership value to be encapsulated within the bounds of an upper and lower rough approximation. If using the voting method, the lower bound would contain the membership values agreed upon by all in the populous, whereas the upper bound would contain membership values agreed upon by at least one member.

Definition 8 (R-fuzzy Set): Let the pair $apr = \Big(J_x, B\Big)$ be an approximation space on a set of values $J_x = \{v_1, v_2, \ldots, v_n\} \subseteq [0,1]$, and let J_x/B denote the set of all equivalence classes of B. Let $\Big(\underline{M}_A(x), \overline{M}_A(x)\Big)$ be a rough set in apr. An R-fuzzy set A is characterised by a rough set as its membership function $\left(\underline{M}_A(x), \overline{M}_A(x)\right)$, where $x \in \mathbb{U}$, given by the formal expression:

$$A = \left\{\left\langle x, \left(\underline{M}_A(x), \overline{M}_A(x)\right)\right\rangle \,\middle|\, \forall x \in \mathbb{U}, \underline{M}_A(x) \subseteq \overline{M}_A(x) \subseteq J_x\right\} \tag{2.5.21}$$

2.5.1 R-fuzzy Approximations

Similar to type-2 fuzzy sets and interval-valued fuzzy, an R-fuzzy set describes its membership using a set itself, which are values that satisfy the membership descriptor. If a membership value has an affinity to the descriptor, then it is included within the R-fuzzy set. For each $x_i \in \mathbb{U}$, there is an associated membership description $d\left(x_i\right)$ which describes the belongingness of the object x_i to the set $A \subseteq \mathbb{U}$. Given that C is the set of available evaluation criteria, each value $v \in J_x$ is evaluated by $c_j \in C$ to determine if it is described by the membership description for x_i with respect to A. The result of the evaluation is given by a simple YES or NO. Understandably, evaluations which return a YES are accepted while evaluations which return a NO are ignored.

$$v \xrightarrow{\left(d(x_i), c_j\right)} \text{YES} \quad \text{or} \quad v \xrightarrow{\left(d(x_i), c_j\right)} \text{NO}$$

For each pair $\left((x_i), c_j\right)$ where $x_i \in \mathbb{U}$ and $c_j \in C$, a set $M_{cj}(x_i) \subseteq J_x$ is created, given by the formal expression:

$$M_{cj}(x_i) = \left\{v \,\middle|\, v \in J_x, v \xrightarrow{\left(d(x_i), c_j\right)} \text{YES}\right\} \tag{2.5.22}$$

The notation used for the approximations of R-fuzzy sets are semantically the same as those presented in Definition 2, however, they differ in their notation. The lower approximation of the rough set $M(x_i)$ for the membership function described by $d(x_i)$ is given by:

$$\underline{M}(x_i) = \bigcap_j M_{cj}(x_i) \tag{2.5.23}$$

The upper approximation of the rough set $M(x_i)$ for the membership function described by $d(x_i)$ is given by:

$$\overline{M}(x_i) = \bigcup_j M_{cj}(x_i) \tag{2.5.24}$$

Therefore the rough set approximating the membership $d(x_i)$ for x_i is given as:

$$M(x_i) = \left(\bigcap_j M_{cj}(x_i), \bigcup_j M_{cj}(x_i) \right) \tag{2.5.25}$$

For any given $d(x_i)$, it can be easily understood that the R-fuzzy set $M(x_i)$ depends wholly on J_x. For the same $d(x_i)$, assume $M_a(x_i)$ and $M_b(x_i)$ are two R-fuzzy memberships constructed for $A \subseteq \mathbb{U}$ from J_x^a and J_x^b using the same criteria set C. $M_a(x_i)$ and $M_b(x_i)$ may be different if $J_x^a \neq J_x^b$. If for the same criteria set C, $M_a(x_i) = M_b(x_i)$ holds if $J_x^a = J_x^b$. In addition to J_x, a change in the criteria set C will bring a change and result in a different $M(x_i)$. Considering, $v \xrightarrow{(d(x_i), c_j)}$ YES, it is perfectly possible for a different criteria set C to produce an entirely different $M_{cj}(x_i)$ and hence an entirely different $M(x_i)$. Which alludes to the fact that an R-fuzzy set $A \subseteq \mathbb{U}$ can only be created if J_x and the criteria set C are known.

An individual may have a unique perception which differs from another individual about the same observation. This different set of perceptions is an example of the criteria set C, where each individual will have their own criteria $c_j \in C$. An R-fuzzy set approach allows for a multitude of different perceptions of an observation to be encapsulated and collected. The membership of an object in an R-fuzzy set is defined as a rough set, hence its operational result is defined by a pair of definable sets for the rough approximation of its membership.

2.5.2 An R-fuzzy Example

Example 1: Assume that $F = \{f_1, f_2, \ldots, f_{10}\}$ is a set containing 10 flights, whose noise levels, given in decibel (dB) were recorded at a particular airport, and are given as N $= \{10, 20, 30, 50, 40, 70, 20, 80, 30, 60\}$. The collected decibel levels are presented in Table 2.3. Each noise N_i value corresponds to flight F_i, for example flight f_4 has a recorded value of 50(dB), whereas f_5 has a value of 40(dB), and so on. Assume that set $C = \{p_1, p_2, \ldots, p_6\}$ represents 6 individuals at the same location, all of whom gave their perceived perception of the noise levels for each of the 10 flights. These values have been collected and are presented in Table 2.4.

Table 2.3: Recorded Noise of Flights

#	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
dB	10	20	30	50	40	70	20	80	30	60

Table 2.4: The Collected Subjective Perceptions of Individuals

#	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
p_1	*NN*	*NN*	*NN*	*AC*	*AC*	*BN*	*NN*	*VN*	*NN*	*BN*
p_2	*NN*	*NN*	*AC*	*AC*	*AC*	*BN*	*NN*	*VN*	*AC*	*BN*
p_3	*NN*	*AC*	*AC*	*BN*	*AC*	*VN*	*AC*	*VN*	*AC*	*BN*
p_4	*NN*	*NN*	*NN*	*AC*	*AC*	*BN*	*NN*	*VN*	*NN*	*BN*
p_5	*NN*	*AC*	*AC*	*AC*	*AC*	*BN*	*AC*	*VN*	*AC*	*BN*
p_6	*NN*	*NN*	*AC*	*AC*	*AC*	*VN*	*NN*	*VN*	*AC*	*BN*

The abbreviated terms contained within the table can be understood as meaning:

$NN \rightarrow$ Not Noisy $\quad AC \rightarrow$ Acceptable $\quad BN \rightarrow$ A Bit Noisy $\quad VN \rightarrow$ Very Noisy

To construct a fuzzy set, one takes the values contained in N, and then inserts them into a simple linear function given as follows:

$$\mu(f_i) = \frac{l_i - l_{\min}}{l_{\max} - l_{\min}} \tag{2.5.26}$$

Where l_i is reference to the noise level of flight f_i and $l_{\max}$ and $l_{\min}$ provide the normalising scope and are the maximum and minimum values contained in N. After completion, one is presented with a fuzzy set containing precise fuzzy membership values for each of the 10 flights:

$$\mu(f_1) = 0.00 \quad \mu(f_2) = 0.14 \quad \mu(f_3) = 0.29 \quad \mu(f_4) = 0.57 \quad \mu(f_5) = 0.43$$
$$\mu(f_6) = 0.86 \quad \mu(f_7) = 0.14 \quad \mu(f_8) = 1.00 \quad \mu(f_9) = 0.29 \quad \mu(f_{10}) = 0.71$$

It is not always possible to know the exact noise level of a particular flight, nor do people need to know the exact levels in their communication (Yang and Hinde, 2010). An R-fuzzy approach provides an answer to the question, how to express a fuzzy membership function if the exact noise level is unknown or not given? If for example, we know that flight f_{11} has an associated abstract description of being AC, how can this be encapsulated using fuzzy membership values? A standard fuzzy type-1 approach would assign it a precise value, but this does not fully appreciate the differences in perception individuals may have for a given flight, as can be seen from Table 2.4. A particular flight can be perceived in a multitude of ways; what is acceptable to some may not be acceptable to all. Rather than neglecting to include a particular response as it may go against the grain of common held interpretations, it would be more ideal to encapsulate all involved perceptions.

Once the fuzzy membership values have been established, based on the recorded noise levels and the linear function given in Eq. (2.5.26), the fuzzy membership set derived is given as follows:

$$J_x = \Big\{0.00, 0.14, 0.29, 0.57, 0.43, 0.86, 0.14, 1.00, 0.29, 0.71\Big\}$$

If it is known that f_{11} is AC, the descriptor can be set to $d(f_{11}) = AC$, where its membership has to satisfy this description. The evaluation criterion is decided by each individual from set C. Each value $v \in J_x$ is evaluated against $p_j \in C$ to conclude if it fits with the description given for $d(f_{11})$ and for $f_{11} \in \mathbb{U}$, using:

$$v \xrightarrow{\big(d(x_i), c_j\big)} \text{YES}$$

For each $p_i \in C$ there is a corresponding row in Table 2.4, for the columns where there is a match with the descriptor given for $d(f_{11})$, its corresponding flights will provide the membership values in accordance to their noise levels. Using the description of AC for flight f_{11} and the values provided by the individuals in set C, a subset of values can be constructed, $M_{pj}(f_{11}) \subseteq J_x$. For example, inspecting p_1, one can see that flights f_4 and f_5 are the only flights that satisfy the descriptor $d(f_{11}) = AC$. As a result, we take the corresponding membership values from J_x for f_4 and f_5, which gives a subset of values, $M_{p1}(f_{11}) = \{0.57, 0.43\}$. For p_2, the descriptor is satisfied by flights f_3, f_4, f_5 and f_9. This results in subset $M_{p2}(f_{11}) = \{0.29, 0.57, 0.43\}$, where instances of duplication are ignored. This process is repeated for all objects of set C, the results of which are given as follows:

$$M_{p1}(f_{11}) = \{0.57, 0.43\} \quad M_{p2}(f_{11}) = \{0.29, 0.57, 0.43\} \quad M_{p3}(f_{11}) = \{0.14, 0.29, 0.43\}$$
$$M_{p4}(f_{11}) = \{0.57, 0.43\} \quad M_{p5}(f_{11}) = \{0.14, 0.29, 0.57, 0.43\} \quad M_{p6}(f_{11}) = \{0.29, 0.57, 0.43\}$$

Once the subsets have been computed and collected $M_{pj}(f_{11})$, one can now apply the notion of approximations. Starting with the lower approximation and using Eq. (2.5.23), one can inspect each subset to find any membership value that occurs in each and every subset. As a result $\{0.43\}$ is the only membership value that satisfies this requirement, therefore it is the only value to be contained in the lower approximation of the rough set. If no such membership value existed, whereby it was not included in all generated subsets, then the resulting rough set would contain an empty lower approximation. The upper approximation Eq. (2.5.24), contains values that have been considered to be valid with relation to the descriptor. This essentially means that all instances contained within the subsets are placed into the upper approximation, where the duplications are removed $\{0.14, 0.29, 0.43, 0.57\}$. One will notice that the lower approximation value of $\{0.43\}$ is also contained, this is understandable as Eq. (2.5.21) clearly states that $\underline{M}_A(x) \subseteq \overline{M}_A(x)$. Any value contained in the lower approximation will also be contained within the upper approximation. The actual rough set approximating the uncertain membership for $d(f_{11})$ is constructed using Eq. (2.5.25), therefore we are presented with:

$$M(f_{11}) = \Big(\{0.43\}, \{0.14, 0.29, 0.43, 0.57\}\Big)$$

This result alludes to the fact that the membership value 0.43 is agreed upon by all and that its corresponding flights are *AC*, as it is the only value to be contained within the lower approximation. Also, the flights associated with 0.14, 0.29 and 0.57 are also considered as *AC* by some, but absolutely not all.

If one was to use a traditional type-1 fuzzy approach to define an appropriate membership for f_{11} when described as *AC*, the average may be taken to represent it, in which case the returned value would be:

$$\mu(f_{11}) = \frac{1}{17} \sum_{x \in \mathbb{U}} \mu(x)/x = 0.40$$

This is the summation of each membership value contained in all the generated subsets $M_{pj}(f_{11}) \in J_x$. Inspecting the value 0.40, it is slightly less than the accepted and more reasonable 0.43, and considerably less than the 0.57, which was agreed upon by 5 of the 6 individuals in the criteria set.

If one was to apply the notion of interval-valued fuzzy sets, where the scope would be the most pessimistic lower bound and optimistic upper bound [0.14,0.57], it is not possible to tell which values were agreed upon unanimously. Simply inspecting the values themselves, one can see that a significantly large area of distribution is given. The value of 0.14 was only agreed upon by 2 of the 6 individuals, but as an interval approach does not concern itself with the *significance* of captured values, it will treat any value as a viable candidate with equal weighting. This same problem is shared with Atanassov intuitionistic fuzzy sets, a value once contained in the interval loses its uniqueness. This harkens back to the initial problem of there being no distinction between the values contained within interval regions or shadow areas, the value itself loses its individuality. With regards to type-2 fuzzy sets the example is too small, such that a reliable membership distribution cannot be created, which gives additional credence for the concept of an R-fuzzy and significance measure pairing. If the shadowed set was to be used, the values for f_{11} would be placed in the *unknown* region. One can see that in this instance, an R-fuzzy set is an ideal concept to use, maintaining uniqueness for all viable membership values.

R-fuzzy sets allow for a greater breadth of uncertainty encapsulation, as compared to the other alternative uncertainty models, the R-fuzzy approach is the most well equipped. This thesis will now go onto to describe grey theory, from which the analysis component is utilised for the new proposed framework.

2.6 Grey Theory

Grey theory is yet another approach for handling uncertainty, which was first initiated in the 1980s by Ju-Long (1982). The paradigm places particular emphasis on domains associated with small samples and poor information, where the information may be partially known and partially unknown, a common trait of uncertain systems (Liu et al., 2012). The term *grey* can be interpreted as a halfway house between a *black* system and a *white* system. Where in a black system absolutely nothing is known, whereas in a white system absolutely everything is known. The purpose of which is to garner an informed and accurate conclusion based on what little, uncertain information is available. This is generally achieved through the processes of generating, excavating and extracting meaningful content. In doing so, the system's operational behaviours and its laws governing its evolution can be accurately described and acutely monitored (Liu and Forrest, 2010).

While undertaking the literature review, it was advised that one should consider investigating the concept of grey theory to see if it was viable for the incorporation into the overall research. The initial stages of the investigation lead to the publication of an improved method for grey model forecasting, presented in Khuman et al. (2013). Although not primarily concerned with the main research output of this thesis, the domain of grey forecasting provided the insight into the various facets of the paradigm, one of which was grey analysis and by proxy, the absolute degree of grey incidence. As progression through the literature developed, so too did the understanding of the concept. It became apparent that there was a relationship to that of probability theory, in particular, the theory of imprecise probability (Walley, 2000), which has more of an association to grey theory than traditional precise probability.

Imprecise probability can be seen as a generalisation of probability theory, where its main application is with regards to vague, partially known systems, two major characteristic traits shared by grey theory. With this being the case, identifying a unique probability distribution will often be difficult. The main ethos of imprecise probability is to take what little available, partial information one has, and make efforts to represent it more accurately, which is very closely aligned to the mantra of grey. The use of interval representation to capture the imprecise probability, and the use of grey numbers to represent an interval of unknown content, are very similar. However, the lower and upper bounds of a grey interval are replaced with lower and upper probabilities, or more generally, lower and upper expectations (Weichselberger, 2000; Williams, 2007). If the lower probability is equal to the upper probability then precise probability is known, therefore a traditional perspective can be adhered to, which differs from a grey perspective, as one may not know all about the information involved. However, if more information is presented to a grey system, whereby the knowledge of the information reduces the scope of the bounding of the interval, steps are made towards a more probabilistic interpretation. Only when the scope of the bound and the value of all possible outcomes are known does one have the equivalence of a true probabilistic precise interpretation. However, in this instance the *grey* number would be regarded as a white number, as there are no unknown granules of information.

Continuing with the notion of probability there is also the theory of possibility Zadeh (1999b); Dubois and Prade (2012), an alternative mathematical theory to probability.

Table 2.5: A Digestive Summary Between Grey, Fuzzy & Prob. Statistics

Criteria	*Grey*	*Fuzzy*	*Prob. Statistics*
Research Objects	Poor Information	Cognitive Uncertainty	Stochastic
Basic Sets	Grey Sets	Fuzzy Sets	Cantor Sets
Methods	Information Coverage	Function of Affiliation	Probability Distribution
Procedures	Sequence Operator	Cut Set	Frequency Distribution
Data Requirement	Any Distribution	Known Membership	Typical Distribution
Emphasis	Intention	Extension	Intention
Objective	Laws of Reality	Cognitive Expression	Historical Laws
Characteristics	Small Sample	Experience	Large Sample

Possibility theory was first put forward by Zadeh in 1978 as an extension to his work on fuzzy, which distinguished itself from probability. Enhancements to the theory by Dubois and Prade have evolved it further still. Whereas probability theory uses a single value to describe how likely an event is to occur, possibility theory makes use of two measures; the possibility and the necessity of an event. The subsets of the universe are measurable, and the resulting distribution of possibility is a function. This allows for the beliefs of an event to be described by a number and a bit.

Dubois and Prade (2012); Hllermeier (2014), have shown fuzzy sets to be possibility distributions under certain considerations, that present the degree of plausibility for each object in the set. A grey set could also be described using possibility distributions, if that grey set was interpreted as a white set. The theory of evidence, also known as the theory of belief functions and the Dempster-Shafer theory, proposed by Shafer et al. (1976); Shafer (1982, 1990), is another framework for dealing with and understanding uncertainty. There already exists logical connections between the theory of evidence to that of probability, possibility and imprecise probability theory, therefore so too is there a relationship with grey theory. The theory allows for the combination of evidence from different sources to arrive at a degree of belief, given by the belief function. As it is a way of representing epistemic plausibility (Couso and Dubois, 2014; Couso et al., 2014), it too can be said that fuzzy has strong links to probability theory. One can describe the probability of a fuzzy event, and equally, the fuzziness of a probability value. Grey and fuzzy are foundationally different from one another, a combination of certain aspects from each theory will allow for a greater amount of uncertainty to be captured. There exists a relationship between grey and possibility theory, only under certain considerations. For example when a grey set is interpreted as a white set described using white numbers. A digestive summary between grey, fuzzy and prob. statistic theory can be seen in Table 2.5.

The perspective that fuzzy adopts contrasts with that of a grey perspective, therefore the hybridisation between the two allows for greater swathes of uncertainty to be captured. A fuzzy perspective is with regards to a clear intention with an unclear extension. A grey perspective is a polar opposite, an unclear intention with a clear extension. The progressive review of grey systems lead to the publication of a divergence paper Khuman et al. (2014). This highlighted some of the intrinsic differences that exist between fuzzy and grey. The publication was mainly concerned itself with the foundational elements such as numbers and sets, and as such, this thesis will now describe those core elements of grey theory to better provide an more in-depth understanding.

2.6.1 Grey Numbers

Much like that of fuzzy set theory, grey theory also makes use of sets and numbers. A grey set makes use of grey numbers $g^{\pm}$, and considers the characteristic function values of a grey set as grey numbers. There are several classes of grey numbers that the reader should be made aware of: The lower limit grey number $g^{\pm} \in [g^{-}, \infty)$. The upper limit

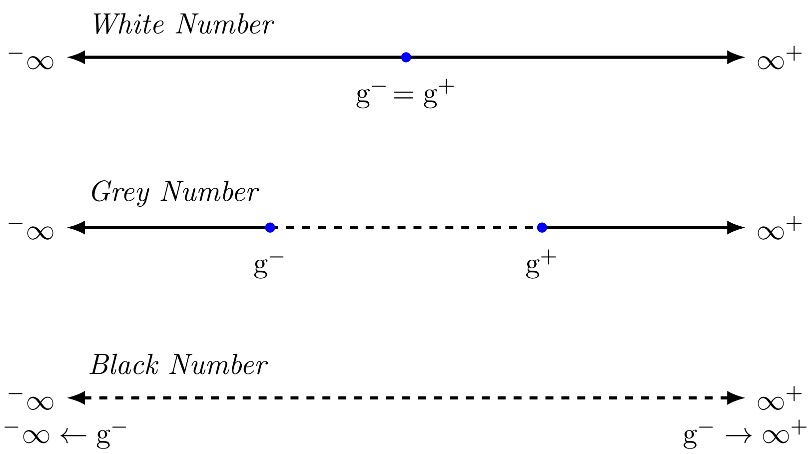


Figure 2.9: A White, Grey & Black Number

grey number $g^{\pm} \in \left({}^{-}\infty, g^{+}\right]$. An interval grey number $g^{\pm} \in \left[g^{-}, g^{+}\right]$. A black number $g^{\pm} \in \left({}^{-}\infty, \infty^{+}\right)$. Finally, a white number: $g^{\pm} \in \left[g^{-}, g^{+}\right] \bigwedge g^{-} = g^{+}$. With the aforementioned classes one can interpret what is meant by white, black and grey. A white number is absolutely known and has an associated exact value. A black number is absolutely unknown, both in its exactness and the range in which it is to be found. A grey number is a halfway house between the two, where an exact value is unknown but the bounds of its range in which it is to be found is known. A visualisation of the various grey number interpretations is presented in Figure 2.9.

Definition 9 (Grey Number): A grey number is a number with an unknown value, but contained within a known bound. The formal expression is given as follows:

$$g^{\pm} \in \left[g^{-}, g^{+}\right] = \left\{g^{-} \leq t \leq g^{+}\right\} \tag{2.6.27}$$

Where $g^{\pm}$ is the grey number, t is the information, and g^{-} and g^{+} are the lower and upper limits of the bound, respectively.

It's noteworthy to mention that grey numbers could be discrete if the candidate values are finite. For example, if a number can only be one value among the values it contains; $2, 4, 6, 8, 10$, it would be given as a set: $g^{\pm} \in \{2, 4, 6, 8, 10\}$. This has a close relationship to the epistemic interpretation of a set, both from the conjunctive and disjunctive perspective (Couso and Dubois, 2014). It can be interpreted from the conjunctive perspective as a piece of precise information, such as the range of ages for a collection of children. But it can be also interpreted from the disjunctive viewpoint as a collection of possible values for a particular number x, according to our incomplete information about it; we know that x belongs to the set, and nothing else. A candidate value inside the conjunctive set is a value that could be the correct value, any value not contained in the set is unequivocally incorrect. As a result its characteristic function could be interpreted as a possibility distribution function, as described by Zadeh (1999b).

Clearly defined membership or characteristic function values are needed when using fuzzy sets (Deschrijver and Kerre, 2003). When information may be partial or incomplete, finding a clearly defined number can prove difficult. How to define a crisp membership value becomes a problem. As was remarked by Yang and Hinde (2010), for epistemic uncertainty, an interval representation infers that any value within the interval may be a possible value. However, one may know that the possible value can only be one of a finite number of values within the interval. As a result, an interval representation is not useful. An extension to the standard grey number is that of the generalised grey number presented in Yang (2007); Yang and John (2012a,b); Yang et al. (2014). Where the allowance of both discrete and continuous values are perfectly acceptable. The notation that follows uses closed brackets, but it should be understood that both closed and open intervals are perfectly acceptable.

Definition 10 (Generalised Grey Number): Let $\mathrm{g}^{\pm} \in \mathbb{R}$ be an unknown real number within a union set of closed or open intervals, presented as follows:

$$\mathrm{g}^{\pm} \in \bigcup_{i=1}^{n} \left[a_i^-, a_i^+\right] \tag{2.6.28}$$

Where $i = 1, 2, \ldots, n$ and n is an integer value and $0 < n < \infty$, $a_i^-, a_i^+ \in \mathbb{R}$ and $a_{i-1}^+ \leq a_i^- \leq a_i^+ \leq a_{i+1}^-$. For any interval $\left[a_i^-, a_i^+\right] \subseteq \bigcup_{i=1}^{n} \left[a_i^-, a_i^+\right]$, p_i is the probability for $\mathrm{g}^{\pm} \in \left[a_i^-, a_i^+\right]$. If the following two conditions hold true: $p_i > 0$ and $\sum_i^n p_i = 1$. If so, $\mathrm{g}^{\pm}$ is a generalised grey number, $\mathrm{g}^- = \inf_{a_i^- \in \mathrm{g}^{\pm}} a_i^-$ and $\mathrm{g}^+ = \sup_{a_i^+ \in \mathrm{g}^{\pm}} a_i^+$ are respectively referred to as the lower and upper limits of the grey number $\mathrm{g}^{\pm}$. Based on this interpretation, it is impossible for there to be more than one number that is the underlying white number, contained within its candidate set.

The well-known notion of basic mass assignment from the theory of evidence Shafer et al. (1976), has similarities to the generalised grey number. As a basic mass assignment is a mapping m defined on the power set of a universe $\mathbb{U}$, assigning a non-negative mass to each set. The collection of focal sets is the collection of those sets with strictly positive masses, $\{A_1, \ldots, A_n\}$, such that $m(A_i) = p_i > 0$. The sum of the masses associated to the focal sets is an absolute 1. The notion of the generalised grey number is formally the same as the notion of the basic mass assignment. In particular, under the epistemic perspective, the generalisation of the conjunctive view of sets, fuzzy sets can be understood as a particular mass assignments where the focal sets are nested. In this case, the associated plausibility measure is, in particular, a possibility measure.

The generalised grey number is particularly important as it allows for the inclusion of both discrete and continuous data to be contained in the candidate set. The very fact that both open and closed intervals may be included means that gaps, or *hesitations* may also be included, which would otherwise not be the case if using a standard interval representation. As grey systems are associated with partially known, incomplete information, grey numbers are associated to partially known numbers. Therefore, intervals can be considered a special case of grey numbers, where the bounds of the number are known, but not its exact position within the bound. However, the candidate of a partial known number could be selected from a finite set of numbers, or a set of intervals, as stated by Yang and John (2012a).

The epistemic disjunctive interpretation is more vivid when considering the generalised grey number, as there can only be one underlying white, correct, value from the candidate set. As grey systems are closely related to incomplete and partially known systems, so too is the perspective of epistemic disjointness (Couso and Dubois, 2014). However, with regards to the interval grey number, a conjunctive perspective may still be prevalent, but as there can only be one underlying white number, a disjoint interpretation is ultimately the final perspective when describing grey numbers.

Much like a fuzzy membership function value, a grey number $\mathrm{g}^{\pm}$ can also be attributed to a membership like value. In this instance the value is referred to as the degree of *greyness*. The degree of greyness depends only on the two limits of the grey number and has nothing to do with the cardinality of its candidate set. The degree of greyness of a grey number measures the significance of uncertainty in a grey number.

Definition 11 (Degree of Greyness): Quantifies the significance of the unknown grey number to the interval containing the information. The formal expression is given as:

$$\mathrm{g}^{\circ}(\mathrm{g}^{\pm}) = f(\mathrm{g}^{-}, \mathrm{g}^{+}) \tag{2.6.29}$$

Where f is a function to determine the significance of the grey interval to the grey number $\mathrm{g}^{\pm}$. Let $D = [d_{min}, d_{max}]$ be the domain of values represented by a grey number $\mathrm{g}^{\pm} \in [\mathrm{g}^{-}, \mathrm{g}^{+}]$. We then have $d_{min} \leq \mathrm{g}^{-}$, $\mathrm{g}^{+} \leq d_{max}$. The degree of greyness can then be obtained using the following expression:

$$f(\mathrm{g}^{-}, \mathrm{g}^{+}) = \frac{|\mathrm{g}^{+} - \mathrm{g}^{-}|}{|d_{\max} - d_{\min}|} \tag{2.6.30}$$

As this is a measure for greyness of a grey number $g^{\pm}$, the value returned can be understood as meaning either a white number: $g^{\circ} = 0 \Leftrightarrow g^{-} = g^{+}$. A grey number : $g^{\circ} \in \left(0,1\right) \Leftrightarrow \left\{g^{-} \leq g^{\pm} \leq g^{+}\right\}$. Or a black number : $g^{\circ} = 1 \Leftrightarrow g^{-} = d_{\min} \bigwedge g^{+} = d_{\max}$.

2.6.2 Grey Sets

Similar to that of a fuzzy set, there is also the notion of sets from a grey perspective. Understandably, if the degree of greyness is 0 then no uncertainty exists, meaning that the value is white, absolutely known and crisp. If the degree of greyness is 1, absolute uncertainty exists and therefore deemed a black number. Any real value returned that falls in the range $[0,1]$ is a grey number. In the same way a fuzzy set is an extension to the idea of a fuzzy number, a grey set is an extension to the idea of a grey number. The notion of grey sets themselves could also be described as either being; white, black or grey.

Definition 12 (White Set): For a set $A \subseteq \mathbb{U}$, if the characteristic function value for all objects x_i with respect to A can be expressed as crisp white numbers, belonging to $v \in [0,1]$. Where $\chi_A : \mathbb{U} \to [0,1]$. If so, then set A can be assumed to be a white set.

Definition 13 (Black Set): For a set $A \subseteq \mathbb{U}$, if the characteristic function value for all objects x_i with respect to A can only be expressed as black numbers, then set A can be assumed to be a black set.

Definition 14 (Grey Set): For a set $A \subseteq \mathbb{U}$, if the characteristic function value for all objects x_i with respect to A can be expressed by grey numbers, $g_A^{\pm} \in \bigcup_{i=1}^{n} \left[a_1^{-}, a_1^{+}\right] \in D\left[0,1\right]^{\pm}$. Where $D\left[0,1\right]^{\pm}$ is the set of all grey numbers within the interval $[0,1]$, and $\chi_A : \mathbb{U} \to [0,1]^{\pm}$. If so, then set A can be assumed to be a grey set.

Similar to the expression of a fuzzy set as defined in Definition 1, the objects constituting the set are presented with their associated grey numbers, as ordered pairs:

$$A = \left\{\left\langle x, g_A^{\pm}(x)\right\rangle \,\middle|\, x \in \mathbb{U}\right\}$$

The characteristic function in this instance is a general expression, one which does not exclude any relevant criteria in defining a set. Therefore, it can be replaced by a probability function, membership function, possibility function and so on (Yang and John, 2012b).

The degree of greyness for an object belonging to a set can also be calculated, much like Definition 11, which presented the degree of greyness of the grey number with relation to its interval. The greyness measure for an object in relation to its set is presented as follows:

Definition 15 (Degree of Greyness for an Object): Assume that $\mathbb{U}$ is a finite universe of discourse and $x \in \mathbb{U}$. For a set $A \subseteq \mathbb{U}$, the characteristic function value of object x with regards to A is $\mathrm{g}^{\circ}_{A}(x) \in D[0,1]^{\pm}$. The degree of greyness for an object is given by the following expression:

$$\mathrm{g}^{\circ}_{A}(x) = \left| \mathrm{g}^{+} - \mathrm{g}^{-} \right| \tag{2.6.31}$$

Based on the degree of greyness for an object, the degree of greyness for a set can also be computed, which is presented as follows:

Definition 16 (Degree of Greyness for a Set): Assume $\mathbb{U}$ is a finite universe of discourse, A is a grey set such that $A \subseteq \mathbb{U}$. Each x_i object is with regards to the grey set A, $x_i \in \mathbb{U} = i = 1, 2, \ldots, n$, where n represents the cardinality of $\mathbb{U}$. The degree of greyness of a set is given by the following expression:

$$\mathrm{g}^{\circ}_{A} = \frac{\sum_{i=1}^{n} \mathrm{g}^{\circ}_{A}(x_i)}{n} \tag{2.6.32}$$

2.6.3 Grey Analysis

Grey theory has been widely applied to analysis, modelling, forecasting, decision making and control. New hybrid approaches have since been developed such as; grey hydrology, grey geology and grey regional economics (Liu et al., 2012). The first furore for grey systems was mainly associated with agriculture, ecology, economy, environmental sciences, financial, management, material sciences, and the like. The introduction of the weakening and strengthening operators for grey sequencing were introduced relatively early on in the creation of the theory. These were created to alleviate the uncontrollable shockwaves of interference that the data may contain, which would skew the raw intent and the underlying meaning. In addition, the creation of the generalised degrees of grey incidence were put forward; the absolute degree, the relative degree and the synthetic degree. Each one taking a particular perspective when analysing the correlation and similarity of the geometrical patterns of the sequence curves. Therefore, allowing for a better and more detailed quantification between characteristic sequences and possible behavioural impact factors. Allowing one to ascertain which factors, triggered *what* changes in the characteristic sequence of the system. Understanding which behavioural factors, and how much of a significance such behavioural factors have in shaping the geometric curve of the characteristic sequence, allows one to better understand the intricacy of the overall model. It is this absolute degree of grey incidence that this thesis adopts for the analysis of the results returned from the R-fuzzy set and significance measure pairing. The use of a traditional grey analysis approach with an un-traditional application, as it will be shown, has provided for additional functionality and versatility with regards the R-fuzzy framework.

The fundamental aspects of grey theory involves the use of grey numbers and grey sets, and as such provides one the basis from which to extend and implement. Apart of the repertoire of grey theory is the concept of grey relational analysis. The traditional degree of grey incidence provides the basis for all variances of the degree of incidence $\Gamma = [\gamma_{ij}]$, where each entry in the i^{th} row of the matrix is the degree of grey incidence for the corresponding characteristic sequence Y_i, and relevant behavioural factors $X_1, X_2, \ldots, X_m$. Each entry for the j^{th} column is reference to the degrees of grey incidence for the characteristic sequences $Y_1, Y_2, \ldots, Y_n$ and behavioural factors X_m. The notion of the absolute degree of grey incidence is one such variation.

It is the absolute degree of grey incidence that this thesis adopts for the analysis of the results, when using an R-fuzzy and significance measure pairing. Grey relational analysis in general has no special requirements regarding data distribution, and has relatively simple associated computation. There have been several proposed novel variations of the degree of grey incidence (Liu et al., 2011b,a). Although this thesis adopts the traditional notion of grey incidence, it is employed in an un-traditional way, as will be seen in Chapter 3. Before the decision was made to adopt the absolute degree of grey incidence into the R-fuzzy framework, the concept was first implemented with regards to natural language processing, presented in Khuman et al. (2015b, 2016d). In the attributed published works, it was shown that the concept of grey incidence could be utilised to measure the metric spaces of the sequence curves for an optimal string, against input strings. With no adjustment in the configuration, only a change in the domain of execution, the same ethos of analysis can be applied to perception based domains.

2.6.4 Absolute Degree of Grey Incidence

The traditional degree of grey incidence provides the basis for all variances of the degree of incidence; $\Gamma = [\gamma_{ij}]$, where each entry in the i^{th} row of the matrix is the degree of grey incidence for the corresponding characteristic sequence Y_i, and relevant behavioural factors $X_1, X_2, \ldots, X_m$. Each entry for the j^{th} column is reference to the degrees of grey incidence for the characteristic sequences $Y_1, Y_2, \ldots, Y_n$ and behavioural factors X_m. The absolute degree of grey incidence $A = [\epsilon_{ij}]_{n\times m}$, is defined as follows:

Definition 17 (Absolute Degree of Grey Incidence): Assume that X_i and $X_j \in \mathbb{U}$ are two non-negative time series sequences of the same magnitude, that are defined as the sum of the distances between two consecutive time points, whose zero starting points have already been computed. The absolute degree of grey incidence can be given as:

$$\epsilon_{ij} = \frac{1 + |s_i| + |s_j|}{1 + |s_i| + |s_j| + |s_i - s_j|} \tag{2.6.33}$$

Where the comparable sequences are given as:

$$\Big|s_i\Big| = \Big|\sum_{k=2}^{n-1} x_i^{(0)}(k) + \frac{1}{2}x_i^{(0)}(n)\Big|$$

$$\Big|s_j\Big| = \Big|\sum_{k=2}^{n-1} x_j^{(0)}(k) + \frac{1}{2}x_j^{(0)}(n)\Big|$$

Such that:

$$\Big|s_i - s_j\Big| = \Big|\sum_{k=2}^{n-1} \Big(x_i^{(0)}(k) - x_j^{(0)}(k)\Big) + \frac{1}{2}\Big(x_i^{(0)}(n) - x_j^{(0)}(n)\Big)\Big|$$

2.6.5 Relative Degree of Grey Incidence

The relative degree of grey incidence is obtained in much the same way, however, the sequences X_i' and X_j' are the initial images of X_i and X_j, defined as follows:

DEFINITION 18 (Relative Degree of Grey Incidence): Using the *initial* images of X_i and X_j, from which the zero images are then derived, the relative degree of grey incidence $B = [r_{ij}]_{n \times m}$, is given as follows:

$$r_{ij} = \frac{1 + |\acute{s}_i| + |\acute{s}_j|}{1 + |\acute{s}_i| + |\acute{s}_j| + |\acute{s}_i - \acute{s}_j|} \tag{2.6.34}$$

Where the initial sequences are given as:

$$X_i^{'(0)} = \left(x_i^{'(0)}(1), x_i^{'(0)}(2), \ldots, x_i^{'(0)}(n)\right) = \left(\frac{x_i^{(0)}(1)}{x_i^{(0)}(1)}, \frac{x_i^{(0)}(2)}{x_i^{(0)}(1)}, \ldots, \frac{x_i^{(0)}(n)}{x_i^{(0)}(1)}\right)$$

$$X_j^{'(0)} = \left(x_j^{'(0)}(1), x_j^{'(0)}(2), \ldots, x_j^{'(0)}(n)\right) = \left(\frac{x_j^{(0)}(1)}{x_j^{(0)}(1)}, \frac{x_j^{(0)}(2)}{x_j^{(0)}(1)}, \ldots, \frac{x_j^{(0)}(n)}{x_j^{(0)}(1)}\right)$$

Therefore, the comparable sequences are given as:

$$\left|s_i^{'}\right| = \left|\sum_{k=2}^{n-1} x_i^{'(0)}(k) + \frac{1}{2} x_i^{'(0)}(n)\right|$$

$$\left|s_j^{'}\right| = \left|\sum_{k=2}^{n-1} x_j^{'(0)}(k) + \frac{1}{2} x_j^{'(0)}(n)\right|$$

Such that:

$$\left|s_i^{'} - s_j^{'}\right| = \left|\sum_{k=2}^{n-1} \left(x_i^{'(0)}(k) - x_j^{'(0)}(k)\right) + \frac{1}{2}\left(x_i^{'(0)}(n) - x_j^{'(0)}(n)\right)\right|$$

2.6.6 Synthetic Degree of Grey Incidence

The synthetic degree of grey incidence is an amalgamation of both absolute and relative degrees of incidence, which is associated with the overall *closeness* between sequences. The value given for θ is typically set to 0.5, as this allows for equal measure of absolute and relative to be incorporated. If a higher value is chosen for θ, more emphasis is placed on the relationship between absolute quantities. If a smaller value is chosen, then more emphasis is placed on the relative rates of change between sequences of a system. It is given as follows:

DEFINITION 19 (Synthetic Degree of Grey Incidence): Assuming that both absolute and relative degrees of grey incidence have already been computed, the synthetic degree of grey incidence is calculated as follows:

$$p_{ij} = \theta_{\epsilon_{ij}} + (1-\theta)r_{ij} \tag{2.6.35}$$

Each variance of incidence takes a different perspective when inspecting the relationships between characteristics and behaviours. Based on the matrix of grey incidence Γ, if there exists k and $i \in \{1, 2, \ldots, n\}$ which satisfies $\gamma_{ki} \geq \gamma_{ij}$, one can infer that the grey system's characteristic Y_k is more favourable than the system's characteristic Y_i, denoted as $Y_k \succ Y_i$. If for any $i = \{1, 2, \ldots, n\}$ with $i \neq k$, there will always be $Y_k \succ Y_i$, which indicates that Y_k is the most favourable characteristic in the system. If there exists l and $j \in \{1, 2, \ldots, m\}$ satisfying $\gamma_{il} \geq \gamma_{ij}$, for $i = \{1, 2, \ldots, n\}$, then it can be understood that the behavioural factor X_l is more favourable than X_j, denoted as $X_l \succ X_j$. If for any $j = \{1, 2, \ldots, m\}$, $j \neq l$, one will always have $X_l \succ X_j$. Therefore X_l is the most favourable factor in the system. In the context of human based perception, the absolute degree of grey incidence was ultimately chosen. The entire sequence needs to be compared to its contemporary from the summated discretised points. If a synthetic degree of grey incidence was to be used, then θ would have to be set to 1, which would mean that it was acting exactly as you would expect the absolute degree of grey incidence to behave.

2.7 Closing Remarks

CHAPTER 2 has provided the background into the concepts that will constitute the new proposed framework. Understanding the notion of fuzzy sets, fuzzy membership values and rough approximations, provides one with the core components needed to fully understand an R-fuzzy approach. EXAMPLE 1 demonstrated the capability of an R-fuzzy set, whereby a membership was defined for an unknown flight, using the criteria set and the descriptor that described it. This thesis itself takes the foundational understanding of R-fuzzy sets, but applies it in a more generalised adaptation. Instead of creating an R-fuzzy set for unknown observations, the R-fuzzy sets created from hereon are generated for the descriptors that constitute the known observations of the criteria set C. As it will be shown in the successive examples, R-fuzzy sets based on what is known, to infer on membership values of unknowns, provides for a high level of detail, especially when considering continuous representation.

Also described was the paradigm of grey theory from its foundational understanding of numbers and sets, and its relationship to that of probability theory. To understand its perspective on uncertainty is to understand its perspective on analysis. The understanding of the absolute degree of grey incidence is paramount, as it is this aspect that provides the new framework the functionality of better inferring to a higher level of detail. Adapting it from its traditional deployment, and executing it in a novel way, will allow for the original R-fuzzy framework to be enhanced.

3

Novel Contributions

"It is in the admission of ignorance and the admission of uncertainty that there is a hope for the continuous motion of human beings in some direction that doesn't get confined, permanently blocked, as it has so many times before in various periods in the history of man."

– Richard P. Feynman

3.1 Introduction

The List of Publications documents the peer-reviewed and academically recognised works attributed to the research throughout its duration. However, not all of the publications are directly related to the thesis with regards to the proposed new framework. The work regarding an improved grey model for forecasting Khuman et al. (2013), although provided the initial insight into grey theory, does not play heavily in the main contributions of this thesis. One of the first incarnations of the research was to develop a framework that utilised a grey model of the first order (GM(1,1)), which used a single input sequence. A GM(1,1) grey model was used to forecast the perceived perceptions of differing age clusters. From the initial testing it was proved that such a framework was not very effective, as one condition of the GM(1,1) is to have monotonic data. This cannot always be guaranteed, nor is it a realistic expectation, especially when involving cohorts of different age groups regarding varying perceptions. Nonetheless, the initial use of sequencing provided one with a foundation from which to explore further. This chapter is structured as follows, highlighting the novel contributions that are directly related to this thesis, and connected to the hypothesis.

Section 3.2 proposes the significance measure, the intermediary component to allow for the information collected using an R-fuzzy approach, to be analysed using grey analysis. Without the significance measure the degree of grey incidence would not be able to be implemented, as there would be no sequence to be acted upon. Section 3.3 puts forward the streamlined encapsulation approach, for providing the minimal configuration needed to correctly contain all relative and triggered uncertain fuzzy membership values. Section 3.4 describes and demonstrates the enhanced R-fuzzy framework. Bridging together the additional uncertainty encapsulation capabilities of the standard R-fuzzy approach, along with the analysis that grey sequencing allows for. As it will be shown, the enhanced framework provides for an additional dimension of complexity to be inferenced from. Section 3.5 discusses the similarity and distance measure that one could make use of when considering the enhanced R-fuzzy framework. Section 3.6 describes the notion of an R-fuzzy α-cut for post processing of the returned results, providing a more humanistic interpretation. Such is the subjective nature of perception, an unexpected response can impact upon the containment of a membership value. The significance measure in this respect can stipulate where one could implement the use of α-cuts to either reduce or increase the degree of significance for *problematic* membership values. In addition, Section 3.6.1 describes how when using an α-cut, one can replicate a shadowed set approach, whereby the three possible areas indicate three-valued logic. Section 3.7 remarks upon the chapter, and summarises on the contributions of this thesis.

As one can see from Definition 8, the membership function of an R-fuzzy set is a rough set of contained viable uncertain fuzzy membership values. What is not clear is the strength associated to each of these captured memberships, ergo the creation of the significance measure, introduced in Khuman et al. (2015a, 2016a). Section 3.2 will now go on to describe the concept of the significance measure more precisely, demonstrating it through a worked example.

3.2 The Significance Measure

The significance measure is the main contribution from this thesis, and as such it is paramount to the overall new proposed framework. It can be seen as a logical enhancement to that of a traditional R-fuzzy set. It can also act as the facilitator that allows for a sequence to be generated so that grey analysis can be undertaken. The significance measure itself is based on and takes its inspiration from the certainty factor employed by traditional rough set theory, which itself is based on a variation of Bayes' theorem (Pawlak, 1998). Rough set theory allows one to employ the use of strength, certainty and coverage factors, each one providing an insight into a particular rule or fact. As they are associated with decision systems indicative of Table 2.2, and rule induction, as such they cannot be translated over to R-fuzzy sets without modification. With every decision rule $A \rightarrow_x D$ there is an associated certainty factor of the decision rule, where A is the rule, D is the decision, and *supp* is a frequency count. The certainty factor presented from a rough set perspective is given as follows:

$$cer_x\Big(A, D\Big) = \frac{|A(x) \cap D(x)|}{|A(x)|} \tag{3.2.1}$$

It can be viewed as a conditional probability that y belongs to $D(x)$ given that y belongs to $A(x)$, symbolically $\pi_x(D|A)$. If $cer_x(A, D) = 1$, then $A \rightarrow_x D$ will be referred to as a certain decision rule, if $0 < cer_x(A, D) < 1$ the decision rule will be referred to as an uncertain decision rule. To allow for the use of a *certainty* factor in an R-fuzzy context, one has to remove the notion of rules, induction of rules or quantification of rules. Therefore making the new significance measure relative to the subset of all values based on $M_{pj}(x) \subseteq J_x$.

Definition 20 (Significance Measure): Using the same notations presented in Definition 8 that described an R-fuzzy set, assume that an R-fuzzy set $M(x_i)$ has already been created, and that a membership set of values J_x and a criteria set C are known. Given that the total number of subsets generated for a given R-fuzzy set is given by $|N|$, and that S_v is the number of subsets that contain the specified membership value being inspected. As each value $v \in J_x$ is evaluated by $c_j \in C$, the significance measure therefore counts the number of instances that v occurred over $|N|$, given by the formal expression:

$$\gamma_{\bar{A}}\{v\} = \frac{S_v}{|N|} \tag{3.2.2}$$

The significance measure expresses the conditional probability that $v \in J_x$ belongs to the R-fuzzy set $M(x_i)$, given by its descriptor $d(x_i)$. The value returned will be presented as a standard fraction, where the denominator $|N|$ represents the cardinality of the total number of generated subsets. The numerator S_v is indicative of the number of instances that the inspected membership value $v \in J_x$ occurred. As $|N|$ provides the scope of the domain in terms of magnitude, S_v will never exceed $|N|$. In essence providing a normalised output which in turn can be translated into any real value, where $\gamma_{\bar{A}}\{v\} : J_x \to [0, 1]$ is the membership function, much like that of Eq. (2.2.2).

3.2.1 Validation

If the value returned by $\gamma_{\bar{A}}\{v\} = 1$, then that particular membership value has been agreed upon by all in the criteria set C. As a result one will know that it absolutely belongs to the lower approximation, as for it to be included the entire populous must agree:

$$\underline{M}_A = \left\{ \gamma_{\bar{A}}\{v\} = 1 \;\middle|\; v \in J_x \subseteq [0, 1] \right\} \tag{3.2.3}$$

Much like before how a lower approximation is a subset of the upper approximation, any membership value with a significance degree of $\gamma_{\bar{A}}\{v\} = 1$, will also be included within the upper approximation. If $\gamma_{\bar{A}}\{v\} = 0$, it can be concluded that absolutely no one perceived that particular membership value to satisfy the descriptor. If $0 < \gamma_{\bar{A}}\{v\} < 1$, then that particular membership value has some significance to some degree relative to the descriptor $d(x_i)$. As a result, this particular value will knowingly be contained within the upper approximation:

$$\overline{M}_A = \left\{ \gamma_{\bar{A}}\{v\} > 0 \;\middle|\; v \in J_x \subseteq [0, 1] \right\} \tag{3.2.4}$$

These interpretations echo the sentiments of fuzzy set theory as presented in Definition 1, whereby an element can be described by its membership function such that it returns any real number in the range $[0, 1]$. Except instead of representing the belongingness of an object to a particular set, the significance degree returns the measure of significance, with relation to its descriptor $d(x_i)$, based on its conditional probability of distribution. Eq. (3.2.2) can be rewritten so that the collected significance degrees constitute a set, given by the following expression:

$$\bar{A} \ = \ \left\{ \left\langle v, \gamma_{\bar{A}}\{v\} \right\rangle \ \middle| \ v \in J_x \subseteq [0,1] \right\} \tag{3.2.5}$$

Where $\bar{A}$ is a set describing the distribution of a specified descriptor $d(x)$, for which the generated R-fuzzy set was created for. It must be understood that fuzzy sets are not concerned with probability, merely the degree of belongingness (Drakopoulos, 1995). As the newly derived significance measure itself is based on Eq. (3.2.1), then the significance is associated with the conditional probability of distribution. The significance measure will in affect validate any value contained in the lower approximation as $\gamma_{\bar{A}}\{v\} = 1$, as this is considered as an absolute truth agreed upon by all. Whereas $\gamma_{\bar{A}}\{v\} = 0$, will validate that the inspected membership value was not considered significant to any degree by anyone. The greater the returned value for $\gamma_{\bar{A}}\{v\}$ the greater its significance with regards to the descriptor that the R-fuzzy set is being modelled for. Understanding the significance of any inspected membership value, can be used to understand the perception of the populous that it was generated from. R-fuzzy sets allow for every conceivable perception to be incorporated, that includes all possible outliers. The associated degrees of significance will quantify just how important or unimportant a membership value truly is, based on the perceptions collected.

3.2.2 Intended Use

Considering Example 1, the use of the R-fuzzy approach was to provide one with a membership value for flight f_{11}, which did not have an associated noise level, only the notion of it being *AC*. Based on the collected perceptions in the criteria set C for all observers, a final generated R-fuzzy set was created for f_{11}, given as:

$$M\big(f_{11}\big) = \Big(\{0.43\}, \{0.14, 0.29, 0.43, 0.57\}\Big)$$

This result implied that 0.43 be an indicative representative of *AC*, as it was the only value contained within the lower approximation. The values contained in the upper approximation were indicative of *AC* to some unknown degree. The originally intended use of the significance measure was not necessarily meant to explicitly define R-fuzzy sets for unknown flights, but rather to quantify the predefined descriptors to represent the entire criteria set C. The noise flight example provided an R-fuzzy set, although based on all collected perceptions of C, it was used to describe f_{11}, a single unknown instance of a flight. The significance measure itself can be used in conjunction with any generated R-fuzzy set, however, it is best utilised when the generated R-fuzzy sets are modelling the descriptors for the entire populous. Given that the prerequisite for an R-fuzzy approach is to have a populated criteria set of descriptors, and a valid fuzzy membership set J_x, adjusting the traditional use of the R-fuzzy approach so that the generated sets themselves are indicative for the whole of C, rather than an unknown instance, pays dividends. It would be more fruitful if the R-fuzzy sets generated were based on the descriptors for the entire collected population. The significance measure when applied to these *universal* R-fuzzy sets of C, allows for the entire universe of discourse to be encapsulated and modelled. The fuzzy membership values constituting J_x, provide the discretised points on the universe, as they remain the same for all computed R-fuzzy sets, a complete comprehensive understanding of the perceptions of all involved is the result. The returned degrees of significance provide the amplitude from which one can use to create a continuous representation for the entire populous of C. This in turn can then be used to be inferred from when considering observations where only the descriptors are given. This will allow for the same resultant sentiment of Example 1 to be maintained, with the adage that the enhanced framework is more thorough and exhaustive.

3.2.3 A Significance Measure Example

Example 2: Referring back to Example 1, it would be meaningful if one was able to obtain the importance of the membership values contained within the upper approximation. By using Eq. (3.2.2), we are able to provide a means of quantification which in turn provides a significance coefficient equal to or within the range of $[0, 1]$. The higher the significance, the more individuals that agreed to it being an indicative representation for their subjective perception. The newly derived significance measure is relatively simple, in that it is a statistical method for counting the significance of a particular membership in relation to J_x and its descriptor. For each membership value contained in J_x, one simply

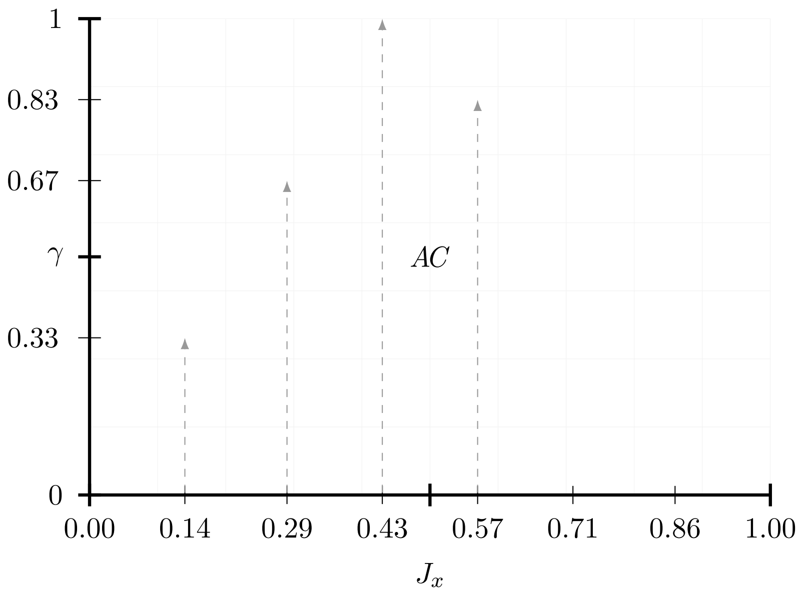


Figure 3.1: A Discrete Visualisation for AC

counts the number of occurrences each instance occurred for each of the subsets created for M_{pj} for a given descriptor. One can obtain the following significance coefficient values for each of the membership values contained in J_x, the results are given as follows:

$$\gamma_{\overline{AC}}\{0.00\} = \frac{0}{6} = 0.00 \qquad \gamma_{\overline{AC}}\{0.14\} = \frac{2}{6} = \frac{1}{3} = 0.33$$

$$\gamma_{\overline{AC}}\{0.29\} = \frac{4}{6} = \frac{2}{3} = 0.67 \qquad \gamma_{\overline{AC}}\{0.43\} = \frac{6}{6} = 1.00$$

$$\gamma_{\overline{AC}}\{0.57\} = \frac{5}{6} = 0.83 \qquad \gamma_{\overline{AC}}\{0.71\} = \frac{0}{6} = 0.00$$

$$\gamma_{\overline{AC}}\{0.86\} = \frac{0}{6} = 0.00 \qquad \gamma_{\overline{AC}}\{1.00\} = \frac{0}{6} = 0.00$$

A visualisation for when the descriptor is set to AC, based on its returned degrees of significance for the values contained in the membership set J_x, is presented in Figure 3.1. The membership value 0.43 returns a degree of significance of $\gamma_{\overline{AC}}\{0.43\} = 1$, echoing the fact that it was agreed upon by all. Figure 3.1 does indeed indicate and also validates that the membership value 0.43 was included within the lower approximation. For any value to score a significance of 1 satisfies the requirement given by Eq. (3.2.3). The membership value of 0.57 returns a degree of significance of $\gamma_{\overline{AC}}\{0.57\} = 0.83$, a relatively high score. If 5 out of the 6 individuals agreed for this membership to be a suitable value for AC, then one may be inclined to treat it as such. Inspecting the membership values that returned

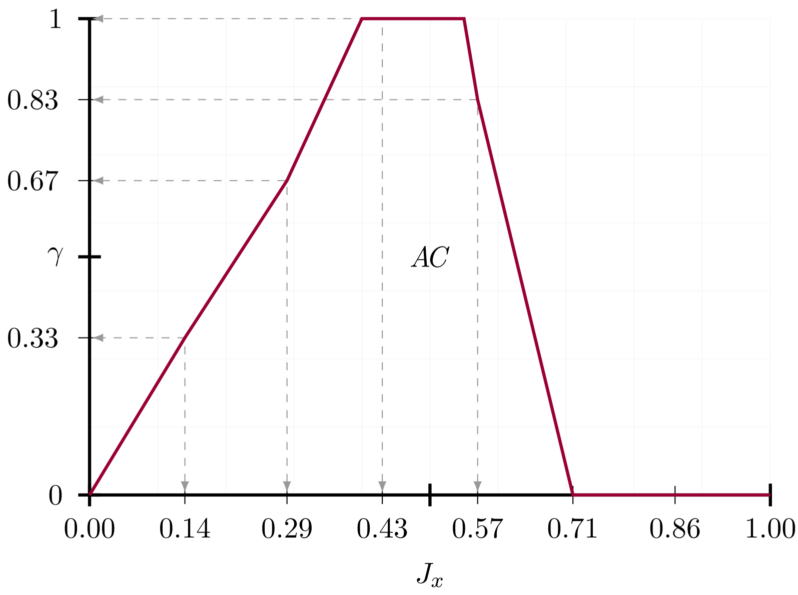


Figure 3.2: A Possible Continuous Representation for *AC*

a significance of 0, those memberships were absolutely disregarded as being candidates to represent the descriptor *AC*.

The projection of the histogram lines for the triggered membership values, provides what could be a convex hull to form the parameters of a set. As was stated in Section 3.2, the significance degree returns the significance of the membership value relative to its descriptor based on the criteria set C. If one was to provide a set to encapsulate the significance degree values, one has then inadvertently created the equivalence of a fuzzy set as defined in Definition 1. Such is the equivalence, the returned degree of significance is the degree of membership to the set representing the descriptor, if and only if the values for the significance degree provide the values for the membership function parameters. A more detailed explanation of this equivalence is provided in Section 4.7.1.

Figure 3.1 presents a discretised interpretation of the degree of significance for the descriptor *AC*, whereas Figure 3.2 provides a *possible* continuous interpretation for the descriptor *AC*. Inspecting the continuous plot, one can see that the plot shares similarities to the trapezoidal membership function given in Eq. (2.2.4). One can easily infer from this continuous representation that the significance degrees are also the degrees of membership, akin to a fuzzy perspective. The parameters for the set were based on the returned degrees of significance, hence why the lines intersect through the apex of the stick heights for the generated degrees of significance. This can be seen as an intuitive dot-to-dot, where the joining of the apex of each triggered fuzzy membership value creates the shape of the membership function.

Notice how Figure 3.2 utilises a trapezoidal-esque membership rather than the expected triangular membership. This was done to show that variation is still allowed and the choice for a trapezoidal was more in keeping with human perception. Using a triangular membership would result in the membership value 0.42 not scoring a degree of 1 for its significance, as would be the case for 0.44 and so forth. It would seem more likely and indeed more plausible that a membership value of 0.42 and 0.44 would indeed be agreed upon by all, especially if the triggered 0.43 was, as a result a trapezoidal membership was chosen. However, this is only an assumption, as the membership value 0.42 was not a discrete viable option according to the fuzzy membership set J_x.

At what point does one define the interval where any value would return a degree of 1? This is open for debate and as such, an assumption is just an assumption, regardless of how it is derived. Nonetheless, a logical assumption, indicative of human intuition is still better than a wild stab in the dark.

The apex of the interval is given by the range $[0.43, 0.55]$, a completely arbitrary assumption, one which could be further enhanced by a human intuition. By utilising the returned degrees of significance for the descriptor being used, one is able to construct a set which returns the correct degrees of significance, and the equivalent degree of inclusion as seen from a fuzzy perspective, only if the parameters of the set are formed from the apex of the stick heights. The provided continuous representation was derived by simple means, and as a result allows one to visualise a possible interpretation of how human perception given by a descriptor, propagates through the membership values.

Inspecting the significance of each of the membership values, one is able to apply a linguistic assumption to describe each contained uncertain fuzzy value, such as:

$$\gamma_{\overline{AC}}\{0.00\} = 0.00 \rightarrow \text{Agreed upon by none}$$
$$\gamma_{\overline{AC}}\{0.14\} = 0.33 \rightarrow \text{Agreed upon by a few}$$
$$\gamma_{\overline{AC}}\{0.29\} = 0.67 \rightarrow \text{Agreed upon by the majority}$$
$$\gamma_{\overline{AC}}\{0.43\} = 1.00 \rightarrow \text{Agreed upon by all}$$
$$\gamma_{\overline{AC}}\{0.57\} = 0.83 \rightarrow \text{Agreed upon by the majority}$$
$$\gamma_{\overline{AC}}\{0.71\} = 0.00 \rightarrow \text{Agreed upon by none}$$
$$\gamma_{\overline{AC}}\{0.86\} = 0.00 \rightarrow \text{Agreed upon by none}$$
$$\gamma_{\overline{AC}}\{1.00\} = 0.00 \rightarrow \text{Agreed upon by none}$$

Using simple statements, one is able to easily ascertain the intent of any membership value by referring to its description. The statements chosen were trivial, but further consideration could be undertaken. This would further enhance the ability to understand the significance of any membership value inspected.

If the descriptor for the example was changed to NN, the following subsets $M_{pj}(NN) \subseteq J_x$ would be generated:

$$M_{p1}(NN) = \{0.00, 0.14, 0.29\} \qquad M_{p2}(NN) = \{0.00, 0.14\}$$
$$M_{p3}(NN) = \{0.00\} \qquad M_{p4}(NN) = \{0.00, 0.14, 0.29\}$$
$$M_{p5}(NN) = \{0.00\} \qquad M_{p6}(NN) = \{0.00, 0.14\}$$

This would create the following R-fuzzy set:

$$M(NN) = \Big(\{0.00\}, \{0.00, 0.14, 0.29\}\Big)$$

This generated R-fuzzy set implies that the fuzzy membership value 0.00 is the only membership value to have been agreed upon by all in the criteria set C. It also implies that the values 0.14 and 0.29 were also agreed upon to some extent.

Table 3.1: The Collected Degrees of Significance for Noise Perception

NN		*AC*		*BN*		*VN*	
J_x	γ	J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{NN}}\{0.00\} =$	1.00	$\gamma_{\overline{AC}}\{0.00\} =$	0.00	$\gamma_{\overline{BN}}\{0.00\} =$	0.00	$\gamma_{\overline{VN}}\{0.00\} =$	0.00
$\gamma_{\overline{NN}}\{0.14\} =$	0.67	$\gamma_{\overline{AC}}\{0.14\} =$	0.33	$\gamma_{\overline{BN}}\{0.14\} =$	0.00	$\gamma_{\overline{VN}}\{0.14\} =$	0.00
$\gamma_{\overline{NN}}\{0.29\} =$	0.33	$\gamma_{\overline{AC}}\{0.29\} =$	0.67	$\gamma_{\overline{BN}}\{0.29\} =$	0.00	$\gamma_{\overline{VN}}\{0.29\} =$	0.00
$\gamma_{\overline{NN}}\{0.43\} =$	0.00	$\gamma_{\overline{AC}}\{0.43\} =$	1.00	$\gamma_{\overline{BN}}\{0.43\} =$	0.00	$\gamma_{\overline{VN}}\{0.43\} =$	0.00
$\gamma_{\overline{NN}}\{0.57\} =$	0.00	$\gamma_{\overline{AC}}\{0.57\} =$	0.83	$\gamma_{\overline{BN}}\{0.57\} =$	0.17	$\gamma_{\overline{VN}}\{0.57\} =$	0.00
$\gamma_{\overline{NN}}\{0.71\} =$	0.00	$\gamma_{\overline{AC}}\{0.71\} =$	0.00	$\gamma_{\overline{BN}}\{0.71\} =$	1.00	$\gamma_{\overline{VN}}\{0.71\} =$	0.00
$\gamma_{\overline{NN}}\{0.86\} =$	0.00	$\gamma_{\overline{AC}}\{0.86\} =$	0.00	$\gamma_{\overline{BN}}\{0.86\} =$	0.67	$\gamma_{\overline{VN}}\{0.86\} =$	0.33
$\gamma_{\overline{NN}}\{1.00\} =$	0.00	$\gamma_{\overline{AC}}\{1.00\} =$	0.00	$\gamma_{\overline{BN}}\{1.00\} =$	0.00	$\gamma_{\overline{VN}}\{1.00\} =$	1.00

If the descriptor was changed to *BN*, the following subsets $M_{pj}(BN) \subseteq J_x$ would be generated:

$$M_{p1}(BN) = \{0.86, 0.71\} \qquad M_{p2}(BN) = \{0.86, 0.71\}$$
$$M_{p3}(BN) = \{0.57, 0.71\} \qquad M_{p4}(BN) = \{0.86, 0.71\}$$
$$M_{p5}(BN) = \{0.86, 0.71\} \qquad M_{p6}(BN) = \{0.71\}$$

This would ultimately create the following R-fuzzy set:

$$M(BN) = \Big(\{0.71\}, \{0.57, 0.71, 0.86\}\Big)$$

This generated R-fuzzy set implies that the fuzzy membership value 0.71 is the only membership value to have been agreed upon by all in the criteria set C. It also implies that the values 0.57 and 0.86 were also agreed upon to some extent.

If the descriptor was changed to VN the following subsets $M_{pj}(VN) \subseteq J_x$ would be generated:

$$M_{p1}(VN) = \{1.00\} \qquad M_{p2}(VN) = \{1.00\}$$
$$M_{p3}(VN) = \{0.86, 1.00\} \qquad M_{p4}(VN) = \{1.00\}$$
$$M_{p5}(VN) = \{1.00\} \qquad M_{p6}(VN) = \{0.86, 1.00\}$$

The resultant R-fuzzy set would be of the form:

$$M(VN) = \Big(\{1.00\}, \{0.86, 1.00\}\Big)$$

This generated R-fuzzy set implies that the fuzzy membership value 1.00 is the only membership value to have been agreed upon by all in the criteria set C. In much the same way as the previously computed R-fuzzy sets, it also implies that the value 0.86 was also agreed upon to some extent.

The associated degrees of significance for each generated R-fuzzy set, using Eq. (3.2.2), have been computed and collected, and are presented in Table 3.1. If one inspects the table, the clustering of triggered fuzzy membership values is apparent for each generated R-fuzzy set. By that, where there is a membership value with a returned degree of significance of 1, the fuzzy membership values either side of that, if applicable, are also triggered, to some degree.

A discrete and continuous plot for NN is presented in Figure 3.3 and Figure 3.4, respectively. As the R-fuzzy set NN is the left most set, it is logical to assume that it would only be involved with the left most fuzzy membership values. With 0.00 being the only value to score a 1 for its significance, the successive triggered membership values directly to the right, taper off with lesser degrees of significance, as it clearly shown in the plots.

A discrete and continuous plot for BN is presented in Figure 3.5 and Figure 3.6, respectively. Much like the previous plots, the stick heights of the triggered membership functions values are intersected. The interval for the apex could be reduced or even exaggerated further, as too could the left and right most anchor points of the set.

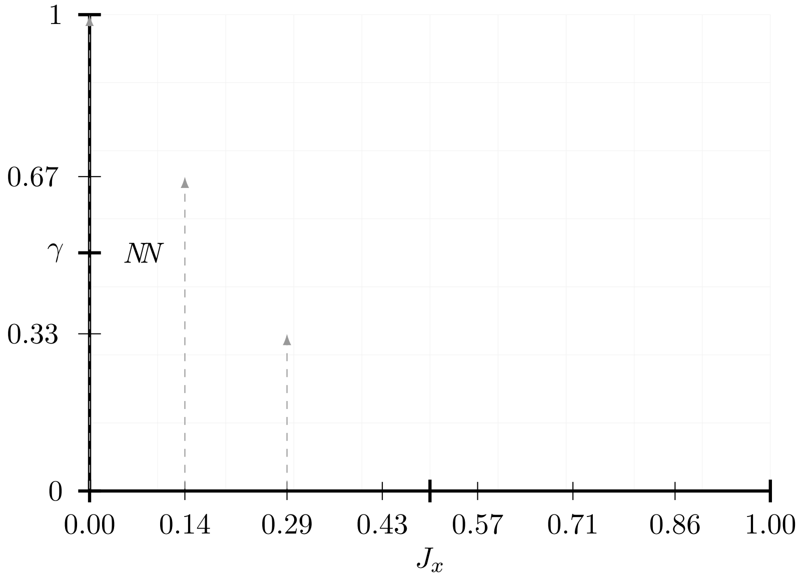


Figure 3.3: A Discrete Visualisation for *NN*

A discrete and continuous plot for *VN* is presented in Figure 3.7 and Figure 3.8, respectively. Much like the plot for *NN*, the *VN* is the right most plot and would logically be associated to the right most fuzzy membership values, which indeed it is. With the membership value of 1.00 being the only membership to score a 1 for its significance, the membership value directly to the left, 0.86, has a lesser degree of significance $\gamma_{\overline{VN}}\{0.86\} = 0.33$.

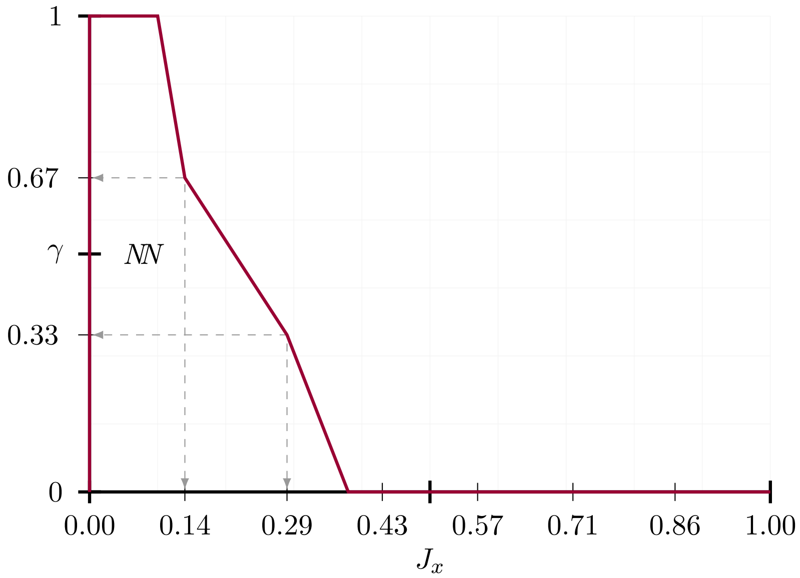


Figure 3.4: A Possible Continuous Visualisation for *NN*

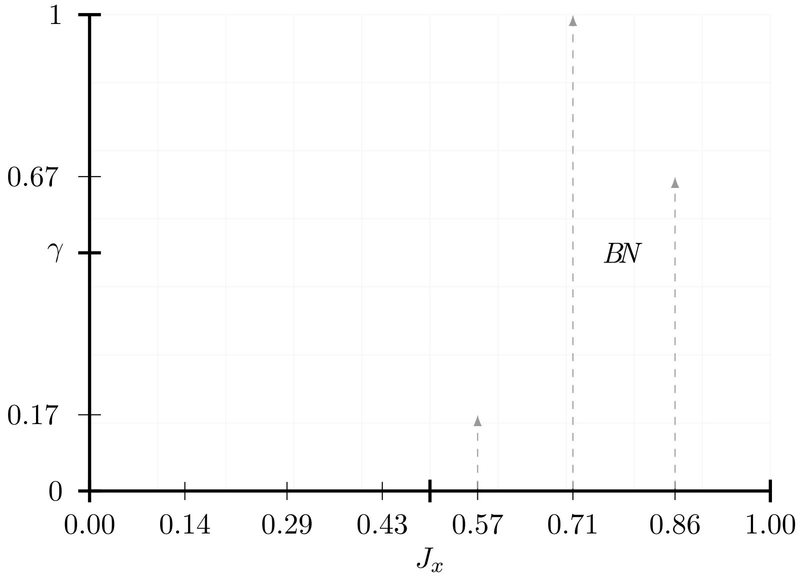


Figure 3.5: A Discrete Visualisation for *BN*

The addition of utilising the significance measure to inspect individual membership values allows one to better appreciate and conserve the diversity of the perceptions being collected. The values contained within Table 3.1 are the collective significance measures for each of the specified descriptors. By placing the membership values contained in J_x in ascending order, and so too the descriptors, starting with *NN* through to *VN*, one can see how the collected human perception of criteria set C, propagates through from one

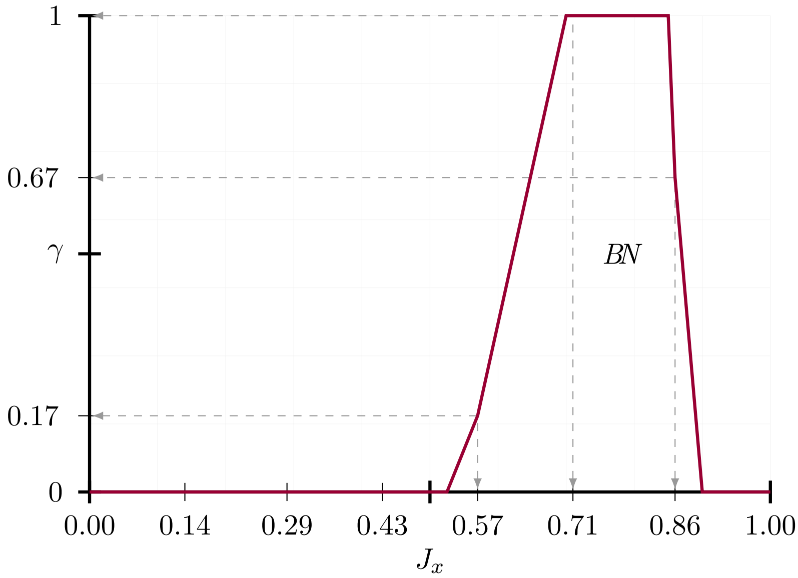


Figure 3.6: A Possible Continuous Visualisation for *BN*

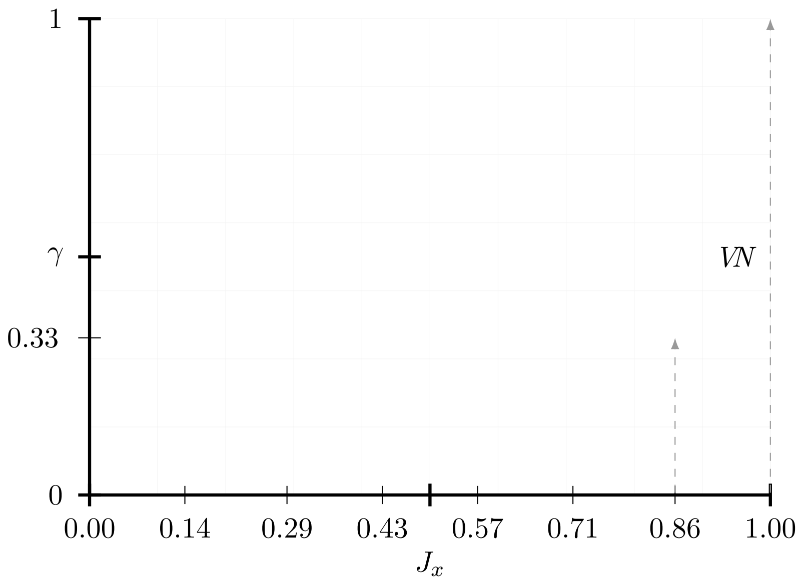


Figure 3.7: A Discrete Visualisation for *VN*

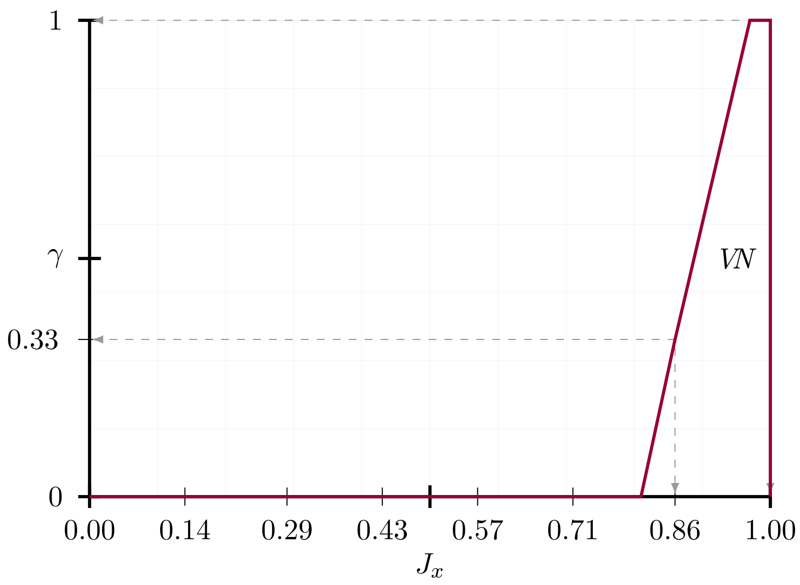


Figure 3.8: A Possible Continuous Visualisation for *VN*

descriptor to the next. This gives credence to the fact that the example collectively and correctly, was able to provide a realistic assumption of the various descriptors based on what was known. The degrees of significance can be used to construct customised sets, such that the degree of significance is also the degree of membership to the descriptor the set is being modelled for. Figures 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8 all show how this can be visualised.

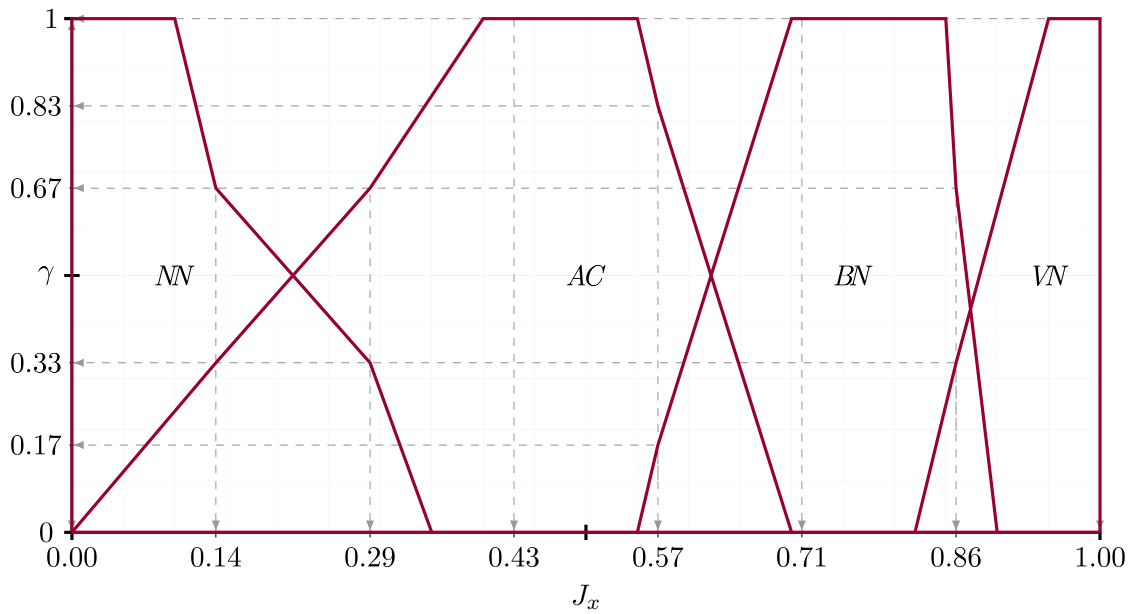


Figure 3.9: A Continuous Visualisation for *NN*, *AC*, *BN* & *VN*

The plot presented in Figure 3.9 shows the collective visualisation of sets based on all the generated R-fuzzy sets. This provides a means to understand the entire collective perceptions of all involved, for all descriptors, for the entire domain. As the plot shows, each descriptor or perspective, does not follow uniformed symmetry. The sets themselves are not equal to one another, especially with regards to area and distribution. This synthesis of human perception is more probable and in keeping, than strict symmetrical uniformed sets. Paying particular interest to the set representing the descriptor *AC*, one can see that the membership value 0.14 has a degree of significance of $\gamma_{\overline{AC}}\{0.14\} = 0.33$. The same membership value has a significance of $\gamma_{\overline{NN}}\{0.14\} = 0.67$ to the set describing the descriptor *NN*. The fact that the sets themselves are constructed via the stick heights of the triggered degrees of significance, allows one to establish an equivalence between a fuzzy set and the generated significance set. Albeit, an equivalence in the returned value and not the perspective the values are investigated from.

Example 2 has shown that the significance measure allows for a higher order of detail to be inferred from. As an R-fuzzy set recognises and more importantly conserves the diversity inherent in human perception, the returned degrees of significance provide valuable information for airports and for people living within proximity. Conventional models would not be able to, or find it very difficult to conserve such diversity. Ignoring the importance of any relative membership value may lead to false noise exposure annoyance, which may cause even more problems for residents. An individual value is not lost within an interval nor a shadowed area, it can be readily inspected and inferred from. The various contained examples could be further enhanced in the aiding of diagnosing visual impairments. Any human perception based environment becomes an applicable domain for an R-fuzzy and significance measure coupling. Even with a small sample, a consensus can still be achieved, providing a generalisation all the while still conserving each and every specific perception. In reality an R-fuzzy approach can be applied to a minimum of 1 observer, from which the significance degree can create telling sets to encapsulate perceptions. Any number of additional observers and their perceptions can be integrated, forever evolving the significance degree measures and the shape of the encapsulating sets themselves.

In the case for Example 2, there is only one membership value for each generated R-fuzzy set to score 1 for its significance. However, this may not always be the case, as it will be shown in Chapter 4, a universal acceptance of a specific value may not always hold true for all involved. Nor can it always be expected that successive membership values be triggered, there may be a gap, in which case a single R-fuzzy set visualisation make look like two or several disjoint membership functions. The significance measure itself is purely objective and can only reiterate the significance of the distribution of each R-fuzzy set. As the criteria set is populated by subjectiveness, there may indeed be instances where expected interpretations are not met. Additionally, the variance between individual observer responses may be rather large or incredibly small, as will be seen in Example 4. Assuming that perceptions collected are truthful, the significance measure provides a metric. One which is not biased nor influenced by external subjectiveness; it does what it does, on what it was presented with. The output of which can then be used as one sees fit.

When considering that the R-fuzzy set computed in Example 1, was for a single unknown flight, associated to the descriptor *AC*, one can see that Example 2 provides for an enhanced amount of detail. Figure 3.9 provides a clear template based upon the perceptions contained in the criteria set C. One can inspect the plot and now make informed decisions based on unknown flights described by just their description. For example, in much the same way that flight f_{11} was described as *AC*, assume that another unknown flight was described using *BN*. In this case, referring to the plot in Figure 3.9, one may be inclined to provide a fuzzy membership value of 0.71, or give an interval of $[0.70, 0.85]$, as implied by the arbitrary plateau given for the *BN*. If the flight was partially described as *AC* and *BN* by an even mix, then according to the plot, one may be inclined to give a value halfway between the footprints of the two descriptors, which would be 0.625, or simply an interval of $[0.55, 0.70]$.

3.3 Streamlined Encapsulation

While investigating the concept of grey theory, the notion of the *typical grey whitenisation weight function* was touched upon, as presented in Definition 21. One of the publications attributed to this thesis Khuman et al. (2016b), involved the use of said weight function for the configuration of the membership function, which encapsulates the returned degrees of significance for any given R-fuzzy set. By using the whitenisation weight function to plot the returned degree of significance, one is able to provide a more specific, versatile heuristic, as compared to the arbitrary selection of points given in the original paper which proposed the significance measure Khuman et al. (2013). Not only are the significance degrees correctly intersected, the minimal number of parameter values are always used, instead of the various and arbitrary points that the original worked adopted. Using grey techniques, the result is a high precision set for encapsulation, with the minimal configuration of parameter values.

The use of whitenisation weight functions from grey theory provides for a heuristic based approach that the original R-fuzzy and significance framework lacks. Much like the original work, the whitenisation functions themselves are based on the returned degrees of significance for each R-fuzzy set that has been modelled. However, unlike the original arbitrary values decided for the function points, the grey approach provides a more streamlined perspective. By using an iterative process of optimisation and a combination of traditional triangular and trapezoidal membership functions, the encapsulated degrees of significance are precisely modelled using less parameter overhead. There are several types of whitenisation functions, but this research concerns itself only with the *typical* whitenisation function with fixed starting points. These starting points will be indicative of the starting and end points of the encapsulated candidate membership values contained within its R-fuzzy set. Rather than intersecting each apex height for each triggered membership value for a given set, the grey whitenisation function uses the minimum parameters required to contain all membership values with their correct corresponding intersections, to a high degree of resolution. The incorporation of a threshold value e to regulate the error rate of each whitenisation configuration is used to secure preciseness.

3.3.1 Whitenisation

Definition 21 (Typical Weight Function of Whitenisation): Assume a continuous function with fixed end points, which are increasing on the left $L(x)$ and decreasing on the right $R(x)$, this is described as a typical weight function of whitenisation.

$$f(x) = \begin{cases} L(x) = \dfrac{x - x_1}{x_2 - x_1}, & \text{if } x \in \big[x_1, x_2\big) \\ 1, & \text{if } x \in \big[x_2, x_3\big] \\ R(x) = \dfrac{x_4 - x}{x_4 - x_3}, & \text{if } x \in \big(x_3, x_4\big] \end{cases} \tag{3.3.6}$$

Eq. (3.3.6) describes the typical trapezoidal whitenisation function with fixed weights. For this to be transformed into a triangular whitenisation weight function, one simply replaces the interval at the apex $[x_2, x_3]$, with a single value. The whitenisation function itself can be seen to act exactly like that of a traditional fuzzy triangular, or trapezoidal membership function, indicative of Eq. (2.2.3) and Eq. (2.2.4), respectively.

3.3.2 A Whitenisation Example

Example 3: Consider the original discrete plot for AC given in Figure 3.1, and also consider the possible continuous representation given in Figure 3.2. Thus far, this thesis has made use of arbitrary values to provide the convex hull of the encapsulated significance values. One can instead make use of the whitenisation weight function to provide for

Table 3.2: The Whitenisation of AC

J_x	γ	*Weight Function* $[0, 0.43, 0.71]$	*Error* e	*Weight Function* $[0, 0.43, 0.85]$	*Error* e	*Weight Function* $[0, 0.43, 0.54, 0.71]$	*Error* e
0.00	0	0	0.0000	0	0.0000	0	0
0.14	0.33	0.3256	0.0044	0.3256	0.0044	0.3256	0.0044
0.29	0.67	0.6744	0.0044	0.6744	0.0044	0.6744	0.0044
0.43	1	1	0.0000	1	0	1	0
0.57	0.83	0.5	0.3300	0.6667	0.1633	0.8235	0.0065
0.71	0	0	0.0000	0.3333	0.3333	0	0
0.86	0	0	0.0000	0	0.0000	0	0
1.00	0	0	0.0000	0	0.0000	0	0

a continuous representation. We begin by structuring a triangular based whitenisation function using the values given in the membership set J_x from the previous example. Consider the R-fuzzy set AC given in Figure 3.2 and its associated degrees of significance, as a result the initial left anchor point is at 0.00. It would not be viable to have the left anchor point at 0.43, as this would indicate its returned degree of significance is 0. As it is already known that the fuzzy membership value of 0.43 returned a significance degree of 1, this can therefore act as the apex index. Once the membership value reaches

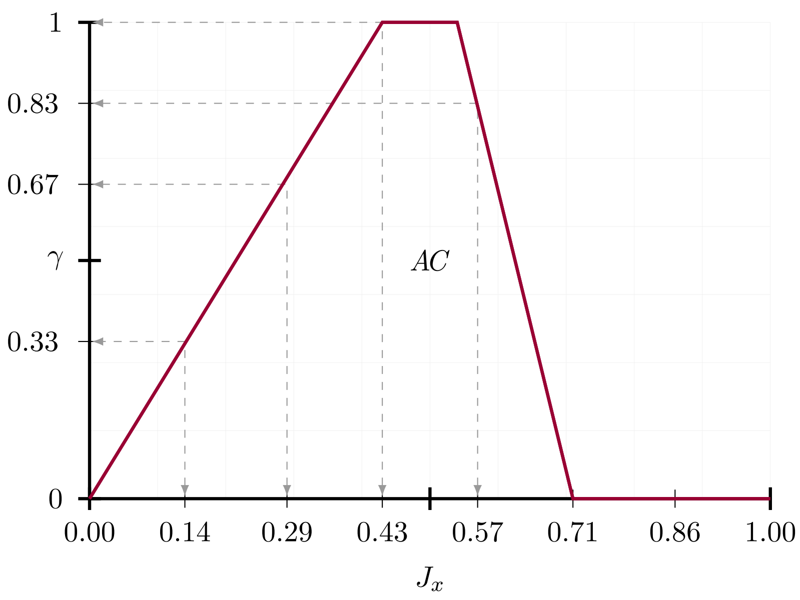


Figure 3.10: A Continuous Streamlined Representation for AC

0.71, the significance degree returns a value of 0, so therefore this becomes the right most anchor point. For this initial state, the parameter values are given as [0,0.43,0.71], these are purely random, although within reason, given the constraints of the set. Using a triangular membership characteristic function, similar to the one presented in DEFINITION 21, one can cross check the degree of membership to the new membership set, to see if it is the same as the returned degree of significance, the results of which are presented in Table 3.2.

The first column contains the membership values which belong to the membership set J_x. The second column presents the degrees of significance for each of the corresponding membership values. The third column puts forward the returned degree of inclusion based on the initial state of the parameter values chosen for the whitenisation weight function. The membership values in J_x are in turn passed through to the weight function from which the results are recorded. The fourth column calculates the error given by e, this is simply the absolute difference between the degree of significance known to be true, against the values returned by the whitenisation weight function. Generally speaking, the returned error values will often be small, however, relative to the error threshold, this will determine whether they need to be adjusted.

The main goal of the weight function is to encapsulate the returned degrees of significance such that they stay true to the original values, and that the error is as small as possible. The error rate for this example has been set to $e = 0.01$, so if any value after the absolute difference has been calculated has exceeded this threshold, the weight function parameters will need to be readjusted and recalculated accordingly. According to the fourth column, the error value for the membership value 0.57, returned an error of 0.3300, highlighted in red, this far exceeds the error threshold of 0.01.

The fifth column indicates a change in the parameter values for the weight function, as such the new error values are presented in the sixth column. These new parameter values have reduced the error for the membership value to 0.1633, still unacceptable according to the threshold of 0.01. But more alarming, the membership value of 0.71 has now also registered as a viable candidate, even though the returned degree of significance was an absolute 0 for this membership value. This implies that another pass is needed. Based on the proximity of the membership values to one another, a triangular based

Table 3.3: The Whitenisation of *NN*

J_x	γ	*Weight Function* $[0, 0, 0.43]$	*Error* e
0.00	1.0	1	0
0.14	0.67	0.6744	0.0044
0.29	0.33	0.3256	0.0044
0.43	0	0	0
0.57	0	0	0
0.71	0	0	0
0.86	0	0	0
1.00	0	0	0

whitenisation function will not be able to effectively encapsulate the membership values according to their associated returned degrees of significance. Therefore, the seventh column indicates that a trapezoidal membership function be used with the parameters set at $[0, 0.43, 0.54, 0.71]$. With these new values the returned errors are below the error threshold of $e = 0.01$, making this configuration the final configuration. The continuous streamlined representation of *AC* is presented in Figure 3.10.

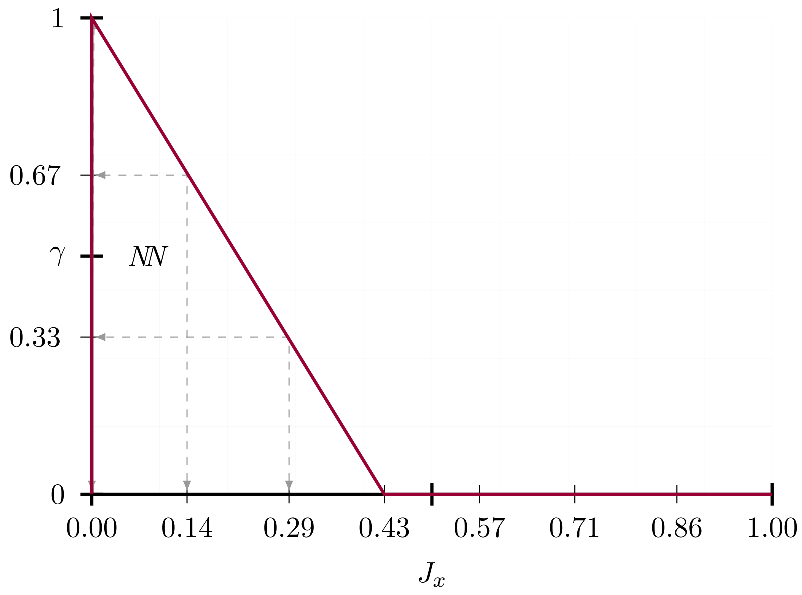


Figure 3.11: A Continuous Streamlined Representation for *NN*

Table 3.3 presents the whitenisation results for the R-fuzzy set *NN*, as it can be seen the first initial and only configuration was satisfactory, none of the returned degrees of significance exceeded the threshold value $e = 0.01$. The configuration of $[0, 0, 0.43]$ also implies that it was the triangular variation of the whitenisation weight function. As the triangular representation has one less parameter value compared to that of the trapezoidal function, of which a variation was given for the original encapsulated set Figure 3.4, the final configuration is indeed more streamlined than the arbitrary setup. Figure 3.11 provides one with a visualisation of the streamlined encapsulating set for the *NN* R-fuzzy set.

Table 3.4 presents the findings for the R-fuzzy set *BN*, unlike that of *NN*, several iterations were needed to decide upon a final configuration. The initial state was indicative of a triangular membership function with the parameters $[0.43, 0.71, 1.00]$, however, this incorrectly applied the corresponding degrees of significance for the fuzzy membership values 0.57 and 0.86, although as its apex was at 0.71, this did return the correct value for the membership value 0.71. The final configuration for the streamlined encapsulating set *BN* can be seen in Figure 3.12

Table 3.5 presents the findings for the R-fuzzy set *VN*, much like that of *AC* and *BN*, several iterations were needed before a final configuration satisfied the error threshold. In addition, much like all the generated streamlined functions, with the exception of *NN*, they all make use of a trapezoidal function for encapsulation. The final configuration for the streamlined encapsulating set *VN* can be seen in Figure 3.13

Simply put, if a parameter value for the whitenisation function is not allowing for the correct expected response, increase it, whether it be the left or right anchor points. If the use of a triangular whitenisation function does not unequivocally encapsulate the returned degrees of significance with the correct intersections, then make use of a trapezoidal function instead, and repeat the process. As was the case with the original work, using the grey whitenisation weight functions has allowed for a continuous representation, where the degrees of significance are also the degrees of membership, akin to a fuzzy perspective.

A continuous visualisation of Example 3 using the grey whitenisation method can be seen in Figure 3.14. Comparing this new plot to that of Figure 3.9, one can see that the

Table 3.4: The Whitenisation of *BN*

J_x	γ	*Weight Function* $[0.43, 0.71, 1.00]$	*Error* e	*Weight Function* $[0.54, 0.71, 1.00]$	*Error* e	*Weight Function* $[0.54, 0.71, 0.80, 0.98]$	*Error* e
0.00	0	0	0.0000	0	0.0000	0	0.0000
0.14	0	0	0.0000	0	0.0000	0	0.0000
0.29	0	0	0.0000	0	0.0000	0	0.0000
0.43	0	0	0.0000	0	0.0000	0	0.0000
0.57	0.17	0.5	0.3300	0.1765	0.0065	0.1765	0.0065
0.71	1	1	0.0000	1	0.0000	1	0.0000
0.86	0.67	0.4828	0.1872	0.4828	0.1872	0.6667	0.0033
1.00	0	0	0.0000	0	0.0000	0	0.0000

membership functions given by the grey whitenisation functions, are smoother and use far less parameters in their construction as compared to the arbitrary values that were used in the original plot, all the while maintaining a high resolution of accuracy.

The use of grey whitenisation weight functions allows for a more simplistic membership set using the minimal number of parameter points. It can indeed be argued that in fact the whitenisation function in this context is simply an equivalent fuzzy triangular, or trapezoidal membership function. In which case, the same procedural steps can be undertaken

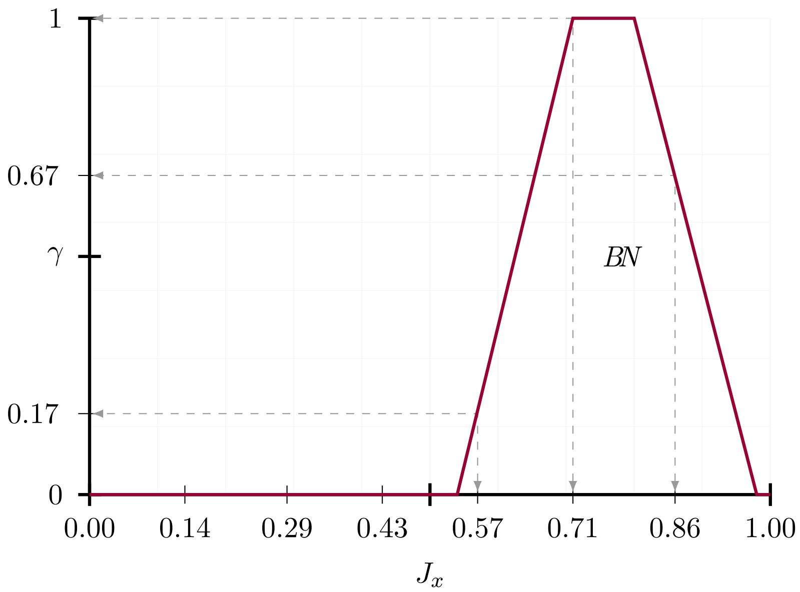


Figure 3.12: A Continuous Streamlined Representation for *BN*

Table 3.5: The Whitenisation of *VN*

J_x	γ	*Weight Function* $[0.71, 1.00, 1.00]$	*Error* e	*Weight Function* $[0.78, 1.00, 1.00]$	*Error* e	*Weight Function* $[0.81, 0.96, 1.00, 1.00]$	*Error* e
0.00	0	0	0.0000	0	0.0000	0	0.0000
0.14	0	0	0.0000	0	0.0000	0	0.0000
0.29	0	0	0.0000	0	0.0000	0	0.0000
0.43	0	0	0.0000	0	0.0000	0	0.0000
0.57	0	0	0.0000	0	0.0065	0	0.0000
0.71	0	0	0.0000	0	0.0000	0	0.0000
0.86	0.33	0.5172	0.1872	0.3636	0.0336	0.3333	0.0033
1.00	1.00	1	0.0000	1	0.0000	1	0.0000

to create a streamlined function that correctly encapsulates all fuzzy membership values. However, if the significance measure degrees for the computed R-fuzzy sets indicate that the sets themselves are not defined as triangular or trapezoidal, then this approach will not be ideally suited. This approach is more favourable with regards to perceptions that follow a general rule of thumb, whereby the expected assumptions are the returned perceptions, with a slight amount of variance.

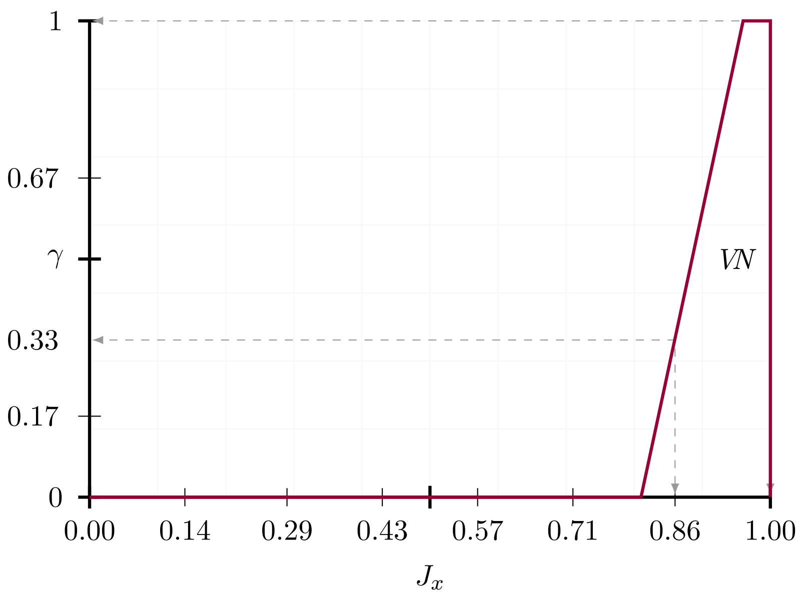


Figure 3.13: A Continuous Streamlined Representation for *VN*

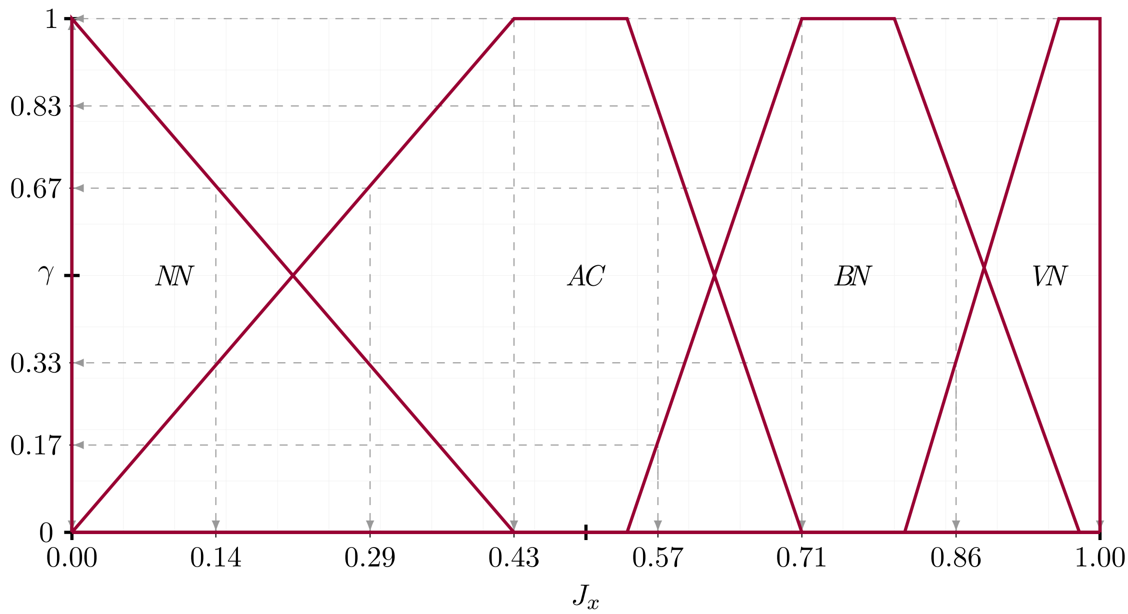


Figure 3.14: A Continuous Streamlined Visualisation for *NN*, *AC*, *BN* & *VN*

3.4 The Enhanced R-fuzzy Framework

In this section the newly proposed framework for the quantification of perception based uncertainty is described and demonstrated. The work contained in Khuman et al. (2016c) first proposed the notion of using an R-fuzzy and significance measure pairing, in addition to grey analysis. The contribution of the significance measure allows for a link, so that R-fuzzy sets and grey analysis can be integrated into the same framework. For it is the significance measure that provides the sequences that allow for the absolute degree of grey incidence to be undertaken.

Using the concept of R-fuzzy sets, one is able to encapsulate a broader sense of uncertainty, catering for both general and specific held subjective assumptions for any given abstract concept. As it has been stated in Section 2.4.8, the various alternative approaches such as; interval-valued fuzzy sets, Atanassov intuitionistic fuzzy sets, shadowed sets and type-2 fuzzy sets, all suffer from certain restrictions. Often an object value will lose its uniqueness if contained within a set or shadowed region. An R-fuzzy approach can absolutely ensure that each and every *viable* uncertain fuzzy membership value contained within its rough set membership, will be distinguishable from other encapsulated values. This guarantee of no loss of information regarding encapsulation was a major advantage over other alternative approaches. The introduction of the significance measure allows for the quantification of contained membership values, and equally, the validation of values not contained. Therefore, an additional level of detail can be garnered, all the while still maintaining the uniqueness that a traditional R-fuzzy set allows for.

The introduction of grey theory for the analysis can only be possible if the absolute degree of grey incidence is used on a sequence of data, indicative of a time series. If one was to simply use an R-fuzzy set, the objects of said set would just be objects and would not be able to be passed to the absolute degree of grey incidence, as it would not constitute a sequence. With the addition of the significance measure, one is able to translate the contained fuzzy membership values along with their correlated degrees of significance, such that it becomes a sequence. The discretised points along the x-axis in this instance are the contained membership values of J_x. The membership set of J_x will always be the same for each generated R-fuzzy set. Therefore, if the data set contains clusters, one is able to generate R-fuzzy subsets with the significance measure, such that it can be passed to the absolute degree of grey incidence. After which, the metric divergences of comparable sequences of computed R-fuzzy subsets can be quantified and collected.

As Example 4 will demonstrate, the enhanced R-fuzzy framework allows for one to garner a great deal more information. Regardless of the amount of variance in the collected perceptions, the sequences generated from the R-fuzzy and significance measure pairing, retain the subjective nature of the populous. Nothing is diluted nor ignored, each and every contained perception can be captured and maintained. It is precisely this aspect of the standard R-fuzzy approach that made it a standout candidate from the various uncertainty models that were surveyed. All subjectivity is collected from which an objective approach can then be taken. The subjective nature of perception based uncertainty provides the foundation from which to extract meaningful sentiment from. If such a foundation did not accurately convey what it was collecting, then any inference from it would not be a true reflection. An R-fuzzy approach ensures that no loss of information is suffered, as a result, the objective nature of the analysis makes direct reference to a truthful representation of the perceptions collected.

3.4.1 An Enhanced R-fuzzy Example

EXAMPLE 4: Assume that $F = \{f_1, f_2, \ldots, f_9\}$ is a set containing 9 different colour swatches based on the color red:

$$f_1 \rightarrow \quad [204, 0, 0] \rightarrow$$

$$f_2 \rightarrow \quad [153, 0, 0] \rightarrow$$

$$f_3 \rightarrow \quad [255, 102, 102] \rightarrow$$

$$f_4 \rightarrow \quad [51, 0, 0] \rightarrow$$

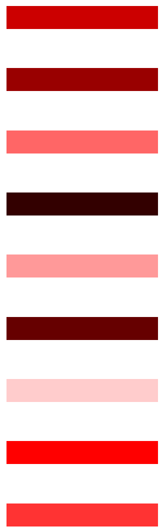
$$f_5 \rightarrow \quad [255, 153, 153] \rightarrow$$

$$f_6 \rightarrow \quad [102, 0, 0] \rightarrow$$

$$f_7 \rightarrow \quad [255, 204, 204] \rightarrow$$

$$f_8 \rightarrow \quad [255, 0, 0] \rightarrow$$

$$f_9 \rightarrow \quad [255, 51, 51] \rightarrow$$



The colours themselves are given by their [RGB] values, from which the average is calculated and stored in $N = \{68, 51, 153, 17, 187, 34, 221, 85, 119\}$. Each average N_i value will correspond to a specific colour swatch F_i. For example, the swatch associated with f_3 has a value of 153, f_5 will be related to 187, and so on. Given that the criteria set $C = \{p_1, p_2, \ldots, p_{15}\}$ contains the perceptions of 15 individuals, all of whom gave their own opinions based on the available descriptors and the swatches themselves. These values have been collected, along with their ages and are presented in Table 3.6.

The descriptor terms contained within the table can be understood as meaning:

$DR \rightarrow$ Dark Red $\qquad R \rightarrow$ Red $\qquad LR \rightarrow$ Light Red

Using the same linear function given in Eq. (2.5.26), which has been used for all the contained examples, the resulting fuzzy membership set is given as:

$$J_x = \{0.25, 0.17, 0.67, 0.00, 0.83, 0.08, 1.00, 0.33, 0.50\}$$

Table 3.6: Human Perception Based on the Variations for the Colour Red

#	*Age*	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
p_1	20	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_2	30	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*
p_3	20	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_4	25	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_5	25	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_6	20	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_7	20	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_8	25	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*
p_9	25	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*
p_{10}	30	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*
p_{11}	20	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_{12}	25	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*LR*
p_{13}	30	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*
p_{14}	30	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*
p_{15}	30	*R*	*DR*	*LR*	*DR*	*LR*	*DR*	*LR*	*R*	*R*

Using DEFINITION 8, the final generated R-fuzzy sets based on the collected subsets for *LR*, *R* and *DR*, respectively, are given as:

$$DR = \Big(\{0.00, 0.08, 0.17\}, \{0.00, 0.08, 0.17\}\Big)$$

$$R = \Big(\{0.25, 0.33\}, \{0.25, 0.33, 0.50\}\Big)$$

$$LR = \Big(\{0.67, 0.83, 1.00\}, \{0.50, 0.67, 0.83, 1.00\}\Big)$$

What's striking when considering the perceptions contained within the table is the level of exactness. For all observers from $p_1, \ldots p_{15}$, for colour swatches $f_1, \ldots f8$, the exact same level of responses were recorded. The colour swatch which divided opinion was that of f_9. As is the subjective nature of perception based uncertainty, one cannot simply *expect* certain results. The expectant volatility was simply not there, however, what will be shown, even from what little variance that does exist, a wealth of understanding can

Table 3.7: The Degrees of Significance for Each of the Generated R-fuzzy Sets

DR		*R*		*LR*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DR}}\{0.00\} =$	1.00	$\gamma_{\overline{R}}\{0.00\} =$	0.00	$\gamma_{\overline{LR}}\{0.00\} =$	0.00
$\gamma_{\overline{DR}}\{0.08\} =$	1.00	$\gamma_{\overline{R}}\{0.08\} =$	0.00	$\gamma_{\overline{LR}}\{0.08\} =$	0.00
$\gamma_{\overline{DR}}\{0.17\} =$	1.00	$\gamma_{\overline{R}}\{0.17\} =$	0.00	$\gamma_{\overline{LR}}\{0.17\} =$	0.00
$\gamma_{\overline{DR}}\{0.25\} =$	0.00	$\gamma_{\overline{R}}\{0.25\} =$	1.00	$\gamma_{\overline{LR}}\{0.25\} =$	0.00
$\gamma_{\overline{DR}}\{0.33\} =$	0.00	$\gamma_{\overline{R}}\{0.33\} =$	1.00	$\gamma_{\overline{LR}}\{0.33\} =$	0.00
$\gamma_{\overline{DR}}\{0.50\} =$	0.00	$\gamma_{\overline{R}}\{0.50\} =$	0.47	$\gamma_{\overline{LR}}\{0.50\} =$	0.53
$\gamma_{\overline{DR}}\{0.67\} =$	0.00	$\gamma_{\overline{R}}\{0.67\} =$	0.00	$\gamma_{\overline{LR}}\{0.67\} =$	1.00
$\gamma_{\overline{DR}}\{0.83\} =$	0.00	$\gamma_{\overline{R}}\{0.83\} =$	0.00	$\gamma_{\overline{DR}}\{0.83\} =$	1.00
$\gamma_{\overline{DR}}\{1.00\} =$	0.00	$\gamma_{\overline{R}}\{1.00\} =$	0.00	$\gamma_{\overline{LR}}\{1.00\} =$	1.00

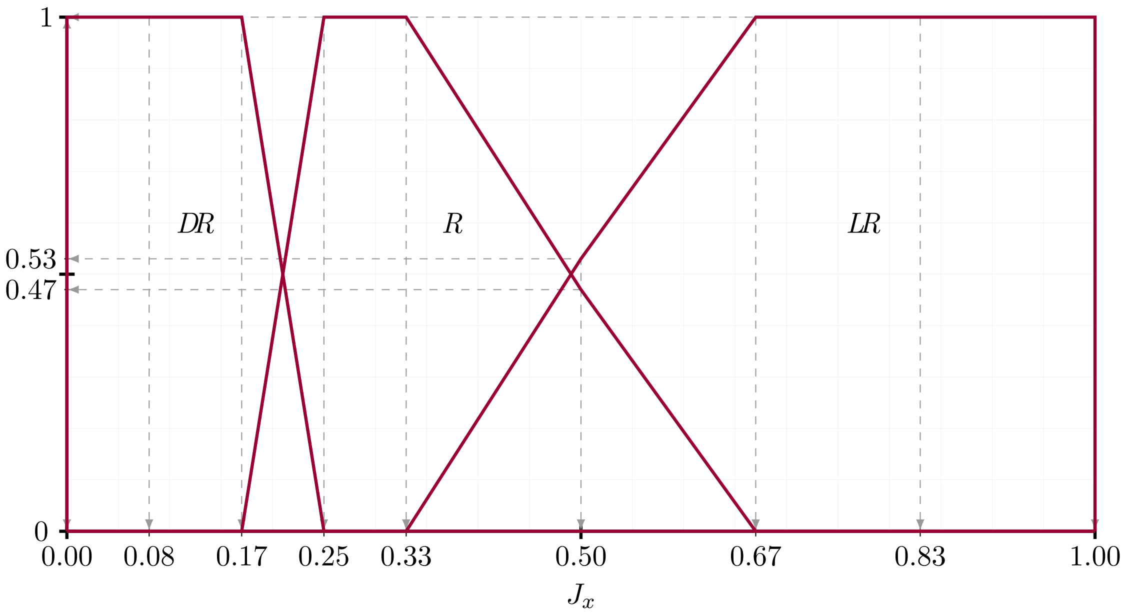


Figure 3.15: A Continuous Visualisation for *DR*, *R* & *LR*

be garnered. By using Eq. (3.2.2), one is able to calculate the degree of significance for each and every encapsulated fuzzy membership value from J_x. The returned degree of significance for all generated R-fuzzy sets are presented in Table 3.7.

Figure 3.15 collectively displays all the generated R-fuzzy sets for EXAMPLE 4, along with the degree of significance for each fuzzy membership value, in accordance to its relative association to each R-fuzzy set. Referring back to Table 3.6, one will see that the age of each individual was also collected. There are three ages of note: 20, 25 and 30, 5 from each criterion. As these are all from the same original criteria set C, the membership set J_x will not change, this is guaranteed, one does not need to recalculate. The creation of R-fuzzy subsets can now be undertaken, such that they represent the three individual age groups collected. From this they can be compared with regards to the perspective of each cluster to that of another, gaining a more detailed understanding of the concept being modelled. Furthermore, as the membership set J_x does indeed remain the same, this can be used as the sequence needed for the absolute degree of grey incidence component.

Table 3.8: The Degrees of Significance for Each of the Generated R-fuzzy Sets for the Age Cluster 20 Year Olds

DR		*R*		*LR*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DR}}\{0.00\} =$	1.00	$\gamma_{\overline{R}}\{0.00\} =$	0.00	$\gamma_{\overline{LR}}\{0.00\} =$	0.00
$\gamma_{\overline{DR}}\{0.08\} =$	1.00	$\gamma_{\overline{R}}\{0.08\} =$	0.00	$\gamma_{\overline{LR}}\{0.08\} =$	0.00
$\gamma_{\overline{DR}}\{0.17\} =$	1.00	$\gamma_{\overline{R}}\{0.17\} =$	0.00	$\gamma_{\overline{LR}}\{0.17\} =$	0.00
$\gamma_{\overline{DR}}\{0.25\} =$	0.00	$\gamma_{\overline{R}}\{0.25\} =$	1.00	$\gamma_{\overline{LR}}\{0.25\} =$	0.00
$\gamma_{\overline{DR}}\{0.33\} =$	0.00	$\gamma_{\overline{R}}\{0.33\} =$	1.00	$\gamma_{\overline{LR}}\{0.33\} =$	0.00
$\gamma_{\overline{DR}}\{0.50\} =$	0.00	$\gamma_{\overline{R}}\{0.50\} =$	0.00	$\gamma_{\overline{LR}}\{0.50\} =$	1.00
$\gamma_{\overline{DR}}\{0.67\} =$	0.00	$\gamma_{\overline{R}}\{0.67\} =$	0.00	$\gamma_{\overline{LR}}\{0.67\} =$	1.00
$\gamma_{\overline{DR}}\{0.83\} =$	0.00	$\gamma_{\overline{R}}\{0.83\} =$	0.00	$\gamma_{\overline{LR}}\{0.83\} =$	1.00
$\gamma_{\overline{DR}}\{1.00\} =$	0.00	$\gamma_{\overline{R}}\{1.00\} =$	0.00	$\gamma_{\overline{LR}}\{1.00\} =$	1.00

The membership values themselves act as the discretised points along the x-axis, whereas, the *varying* significance degrees give the associated amplitude. For example, if one refers to Table 3.8 which contains the data for the age cluster 20 year olds, the membership value 0.50 for the R-fuzzy set *LR*, has a returned degree of significance of 1.00. The same membership value and R-fuzzy set for the age cluster 25 year olds in Table 3.9, returns a degree of significance of 0.60. For the age cluster 30 year olds in Table 3.10, the returned degree of significance for the same fuzzy membership value is 0.00.

Table 3.9: The Degrees of Significance for Each of the Generated R-fuzzy Sets for the Age Cluster 25 Year Olds

DR		*R*		*LR*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DR}}\{0.00\} =$	1.00	$\gamma_{\overline{R}}\{0.00\} =$	0.00	$\gamma_{\overline{LR}}\{0.00\} =$	0.00
$\gamma_{\overline{DR}}\{0.08\} =$	1.00	$\gamma_{\overline{R}}\{0.08\} =$	0.00	$\gamma_{\overline{LR}}\{0.08\} =$	0.00
$\gamma_{\overline{DR}}\{0.17\} =$	1.00	$\gamma_{\overline{R}}\{0.17\} =$	0.00	$\gamma_{\overline{LR}}\{0.17\} =$	0.00
$\gamma_{\overline{DR}}\{0.25\} =$	0.00	$\gamma_{\overline{R}}\{0.25\} =$	1.00	$\gamma_{\overline{LR}}\{0.25\} =$	0.00
$\gamma_{\overline{DR}}\{0.33\} =$	0.00	$\gamma_{\overline{R}}\{0.33\} =$	1.00	$\gamma_{\overline{LR}}\{0.33\} =$	0.00
$\gamma_{\overline{DR}}\{0.50\} =$	0.00	$\gamma_{\overline{R}}\{0.50\} =$	0.40	$\gamma_{\overline{LR}}\{0.50\} =$	0.60
$\gamma_{\overline{DR}}\{0.67\} =$	0.00	$\gamma_{\overline{R}}\{0.67\} =$	0.00	$\gamma_{\overline{LR}}\{0.67\} =$	1.00
$\gamma_{\overline{DR}}\{0.83\} =$	0.00	$\gamma_{\overline{R}}\{0.83\} =$	0.00	$\gamma_{\overline{LR}}\{0.83\} =$	1.00
$\gamma_{\overline{DR}}\{1.00\} =$	0.00	$\gamma_{\overline{R}}\{1.00\} =$	0.00	$\gamma_{\overline{LR}}\{1.00\} =$	1.00

Regardless of how small or large the difference between the returned degrees of significance for comparable fuzzy membership values, the fact that there can be a difference should provide one the motivation to explore further. It is precisely this aspect of wanting to investigate that warrants the use of grey analysis. Given the tiny amount of variance that exists in Example 4, one would be forgiven to assume that not a lot of further information could be garnered. However, as it will now be demonstrated when applying the notion grey analysis, there is yet a wealth of understanding still to be obtained.

Given the fuzzy membership values of J_x and the correlated degrees of significance, indicative of an R-fuzzy set, one can now make efforts to quantify the differences between the metric spaces of comparable sequences. Figure 3.16 provides a visualisation of the comparison between the R-fuzzy set and significance measure sequence generated for *LR*,

Table 3.10: The Degrees of Significance for Each of the Generated R-fuzzy Sets for the Age Cluster 30 Year Olds

DR		*R*		*LR*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DR}}\{0.00\} =$	1.00	$\gamma_{\overline{R}}\{0.00\} =$	0.00	$\gamma_{\overline{LR}}\{0.00\} =$	0.00
$\gamma_{\overline{DR}}\{0.08\} =$	1.00	$\gamma_{\overline{R}}\{0.08\} =$	0.00	$\gamma_{\overline{LR}}\{0.08\} =$	0.00
$\gamma_{\overline{DR}}\{0.17\} =$	1.00	$\gamma_{\overline{R}}\{0.17\} =$	0.00	$\gamma_{\overline{LR}}\{0.17\} =$	0.00
$\gamma_{\overline{DR}}\{0.25\} =$	0.00	$\gamma_{\overline{R}}\{0.25\} =$	1.00	$\gamma_{\overline{LR}}\{0.25\} =$	0.00
$\gamma_{\overline{DR}}\{0.33\} =$	0.00	$\gamma_{\overline{R}}\{0.33\} =$	1.00	$\gamma_{\overline{LR}}\{0.33\} =$	0.00
$\gamma_{\overline{DR}}\{0.50\} =$	0.00	$\gamma_{\overline{R}}\{0.50\} =$	1.00	$\gamma_{\overline{LR}}\{0.50\} =$	0.00
$\gamma_{\overline{DR}}\{0.67\} =$	0.00	$\gamma_{\overline{R}}\{0.67\} =$	0.00	$\gamma_{\overline{LR}}\{0.67\} =$	1.00
$\gamma_{\overline{DR}}\{0.83\} =$	0.00	$\gamma_{\overline{R}}\{0.83\} =$	0.00	$\gamma_{\overline{LR}}\{0.83\} =$	1.00
$\gamma_{\overline{DR}}\{1.00\} =$	0.00	$\gamma_{\overline{R}}\{1.00\} =$	0.00	$\gamma_{\overline{LR}}\{1.00\} =$	1.00

with relation to the age cluster 20 year olds from Table 3.8, against the R-fuzzy set and significance measure sequence generated for *LR*, with relation to the age cluster 25 year

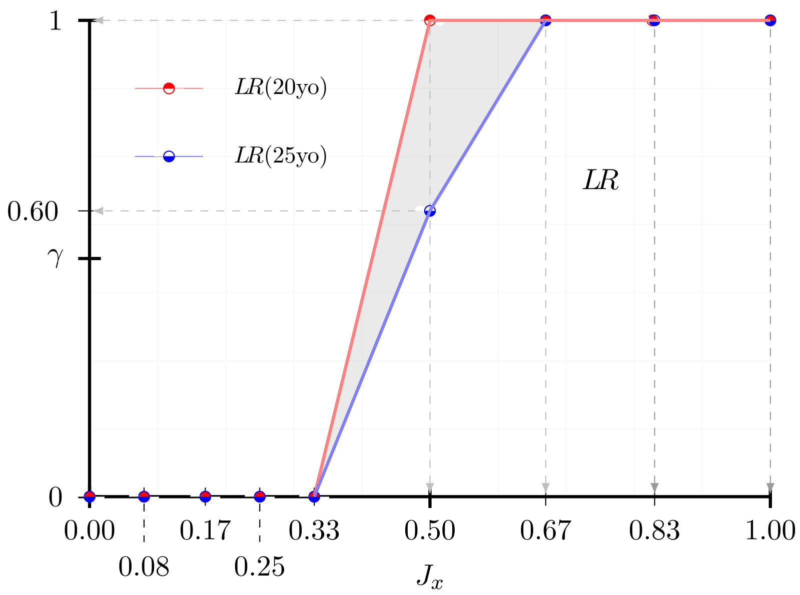


Figure 3.16: The Comparability Between 20 and 25 Year Olds for *LR*

olds from Table 3.9. The shaded area of the plot signifies the difference in the metric spaces between the two sequences.

The greater the difference between sequences the smaller the returned absolute degree of grey incidence, equally, the more similar the sequences, the greater the value. Using Definition 17 one can now apply the absolute degree of grey incidence to quantify the overall difference between the two sequences based on their metric spaces. As the degree of grey incidence uses *absolute* values, the order in which the sequences are passed through will not impact on the result. The sequence associated with LR for 20 year olds is given as: $s_i = \{0.00, 0.00, 0.00, 0.00, 0.00, 1.00, 1.00, 1.00, 1.00\}$. The sequence associated with LR for 25 year olds is given as: $s_j = \{0.00, 0.00, 0.00, 0.00, 0.00, 0.60, 1.00, 1.00, 1.00\}$. Using Eq. (2.6.33), the returned result is $\epsilon(0.950)$. A very high scoring value, indicating that the two sequences are indeed very similar. Referring back to Figure 3.16, it can quite easily be inferred that the similarities are indeed there, the sequences score almost exactly the same for their returned degrees of significance, except for one point. The fuzzy membership value of $J_x\{0.50\}$ differs between the two R-fuzzy sets for the two different age clusters. Comparing the R-fuzzy sets and associated significance measure sequences with that of differing age clusters, one can quantify the rate of change in perception as one propagates through successive clusters.

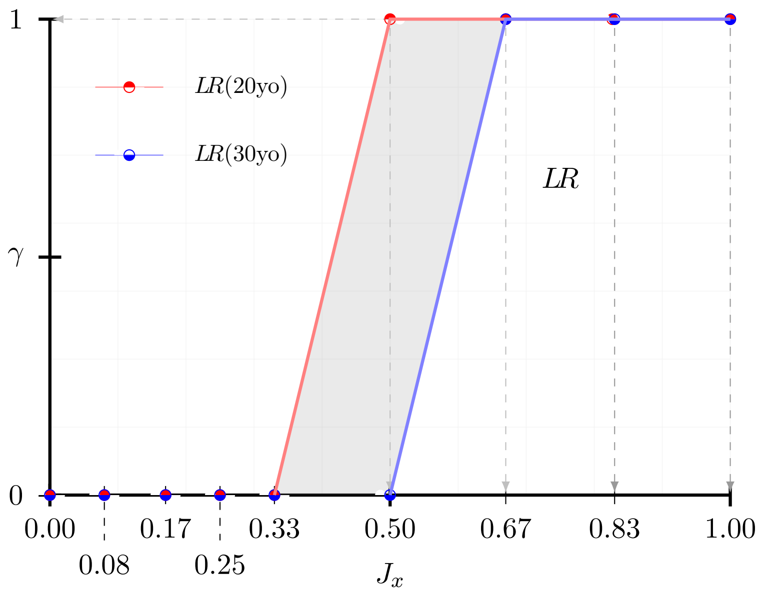


Figure 3.17: The Comparability Between 20 and 30 Year Olds for LR

Figure 3.17 provides a visualisation of the comparison between the R-fuzzy set and significance measure sequence generated for LR, with relation to age cluster for 20 year olds, against the R-fuzzy set and significance measure sequence generated for LR, with relation to the age cluster for 30 year olds from Table 3.10. In exactly the same way, the sequences are compared using Eq. (2.6.33). In this case, the sequence associated with LR for 30 year olds is given as: $s_j = \{0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 1.00, 1.00, 1.00\}$, and the returned absolute degree of grey incidence is given as $\epsilon(0.875)$. Another high scoring value but not as high as the comparison between the R-fuzzy set and significance measure sequences computed for 20 and 25 year olds. Simply inspecting the plot, one can see that the area of divergence is greater than the comparison between 20 and 25 year olds.

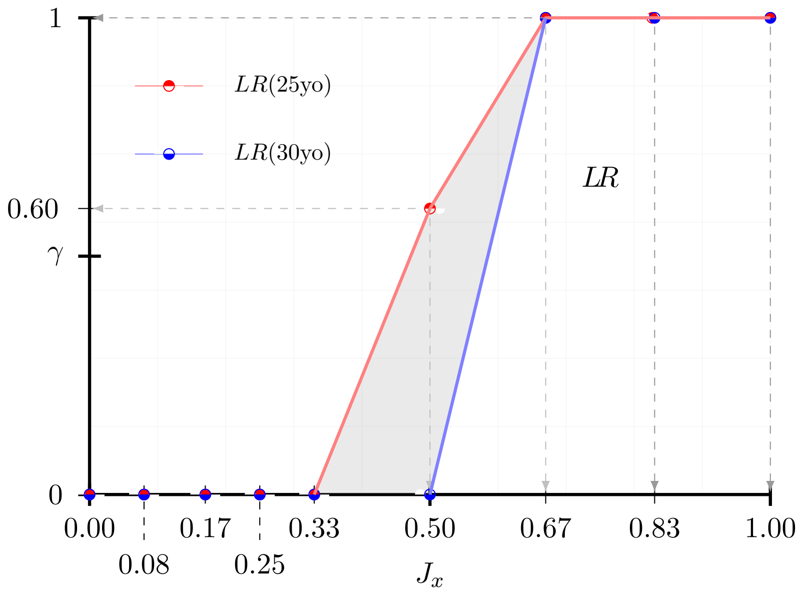


Figure 3.18: The Comparability Between 25 and 30 Year Olds for LR

This comparing and contrasting is repeated for all clusters against all generated R-fuzzy sets and significance measure sequences. Table 3.11 provides a summary of the collected absolute degree of grey incidence values for all comparisons, against each of the three age clusters. One can see that the age clusters of 20 and 25 share the closest correlated perception for the concept LR, with an absolute degree of grey incidence value of $\epsilon(0.950)$. The age clusters of 20 and 30 share the least correlated perception with an absolute degree of grey incidence value of $\epsilon(0.875)$. The age clusters of 25 and 30 share an intermediary correlation with a returned value of $\epsilon(0.916)$, the visualisation can be seen in Figure. 3.18.

Table 3.11: A Comparable Summary of the Returned Absolute Degree of Grey Incidence for *LR*, *R* & *DR*

LR	20yo	25yo	30yo	*R*	20yo	25yo	30yo	*DR*	20yo	25yo	30yo
20yo	$\epsilon(1.00)$	$\epsilon(0.950)$	$\epsilon(0.875)$	20yo	$\epsilon(1.00)$	$\epsilon(0.931)$	$\epsilon(0.857)$	20yo	$\epsilon(1.00)$	$\epsilon(1.00)$	$\epsilon(1.00)$
25yo	-	$\epsilon(1.00)$	$\epsilon(0.916)$	25yo	-	$\epsilon(1.00)$	$\epsilon(0.914)$	25yo	-	$\epsilon(1.00)$	$\epsilon(1.00)$
30yo	-	-	$\epsilon(1.00)$	30yo	-	-	$\epsilon(1.00)$	30yo	-	-	$\epsilon(1.00)$

Applying the same level of inspection, one can see that the correlation for each *R* significance measure sequence, between the age clusters of 20 and 25 year olds, is again stronger than that of 20 and 30 year olds. With regards to the *DR*, it is logical to infer that all clusters correlated exactly with one another. Simply inspecting Tables 3.8, 3.9 and 3.10, one can see that each significance measure sequence is absolutely the same, meaning, for the notion of *DR*, all age clusters agreed upon the same fuzzy membership values. The Figures 3.19, 3.20, 3.21, 3.22, 3.23 and 3.24 provide the visualisations for the remaining combinatorial comparisons, between R-fuzzy sets and the significance measure sequences, to that of varying age clusters.

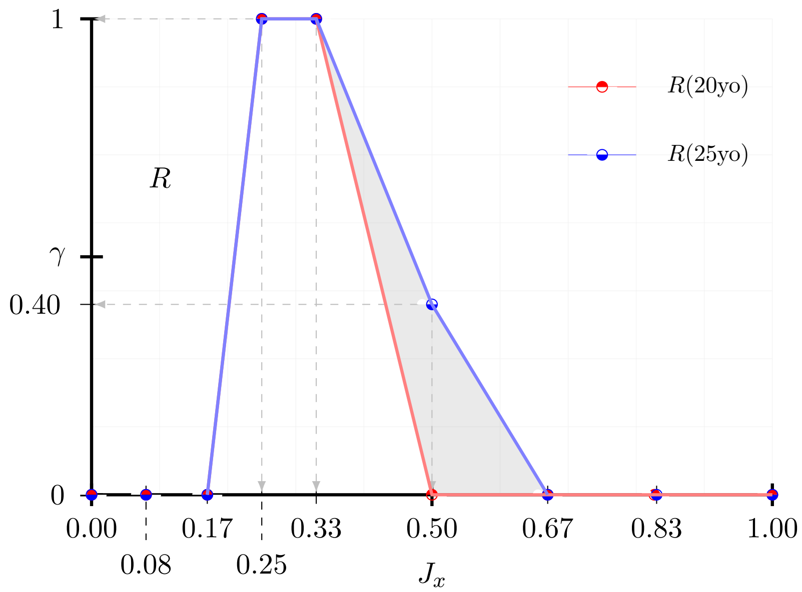


Figure 3.19: The Comparability Between 20 and 25 Year Olds for *R*

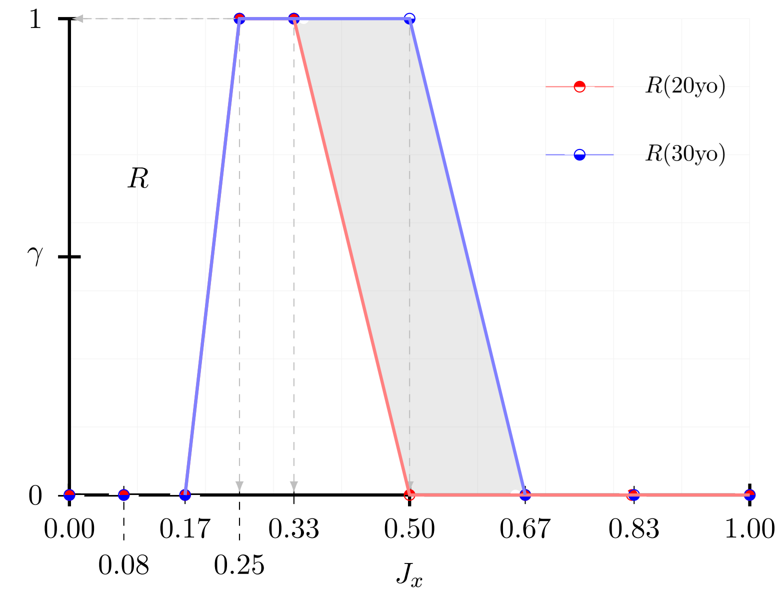


Figure 3.20: The Comparability Between 20 and 30 Year Olds for R

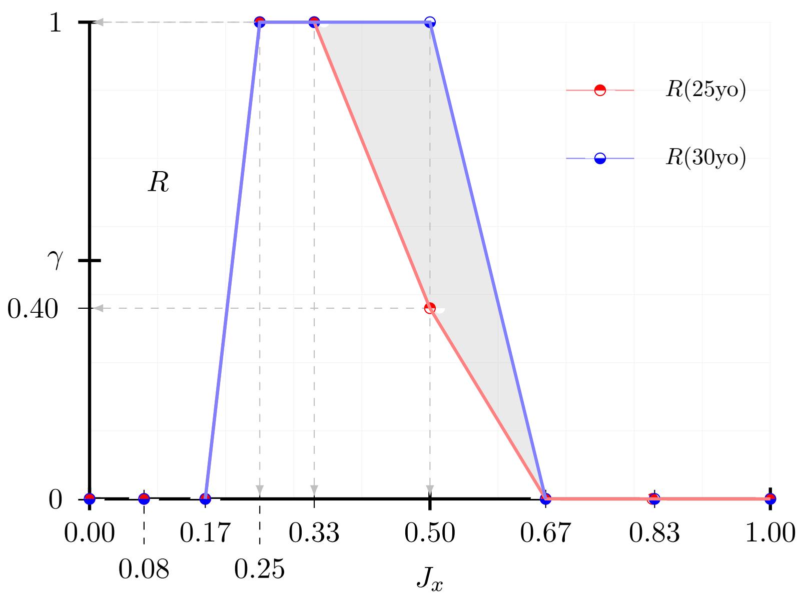


Figure 3.21: The Comparability Between 25 and 30 Year Olds for R

3.5 Similarity and Distance Measures

The enhanced R-fuzzy approach makes use of sequences for the comparison of computed R-fuzzy subsets. The generated sequences related to this thesis are indicative of the returned degrees of significance, for all membership values belonging to the fuzzy membership set J_x, which is the universe of discourse. However, there are indeed similarities

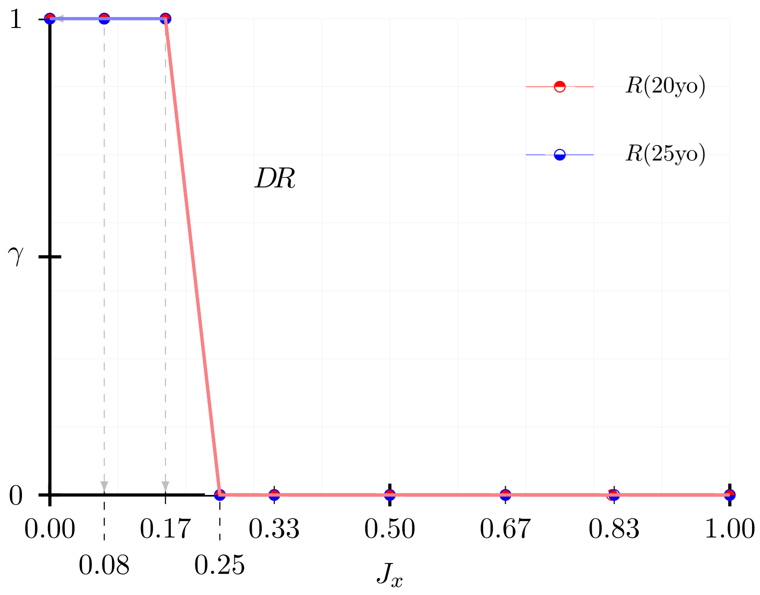


Figure 3.22: The Comparability Between 20 and 25 Year Olds for *DR*

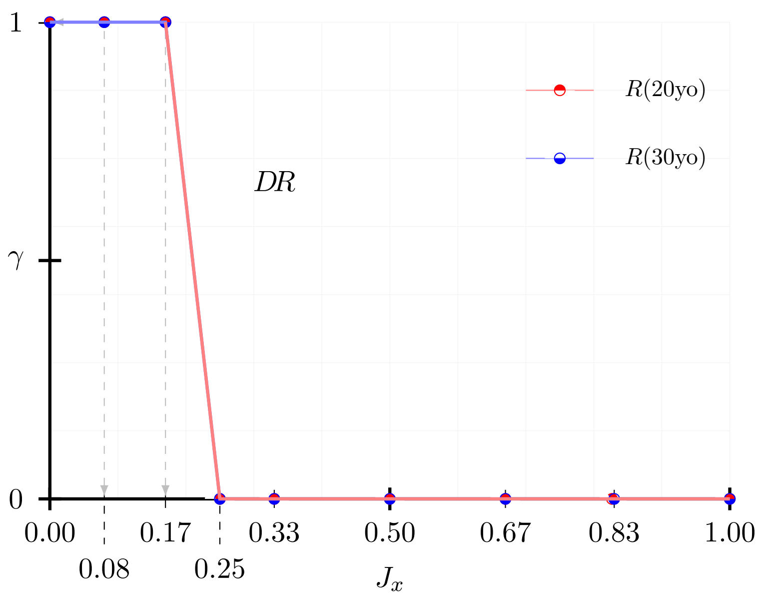


Figure 3.23: The Comparability Between 20 and 30 Year Olds for *DR*

that exist with the sequences computed in relation to EXAMPLE 4, to that of what one would consider a type-1 fuzzy set. In SECTION 4.7.1, the relationship between that of the significance measure when used in conjunction with an R-fuzzy set, to that of a type-1 fuzzy set is described. In which case, one would be able to treat two R-fuzzy sets accompanied with two significance measure sequences, as two type-1 fuzzy sets, and provide a similarity index.

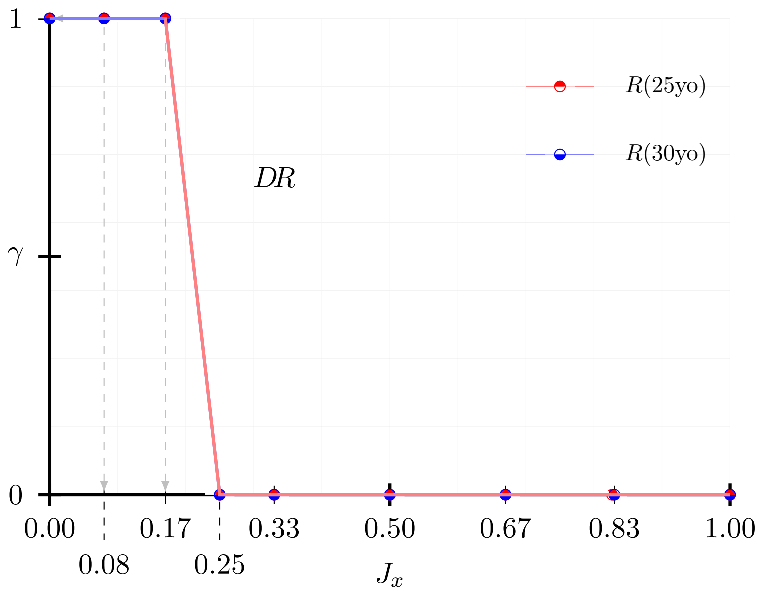


Figure 3.24: The Comparability Between 25 and 30 Year Olds for *DR*

Many current similarity and distance measures, of which there are many, do not take into consideration the direction of change between comparable fuzzy sets (McCulloch et al., 2014). However, the rate of change in propagating fuzzy sets in the context of perception modelling is not necessarily needed. The similarity measure in the context of perception modelling is far more informative, as it is the similarity of the divergences that are being computed. It can indeed be argued that instead of the absolute degree of grey incidence, one could make use of a fuzzy similarity relation, comparing the returned degrees of significance for each triggered membership value for each R-fuzzy set. Given the functionality of the absolute degree of grey incidence, demonstrated in Example 4, it is indeed a similarity measure. Instead of measuring the degree of membership of a fuzzy set, the degree of significance of each fuzzy membership value is measured and summated. The family that the absolute degree of grey incidence belongs to, has many other variations (Liu and Lin, 2006), some of which where presented in Section 2.6.3. When considering the examples presented in this thesis, it is more favourable to make use of an absolute method, such as the one adopted.

Considering the well known Jaccard similarity measure, given as:

$$s(A,B) = \frac{\sum_{i=1}^{n} \min\Big(\mu A(x_i), \mu B(x_i)\Big)}{\sum_{i=1}^{n} \max\Big(\mu A(x_i), \mu B(x_i)\Big)} \tag{3.5.7}$$

Where A and B are representative of two fuzzy sets, in this case two R-fuzzy sets, where the significance measure degree for an object, is the degree of membership which one would associate to an object of a fuzzy set. Applying the Jaccard similarity measure to EXAMPLE 4 would return the scores contained within Table 3.12. The values returned by the Jaccard measure, generally agree to the sentiment returned by the absolute degree of grey incidence, as the ranking order is maintained. However, inspecting the returned metrics, one will see that the Jaccard method does not have the same resolution as the grey method. What's more, the values although agree with the grey approach, the difference in the values themselves is varied.

Table 3.12: The Returned Jaccard Similarity Index For *LR*, *R* & *DR*

LR	20yo	25yo	30yo	*R*	20yo	25yo	30yo	*DR*	20yo	25yo	30yo
20yo	$\epsilon(1.00)$	$\epsilon(0.900)$	$\epsilon(0.750)$	20yo	$\epsilon(1.00)$	$\epsilon(0.833)$	$\epsilon(0.667)$	20yo	$\epsilon(1.00)$	$\epsilon(1.00)$	$\epsilon(1.00)$
25yo	-	$\epsilon(1.00)$	$\epsilon(0.833)$	25yo	-	$\epsilon(1.00)$	$\epsilon(0.800)$	25yo	-	$\epsilon(1.00)$	$\epsilon(1.00)$
30yo	-	-	$\epsilon(1.00)$	30yo	-	-	$\epsilon(1.00)$	30yo	-	-	$\epsilon(1.00)$

Table 3.13 presents the absolute difference between the Jaccard similarity measure to that of the absolute degree of grey incidence. The values although generally small, are still different throughout. If one was to inspect the plots for EXAMPLE 4 visually, they would probably be inclined to agree with the values returned by the absolute degree of grey incidence, to be a more humanistic interpretation of the results.

Table 3.13: The Absolute Difference Between the Grey and the Jaccard Measure

LR	20yo	25yo	30yo	*R*	20yo	25yo	30yo	*DR*	20yo	25yo	30yo
20yo	$\epsilon(0.00)$	$\epsilon(0.05)$	$\epsilon(0.125)$	20yo	$\epsilon(0.00)$	$\epsilon(0.098)$	$\epsilon(0.190)$	20yo	$\epsilon(0.00)$	$\epsilon(0.00)$	$\epsilon(0.00)$
25yo	-	$\epsilon(0.00)$	$\epsilon(0.083)$	25yo	-	$\epsilon(0.00)$	$\epsilon(0.114)$	25yo	-	$\epsilon(0.00)$	$\epsilon(0.00)$
30yo	-	-	$\epsilon(0.00)$	30yo	-	-	$\epsilon(0.00)$	30yo	-	-	$\epsilon(0.00)$

A distance measure in the context of this thesis would not be an ideal approach to use, as the distance measure provides a value of distance between two *differing* R-fuzzy sets. The premise of this thesis is to compare the geometric patterns generated from like R-fuzzy sets based on varying permutations of clusters. For example, the significance sequence generated for *LR* with relation to 20 year olds, was compared to that of *LR* with relation to 25 year olds. This provided one with a metric that indicated the strength of how a particular cluster for a given abstract notion, compared to that of a different cluster, with regard to the same abstract notion. As the computed R-fuzzy subsets are themselves generated from the same original criteria set, it is completely possible to compare different R-fuzzy sets, such as *LR* to that of *DR*, as they are all based on the same universe of discourse. However, this would not prove to be helpful, as the comparison between different R-fuzzy sets does not provide an insight into the sub-clusters contained within the criteria. Nor are we interested in the shift from one R-fuzzy set to the next, we are simply concerned with the difference of the same R-fuzzy sets for different clusters.

3.6 An R-fuzzy α-Cut

Once the significance measure has been applied to the contained fuzzy membership values of an R-fuzzy set, one could then use what would be equivalent to an α-cut to provide horizontal slices. An α-cut from a fuzzy perspective can be seen as nested sets of non-fuzzy crisp sets, whose degree of memberships within the set are equal to or greater than α. The α-cut is defined as:

$$A_\alpha = \left\{ x \;\middle|\; \mu_A(x) \geq \alpha \right\} \tag{3.6.8}$$

As the intended α-cut will be applied on an R-fuzzy and significance pairing, the α-cut can instead be defined as the following, using the same notation that defined the significance measure:

$$\bar{A}_\alpha = \left\{ v \;\middle|\; \gamma_{\bar{A}}\{v\} \geq \alpha \right\} \tag{3.6.9}$$

Where the α-cut of the R-fuzzy set $\bar{A}$ is comprised of the possible degrees of significance that can be computed given the criteria set. Assuming that the criteria set C is of a fixed size for all generated R-fuzzy sets associated to it, a change in C would impact on successively computed R-fuzzy sets. Therefore, given that the same number of observers in C remains the same, this will ensure that the same increments of significance can be adhered to. Referring back to Examples 1, 2 and 3, which were all based on the same original criteria set, all involving the use of 6 individuals. With that being the case, only the following possible degrees of significance can be registered:

$$\frac{0}{6} = \gamma\{0\},\ \frac{1}{6} = \gamma\{0.17\},\ \frac{2}{6} = \gamma\{0.33\},\ \frac{3}{6} = \gamma\{0.50\},\ \frac{4}{6} = \gamma\{0.67\},\ \frac{5}{6} = \gamma\{0.83\},\ \frac{6}{6} = \gamma\{1\}$$

As the generated subsets that constitute an R-fuzzy are indicative of the cardinality of the number of observers contained in C, so too is the expected possible degrees of significance that can be produced. With the exception of 0.50, all the possible values were triggered in some way.

Consider the extremes 0 and 1 for the returned degrees of significance, to obtain a significance of 1, all in the populace must agree, therefore, one should take the possible degree of significance of 0.83, as the extreme for the highest grade before absolute inclusion. Also, one should consider 0.17 as the other extreme, as a value of 0 indicates absolute non-inclusion. Therefore, 0.17 can be seen as the lowest possible grade before exclusion from the R-fuzzy set. Given the possible instances of significance, one can create an α-cut with a horizontal slice at each instance. In much the same way that each α-cut signifies the strength of significance, one could be interested in the extremes, 0.17 and 0.83. As the membership of an R-fuzzy set is a rough set with an accompanying lower and upper approximation, the membership values associated to the lowest extreme α-cut could be resigned to non-inclusion. Whereas the membership values assigned to the highest extreme α-cut could be elevated to the lower approximation, which is an absolute agreement and inclusion.

Figure 3.25 shows the highest and lowest extreme α-cuts at 0.83 and 0.17, respectively, with regards to EXAMPLE 2. The plot itself indicates that only the membership value of 0.57 scored a returned degree of significance of 0.83 with relation to the R-fuzzy set *AC*, and a degree of significance of 0.17 with relation to the R-fuzzy set *BN*, hence why these two R-fuzzy sets are more visible. To return a degree of significance of 0.17, only a single observer would have had to agree to its sentiment. If one refers back to Table 2.4, one can see that observer at p_3 gave their perception of f_4, which was associated to 50dB, which was presented as $\mu(f_4) = 0.57$, as being *BN*, the reason why 0.57 was triggered. The same observer was the reason why the value 0.57 was not unanimously

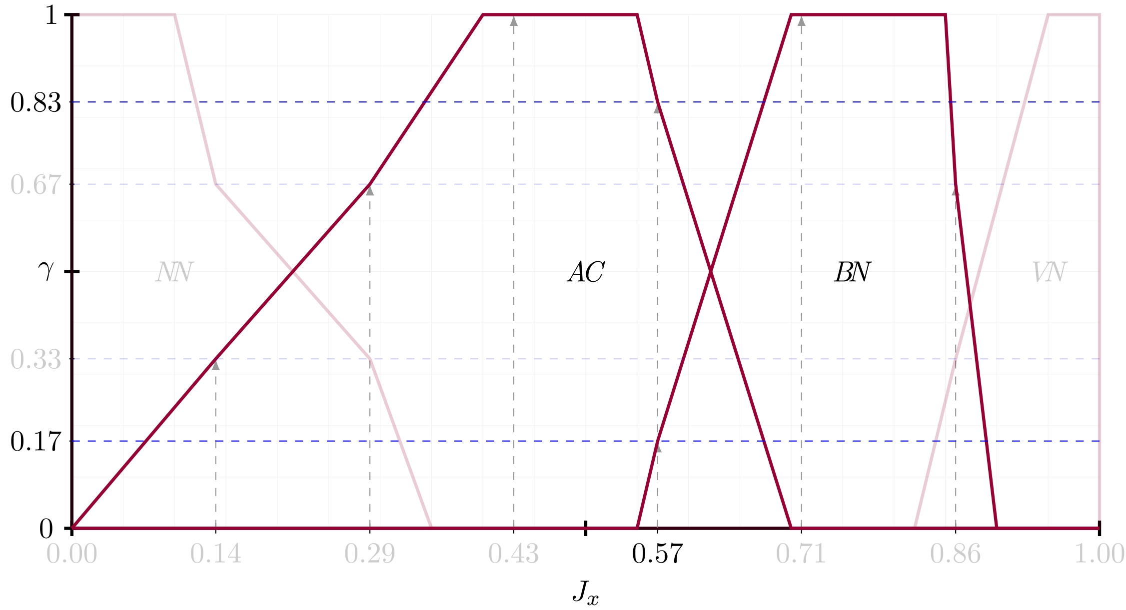


Figure 3.25: α-Cut Representation

agreed upon as being completely indicative of *AC*, hence why 0.57 scored a degree of significance of 0.83. If one was to question the validity of a particular observer or set of observers, one may be inclined to instead make use of the α-cuts, to either promote or demote the significance value for a particular fuzzy membership value. In this case, the impact of changing the degree of significance will change the shape of the membership of

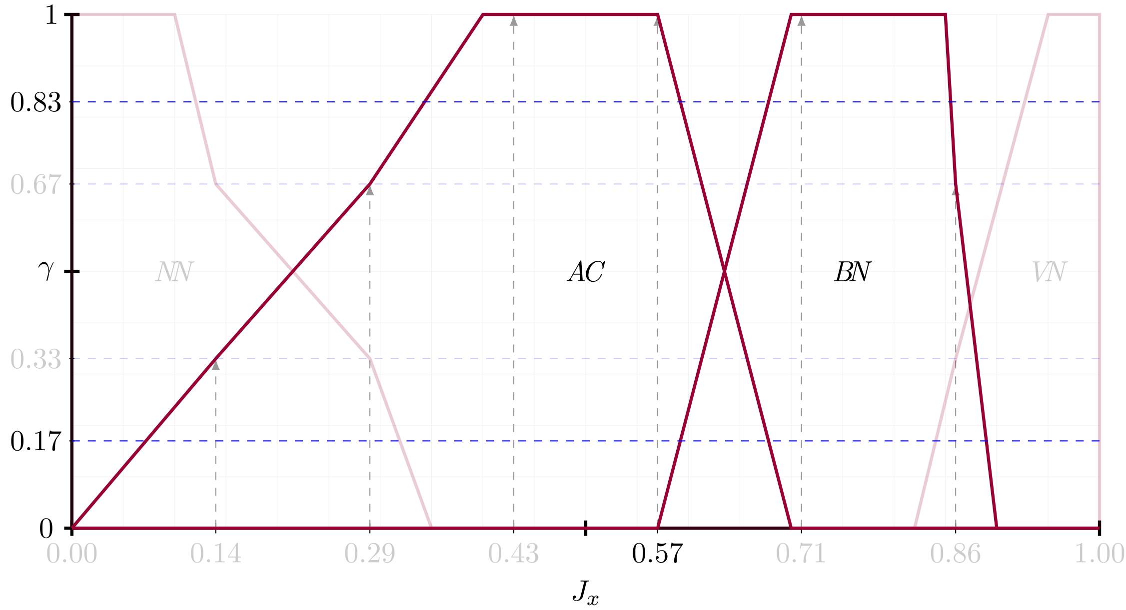


Figure 3.26: A Possible Readjusted Visualisation for *AC* & *BN*

the encapsulated set. Figure 3.26 provides one with a possible visualisation of the adjusted encapsulation for both *AC* and *BN*. One can see from the plot the membership value 0.57 now returns a degree of significance of absolute 1 with regards to *AC*, and a returned degree of significance of absolute 0 with relation to *BN*. All other remaining membership values and their associated scores remain the same, as do the unaltered R-fuzzy sets *NN* and *VN*. This process can be repeated for any instances where the returned degrees of significance encroach onto that of the extremes indicated by the α-cuts.

The use of α-cuts to allow for post result adjustment, provides one the means to apply subjectiveness to the overall returned results. If a single, or small amounts of observers contained in C, gave their perceptions which greatly conflicted against that of the general consensus, then one may be inclined to shift the significance accordingly. However, this raises an interesting question, assuming that the perceptions of those involved are genuine, therefore certain returned perceptions, no matter how outlandish may be indicative of underlying conditions. Colour blindness could be identified as too could varying degrees of deafness or hard of hearing. The application of perception based modelling is vast, its subjective nature makes it a challenging domain in which to engage. The use of an R-fuzzy approach coupled with the significance measure provides one with a framework in which to explore such domains, and return a high level of detail.

3.6.1 An R-fuzzy Shadowed Set

By using the perspective of α-cuts to be indicative of extreme bounds, whereby one is presented with three criteria; non-inclusion (0), total acceptance (1) and upper approximation inclusion to a varying degree $[0, 1]$, one has replicated a shadow set approach as presented in Definition 7. Where possible membership values in the extremes are given either as a 1 or 0, and the values in the shadow region simply belong to the interval $[0, 1]$. This new R-fuzzy shadowed set can be given by the following expression:

$$\bar{A} = \Big\{ \Big\langle v, \gamma_{\bar{A}}(v) \Big\rangle \ \Big| \ v \in \mathbb{U} \Big\} \tag{3.6.10}$$

Where its membership function is indicative of the three contained regions:

$$\gamma_{\bar{A}}(v) : V \to \Big\{ 0, 1, [0, 1] \Big\} \tag{3.6.11}$$

One could argue the point, why would you go to the effort of calculating the significance for each of the contained fuzzy membership values if you are not going to use them? Given certain based perception domains, this would be beneficial, knowing what membership values would be deemed acceptable may be all the information one needs rather than the exact degree of its significance. In particular, one would be more concerned with the values that were on the precipice of totally being agreed upon by all. Simply using the idea of α-cuts having their horizontal slices at the varying possible degrees of significance, would provide one with a great deal of detail regarding the domain the R-fuzzy set was modelled for. This undoubtedly allows for better uncertainty management, a goal of higher uncertainty based approaches (Zadeh, 1983; Brashers, 2001).

3.7 Closing Remarks

Chapter 3 has described and demonstrated the main contributions of this thesis. The pairing of the R-fuzzy approach with that of the significance measure, allows for a greater level of detail to be inferred from. The adaptation from the traditional use of an R-fuzzy set to provide an indication for an unknown observation, as presented in Example 1, compared to the more detailed approach as described in Example 2, shows that an R-fuzzy approach when paired with the significance measure, allows for a universal encapsulation of the criteria set C. This configuration provides for a more informative approach, from which a greater amount of information can be inferenced from. The significance measure caters for not only a discretised understanding, but it also allows for a continuous representation, allowing one to infer the significance of unseen fuzzy membership values.

The addition of the streamlined concept provided the means to generate encapsulating sets for the returned degrees of significance, using the most minimal parameter values. The whitenisation functions were able to provide for a high precision set which adhered to the correctly captured and triggered membership values. Such an approach concentrates on a purely objective perspective and does not take into account a subjective understanding. For example, the plateau regions for instances when a trapezoidal membership function was used, did not extend past the triggered membership values. The streamlined approach does away with arbitrary assumptions, and presents the set as computed based on the returned degrees of significance, a more derivable approach.

With regards to the enhanced R-fuzzy framework as demonstrated in Example 4, the enhanced streamlined sets are not needed, as the returned degrees of significance provides the index points needed for the sequences. The enhanced proposal allows for each member of the fuzzy set J_x, to act as the discretised point along the x-axis. The returned degrees of significance therefore act as the y co-ordinates, connecting the points creates a sequence indicative of its R-fuzzy set. If the data contains clusters, then R-fuzzy subsets can be computed on the same initial fuzzy membership set J_x. In which case, this caters for the functionality to be able to compare and contrast any permutation of R-fuzzy sets and significance measure sequences, to that of conditional attributes.

As Example 4 made use of real test subjects with regards to their perceptions, it is noteworthy to mention that it is plausible that the same experiment could be repeated on the same population, with different results being collected. When modelling perceptions, one is simply taking a snapshot of that very moment, collecting the perceptions of individuals that may or may not have been influenced by their surroundings, mood, time of the day and so on. This does warrant further investigation into how best to record and collect the data, but as is the nature of perception, it is fundamentally subjective to the individual.

4

Empirical Observations

"Il n'est pas certain que tout soit incertain." - Translation: "It is not certain that everything is uncertain."

– Blaise Pascal

4.1 Introduction

This chapter will take on an empirical perspective and provide one with observations regarding varying aspects of the R-fuzzy and the significance measure pairing. The previous examples contained within this thesis have all indicated that a universal value can be agreed upon, however this may not always be the case. Confliction; where contrasting perspective based on the same initial observation can reduce the significance of any contained value, such that a single value cannot be used to indicate the subjectiveness of a given concept. The addition of extra descriptors allows observers more choice in relaying their perceptions, which impacts on the likelihood of membership values attaining absoluteness. The previous examples have also implied the distribution of each computed R-fuzzy set be *neatly* contained, as Example 5 will show, disjoint areas of distribution can occur. Nonetheless, the enhanced R-fuzzy approach can still be undertaken to return a metric, indicative of the divergence between geometric patterns of the sequences obtained using the significance measure.

Section 4.2 describes the restrictions that face a non-enhanced R-fuzzy approach, this ultimately provided the rationale for the work contained in this thesis. Section 4.3 provides an insight when considering how confliction can impact upon the returned degrees of significance. Section 4.4 looks into the impact the number of descriptors can have on the distribution and significance scores for the contained fuzzy membership values. Section 4.5 explains the use of thresholds to better articulate on a given concept. Section 4.6 provides an example of disjoint distribution and how it can be captured and inferred from, using the enhanced approach. Section 4.7 describes the relationship that an R-fuzzy and significance measure pairing has to that of traditional fuzzy set theory. Finally, Section 4.8 concludes the chapter with closing remarks.

4.2 R-fuzzy Restrictions

Consider Definition 8 and Section 2.5.1, which gave the definition for an R-fuzzy set, and described its approximations, respectively. The traditional R-fuzzy approach when compared to other uncertainty models, such as those contained in Section 2.4, has clear advantages. Its ability to retain uniqueness for all captured object values, with the addition of a greater breadth of uncertainty encapsulation, is the reason why it was decided upon to provide for the foundation of this thesis. However, an R-fuzzy set without the use of the significance measure does have restrictions. Consider a criteria set C of 10, 20 or 50 individuals, regardless of size the same vulnerability exists. Contained within C are their correlated collected perceptions. As it has been stipulated in Definition 8, the lower approximation is associated to absoluteness, for any value contained would have to be agreed upon by all in C. Depending on the concept being modelled, a specific value to represent a specific descriptor may be universally agreed upon. However, if 1 out of 10, 1 out of 20, or 1 out of 50, if a single individual disregarded a specific value to be indicative of their subjectiveness, then that uncertain fuzzy membership value will be blocked from belonging to the lower approximation. As a result, the general consensus, regardless of how strongly agreed with, will not have at least a single value to act as its representative.

The same problematic sentiment is prevalent with regards to the upper approximation. A single vote from any individual from any size criteria set C, can allow for any instance of an uncertain fuzzy membership value to be contained within the upper approximation. As an upper approximation contains all values with at least a single voter, all the way up to the entire criteria set minus one, $C-1$, depending on the size of C, this scope can be relatively small or extremely drastic. This was the motivation for the creation of the significance measure, to allow for one to be able to provide a metric for the contained membership values of any computed R-fuzzy set. As it has already been shown, the significance degree for any value belonging to the lower approximation will score an absolute 1, whereas any value belonging to the upper approximation will score any real value within the interval $[0, 1]$. Any membership value not included in the computed R-fuzzy set will score a value of 0.

With this understanding of the vulnerability with regards to an R-fuzzy set, the significance measure simply allows for each and every uncertain fuzzy membership value to be quantified. The same original problem of voter subjectiveness can still occur, however the significance measure can indicate which values can be treated *differently* if need be. This gave rise to the notion of an R-fuzzy α-cut, which will now be described.

4.3 Confliction

One of the benefits of utilising an R-fuzzy approach with regards to how it has been described in this thesis, is that it can be deployed on data sets of varying size. The previous examples, with regard to the cardinality of the criteria set C have been relatively small. However, an R-fuzzy set can be created with just a single observer. In this instance, as there are no other observers to conflict with, any response given by a single individual, regardless of how outlandish, will always be given as absoluteness. By that, any returned membership value for any generated descriptor will always be contained in the lower approximation, by proxy, the upper approximation too. The visualisation of such an instance could take the form of an interval, or in keeping with a more humanistic interpretation, something similar to the continuous representation on the plots from the previous examples. Consider the following R-fuzzy set:

$$M = \Big(\big\{0.25, 0.33, 0.50\big\}, \big\{0.25, 0.33, 0.50\big\}\Big)$$

The membership values themselves are not important, but consider how one could provide for a continuous visualisation for such an R-fuzzy set, where there would be no need to implement the significance measure, as the only values contained in the set, are given an absolute degree of significance of 1. If considering a strict interval representation, a visualisation could look like the one presented in Figure 4.1.

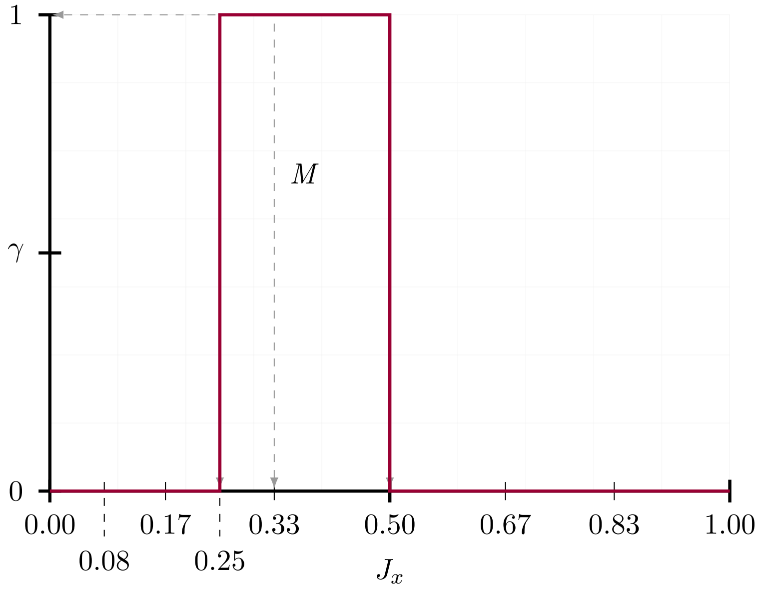


Figure 4.1: A Strict Interval Visualisation for M

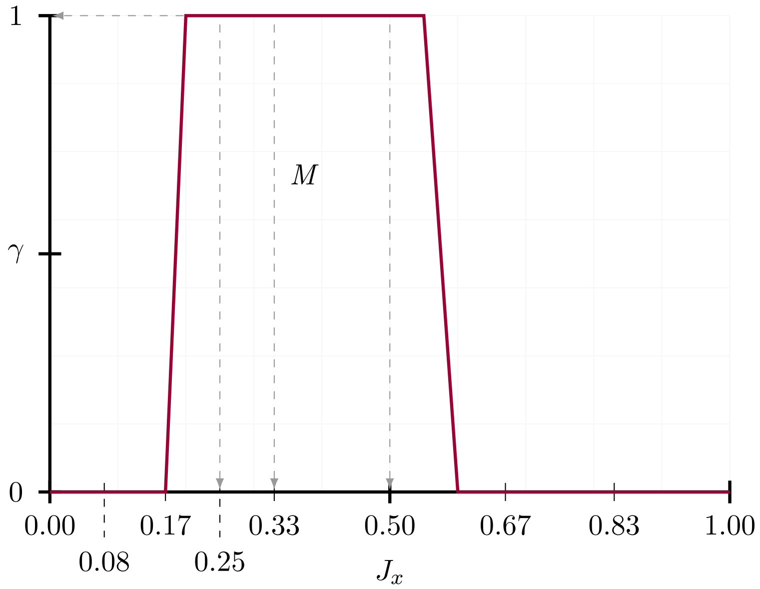


Figure 4.2: Another Possible Visualisation for M

One may be inclined to instead represent the R-fuzzy set more in keeping with how they have been implemented in this thesis, as such, the plot contained in Figure 4.2 provides a possible visualisation of this. Now consider that an additional observer has been added to the criteria set, and there is a level of confliction such that the returned R-fuzzy set is now given as follows:

$$M = \Big(\{0.25, 0.33\}, \{0.25, 0.33, 0.50\}\Big)$$

The returned degree of significance for the contained R-fuzzy are given as:

$$\gamma_{\overline{M}}\{0.00\} = \frac{0}{2} = 0.00 \qquad \gamma_{\overline{M}}\{0.08\} = \frac{0}{2} = 0.00 \qquad \gamma_{\overline{M}}\{0.17\} = \frac{0}{2} = 0.00$$

$$\gamma_{\overline{M}}\{0.25\} = \frac{6}{2} = 1.00 \qquad \gamma_{\overline{M}}\{0.33\} = \frac{2}{2} = 1.00 \qquad \gamma_{\overline{M}}\{0.50\} = \frac{1}{2} = 0.50$$

$$\gamma_{\overline{M}}\{0.67\} = \frac{0}{2} = 0.00 \qquad \gamma_{\overline{M}}\{0.83\} = \frac{0}{2} = 0.00 \qquad \gamma_{\overline{M}}\{1.00\} = \frac{0}{2} = 0.00$$

As a result of this additional observer, the possible returned degrees of significance have also been impacted upon, as now one can anticipate a possible returned degree of 0.50, not just absolute 1 or 0. The shape of the encapsulated set of the significance values is also impacted upon, as such, a possible visualisation may be indicative of that of Figure 4.3.

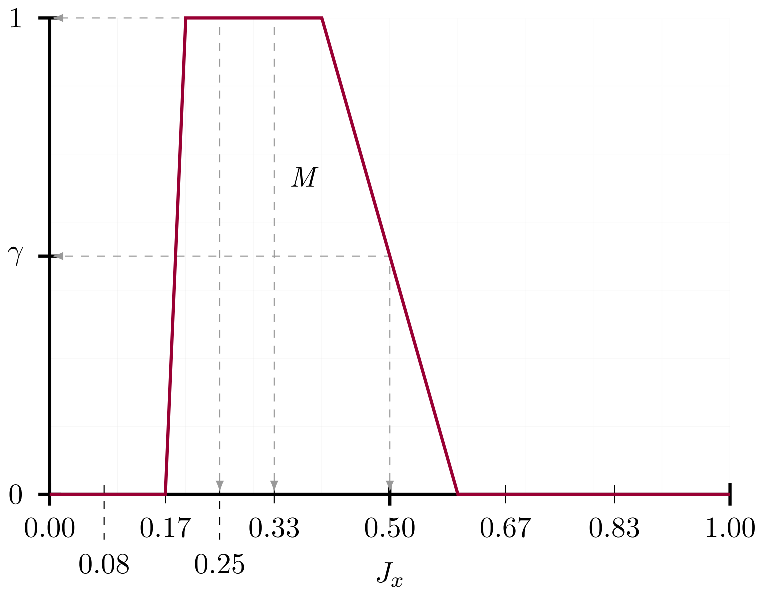


Figure 4.3: A Possible Visualisation for M With Included Confliction

It stands to reason that a confliction of interest can only impact upon a membership value by demoting its significance. The addition of a third observer, when considering the membership value 0.50, would either return a significance degree of $\gamma_{\bar{M}}\{0.50\} = 0.33$, or $\gamma_{\bar{M}}\{0.50\} = 0.67$, but as confliction is still prevalent, the membership value will never return a degree of significance equal to 1, nor 0. This is the same case when considering criteria sets that contain a larger number of observers. The more subjectiveness involved the more likely confliction will arise, as such, the shape of the encapsulating set will have to change accordingly. As it was shown in Example 4, there may indeed be cases where the variation contained within the criteria set is minimal. It is perfectly plausible that the collected perceptions of a populous may indeed be identical. The dichotomy of a little variation to that of a huge amount, is the subjective nature of perception based uncertainty. The responses of those collected may follow a norm, however they may not. Regardless, an R-fuzzy approach is able to capture the general consensus and equally the slight nuances prevalent. The significance measure is able to provide a numeric quantitative value to its importance, with no loss of information.

4.4 Descriptors

The procedural steps involved when considering an R-fuzzy approach require that the descriptors already be decided upon. The various examples contained in this thesis have all involved the use of at least 3 descriptors. Some logical consideration should be applied to the number of descriptors one should employ, having too few will not allow for a comprehensive overview, equally, having too many may reduce the significance of contained membership values. Given that a continuous visualisation of Example 4 was presented in Figure 3.15, assume that the perception of the same individuals were collected, but this time they were given two additional descriptors to choose from. *Very Dark Red* (*VDR*) and *Very Light Red* (*VLR*) along with *DR*, *R* and *LR* provides for a total of 5 descriptors. Modelling the returned degrees of significance for this resulted in the plot contained in Figure 4.4. One can see from the plot that the number of descriptors given as a choice, greatly impacts upon the significance a membership value can score. Having too few descriptors does not convey the concept being modelled in any great detail. Having too many descriptors reduces the likelihood of generating high scoring membership values. The trade-off between generalisation and increased resolution, can be seen as more of an art than an exact science, although common sense should be used.

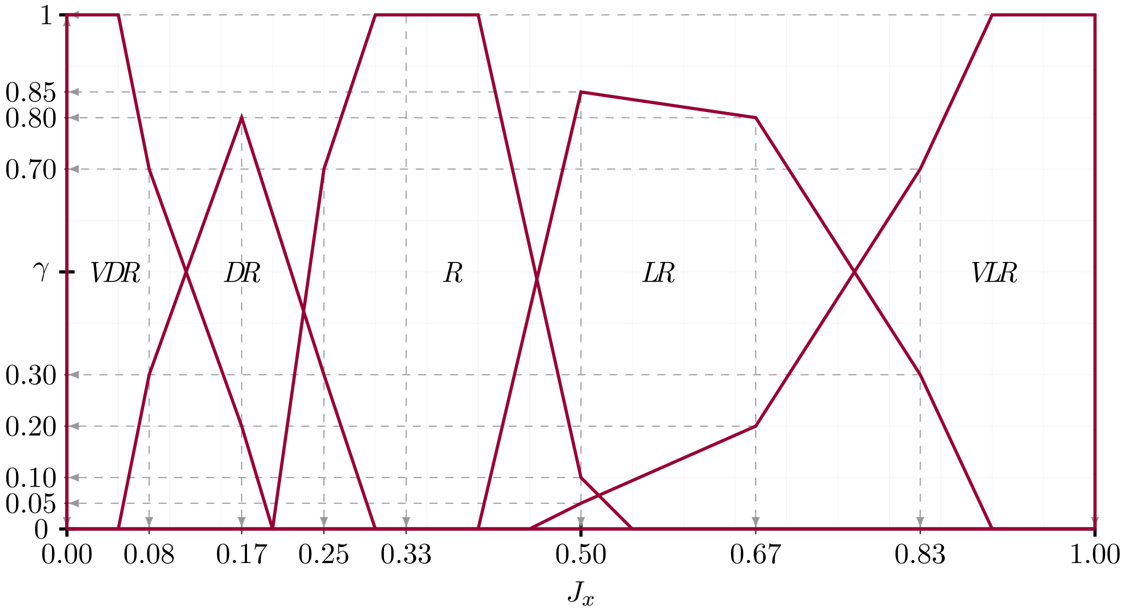


Figure 4.4: A Possible Continuous Visualisation for *VDR*, *DR*, *R*, *LR* & *VLR*

With the addition of the two extra descriptors to choose from, the distribution of *DR* and *LR* are noticeably impacted upon. The distribution of *R* is affected to some degree, but as it is predominately based in the middle, it is only altered slightly. The membership value of 0.83 scored a significance of absolute 1 with regards to *LR* in the original plot Figure 3.15, whereas in the new plot the membership value has a significance of $\gamma_{\overline{LR}}\{0.83\} = 0.30$ and $\gamma_{\overline{VLR}}\{0.83\} = 0.70$. As the additional descriptors are the new left and right most extremes, the old extremes are *DR* and *LR* are not as prominent. The membership values that scored highly in the original plot are reduced by this addition of extra choice.

As it has been seen throughout this thesis, the majority of the plots themselves are based on *possible* configurations, which are arbitrary and subjective. If one was to employ more derivable means, one would most likely do away with the notion of plateau regions, and instead simply connect-the-dots for each uncertain fuzzy membership value, along with its returned degree of significance. In which case, a derivable plot based on Figure 4.4, would look exactly like that of Figure 4.5. Here, one can see that there are no arbitrarily defined plateau regions, instead the plot for each of the contained R-fuzzy sets, is purely based on

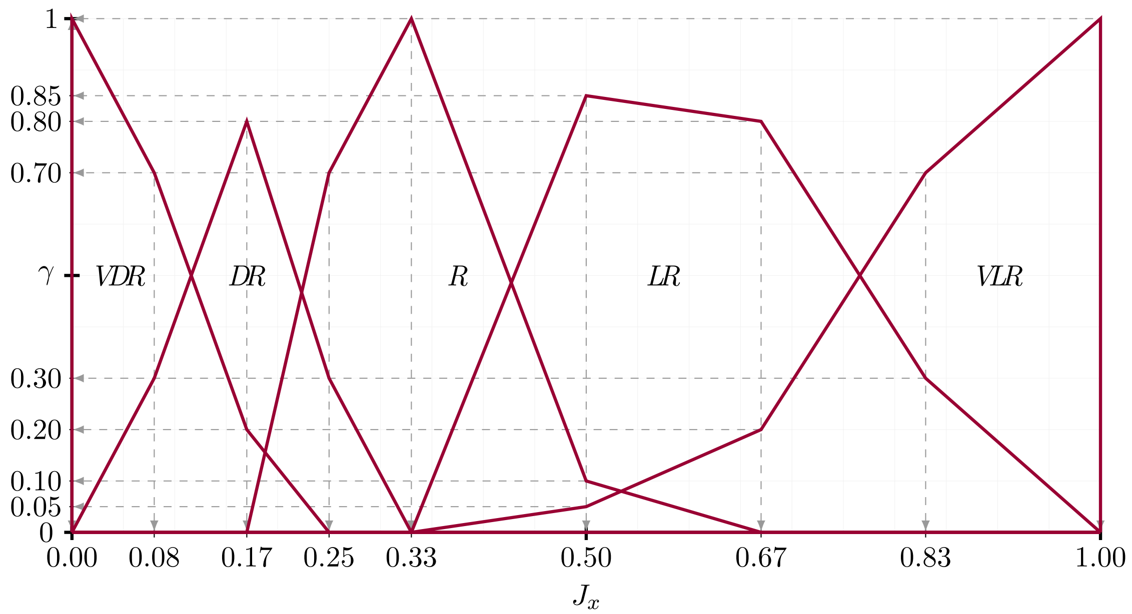


Figure 4.5: A Derivable Continuous Visualisation for *VDR*, *DR*, *R*, *LR* & *VLR*

the significance measure values for each triggered fuzzy membership.

4.5 Thresholds

When using the significance measure to create a continuous representation, from which an entire abstract concept can be modelled and rendered as seen in Figures 3.9, 3.14 and 3.26. This can be used as the basis from which to infer from, using varying membership values not necessarily included from the original example. It provides one with a reference point, where a membership value might fall when speculating with regards to an unknown membership. For example, the R-fuzzy set generated for AC when using the significance measure, presented in Figure 3.2, contained the fuzzy membership value 0.43 in its lower approximation, hence why it has an absolute value of 1 for its significance. The value 0.57 was contained within the upper approximation with a significance of 0.83, the most extreme α-cut. If one was to speculate on what the returned result would be for the membership value of 0.44, a membership not included in the original fuzzy membership set J_x, one would be inclined to score it an absolute 1. One would also be inclined to give a score of 1 to the membership value of 0.45, however these are arbitrary assumptions as one can only assume this would be the case. It would be logical to assume that any membership value greater than 0.50 would need to lose its degree of significance gradually, as the known membership value of 0.57 returned a degree of significance of 0.83.

R-fuzzy sets could be used more intuitively if say, one wanted to know which membership values surpassed a particular threshold of acceptance. Assume $\Phi(v) > 80$, where Φ is representative of a subset of membership values which evaluate to *true* based on the condition of exceeding a significance degree measure of 80. As this would include all absolute lower approximation values, if such existed, it would also include membership values that were incredibly close to being included within the lower approximation. The fuzzy membership value 0.57 returned a degree of significance of 0.83, with regard to the R-fuzzy set AC, according to Examples 2 and 3. This would exceed the threshold of $\Phi(v) > 80$, and as a result would be treated as a lower approximation. Providing a threshold would alleviate the problem of allowing an outlier to stop a particular membership value from agreeing to the general consensus. The threshold value itself could be derived from subjective means, more relative to the notion being modelled.

In much the same way α-cuts provide crisp cut off points, the use of thresholds such as Φ, allows for post result adjustment. To be able to ascertain which values exceeded a threshold, one will be able to identify the membership values that were of greater importance relative to a particular R-fuzzy set.

4.6 Disjoint Distribution

As indicated by previous examples, if such a membership value exists to describe a particular descriptor, such that all collected perceptions agree, then the registered membership values either side will often also be triggered to some degree. Examples 2, 3 and 4, have all indicated that neighbouring membership values are also triggered, as the upper approximation in all instances has more than one contained value. With that being the case one would assume that there would be no contained voids between triggered memberships, however, that may not always be true. From the previous examples, the propagation from one generated R-fuzzy set to the next, tends to overlap, this implies that shared instances of membership values to more than one R-fuzzy set exist. This follows a very harmonic understanding of shifting from one descriptor to the next. Once the overall concept has been visualised, the various R-fuzzy sets based upon their returned degrees of significance, simply provide a reference of how the populous perceived the concept being modelled. Understandably, certain perceptions will involve less subjective variation than others. In which case, conformity will more likely be achieved, such that a value can be used indicatively for a given descriptor.

It is quite possible, and completely plausible to have instances where disjoint distribution occurs, in which case a single membership value or a range of values for a particular R-fuzzy set may be ignored, as they were not voted to be indicative of any perception collected. When dealing with subjectiveness, one cannot simply imply that *their* perception is wrong, as the R-fuzzy set and significance approach is purely objective, based on the collected subjectiveness of the criteria set C. The enhanced R-fuzzy approach will still allow for a sequence to be generated, as the value 0 is still a viable score for the returned degree of significance, and as such it can be collected.

A disjoint distribution for a particular R-fuzzy set will often also be contained within generated R-fuzzy subsets based on the same criteria set. Therefore, when comparisons between sequences using the absolute degree of grey incidence is undertaken, the similarities will be relatively high. However, the subtlety of the divergences is the motivation for the enhanced R-fuzzy framework. If there does exist a difference in perception from contained clusters, then quantify and provide a metric. This numeric index will be indicative of how severe or similar a cluster of cohorts is, with regards to their subjectiveness given any modelled concept.

4.6.1 A Disjoint Enhanced R-fuzzy Example

EXAMPLE 5: Assume that $F = \{f_1, f_2, \ldots, f_9\}$ is a set containing 9 different colour swatches based on the color blue:

$f_1 \rightarrow$	$[25, 25, 112] \rightarrow$
$f_2 \rightarrow$	$[204, 204, 255] \rightarrow$
$f_3 \rightarrow$	$[42, 82, 190] \rightarrow$
$f_4 \rightarrow$	$[0, 51, 153] \rightarrow$
$f_5 \rightarrow$	$[0, 0, 128] \rightarrow$
$f_6 \rightarrow$	$[0, 0, 255] \rightarrow$
$f_7 \rightarrow$	$[40, 146, 172] \rightarrow$
$f_8 \rightarrow$	$[0, 128, 255] \rightarrow$
$f_9 \rightarrow$	$[36, 186, 255] \rightarrow$

In much the same way as seen in EXAMPLE 4, the enhanced R-fuzzy framework example, EXAMPLE 5 investigates the collected perception of 20 individuals, with regards to the colour blue. As the criteria has increased to 20, so too does the chances of disjoint distribution. The more individuals to give their perception, the more likely confliction may arise. These collected perceptions are presented in Table 4.1. Notice the inclusion of the *Sex* attribute, of which C contains 15 males and 5 females.

The descriptor terms contained within the table can be understood as meaning:

$DB \rightarrow$ Dark Blue $\qquad B \rightarrow$ Blue $\qquad LB \rightarrow$ Light Blue

Using the same linear function given in Eq. (2.5.26), which has been used for all the contained examples, the resulting fuzzy membership set is given as:

$$J_x = \left\{0.06, 1.00, 0.35, 0.14, 0.00, 0.24, 0.43, 0.48, 0.65\right\}$$

Table 4.1: Human Perception Based on the Variations for the Colour Blue

#	*Sex*	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
p_1	*M*	*DB*	*LB*	*LB*	*DB*	*DB*	*B*	*LB*	*LB*	*B*
p_2	*F*	*DB*	*LB*	*LB*	*B*	*DB*	*B*	*LB*	*LB*	*LB*
p_3	*M*	*DB*	*LB*	*LB*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_4	*M*	*DB*	*LB*	*B*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_5	*M*	*DB*	*LB*	*B*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_6	*M*	*DB*	*LB*	*B*	*DB*	*DB*	*B*	*DB*	*LB*	*LB*
p_7	*F*	*DB*	*LB*	*B*	*DB*	*DB*	*B*	*LB*	*B*	*LB*
p_8	*F*	*DB*	*LB*	*DB*	*DB*	*DB*	*B*	*DB*	*LB*	*LB*
p_9	*M*	*DB*	*LB*	*B*	*B*	*DB*	*B*	*LB*	*LB*	*LB*
p_{10}	*M*	*DB*	*LB*	*B*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_{11}	*M*	*DB*	*LB*	*B*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_{12}	*M*	*DB*	*LB*	*B*	*B*	*DB*	*B*	*LB*	*B*	*LB*
p_{13}	*M*	*DB*	*LB*	*B*	*B*	*DB*	*B*	*LB*	*B*	*LB*
p_{14}	*M*	*DB*	*LB*	*DB*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_{15}	*M*	*DB*	*LB*	*DB*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_{16}	*F*	*DB*	*LB*	*DB*	*DB*	*DB*	*B*	*LB*	*B*	*LB*
p_{17}	*F*	*DB*	*LB*	*B*	*B*	*DB*	*B*	*LB*	*LB*	*LB*
p_{18}	*M*	*DB*	*LB*	*DB*	*DB*	*DB*	*B*	*LB*	*B*	*LB*
p_{19}	*M*	*DB*	*LB*	*DB*	*DB*	*DB*	*B*	*LB*	*LB*	*LB*
p_{20}	*M*	*DB*	*LB*	*B*	*B*	*DB*	*B*	*LB*	*LB*	*LB*

Using Definition 8, the final generated R-fuzzy sets based on the collected subsets for *LB*, *B* and *DB*, respectively, are given as:

$$DB = \Big(\{0.00, 0.06\}, \{0.00, 0.06, 0.14, 0.35, 0.43\}\Big)$$

$$B = \Big(\{0.24\}, \{0.14, 0.24, 0.35, 0.48, 0.65\}\Big)$$

$$LB = \Big(\{1.00\}, \{0.35, 0.43, 0.48, 0.65, 1.00\}\Big)$$

Table 4.2: The Degrees of Significance for Each of the Generated R-fuzzy Sets

DB		*B*		*LB*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DB}}\{0.00\} =$	1.00	$\gamma_{\overline{B}}\{0.00\} =$	0.00	$\gamma_{\overline{LB}}\{0.00\} =$	0.00
$\gamma_{\overline{DB}}\{0.06\} =$	1.00	$\gamma_{\overline{B}}\{0.06\} =$	0.00	$\gamma_{\overline{LB}}\{0.06\} =$	0.00
$\gamma_{\overline{DB}}\{0.14\} =$	0.70	$\gamma_{\overline{B}}\{0.14\} =$	0.30	$\gamma_{\overline{LB}}\{0.14\} =$	0.00
$\gamma_{\overline{DB}}\{0.24\} =$	0.00	$\gamma_{\overline{B}}\{0.24\} =$	1.00	$\gamma_{\overline{LB}}\{0.24\} =$	0.00
$\gamma_{\overline{DB}}\{0.35\} =$	0.30	$\gamma_{\overline{B}}\{0.35\} =$	0.55	$\gamma_{\overline{LB}}\{0.35\} =$	0.15
$\gamma_{\overline{DB}}\{0.43\} =$	0.10	$\gamma_{\overline{B}}\{0.43\} =$	0.00	$\gamma_{\overline{LR}}\{0.43\} =$	0.90
$\gamma_{\overline{DB}}\{0.48\} =$	0.00	$\gamma_{\overline{B}}\{0.48\} =$	0.25	$\gamma_{\overline{LB}}\{0.48\} =$	0.75
$\gamma_{\overline{DB}}\{0.65\} =$	0.00	$\gamma_{\overline{B}}\{0.65\} =$	0.05	$\gamma_{\overline{LB}}\{0.65\} =$	0.95
$\gamma_{\overline{DB}}\{1.00\} =$	0.00	$\gamma_{\overline{B}}\{1.00\} =$	0.00	$\gamma_{\overline{DB}}\{1.00\} =$	1.00

By using Eq. (3.2.2), one is able to calculate the degree of significance for each and every encapsulated fuzzy membership value, from J_x that has an affinity to its descriptor. The returned degree of significance for all generated R-fuzzy sets are presented in Table 4.2.

If one inspects the returned degrees of significance contained in Table 4.2, one can see where certain membership values have not been registered nor triggered. Not only that, but consider the R-fuzzy set *LB*, and notice how the returned degree of significance for the membership value $\gamma_{\overline{LB}}\{0.35\} = 0.15$, this increases for $\gamma_{\overline{LB}}\{0.43\} = 0.90$, but then drops for $\gamma_{\overline{LB}}\{0.48\} = 0.75$, to then rise again for $\gamma_{\overline{LB}}\{0.65\} = 0.95$, to finally $\gamma_{\overline{LB}}\{1.00\} = 1.00$. All the previous examples have not shown that the returned degrees of significance can fluctuate. Generally speaking the values increase to a point of absolute inclusion, to then gradually decrease to non-inclusion. Nonetheless, the fluctuations are completely acceptable, as the significance measure merely objectively quantifies the collected subjective perceptions of any criteria set.

Figure 4.6 provides one with a possible continuous representation of the generated R-fuzzy sets and the returned degrees of significance. As one can see, the plot differs greatly

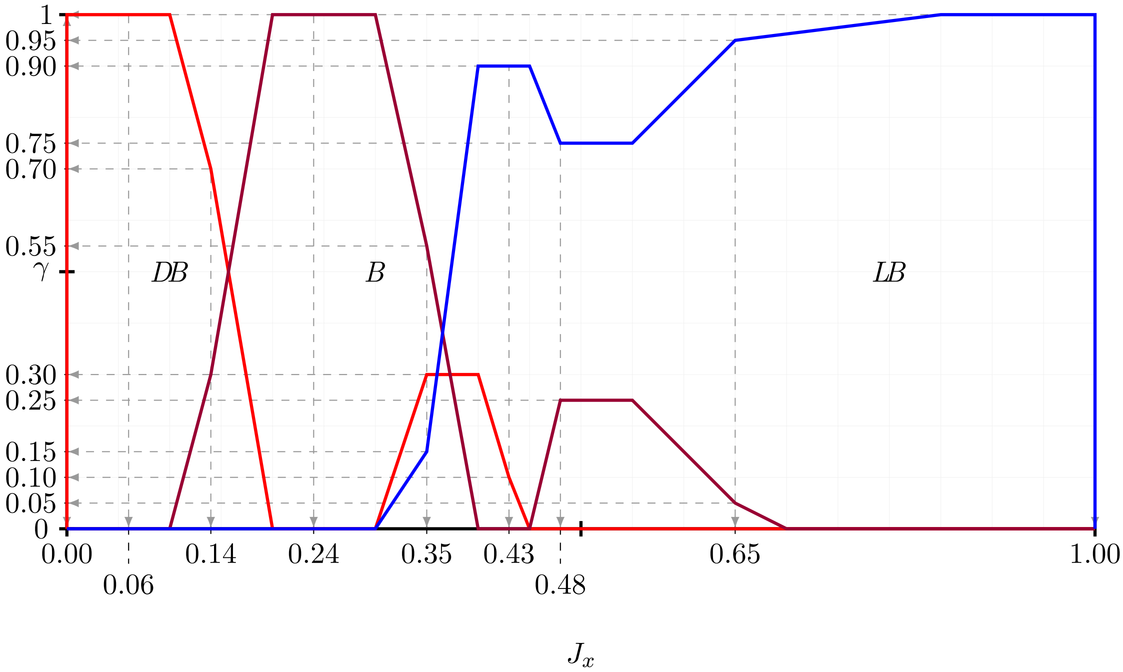


Figure 4.6: A Possible Continuous Visualisation for *DB*, *V* & *LB*

from previous plots, nonetheless, the returned degrees of significance are correctly classified based on what was presented in Table 4.2. Given that $\gamma_{\overline{DB}}\{0.24\} = 0.00$, there exists an area of disjointness between 0.14 and 0.35. As this is a possible visualisation, *DB* has the left most part of the set end at 0.20, and restarts for the right most part of the set at 0.30. The R-fuzzy set *B* was constructed in much the same way, the disjointness itself was declared using arbitrary values which covered the void correctly. The R-fuzzy set *LB* has considerable variance throughout its duration, as can be seen in its fluctuations, it does not however have an area of disjointness. All three generated R-fuzzy sets do have at least one value which returned a significance degree of 1, so even with the extra members for the criteria set, there is still a value that exists indicative of the collective perception held.

In much the same way that the plot contained in Figure 4.5 is the derivable plot for Figure 4.4, the plot contained in Figure 4.7 is the derivable plot for Figure 4.6. Regardless of which plot one draws their attention to, the sentiment of what the visualisation is conveying is maintained, whether it be through a subjective, arbitrary configuration, or a repeatable, derivable configuration for the contained R-fuzzy sets. If one was to implement the use of plateaus, the start and the end points of distribution will be purely subjective and arbitrary. If a more precise and derivable plot is needed, one which is more stricter

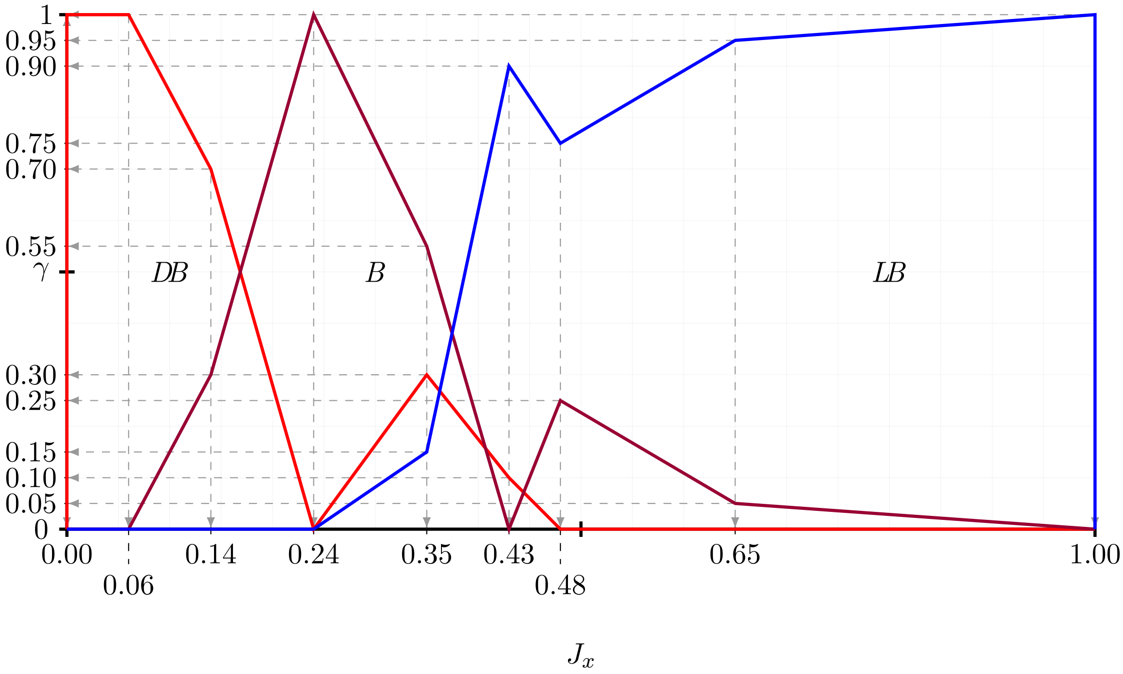


Figure 4.7: A Derivable Continuous Visualisation for *DB*, *V* & *LB*

in terms of membership value coverage, than the means of connecting-the-dots should be chosen, as this can be exactly replicated by third parties assuming it is configured on the same initial data set.

As the data contained in Table 4.2 contains a *Sex* attribute, one can generate 2 subsets with relation to male and female. This in, much the same way as demonstrated in Example 4, allows for one to create R-fuzzy subsets, from which the returned degrees of significance can be used to generate the comparable sequences needed, for the absolute degree of grey incidence. Regardless of how the encapsulation of the returned degrees of significance look, it still provides a valid sequence for comparisons to be undertaken, as any generated subset will be indicative of the overall R-fuzzy sets generated from it, along with disjointness and all.

Table 4.3 and Table 4.4, contain the returned degrees of significance relating to males and females, respectively, all collected from Table 4.1. The plots contained in Figures 4.8, 4.9 and 4.10, show the comparable sequences computed from the returned degrees of significance, for *DB*, *B* and *LB*.

Table 4.3: The Degrees of Significance for Males

DB		*B*		*LB*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DB}}\{0.00\} =$	1.00	$\gamma_{\overline{B}}\{0.00\} =$	0.00	$\gamma_{\overline{LB}}\{0.00\} =$	0.00
$\gamma_{\overline{DB}}\{0.06\} =$	1.00	$\gamma_{\overline{B}}\{0.06\} =$	0.00	$\gamma_{\overline{LB}}\{0.06\} =$	0.00
$\gamma_{\overline{DB}}\{0.14\} =$	0.73	$\gamma_{\overline{B}}\{0.14\} =$	0.27	$\gamma_{\overline{LB}}\{0.14\} =$	0.00
$\gamma_{\overline{DB}}\{0.24\} =$	0.00	$\gamma_{\overline{B}}\{0.24\} =$	1.00	$\gamma_{\overline{LB}}\{0.24\} =$	0.00
$\gamma_{\overline{DB}}\{0.35\} =$	0.27	$\gamma_{\overline{B}}\{0.35\} =$	0.60	$\gamma_{\overline{LB}}\{0.35\} =$	0.13
$\gamma_{\overline{DB}}\{0.43\} =$	0.07	$\gamma_{\overline{B}}\{0.43\} =$	0.00	$\gamma_{\overline{LR}}\{0.43\} =$	0.93
$\gamma_{\overline{DB}}\{0.48\} =$	0.00	$\gamma_{\overline{B}}\{0.48\} =$	0.20	$\gamma_{\overline{LB}}\{0.48\} =$	0.80
$\gamma_{\overline{DB}}\{0.65\} =$	0.00	$\gamma_{\overline{B}}\{0.65\} =$	0.07	$\gamma_{\overline{LB}}\{0.65\} =$	0.93
$\gamma_{\overline{DB}}\{1.00\} =$	0.00	$\gamma_{\overline{B}}\{1.00\} =$	0.00	$\gamma_{\overline{LB}}\{1.00\} =$	1.00

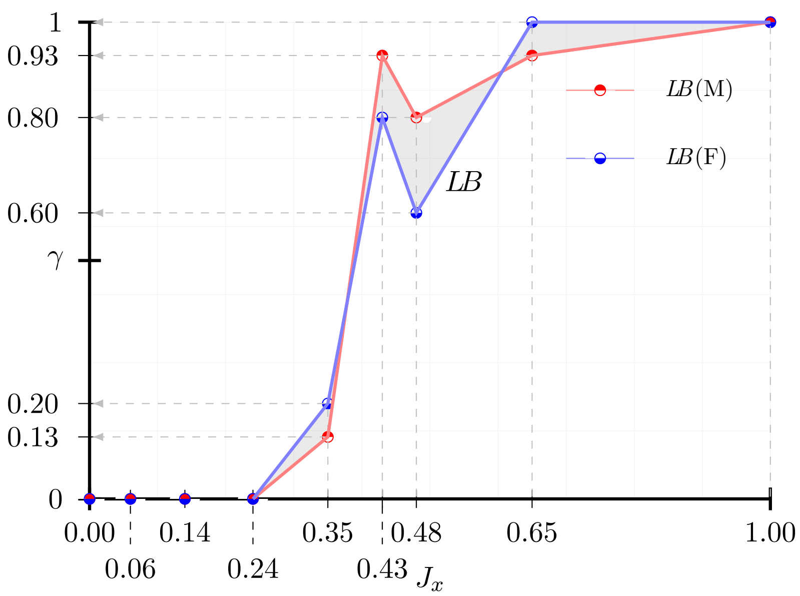


Figure 4.8: The Comparability Between Males and Females for *LB*

Table 4.5 provides a summary of the collected absolute degree of grey incidence values for all generated R-fuzzy sets and significance measure sequences, for each comparable permutation, of which there are 3. Inspecting the table one can see that sequences gen-

Table 4.4: The Degrees of Significance for Females

DB		*B*		*LB*	
J_x	γ	J_x	γ	J_x	γ
$\gamma_{\overline{DB}}\{0.00\} =$	1.00	$\gamma_{\overline{B}}\{0.00\} =$	0.00	$\gamma_{\overline{LB}}\{0.00\} =$	0.00
$\gamma_{\overline{DB}}\{0.06\} =$	1.00	$\gamma_{\overline{B}}\{0.06\} =$	0.00	$\gamma_{\overline{LB}}\{0.06\} =$	0.00
$\gamma_{\overline{DB}}\{0.14\} =$	0.60	$\gamma_{\overline{B}}\{0.14\} =$	0.40	$\gamma_{\overline{LB}}\{0.14\} =$	0.00
$\gamma_{\overline{DB}}\{0.24\} =$	0.00	$\gamma_{\overline{B}}\{0.24\} =$	1.00	$\gamma_{\overline{LB}}\{0.24\} =$	0.00
$\gamma_{\overline{DB}}\{0.35\} =$	0.40	$\gamma_{\overline{B}}\{0.35\} =$	0.40	$\gamma_{\overline{LB}}\{0.35\} =$	0.20
$\gamma_{\overline{DB}}\{0.43\} =$	0.20	$\gamma_{\overline{B}}\{0.43\} =$	0.00	$\gamma_{\overline{LR}}\{0.43\} =$	0.80
$\gamma_{\overline{DB}}\{0.48\} =$	0.00	$\gamma_{\overline{B}}\{0.48\} =$	0.40	$\gamma_{\overline{LB}}\{0.48\} =$	0.60
$\gamma_{\overline{DB}}\{0.65\} =$	0.00	$\gamma_{\overline{B}}\{0.65\} =$	0.00	$\gamma_{\overline{LB}}\{0.65\} =$	1.00
$\gamma_{\overline{DB}}\{1.00\} =$	0.00	$\gamma_{\overline{B}}\{1.00\} =$	0.00	$\gamma_{\overline{LB}}\{1.00\} =$	1.00

erated for *DB* shared the most similarities with a returned metric of $\epsilon(0.968)$. This was then followed by *LB*, with a metric of $\epsilon(0.940)$, therefore the greatest divergence exists for *B*, with a metric of $\epsilon(0.841)$.

With the increase in observers contained within the criteria set, comes a greater chance of disjointness. The sequences generated for Example 4 were relatively self-explanatory and easily interpreted. The sequences generated for Example 5 are considerably more volatile, as indicated by the plots. The sequences themselves often crisscross with one another. This is not a problem, for the absolute degree of grey incidence uses the absolute

Table 4.5: A Comparable Summary of the Returned Absolute Degree of Grey Incidence for *LB*, *B* & *DB*

LB	M	F	*B*	M	F	*DB*	M	F
M	$\epsilon(1.00)$	$\epsilon(0.940)$	M	$\epsilon(1.00)$	$\epsilon(0.841)$	M	$\epsilon(1.00)$	$\epsilon(0.968)$
F	-	$\epsilon(1.00)$	F	-	$\epsilon(1.00)$	F	-	$\epsilon(1.00)$

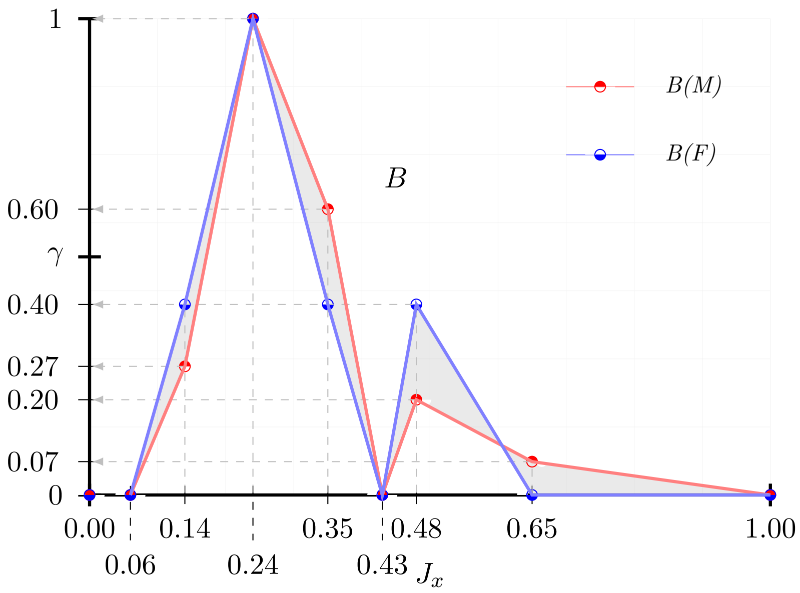


Figure 4.9: The Comparability Between Males and Females for *B*

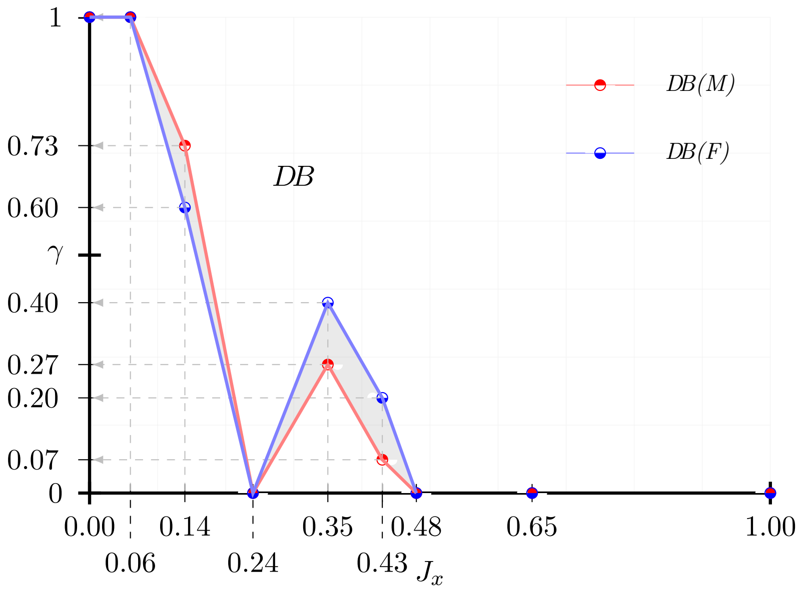


Figure 4.10: The Comparability Between Males and Females for *DB*

values for the area encapsulated between discretised points. As such, a reliable and detailed metric can be returned from which an insight provides the perceived perceptions of males and females. As there were 3 times more males as compared to females, the overall similarities between sequences as indicated by the high returned metrics, would indicate that with more females, the value would only increase slightly, if at all. The robustness that the enhanced R-fuzzy approach has, allows for it to be executed on clusters with uneven frequency, with relatively small amounts of data. This enables whatever mix is contained within the criteria set, to be able to be compared. As the comparisons are of the same R-fuzzy descriptors, the general overall similarity will be prevalent, however, the enhanced R-fuzzy approach is best utilised to provide a metric to the small divergences that exist.

Although Example 4 and Example 5 are both based on the visual perception of colour, one can see that there are noticeable differences in their visualisations. Example 4 presented the population with varying swatches of the colour red, whereas Example 5 considered variations of the colour blue. The *red* experiment compared and contrasted the perceptions of the contained age groups; 20, 25 and 30, of which there was a total of 5 each, making the population size 15. The *blue* experiment used 20 individuals, the majority of which came from the red experiment, but this time only the sexes of the individuals were compared and contrasted with; the population total was 20. The increase in observers increases the possible likelihood of disjointness.

The red experiment followed a more general sense of conformity, as there were no areas of disjointness. The blue experiment however, created areas of disjointness between the generated R-fuzzy sets. Consider Figure 4.6, where the plot is indicative of all the collected perceptions of Table 4.2. Paying particular interest to the R-fuzzy set *DB*, one can see that the area of disjointness occurs between the fuzzy membership values of 0.14 and 0.35, the fuzzy membership value of 0.24 returns a significance of 0. If one moves to the R-fuzzy set *B*, the area of disjointness occurs between fuzzy membership values of 0.35 and 0.48, the fuzzy membership value of 0.43 returns a significance of 0. What is particularly interesting is that when the R-fuzzy set for *DB* is re-engaged, by that, the next time a fuzzy value is triggered for *DB*, is at 0.35, the majority of the R-fuzzy set *B* is contained in the area of its disjointness. Although *DB* does not continuously cover the triggered membership values unlike that of the R-fuzzy sets for Example 4, they still have areas of distribution that were agreed upon by some, hence why the significance values for memberships after the area of disjointness are greater than 0.25.

Considering the possible plot of *LB* in Figure 4.6, one will notice an indent in the plateau between the fuzzy membership values 0.48 and 0.65. This non-convex shape is in stark contrast to the typical convex sets seen for the R-fuzzy sets from the previous examples. This can be attributed to the re-engaged area of distribution related to the R-fuzzy set

B, when it is re-engaged between the fuzzy membership values of 0.48 and 0.65. It echoes the sentiment of an Atanassov intuitionistic fuzzy set as given in Definition 5, although purely from a semantic point of view. The way the experiments were set up, did not allow for an individual to choose a possible R-fuzzy set that would have been indicative of; *I do not agree with*. They were essentially *made* to choose from an approved options list. This undoubtedly effects the distribution of created R-fuzzy sets, as when individuals do not cohesively agree with one another, the factions this causes will stop certain uncertain fuzzy membership values from obtained a significance of absolute 1. For areas of disjointness, it can be seen that when an R-fuzzy set is re-engaged, the *second* part to it, will most likely be roughly the same area of distribution from the neighbouring R-fuzzy set, where the values score significantly less.

4.7 Relationship to Fuzzy

R-fuzzy sets are different from traditional fuzzy sets, in that their membership values are expressed as a set rather than a value. However, there are overlaps with other fuzzy variations under special considerations. In the original R-fuzzy paper by Yang and Hinde (2010), the relationships between R-fuzzy and type-1 fuzzy sets, interval-valued fuzzy sets and Atanassov intuitionistic fuzzy sets were given, with accompanying theorems and proofs. This section will put forward the relationship that exists between that of the significance measure defined in Definition 20, to that of a type-1 fuzzy set. Also described is the relationship between an R-fuzzy set and significance measure pairing, to that of a special condition type-2 fuzzy set. It was remarked by Yang and Hinde, that if the distribution of a membership function could be modelled, it may then be used to derive a fuzzy set, which would give rise to an equivalent type-2 fuzzy set. As it has been shown in Section 3.2, the significance degree measure given in Eq. (3.2.2) does indeed act as an equivalent fuzzy set, when describing its descriptor. Albeit, with regards to its conditional probability of distribution.

4.7.1 Relationship to a Type-1 Fuzzy Set

The significance measure described in Definition 20 is equivalent to a type-1 fuzzy set, if it can be described in the same way as presented in Definition 1. Whereby its membership function must satisfy the restriction imposed upon it, such that an object is assigned a degree of inclusion either equalling or within the range of [0,1]. Also for equivalence to be satisfied, the continuous set representation must be based upon the apex stick heights of the returned degrees of significance for the triggered membership values. From Definition 8 and Definition 20, assume that set A is a descriptor for a particular R-fuzzy set. Given that a type-1 fuzzy set is a collection of ordered pairs. The degree of significance for each membership value belonging to a particular R-fuzzy set is quantified by its membership function $\gamma_{\bar{A}}\{v\} : J_x \to [0,1]$, such that it can be given by the expression:

$$\bar{A} \;=\; \Big\{ \Big\langle v, \gamma_{\bar{A}}\{v\} \Big\rangle \;\Big|\; v \in J_x \Big\} \qquad (3.2.5 \textit{ revisited})$$

Therefore, based on its descriptor the set will contain ordered pairs of membership values and their associated degrees of significance. One can see this expressions is equivalent to the notation given in Eq. (2.2.1):

$$A \;=\; \Big\{ \Big\langle x, \mu_A(x) \Big\rangle \;\Big|\; x \in \mathbb{U} \Big\} \qquad (2.2.1 \textit{ revisited})$$

Where an object is provided with a degree of inclusion relative to the set being inspected. Here we have $v \in J_x$ which is the membership set of membership values instead of $x \in \mathbb{U}$. As J_x provides what essentially is the universe of discourse, the significance degree measure does indeed act as an equivalent type-1 fuzzy set, when the set is representative of the descriptor the R-fuzzy set was created for.

4.7.2 Relationship to a Type-2 Fuzzy Set

A type-2 fuzzy set as given in Definition 6, is a logical extension to that of a type-1 fuzzy set, whereby the addition of a secondary grade of membership is used. The secondary grade itself is a type-1 fuzzy membership, and provides a three-dimensional perspective, allowing for greater encapsulation of uncertainty. The general type-2 fuzzy set is given as follows:

$$\tilde{A} = \left\{ \left\langle (x,u), \mu_{\tilde{A}}(x,u) \right\rangle \,\middle|\, \forall x \in \mathbb{U}, \forall u \in J_x \subseteq [0,1] \right\} \tag{2.4.12 revisited}$$

An R-fuzzy set A is equivalent to a special type-2 fuzzy set as presented in Definition 6, only when one considers the probability distribution of the significance degree measure as a fuzzy membership, only then is an R-fuzzy set equivalent to a special condition type-2 fuzzy set with discrete secondary membership functions.

From Definition 6, we have (x,u) and $\mu_{\tilde{A}}(x,u)$, where (x,u) is indicative of an intersection, and where $\mu_{\tilde{A}}(x,u)$ represents the amplitude, or stick height of objects for said intersection. From Definition 8, an R-fuzzy set uses a rough set to describe its membership, as a result we have $(\underline{M}_A, \overline{M}_A)$, where the lower and upper approximations, $\underline{M}$ and $\overline{M}$, respectively, provide the bounds of the set being approximated, which is the descriptor for set A. The degree of significance as presented in Definition 20, describes the conditional distribution of triggered membership values for its descriptor, given by Eq. (3.2.5). Where the collection of $\gamma_{\bar{A}}\{x\}$ provides the degree of significance of each and every membership value that satisfied the descriptor. As $\mu_{\tilde{A}}(x,u)$ provides one with the amplitude of objects over the footprint-of-uncertainty, $\bar{A}$ provides one with the degree of significance for all triggered membership values satisfying the requirements given by the descriptor. Both approaches make use of a set, which itself describes the distribution of that set.

4.8 Closing Remarks

The strengths of the R-fuzzy and significance measure pairing, allows for the synthesis of a type-2 fuzzy approach, as it has been shown in Section 4.7.2. Such is the connection, the dilemma of excessive precision is no longer such a burden. Type-2 fuzzy is often associated with high computational overhead and complexity of representation, while using crisp values to describe its secondary membership function. As it has already been stated, one may not know all regarding the problem, so securing exactness from which one can base a foundation, could be seen as an unrealistic expectation. An R-fuzzy set on the other hand, allows one to model perception where exactness is not fulfilled. With the addition of the significance measure, one is then able to model the distribution of the encapsulated membership set, so to provide varying amplitude to the stick heights of the captured viable fuzzy membership values. Using this method to get type-1 membership values, R-fuzzy sets can setup an equivalent representation which would require a type-2 membership in a type-2 fuzzy set representation. Type-1 membership can be linked with objective measurements, but type-2 membership is much more subjective and lacks a reliable methodology to set it up. There is no one definitive method to define the secondary grade of membership of a type-2 fuzzy set. Therefore, the work contained in this thesis presents a practical and an effective methodology to conduct the work requiring a type-2 representation. From this perspective, an R-fuzzy and significance measure approach allows for an equivalency, an intermediary approach, to ascertain higher details of resolution that a type-2 fuzzy approach could capture, without the burden of high computational overhead and complexity of representation.

As previous examples have demonstrated, perception does not necessarily follow a strict and stringent train of thought, an individual may hold a perception that goes against the grain of the general consensus. An informative representation of a descriptive term should satisfy not only the requirements of the imprecise representation, but also convey both the common perceptions and individual subjectiveness. The very fact that an R-fuzzy approach caters for this provided the relevance for deciding upon an R-fuzzy foundation from which to expand upon.

The problem of having to use excessive precision to describe increasingly imprecise phenomena has seen several approaches created to try to resolve this dilemma. To some extent these new approaches do offer viable solutions but not wholeheartedly, as several questions still remain. The likes of interval-valued fuzzy sets, Atanassov intuitionistic fuzzy sets and shadowed sets, all allow for the means of encapsulating the uncertainty involved concerning the membership values of a fuzzy set. However, an interval-valued fuzzy set implies that the values contained within its interval are equally distributed, this is an unrealistic assumption for perception based domains. The examples contained within this thesis have shown that unified distribution is not always the case. Values may have a different relationship with the membership values concerned, whereas others may involve uncertainty. Other pitfalls are that a value can lose its inherent meaning if placed in shadowed regions or intervals, once placed in such a container, its uniqueness is diminished.

As seen from the examples, one can see that perception based perspectives may not follow a universal interpretation, individuals may give varying results based on the same observations. These differences and similarities in their perceptions should all be taken into account. With this being the case, a single fuzzy membership value cannot be used to represent a descriptive object, doing so would skew the underlying intent of the perceptions involved. The general consensus and the individual perceptions need to be taken into consideration. A type-2 fuzzy approach extends into the third dimension by using a type-1 fuzzy set to replace the use of crisp membership values. Nevertheless, the secondary membership function itself would still be using crisp values, as a result, the same initial problem still exists. However, the higher levels of type-n one could implement, the closer one gets to precision and an agreed upon model, but not without consequence, the burden of complexity and computation becomes too costly.

As the membership of an R-fuzzy set itself is a set which contains discrete data, there is no loss of detail, unlike that of an interval valued fuzzy set approach. Once the interval has been entered, there is no sense of how close to the bound of that interval the object may be, extremely pessimistic or overtly optimistic. The interval assumes generality and uniformed distribution throughout. As the membership set of an R-fuzzy set is a rough set, the contents of which are crisply defined possible fuzzy membership values, that have an affinity to the descriptor it is being modelled for, no loss of information is suffered. Therefore, one can then quantify the distribution of that R-fuzzy set using the significance measure.

The introduction of the absolute degree of grey incidence for the analysis component is only applicable if one can generate a sequence. As the significance measure computes the significance for all members of the membership set J_x, the returned values apply the amplitude to the discretised points that constitute the universe of discourse. Assuming that all the derived R-fuzzy sets are generated from the same criteria set C, the returned degrees of significance does indeed create a sequence. As Examples 4 and 5 have shown, if the data contains attributes which can create R-fuzzy subsets, one can compute the metrics of divergence based on any number of combinatorial permutations.

5

Conclusion

> *"I am absolutely uncertain about the uncertain amount of uncertainty involved..."*
>
> – Me

5.1 Introduction

This labour of love has documented the journey undertaken throughout the duration of the research, this chapter concludes upon its findings. The foundational underpinnings that constitute the main body of this work is that of the R-fuzzy concept, originally proposed by Yang and Hinde (2010). The functional ability of R-fuzzy when compared to that of other uncertainty modelling concepts, proved itself to be the best-suited candidate from which to expand upon. To echo the sentiment of Section 2.4, the problems associated with the use of increased amounts of precision to model ever increasing amounts of imprecise uncertainty, has seen the creation of several approaches. To varying degrees of success, some of these alternatives do allow for the encapsulation of uncertainty from which one can then further inference. The likes of interval-valued fuzzy sets, Atanassov intuitionistic fuzzy sets and shadowed sets, all allow for the means of encapsulating the uncertainty involved concerning the membership values of a fuzzy set. However, an interval-valued fuzzy set implies that the values contained within its interval are equally and uniformly distributed. This uniformed distribution is an unrealistic assumption to place when considering perception based uncertainty, the very uncertainty that this thesis is associated with. The variance of subjectiveness should be modelled as precisely as possible, allowing for not only a general consensus, but catering for subjective nuances. As it has been demonstrated in the examples, confliction can be prevalent, from an insignificant amount to vast amounts of contradiction. Considering the variations of uncertainty models highlighted in this thesis, a value will lose its inherent meaning if placed within an interval or shadowed region, as once placed in such a container, its uniqueness is dismissed, as it can no longer be distinguished from.

Perception based uncertainty may not involve a universal interpretation of a given concept. The humanistic nature of subjectiveness can be a cause for concern, especially when attempting to define a crisp value indicative of a notion. These possible conflictions of interest, regardless of how similar or dissimilar, should all be taken into consideration. With this understanding, a singleton value, such a fuzzy membership value, may not always be an ideal choice when representing a descriptive object. Doing so, may provide a value that is not truly representative of the populous it was garnered from. This dilemma gives credence to notion of a type-2 fuzzy set, whereby its secondary grade of membership is that of a type-1 fuzzy set. However, the secondary membership function itself would still be using crisp values, harking back to the same problem. Nonetheless, the higher levels of type-n one could implement, the closer one gets to precision and an agreed upon model, but not without consequence, the burden of complexity and computation becomes too costly.

5.2 The Significance Measure

As the membership function of an R-fuzzy set is a rough set, comprised of a lower and upper approximation, no loss of information is suffered, all contained objects retain their discernibility. Unlike the other investigated models, an object is not lost when contained within its interval or shadowed region. The effectiveness of using an R-fuzzy approach, coupled with that of the significance measure, as presented in Section 3.2, provides one with the means for encapsulating fuzzy membership values, when precise memberships are not known.

The original deployment of an R-fuzzy set was to provide a set of values for an unknown observation. With the introduction of the significance measure, the intended use as described in SECTION 3.2.2, provided the means to quantify the significance for all generated fuzzy membership values of J_x, relative to the associated descriptors the R-fuzzy sets were being computed for. In doing so, one can now comprehensively model all the collected perceptions contained in the criteria set C, from which to further query. With the continuous representation that is now possible, provides for a more humanistic interpretation of the perceptions collected. In addition it also provides the significance for values not necessarily contained within the original membership set J_x. As EXAMPLE 2 has demonstrated, a complete and comprehensive understanding of the perception is produced, providing one with additional insight into the abstraction being modelled. The significance measure itself based on that of the standard rough certainty factor described in Eq. (3.2.1), from which a relative significance measure can be derived, Eq. (3.2.2). One which is indicative of J_x and the descriptor based on the rough membership set generated. Understanding the importance of the membership values contained within the upper approximation allows for a better understanding on the perception being modelled.

Given an R-fuzzy and significance measure pairing, there is indeed a relationship to that of a generalised type-2 fuzzy set approach as described in SECTION 4.7.2. As the membership of an R-fuzzy set itself is a set which contains discrete data, there is no loss of detail, unlike that of an interval valued fuzzy set approach (Yang and Hinde, 2010; Khuman et al., 2015a, 2016a). Once an object has entered the interval, there is no sense of how close to the bound of that interval it is; extremely pessimistic or overtly optimistic, the interval assumes generality and uniformed distribution. As the membership set of an R-fuzzy set is a rough set, the contents of which are crisply defined possible fuzzy membership values, that have an affinity to the descriptor it is being modelled for, no loss of information is suffered. The significance measure provides the means to apply an amplitude to the contained membership values. Understandably, it is not an absolute equivalence, rather than see an R-fuzzy and significance measure equate to a generalised type-2 fuzzy set, it is more understandable to see how certain conditional restrictions of a type-2 fuzzy set relate to that of an R-fuzzy and significance measure pairing.

Uncertainty with regards to fuzzy membership values is a common problem in the application of fuzzy sets, which has led to the concept of type-2 fuzzy sets. Type-2 fuzzy sets have a strong theoretical capacity in uncertainty representation (Zadeh, 1975a,b,c), but its associated difficulty is with regards to its highly subjective type-2 membership. In addition, its computational complexity limits its applicability and application domain. By connecting R-fuzzy sets and type-2 fuzzy sets through the significance measure, a new method to solve the challenge of type-2 fuzzy sets in applications, is the result. This was done by replacing the subjective membership with a collection of objective type-1 membership values in an R-fuzzy set. In this way, the precision of type-2 fuzzy sets are preserved but its difficulty in defining type-2 membership is removed. As there is no absolute one method to define the secondary membership of a type-2 fuzzy set, it will be interesting to see how the enhanced R-fuzzy framework can be utilised in this regard.

5.3 Streamlined Encapsulation

In addition to the creation of the significance measure, this thesis also introduced the notion of streamlined encapsulation as presented in SECTION 3.3. The use of the grey whitenisation function to provide an encapsulation of the returned degrees of significance, for any computed R-fuzzy set, still allows for a continuous representation to be inferred from. Unlike the original arbitrary values chosen for the left and right most footprints of a set, or the magnitude of the unit interval at the apex of the trapezoidal memberships, the streamlined approach provides for a high detailed encapsulating set using only the minimal parameters required.

It can be argued that in the context of this research and the deployment of the whitenisation weight function, it is indeed a typical triangular or trapezoidal membership. In which case using the expected notation and the methodology of the streamlined approach, would return the same results. The use of grey theory techniques was done to demonstrate the effectiveness of marrying together the likes of grey, R-fuzzy and the significance measure. The originally intended use of whitenisation from a purely grey perspective is mainly with regards to clustering (Sifeng et al., 2011; Liu and Lin, 2006). Using a grey whitenisation weight function allows for the classification of observations or objective indices, into definable classes. As it has been shown, taking a current methodology and adapting it in an unconventional manner, can provide for promising results.

Comparing the streamlined results with that of arbitrary values decided upon for Example 3, the whitenisation plot can be seen in Figure 3.14, whereas the arbitrary values can be seen in Figure 3.9. The whitenisation approach does away, wherever possible, with the use of a trapezoidal membership, unless absolutely required. However, it can be argued that where a trapezoidal is favourable, the apex of its interval can be used to better model humanistic understandings. The whitenisation approach, although streamlined, can be seen as too stringent regarding the domain it is being modelled for. It becomes specific on the criteria set it computed from, whereas the arbitrary configuration allows for more of a generalisation. There may indeed be domains that would ideally suit the interest of a streamlined methodology, this could be further investigated.

5.4 The Enhanced R-fuzzy Framework

The main contribution of this thesis is that of the enhanced R-fuzzy framework as described in Section 3.4. In much the same with regards to the streamlined approach, the use of grey techniques were integrated. Again, in much the same way, these techniques where adapted from their expected use to that of an alternative. The enhanced R-fuzzy approach utilises the heightened resolution and detail garnered when implementing the significance measure. The application of grey analysis, specifically, the absolute degree of grey incidence is then used to quantify the divergence between comparable R-fuzzy sets for a given abstract concept. Even when only a tiny amount of variance is contained within the criteria set, a wealth of knowledge can still be obtained, as was demonstrated in Example 4. The use of the absolute degree of grey incidence allowed for one to further inspect the change in perception as one propagated through each cluster. If changes did occur, the corresponding significance measure sequence when compared to that of another, would quantify the amount of difference between the metric spaces of the sequences. As the sequences are all based on the degrees of significance for the same fuzzy membership values in J_x, the magnitude for each sequence is guaranteed to be the same with the one to be compared against.

Given that Example 4 compared the divergence in perception with regards to age clusters, Example 5 compared the divergence in perception with regards to sex. In addition, the perceptions collected in the criteria set contained a huge amount of variance, which created disjoint areas of distribution. With disjoint areas of distribution, the enhanced framework can still be employed. Even if a discretised point scored a significance of 0, it is still a valid score and can be compared against. The versatility and robustness of the enhanced approach allows for it to be deployed on domains inherently associated with great variance in perceptions. The use of R-fuzzy, the significance measure and the absolute degree of grey incidence, allows for even greater levels of detail to be obtained and inferred from. In this configuration, the approach provides a more detailed similarity measure, the resolution of which may be favourable to some.

5.5 Future Work

The work contained in this thesis is predominantly theoretical, although the examples use real world data, the concepts that this thesis puts forward would benefit immensely from more real world applicability. As is the nature of perception based domains, the application of the proposed framework will be investigated with other domain specific environments. If it can be shown that such an approach is more effective than other standards, this will provide additional credence to the R-fuzzy notion. With regards to the structure of the R-fuzzy framework, the sensitivity of the result of R-fuzzy sets with regards to the criteria set C and the domain J_x, may act as a restriction to its application in some real world problem domains, where it might be hard to construct a comprehensive C and J_x in the early stages of the inspection. On the other hand, there is no standard framework to construct C and J_x for R-fuzzy sets at the moment.

Therefore, a number of related future research directions can be foreseen from this research. A systematic investigation into the methodology to establish the criteria set C and domain J_x needs to be carried out to establish a practical framework under different application domains. The sensitivity of R-fuzzy sets on its associated C and J_x needs to be further studied to identify a suitable strategy to minimise its impact on the results. Although SECTION 4.7.2 described the relationship between R-fuzzy sets and type-2 fuzzy sets, the relationship between R-fuzzy sets and type-n fuzzy sets can be further explored. In addition to the higher precision for uncertain memberships, the application of R-fuzzy sets to group decision making is also an interesting research topic for further investigation. The application domain of type-2 fuzzy, where either the general or interval interpretation is used, would also be an ideal research interest. Type-2 fuzzy has an excellent capacity in dealing with uncertainty, but the difficulty in defining its fuzzy sets of membership limits its applicability. With the connections between R-fuzzy sets and the significance measure, to that of type-2 fuzzy sets, a type-2 fuzzy problem can be converted to an R-fuzzy set problem. In which case, R-fuzzy sets can be applied to solve problems where type-2 fuzzy sets are currently used, or where it would be preferable to use a type-2 approach. Therefore a typical type-2 application domain, once converted to an R-fuzzy set, becomes an application domain of an R-fuzzy approach.

Bibliography

Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 20(1):87 – 96.

Black, M. (1937). Vagueness. an exercise in logical analysis. *Philosophy of Science*, 4(4):427–455.

Bodjanova, S. (2007). Granulation of a fuzzy set: Nonspecificity. *Information Sciences*, 177(20):4430 – 4444.

Boole, G. (1847). *The mathematical analysis of logic*. Philosophical Library.

Boole, G. (1854). *An investigation of the laws of thought: on which are founded the mathematical theories of logic and probabilities*, volume 2. Walton and Maberly.

Brashers, D. E. (2001). Communication and uncertainty management. *Journal of communication*, 51(3):477–497.

Bustince Sola, H., Fernandez, J., Hagras, H., Herrera, F., Pagola, M., and Barrenechea, E. (2015). Interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: Toward a wider view on their relationship. *Fuzzy Systems, IEEE Transactions on*, 23(5):1876–1882.

Cantor, G. (1895). Beitrge zur begrndung der transfiniten mengenlehre. *Mathematische Annalen*, 46(4):481 – 512.

Chance, B. L. and Rossman, A. J. (2006). *Investigating statistical concepts, applications, and methods*. Cengage Learning.

Couso, I. and Dubois, D. (2014). Statistical reasoning with set-valued information: Ontic vs. epistemic views. *International Journal of Approximate Reasoning*, 55(7):1502 – 1518. Special issue: Harnessing the information contained in low-quality data sources.

Couso, I., Dubois, D., and Snchez, L. (2014). *Random Sets and Random Fuzzy Sets as Ill-Perceived Random Variables.* Springer International Publishing.

Deng, T., Chen, Y., Xu, W., and Dai, Q. (2007). A novel approach to fuzzy rough sets based on a fuzzy covering. *Information Sciences*, 177(11):2308 – 2326.

Deschrijver, G. and Kerre, E. E. (2003). On the relationship between some extensions of fuzzy set theory. *Fuzzy Sets and Systems*, 133(2):227 – 235.

Despi, I., Opris, D., and Yalcin, E. (2013). Generalised atanassov intuitionistic fuzzy sets. In *Proceeding of the Fifth International Conference on Information, Process, and Knowledge Management.*

Drakopoulos, J. A. (1995). Probabilities, possibilities, and fuzzy sets. *Fuzzy Sets and Systems*, 75(1):1 – 15.

Dubois, D. and Prade, H. (1990). Rough fuzzy sets and fuzzy rough sets*. *International Journal of General Systems*, 17(2-3):191–209.

Dubois, D. and Prade, H. (2012). *Possibility theory: an approach to computerized processing of uncertainty.* Springer Science & Business Media.

Dubois, D. J. (1980). *Fuzzy sets and systems: theory and applications*, volume 144. Academic press.

Feller, W. (1968). *An introduction to probability theory and its applications: volume I*, volume 3. John Wiley & Sons London-New York-Sydney-Toronto.

Hllermeier, E. (2014). Learning from imprecise and fuzzy observations: Data disambiguation through generalized loss minimization. *International Journal of Approximate Reasoning*, 55(7):1519 – 1534. Special issue: Harnessing the information contained in low-quality data sources.

Huynh, V.-N. and Nakamori, Y. (2005). A roughness measure for fuzzy sets. *Information Sciences*, 173(13):255 – 275.

Jensen, R. and Shen, Q. (2008). *Computational intelligence and feature selection: rough and fuzzy approaches*, volume 8. John Wiley & Sons.

Jensen, R. and Shen, Q. (2009). New approaches to fuzzy-rough feature selection. *Fuzzy Systems, IEEE Transactions on*, 17(4):824–838.

Ju-Long, D. (1982). Control problems of grey systems. *Systems & Control Letters*, 1(5):288 – 294.

Khuman, A., Yang, Y., and John, R. (2016a). Quantification of r-fuzzy sets. *Expert Systems with Applications*, 55:374 – 387.

Khuman, A. S., Yang, Y., and John, R. (2013). A new approach to improve the overall accuracy and the filter value accuracy of the gm (1,1) new-information and gm (1,1) metabolic models. In *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pages 1282–1287.

Khuman, A. S., Yang, Y., and John, R. (2014). A commentary on some of the intrinsic differences between grey systems and fuzzy systems. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2032–2037.

Khuman, A. S., Yang, Y., and John, R. (2015a). A significance measure for r-fuzzy sets. In *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*, pages 1–6.

Khuman, A. S., Yang, Y., and John, R. (2016b). R-fuzzy sets and grey system theory. In *2016 IEEE International Conference on Systems, Man, and Cybernetics*, page in press.

Khuman, A. S., Yang, Y., John, R., and Liu, S. (2016c). Quantification of perception clusters using r-fuzzy sets and grey analysis. In *2016 International Conference on Grey Systems and Uncertainty Analysis (GSUA2016)*, page in press.

Khuman, A. S., Yang, Y., and Liu, S. (2015b). Grey relational analysis and natural language processing. In *2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS)*, pages 107–112.

Khuman, A. S., Yang, Y., and Liu, S. (2016d). Grey relational analysis and natural language processing to: Grey language processing. *Journal of Grey System*, 28(1):374–387.

Klir, G. and Folger, T. (1988). *Fuzzy Sets, Uncertainty, and Information*. Physica-Verlag.

Klir, G. and Wierman, M. (1999). *Uncertainty-Based Information: Elements of Generalized Information Theory*. Physica-Verlag.

Klir, G. J., Clair, U. S., and Yuan, B. (1997). *Fuzzy set theory: foundations and applications*. Prentice-Hall.

Komorowski, J., Pawlak, Z., Polkowski, L., and Skowron, A. (1999). Rough sets: A tutorial. *Rough fuzzy hybridization: A new trend in decision-making*, pages 3–98.

Lesniewski, S. (1929). Grundzge eines neuen systems der grundlagen der mathematik. *Fundamenta Mathematicae*, 14(1):1–81.

Li, Q., Wang, P., and Lee, E. (1996). r-fuzzy sets. *Computers & Mathematics with Applications*, 31(2):49 – 61.

Lin, T. Y. and Cercone, N., editors (1996). *Rough Sets and Data Mining: Analysis of Imprecise Data*. Kluwer Academic Publishers, Norwell, MA, USA.

Liu, S., Cai, H., Cao, Y., and Yang, Y. (2011a). Advance in grey incidence analysis modelling. In *SMC*.

Liu, S., Forrest, J., and Yang, Y. (2012). A brief introduction to grey systems theory. *Grey Systems: Theory and Application*, 2(2):89–104.

Liu, S. and Forrest, J. Y. L. (2010). *Grey systems: theory and applications*. Springer.

Liu, S. and Lin, Y. (2006). *Grey Information: Theory and Applications*. Advanced Information and Knowledge Processing. Springer-Verlag London, 1 edition.

Liu, S., Xie, N., and Forrest, J. (2011b). Novel models of grey relational analysis based on visual angle of similarity and nearness. *Grey Systems: Theory and Application*, 1(1):8–18.

McCulloch, J., Wagner, C., and Aickelin, U. (2014). Analysing fuzzy sets through combining measures of similarity and distance. In *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 155–162.

Melin, P. and Castillo, O. (2014). A review on type-2 fuzzy logic applications in clustering, classification and pattern recognition. *Applied Soft Computing*, 21(0):568 – 577.

Melin, P., Gonzalez, C., Castro, J., Mendoza, O., and Castillo, O. (2014). Edge-detection method for image processing based on generalized type-2 fuzzy logic. *Fuzzy Systems, IEEE Transactions on*, 22(6):1515–1525.

Mendel, J., Hagras, H., Bustince, H., and Herrera, F. (2016). Comments on "interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: Towards a wide view on their relationship";. *Fuzzy Systems, IEEE Transactions on*, 24(1):249–250.

Mendel, J., Liu, F., and Zhai, D. (2009). α -plane representation for type-2 fuzzy sets: Theory and applications. *Fuzzy Systems, IEEE Transactions on*, 17(5):1189–1207.

Nanda, S. and Majumdar, S. (1992). Fuzzy rough sets. *Fuzzy Sets and Systems*, 45(2):157 – 160.

Newell, A., Shaw, J. C., and Simon, H. A. (1959). Report on a general problem-solving program. In *Proceedings of the International Conference on Information Processing*, pages 256–264.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences*, 11(5):341–356.

Pawlak, Z. (1998). Rough set theory and its applications to data analysis. *Cybernetics & Systems*, 29(7):661–688.

Pawlak, Z. and Skowron, A. (2007). Rough sets: Some extensions. *Information Sciences*, 177(1):28 – 40. Zdzis?aw Pawlak life and work (19262006).

Pedrycz, W. (1998). Shadowed sets: representing and processing fuzzy sets. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 28(1):103–109.

Radzikowska, A. M. and Kerre, E. E. (2002). A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems*, 126(2):137 – 155.

Russell, B. (1996). *The principles of mathematics*. WW Norton & Company.

Sambuc, R. (1975). Fonctions ϕ foues', application a l'aide au diagnostic en patholologie thyroidienne. In *PhD thesis*.

Shafer, G. (1982). Belief functions and parametric models. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 322–352.

Shafer, G. (1990). Perspectives on the theory and practice of belief functions. *International Journal of Approximate Reasoning*, 4(5):323 – 362.

Shafer, G. et al. (1976). *A mathematical theory of evidence*, volume 1. Princeton university press Princeton.

Sifeng, L., Forrest, J., and Yingjie, Y. (2011). A brief introduction to grey systems theory. In *Grey Systems and Intelligent Services (GSIS), 2011 IEEE International Conference on*, pages 1–9.

Sun, B., Ma, W., and Chen, D. (2014). Rough approximation of a fuzzy concept on a hybrid attribute information system and its uncertainty measure. *Information Sciences*, 284(0):60 – 80. Special issue on Cloud-assisted Wireless Body Area Networks.

Wagner, C. and Hagras, H. (2010). Toward general type-2 fuzzy logic systems based on zslices. *Fuzzy Systems, IEEE Transactions on*, 18(4):637–660.

Wagner, C. and Hagras, H. (2013). *Advances in Type-2 Fuzzy Sets and Systems: Theory and Applications*, chapter zSlices Based General Type-2 Fuzzy Sets and Systems, pages 65–80. Springer New York, New York, NY.

Walley, P. (2000). Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2):125 – 148.

Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24(2):149 – 170.

Williams, P. (2007). Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44(3):366 – 383.

Wu, W.-Z., Mi, J.-S., and Zhang, W.-X. (2003). Generalized fuzzy rough sets. *Information Sciences*, 151(0):263 – 282.

Xu, W., Liu, Y., and Sun, W. (2012). Intuitionistic fuzzy rough sets model based on (θ, ϕ)-operators. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*, pages 234–238. IEEE.

Yang, Y. (2007). Extended grey numbers and their operations. In *2007 IEEE International Conference on Systems, Man and Cybernetics*, pages 2181–2186.

Yang, Y. and Hinde, C. (2010). A new extension of fuzzy sets using rough sets: R-fuzzy sets. *Information Sciences*, 180(3):354 – 365.

Yang, Y. and John, R. (2012a). Grey sets and greyness. *Information Sciences*, 185(1):249 – 264.

Yang, Y. and John, R. (2012b). Some extended operations of grey sets. *Kybernetes*, 41(7/8):860–873.

Yang, Y., Liu, S., and John, R. (2014). Uncertainty representation of grey numbers and grey sets. *IEEE Transactions on Cybernetics*, 44(9):1508–1517.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(1):28–44.

Zadeh, L. A. (1975a). The concept of a linguistic variable and its application to approximate reasoningi. *Information sciences*, 8(3):199–249.

Zadeh, L. A. (1975b). The concept of a linguistic variable and its application to approximate reasoningii. *Information sciences*, 8:301–357.

Zadeh, L. A. (1975c). The concept of a linguistic variable and its application to approximate reasoningiii. *Information sciences*, 9:43–80.

Zadeh, L. A. (1983). The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy sets and systems*, 11(1):199–227.

Zadeh, L. A. (1999a). From computing with numbers to computing with words. from manipulation of measurements to manipulation of perceptions. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 46(1):105–119.

Zadeh, L. A. (1999b). Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 100:9–34.

Zeng, A., Li, T., Liu, D., Zhang, J., and Chen, H. (2015). A fuzzy rough set approach for incremental feature selection on hybrid information systems. *Fuzzy Sets and Systems*, 258(0):39 – 60. Special issue: Uncertainty in Learning from Big Data.