A Mathematician's Guide to Fuzzy Logic

with Applications in Fuzzy Additive Systems

A Thesis

Presented to the Faculty of the Department of Mathematical Sciences

Middle Tennessee State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Mathematical Sciences

---

by

Zachariah Thomas

March 29th, 2021

Thesis Committee:

Dr. Mary Martin, Chair

Dr. Wandi Ding

Dr. Joshua Phillips

Dr. James Hart

ACKNOWLEDGMENTS

I must give special thanks to my educators of all levels that made a significant difference in my academic development. This includes Dr. Mary Martin, Dr. James Hart, Dr. Wandi Ding, Dr. Zachariah Sinkala, Mr. James Reed, and Ms. Maria Smith. Additionally, I give thanks to my parents, Ronald and Donna Thomas, who homeschooled me for the majority of my grade school education.

ABSTRACT

Traditional set theory, or crisp set theory, is built on the concept of crisp sets. These are sets for which the membership of an element within a set is defined to be either true or false; in or out; 1 or 0. This construction is extremely useful, as the vast majority of mathematics has shown, but struggles to model concepts of our world which possess vagueness or uncertainty. Therefore, as in the style of Lotfi Zadeh [51], we explore an expansion of set theory to allow an element to be partially within a set, thus constituting what is known as a fuzzy set. These fuzzy sets are namely used in modelling this vagueness. Definitions of core mathematical constructions can be expanded to be defined for these fuzzy sets. These expanded definitions prove to be equivalent to traditional, crisp definitions when crisp sets are used.

Throughout this paper, we explore the results of fuzzy research in set theory, algebra, and analysis; as well as the selected topics of fuzzy systems, an application of fuzzy logic in computer science. It is our aim that the reader, with a moderate background in theoretical mathematics, will be able to read this paper as a guided entry into the world of fuzzy mathematics. There are many reviews of fuzzy logic that have a similar goal. These reviews frequently focus on a fuzzy set theory introduction along with a specified application (see [29] and [31]). In contrast, we aim in this paper to provide a comprehensive, concise exploration of recent fuzzy research in a variety of mathematical fields, while illustrating the parallels of our fuzzy constructions to corresponding traditional ones. We do this following the recent research of the international journal, *Fuzzy Sets and Systems*. Additionally, we conclude this paper by illustrating the logical structure of fuzzy inference systems, as an exploration of preliminary information needed by a researcher to interact with the recent work, “Fuzzy Additive Systems” [23], by Bart Kosko.

CONTENTS

List of Tables

List of Figures

CHAPTER 1

Introduction

How well do we understand the logic of our own language? If asked, would we be able to concretely outline the logical decisions we make regularly without much thought? Let us consider the following task. "Collect a set of apples from a fruit basket". Most humans, speaking the language of the task, would have no trouble identifying and separating the apples from the rest of the fruit. If we ask several people to complete the same task, they would come to the same end result; a subset, A of the universal set of fruit X, consisting of some number, n, apples.

Now let us alter the task slightly. The command is now, "Collect all ripe apples". In this case, the vague qualifier "ripe" has been added. Let us ask different people to collect the subset, A_R, of ripe apples; participants still perform the task with little difficulty, but may come to different end results. What one person considers ripe may not look quite ripe to another. This is despite the fact the word, "ripe" has a commonly understood definition. This is because "ripe" is vague, and while it has a universally agreed upon definition, people sometimes disagree as to what meets the definition and what does not. There is generally uniform agreement in extremal cases, such as a bright red apple or a hard, prematurely green apple; however, intermediate cases, such as an apple that has just begun to redden, may be some cause for argument.

We use words like "ripe", "tall", "bright", "cold", and "long" every day in our spoken language, and generally function with a common understanding of what is meant by these terms. Yet, in our mathematical language, we choose to eliminate words like this from our academic vocabulary. We do this as a means to be more precise, and to make for more common agreement about decisions and reasoning. This allows for as little disagreement among participants as possible. For example, instead of "cold", we might say "below 45° Fahrenheit", or instead of "bright" we say "exhibiting 80 or more lumen". This increases precision, but at the cost of altering the commonly understood definition behind of the original descriptor, "cold".

When asked to qualify two days, one after the other, with one day being $44°F$ and the next day being $45°F$, participants would likely find that the same people would consider both days to be cold, or both days to be not cold. However, let us instead give participants a thermometer, and ask them to determine if the temperature is strictly below $45°F$. Now the responses to the survey are more uniform, with less disagreement, but each participant would say that the first day was below the threshold, and the second was not. These responses match little to none of the original responses. This implies that our innate understanding of the word "cold" does not indeed match up perfectly with a precise qualifier like "below $45°F$". We are unable to model a set of cold days, or a set of bright light bulbs. Therefore we have sacrificed modelling a large number of commonly-used, intuitive concepts for the sake of uniform agreement and precision.

In terms of traditional set theory (and consequently algebra, analysis, graph theory, and topology) most vague qualifiers such as "cold", are often redefined into some crisp statement. These crisp statements generate what we understand to be crisp subsets of the universal set. For each of these subsets, every element is either "in" the set, or "not in" the set. No element is ever both "in" and "not in" at the same time. The characteristic function of this set maps the universal set to the collection $\{0, 1\}$. However, consider a simple expansion of the codomain of the characteristic function to the continuum, $[0, 1]$; we thus gain a method of modelling these vague qualifiers. It is this vagueness of truth values which we call, fuzzy logic; the alternative, traditional logic of $\{0, 1\}$, we call crisp logic. Using this fuzziness, we are able to mathematically model the elements of our language which do not so easily fit into crisp logic.

Throughout this chapter, as our introduction to the beginning of fuzzy logic, we further explore the natural way we use fuzzy logic in common language, as well as in our functional psychology. We answer the question of whether we can "round off" values of "in" or "out", and in so doing show how fuzzy logic resolves the Sorites Paradox. Finally, we formally define a fuzzy set, and give some basic examples of modelling with fuzzy sets.

1.1 How the Brain Uses Fuzzy Logic

Perhaps the best argument for the application of fuzzy logic is the human brain itself. Often without being conscious of it, our brains classify objects and ideas. We look at new objects in our environment, and try to understand them in terms of objects we have already interacted with. This is done more formally by scientists. When discovering a new species, one of the first tasks to complete is that of classifying the species into a group, such as class, family, or genus. In this way, scientists can make further predictions about what qualities the species may have by comparing it to other members of its classifications. To try to understand the animal kingdom without this classification system would be virtually impossible. With over 8.7 million different species on earth, trying to become a veterinarian would be quite a task indeed.

Less formally, our brains operate by subconsciously classifying everything around us, whether for good or for bad. In pedestrian life, we classify to operate in the world around us. We may decide whether to watch a movie based on characteristics we match with previous movies we liked. Part of our reasoning to predict why we might like a movie is the classification into genres that we may be a fan of; horror, comedy, action, romance. We may also consider whether these are movies that have actors we enjoy or not. We constantly label groups of objects and try to put every object into some group. This mental classification is the subject of study of many psychologists, including (Bornstein, 2010)[8] and (Rosch, 1973)[42], who discuss in detail the inner workings of our mental classification system.

Without this system, our brains would not be able to function, as we would need to remember every detail of our memories independently (Antzoulatos & Miller, 2011)

[4]. In the wild, we would use classifications such as “predator” to identify a tiger, and know not to go near it. In the developed world, we classify streets to identify whether it is safe to stand in the middle of one. This classification, or sorting into sets, is not a concept invented by mathematicians. Rather, it is an innate function of the brain. We classify both for efficiency, and to be able to make relatively dependable predictions about the future. This also helps us to maximize the speed with which we make decisions that result in beneficial outcomes.

We attempt to mimic this process in our construction of mathematics. However, to fully model these classifications, we must be able to model even the vague and uncertain classifications that the brain makes. Traditionally, to model vague concepts with sets, we would create a cut-off point of some artificial threshold, such as $45°F$. We are often asked to do this in every-day language without us realizing. For example, we have a mental classification of what books we enjoy. This is remembered as a fuzzy set, as we obviously have some books which we enjoy more than others. However, suppose we are presented a specific book, and asked a crisp question, “Do you like this book?”. If not asked to expound on our answer, we often give a crisp answer, “yes” or “no”. In this way, we round off our response to an extremal value based on a subjective benchmark. We do this commonly in crisp mathematical modeling, such as in the example of redefining “cold” to being “below 45°”. This method is useful in some cases, even in fuzzy logic. We formally define this in chapter 2 as creating an $\alpha-$cut of a fuzzy set.

The same concept of rounding off to an idealization of mental categories is shown in an experiment at UC Berkley in 1973 [42]. In this experiment, Eleanor Rosch provided students with a task of categrorizing objects as being “in” a set, or “not in” a set. These categories were well-known, with definitions internalized (i.e. “vehicle”, “fruit”, “crime”, etc.). In the experiment, Rosch found that students universally agreed upon focal examples of each category, such as “car”, “apple”, and “murder” for their respective categories. However, when it came to other objects, such as tricycle, olive, and blackmail, there was a fair degree of disagreement as to whether these items were in their respective categories, or sets. It is our assertion that when students are asked to crisply categorize these objects, they are rounding off their internalized membership values for these objects within the sets. For example, many view a tricycle as being "somewhat" a vehicle, but if asked if it is, round off to either a crisp yes or no. This rounding off creates a larger mean squared deviation among the responses. Furthermore, when formally stated and pushed to its limits, this method of rounding off leads to a quite glaring paradox.

1.2 Sorites Paradoxes

Let x be an object of our universe, such that x is in a classification A. In many applications, this object x can be broken down into components. The Sorites paradox arises from the question, "If we remove one of these components which make the object x, is the resulting object still in the set A?". This is best illustrated through examples.

Example 1.1. Imagine we have just received an order of gravel. It is deposited in a 'heap' next to where we would like our new driveway. We would all say that the object is a heap, or pile, and if we were to record a set $H = \{heaps\}$, we would say that this object is in the set of heaps. In other words, the membership value of this object in the set, H, is 1.

Now we begin the process of shovelling the gravel where we want our new driveway to be. Begin by taking one shovel full of gravel, and toss it onto the desired area. At this point, do we have a heap, or a driveway? Most people would say that we still have a heap. However, we have obviously made it "less of a heap" with the first shovel-full. Yet in crisp set theory, there is no way of modeling the statement, "This object is less of a heap than the previous object". Thus we say that the membership of the object remains 1.

Now continue to shovel a second time. Do we still have a heap? We would say yes, and thus the membership of the object is still 1. How about after the third shovel? The fourth? The fiftieth? By this inductive process, we would say that for an object which is a heap, the membership of the object in H is not diminished by the removal of a single shovel of gravel, or even the removal of a single piece of gravel. Therefore, if we fast forward to the end of the process, we are forced to say that the completed driveway is a heap. This contradicts our natural understanding of a heap, and exemplifies the Sorites paradox. We are unable to model the gradual diminishing value of "heap-ness" of the object, and thus our crisp mathematical logic, through induction, leads us to make statements which contradict the physical world we are trying to model.

Example 1.2. We know what a chair is because we have seen many examples of chairs and have labeled them as such. If we happen upon a chair that we have never seen before, we do not have to wait for someone to tell us it is a chair before we can sit on it. We are able to recognize the object and classify it into the set of "chairs" in our brain, based on the objects likeness to other elements already in our mental set "chairs". This is sometimes despite the fact that some chairs do not perfectly match-up with previously learned examples of chairs. There are new styles of chairs being designed quite frequently. Yet even if we have never seen a specific type of chair, we can still recognize it as a chair by specific characteristics we have learned by trial and error in our formative years.

This is thus far straightforward, and does not contradict anything we have learned from bivalent logic and crisp set theory. However, if this were all our brains did, we would operate a lot more like computers than we actually do. We would be purely beings of zeros and ones. Of course, we know this is not the case.

To illustrate, consider a wooden chair, c in the set of all chairs C. What does our brain tell us if we take that chair, and a chisel, and chip a small sliver of wood from the chair. Is the object that is left still a chair? If so, chip away another sliver. Is it still a chair? Continue the process until all that remains is a pile of wood shavings. When does the object stop being a chair and start being a pile of kindling?

During this process of chiseling and reclassifying, most humans quickly start to use words like "kinda" ," somewhat", " mostly" and "partially", when classifying the ever decreasing "chair-ness" of the object. After 10 slivers have been taken, we might say that the object is a chair, but when asked to compare it to what it once was, say that the original object was "more of" a chair. After further chiselling, we say that the object is "partially " a chair, perhaps when the object starts to be structurally unsound. Keep chiseling and we start saying that we have "a half chair". Continue and we say we have "a piece of a chair" and finally we just have a pile of chips, which we do not classify as chair at all.

At what point is the object no longer in the set C? We know that we cannot pinpoint an exact moment in time, or a specific molecule of the wood structure, that was responsible for the object loosing its "chair" status. If we could, then this molecule would be the defining particle of the chair. We do not measure any number of molecules or wood slivers when using the word "chair" while interacting with the world. However, this does not usually keep people up at night wondering how could anything ever be a chair, but not be "as chairy" as it once was. This degree of "chair-ness" doesn't fit well into crisp set theory. Yet, in general situations, the common pedestrian understands that words like "kinda" and "somewhat" emphasize a real phenomenon of degree to which we can use our seemingly bivalent classifiers such as "chair". In set theory, we refer to these classifiers as sets, and in the quest to model the vagueness of classifiers, we expand set theory to create a fuzzy set theory.

1.3 Application using Vagueness of Non-vague Classifiers

We have seen that we may apply fuzzy logic to model conceptual problems in mathematics with vagueness or uncertainty. We thus find immediate applications of this new capability. Many of these applications arrive in increasing the efficiency of automated machines, or the accuracy of smart devices. For example, how do we model the statement, “Is the helicopter balanced?”. We have seen that certain qualifiers such as “cold” are by their nature, fuzzy. However, “balanced”, in this context, has a crisp definition. It is the state in which the edges of an object’s specified base, or plane, is equidistant from the center of Earth’s gravitational pull. Even though this concept has a crisp definition, most of us do not use it 100% crisply. A helicopter technician would consider $0.5°$ of tilt to be sufficient for stable flight, and would not waste energy attempting to achieve the perfect balance. Indeed, the further we dissect degree measures of tilt, the more it is evident that true, perfect “level” is practically impossible in real applications. Traditionally, we would not use the term “level” or “balanced”, but would opt to program a range of tilt for which we would consider the helicopter “level enough”. If the helicopter was in this range, then the automatic program would consider the object “stable”, and thus not put forth the algorithm to correct for imbalance. If it was outside this range, even by a milli-degree, it would begin the stabilization procedures it has been programmed for.

This method can be inefficient and clunky. A skilled helicopter pilot does not wait until the helicopter is at a certain tilt before correcting suddenly. The pilot continuously makes appropriate, proportional corrections based on the slight changes in tilt, or even make corrections to counteract anticipated changes in the tilt. We may say to the question, “Is the helicopter balanced?”, merely “yes” or “no”, but we know there is always some measure of tilt to the aircraft. We as humans “round off” the meaning of “balanced” and say “yes” when the aircraft is “balanced enough”. Hence, despite the fact that the qualifier “balanced” has a crisp definition, we use it as a fuzzy qualifier.

The above example is an over-simplification of one of the applications of fuzzy logic. In (Phillips et al., 1996) [40], Phillips, Karr, and Walker create a system of fuzzy rules by genetic algorithms, that can pilot a helicopter.

1.4 The Fuzzy Goal

Our claim is not that every definition, and every statement, is only ever “kind of” true; rather that the functional definition, and sometimes the technical definition, is often not so crisply defined as we would like to think. Indeed, there are objects that clearly meet the technical definition of words like “disease”, and objects that do not. However, the everyday, functional understanding of “disease” may not classify exactly the same way that the technical definition would. One key reason for this is that we often rate objects in our minds according to what is “sort-of diseas-ish” (i.e. the common cold) and what is “ super diseas-y” (i.e. lung cancer). Classes, even ones that have very crisp technical definitions, often still function in fuzzy, ranked ways in the human mind and in everyday language. Any useful classification that exists must be one we can mentally model. Otherwise, what is the purpose of mathematical classification?

Einstein once said in his lecture, “Geometry and Experience” [11], “As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.” It is our desire with fuzzy logic to follow the first half of this statement. We sacrifice certainty in the name of modelling reality with more accuracy; we maintain that the fuzzier the formulation, the better it models the physical world as we understand it through our internalized understanding within our language.

1.5 Crisp Set

We need begin our discussion of fuzzy logic with its core, a fuzzy subset. To define a fuzzy subset, we first need define a simpler subset, or crisp subset of a universal set. We use notation that is not standard for crisp sets. Envision a universal class or domain, labeled X, of every object being discussed in a system.

Definition 1.3. Crisp Subset: We call A a crisp subset of X, or simply a crisp set, if A is a function
$A : X \to \{0, 1\}$.

To give a general sense of this definition in more traditional terms, we can interpret A as

$$A(x) = \left\{ \begin{array}{cc} 0 & x\ not\ in\ A \\ 1 & in\ A \end{array} \right\}$$

This function is traditionally called the characteristic function of the set A. For ease of expansion into fuzzy sets, we use the characteristic function, referred to as the membership value function for fuzzy sets, as the definition of a set. For finite sets, it is sometimes efficient to also write A using set function notation; that is, a set of coordinate pairs,
$\{(x, A(x))\ |x \in X$ and $A(x)$ is the membership of x in $A\}$.

Example 1.4. Consider the universal set of base 10 digits, $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Now consider the set of all nonnegative, even, single-digit integers, denoted E. Then we may define E as

$$E = \{(x, E(x))\} = \{(0, 1), (1, 0), (2, 1), (3, 0), (4, 1), (5, 0), (6, 1), (7, 0), (8, 1), (9, 0)\}$$

Example 1.5. Consider a universal set of people,
$X = \{John,\ Barbara,\ Stacey,\ Carmen,\ Michael, Suzy\}$.
Let L denote the set of people taller than 6 feet. Suppose we know the heights of our elements are 5’ 4”, 6’ 1” , 4’ 6”, 6’ 2”, 5’ 7.99”, and 5’ 9” respectively. We therefore see that

$$L = \{(John, 0), (Barbara, 1), (Stacey, 0), (Carmen, 1), (Michael, 0), (Suzy, 0)\}$$

We can view the universal set, X, as the set with the membership function $X(x) = 1$, $\forall x \in X$. The empty set as the set with membership value $\emptyset(x) = 0 \ \forall x \in X$.

1.6 Fuzzy Set

We have previously seen that a crisp set corresponds to a function whose codomain is $\{0, 1\}$. In a similar fashion, we define a fuzzy set by exchanging the codomain of $\{0, 1\}$ with the interval $[0, 1]$.

Definition 1.6. Fuzzy Set
Let X be the universal set. The object A is a fuzzy subset of X if A is a function $A : X \to [0, 1]$.

In the next chapter, we explore how this definition of a set, in combination with new set operation definitions, is compatible with crisp set theory definitions using fuzzy notation. As such, this new definition of a subset is not a new set theory, but rather an expansion of our traditional set theory. Hence, this new definition gives us the ability to model intuitive concepts, which are traditionally problematic for crisp sets, without loosing any results that have been obtained by crisp set theory.

Example 1.7. Consider Example 1.5 of crisp sets above, with the same universal set and heights for the elements. In this example, we clearly defined a binary partition of the elements such that each person is either taller than 6 ft. or not. Now, instead of collecting all people over 6 ft. in height, let us construct the set of people that are tall. To do this, we could define a cutoff based on the writers intuition. For example, suppose we make a cutoff for being tall at 5' 8". In this case, we have a means to assign a membership value of 1 to people taller than this height. However, what do we do with Michael? He is 5' 7.99" in height. We can continue to "round off" people's membership values to 1 and 0, and in the case of Michael assign a value of 0. However, this would not accurately model the task at hand. We are trying to classify based on a vague, fuzzy concept, "tall". Suppose we are constructing this list for a casting committee in order to hire a person for a role requiring a certain height. Most humans would give Michael a membership value of 1 for this purpose. This is despite the fact he is below the assigned threshold of 5'8". In this way, we have temporarily resolved our dilemma. We thus force a binary classification (in TALL or not in TALL); we do this either based on artificial mathematical measurement, such as inches, or by general understanding of task at hand, in which case we count Michael as "close enough". This "close enough" is where fuzzy logic was used, whether we refer to it as such or not. The only way to not use fuzzy, and still have Michael in the set, would be to lower the threshold to a threshold Michael would crisply meet. Let us say, for this example, a threshold of 5' 7.99".

While this ends the logical inconsistencies for this example, what if we had a person of 5" 7.98"? Would we still give them a 1 for a membership value? If not, why? They are as close to the new threshold as Michael was to the original threshold. If we do assign them a 1, then how about a person of 5' 7.97"?, 5' 7.96" ? ,5' 7.95" ? , 5' 7.94"? Do you continue to give each new slightly shorter person a value of 1, until the universal set, or hiring pool, is classified as tall? In such a case, we yield at the Sorites paradox again, and the classification looses all meaning. The fact remains, if the purpose of this exercise is to construct a set of tall people, it is impossible to create a crisp method of assigning a 1 or 0 to applicants, which is both applicably effective, and universally agreed upon regardless of the person assigning the membership values.

On another point, consider two elements of our set, Barbara and Carmen. Barbara has height of 6'1", and Carmen has a height of 6'2". In common speech, we would say that Carmen is taller than Barbara. However according to the crisp, 0 vs 1 assignment, they both have a membership value of 1. We thus loose any information to compare these two with a binary construction. All the information retained is that they are both taller than 5'8". An observer of our construction would not be able to discern who was taller than the other.

Our answer to both of the previous dilemmas is to assign fuzzy membership values to the elements of the universal set. For the set $T = \{x \in X \mid \text{x is tall}\}$, we may assign a membership value to Michael of 0.8, Barbara of 0.95, and Carmen 0.96, and so on. In this way we may construct the full fuzzy set of tall people.

$$T = \{(John, 0.2), (Barbara, 0.95), (Stacey, 0.1), (Carmen, 0.96), (Michael, 0.8), (Suzy, 0.87)\}$$

In this fashion, we retain information to compare items within the set. We also have created a clear way of defining membership values such that the membership values of elements is more uniform among those tasked with assigning membership values. That is, the discrepancy of membership values can be minimized. This discrepancy is often modeled by a form of a sum-of-squares function, where

$$Discrepancy = \sum_{x \in U} (A_i(x) - A_j(x))^2$$

such that A_i and A_j represent two constructions of a fuzzy set based on a fuzzy quantifier.

Example 1.8. Membership Assignment by Rule: Consider the set S of small, non-negative integers. We can define the membership values of S by a rule, such that;

$$S(n) = \begin{cases} 1 & n = 0 \\ \min\{1, \frac{1+n}{n^2}\} & n \geq 1 \end{cases}$$

Thus we see,

$$S = \{(0, 1), (1, 1), (2, 0.75), (3, 0.44), (4, 0.31), (5, 0.24), (6, 0.19), (7, 0.16), (8, 0.14), ...\}$$

Example 1.9. A Smart Fuzzy Washing Machine:

A smart washing machine needs to determine whether to extend the wash cycle for the load of clothes it has been washing, and for how long. To do this, we define a set $D = \{b \mid \text{b is a dirty load of clothes}\}$, and we define a function d such that $d(x_b) = D(b)$ Where $d(x_b) = \frac{1}{\pi}tan^{-1}(x_b - 5) + \frac{1}{2}$ and x_b is the concentration of impurities, measured in grams per liter, of the water the machine has most recently drained from the load, b. The graph of this function is displayed below.

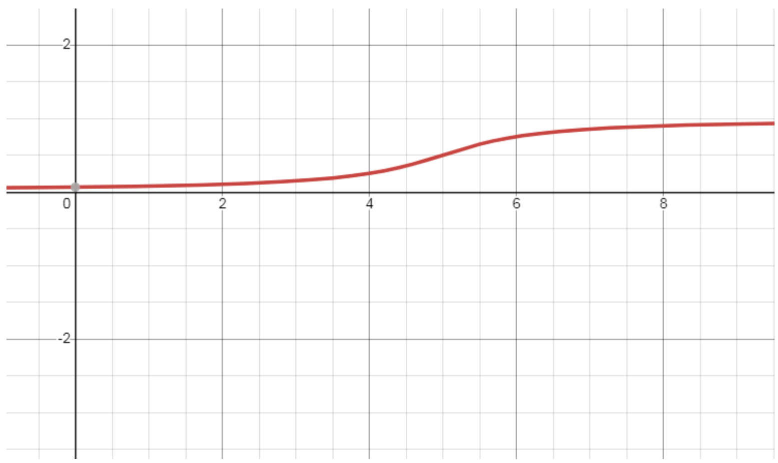

Figure 1: $d(x_b) = \frac{1}{\pi}tan^{-1}(x_b - 5) + \frac{1}{2}$

Notice that as x_b increases, $D(b)$ approaches 1. If the load, b_1, has $x_{b_1} = 10$, then we assign $D(b_1) = 0.94$. Therefore, the machine can then interpret that the load is 94% dirty. In this way, we then perform a secondary wash cycle, whose time duration is established by a defuzzifying time function, such as $T(b) = max\{(D(b) - 0.2), 0\} \cdot 45$. The machine measures the water flow continually throughout the wash cycle, and makes adjustments to the wash time as needed. It does so until the membership value of b is below a specified threshold, say, 0.2. In fact, a system of fuzzy sets and defuzzifying functions may be created so as to account for more variables than just water impurities and cycle time. In so doing, we optimize resources such as water and electricity usage.

This explanation is highly oversimplified. A fuzzy washing machine was one of the first consumer applications of fuzzy logic, and achieved far better efficiency than any of its competitors at the time. This washing machine works by way of a fuzzy associative memory system, or FAM [21]. This FAM integrates over several values, each with a specified fuzzy membership. We then use this "centroid" to make a non-fuzzy decision.

1.7 Moving Forward

At this point, we reiterate that some of the language in this paper is not typical for traditional mathematical writing. Words such as, "very", "really" or "kinda" (and "kinda" is not even a word) are vague, and often avoided in traditional mathematics. This is because, for traditional mathematics, we strive for certain, logically bivalent answers to difficult questions. Yet, we must be able to use these words in mathematics, because they are words that we as humans understand as real concepts. This is the driving force behind our study of fuzziness.

The remainder of this paper aims to give a comprehensive overview of the most core studies of fuzzy logic. The topics of the remaining chapters include:

Ch.	Fuzzy Topic	Inspiring Question to Model	Research Papers
2	Fuzzy Set Operations, Partitions and Similarity Relations	"To what degree is x both fast and large?" and "How similar is 3 to 7?"	[32] [43] [35]
3	T-norms Negations, T-conorms (generalized fuzzy set operations)	"How much like a deer, and like a dog, is the image, based on a vector of fuzzy truth values?"	[35] [17]
4	Fuzzy Groupoids, Fuzzy Groups, Fuzzy Quotient Groups	"How well does an element stay in a group?"	[26] [43] [12] [1]
5	Fuzzy Metric Spaces	"How certain are we two items are within 5 units apart?"	[49] [13] [14]
6	Specificity Measures and Fuzzy (Non-additive) Measures	"How crisp is a fuzzy set?", and "If we are more than the sum of our parts, how might we measure our parts?"	[48] [47] [20] [25]
7	Fuzzy Additive Systems	"Computer, tell me if this is a goat."	[23] [36]

Table 1: Fuzzy Topics of Each Chapter

CHAPTER 2

Fuzzy Set Theory

Let us now build on our core understanding of a fuzzy set to form a fuzzy set theory. We do this following the model of crisp set theory. However, traditional definitions of set operations do not accept fuzzy sets as their inputs. Therefore, we redefine these set operations, and hence extend our traditional set theory to be defined for fuzzy sets.

While redefining these operations, we wish to inherit the results of crisp set theory. To do this, we must ensure that, if they are limited to act on crisp sets, these new definitions are equivalent to traditional set theory definitions. In this way, we construct a fuzzy set theory which is not a contradictory set theory, but rather an extension of crisp set theory. We begin this process by presenting the following fundamental fuzzy set theory definitions (definition 2.1), and consequently explore these definitions in detail with examples. Additionally, we continue throughout this chapter by showing that each of these definitions, if restricted to inputs of crisp sets, yield equivalent results to the those reached by the definitions we know from crisp set theory.

2.1 Fuzzy Set Operations

Recall that the membership of an element in a fuzzy set is denoted with function notation. To illustrate, when we use the notation, A, we refer to the fuzzy set itself (the function from X into $[0, 1]$). However, when using the notation $A(x)$, we refer to the single number in the interval $[0, 1]$ which quantifies the membership of x in A.

Definition 2.1. (Walker, 2018) [35] Let A and B be fuzzy subsets of a crisp set X.

(i)Fuzzy Subset: We say A is a subset of B iff for every element, x in X, x has membership in B at least as strong as the membership of x in A.

$A \subseteq B$ iff $A(x) \leq B(x)\ \forall x \in X$

(ii) Fuzzy Set Equality: We consider two fuzzy subsets of X to be equal iff for all elements x in X, x has equal membership in A as B. That is, $A = B$ iff $A(x) = B(x)\ \forall x \in X$. Equivalently, $A = B$ iff $A \subseteq\supseteq B$.

(iii) Fuzzy Power Set: We denote the crisp set of all fuzzy subsets of X as $[0,1]^X$; that is, $[0,1]^X = \{A \mid A \subseteq X\}$

(iv) Fuzzy Set Intersection: The intersection of two fuzzy sets can be defined using a variety of operations meeting certain qualifications (see Ch.3; T-norms). The most standard operation used is the min operation. That is, the intersection of two fuzzy sets A and B is a new fuzzy set who's membership values are defined as ,
$(A \cap B)(x) = min\{A(x), B(x)\}\ \forall x \in X$.

(v) Fuzzy Set Union: The union of two fuzzy sets can similarly be defined using a variety of operations meeting certain qualifications (see Ch.3; T-conorms). The most standard operation used is the max operation. That is to say that the union of two fuzzy sets A and B is a new fuzzy set who's membership values are determined as
$(A \cup B)(x) = max\{A(x), B(x)\}\ \forall x \in X$

(vi) Fuzzy Set Complement: The set complement of a fuzzy set can similarly be defined by any of a class of unary operations meeting certain qualifications (Ch. 3; Negations). Most commonly, The set complement of A has membership values defined by $A^c(x) = 1 - A(x)$, $\forall x \in X$.

(vii) Fuzzy Set Difference: We define fuzzy set difference in the same way we do in crisp set theory, as the intersection of a set and the complement of a second set. By using the above definitions of intersection and complement, this yields an intuitive definition of set difference, as applied to fuzzy sets. For an element x in X,
$(A \sim B)(x) = (A \cap B^c)(x) = \min\{A(x), 1 - B(x)\}$.

(viii) Bounded Difference: Similarly, the fuzzy bounded difference is denoted $A \ominus B$ and its membership value for element x is defined by $(A \ominus B)(x) = \max\{0, A(x) - B(x)\}$. When applied to crisp sets, this operation is equivalent to set difference.

Example 2.2. To illustrate an intuition of the fuzzy subset definition, allow us to consider the fuzzy sets Y, of young people, and A, of adolescents. In crisp set theory, we would round off the membership values of people for these sets, and would say that $A \subseteq Y$. However, the same holds when we maintain the fuzziness of A and Y.

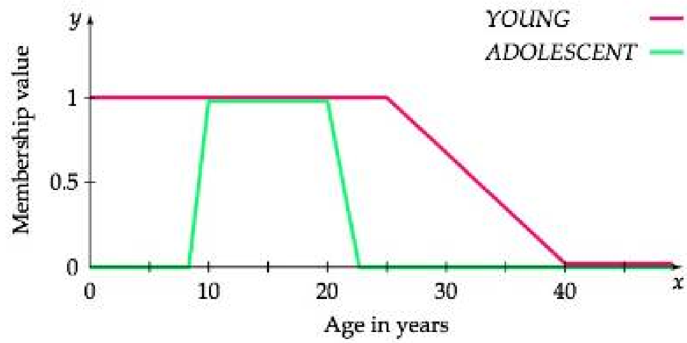

Figure 2: Young $\supseteq$ Adolescent: (Aufman, 2017, pg.63)[5]

From this figure, we see the more natural membership functions for A and Y. As most would describe it, one does not jump from being 100% young to 100% old. Thus a gradual decrease in youthfulness is shown. Similarly we see that most would consider an adolescent to be someone around the ages of 10 to 20, with some disagreement, or fuzziness, around the edges. The change in the membership function of A is steeper than that of Y, but it is still not immediate, thus the graph is still not crisp. Also, from this graph, we see that A is indeed a fuzzy subset of Y, as for any element of the domain x(age), $A(x) \leq Y(x)$. In the terminology of functions, a fuzzy set A is a fuzzy subset of B iff the function B dominates the function A.

Example 2.3. To illustrate a non-example, consider the fuzzy sets of fast computers, F, and computers with large memory, M, where F and M are fuzzy subsets of the universal set of computer models, $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$. We have listed the membership values of F and M below.

	F	M
x_1	0.4	0.6
x_2	0.8	0.5
x_3	0.9	0.5
x_4	0.3	0.45
x_5	0.55	0.7
x_6	0.8	0.7

Table 2: Fast and Large-Memory Computers

It is evident through examination of the membership values of any element of X, and the definition of fuzzy set equality, that these two sets are not equal. Exemplar; $F(x_1) = 0.4 \neq 0.6 = M(x_1)$. Therefore $F \neq M$ by definition of fuzzy set equality.

Additionally, neither set is a fuzzy subset of the other. Consider elements x_1 and x_2. $F(x_1) < M(x_1)$, thus M cannot be a fuzzy subset of F by definition of fuzzy subsethood. Also, $F(x_2) > M(x_2)$, thus F cannot be a fuzzy subset of M.

Now let us consider the set of computers which are both fast, and have large memory. We may model this set with a set intersection, just as we would in crisp logic. We calculate the intersection of fuzzy sets by taking the minimum of each element's membership value in the respective sets. For example,
$(F \cap M)(x_1) = \min\{F(x_1), M(x_1)\} = \min\{0.4, 0.6\} = 0.4$. The remaining membership values are calculated in the same manner. We have listed these membership values using the coordinate pair notation, introduced in chapter 1.

$$F \cap M = \{(x_1, 0.4), (x_2, 0.5), (x_3, 0.5), (x_4, 0.3), (x_5, 0.55), (x_6, 0.7)\}$$

Alternatively, the set of computers which are fast or have large memory, may be modeled with a set union, similar to what we would expect from crisp logic. We derive the union of two fuzzy sets by taking the maximum of the membership value of each element. We thus find that

$$F \cup M = \{(x_1, 0.6), (x_2, 0.8), (x_3, 0.9), (x_4, 0.45), (x_5, 0.7), (x_6, 0.8)\}$$

Now let us consider the complement of a fuzzy set. The set of computers that are not fast may be modeled by the complement of F, denoted F^c. We calculate the membership values of the elements of this set by using the negation function $F^c(x) = 1 - F(x)$. These membership values correlate inversely with the degree to which a computer is considered "fast".

$$F^c(x) = \{(x_1, 0.6), (x_2, 0.2), (x_3, 0.1), (x_4, 0.7), (x_5, 0.45), (x_6, 0.2)\}$$

Remark 2.4. We are careful to refer to this set as the set of "not fast" computers, and do not necessarily call it the set of "slow" computers. It is sometimes the case in fuzzy logic that opposite classifiers, such as fast and slow, are not necessarily modeled by set complement. That is, it may be the case that $F^c \neq S$ and $S^c \neq F$. The likelihood that a person would call a computer "fast" may not be equal to the likelihood a person may refer to a computer as "not slow". When this is the case, we refer to this model as an intuitionistic fuzzy set (Atanassov, 1986) [6].

Remark 2.5. We may also model a set of fast computers without large memory. However, we may use two methods to model this, set difference and bounded set difference. In crisp logic, when all truth values are of $\{0, 1\}$, these two operations are equivalent. However, in fuzzy logic, $A \sim B$ is not necessarily equal to $A \ominus B$. This is evident through the membership values of x_1 in the above example.
Notice $(F \ominus M)(x_1) = \max\{0, F(x_1) - M(x_1)\} = \max\{0, -0.2\} = 0$.
However, $(F \sim M)(x_1) = (F \cap M^c)(x_1) = \min\{F(x_1), 1 - M(x_1)\} = \min\{0.4, 0.4\} = 0.4$. Thus, $(F \sim M)(x_1) \neq (F \ominus M)(x_1)$, and hence, by definition of fuzzy set equality, $F \sim M \neq F \ominus M$.

We wish to prove that if we restrict the membership values of a fuzzy set A and B to values of 0 and 1 (crisp sets), then our definitions of fuzzy set difference and fuzzy bounded difference do indeed yield identical results. That is, we need prove fuzzy set equality, $A \sim B = A \ominus B$. Recall that in the context of this chapter's set definitions, we treat sets as functions from a universal set of elements, X, into a set of truth values. When restricted to crisp sets, this set of truth values is limited to $\{0, 1\}$. Thus proving fuzzy set equality is equivalent to proving function equality. That is, we need prove that any arbitrary element in the universal set, this element has equal membership in both fuzzy sets. Often when restricting ourselves to crisp sets, this proof follows a case structure by the finite set of possible combinations of truth values any element may hold. We illustrate this in the following proposition.

Proposition 2.6. Let X be a universal set, and let $A, B \in [0, 1]^X$. If A and B are crisp sets, then $A \ominus B = A \sim B$

Proof. Let A and B be fuzzy subsets of universal set X. Let $x \in X$. We need show $(A \ominus B)(x) = (A \sim B)(x)$.

Case 1) Let $A(x) = B(x) = 0$. Therefore, the values of $(A \ominus B)(x)$ and $(A \sim B)(x)$ are determined as, $(A \ominus B)(x) = \max\{0, A(x) - B(x)\} = 0$ and $(A \sim B)(x) = (A \cap B^c)(x) = \min\{A(x), 1 - B(x)\} = \min\{0, 1\} = 0$

Case 2) Let $A(x) = B(x) = 1$. Therefore, the values of $(A \ominus B)(x)$ and $(A \sim B)(x)$ are $(A \ominus B)(x) = \max\{0, 1 - 1\} = 0$ and $(A \sim B)(x) = (A \cap B^c)(x) =$
$\min\{A(x), 1 - B(x)\} = \min\{1, 1 - 1\} = 0$.

Case 3) Let $A(x) = 1,\ B(x) = 0$. Therefore, the values of $(A \ominus B)(x)$ and $(A \sim B)(x)$ are, $(A \ominus B)(x) = \max\{0, A(x) - B(x)\} = \max\{0, 1 - 0\} = 1$ and $(A \sim B)(x) = (A \cap B^c)(x) = \min\{1, 1 - B(x)\} = \min\{1, 1\} = 1$

Case 4) Let $A(x) = 0,\ B(x) = 1$. Therefore, the values of $(A \ominus B)(x)$ and $(A \sim B)(x)$ are, $(A \ominus B)(x) = \max\{0, A(x) - B(x)\} = \max\{0, -1\} = 0$ and $(A \sim B)(x) = (A \cap B^c)(x) = \min\{A(x), 1 - B(x)\} = \min\{0, 1 - 1\} = 0$.

Thus, for any case of membership values of x in A and B, we yield equality of membership values of x in $A \ominus B$ and $A \sim B$. Because this equality holds for arbitrary $x \in X$, by the definition of set equality, $A \ominus B = A \sim B$.

$\square$

Following a similar pattern to the previous proof, we may show that the remaining fuzzy set theory definitions are equivalent to their crisp set theory counterparts whenever the truth values of elements are limited to $\{0, 1\}$. Let us consider a definition of subset from crisp set theory. Let us then limit our definition of subsethood of fuzzy sets to $\{0, 1\}$. We thus have two equivalent definitions. Suppose $A \subseteq B$ in the fuzzy sense, and every element of A and B has membership 0 or 1. If $A(x) \leq B(x)$ for all $x \in X$, then $(A(x) = 1) \implies (B(x) = 1)$. Equivalently, in crisp terminology, this translates to, $(x \in A) \implies (x \in B)$. Thus $A \subseteq B$ in the crisp sense, when supposing subsethood in the fuzzy sense. Conversely, using the same reasoning, if $(A \subseteq B)$ in the crisp set, then $(A \subseteq B)$ in the fuzzy sense.

Now suppose that for crisp subsets of X, A and B, that $A(x) = B(x)\ \forall x \in X$. That is, $A = B$ in the fuzzy sense. Let $x \in A$. Thus $A(x) = 1$, and hence $B(x) = 1$. Therefore, $x \in B$. Thus $A \subseteq B$ in the crisp sense. Similarly, we see that $B \subseteq A$ in the crisp sense, and thus $A = B$ in the crisp sense. Therefore, if $A = B$ using the fuzzy definition, then $A = B$ using the crisp definition.

Through a case proof, as in the proof of proposition 2.6, we may verify that fuzzy set intersection, union, and complement definitions are equivalent to their corresponding crisp definitions if the truth values of elements are limited to truth values of crisp sets. Hence, fuzzy set theory is not a distinct set theory; rather, it contains the original crisp set theory within the fuzziness; $Crisp \subset Fuzzy$. It is this mantra that we maintain throughout this paper.

2.2 Fuzzy subsets with Adverb Modifiers

Let us now examine examples of fuzzy subset operations through the use of vague modifiers. It is possible when using fuzzy sets, to create a decreasing (increasing) sequence of fuzzy sets, in which each set in the sequence is a fuzzy subset (superset) of the previous set in the sequence. This sequence occurs whenever we apply adverbs such as "slightly", "very", and "extremely", with fuzzy classifiers, such as "hot", "tall", or "fast", for which an original fuzzy set is defined to model.

Example 2.7. Consider a new universal set of computers, X, and the set F_1, the set of "fast" computers. Now consider the set F_2 of "very fast computers". These three sets and each elements membership value are listed below.

X	speed in GHz	$F_1(x)$	$F_2(x)$
Adobe	3.2	0.2	0.05
Microsoft	4.5	0.85	0.70
Apple	4.7	0.95	0.85
Bojangle	1.9	0.05	0.001

Table 3: Fast and Faster Computers

In this context, we see that $F_2 \subseteq F_1$, as for every computer, the membership value for the computer in F_2 is less than or equal to that in F_1. We can of course, use this modifier again to construct a fuzzy set F_3 of "very, very fast" computers. We can then continue the process to create a decreasing sequence of fuzzy sets, $\{F_i\}$, by order of subsethood. These are sets of ever faster computers, which converge ultimately to the empty set.

2.3 Laws of Fuzzy Set Theory

Using traditional definitions from set theory, mathematicians have derived a number of set laws commonly used in all branches of mathematics. We have previously illustrated the equivalence of fuzzy set theory definitions to those of crisp set theory when these definitions are restricted to crisp sets. Therefore, we infer all known laws of set theory hold in fuzzy set theory when limited to the extremal cases of crispness. We need determine which laws hold for true fuzzy sets, and which laws are only true when restricted to crisp sets.

2.3.1 Distributive Laws of Fuzzy Union and Intersection

We begin with the distributive laws of set union and intersection. Recall that for crisp sets A, B, and C, we know that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Proposition 2.8. Distributive Laws of Union and Intersection hold for fuzzy sets.

Proof. Let A, B and C be fuzzy subsets of the universal set, X. Let $x \in X$. We need show $A \cap (B \cup C)(x) = ((A \cap B) \cup (A \cap C))(x)$.

By fuzzy definitions of union and complement,
$(A \cap (B \cup C))(x) = \min\{A(x), (B \cup C)\} = \min\{A(x), \max\{B(x), C(x)\}\}$. Similarly,
$((A \cap B) \cup (A \cap C))(x) = \max\{(A \cap B)(x), (A \cap C)(x)\} = \max\{\min\{A(x), B(x)\}, \min\{A(x), C(x)\}\}$

Case 1) Suppose $A(x) \leq B(x) \leq C(x)$.
Thus, $\min\{A(x), \max\{B(x), C(x)\}\} = \min\{A(x), C(x)\} = A(x)$.
Additionally, $\max\{\min\{A(x), B(x)\}, \min\{A(x), C(x)\}\} = \max\{A(x), A(x)\} = A(x)$.
Therefore, $(A \cap (B \cup C))(x) = ((A \cap B) \cup (A \cap C))(x)$. (The case $A(x) \leq C(x) \leq B(x)$ is equivalent by use of the symmetric property of the max operation).

Case 2) Suppose $B(x) \leq C(x) \leq A(x)$
Thus, $\min\{A(x), \max\{B(x), C(x)\}\} = \min\{A(x), C(x)\} = C(x)$.
Additionally, $\max\{\min\{A(x), B(x)\}, \min\{A(x), C(x)\}\} = \max\{B(x), C(x)\} = C(x)$.

Therefore, $(A\cap(B\cup C))(x) = ((A\cap B)\cup(A\cap C))(x)$. (The case $C(x) \leq B(x) \leq A(x)$ is equivalent by use of the symmetric property of the max operation).

Case 3) Suppose $C(x) \leq A(x) \leq B(x)$.
Thus, $\min\{A(x), \max\{B(x), C(x)\}\} = \min\{A(x), B(x)\} = A(x)$.
Additionally, $\max\{\min\{A(x), B(x)\}, \min\{A(x), C(x)\}\} = \max\{A(x), C(x)\} = A(x)$.
Therefore, $(A\cap(B\cup C))(x) = ((A\cap B)\cup(A\cap C))(x)$. (The case $B(x) \leq A(x) \leq C(x)$ is equivalent by use of the symmetric property of the max operation).

Consequently, in all cases $(A \cap (B \cup C))(x) = ((A \cap B) \cup (A \cap C))(x)$. This is for all $x \in X$, thus we have proven fuzzy set equality. The proof for $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ is nearly identical to this. We thus conclude that distributive laws of set intersection and union hold for fuzzy sets.

$\square$

2.3.2 Demorgan's Laws

In addition to the distributive laws, DeMorgan's Laws are some of the most core to crisp set theory, and serve as a base on which many other results are built. Consequently, we wish to be able to extend these rules to our construction of a comparable fuzzy set theory. Thus, using our fuzzy definitions for fuzzy intersection, union, complement, and fuzzy set equality, let us show that, for a universal set X, and $A, B \in [0,1]^X$, $(A \cap B)^c = A^c \cup B^c$ and $(A \cup B)^c = A^c \cap B^c$.

Proposition 2.9. Demorgan's Laws hold for true fuzzy sets.

Proof. Let X be a universal set, and $A, B \in [0,1]^X$. Let $x \in X$. We need show $((A \cap B)^c)(x) = (A^c \cup B^c)(x)$.

$$
\begin{aligned}
((A \cap B)^c)(x) &= 1 - (A \cap B)(x) && \textbf{definition of set complement} \\
&= 1 - \min\{A(x), B(x)\} && \textbf{definition of set intersection} \\
&= 1 + \max\{-A(x), -B(x)\} && \textbf{order reversing property of multiplication by -1} \\
&= \max\{1 - A(x), 1 - B(x)\} && \textbf{translation invariance of order } \leq \\
&= \max\{A^c(x), B^c(x)\} && \textbf{definition of set complement} \\
&= (A^c \cup B^c)(x) && \textbf{definition of set union}
\end{aligned}
$$

Thus, $((A \cap B)^c)(x) = (A^c \cup B^c)(x)$. As this equality holds for arbitrary $x \in X$, we have shown fuzzy set equality, $(A \cap B)^c = A^c \cup B^c$.

□

Remark 2.10. One may prove Demorgan's second law in a similar manner. However, we may also prove this law using Demorgan's first law, in combination of the involution property of complement, $((A^c)^c = A)$. To show this, let $x \in X$, and $A \in [0,1]^X$. We see that $(A^c)^c(x) = 1 - A^c(x) = 1 - (1 - A(x)) = A(x)$. Thus $(A^c)^c = A$ for any fuzzy set $A \in [0,1]^X$. Thus the involution property is verified for the standard fuzzy complement defined in definition 2.1.

Now, using Demorgan's first law on fuzzy sets A^c and B^c, we see $(A^c \cap B^c)^c = (A^c)^c \cup (B^c)^c = A \cup B$. Thus, by taking the complement of both sides, we see $((A^c \cap B^c)^c)^c = (A \cup B)^c$. We then invoke the involution property again, and yield $A^c \cap B^c = (A \cup B)^c$. Thus Demorgan's second law holds for true fuzzy sets.

2.4 Broken Laws; Extremal Laws

Some rules of crisp set theory hold only for truly crisp sets. That is, some set equations hold for both crisp and fuzzy definitions only if inputs are limited to crisp sets. This does not nullify our claim that $Crisp \subset Fuzzy$, but rather illustrates the principle that crispness is a special case of fuzziness, and as such holds special properties.

Proposition 2.11. The following laws hold true for crisp sets, using either the crisp or fuzzy set operation definitions. However, these rules do not hold for truly fuzzy sets.

For $A, B \in [0,1]^X$,

(i) $A \sim B = A \ominus B$

(ii) $A \cup A^c = X$

(iii) $A \cap A^c = \emptyset$ (Law of Excluded Middle)

We have illustrated in Remark 2.5, that the first equation does not hold for truly fuzzy sets. Furthermore, through counterexamples in Example 2.12 below, we see that (ii) and (iii) also do not hold for true fuzzy sets.

Example 2.12. Consider, $X = \{x\}$, and a fuzzy subset $A = \{(x, 0.7)\}$. Thus $A^c = \{x, 0.3\}$. For $A \cup A^c = X$, we would expect $(A \cup A^c)(x) = 1$. However, using the fuzzy definition of union, we see that $(A \cup A^c)(x) = \max\{A(x), A^c(x)\} = \max\{0.7, 0.3\} = 0.7 \neq 1$. Thus, $A \cup A^c \neq X$.

Similarly, consider (iii), the Law of Excluded Middle. For the same set A above, $(A \cap A^c)(x) = \min\{0.7, 0.3\} = 0.3 \neq 0 = \emptyset(x)$. Thus for proper fuzzy sets, the law of excluded middle does not hold. Of course, laws (ii) and (iii) do hold for crisp set theory, and thus, when using equivalent fuzzy definitions on crisp sets, these laws hold.

To show this, let A be a crisp subset of universal set X, and let $x \in X$. Then either ($A(x) = 1$ and $A^c(x) = 0$) or ($A(x) = 0$ and $A^c(x) = 1$). Therefore $\max\{A(x), A^c(x)\} = \max\{1, 0\} = 1$. Hence $(A \cup A^c)(x) = 1 = X(x)$ for all $x \in X$. Thus, we have shown $A \cup A^c = X$.

Additionally, we see that $\min\{A(x), A^c(x)\} = \min\{1, 0\} = 0$. Therefore, $(A \cap A^c)(x) = 0 = \emptyset(x)$ for all $x \in X$. Thus, we have shown $(A \cap A^c) = \emptyset$.

Thus, we further justify our assertion that any fuzzy set definition, if applied to only crisp sets, yields identical results as to that of traditional crisp set theory. Fuzzy logic is not a reinvention of the set theory wheel, but rather an extension of that wheel, in order to include the ability to model vagueness. Our final section further exemplifies this principle.

2.5 Partitions and Similarity

In this section, we explore a fuzzification of the crisp topic, equivalence relations. Rather than an element being in an equivalence class, we now speak about an element being in a similarity class. A fuzzy equivalence relation, or similarity relation, models the vague concept, "These elements are similar, or comparable". This is in contrast to the crisp counterpart, "These elements are identical, or equivalent". In crisp set theory, the existence of an equivalence relation is tandem to the existence of a partition. We are able to maintain this principle through a partition of a set into a collection of fuzzy subsets.

We begin by defining a fuzzy partition of a crisp set. It is worth noting that there are many other ways to define a more fuzzy, or "weak" fuzzy partition of sets. Several of these are explored by M. Shakhatreh. [46]. The following is an interpretation from (Morsi, 1994) [32] of a definition introduced in (Ochinnikov, 1991) [37].

Definition 2.13. Fuzzy Partition: Let $P = \{P_i\}$ be a collection of fuzzy subsets of universal set X. We call P a fuzzy partition of X iff

(i) Every element, P_i, of P has at least one element, x of X which fully belongs to P_i. That is, $\forall P_i \in P$, $\exists x \in X \ s.t. P_i(x) = 1$.

(ii) Every element of X fully belongs to exactly one element of P. That is, $\forall x \in X$, $\exists! \ P_i \in P$ such that $P_i(x) = 1$.

(iii) The elements of P are linked in that, the intersection of any two elements of P achieves maximal degree at the two points which fully belong to each set respectively.

Formally, for $P_i, P_j \in P$, and $x, y \in X$, such that $P_i(x) = P_j(y) = 1$, then $P_i(y) = P_j(x) = \sup_{z \in X}(P_i \cap P_j)(z)$

Remark 2.14. In crisp logic, (i) and (ii) constitute a partition. This is because in crisp logic, (i) translates into "Every element of P is non-empty". Additionally, (ii) translates into, "Every element of X belongs to exactly one element of P. Equivalently, the intersection of any two distinct elements of P is the empty set, and the union of all such P_i is the universal set".

Example 2.15. Consider the real numbers, and let us consider four fuzzy subsets of $\mathbb{R}$, named E, O, Q, and I, with membership values defined below $\forall x \in \mathbb{R}$.

$$E(x) = \begin{cases} 1 & x \textbf{ is an even integer} \\ 0.5 & x \textbf{ is an odd integer} \\ 0.3 & x \textbf{ is a rational, non-integer} \\ 0 & x \textbf{ is irrational} \end{cases}$$

$$O(x) = \begin{cases} 0.5 & x \textbf{ is an even integer} \\ 1 & x \textbf{ is an odd integer} \\ 0.3 & x \textbf{ is a rational, non-integer} \\ 0 & x \textbf{ is irrational} \end{cases}$$

$$Q(x) = \begin{cases} 0.3 & x \textbf{ is an even integer} \\ 0.3 & x \textbf{ is an odd integer} \\ 1 & x \textbf{ is a rational, non-integer} \\ 0 & x \textbf{ is irrational} \end{cases}$$

$$I(x) = \begin{cases} 0 & x \textbf{ is an even integer} \\ 0 & x \textbf{ is an odd integer} \\ 0 & x \textbf{ is a rational non-integer} \\ 1 & x \textbf{ is irrational} \end{cases}$$

Notice then that each of the sets above, there exist elements of $\mathbb{R}$ which has membership of 1 in that set (for example, $2, 3, \frac{2}{3}, \pi$ respectively). Therefore, condition (i) is met. Additionally, for any $x \in \mathbb{R}$, x has membership of 1 in exactly one of the above fuzzy sets, thus, condition (ii) is met.

Condition (iii) is slightly less-trivial to show, as we must consider an arbitrary intersection of two of the created sets. For this purpose, let $a, b, c, d \in \mathbb{R}$ such that

a is an even integer, b is an odd integer, c is a rational non-integer, and d is an irrational number. Thus $E(a) = O(b) = Q(c) = I(d) = 1$. We thus determine each of the following cases.

Case 1) $\bigvee_{x \in \mathbb{R}} (E \cap O)(x) = 0.5 = E(b) = O(a)$

Case 2) $\bigvee_{x \in \mathbb{R}} (E \cap Q)(x) = 0.3 = E(c) = Q(a)$

Case 3) $\bigvee_{x \in \mathbb{R}} (E \cap I)(x) = 0 = E(d) = I(a)$

Case 4) $\bigvee_{x \in \mathbb{R}} (O \cap Q)(x) = 0.3 = O(c) = Q(b)$

Case 5) $\bigvee_{x \in \mathbb{R}} (O \cap I)(x) = 0 = O(d) = I(b)$

Case 6) $\bigvee_{x \in \mathbb{R}} (Q \cap I)(x) = 0 = Q(c) = I(a)$

Thus, in all cases, condition (iii) holds. We hence conclude that $\{E, O, Q, I\}$ is a fuzzy partition of $\mathbb{R}$.

We see now the extension of a partition to fuzzy logic. The famous corresponding topic is of course, equivalence relations. A fuzzy equivalence relation, however, is not quite as straightforward. We wish to define such a relation, and show that the existence of one is equivalent to the existence of a partition, just as in crisp set theory. Recall, these fuzzy equivalence relations, or called similarity relations, model the phrase, "These elements are similar". To define such similarity relations, we need first define a fuzzy relation.

Definition 2.16. (Zimmerman, 1991)[53] Fuzzy Relation of a Fuzzy Set. Let $A \in [0,1]^X$ and $B \in [0,1]^Y$, for universal sets X, Y. $R : X \times Y \to [0,1]$ is a fuzzy relation on between A and B, iff $R(x,y) \leq \min\{A(x), B(y)\}$, for all $x \in X$ and $y \in Y$.

Note: If R is a fuzzy relation between A and A, we yield an equivalent definition to a fuzzy graph (soft graph). It is on this structure that the entire study of fuzzy graph theory is built. [30].

Example 2.17. Suppose that we are given a universal set, X, of people within an organization. Suppose x_1 knows x_2, and x_2 knows x_3, and so on. We recognize that in our social lives, there are different levels to which one can know a person. We would say that we 'know' a coworker that joined the company 2 weeks ago. However, we 'know' our cousin that works in the other department, more than we know the new coworker. We would hopefully say that we know our spouse more than any of our coworkers. We can model these degrees of 'knowing' as a fuzzy relation. Let S be a function from X^2 into $[0,1]$. Consider S to model the degree to which a person knows another. That is, we would say $S(x,y) = 0$ if x knows nothing of y, $S(x,y) = 0.3$ if x has met y recently, $S(x,y) = 0.6$ if x knows y well, $S(x,y) = 0.8$ if x is good friends with y, $S(x,y) = 0.9$ if x is a childhood friend or relative of y, and $S(x,x) = 1$. In this way, we have defined a fuzzy subset of the set $X \times X$, and thus R is a fuzzy relation from X to X, or simply, a fuzzy relation on X.

In this example, we used crisp universal set X. We can construct a fuzzy subset of X, A such that $A(x) = 1$ if x is a full-time employee, $A(x) = 0.5$ if x is a part time employee, and $A(x) = 0.2$ if x is an independent contractor of the company. In this way, we may construct another fuzzy relation S_1 on A such that
$S_1(x,y) = \min\{S(x,y), A(x), A(y)\}$, such that the inequality $S_1(x,y) \leq \min\{A(x), A(y)\}$ holds. Thus, we have shown that S_1 as defined above, is a function from $X \times X$ into $[0,1]$, and holds the inequality of a fuzzy relation. Therefore, S_1 is indeed a fuzzy relation on A.

The above is an example of a relation on A, and little more. In expanding this concept to a fuzzy similarity relation, we need define properties of fuzzy relations. These are twin to the properties of a crisp equivalence relation; reflexive, symmetric, and transitive.

Definition 2.18. Reflexive:(Zadeh, 1971) [52] The fuzzy relation R of a fuzzy set $A \in [0,1]^X$ is reflexive (or strong reflexive) iff $R(x,x) = 1, \ \forall x \in X$.

S on X, as defined in the previous example, is reflexive by its construction.

Definition 2.19. Weak Reflexive:(Yeh, 1975) [50] The fuzzy relation R of fuzzy set $A \in [0,1]^X$ is weak reflexive iff $R(x,y) \leq min\{R(x,x), R(y,y)\}\ \forall x, y \in X$.

S_1 on A above is weak reflexive, without necessarily being reflexive (strong reflexive). To illustrate, consider an independent contractor, z. Then $S_1(z,z) = \min\{S(x,x), A(z), A(z)\} = \min\{1, 0.2, 0.2\} = 0.2$. Thus S_1 is not strong reflexive. However, by its definition, for arbitrary $x, y \in X$, we see $S_1(x,x) = \min\{S(x,x), A(x), A(x)\} = \min\{1, A(x)\} = A(x)$. Therefore, we also note that $\min\{S(x,y), A(x), A(y)\} \leq A(x)$, and similarly, $\min\{S(x,y), A(x), A(y)\} \leq A(y)$. Thus,
$S_1(x,y) \leq \min\{A(x), A(y)\} = \min\{S_1(x,x), S_1(y,y)\}$. Therefore, S_1 is weak reflexive.

Remark 2.20. The weak form of fuzzy reflexive is mentioned here because it is necessary to expand the following results to relations on fuzzy subsets of a universal set. If a relation, R, exists on a fuzzy subset, A, of universal set, X, such that R is strong reflexive, thus $R(x,x) = 1 \forall x \in X$. Consequently, by the relation inequality, we see that $R(x,x) \leq A(x)\ \forall x \in X$. Thus, $A(x) = 1 = X(x) \forall x$, and hence, $A = X$. Thus, a relation which is strong reflexive can only exist on a crisp, universal set. In defining a fuzzy equivalence relation on any fuzzy (or even a crisp) proper subset of a universal set, this reflexive must be weakened to this secondary form of reflexive. However, for the purposes of this paper, we explore results of strong reflexive relations on the universal set.

Definition 2.21. Symmetric:(Zimmerman, 1991)[53] The fuzzy relation R on a fuzzy set $A \in [0,1]^X$ is symmetric iff $R(x,y) = R(y,x)$ for all $x, y \in X$

Both S and S_1 are symmetric by their construction.

Definition 2.22. Transitive : The fuzzy relation R on a fuzzy set $A \in [0,1]^X$ is transitive iff for any $x, z \in X$,
$\bigvee_{y \in X} \min\{R(x,y), R(y,z)\} \le R(x,z) \ \forall y \in X$.

Neither S nor S_1 may necessarily be transitive. For example, consider $x, y, z \in X$ such that y has a friendly relationship with both x and z, but x and z don't necessarily know each other in a typical social web. Thus these may not be transitive.

Definition 2.23. Similarity Relation: (Zimmerman, 1991) [53] A fuzzy relation R on a set X is a similarity relation on X if it is reflexive, symmetric, and transitive.

Remark 2.24. If we suppose R meets the definition of a similarity relation on X, and that the values of R are limited to $\{0,1\}$, then R is a crisp equivalence relation. Since R is fuzzy reflexive, we see that $R(x,x) = 1 \ \forall x \in X$. Therefore, in crisp notation, $xRx \ \forall x \in X$; R is crisp reflexive. Additionally, suppose that for some $x, y \in X$, that $R(x,y) = 1$. Therefore, by fuzzy symmetric property, $R(y,x) = 1$. Thus, in crisp notation, we have shown, $xRy \implies yRx$; R is crisp symemtric. Furthermore, suppose now that $x, y, z \in X$ such that $R(x,y) = R(y,z) = 1$. Thus, by fuzzy transitive property, $\bigvee_{w \in X} \min\{R(x,w), R(w,z)\} \le R(x,z)$. By choosing $w = y$, we see $1 = \min\{R(x,y), R(y,z)\} \le \bigvee_{w \in X} \min\{R(x,w), R(w,z)\} \le R(x,z)$. Thus, $R(x,z) = 1$. Thus, in crisp notation, we have shown, $xRy \wedge yRz \implies xRz$; R is crisp transitive.

Therefore, a similarity relation, when limited to crisp truth values, is a crisp equivalence relation. Consequently, we know that a similarity relation, limited to crisp truth values, correlates to a crisp partition. We need show that an unrestricted fuzzy similarity relation correlates to a fuzzy partition. We later use this proposition in our construction of fuzzy quotient groups (chapter 4).

Proposition 2.25. There exists a fuzzy similarity relation on a universal set X, iff there exists a corresponding fuzzy partition of the set X.

Proof. Let X be the universal set. Let $x, y, z \in X$.

$\Longrightarrow$: Let $R : X \times X \to [0, 1]$ be a fuzzy similarity relation on X. Consider the collection of fuzzy subsets of X, $\{A_x \,:\, x \in X\}$, who's membership values are defined such that, for fixed $x \in X$, $A_x(y) = R(x, y)\ \forall y \in X$. This is often referred to as the similarity class of x with respect to R.

(i) Condition (i) is met by the reflexive property of R, as for every $x \in X$, there is a set, namely A_x, such that $A_x(x) = R(x, x) = 1$.

(ii) Now let us consider that there exists a secondary set A_y such that $A_y(x) = 1$ as well. By transitive property of R, we see that, for any $w \in X$, $A_x(w) = R(x, w) \geq \min\{R(x, y), R(y, w)\}$. By symmetric property of R, $R(x, y) = R(y, x)$. By hypothesis, $R(y, x) = 1$.
Therefore, $A_x(w) \geq \min\{R(x, y), R(y, w)\} = \min\{1, R(y, w)\} = R(y, w) = A_y(w)$.
Thus $A_x(w) \geq A_y(w)$.

Also by transitive property,
$A_y(w) = R(y, w) \geq \min\{R(y, x), R(x, w)\}$, but again by hypothesis, $R(y, x) = 1$.
Therefore, $A_y(w) \geq \min\{R(y, x), R(x, w)\} = \min\{1, R(x, w)\} = R(x, w) = A_x(w)$.
Thus, $A_y(w) \geq A_x(w)$.

Therefore, $A_y(w) = A_x(w)$. Since this is for arbitrary $w \in X$, we have fuzzy set equality, $A_x = A_y$. Therefore,$\forall x \in X$, there exists a unique fuzzy set A_x such that $A_x(x) = 1$. If there also exists a secondary set A_y such that $A_y(x) = 1$, then $A_x = A_y$.

By extension, suppose there exists any element $x \in X$, and any two sets A_y, A_z such that $A_y(x) = A_z(x) = 1$. By by property (i), $A_x(x) = 1$. Therefore, by the proof above, $A_y = A_x = A_z$. Thus, (ii) holds.

(iii) Let $a, x, b, z \in X$ such that $A_x(a) = A_z(b) = 1$. We need show that $\bigvee_{y \in X} (A_x \cap A_z)(y) = A_x(b) = A_z(a)$. We know that $A_x(x) = A_z(z) = 1$. Therefore, by property (ii), this is equivalent to showing that $\bigvee_{y \in X} (A_x \cap A_z)(y) = A_x(z) = A_z(x)$.

By definition of fuzzy set intersection, this indicates we need show that $\bigvee_{y \in X} \min\{A_x(y), A_z(y)\} = A_x(z) = A_z(x)$.

By transitive and symmetric properties of R, we see for any $y \in X$, $\{R(x, y), R(z, y)\} \leq R(x, z)$. Therefore, $\bigvee_{y \in X} \{R(x, y), R(z, y)\} \leq R(x, z)$. Hence, by the definition of A_x, we see that $\bigvee_{y \in X} \{A_x(y), A_z(y)\} \leq A_x(z)$.

To prove (iii), it remains to show $\bigvee_{y \in X} \{A_x(y), A_z(y)\} \geq A_x(z)$. To do this, consider $y = z$. Thus we see that $\bigvee_{y \in X} \min\{A_x(y), A_z(y)\} \geq \min\{A_x(z), A_z(z)\}$. By reflexive property of R, $\min\{A_x(z), A_z(z)\} = \min\{A_x(z), 1\} = A_x(z)$. Therefore, $\bigvee_{y \in X} \{A_x(y), A_z(y)\} \geq A_x(z)$.

Thus we conclude that $\bigvee_{y \in X} \{A_x(y), A_z(y)\} = A_x(z)$. The proof that $\bigvee_{y \in X} \{A_x(y), A_z(y)\} = A_z(x)$ is identical. Therefore, $\bigvee_{y \in X} (A_x \cap A_z)(y) = A_x(z) = A_z(x)$. As stated in the introduction to this proof, since this equality holds for x and z, by property (ii), the equality holds for any $a, b \in X$; if $A_x(a) = A_z(b) = 1$ then $\bigvee_{y \in X} (A_x \cap A_z)(y) = A_x(b) = A_z(a)$. Hence, we conclude that condition (iii) is met.

Thus, we conclude that the existence of a fuzzy equivalence relation, R, on X, implies the existence of a fuzzy partition of X.

$\Longleftarrow$: Let $\{A_i\}$ be a fuzzy partition of X. Consider the relation R as defined by (Morsi, 1994)[32].

$$R(x,y) = \bigvee_{w\in X} \min\{A_x(w), A_y(w)\}$$

where A_x denotes the unique element of the fuzzy partition for which $A_x(x) = 1$. We provide the following proof that R defines a similarity relation on X.

Reflexive: The reflexive property is a direct result of the definition of R. Since $R(x,x) = \bigvee_{y\in Y} \min\{A_x(y), A_x(y)\}$ We then choose $y = x$, and find that
$R(x,x) \geq \min\{A_x(x), A_x(x)\} = A_x(x) = 1$. Thus, $R(x,x) = 1$; R is reflexive.

Symmetric: By the symmetric property of the min operation, we see
$R(x,y) = \bigvee_{w\in X} \min\{A_x(w), A_y(w)\} = \bigvee_{w\in X} \min\{A_y(w), A_x(w)\} = R(y,x)$. Thus, R is symmetric.

Transitive: To show that R is transitive we need show that for any $y \in X$,
$\min\{R(x,y), R(y,z)\} \leq R(x,z)$. Using the definition of R, $R(x,y) = \bigvee_{w\in X} \min\{A_x(w), A_y(w)\}$, and property (iii) of a fuzzy partition, we infer,

$$A_x(z) = A_z(x) = \bigvee_{w\in X} \min\{A_x(w), A_z(w)\} = R(x,z).$$

Similarly, by property (iii), $R(x,y) = A_x(y) = A_y(x)$, and
$R(y,z) = A_y(z) = A_z(y)$. Therefore, $\min\{R(x,y), R(y,z)\} = \min\{A_x(y), A_z(y)\}$.

Again, by property (iii) of fuzzy partitions, for $y \in X$, $\min\{A_x(y), A_z(y)\} \leq A_x(z)$. Therefore $\min\{R(x,y), R(y,z)\} = \min\{A_x(y), A_z(y)\} \leq A_x(z) = R(x,z)$. Thus, as this is for arbitrary $y \in X$, $\bigvee_{y\in X} \min\{R(x,y), R(y,z)\} \leq R(x,z)$, and hence transitivity holds.

From the above, we see that R defines a similarity relation on X. $\square$

Therefore, we conclude a fuzzy partition is equivalent to a similarity relation. We often refer to the fuzzy sets of a fuzzy partition as similarity classes, and may denote them as $[x]$, where $[x]$ represents the unique similarity class for which $[x](x) = 1$. We use this notation in chapter 4 to construct fuzzy quotient groups.

2.6 Places to Go Next

In this chapter, we have explored elementary properties of fuzzy set theory. We have selected these topics based on their prevalence in fuzzy logic research, and their usefulness to later chapters of this work. The following works explore further principles of fuzzy set theory.

(i) (Nguyen, Walker & Walker, 2018)[35] Nguyen, Walker and Walker have written a textbook of fuzzy logic,*A First Course in Fuzzy Logic*. This work contains chapters dedicated to fuzzy set logic in both the standard $[0, 1]$ set of truth values, as well as a generalization of such for the use of fuzzy truth values in an arbitrary complete lattice.

(ii) (Zimmerman, 2001)[53] In *Fuzzy Set Theory, and Its Applications*, Zimmerman describes the elementary fuzzy set operations, and extends this to type 2 fuzzy sets. Zimmerman later extends the principles of fuzzy set theory to explore applications through fuzzy control theory.

(iii) (Mohinta and Samanta, 2015)[30] in *An Introduction to Fuzzy Soft Graphs*, Mohinta and Samanta introduce the graphical interpretation of fuzzy relations on a fuzzy set. They explore much of the core constructions of fuzzy graph theory, which are built on weak reflexive, fuzzy relations.

CHAPTER 3

Fuzzy Connectives and $L-$Fuzzy Set Theory

In the previous chapter, we used specific operations to define fuzzy set operations for intersection, complement, and union; min, $1 -$ " ", and max, respectively. These operations are examples of connectives (defined in 3.1). These specific connectives work well with the interval of truth values $[0, 1]$, and are the most commonly used connectives in fuzzy logic research.

Definition 3.1. Connective: Let X be a nonempty set of elements known as truth values. An operation $*$ is an n-ary connective on X iff $* : X^n \to X$.

Example 3.2. Consider the interval $[0, 1]$. Let f be a function defined $f : [0, 1]^3 \to [0, 1]$ such that for $x, y, z \in [0, 1]$, $f(x, y, z) = max\{0, x - y - z\}$. f is a trinary connective, as the range of f is $[0, 1]$.

Example 3.3. For the interval $[0, 1]$, the binary operations min and max, as well as the unary operation, $' : x \mapsto 1 - x$, are connectives of $[0, 1]$.

It is possible, to generalize many of the properties of min, max and $1 -$ "" respectively, so as to model a system of truth values, other than that of $[0, 1]$. When doing so, we place the connectives used to model the standard set operations into three distinct categories; T-norms, negations, and T-conorms. A T-norm is used when defining a generalized set intersection; a negation is used for a generalized set complement; a T-conorm is used for a generalized set union. Using these groups of connectives, we can generalize our truth values from $[0, 1]$, to any complete lattice, L. As a reminder, we define a complete lattice L below.

Definition 3.4. Let L be a nonempty crisp set, with a partial order denoted $\leq$. $(L, \leq)$, or briefly, L if the order is clear, is a complete lattice iff for any $A \subseteq L$, there exist $\bigvee A \in L$ and $\bigwedge A \in L$ such that, $\bigvee A = \sup\limits_{a \in A} a$ and $\bigwedge A = \inf\limits_{a \in A} a$.

Example 3.5. Consider the interval $[0, 1]$ and a set $V_n = \{v \mid v \text{ is a n-dimensional vector, with entries from } [0, 1]\}$. Let $a, b \in V_n$ such that $a = \langle a_i \rangle$ and $b = \langle b_i \rangle$. Define a partial order $\trianglelefteq$ such that $a \trianglelefteq b$ iff for all entries a_i and b_i of a and b, $a_i \leq b_i$.

Thus, $(L, \trianglelefteq)$ defines a complete lattice, as for any subset A of vectors in V_n, there exists $\bigvee A = \langle \sup_{\langle a \rangle \in A} a_i \rangle$, the vectors of supremums of the ith entry of the vectors of A. Similarly, $\bigwedge A = \langle \inf_{\langle a \rangle \in A} a_i \rangle$.

3.1 T-norms: Class 1 of Fuzzy Connectives

Let us begin our exploration of connectives with the first of three categories of connectives, T-norms. We begin our construction of T-norms with a goal to generalize the connective, min, in order to model set intersection of fuzzy sets whose elements have truth values in a general complete lattice. In order to do this, we define a T-norm on a general lattice L to be any binary operation on L which preserves certain properties of min.

Definition 3.6. (Nguyen et al., 2018) [35] Let L be a complete lattice with order $\leq$, and let $1 = \bigvee L$ and $0 = \bigwedge L$. A binary connective $T : L \times L \to L$, is a T-norm if it satisfies the following conditions. For any $w, x, y, z \in L$,

(i) $1 \; T \; x = x$ (identity)

(ii) $x \; T \; y = y \; T \; x$ (symmetric)

(iii) $x \; T \; (y \; T \; z) = (x \; T \; y) \; T \; z$ (associative)

(iv) $(w \leq x)$ and $(y \leq z) \implies (w \; T \; y) \leq (x \; T \; z)$ (order preserving)

Example 3.7. Let us verify that $T = \min$ is indeed a T-norm on the lattice $[0, 1]$.

(i) For all $x \in [0, 1]$, $x \leq 1$, thus $x \; T \; 1 = min\{x, 1\} = x$ for all x.

(ii) Let $x, y \in [0, 1]$. By commutative property of the min operation,
$x \; T \; y = min\{x, y\} = min\{y, x\} = y \; T \; x$.

(iii) Let $x, y, z \in [0,1]$. We wish to show that $xT(yTz) = \min\{x, \min\{y, z,\} = \min\{\min\{x, y\}, z\} = (xTy)Tz$. This is shown through cases.

Case 1) Suppose $\min\{x, \min\{y, z,\}\} = x$. Thus, $x \leq \min\{y, z\}$, and thus by transitivity of $\leq$, $x \leq y$ and $x \leq z$. Therefore, $\min\{x, y\} = x$ and $\min\{x, z\} = x$. Therefore, $\min\{\min\{x, y\}, z\} = \min\{x, z\} = x$. Hence we conclude, $\min\{x, \{y, z\}\} = \min\{\min\{x, y\}, z\}$.

Case 2) Suppose $\min\{x, \min\{y, z\}\} = y$. Thus $y \leq z$, and consequently $y \leq x$. Therefore, $\min\{x, y\} = y$ and $\min\{y, z\} = y$. Thus, $\min\{\min\{x, y\}, z\} = y$. Hence we conclude $\min\{x, \{y, z\}\} = \min\{\min\{x, y\}, z\}$

Case 3) Suppose $\min\{x, \min\{y, z\}\} = z$. Thus, in a similar fashion to case 2, $z \leq y$ and $z \leq x$. Therefore, $z \leq \min\{x, y\}$, and thus $\min\{\{x, y\}, z\} = z$. Hence we conclude, $\min\{x, \{y, z\}\} = \min\{\min\{x, y\}, z\}$.

Thus, in all cases, $xT(yTz) = \min\{x, \min\{y, z\}\} = \min\{\min\{x, y\}, z\} = (xTy)Tz$, and hence, T is associative.

(iv) Let $w, x, y, z \in [0,1]$ such that $w \leq x$ and $y \leq z$. Thus, by definition of min,
$w\ T\ y = min\{w, y\} \leq w \leq x$. Similarly, $\min\{w, y\} \leq y \leq z$.
Therefore, $\min\{w, y\} \leq \min\{x, z\}$ and thus $w\ T\ y \leq x\ T\ z$; $T = \min$ is order preserving.

Therefore, we have shown that $([0,1],\ \min\ )$ satisfies the properties of a T-norm. From here on, we denote $\wedge = \min$.

Example 3.8. Consider the complete lattice V_n of Example 3.5. Let $w = \langle w_i \rangle$, $x = \langle x_i \rangle$, $y = \langle y_i \rangle$, and $z = \langle z_i \rangle$ be vectors in V_n. Define $a * b = \langle a_i \wedge b_i \rangle$. In this way, we see that

(i) $x * \langle 1 \rangle = \langle x_i \wedge 1 \rangle = \langle 1 \wedge x_i \rangle = \langle x_i \rangle = x$. Thus (i) is satisfied.

(ii) $*$ is symmetric as a direct result of the symmetric property of min.

(iii) Associativity is a direct result of associativity of min in each coordinate, i.

(iv) Let $w \trianglelefteq y$ and $x \trianglelefteq z$. Thus, by construction of V_n and its partial order $\trianglelefteq$, $w_i \leq y_i$ $\forall i$, and $x_i \leq z_i \forall i$. Therefore, $w_i \wedge x_i \leq y_i \wedge z_i$ $\forall i$ by order preserving property of min. Thus, $w * x \trianglelefteq y * z$. Hence, $*$ is order preserving.

Therefore, we conclude that $*$ as defined above, is a T-norm on V_n.

Example 3.9. (Nguyen, Walker & Walker , 2018) [35] Consider the binary connective defined below on $[0, 1]$, with the standard order.

$$x \bigtriangleup \ y = \begin{cases} \text{x} \wedge y & \text{if } x \vee y = 1 \\ \\ 0 & \text{otherwise} \end{cases}$$

This is a non-continuous norm, but it is a T-norm; though it is an extremal one.

Proof. Due to the piece-wise construction of this operation, it is necessary to use case proofs to show that $\bigtriangleup$ is indeed a T-norm.

(i)Identity: Let $x \in [0, 1]$. Thus, we see that $1 \vee x = 1$ $\forall x \in [0, 1]$. Thus, 1 $\bigtriangleup x$=1 $\wedge$ $x = x$, $\forall x \in [0, 1]$. Thus the idnetity condition is satisfied.

(ii) Symmetric: Let $x, y \in [0, 1]$. We know that $\wedge$ and $\vee$ are symmetric. Therefore, if $x \vee y = 1$, then $y \vee x = 1$. Thus, if $x \bigtriangleup y = x \wedge y$, then $x \bigtriangleup y = x \wedge y = y \wedge x = y \bigtriangleup x$. Similarly, if $x \vee y \neq 1$, then $y \vee x \neq 1$. Hence, $x \bigtriangleup y = 0 = y \bigtriangleup x$. Therefore, $\bigtriangleup$ has the symmetric property is any case.

(iii) Associative: Let $x, y, z \in [0.1]$.

Case 1) Suppose all elements have value strictly less than 1. Therefore, $(x \bigtriangleup y) \bigtriangleup z = (0) \bigtriangleup z = 0 = x \bigtriangleup (0) = x \bigtriangleup (y \bigtriangleup z)$.

Case 2) Suppose that any single element has value of 1, but the other two do not.

Case 2a) $x = 1$ and $y, z < 1$. Therefore, $y \vee z < 1$, and thus $y \bigtriangleup z = 0$. Also, $x \vee (y \bigtriangleup z) = 1$, and thus $x \bigtriangleup (y \bigtriangleup z) = 1 \wedge 0 = 0$. Similarly, $x \bigtriangleup y = x \wedge y = y < 1$, thus $(x \bigtriangleup y) \vee z < 1$. Therefore, $(x \bigtriangleup y) \bigtriangleup z = 0$. Thus associativity holds.

Case 2b) Suppose that $y = 1$ and $x, z < 1$. Thus $y \vee z = 1$ and hence $y \bigtriangleup z = y \wedge z = z < 1$. Thus, $x \vee (yz) < 1$. Therefore, $x \bigtriangleup (y \bigtriangleup z) = 0$. Similarly, $x \vee y = 1$ and hence $x \bigtriangleup y = x \wedge y = x < 1$. Thus, $(x \bigtriangleup y) \vee z < 1$, and thus $(x \bigtriangleup y) \bigtriangleup z = 0$. Thus, associativity holds.

Case 2c) $z = 1$. Therefore, $yz = y < 1$. Thus $x \vee (y \bigtriangleup z) < 1$, and hence $x \bigtriangleup (y \bigtriangleup z) = 0$. Similarly $x \bigtriangleup y = x \wedge y = x < 1$. Therefore, $(x \bigtriangleup y) \vee z < 1$, and thus $(x \bigtriangleup y) \bigtriangleup z = 0$. Therefore, associativity holds.

Thus, in any of the three cases, associativity holds when exactly one element is equal to 1.

Case 3) Suppose that exactly one element is not 1.

Case 3a) $x < 1$: $x \bigtriangleup (y \bigtriangleup z) = x \bigtriangleup (1 \bigtriangleup 1) = x = x \bigtriangleup 1 = (x \bigtriangleup 1) \bigtriangleup 1 = (x \bigtriangleup y) \bigtriangleup z$

Case 3b) $y < 1$: $x \bigtriangleup (y \bigtriangleup z) = 1 \bigtriangleup (y \bigtriangleup 1) = 1 \bigtriangleup y = (1 \bigtriangleup y) \bigtriangleup 1 = (x \bigtriangleup y) \bigtriangleup z$

Case 3c) $z < 1$: $x \bigtriangleup (y \bigtriangleup z) = 1 \bigtriangleup (1 \bigtriangleup z) = 1 \bigtriangleup z = (1 \bigtriangleup 1) \bigtriangleup z = (x \bigtriangleup y) \bigtriangleup z$

Case 4) Suppose $x = y = z = 1$. Then $(x \bigtriangleup y) \bigtriangleup z = 1 = x \bigtriangleup (y \bigtriangleup z)$
Thus, associativity holds in all cases.

(iv) Let $w, x, y, z \in [0, 1]$, such that $w \leq y$ and $x \leq z$. We need show that $\bigtriangleup$ preserves order, that is, $w \bigtriangleup x \leq y \bigtriangleup z$.

Case 1) $w < 1$ and $x < 1$. Then, $w \bigtriangleup x = 0 \leq y \bigtriangleup z$.

Case 2) $w = 1$. Thus $y = 1$. Therefore, by property (i), $w \bigtriangleup x = x \leq z = y \bigtriangleup z$

Case 3) $x = 1$. Thus, $z = 1$ Therefore, $w \bigtriangleup x = w \leq y = y \bigtriangleup z$.

Thus, $\bigtriangleup$ preserves order of $[0, 1]$.
Therefore, we conclude that $\bigtriangleup$ as defined above satisfies the properties of a T-norm.

□

There are an infinite number of T-norms on the interval $[0, 1]$. However, it is of interest that there are extremal T-norms which serve as bounds on what the values of a T-norm may take. To prove this, let us first consider the following lemma, which is a generalization of $0 \wedge x = 0 \quad \forall x \in [0, 1]$. We can expand this property to any T-norm, and say that the infimum of a lattice, or the 0, absorbs elements through T.

Lemma 3.10. For any T-norm, T, on complete lattice L, with $\bigwedge L$ denoted 0, $0 \; T \; x = x \; T \; 0 = 0$.

Proof. Let $x \in L$. Note that $x \leq 1$. Therefore, by the order preserving property of T-norms, we see $0 \; T \; x \; \leq 0 \; T \; 1$. By the first property of T-norms, $0 \; T \; 1 = 0$. Thus,

$0 \; T \; x \leq 0$. However, T is a function into $[0,1]$, thus, $0 \; T \; x \geq 0$. Therefore, $0 \; T \; x = 0$; by symmetry of T, $x \; T \; 0 = 0$.
Thus we conclude, $0 \; T \; x = x \; T \; 0 = 0 \; \forall x \in [0,1]$. $\square$

As stated in (Nguyen et al., 2018) [35], the previous Example 3.9 of a T-norm, $\triangle$, and the min T-norm, denoted $\wedge$, are extremes of the class of of connectives known as T-norms. This tells us that when working with any T-norm on the standard lattice $[0,1]$, this T-norm exists in a limited range between $\triangle$ and $\wedge$. Consequently, we may return to our intuitive understanding of these truth values as degrees of membership, of "in-ness" within a set. With this proposition, we interpret that no matter what T-norm we use to model fuzzy set intersection, an element can never be "more" in an intersection of sets than in either of the original sets. With this aim in mind, we provide the following proof.

Proposition 3.11. For any T-norm, T, and $x, y \in [0,1]$,

$$x \triangle y \leq x \; T \; y \leq x \wedge y$$

Proof. Let T be a T-norm on $[0,1]$. Let $x, y \in [0,1]$.

$(\triangle \leq T)$ First we prove that $x \triangle y \leq x \; T \; y$.

Case 1) Suppose that $x = 1$. Then by use of the first property of T-norms, $x \triangle y = x \wedge y = y = 1 \; T \; y = x \; T \; y$. By symmetry of both $\triangle$ and T, the case $y = 1$ is identical. Hence, $x \triangle y = x \; T \; y$.

Case 2) $x < 1$ and $y < 1$. Therefore, $x \triangle y = 0 \leq x \; T \; y$.

Thus, $x \triangle y \leq x \; T \; y$.

$(T \leq \wedge)$ Now we prove that $x\ T\ y \leq x \wedge y$. We use the identity and order preserving properties of T-norms to find that $x \bigtriangleup y \leq x \bigtriangleup 1 = x$; $x \bigtriangleup y \leq 1 \bigtriangleup y = y$. Therefore, $x \bigtriangleup y \leq \min\{x, y\} = x \wedge y$.

Therefore, we conclude that the T-norms $\wedge$ and $\bigtriangleup$ form upper and lower bounds on any arbitrary T-norm of the interval $[0, 1]$. These particular bounds imply that an arbitrary T-norm has a very weak lower bound, but a relatively strong upper bound, min. □

There are many properties and principles of specific T-norms that can be defined; such as nilpotent, convex, Archimedean, and more (Nguyen et al., 2018) [35]. These properties induce different interpretations of special types of lattice-T-norm constructions. We include an exploration of nilpotency as an example of such properties. Nilpotency can be thought of as the property of a T-norm to ultimately be terminal; a T-norm which is anti-nilpotent, or strict, never diminishes, or vanishes to zero for non-zero elements.

Definition 3.12. (Nguyen, Walker & Walker, 2018) [35] Let T be a T-norm on a complete lattice, L. Let $x \in L$. Denote $x^{\overline{n}} = xTxTx...Tx$, or x under T, n times. The T-norm, T is called nilpotent iff for all $x \in L,\ x \neq 1$, there exists positive integer n such that $x^{\overline{n}} = 0$ for some n dependent upon x. T is called strict if $\forall x \in X,\ x \neq 0$, and $\forall n \in \mathbb{N}$, $x^{\overline{n}} \neq 0$.

Note, for a T-norm to be nilpotent, this norm must also have the condition that $\forall x \in L$, $x\ T\ x < x$.This is evident through the result of a contradiction otherwise. If there exists $x \neq 0$ such that $x\ T\ x = x$, then by induction, $x^{\overline{n}} = x\ \forall n \in \mathbb{N}$. Thus, a nilpotent T-norm has an innate cost associated with it; that is the difference of $x - (x\ T\ x)$. Eventually, this cost will outtake the element x itself, rendering the entire lattice, L, (save its supremum 1), anchored down to the infimum, 0.

Example 3.13. The standard T-norm, $\wedge$ is strict. This is easily shown using the definition of $\wedge$. Let $x \in (0, 1]$. Thus, $x^{\overline{2}} = \min\{x, x\} = x$. Therefore, by induction,

$\forall n \in \mathbb{N}$, $x^{\overline{n}} = x > 0$. Therefore, $\wedge$ is strict.

Example 3.14. Consider the T-norm of Example 3.9, $\triangle$. We shall show that this T-norm is nilpotent.

Let $x \in [0, 1)$. By definition of $\triangle$, as $x \vee x = x < 1$, we see that $x \triangle x = 0$. Thus, $\forall x \in L,\ x \neq 1$, $x^{\overline{2}} = 0$. Thus, we conclude $\triangle$ is nilpotent.

Example 3.15. Consider the T-norm Γ defined by $x\ \Gamma\ y = 0 \vee (x + y - 1)$. Γ is nilpotent. As an example, $x = \frac{3}{4}$. We see that $\frac{3}{4}\ \Gamma\ \frac{3}{4} = \frac{1}{2}$. continuing, we get $\frac{1}{2}\ \Gamma\ \frac{3}{4} = \frac{1}{4}$. Finally, one more time, $\frac{1}{4}\ \Gamma\ \frac{3}{4} = 0$. That is, $\frac{3}{4}^{\overline{3}} = 0$. This illustrates the property of Γ to decrease, or have a "cost" associated with using it.

Let $x \in [0, 1)$. To prove that Γ is nilpotent, we need show that there exists $n \in \mathbb{N}$ such that $x^{\overline{n}} = 0$. Consider $n \geq \frac{1}{1-x}$.

We suppose that $x^{\overline{n}} > 0$, and come to a contradiction. Note that

$$x^{\overline{n}} = (x + x - 1) + (x - 1) + (x - 1) .... + (x - 1) = (x + x ... + x) - (1 + 1 + ...1) =$$

$$= \sum_{i=1}^{n} x - \sum_{i=1}^{n-1} 1 = nx - (n - 1) = n(x - 1) + 1$$

By our hypothesis, we assumed $x^{\overline{n}} > 0$, thus, $n(x - 1) + 1 > 0$, and thus, $n < \frac{1}{1-x}$. However, this contradicts our choice of n. We thus conclude that for this choice of n, that $x^{\overline{n}} = 0$. Hence, Γ is indeed nilpotent.

Example 3.16. Consider the T-norm, $\cdot$, of ordinary multiplication in $[0, 1]$. Recall from ring theory that the ring of real numbers under addition and mutliplication, contains no zero-divisors. Thus, $\forall x \neq 0,\ x \cdot x \neq 0$. Therefore, by induction, $x^{\overline{n}} \neq 0\ \forall n \in \mathbb{N}$. Therefore, this T-norm is strict.

3.2 Negations and Involutions

We now turn our attention to the second category of connectives. These connectives are called negations, and are used to model a generalized set complement. Recall from chapter 2, that for a fuzzy set A of universal set X, the set complement A^c has membership values defined by $A^c(x) = 1 - A(x)\ \forall x \in X$. This operation is the prime example of the type of connectives used to model fuzzy set complement. We denote it $'$, such that $x' = 1 - x$ for $x \in [0, 1]$. We call the generalization of this connective a negation. In our construction, we desire the properties of negation to reflect key properties of this prime example. Specifically, we wish for a proposed negation ^{n} to "flip" the interval $[0, 1]$. Specifically, we wish for a general negation ^{n} to map the supremum 1 to the infimum, 0, and map the infimum 0 to the supremum 1. Also, we wish for ^{n} to reverse order, that is, for any $x, y \in L$, if $x \leq y$, then $y^n \leq x^n$.

We may additionally require this operation, ^{n}, to have order 2. That is, if we "flip" the interval $[0, 1]$ twice, then we yield the original interval. This last stipulation defines a special kind of negation, known as an involution . For our purposes, we only consider involutions in the context of norms on fuzzy sets, and refer to them simply as, negations, as in [35].

Definition 3.17. (Nguyen, Walker & Walker, 2018) [35] Let L be a complete lattice with order $\leq$, where $1 = \bigvee L,\ 0 = \bigwedge L$. The unary connective $^n : L \to L$ is a negation iff for any $x, y \in L$,

(i) $0^n = 1,\ 1^n = 0$

(ii) $x \leq y \implies y^n \leq x^n$ (order reversing)

(iii) $(x^n)^n = x$ (Involution Property)

Example 3.18. We have discussed how this definition is inspired by the operation $'$ on $[0, 1]$. Indeed, we see that $'$ does possess the properties of a negation of the lattice

$[0, 1]$.

Proof.

(i) By definition, $0' = 1 - 0 = 1$, and $1' = 1 - 1 = 0$. $'$ maps supremum to infimum, and infimum to supremum.

(ii) Let $x, y \in [0, 1]$ such that $x \leq y$. Therefore, $-x \geq -y$, hence $1 - x \geq 1 - y$; thus $x' \geq y'$. $'$ is order reversing.

(iii) Let $x \in [0, 1]$. $(x')' = 1 - x' = 1 - (1 - x) = 1 - 1 + x = x$. $'$ has the involution property (also shown in Remark 2.10).
Thus, $'$, is a negation. $\square$

Example 3.19. (Nguyen, Walker & Walker , 2018)[35] Consider the function $\eta : [0, 1] \to [0, 1]$ such that $\eta(x) = \frac{1-x}{1+5x}$. We wish to show η is a negation on $[0, 1]$.

Let $x, y \in [0, 1]$

(i) $\eta(0) = \frac{1-0}{1+5\cdot 0} = 1$ and $\eta(1) = \frac{1-1}{1+5\cdot 1} = 0$

(ii) Suppose $x \leq y$ then $1 - x \geq 1 - y$.
Also, $0 < 1 + 5x \leq 1 + 5y$. Therefore, by dividing over these inequalities,
$\eta(x) = \frac{1-x}{1+5x} \geq \frac{1-y}{1+5x} = \eta(y)$. Hence, η is order reversing.
(iii) $\eta(\eta(x)) = \frac{1 - \frac{1-x}{1+5x}}{1+5(\frac{1-x}{1+5x})} = \frac{\frac{6x}{1+5x}}{\frac{6}{1+5x}} = \frac{6x}{6} = x$

Therefore, η as defined above, is a negation.

3.3 T-conorms: The Dual of T-norms

Now that we have discussed the concept of a negation, we may define a T-conorm (sometimes referred to as an S-norm). These T-conorms are generated by a T-norm in combination with a negation, and have many similar characteristics to the operation max. With an understanding of T-conorms, we are able to model set union of fuzzy sets who's truth values exist in any complete lattice. Additionally, we show that the max operation itself is indeed a T-conorm.

Definition 3.20. T-conorm: Let L be a complete lattice. Let $S : L \times L \to L$. S is a T-conorm on L iff $\exists$, T , ^{n}, a T-norm and negation on L such that $\forall x, y \in L$, $x\ S\ y = (x^n\ T\ y^n)^n$. We say that S is dual to T with respect to ^{n}, or simply S is dual to T, if ^{n} is understood.

Example 3.21. Let us prove max is the T-conorm generated by $\wedge$ and $'$ on $[0, 1]$.

Let $x, y \in [0, 1]$. Without loss of generality, suppose $x \leq y$. Hence, $-x \geq -y$, and $1 - x \geq 1 - y$.

Therefore, $(x'\ T\ y')' = 1 - \min\{1 - x, 1 - y\} = 1 - (1 - y) = y = \max\{x, y\}$.

Thus max is generated by min and $'$, and thus is a T-conorm. Moving forward, we denote $\vee = \max$.

Example 3.22. Consider the T-norm $*$ on V_n as in Example 3.8. Let $x \in V_n$. Define a negation operation ” such that $x” = \langle 1 - x_i \rangle$. As an extension of the previous example, we generate a T-conorm, &, defined, for $\langle x \rangle, \langle y \rangle \in V_n$, $\langle x \rangle$ & $\langle y \rangle = \langle x_i \vee y_i \rangle$. Hence, & is dual to $*$ with respect to ”.

By using the properties of a general T-norm on $[0, 1]$, and any negation on $[0, 1]$, we see that T-conorms have certain properties which generalize the properties of max. (Nguyen, Walker & Walker, 2018) [35].

Proposition 3.23. T-conorm: Let ^{n} be a fixed negation of $[0, 1]$. If a connective $S : [0, 1] \times [0, 1] \to [0, 1]$, is a $t - conorm$, dual to a T-norm, T, with respect to ^{n}, then S satisfies the following conditions. For any $w, x, y, z \in [0, 1]$;

(i) $0\ S\ x = x$ identity

(ii) $x\ S\ y = y\ S\ x$ symmetric

(iii) $x\ S\ (y\ S\ z) = (x\ S\ y)\ S\ z$ associative

(iv) $(w \leq x)$ and $(y \leq z) \implies (w\ S\ y) \leq (x\ S\ z)$ order preserving

Proof. Let T be a T-norm, and, ^{n}, a negation, such that S is dual to T with respect to ^{n}. Let $w, x, y, z \in [0, 1]$.

Proof.

(i) Identity

$$
\begin{aligned}
0\ S\ x &= (0^n\ T\ x^n)^n && \text{definition of } S \text{ being dual to } T. \\
&= (1\ T\ x^n)^n && \text{first property of negation} \\
&= (x^n)^n && \text{first property of T-norm} \\
&= x && \text{third property of negation}
\end{aligned}
$$

(ii) Symmetric

$$
\begin{aligned}
x\ S\ y &= (x^n\ T\ y^n)^n && \text{definition of } S\text{, dual of } T \\
&= (y^n\ T\ x^n)^n && \text{commutativity of } T \\
&= y\ S\ x && \text{definition of } S\text{, dual of } T
\end{aligned}
$$

(iii) Associative

$$
\begin{aligned}
x\ S\ (y\ S\ z) &= [x^n\ T\ ((y^n\ T\ z^n)^n)^n]^n && \text{definition of } S \text{ dual of } T \\
&= [x^n\ T\ (y^n\ T\ z^n)]^n && \text{involution property of negation} \\
&= [(x^n\ T\ y^n) T\ z^n]^n && \text{associativity of } T \\
&= [((x^n\ T\ y^n)^n)^n\ T\ z^n]^n && \text{third property of negation} \\
&= (x\ S\ y)\ S\ z && \text{definition of } S \text{ dual of } T
\end{aligned}
$$

(iv) Order Preserving: Let $w \leq x$ and $y \leq z$.

$x^n \leq w^n$ *and* $z^n \leq y^n$	by order reversing property of negation
$(x^n\ T\ z^n) \leq (w^n\ T\ y^n)$	order preserving property of T
$(w^n\ T\ y^n)^n \leq (x^n\ T\ z^n)^n$	order reversing property of negation
$w\ S\ y \leq x\ S\ z$	definition of S dual of T

□

Therefore, if T is any T-conorm, ^{n} is a negation, and S a $T-conorm$ dual to T with respect to ^{n}, then with respect to negation ^{n}, S has the above for properties, twin of those of T-norms. ***

□

Proposition 3.24. Let ^{n} be a fixed negation of $[0, 1]$. Let s be a binary connective of $[0, 1]$ such that s has the above properties of proposition 3.23. Let ^{n} be a negation of $[0, 1]$. If we define a connective t such that $\forall x, y \in [0, 1]$, $x\ t\ y = (x^n\ s\ y^n)^n$, then t is a $T-norm$. Additionally, s is the T-conorm generated by T-norm, t, with respect to negation ^{n}.

The proof of 3.24 is nearly identitcal to the proof of proposition 3.23.

Remark 3.25. From proposition 3.24, a binary connective, S, is a T-conorm iff S meets the conditions of proposition 3.23. Thus, it is equivalent to use these properties as a definition of a T-conorm, as is commonly done in fuzzy research.

From this section, we have seen that these three categories of connectives work in tandem with one another. Indeed, let L be any complete lattice, let $\mathbf{T}$ be the set of all T-norms on L; $\mathbf{N}$ the set of all negations; and $\mathbf{S}$ the set of all T-conorms. In our defining of T-conorms, we see that we may define a function for each negation, $^{\mathbf{n}}$, denoted$f_n : \mathbf{T} \to \mathbf{S}$ such that f_n maps any T-norm to its dual T-conorm. From Propositions 3.23 and 3.24, we see that each f_n is a one to one, onto correspondence.

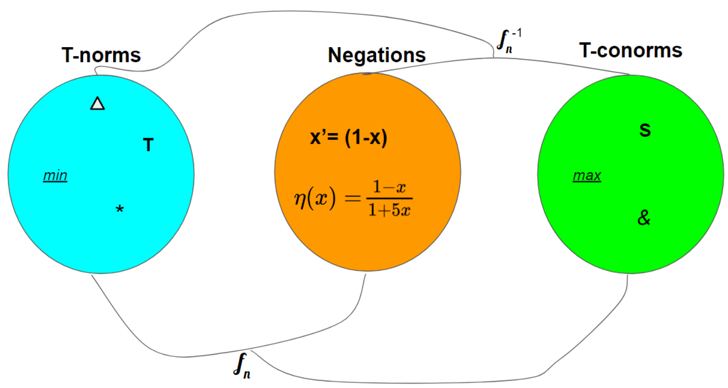

Figure 3: $f_n : \mathbf{T} \to \mathbf{S}$

3.4 A Generalized Vagueness: $L-$Fuzzy Set Theory

Having defined categories of connectives, we gain an ability to model set operations in a variety of combinations of T-norms, T-conorms, and negations. Additionally, we may generalize a fuzzy set, rather than a function from a universal set into $[0, 1]$, as a function into any complete lattice, L.

Definition 3.26. L-Fuzzy Set:(Goguen, 1967) [17] Let X be a crisp set. Let L be a complete lattice. An $L-$fuzzy set is a function $A : X \to L$.

Using this definition, we may generalize fuzzy set theoretical definitions for an arbitrary $L-$fuzzy power set, the intersection by an appropriate T-norm, set complement by a negation, and union by the dual T-conorm. This generalized construction is of particular use when using vectors of truth values in an additive fuzzy system (chapter 7).

Definition 3.27. L-fuzzy Set Theory: Let X be a crisp set, and suppose L is a complete lattice with order $\trianglelefteq$. Let T, ^{n}, and S be a T-norm, negation, and dual T-conorm respectively on L. For $A, B \in L^X$, we define the following operations, yielding similar results to those of chapter 2.

(i)$A \subseteq B$ iff $A(x) \trianglelefteq B(x)$ $\forall x \in X$

(ii) $A = B$ iff $A(x) = B(x)$ $\forall x \in X$

(iii) $\forall x \in X$, $(A \cap B)(x) = A(x)\ T\ B(x)$

(iv) $\forall x \in X$, $(A \cup B)(x) = A(x)\ S\ B(x)$

(v) $\forall x \in X$, $A^c(x) = (A(x))^n$

Example 3.28. Vector Truth Values of $[0, 1]$: Consider the complete lattice ,V_n, in Example 3.5, where $V^n = \{v \mid v$ is a vector of length n, with entries in [0,1]$\}$, is a complete lattice. Let X be a crisp set. Let $A : X \to V_n$. Thus we assign an element $x \in X$, a vector $v = \langle v_i \rangle \in V_n$ as its truth value. We thus have constructed a V_n-fuzzy set, A, by using a corresponding V_n-fuzzy set theory of (X, V_n); (X, V_n) defined by the T-norm $*$, negation ", and dual T-conorm &, as defined in examples 3.8, and 3.22.

Remark 3.29. Demorgan's Laws hold for $L-$fuzzy set constructions. This is a direct result of the construction of a T-conorm's dual nature to a T-norm as in definition 3.20.

Remark 3.30. The distributive laws may not hold for arbitrary T-norm and T-conorm in an arbitrary lattice. If they do, then we call this construction a distributive lattice. (Goguen, 1967) [17]

Example 3.31. The distributive laws hold for the nth-dimensional vector space, V_n with $*$ and & defined as in examples 3.8 and 3.22. This is a direct result of the distributive laws of T-norm and T-conorm $\wedge$ and $\vee$.

Let $\langle a_i \rangle, \langle b_i \rangle \in V_n$. Thus we see,
$(\langle a_i \rangle \ \& \ \langle b_i \rangle) \ * \ \langle c_i \rangle = \langle (a_i \vee b_i) \wedge c_i) \rangle$.
Additionally, by distributive laws of $\wedge$ and $\vee$,
$\langle (a_i \vee b_i) \wedge c_i) \rangle = \langle (a_i \vee c_i) \wedge (b_i \vee c_i) \rangle = (\langle a_i \rangle \ * \ \langle c_i \rangle) \ \& \ (\langle b_i \rangle \ * \ \langle c_i \rangle)$

Thus, $(V_n, \langle 0 \rangle, \langle 1 \rangle, *, \&, ')$ as constructed in examples 3.8 and 3.22, forms a distributive $L-$fuzzy set theory.

3.5 Places to Go Next

From the above, we have explored basic principles of generalized connectives, which are used to construct a generalized $L-$fuzzy set logic. Such principles are the foundation of constructions moving forward in this thesis. For further self-study in connectives and $L-$fuzzy set theory, we have listed the following works of interest.

(i) (Goguen, 1967) [17] In “L-Fuzzy Sets”, Goguen explores the extensions of fuzzy logic to a set L called a “poset”, as a “decision language”. This paper is one of the original depictions of L-fuzzy set theory constructions.

(ii) (Nguyen, Walker, & Walker, 2018) [35] In this book, “A first course in fuzzy logic”, Nguyen, Walker and Walker explore a variety of topics of fuzzy logic and its applications. In particular, concepts of connectives in the interval $[0, 1]$ are explored in relation to particular properties they may hold.

CHAPTER 4

Fuzzy Algebra

In this chapter, we continue our mission of developing fuzzy mathematics to the next useful level, fuzzy algebra. Recall that fuzzy subsets can be defined in any scenario involving a crisp subset. Therefore, structures such as subgroups can be defined using fuzzy subsets. While it would be impossible to explore every fuzzy algebraic structure and all their properties in one paper, this chapter shows a method of modeling a fuzzy "closure" with respect to an arbitrary operation. This method can be expanded to model further fuzzy algebraic structures involving "closure", such as rings. While defining fuzzy subgroups, recall that it is our motivation to show that fuzzy logic is an expansion of traditional logic. That is, if we define a fuzzy algebraic structure, and limit truth values of elements of this structure to values of $\{0, 1\}$, then our new definition is equivalent to the corresponding, traditional (crisp) definition.

When moving from crisp to fuzzy logic, particularly for the purpose of studying fuzzy algebraic structures, it is helpful to think of a fuzzy set as a mathematical cloud. Clouds exist, and we may categorize different types and sizes of clouds; however, they are not easy to grab. Their boundaries are difficult to define, as they are merely a loosely defined area of the atmosphere with a high density of water vapor. These clouds are not point-mass objects; that is, their mass cannot be acted upon as if it exists solely on a central point. Rather, we must consider the cloud in its entirety, with all its structure defined by its fuzzy vapor, and elements moving through the vapor. In the context of fuzzy groupoids, the discussion of such clouds focuses on closure of a cloud under a given operation. If we as an element traverse through a cloud A through a given operation $*$, what is the behavior of $*$? If A was a crisp set, we would ask the question, "Can an object leave A through $*$?". However, we cannot determine a crisp boundary at which we move from fully in A to not in A; thus the question we instead ask is, "To what degree does $*$ keep an object within A?".

4.1 From the Seed of Fuzzy Subsets, Comes the Sapling of Fuzzy Subgroupoids

To derive a definition of a fuzzy subgroupoid, we first recall the definition of a crisp groupoid and subgroupoid. For a nonempty set G, and a binary operation on G, $*$, we say that the pair, $(G, *)$ is a groupoid iff for any $x, y \in G$, $x * y \in G$. When the operation is clear, we often denote $x * y$ as simply, xy. For a groupoid, $(G, *)$, with a crisp subset A of G, we say that A is a crisp subgroupoid of G iff for any $x, y \in A$, $x * y \in A$. Using the nominclature of fuzzy logic, this definition is equivalent to the following.

Definition 4.1. Groupoid: Let G be a non-empty crisp set, and $*$ a binary operation on G. The pair, $(G, *)$ is referred to as a groupoid, iff $\forall x, y \in G$,
$(G(x) = G(y) = 1) \implies (G(x * y) = 1)$.

Definition 4.2. Subgroupoid: Let $(G, *)$ be a goupoid, and suppose A is a crisp subset of G. We refer to $(A, *)$ as a subgroupoid iff $(A, *)$ is a groupoid. That is, A is non-empty, and $\forall x, y \in G$, $(A(x) = A(y) = 1) \implies (A(x * y) = 1)$.

This refers to the classic concept of closure of a subset with respect to an operation. This closure can be paraphrased as, "The output of an operation is in A when the inputs are in A". Conversely, if either element is not in A, or has membership of 0 in the set A, then the result may or may not have membership 1 in A. To put it another way, an element x has membership in A at least as strong as the minimum membership of any pair of factors of x. Therefore, we may equivalently define a crisp subgroupoid of G, as a subset A of G such that $\forall x, y \in G$, $\min\{A(x), A(y)\} \leq A(x*y)$.

Notice that this definition, which we refer to as the subgroupoid inequality, is compatible with fuzzy membership values. Therefore, we may define a fuzzy subgroupoid by this inequality, and note that this is equivalent to the traditional definition of subgroupoids when the membership values of A are of $\{0, 1\}$. However, this is not the only definition used in fuzzy research. In fact, there are at least three definitions of a fuzzy subgroupoid which we may derive, as outlined in (Liu, 1982) [26]. In the next section, we explore the mathematical components of these three definitions, and show that these definitions are equivalent; thus they serve as alternative interpretations to the same fuzzy structure.

4.2 Deriving the Fuzzy Subgroupoid Inequality

To explore these alternative definitions of fuzzy subgroupoids, we need first introduce a new operation, fuzzy set composition. This is a binary operation on the fuzzy power set of a crisp groupoid, G. While this composition is on fuzzy sets, which are often thought of as functions, this composition is not traditional function composition.

Let $\cdot$ be a binary operation on a crisp set G. Let $A, B \in [0,1]^G$. We define the product of A and B, denoted $A \circ B$, as in (Liu, 1982) [26]. For $z \in G$,

$$(A \circ B)(z) = \begin{cases} \bigvee_{x \cdot y = z} min\{A(x), B(y)\} & \exists\, x, y \in G \ni x \cdot y = z \\ 0 & \nexists\, x, y \in G \ni x \cdot y = z \end{cases}$$

This definition is equivalent to stating that an element is exactly as strong in the composition of two sets, as any pair of its factors are within the two factor sets respectively. If we compose a fuzzy subset A of a crisp groupoid G, this definition is similar to our previous fuzzy subgroupoid inequality; any element of G is as strong in A as the minimum of either of its factors.

For fluidity in the following propositions, we need also define a fuzzy singleton in a fuzzy set. A fuzzy singleton is a crisp singular element of the universal set, paired with a designated strength, $\lambda \in [0,1]$. Rather than being "in" the set in the traditional sense, we say that the fuzzy element is "in" a fuzzy subset if the stated strength of said element is no stronger than the membership of said element in the fuzzy subset.

Definition 4.3. (Liu, 1982) [26] Fix $x \in G$, and let $\lambda \in [0,1]$. We define a fuzzy singleton as a fuzzy subset of G, denoted x_λ, for $x \in G$, with membership values defined as below for any $y \in G$.

$$x_\lambda(y) = \begin{cases} \lambda & y = x \\ 0 & y \neq x \end{cases}$$

We may interpret these fuzzy singletons as the fuzzy equivalence to singular elements. However, as with fuzzy sets, it is difficult to concretely grasp the individual elements. These elements are not particles that we can easily touch. Rather, they are more closely related to photons in some aspects. Photons are not easily trapped or handled, but they can still be acted upon. When considering a fuzzy subgroupoid, visualize light photons (fuzzy singletons) moving throughout a cloud (fuzzy set). Additionally, consider an operation $*$ representing the structure of the water vapor of the cloud which reflects the photons to different places throughout the cloud. If this is an arbitrary cloud, the light may traverse further outside of the cloud, and be reflected elsewhere in the universe. We say the cloud is closed under the reflection operation $*$ iff the reflection operation $*$ either causes the photons to travel deeper into the cloud, or at least does not allow photons further outside of the cloud. That is, even though the boundaries of the cloud are not crisp, the cloud acts effectively as a trap which no light may 'escape'. $*$ need not be a unary operations; rather, $*$ may take multiple photons, two in the case of fuzzy subgroupoids, and combine them to one photon. This resulting photon cannot exist further outside of the cloud than all of the original photons.

Additionally, each photon has a wavelength by which it travels, similar to the strength associated with a fuzzy singleton. Suppose these wavelengths are scaled to the interval $[0,1]$. If this scaled wavelength is less than or equal to the degree to which the photon is in the cloud, we merely say that the photon is "in" the cloud. We then envision the fuzzy singletons we consider to be "in" the fuzzy groupoid. This yields a similar figure to the cloud displayed below. Photons of longer wavelengths, such as red and orange photons, exist deeper within the cloud; photons of shorter wavelength, such as blues and violets, exists towards the fuzzy boundaries of the cloud.

Figure 4: Rainbow Clouds Above Belukha Mountain [45]

For a fuzzy subset A of G, and $x \in G$, we often use the crisp symbol, $\in$ to denote the inclusion of a fuzzy singleton in A. That is, we say "$x_\lambda \in A$" to indicate $\lambda \leq A(x)$. Consequently, for $y \in G$, $A(y) = \bigvee_{x_\lambda \in A} x_\lambda(y)$. Note that the crisp definition of $\in$ is equivalent to the inclusion of a fuzzy singleton in a set when $\lambda = 1$. That is, for crisp set A, the crisp element $x \in A$ iff $x_1 \in A$ in the fuzzy sense.

Using the above definitions, we provide proofs of the following equations, which are listed in (Liu, 1982) [26]. These equations are used later in this section to prove the equivalence of three definitions of a fuzzy subgroupoid.

Lemma 4.4. Let G be a crisp groupoid.

(i) Let x_λ and y_μ be fuzzy singletons in G. Then $x_\lambda \circ y_\mu = (x \cdot y)_{min(\lambda,\mu)}$

(ii) Let $A, B \in [0,1]^G$. Then $A \circ B = \bigcup_{x_\lambda \in A, y_\mu \in B} (x_\lambda \circ y_\mu)$

Proof. (i) Let x_λ, y_μ be fuzzy singletons of G for crisp groupoid $(G, \cdot)$, and let $w \in G$.

Thus, by definition of set composition,

$$(x_\lambda \circ y_\mu)(w) = \bigvee_{u \cdot v = w} min\{x_\lambda(u), y_\mu(v)\}$$

By definition of a fuzzy singleton, we see that

$$min\{x_\lambda(u), y_\mu(v)\} = \left\{ \begin{array}{cc} min\{\lambda, \mu\} & u = x, y = v \\ 0 & otherwise \end{array} \right.$$

However, if $u = x$ and $y = v$, then $w = x \cdot y$. Hence by definition of a fuzzy singleton

$$(x_\lambda \circ y_\mu)(w) = \left\{ \begin{array}{cc} min\{\lambda, \mu\} & w = x \cdot y \\ 0 & otherwise \end{array} \right\} = (x \cdot y)_{\min\{\lambda, \mu\}}(w)$$

Since this equality is for any $w \in G$, we have proven fuzzy set equality, $(x_\lambda \circ y_\mu) = (x \cdot y)_{\min\{\lambda, \mu\}}$.

(ii) Let $A, B \in [0,1]^G$ and $w \in G$. We need show $(A \circ B)(w) = (\bigcup_{x_\lambda \in A,\ y_\mu \in B} (x_\lambda \circ y_\mu))(w)$.

For the trivial case, suppose there does not exist $u, v \in G$ such that $u \cdot v = w$. Thus $(A \circ B)(w) = 0$. Also, for any two elements x, y, $x \cdot y \neq w$, and therefore, $(x \cdot y)_{\min\{\lambda, \mu\}}(w) = 0\ \forall \lambda, \mu \in [0, 1]$. Consequently, by part (i) $(x_\lambda \circ y_\mu)(w) = 0$. Therefore, the membership of w in the union of all such sets is 0: $\bigcup_{x_\lambda \in A,\ y_\mu \in B} (x_\lambda, y_\mu)(w) = 0$. Thus the equality holds when there does not exist two elements which combine to w.

Now let us consider the other case, that there exists at least one pair of elements $u, v \in G$ such that $u \cdot v = w$.

By definition of fuzzy set composition,

$$(A \circ B)(w) = \bigvee_{u \cdot v = w} min\{A(u), B(v)\}$$

Additionally, by definition of fuzzy singletons in a fuzzy set,

$$\bigvee_{u\cdot v=w} min\{A(u), B(v)\} = \bigvee_{u\cdot v=w} min\{\bigvee_{x_\lambda\in A} x_\lambda(u), \bigvee_{y_\mu\in B} y_\mu(v)\}$$

Thus for any fixed pair of singletons x_λ and y_μ, $(A\circ B)(w) \geq \bigvee_{u\cdot v=w} \min\{x_\lambda(u), y_\mu(v)\}$. As this holds for any fixed pair of singletons, the supremum over the choice of such singletons in A and B respectively holds the following inequality.

$$(A\circ B)(w) \geq \bigvee_{x_\lambda\in A,\ y_\mu\in B} \left(\bigvee_{u\cdot v=w}\{\min\{x_\lambda(u), y_\mu(v)\}\right)$$

Additionally, by part (i),

$$\bigvee_{x_\lambda\in A,\ y_\mu\in B} \left(\bigvee_{u\cdot v=w}\{\min\{x_\lambda(u), y_\mu(v)\}\right) = \bigvee_{x_\lambda\in A\ ,\ y_\mu\in B} (x_\lambda\circ y_\mu)(w)$$

Also, by definition of fuzzy set union,

$$\bigvee_{x_\lambda\in A\ ,\ y_\mu\in B} (x_\lambda\circ y_\mu)(w) = \bigcup_{x_\lambda\in A\ ,\ y_\mu\in B} (x_\lambda\circ y_\mu)(w)$$

Therefore, we have the inequality $(A\circ B)(w) \geq \bigcup_{x_\lambda\in A\ ,\ y_\mu\in B} (x_\lambda\circ y_\mu)(w)$.

We need prove inequality in the other direction;
$(A\circ B)(w) \leq \bigcup_{x_\lambda\in A\ ,\ y_\mu\in B} (x_\lambda\circ y_\mu)(w)$. By use of the definition of set composition and fuzzy set union, we have the equality,

$$\bigcup_{x_\lambda\in A,\ y_\mu\in B} (x_\lambda\circ y_\mu)(w) = \bigvee_{x_\lambda\in A,\ y_\mu\in B} \left(\bigvee_{u\cdot v=w} min\{x_\lambda(u), y_\mu(v)\}\right)$$

Therefore, choose $x_\lambda = u_{A(u)}$, and $y_\mu = v_{B(v)}$. Therefore,

$$\bigvee_{x_\lambda\in A,\ y_\mu\in B} \left(\bigvee_{u\cdot v=w} min\{x_\lambda(u), y_\mu(v)\}\right) \geq \bigvee_{u\cdot v=w} \min\{u_{A(u)}(u), v_{B(v)}(v)\}$$

Also, by definition of fuzzy singleton,

$$\bigvee_{u \cdot v = w} \min\{u_{A(u)}(u), v_{B(v)}(v)\} = \bigvee_{u \cdot v = w} min\{A(u), B(v)\}$$

Additionally, by definition of fuzzy set composition,

$$\bigvee_{u \cdot v = w} min\{A(u), B(v)\} = (A \circ B)(w)$$

Therefore, $(A \circ B)(w) \leq \bigcup_{x_\lambda \in A \ , \ y_\mu \in B} (x_\lambda \circ y_\mu)(w)$.

Therefore, in combination with inequality $\geq$ proven previously, we then conclude $(A \circ B)(w) = \bigcup_{x_\lambda \in A, y_\mu \in B} (x_\lambda \circ y_\mu)(w)$. Consequently, as this is for any $w \in G$, we have thus shown fuzzy set equality;

$$A \circ B = \bigcup_{x_\lambda \in A, y_\mu \in B} (x_\lambda \circ y_\mu)$$

$\square$

Let us now consider the multiple definitions of a fuzzy subgroupoid. Liu [26] defined a fuzzy subgroupoid to be a fuzzy set A of a crisp group $(G, *)$ such that $A \circ A \subseteq A$. However this definition is equivalent, even in the truly fuzzy case, to the fuzzy subgroupoid inequality, $\min\{A(x), A(y)\} \leq A(x \cdot y)$ $\forall x, y \in G$. Additionally, we have discussed a fuzzification of elements being "in" a set using fuzzy singletons. We show that for a fuzzy subgroupoid, A, if fuzzy singletons, $x_\lambda, y_\mu \in A$, then $x_\lambda \circ y_\mu \in A$. This definition has parallels to the traditional definition of a subgroupoid. We provide the following proofs showing these definitions are equivalent, as they are listed in [26].

Proposition 4.5. Let $(G, \cdot)$ be a crisp groupoid. Let $A \in [0,1]^G$. The following possible properties of A are equivalent. If A possesses any of these properties, and thus all of them, we call A a fuzzy subgroupoid of G.

(i) $min\{A(x), A(y)\} \leq A(x \cdot y) \;\; \forall x, y \in G$ (Subgroupoid Inequality)

(ii) $x_\lambda \circ y_\mu \;\in A \; \forall x_\lambda, \; y_\lambda \in A$

(iii) $A \circ A \subseteq A$

Proof. Let $(G, *)$ be a crisp groupoid. Let A be a fuzzy subset of G. Let $x, y \in G$.

Part 1) Suppose (i) holds for A, and let $x_\lambda, y_\mu \in A$. Let us show (ii); $x_\lambda \circ y_\mu \in A$. By definition of fuzzy singleton, this is equivalent to showing that $(x_\lambda \circ y_\mu)$ is a fuzzy singleton and that $(x_\lambda \circ y_\mu)(z) \leq A(z) \; \forall z \in G$. We know $(x_\lambda \circ y_\mu)$ is a fuzzy singleton by Lemma 4.4. Thus, we need only show $(x_\lambda \circ y_\mu)(z) \leq A(z) \; \forall z \in G$. Hence, let $z \in G$.

Case 1) Suppose $z = x \cdot y$. By Lemma 4.4, $(x_\lambda \circ y_\mu)(x \cdot y) = (x \cdot y)_{min\{\lambda,\mu\}}(x \cdot y)$

Thus, by definition of fuzzy singletons in A, $(x_\lambda \circ y_\mu)(x \cdot y) = \min\{\lambda, \mu\}$ Therefore, by supposition that $x_\lambda, y_\mu \in A$, we see $\min\{\lambda, \mu\} \leq min\{A(x), A(y)\}$.

Also, by hypothesis, $\min\{A(x), A(y)\} \leq A(x \cdot y)$. It then follows that, $(x_\lambda \circ y_\mu)(z) = (x_\lambda \circ y_\mu)(x \cdot y) \leq A(x \cdot y)$. However, $x \cdot y = z$; thus $A(x \cdot y) = A(z)$. Therefore, $(x_\lambda \circ y_\mu)(z) \leq A(z)$.

Case 2) $z \neq x \cdot y$, Then, also by Lemma 4.4,
$(x_\lambda \circ y_\mu)(z) = (x \cdot y)_{\min\{\lambda,\mu\}}(z) = 0 \leq A(z)$.

Therefore, $(x_\lambda \circ y_\mu)(z) \leq A(z)$ for any case. This is for any $z \in G$, thus $(x_\lambda \circ y_\mu) \in A$; (ii) holds in A.

Part 2) Suppose (ii) holds for A. Let us show (iii) $(A \circ A) \subseteq A$.

By this hypothesis, we see for any $x_\lambda, y_\mu \in A$, that $(x_\lambda \circ y_\mu) \in A$. Thus $(x_\lambda \circ y_\mu)(z) \leq A(z)\ \forall z \in G$. Therefore, as this is for any $x_\lambda, y_\mu \in A$, the inequality holds for the union of all such compositions of fuzzy singletons in A,

$$\bigcup_{x_\lambda, y_\mu \in A} (x_\lambda \circ y_\mu)(z) \leq A(z).$$

Additionally, from Lemma 4.4 ,

$$\bigcup_{x_\lambda, y_\mu \in A} (x_\lambda \circ y_\mu)(z) = (A \circ A)(z).$$

Thus $(A \circ A)(z) \leq A(z)$. This is for any $z \in G$, hence the fuzzy set $(A \circ A)$ is a fuzzy subset of A; $(A \circ A) \subseteq A$; (iii) holds in A.

Part 3) Suppose (iii). Let us show (i); $A(x \cdot y) \geq \min\{A(x), A(y)\}\ \forall x, y \in G$.

Let $x, y \in G$. By our supposition, and definition of fuzzy subset, $A(x \cdot y) \geq (A \circ A)(x \cdot y)$. Also, by definition of fuzzy set composition, $(A \circ A)(x \cdot y) = \bigvee_{a \cdot b = x \cdot y} min\{A(a), A(b)\}$. Specifically, choosing $a = x, b = y$, we find,

$$\bigvee_{a \cdot b = x \cdot y} min\{A(a), A(b)\} \geq min\{A(x), A(y)\}.$$

Therefore, $A(x \cdot y) \geq \min\{A(x), A(y)\}$, and thus, (i) holds in A.

Therefore, we have shown that (i) $\implies$ (ii) $\implies$ (iii) $\implies$ (i); the three definitions listed above of a fuzzy subgroupoid, are equivalent.

$\square$

Recall from the beginning of this section that inequality of (i), $\min\{A(x), A(y)\} \leq A(x \cdot y)\ \forall x, y \in G$, is equivalent to the definition of a crisp subgroupoid when membership values of elements are limited to $\{0, 1\}$. This is easily shown as for any $x, y \in G$, if $x, y \in A$, then $A(x) = A(y) = 1$, and thus by fuzzy subgroupoid inequality, $A(x \cdot y) = 1$. Therefore, in crisp nominclature, $x \cdot y \in A$. Also, by the proposition above, this implies that the other two definitions of fuzzy subgroupoid, when A is crisp, are equivalent to the definition of a crisp subgroupoid. For the remainder of this section, we shall use (i) as our most useful characterization of the definition of a fuzzy subgroupoid. We call this inequality of (i) the subgroupoid inequality. This definition is particularly helpful in the construction of fuzzy quotient groups.

Definition 4.6. (Rosenfeld, 1971)[43], (Liu, 1982) [26] Let $(G, \cdot)$ be a crisp groupoid. Let $A \in [0, 1]^G$, $A \neq \emptyset$. A is a fuzzy sub-groupoid IFF the subgroupoid inequality holds for elements in A.

$$min\{A(x), A(y)\} \leq A(x \cdot y) \; \forall x, y \in G$$

Example 4.7. Consider the groupoid of positive real numbers, with respect to the averaging operator, $\Diamond : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$, such that for $x, y \in \mathbb{R}^+$, $x \Diamond y = \frac{x+y}{2}$. Let $a \in \mathbb{R}^+$ be a fixed point, and represent a threshold. Define a fuzzy subset of $\mathbb{R}^+$ to be R_a such that $R_a(x) = \min\{1, \frac{x}{a}\}$. This models the set of values "almost a". For illustration, consider $a = 7$.

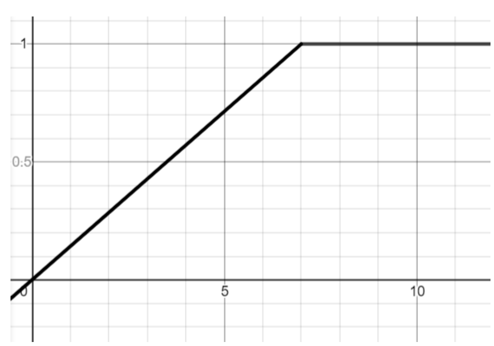

Figure 5: R_7 Membership Function

Note that the graph is non-decreasing. Therefore, the membership of the average of any two elements is greater than, or equal to, the minimum of the memberships of either of the original two, factor elements. Thus, R_7 is a fuzzy subgroupoid of $\mathbb{R}^+$. This is the case for any fixed $a \in \mathbb{R}^+$, as for $x, y \in \mathbb{R}^+$, $(x \Diamond y) \geq \min\{x, y\}$, hence $R_a(x \Diamond y) \geq \min\{R_a(x), R_a(y)\}$. Therefore the pair $(R_a, \Diamond)$ is a fuzzy subgroupoid of $(\mathbb{R}^+, \Diamond)$ for any $a \in \mathbb{R}^+$.

Example 4.8. We know the real numbers under multiplication define a crisp groupoid. Let us define a fuzzy subset of the real numbers as $O(x) = \max\{1 - |x|, 0\}$ for $x \in \mathbb{R}$, as the real numbers close to 0. To show this is a fuzzy subgroupoid of $(\mathbb{R}, \cdot)$, we need show that $\min\{O(x), O(y)\} \leq O(xy)$ $x, y \in \mathbb{R}$.

Case 1) Suppose that either $O(x) = 0$ or $O(y) = 0$. Therefore,
$\min\{O(x), O(y)\} = 0 \leq O(xy)$.

Case 2) Suppose $O(x) > 0$ and $O(y) > 0$. Therefore, $-1 < x < 1$.

Case 2a) Suppose $y > 0$. Thus, $-y < xy < y$, and thus $|xy| < |y|$. Therefore, $1 - |xy| > 1 - |y|$. Thus $O(xy) > O(y)$; the subgroupoid inequality holds.

Case 2b) Suppose $y = 0$. Thus, $xy = 0$, and thus, $O(xy) = O(y)$.

Case 2c) Suppose $y < 0$. Thus, $-y > xy > y$, and therefore, again we see, $|xy| < |y|$. Therefore, $1 - |xy| > 1 - |y|$. Thus $O(xy) > O(y)$; the subgroupoid inequality holds.

Therefore, in any case, $\min\{O(x), O(y)\} \leq O(xy)$. Therefore, $(O, \cdot)$ is a fuzzy subgroupoid of $(\mathbb{R}, \cdot)$.

4.3 From the Sprout of Groupoids, Comes the Tree of Groups

Having discussed fuzzy subgroupoids, we turn our attention now to the next useful extension, fuzzy subgroups. We have already seen a method of defining closure, with respect to an operation, when creating a fuzzy subgroupoid. We can use a similar inequality to define help us define a fuzzy subgroup. First, recall the following definitions.

Definition 4.9. Let $(G, *)$ be a crisp groupoid. We say, $(G, *)$ is a group, or simply G is a group, iff

(i) $*$ is associative

(ii) There exists a unique neutral element in G, that is
$\exists! e \in G \ s.t. \ e * x = x * e = x \ \forall x \in G$

(iii) Every element of G has an inverse with respect to $*$ in G. That is
$\forall x \in G, \ \exists y \ni x \ s.t. \ x * y = y * x = e.$

Definition 4.10. Let G be a crisp group. The pair $(A, *)$ is a crisp subgroup iff A is a subgroupoid, and $\forall x \in A, \ x^{-1} \in A$.

That is, A is a crisp subgroup if A is closed with respect to both $*$ and $(^{-1})$. We retain the definition we have already derived for closure under the binary operation $*$. The only part that needs exploring is closure with respect to $(^{-1})$, which we define in a similar manner. Recall that in the definition of fuzzy subgroupoid, we modeled the statement, A is closed under $*$ iff any output of $*$ is no less in A than at least one of its corresponding input elements. This constraint keeps us from leaving the fuzzy subgroup under $*$. Similarly, we treat $(^{-1})$ as a unary operation, and define closure under this operation such that the output of this operation is no less in A than the corresponding input; that is, $\forall x \in G, \ A(x) \leq A(x^{-1})$.

Definition 4.11. (Rosenfeld, 1971)[43] Let $(G, *)$ be a crisp group, and $A \in [0, 1]^G$. The pair $(A, *)$ is a fuzzy subgroup of G iff

(i) $(A, *)$ is a fuzzy subgroupoid of G

(ii) $A(x) \leq A(x^{-1})\ \forall x \in G$

Remark 4.12. We know, $(x^{-1})^{-1} = x$ in crisp group G. Therefore, a fuzzy subgroupoid, A, of G, is a fuzzy subgroup of G iff $A(x) = A(x^{-1})\ \forall x \in G$. Thus, this is equivalent, when limited to crisp sets, to the definition of a crisp subgroup.

Lemma 4.13. [43] Let $(G, *)$ be a group with identity, e, and let A be a fuzzy subgroup of G. Then $A(x) \leq A(e)\ \forall x \in G$.

Proof. Let $x \in G$. By definition, $A(x) = \min\{A(x), A(x^{-1})\}$.
Also, by the subgroupoid inequality, $\min\{A(x), A(x^{-1})\} \leq A(x * x^{-1}) = A(e)$. Thus, $A(x) \leq A(e)$.

$\square$

Example 4.14. Consider the group $(\mathbb{R}, +)$. Now let us consider the fuzzy subset, E of $\mathbb{R}$, for which membership values are defined by;

$$E(x) = \begin{cases} 1 & x \textbf{ is an even integer} \\ 0.5 & x \textbf{ is an odd integer} \\ 0 & x \textbf{ is not an integer} \end{cases}$$

First, let us show that $(E, +)$ is a fuzzy subgroupoid. We need show that for any real numbers $x, y \in \mathbb{R}$, that $x + y$ has membership in E at least as strong as either x or y. Thus, let $x, y \in \mathbb{R}$.
Case 1) Suppose that either x or y is not an integer. Thus $\min\{E(x), E(y)\} = 0$, and thus, regardless of the value of $E(x + y)$, we see $\min\{E(x), E(y)\} \leq E(x + y)$.
Case 2) Suppose that x and y are integers, such that one of them is an odd integer. Without loss of generality, suppose that x is any integer, and y is an odd integer. Thus, $x + y$ is an integer, and thus has membership of either 0.5 or 1 in E. Therefore, $E(y) \leq E(x + y)$.
Case 3) Suppose both x, y are both even integers. Thus, $x + y$ is an even integer, and hence, $1 = E(x) = E(y) = E(x + y)$.

Thus in any case, the inequality, $\min\{E(x), E(y)\} \leq E(x+y)$, holds. Therefore $(E, +)$ is a fuzzy subgroupoid of $\mathbb{R}$.

Now, let us show that $(E, +)$ is a fuzzy subgroup. Let $x \in \mathbb{R}$. The inverse of x with respect to addition is $-x$. We know that x is even iff $-x$ is even; x is odd iff $-x$ is odd; and x is not an integer iff $-x$ is not an integer. Thus, in any case, $E(x) = E(-x)$.

Therefore, we may conclude that $(E, +)$ is a fuzzy subgroup of $(\mathbb{R}, +)$.

Example 4.15. Consider the group of $(\mathbb{Z}_6, +_6)$ of integers mod 6, under addition mod 6. Define a fuzzy subset, A, such that $A = \{(0, 1), (1, 0.4), (2, 0.7), (3, 0.4), (4, 0.7), (5, 0.4)\}$. Through a similar case by case analysis as to the previous example, we see that $(A, +_6)$ meets the subgroupoid inequality, as $\forall x, y \in \mathbb{Z}_6$, $\min\{A(x), A(y)\} \leq A(x + y)$.

Additionally, $A(1) = A(5)$, $A(3) = A(3)$, and $A(2) = A(4)$. Thus, $A(x) = A(x^{-1})$, $\forall x \in \mathbb{Z}_6$. Therefore, we conclude that A defines a fuzzy subgroup of $\mathbb{Z}_6$.

Note that the membership value of the identity of G is not required to reach 1 for the identity of the group, but if it does, we say the subgroup is normalized. In the topic of fuzzy subgroups, based on Lemma 4.13, a fuzzy subgroup A of crisp group G is normalized iff $A(e) = 1$ for where e is the identity of the group G.

Definition 4.16. (Nguyen, Walker & Walker, 2019)[35] A fuzzy subset, A, of universal set, X, is said to be normalized iff $\exists x \in X$ such that $A(x) = 1$.

Remark 4.17. Every fuzzy subgroup corresponds to a normalized fuzzy subgroup.

This is evident through translation of the membership values by the value, $1 - A(e)$. Thus for fuzzy subgroup A of a group $(G, *)$, the corresponding normalized fuzzy subgroup is defined as $A^1(x) = (1 - A(e)) + A(x)$. In this definition, A^1 is normalized and maintains the structure of the subgroup A. Therefore, in our upcoming constructions, we may suppose A to be a normalized fuzzy subgroup of G.

There is an interesting, theorem regarding fuzzy subgroups, which is not quite as intuitive as the previous statements on fuzzy groups. We use this theorem later in the chapter to help us define one of the forms of a fuzzy quotient group.

Theorem 4.18. (Akgul, 1988)[1] Let G be a group. Let A be a fuzzy subgroup of G. Suppose $x, y \in G$. If $A(x) \neq A(y)$, then $A(xy) = \min\{A(x), A(y)\}$.

Proof. Let G be a group. Let A be a fuzzy subgroup of G. Suppose $x, y \in G$ such that $A(x) \neq A(y)$. Without loss of generality, we may suppose $A(x) > A(y)$. We know by definition of fuzzy subgroup, that $A(xy) \geq \min\{A(x), A(y)\}$. Thus we need only prove that $A(xy) \leq \min\{A(x), A(y)\}$.

Proof by contradiction:

Suppose that $A(xy) > \min\{A(x), A(y)\}$. Therefore $A(y) = A(x^{-1}(xy)) \geq \min\{A(x^{-1}), A(xy)\}$ by the fuzzy subgroupoid inequality. Also, by Remark 4.12, $(A(x) = A(x^{-1})$, thus $A(y) \geq \min\{A(x), A(xy)\}$.

From the hypothesis, $A(x) > A(y)$, therefore $A(x) > A(y) \geq \min\{A(x), A(xy)\}$. Thus $\min\{A(x), A(xy)\} = A(xy)$. Therefore, $A(y) \geq A(xy)$. However, by supposition, $A(xy) > A(y)$, thus we have a contradiction. Therefore, we conclude $A(xy) = \min\{A(x), A(y)\}$ whenever $A(x) \neq A(y)$. □

4.4 Distance from an Arbitrary Zero

The existence of a fuzzy subgroup in a group yields an ordered crisp set of crisp subgroups by order of subsethood. This also gives us the ability to think of elements in regards to their distance from the identity element, regardless of the group in which they come from. To do this, we first need to express the definition of an $\alpha-$cut, or $\alpha-$ level cut, of a fuzzy set A. For $\alpha \geq 0$, we denote the α cut of a fuzzy set A by A_α. α is an arbitrary threshold of membership imposed on a fuzzy set. If $A \in [0,1]^X$ for universal set X, and $\alpha \in [0,1]$, $A_\alpha = \{x \in X \mid A(x) \geq \alpha\}$ (Nguyen, Walker &

Walker, 2018) [35]. In Example 1.7, a fuzzy set of tall people, we created an $\alpha-$cut when only considering people 'tall' if they are 6" tall or more. This is a common method for defuzzifiying a fuzzy set into a crisp set.

Proposition 4.19. Let A be a fuzzy subset of crisp group G. A is a fuzzy subgroup of G iff A_α is either empty or a crisp subgroup of G, $\forall \alpha \in [0,1]$.

Proof. ($\Longleftarrow$)Let G be a group. Let A be a fuzzy subgroup of G, and let $\alpha \in [0,1]$. Let us show that A_α is a subgroup of G.

Case 1) If $\alpha > A(e)$, then $A_\alpha = \emptyset$ by Lemma 4.13.

Case 2) Suppose $\alpha \leq A(e)$. Thus, $e \in A_\alpha$; $A_\alpha \neq \emptyset$. Let $x, y \in A_\alpha$. By the definition of A_α, $\alpha \leq A(x), A(y)$. Additionally, by the subgroupoid inequality, $\alpha \leq \min\{A(x), A(y)\} \leq A(x * y)$. Thus $x * y \in A_\alpha$ by definition of A_α. Therefore, A_α is closed under $*$. Additionally, by Remark 4.12, for $x \in A_\alpha$, $\alpha \leq A(x) = A(x^{-1})$, thus, $x^{-1} \in A_\alpha$. Therefore, A_α is closed with respect to inverses, and thus we have shown A_α is a subgroup of G whenever $A_\alpha \neq \emptyset$.

($\Longrightarrow$) Let $(G, *)$ be a group with identity e, and suppose A is a fuzzy subset of G such that every nonempty $\alpha-$cut, A_α, is crisp subgroup of G.

(i) To show A is a fuzzy subgroupoid, we need verify
$\min\{A(x), A(y)\} \leq A(x * y)\ \forall x, y \in G$.

Let $x, y \in G$, and let $k = A(x)$, and $j = A(y)$. Therefore, $x \in A_k$ and $y \in A_j$, and hence A_j and A_k are nonempty. Without loss of generality, suppose that $k \geq j$. That is, j is a lower threshold to meet than k. Therefore, $A_k \subseteq A_j$, and thus $x \in A_j$. Hence, $x * y \in A_j$ by hypothesis that A_j is a crisp subgroup closed under $*$. Therefore, $A(x * y) \geq j = \min\{A(x), A(y)\}$ by definition of A_j. This is for arbitrary $x, y \in G$, and thus A is a fuzzy subgroupoid of G.

(ii) Let us show $A(x) \leq A(x^{-1}),\ \forall x \in G$. We prove this by contradiction. Suppose that there exists $x \in G$ such that $A(x) > A(x^{-1})$. Then there exists $\alpha \in [0,1]$ such that $A(x^{-1}) < \alpha < A(x),$. Therefore, $x \in A_\alpha$ and $x^{-1} \notin A_\alpha$. This contradicts our hypothesis that A_α is a group closed under inverses. Therefore, we conclude that $A(x) \leq A(x^{-1})$ for all $x \in G$. Furthermore, $A(x) = A(x^{-1})$ by Remark 4.12.

Therefore, by (i) and (ii), we conclude that A is a fuzzy subgroup of G.

□

Example 4.20. Consider the group $(\mathbb{Z}_6, +_6)$, and A, a fuzzy subset of $\mathbb{Z}_6$, defined by $A(0) = 1$, $A(1) = A(3) = A(5) = 0.3$, and $A(2) = A(4) = 0.5$.
Similar to Example 3.15, A defines a fuzzy subgroup of $\mathbb{Z}_6$, and by the previous theorem, each $\alpha-$cut of A defines a crisp subgroup of G.

$A_1 = \{0\}$

$A_{0.5} = \{0, 2, 4\}$

$A_{0.3} = \mathbb{Z}_6$

Definition 4.21. (Akgul, 1988) [1] If there are only a finite number of values which A attains in the interval $[0, 1]$, we call A a level fuzzy subgroup.

For a level fuzzy subgroup, we may index the values reached by A, $\{t_i\}$, in ascending order; $t_1 < t_2 < ...t_n$. Thus $A_{t_{i+1}} \subset A_{t_i}$ for all i, where $\{e\} \subseteq A_{t_n}$ and $A_{t_1} = G$. Therefore, we have constructed a bounded chain on the alpha cuts of a fuzzy subgroup of a group, with subsethood being the order. This chain is denoted by $\Gamma_A(G)$ and is called the level subgroup chain [1]. Therefore, consider an illustration of the $\alpha-$ cut subgroups of a fuzzy subgroup, with their corresponding order of subsethood.

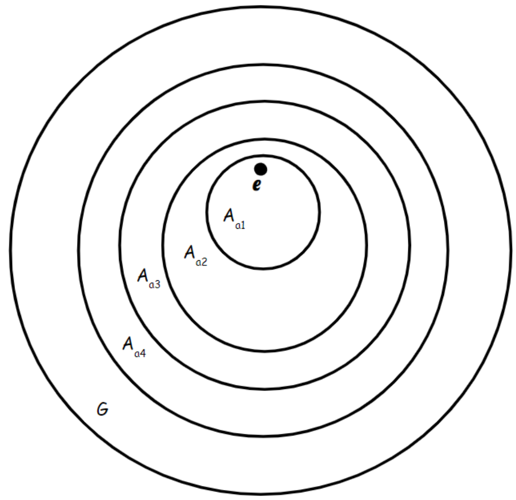

Figure 6: α- Cut Subgroups

Following this pattern, we may view a fuzzy subgroup as a three-dimensional cone, where the "tip" is the identity element, and the base is the full group. Note that it is possible for more than one element to have equal membership with the identity element. If this happens, then the "tip" of the cone is instead a disk, where the disk represents the core of the fuzzy group, and is itself a crisp subgroup. In this case, the shape becomes a frustum.

Illustrating the fuzzy, $\alpha-$cut subrgoups in this manner leads us to an intuitive, three-dimensional map of the group. Thus we can define a pseudometric on a group, induced by an arbitrary fuzzy subgroup. Consequently, as this map is centered around the identity, we may define a distance of an arbitrary element from the identity as a norm on G.

Let A be a fuzzy subgroup of a group $(G, *)$. A pseudometric on G is a nonnegative function $d : G \times G \to [0, \infty)$, such that $d(x, x) = 0\ \forall x \in G$; d is symmetric; d adheres to the triangle inequality, $d(x, z) \leq d(x, y) + d(y, z) \forall x, y, z \in G$. In the context of a fixed fuzzy subgroup A, we are interested in the pseudometric $d(x, y) = |A(x) - A(y)|$. Note that $d(x, x) = 0\ \forall x \in X$. Also, d is symmetric and adheres to the triangle inequality as a direct result of these properties holding for the absolute value metric on $\mathbb{R}$. Therefore d is a pseudometric on G. Thus, we may define a $|| \; || : G \to [0,1]$ such that $||x|| = A(e) - A(x)$. Using this definition, the function $|| \; ||$ meets certain properties that are useful in defining distance in an arbitrary algebraic structure from the identity element e.

$\forall x, y \in G$,

(i) $||e|| = 0$

(ii) $||x * y|| \leq ||x|| + ||y||$

(iii) $||x^{-1}|| = ||x||$

We thus define $|| \; ||$ to be a pseudonorm on G, in a similar style as that in (Emniyet, 2018) [12]. To make this a full norm, such that $||x|| \neq 0 \; \forall x \neq e$, we need to construct the fuzzy quotient group at the core of the fuzzy group. In a normalized fuzzy subgroup A, this core corresponds to the largest crisp subgroup which is a subset of the A.

4.5 The Quest for a Fuzzy Quotient

4.5.1 Preliminary Facts: How to be Normal

We define a fuzzy subgroup to be normal in a way that builds on our work for closure with respect to an operation and with respect to inverses. In crisp logic, we define a subgroup to be normal iff it is closed with respect to conjugates; that is, for group G and subgroup H, H is normal iff for all $a \in H$ and $x \in G$, $x^{-1}ax \in H$. In the spirit of our previous fuzzification, we alternatively state that the membership of the conjugate is at least as strong as that of the central element a. Thus, let us consider the statement $H(a) \leq H(x^{-1}ax)$. If this statement holds for any conjugates in G, then $H(x^{-1}ax) \leq H(x(x^{-1}ax)x^{-1}) = H(a)$. Hence the first inequality implies the full equality.

Definition 4.22. (Akgul, 1988) [1] Let G be a group, and H a fuzzy subgroup of G. H is said to be normal iff $H(a) = H(x^{-1}ax)$ for all $a, x \in G$.

We wish to construct a fuzzy quotient group ${}^{G}\!/_{H}$ for a normal fuzzy quotient group, H of G. We do this in a similar fashion to crisp group theory. There are two possible ways of creating a fuzzy quotient group. The first is as a fuzzy subgroup of a crisp quotient group. This method allows us to extend the pseudonorm, $|| \; ||$, of the previous section, to a true norm. However, we may also construct a fuzzy quotient group to be a crisp group of fuzzy cosets of a subgroup. This construction allows us to model fuzzy homomorphism. We finish this section by exploring both constructions.

4.5.2 Fuzzy Quotient Groups; Type 1

For the first construction, let G be a group, and H a fuzzy, normal subgroup. Consider $\overline{H} = \{h \in G \mid H(h) = 1\}$. If $\overline{H} = \emptyset$, translate H to be normalized, as in Remark 4.17. Therefore, suppose H to be normalized, thus $\overline{H} = core\{H\} \neq \emptyset$, and is thus a crisp subgroup by proposition 4.19. $\overline{H}$ may only be the singleton set, $\{e\}$, or may be a nontrivial subgroup of G itself. This $\overline{H}$ is the cap of the frustum-shaped graph of H in G.

We define the fuzzy group ${}^{G}\!/_{H}$ as a fuzzy subgroup of the crisp quotient group, ${}^{G}\!/_{\overline{H}}$. We define the membership values of cosets in ${}^{G}\!/_{H}$ in the intuitive fashion; ${}^{G}\!/_{H}(\overline{H}a) = H(a)\ \forall a \in G$ (Akgul, 1988) [1]. By definition of the core of H, ${}^{G}\!/_{H}(\overline{H}h) = 1$ iff $h \in \overline{H}$.

Theorem 4.23. (Akgul, 1988) [1] Let $(G, *)$ be a group, and H a fuzzy normalized, normal, subgroup. (i) The fuzzy subset ${}^{G}\!/_{H} \in [0,1]^{{}^{G}\!/_{\overline{H}}}$, as defined in 4.22, is a fuzzy subgroup of the crisp quotient group ${}^{G}\!/_{\overline{H}}$. (ii) Furthermore, for the natural homomorphism $\phi : G \to {}^{G}\!/_{\overline{H}}$ such that $\phi(x) = \overline{H}x$, we see that $H = {}^{G}\!/_{H} \circ \phi$. Here $\circ$ refers to traditional function composition.

Proof. Let G be a group, H a fuzzy normalized, normal subgroup of G, and $a, b \in G$.

(i) ${}^{G}/_{H}(\overline{H}a \cdot \overline{H}b) = {}^{G}/_{H}(\overline{H}ab) = H(ab)$ by crisp coset multiplication, and definition of ${}^{G}/_{H}$. Additionally, $H(ab) \geq \min\{H(a), H(b)\}$ by definition of fuzzy subgroup H. Also, $\min\{H(a), H(b)\} = \min\{{}^{G}/_{H}(\overline{H}a), {}^{G}/_{H}(\overline{H}b)\}$ by definition of ${}^{G}/_{H}$.

Thus, for $a, b \in G$, ${}^{G}/_{H}(\overline{H}a \cdot \overline{H}b) \geq \min\{{}^{G}/_{H}(\overline{H}a), {}^{G}/_{H}(\overline{H}b)\}$, and hence, the inequality of fuzzy subgroupoids, holds.

Additionally, by the inverse membership equality for fuzzy subgroup, H, we see that ${}^{G}/_{H}(Ha) = H(a) = H(a^{-1}) = {}^{G}/_{H}(\overline{H}a^{-1})$. Thus, ${}^{G}/_{H}$ maintains equal membership values for inverse elements of ${}^{G}/_{\overline{H}}$.

Therefore ${}^{G}/_{H}$ is a fuzzy subgroup of the crisp quotient group ${}^{G}/_{\overline{H}}$.

(ii) Let $a \in G$. We need to use both the definition of ${}^{G}/_{H}$ as a function into $[0, 1]$, and the definition of ϕ as the natural homomorphism, to find that $H(a) = {}^{G}/_{H}(\overline{H}a) = {}^{G}/_{H}(\phi(a))$.
Therefore, as this if for any $a \in G$, by definition of standard function composition, $H = {}^{G}/_{H} \circ \phi$. □

Note that this second statement is not the kind of group homomorphism that we would expect from crisp quotient group constructions; ϕ $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$. To establish a homomorphic fuzzy quotient group requires a secondary construction. Instead, the construction above is useful in creating a norm on G.

Recall the pseudonorm from the previous section. We redefine $|| \ ||$ as in (Emniyet, 2018) [12] .

Suppose G is a group, and H any normalized, normal subgroup of G, define $|| \ || : {}^{G}/_{\overline{H}} \to [0, 1]$, where $||\overline{H}a|| = {}^{G}/_{H}(\overline{H}) - {}^{G}/_{H}(\overline{H}a)$. $|| \ ||$ is thus a norm on ${}^{G}/_{\overline{H}}$, such that, for any $x \in {}^{G}/_{\overline{H}}$,

(i) $||x|| = 0$ iff $x = \overline{H}$

(ii) $||x * y|| \leq ||x|| + ||y||$

(iii) $||x^{-1}|| = ||x||$

4.5.3 Fuzzy Quotient Groups; Type 2

The previous method of defining fuzzy quotient groups describes a fuzzy subgroup of a group of crisp cosets. For this second method, we wish to define a fuzzy quotient group in the reverse sense, as a crisp group of fuzzy cosets. In this construction, we parallel the traditional construction of crisp quotient groups. We show that this collection of fuzzy cosets partitions the group G. We also see that the traditional operation of cosets is well-defined for fuzzy cosets, and that this quotient group of fuzzy cosets is homomorphic to the original group G.

Let us consider a group G with identity e. Let H be a normalized, normal fuzzy subgroup of G. We denote H by $[e]$, the similarity class of the identity (see Ch. 2; Fuzzy Similarity Relations). We denote the fuzzy coset Hx as the similarity class $[x]$, whose membership values are defined as $[x](a) = [e](xa^{-1})$, $\forall a, \in G$. First, we wish to show that $P = \{[x] \mid x \in G\}$ defines a fuzzy partition of G, and thus indeed corresponds to similarity classes of G by Proposition 2.25

To do this, let us first recall the definition of a fuzzy partition $\{P_i\}$ on a crisp set G, from chapter 2.

(i) $\forall P_i \in P$, P_i is normalized.

(ii) $\forall x \in G$, there is exactly one $P_i \in P$ such that $P_i(x) = 1$.

(iii) for $P_i, P_j \in P$, and $x, y \in G$, such that $P_i(x) = P_j(y) = 1$, then $P_i(y) = P_j(x) = \sup_{z \in G}(P_i \cap P_j)(z)$

$\star$ Note: To prove P is a fuzzy partition of G, we need first be aware of the following equations, which we use regularly in the following proofs. We supposed $[e]$ is a fuzzy subgroup, and thus, $[e]$ has equal membership with respect to inverses in G. Therefore, for $x, y \in G$, $[x](y) = [e](xy^{-1}) = [e](yx^{-1}) = [y](x)$. Similarly, $[e]$ is normal, thus has equal membership with respect to conjugates in G. Therefore, we see that $\forall x, y \in G$, $[x](y) = [e](xy^{-1}) = [e](y^{-1}(xy^{-1})y) = [e](y^{-1}x)$. Consequently, we conclude,

$$[x](y) = [e](xy^{-1}) = [e](y^{-1}x) = [e](x^{-1}y) = [e](yx^{-1}) = [y](x)$$

Proposition 4.24. Let G be a group, and suppose H is a normal, normalized fuzzy subgroup of G. Denote $[e] = H$ as defined above. Define a collection of fuzzy subsets of G, $P = \{[x] \mid x \in G\}$, as defined above. P forms a fuzzy partition of G.

Proof. (i) Let $[x] \in P$ for $x \in G$. Therefore, by our definition of [ ], $[x](x) = [e](xx^{-1}) = [e](e)$. However, $[e]$ was normalized by hypothesis, thus by Lemma 4.13, $[x](x) = [e](e) = 1$; hence $[x]$ is normalized.

(ii) Suppose that for $x, y, z \in G$, $[x](z) = [y](z) = 1$. We need show that $[x] = [y]$, or equivalently, $[x](a) = [y](a) \forall a \in G$. To do this, we need the following lemma, proven after the proof of this theorem.

Lemma 4.25: Suppose there exists $z \in G$ such that $[x](z) = [y](z) = 1$. Then $1 = [e](ax^{-1}ya^{-1})\ \forall a \in G$.

Now let us show $[x](a) = [y](a)\ \forall a \in G$. By note $\star$, and definition of [ ], this is equivalent to showing $[e](ax^{-1}) = [e](ya^{-1})$. We prove this by contradiction. Suppose $[e](ax^{-1}) \neq [e](ya^{-1})$ for some fixed $a \in G$. We then infer that $\min\{[e](ax^{-1}), [e](ya^{-1})\} < 1$. Also, by Theorem 4.18, we know $[e](ax^{-1}ya^{-1}) =$

$\min\{[e](ax^{-1}), [e](ya^{-1})\}$. Therefore, $[e](ax^{-1}ya^{-1}) < 1$. This contradicts Lemma 4.25 as stated above.

Therefore, we must conclude that indeed, $[e](ax^{-1}) = [e](ya^{-1})$. Additionally, $[e](ax^{-1}) = [e](xa^{-1})$ by note $\star$ above.
Thus $[x](a) = [e](xa^{-1}) = [e](ax^{-1}) = [e](ya^{-1}) = [y](a)$ for all $a \in G$.
Therefore, we have shown fuzzy set equality, $[x] = [y]$.

(iii) Let $w, w, x, y \in G$, and consider sets $[x], [y]$. Suppose $[x](u) = [y](w) = 1$. To prove the third stipulation of a fuzzy partition is equivalent to proving
$[y](w) = [x](u) = \bigvee_{z \in G} \min\{[x](z), [y](z)\}$. From property (i), we know that the specific case exists for $u = x$ and $w = y$; that is $[x](x) = [y](y) = 1$. We first show that $[y](x) = [x](y) = \bigvee_{z \in G} \min\{[x](z), [y](z)\}$. Following, we show that the general case, $[x](u) = [y](w) = 1$, is equivalent to the specific case.

First, let $z \in G$. By the definition of normal subgroup $[e]$, we see that
$[x](y) = [e](xy^{-1}) = [e](z^{-1}xy^{-1}z)$. Also, by the subgroupoid inequality,
$[e](z^{-1}xy^{-1}z) \geq \min\{[e](z^{-1}x), [e](y^{-1}z)\}$

By note $\star$, $\min\{[e](z^{-1}x), [e](y^{-1}z)\} = \min\{[x](z), [y](z)\}$. Therefore,
$[x](y) \geq \min\{[x](z), [y](z)\}$. This is for arbitrary element z, thus $[x](y) \geq \bigvee_{z \in G} \min\{[x](z), [y](z)\}$.

We still need verify $[x](y) \leq \bigvee_{z \in G} \min\{[x](z), [y](z)$. To do this, first recall that $[y](y) = 1$. Therefore, $[x](y) = \min\{[x](y), [y](y)\}$. By choosing $z = y$, we then infer that $\min\{[x](y), [y](y)\} \leq \bigvee_{z \in G} \min\{[x](z), [y](z)\}$. Consequently, $[x](y) \leq \bigvee_{z \in G} \min\{[x](z), [y](z)\}$.

Thus we have shown $\geq$ and $\leq$, and hence $[x](y) = \bigvee_{z \in G} \min\{[x](z), [y](z)\}$. The proof of $[y](x) = \bigvee_{z \in G} \min\{[x](z), [y](z)\}$ is identical to this. Therefore, we may conclude that $[y](x) = [x](y) = \bigvee_{z \in G} \min\{[x](z), [y](z)\}$.

This proves the desired equality for the special case $u = x$ and $w = y$. Recall that there may be other elements besides x and y, say $u, w \in G$, such that $[x](u) = [y](w) = 1$. If this is the case, note that $[u](u) = [w](w) = 1$ as well. Therefore, by (ii), we see that this implies $[x] = [u]$ and $[y] = [w]$. Therefore, this case is shown simply by use of the above proof. substituting u for x, and w for y, we see that

$$[u](w) = [w](u) = \bigvee_{z \in G} \min\{[u](z), [w](z)\}.$$

Therefore, by definition of fuzzy set equality, and note $\star$,

$$\bigvee_{z \in G} \min\{[u](z), [w](z)\} = \bigvee_{z \in G} \min\{[x](z), [y](z)\}.$$

Additionally, $[u](w) = [x](w)$ and $[w](u) = [y](u)$. Hence,

$$[y](u) = [x](w) = \bigvee_{z \in G} \min\{[x](z), [y](z)\}.$$

Therefore, the equation of (iii) holds for any arbitrary elements u, w, x, y such that $[x](u) = [y](w) = 1$.

$\square$

Lemma 4.25. Suppose there exists $x, y, z \in G$ such that $[x](z) = [y](z) = 1$. Then $1 = [e](ax^{-1}ya^{-1})\ \forall a \in G$.

Proof. Notice that, by our hypothesis and the definition of [ ],
$[x](z) = [e](xz^{-1}) = 1 = [e](yz^{-1}) = [y](z)$. Also, by the note $\star$ above, $[e](yz^{-1}) = [e](z^{-1}y)$, and hence, by definition of $[e]$ as a subgroup,

$$\min\{[e](x^{-1}z), [e](z^{-1}y)\} \leq [e]((x^{-1}z)(z^{-1}y)) = [e](x^{-1}y).$$

Therefore, $1 = \min\{[e](x^{-1}z), [e](z^{-1}y)\} \leq [e](x^{-1}y)$, and hence $[e](x^{-1}y) = 1$. Furthermore, by definition of normal, $[e]$ has fuzzy closure with respect to conjugates. Thus for any $a \in G$, $1 = [e](x^{-1}y) = [e](ax^{-1}ya^{-1})$. $\square$

Therefore, we have shown that the definition of fuzzy cosets, as defined above, create a fuzzy partition of the group G. We define the composition of cosets in the same way we have previously defined composition of any fuzzy subsets of a group. For $x, y, z \in G$, $([x] \circ [y])(z) = \bigvee_{ab=z} \min\{[x](a), [y](b)\}$. This leads us to the following lemma, showing that the crisp set of fuzzy cosets, P, is closed under fuzzy set composition, and homomorphic to G. We use this lemma in Theorem 4.27 to prove the fuzzy quotient group has properties parallel to those of a crisp quotient group.

Lemma 4.26. For $x, y \in G$, $[x] \circ [y] = [xy]$.

Proof. Let $x, y, z \in G$. Therefore, by definition of [ ]

$$([x] \circ [y])(z) = \bigvee_{ab=z} \min\{[x](a), [y](b)\} = \bigvee_{ab=z} \min\{[e](xa^{-1}), [e](yb^{-1})\}$$

By the note $\star$ of the previous proof,

$$\bigvee_{ab=z} \min\{[e](xa^{-1}), [e](yb^{-1})\} = \bigvee_{ab=z} \min\{[e](yb^{-1}), [e](a^{-1}x)\}$$

By composing the elements yb^{-1} with $a^{-1}x$, we can use the definition of fuzzy subgroups to conclude:

$\bigvee_{ab=z} \min\{[e](yb^{-1}), [e](a^{-1}x)\} \leq \bigvee_{ab=z} [e]((yb^{-1})(a^{-1}x)) = [e](y(ab)^{-1}x) = [e](yz^{-1}x)$

Fuzzy closure of normal $[e]$ under conjugates then implies

$[e](yz^{-1}x) = [e](x(yz^{-1}x)x^{-1}) = [e](xyz^{-1}) = [xy](z)$. Thus, $([x] \circ [y])(z) \leq [xy](z)$.

To prove inequality $\geq$, recall $([x] \circ [y])(z) = \bigvee_{ab=z} \min\{[e](xa^{-1}), [e](yb^{-1})$. Thus choose $a = zy^{-1}$ and $b = y$. Hence, we see that

$$([x] \circ [y])(z) \geq \min\{[e](x(zy^{-1})^{-1}), [e](yy^{-1})\} = \min\{[e](xyz^{-1}), [e](e)\}.$$

Because $[e]$ is normalized, by Lemma 4.13, $[e](e) = 1$. Therefore,

$$\min\{[e](xyz^{-1}), [e](e)\} = \min\{[e](xyz^{-1}), 1\} = [e](xyz^{-1}) = [xy](z).$$

Thus $([x] \circ [y])(z) \geq [xy](z)$. Therefore, we conclude that $[x] \circ [y](z) = [xy](z)$. As this is for any $z \in G$, we have shown fuzzy set equality: $[x] \circ [y] = [xy]$.

$\square$

From the previous proofs, we have seen that $P = \{[x]\}$ is a fuzzy partition of G; we have additionally defined a binary operation, $\circ$, which is closed in the set $\{[x]\}$. Thus we may prove results equivalent to those stated in (Morsi, 1994)[32], which correspond to properties of crisp quotient groups.

Theorem 4.27. Let G be a group with identity e. Suppose $[e]$ denotes any normal, normalized fuzzy subgroup of G. Let the membership of elements in the composition of cosets be as defined in Lemma 4.26. Thus, the following properties hold.

(i) $[x] \circ [y]$ is well defined

(ii) $[e]$ is the idenity of the operation $\circ$

(iii) $[x^{-1}] = [x]^{-1}$

Proof. Let G be a group with identity e. Suppose $[e]$ denotes a fuzzy, normalized, normal subgroup of G. Let $w, x, y, z \in G$.

(i) Suppose that $[x] = [w]$ and $[y] = [z]$. By definition of fuzzy set equality, this implies that $\forall c \in G$, $[x](c) = [w](c)$ and $[y](c) = [z](c)$. Therefore,

$$([x] \circ [y])(c) = \bigvee_{ab=c} ([x](a), [y](b)) = \bigvee_{ab=c} ([w](a), [z](b)) = [w] \circ [z](c)$$

Thus we have shown fuzzy set equality; $([x] \circ [y]) = ([w] \circ [z])$. Therefore, $\circ$ is well defined. Hence, by Lemma 4.26, $\{[x]\}$ is a groupoid.

(ii) Additionally, by Lemma 4.26, $[x] \circ [e] = [xe] = [x] = [ex] = [e] \circ [x]$. Thus $[e]$ acts as the identity of the fuzzy groupoid $\{[x]\}$.

(iii) Also by Lemma 4.26,
$[x] \circ [x^{-1}] = [xx^{-1}] = [e] = [x^{-1}x] = [x^{-1}] \circ [x]$. Thus $[x^{-1}] = [x]^{-1}$.

$\square$

We thus denote ${}^{G}/_{[e]} = \{[x] \mid x \in G\}$; we then define $({}^{G}/_{[e]}, \circ)$, to be the fuzzy quotient group of G mod $[e]$. Additionally, by may employ the natural homomorphism $\phi : G \to {}^{G}/_{[e]}$, where $\phi(x) = [x]\ \forall x \in G$. Hence, similar to crisp logic, we see that $\forall x, y \in g$, $\phi(x \cdot y) = [x \cdot y] = [x] \circ [y]$. Therefore, G is homomorphic to ${}^{G}/_{[e]}$.

From this work, we see that for any normal, normalized fuzzy subgroup, H, of a crisp group G, we may denote H as the similarity class of the identity, and define a quotient group with similar characteristics as that of a crisp quotient group. This construction is a crisp set of fuzzy sets, rather than a fuzzy set of crisp sets.

4.6 Places to Go Next

In this chapter, we have thoroughly explored fuzzy subgroups in relation to crisp groups. Fuzzy algebra does not end here, and these topics continue to develop. The subject is expansive, and could easily fill a few books. Furthermore, research is regularly being published on the topic. For the reader wishing to learn more, the following are two current research papers that build on the concepts discussed in this chapter.

(i) (Seselja & Tepavcevic), 2020) [44] In "Ω-groups in the Language of Ω-groupoids" (2020), B. Šešelja and A. Tepavčević combine concepts of $L-fuzzy$ sets (discussed in chapter 3), with topics from this chapter, to create a theory of fuzzy $\Omega-$groups. Here Ω is an arbitrary complete lattice.

(ii) (Emniyet & Sahin, 2018) [12] In "Fuzzy Normed Rings"(2018), A. Emniyet and M. Şahin, expand principles discussed in this chapter to develop fuzzy ring theory. This includes norms on said rings. These constructions are formed in a manner which is an intuitive extension of the fuzzy group theory discussed in this chapter.

CHAPTER 5

Fuzzy Metrics

We now turn our attention from algebra, to analysis. In chapters 2 and 4, we have defined constructions on fuzzy sets corresponding to their equivalent constructions on crisp sets. We could continue to do this for fuzzy metrics, and define a metric a fuzzy set; a sense of distance between fuzzy points. This type of construction is explored in detail by (Xia & Guo, 2004) [49]. We presented a similar concept previously in chapter 2, a fuzzy similarity relation. This fuzzy similarity relation models the "nearness " of two elements of a fuzzy set, as defined by their similarity to one another.

In the general fuzzy theory community, a more commonly used fuzzy metric theory creates a vague bound on the value of a crisp metric. Recall from chapter 1, that a prime application of fuzzy logic is in modelling uncertainty. In this chapter, we use a fuzzy metric as a way to model the degree of certainty to which two elements of a crisp set are within a distance, t, apart. These fuzzy metric spaces have properties correlating directly with those of crisp metric spaces. Additionally, a fuzzy metric space can be interpreted as uncountably many crisp semimetrics working in tandem under certain rules. We show that a crisp metric space induces a fuzzy metric space, but the converse is not necessarily true. Using these fuzzy spaces, we define similar structures and find similar results as to those of crisp metric spaces. These include a form of Cauchy and convergent sequences, and a fuzzy topology induced by the fuzzy metric. We also define a form of special types of mappings from a fuzzy metric space into itself; namely a contraction mapping and a continuous mapping. Finally, we show the corresponding fixed point theorem of contraction mappings.

5.1 Uncertain Times Call for Uncertain Metrics

Consider a crisp metric space (X, d). d is understood to be a function from X^2 into $[0, \infty)$, meeting certain properties; namely, for all $x, y, z \in X$,

(i) $d(x, y) = 0$ iff $x = y$

(ii) $d(x, y) = d(y, x)$

(iii) $d(x, z) \leq d(x, y) + d(y, z)$: triangle inequality.

Because d is a function, we may interpret d as a crisp subset of the cross product, $X \times X \times [0, \infty)$. For any structure defined by a crisp subset, our goal is to form a twin construction using a fuzzy subset. Therefore, if d is a subset of a universal set $X \times X \times [0, \infty)$, we may attempt to define M, a fuzzy metric, as a fuzzy subset of $X \times X \times [0, \infty)$.

Of course, merely defining a fuzzy metric, M as any fuzzy subset of $X \times X \times [0, \infty)$ would be too broad, and would yield no useful theorems. Rather consider a crisp, nonempty set X, a fuzzy subset M of $X \times X \times [0, \infty)$, and a continuous T-norm, $*$. We define a triple, $(X, M, *)$, to be a fuzzy metric space if it satisfies the following properties.

Definition 5.1. (George & Veeramani, 1994) [13] Let X be a nonempty, universal set, M a fuzzy subset of $X \times X \times [0, \infty)$, and $*$ a continuous T-norm. We say that $(X, M, *)$ is a fuzzy metric space iff for $x, y, z \in X$ and $t, s \in (0, \infty)$;

(i) $M(x, y, t) \neq 0$

(ii) $M(x, y, t) = 1 \Leftrightarrow x = y$

(iii) $M(x, y, t) = M(y, x, t)$

(iv) $M(x, y, t) * M(y, z, s) \leq M(x, z, t + s)$

(v) for any fixed $x, y \in X$, the function $M(x, y, t) : (0, \infty) \rightarrow [0, 1]$ is continuous; $\forall\ \epsilon > 0, \exists\ \delta > 0$ such that if $|t - s| < \delta$ then $|M(x, y, t) - M(x, y, s)| < \epsilon$.

It is useful to interpret these fuzzified properties of metric spaces as a degree of certainty of an upper bound on distance between two points. For example, suppose we as an observer are in a public setting, but have lost our glasses. Without these glasses, our vision is quite poor.

Figure 7: Uncertainty of Distance: Blurry Vision [24]

We see a person that is possibly our friend in the distance. However, the image is blurry, so blurry in fact that we are unsure of how far away the person is. Even though we are not certain exactly how far away they are, we are nearly certain (certainty of 0.99) that they are less than 100 meters away. Using our blurry vision, we can also make a reasonable assumption, with certainty of 0.8, that they are less than 10 meters away. Continuing, we may say that we have certainty = 0.5 that the person is less then 8 meters away.

By thinking of fuzzy metrics in this manner, we interpret the properties of a fuzzy metric as follows. Property (i) tells us that there is always some degree of certainty that the distance between any two elements is bound above by a specified positive real number. This is the case even if this degree be a vanishingly small certainty. (ii) indicates our universal certainty that for any element has distance 0 from itself, and thus a distance less than any positive upper bound t. (iii) is a direct correlation to the symmetry of metric spaces. That is, any two elements x_1 and x_2 hold the same distance from point x_1 to point x_2 as from point x_2 to point x_1; thus, we have the same certainty that both distances are below a set bound t.

The fourth property, (iv), is a generalization of a weaker form of the triangle inequality. It follows from a more intuitive sense of distance in a space. Consider elements x, y, z in a crisp metric space (X, d). We know that $d(x, y)+d(y, z) \geq d(x, z)$; that is, the path from x to z is minimized when connected directly, rather than routing through a tertiary element y. Therefore, let us suppose we are confident to degree k that $d(x,y) < t$, and to degree j that $d(y,z) < s$; $t, s, \in (0,\infty)$ and $j, k \in [0,1]$. Then we must be at least $j \wedge k$ confident that $d(x,z) < t+s$. In the intuitive understanding, we use the T-norm $\wedge$, but this can be generalized to any continuous T-norm on $[0,1]$.

The fifth property, (v), is a new stipulation to account for the new variable, $t \in (0,\infty)$, and does not directly correspond to a crisp metric space property. It is included in the definition for its logical implications. Continuity gives us the ability to better define a fuzzy openness and fuzzy closedness of subsets in a fuzzy metric space. Thus, for our purposes, we include this stipulation that $M(x,y,t) : (0,\infty) \to [0,1]$ be continuous, $\forall x, y \in X$.

Fuzzy metric spaces may exist on their own, but many common constructions of fuzzy metrics are built from crisp metrics. There are multiple ways of incorporating the fuzzy uncertainty element t into crisp metric theory. One such method is a fuzzification of a crisp metric space as presented in (George Veeramani, 1994)[13]. The following provides an example of this.

Example 5.2. Consider the triple $(\mathbb{R}, M, \cdot)$; $\cdot$ is standard, real number multiplication, and M is the proposed fuzzy metric, defined as $M(x,y,t) = e^{\frac{-|x-y|}{t}}$.

(i) Based on the characteristic horizontal asymptote of the exponential function, for any $t \in (0,\infty)$, $M(x,y,t) > 0$ $\forall x, y \in X$.
(ii) $M(x,y,t) = 1$ iff $\frac{|x-y|}{t} = 0$ iff $x = y$.
(iii) M is symmetric by the symmetric property of the absolute value function.
(iv) By the triangle inequality of the absolute value metric,
$|x-z| \leq |x-y| + |y-z| \leq \left(\frac{t+s}{t}\right)|x-y| + \left(\frac{t+s}{s}\right)|y-z|$.
Therefore, $\frac{|x-z|}{t+s} \leq \frac{|x-y|}{t} + \frac{|y-z|}{s}$, and hence, $e^{\frac{|x-z|}{t+s}} \leq e^{\frac{|x-y|}{t}} \cdot e^{\frac{|y-z|}{s}}$.
Thus, $e^{\frac{-|x-z|}{t+s}} \geq e^{\frac{-|x-y|}{t}} \cdot e^{\frac{-|y-z|}{s}}$. Hence $M(x,z,t+s) \geq M(x,y,t) \cdot M(y,z,s)$.
(v) Let $x, y \in X$ be fixed. By continuity of the rational exponential function, we find

that $M(x, y, t) : (0, \infty) \to [0, 1]$ is continuous. Thus $(\mathbb{R}, M, \cdot)$ is a fuzzy metric space.

We previously interpreted a fuzzy metric as a fuzzy upper bound on a sense of distance. Therefore, we would expect that if a bound increases, the certainty of a fixed value being below that bound increases as well; rather it at least does not decrease. For example, if two elements of a set x and y are, with degree of certainty $\frac{1}{2}$, nearer to one another than 5, then x and y are, with degree greater than or equal to $\frac{1}{2}$, nearer to one another than 6. Thus, let us prove the following lemma from (Grabiec, 1988) [14].

Lemma 5.3. Let $(X, M, *)$ be a fuzzy metric space. Fix $x, y \in X$. Then $M(x, y, t) : (0, \infty) \to [0, 1]$ is non-decreasing as a function of t.

Proof. Let $(X, M, *)$ be a fuzzy metric space. Let $x, y \in X$ be fixed. Let $t, s, \in (0, \infty)$. The case that $t = s$ is trivial; $M(x, y, t) = M(x, y, s)$. Therefore, suppose that $t < s$. Let us show, $M(x, y, t) \leq M(x, y, s)$.

Notice, by property (iv) of fuzzy metric M, we see $M(x, y, s) \geq M(x, y, t) * M(y, y, s - t)$. However, by property (i) of a fuzzy metric, $M(y, y, s - t) = 1$. Additionally, by property (i) in definition 3.5 of T-norm $*$, $M(x, y, t) * M(y, y, s - t) = M(x, y, t) * 1 = M(x, y, t)$. Therefore, $M(x, y, s) \geq M(x, y, t) * M(y, y, s - t) = M(x, y, t) * 1 = M(x, y, t)$. Thus, for any fixed $x, y \in X$, the function $M(x, y, t) : (0, \infty) \to [0, 1]$ is non-decreasing. □

Proposition 5.4. (Paknazar & De La Sen, 2020)[38] A fuzzy metric space can be induced by a crisp metric space.

Proof. Let (X, d) be a metric space; suppose $x, y \in X$ and $t \in (0, \infty)$. Define $M_d(x, y, t) = \frac{t}{t+d(x,y)}$. Consider the T-norm $\cdot$ of real number multiplication. We wish to show that $(X, M_d, \cdot)$ is a fuzzy metric space.

$M_d(x,y,t) > 0$ $\forall t \in (0,\infty)$ by construction of M_d. Thus property (i) of fuzzy metrics is met. Using the symmetric property of crisp metric d, $x = y$ iff $d(x,y) = 0$ iff $M_d(x,y,t) = 1$ $\forall t$, thus condition (ii) is met. Additionally, (iii) is inherited directly by the symmetric property of d. Property (iv) is less trivial to show.

Let $x, y, z \in X$ and $t, s \in (0,\infty)$. We need show:

$$M_d(x,z,t+s) \geq M_d(x,y,t) \cdot M_d(y,z,s)$$

First note that by definition of M_d, $M_d(x,z,t+s) = \frac{t+s}{t+s+d(x,z)}$

$= \frac{(t+s)^2}{(t+s)^2+(t+s)d(x,z)}$ by multiplication of $(t+s)$

$\geq \frac{(t+s)^2}{(t+s)^2+(t+s)d(x,y)+(t+s)d(y,z)}$ by the triangle inequality of d

$\geq \frac{(t+s)^2}{(t+s)^2+(t+s)d(x,y)+(t+s)d(y,z)+d(x,y)d(y,z)}$ by addition of $d(x,y) \cdot d(y,z)$

$= \frac{(t+s)(t+s)}{((t+s+d(x,y))(t+s+d(y,z))} = \frac{t+s}{t+s+d(x,y)} \cdot \frac{t+s}{t+s+d(y,z)}$

Note that $\frac{t}{t+d(x,y)} \leq \frac{t+s}{t+s+d(x,y)}$, and $\frac{s}{s+d(y,z)} \leq \frac{t+s}{t+s+d(y,z)}$. Thus by the order preserving property (iv) of T-norm $\cdot$, $\frac{t+s}{t+s+d(x,y)} \cdot \frac{t+s}{t+s+d(y,z)} \geq \frac{t}{t+d(x,y)} \cdot \frac{s}{s+d(y,z)}$; by definition, $\frac{t}{t+d(x,y)} \cdot \frac{s}{s+d(y,z)} = M_d(x,y,t) \cdot M_d(y,z,s)$. Therefore, property (iv) holds for M_d.

Finally, let us consider property (v). Fix $x, y \in X$. Therefore, $d(x,y)$ is fixed; hence as a function of t, $M_d(x,y,t) = \frac{t}{t+d(x,y)}$. Thus $M_d(x,y,t) : (0,\infty) \to [0,1]$ is continuous, and property (v) holds. Therefore, we see that $(X, M_d, \cdot)$ does indeed define a fuzzy metric space. Moving forward, we refer to this construction of M_d when referring to the fuzzy metric induced by the crisp metric d.

$\square$

Remark 5.5. The converse of Proposition 5.4 is not necessarily true. That is, let $(X, M, *)$ be a fuzzy metric space and fix $t \in (0,\infty)$. Let d_t be a function generated conversely as in 5.4; $d_t(x,y) = \frac{t}{M(x,y,t)} - t$. d_t may not be a metric on X. This construction falls short of a crisp metric in that the triangle inequality may not hold.

As a counterexample, consider the fuzzy metric space defined in Example 5.2; $(\mathbb{R}, M, \cdot)$; $M(x,y,t) = e^{\frac{-|x-y|}{t}}$. Fix $t = 1$, and consider the proposed metric d_1, such that $d_1(x,y) = \frac{1}{M(x,y,1)} - 1 \; \forall x,y \in \mathbb{R}$.

Consider elements $0, 3, 8 \in \mathbb{R}$. By definition of M, we find that $M(0,3,1) = e^{-3}$, $M(3,8,1) = e^{-5}$, and $M(0,8) = e^{-8}$. Thus, by definition of d_1, $d_1(0,3) = e^3 - 1$, $d_1(3,8) = e^5 - 1$, and $d_1(x,z) = e^8 - 1$; that is $d_1(0,8) \approx 2979.96$, and
$d_1(0,3) + d_1(3,5) \approx 166.50$. Therefore, $d_1(0,8) > d_1(0,3) + d_1(3,8)$, and thus the triangle inequality does not hold; d_1 is not a metric on $\mathbb{R}$.

Therefore, to induce a crisp metric space from a fuzzy metric space, we need define an additional property. We call a fuzzy metric triangular if we obtain the triangle inequality for an induced crisp semimetric as defined above; thus we form a full crisp metric.

Definition 5.6. (Paknazar & De La Sen, 2020) [38] A fuzzy metric M of a fuzzy metric space $(X, M, *)$ is called triangular iff $\forall x, y \in X$ and $t \in (0, \infty)$;

$$\frac{1}{M(x,y,t)} - 1 + \frac{1}{M(y,z,t)} - 1 \geq \frac{1}{M(x,z,t)} - 1$$

Proposition 5.7. A triangular fuzzy metric space induces infinitely many crisp metrics.

Proof. Let $(X, M, *)$ be a triangular fuzzy metric space. Fix $t \in (0, \infty)$. Define a real valued function d_t on X^2 such that $\forall x, y \in X$, $d_t(x,y) = \frac{t}{M(x,y,t)} - t$.

For $x, y, \in X$, $M(x,y,t) \leq 1$; therefore, it is true that $d_t(x,y) \geq 0$. Additionally, by property (i) of fuzzy metric M, $M(x,y,t) = 1$ iff $x = y$. Therefore, $d(x,y) = 0$ iff $x = y$. Also, by the symmetric property of M, $d_t(x,y) = d_t(y,x)$. Finally, by supposition that M is triangular, we find that

$$\frac{1}{M(x,y,t)} - 1 + \frac{1}{M(y,z,t)} - 1 \geq \frac{1}{M(x,z,t)} - 1$$

Therefore, by multiplying by t,

$$\frac{t}{M(x,y,t)} - t + \frac{t}{M(y,z,t)} - t \geq \frac{t}{M(x,z,t)} - t$$

Thus, $d_t(x,y) + d_t(y,z) \geq d_t(x,z)$; d_t adheres to the triangle inequality.This is for any fixed $t \in (0, \infty)$, thus we have shown that M induces an infinite set of crisp metrics, $\{d_t \mid t \in (0, \infty)\}$ on X.

□

We have shown in the previous proofs and remarks that a fuzzy metric space does not always directly correlate to a crisp metric; however a crisp metric does correlate to a fuzzy metric on a set. Also, when a fuzzy metric M, does induce a crisp metric, it induces infinitely many crisp metrics $\{d_t \mid t \in (0, \infty)\}$ on the universal set. Therefore, a fuzzy metric space can be interpreted as infinitely many crisp semimetric spaces connected by a single variable of certainty t. Instead of the triangle inequality, we only assume the weaker conditions, (iv) and (v) of fuzzy metrics. These conditions are enough to build a fuzzy metric theory twin to that of crisp metric theory. To begin this process, we first explore some preliminary results of sequences in fuzzy metric spaces.

5.2 Convergence of Certainty; Certainly Converging

We have defined fuzzy metric spaces and considered their relationship with crisp metric spaces. To further develop the theory of fuzzy metrics, we now turn our attention to sequences of points in fuzzy metric spaces. It is useful to remind ourselves that these fuzzy metric spaces are not on a fuzzy set, but a crisp set X; that is, the fuzziness is of the uncertainty of the distance between two points. Thus, when we define Cauchy and convergent sequences in fuzzy metric spaces, we interpret this as a convergence of certainty. Suppose we look into an arbitrary world containing a sequence, and our vision is perfect. We think of this sequence as convergent in a crisp metric space if we see that the sequence approaches a unique point. However, if our vision is blurry, the point we approach may be not be clear. If we as the observer move from point to point along the sequence, we may see a blurry point in the space which the sequence approaches. This point becomes more clear as the observer traverses further along the sequence; thus the observer becomes more certain of how far away the limit point is.

Definition 5.8. (George Veeramani, 1994) [13] Let $(X, M, *)$ be a fuzzy metric space. A sequence $\{x_n\} \subseteq X$ is called Cauchy iff for any fixed $t > 0$, $\forall \epsilon \in (0, 1)$, $\exists N \in \mathbb{N}$ such that $M(x_n, x_m, t) > 1 - \epsilon$ for all $n, m \geq N$. Equivalently, $\{x_n\}$ is Cauchy iff $\lim\limits_{n,m\to\infty} M(x_n, x_m, t) = 1 \; \forall t > 0$.
In this definition, the sequence approaches complete certainty 1 that the distance between two points of the sequence is less than any upper bound. Consequently the inequality is flipped in contrast to the traditional metric theory definition.

Proposition 5.9. Consider a Cauchy sequence $\{x_n\}$ in a crisp metric space (X, d). If M_d is the fuzzy metric induced by d, then the sequence $\{x_n\}$ is Cauchy in the induced fuzzy metric space $(X, M_d, \cdot)$.

Proof. Fix $t \in (0, \infty)$ and let $\epsilon > 0$. We wish to find N such that if $n, m \geq N$, then $M_d(x_n, x_m, t) > 1 - \epsilon$.

Since t and ϵ are fixed, let $\delta = \frac{t}{1-\epsilon} - t$. By hypothesis, $\{x_n\}$ is Cauchy with respect to d. Thus we may choose N such that if $n, m \geq N$, then $d(x_n, x_m) < \delta$.
Therefore, $d(x_n, x_m) < \frac{t}{1-\epsilon} - t$. Thus $d(x_n, x_m) + t < \frac{t}{1-\epsilon}$. Hence $\frac{1}{d(x_n,x_m)+t} > \frac{1-\epsilon}{t}$, and thus $\frac{t}{d(x_n,x_m)+t} > 1 - \epsilon$. Therefore, we conclude $M_d(x_n, x_m, t) = \frac{t}{d(x_n,x_m)+t} > 1 - \epsilon$.

As this is for any such $t > 0$ and $\epsilon \in (0, 1)$, we may choose such an N as above such that $M_d(x_n, x_m, t) > 1 - \epsilon$ for all $t > 0$ and $\epsilon \in (0, 1)$. We conclude that the sequence is M_d-Cauchy. $\square$

Remark 5.10. Through a similar proof, if a crisp metric d induces a fuzzy metric M, and a sequence is Cauchy with respect to M, then the sequence is Cauchy with respect to d.

Example 5.11. From the above proposition, every example of a Cauchy sequence of a crisp metric space is Cauchy with respect to the induced fuzzy metric space.
1. $\{\frac{sin(n)}{n}\}$ is Cauchy in $(\mathbb{R}, M_{|\cdot|}, \cdot)$, where $M_{|\cdot|}$ is the fuzzy metric induced by the absolute value metric.
2. $\{\langle \frac{1}{n}, \frac{2+2n}{n}, \frac{3+3n}{n} \rangle\}$ is Cauchy in $(\mathbb{R}^3, M_E, \cdot)$, where M_E is the fuzzy metric induced by the Euclidean metric E, on $\mathbb{R}^3$.

Definition 5.12. A sequence $\{x_n\}$ in a fuzzy metric space $(X, M, *)$ is said to be convergent to a point $x \in X$ iff $\forall \epsilon > 0$, there exists $N \in \mathbb{N}$ and $t > 0$ such that $M(x_n, x, t) > 1 - \epsilon$. Equivalently, $\{x_n\}$ is convergent to x iff, for all $t > 0$, $\lim_{n \to \infty} M(x_n, x, t) = 1$.

Recall that when discussing fuzzy metrics, we wish to 'converge' in certainty to a point. To converge in certainty, we approach absolute certainty of a state of being. It is necessary that we verify the uniqueness of a point of convergence. That is, if $\{x_n\}$ is a sequence in $(X, M, *)$, then $\{x_n\}$ cannot converge to more than one point in the space X.

Proposition 5.13. Let $\{x_n\}$ be a convergent sequence of points in a fuzzy metric space $(X, M, *)$. Suppose that $\{x_n\}$ converges to points x and y in X. Then $x = y$.

Proof. Let $\epsilon > 0$, and fix $t > 0$.

Notice that by property (iv) of a fuzzy metric space,
$M(x, y, t) \geq M(x_n, x, \frac{t}{2}) * M(x_n, y, \frac{t}{2})$. Recall that the T-norm $*$ is continuous, thus we may choose δ such that if $M(x_n, x, \frac{t}{2}) > 1 - \delta$ and $M(x_n, y, \frac{t}{2}) > 1 - \delta$, then $M(x_n, x, \frac{t}{2}) * M(x_n, y, \frac{t}{2}) > 1 - \epsilon$.

Thus because we have hypothesized that x_n converges to x and to y , we may choose N_x and N_y such that if $n \geq \max\{N_x, N_y\}$, then $M(x_n, x, \frac{t}{2}) > 1 - \delta$ and

$M(x_n, y, \frac{t}{2}) > 1 - \delta$, and therefore, $M(x, y, t) \geq M(x_n, x, \frac{t}{2}) * M(x_n, y, \frac{t}{2}) > 1 - \epsilon$. Since this is for arbitrary $\epsilon > 0$, and x, y and t are fixed, then $M(x, y, t) = 1$. Therefore, by property (ii) of a fuzzy metric space, $x = y$; the limit of a convergent sequence in a fuzzy metric space is unique.

□

We have shown through Proposition 5.9 and Remark 5.10 that there is a twin correspondence of Cauchy sequences in a crisp metric space, and the fuzzy metric space induced by the crisp metric space. That is, suppose $(X, M_d, *)$ is the fuzzy metric induced by a crisp metric d on X. For a sequence $\{x_n\}$ of points in X, $\{x_n\}$ is a Cauchy sequence with respect to M_d iff $\{x_n\}$ is Cauchy with respect to d. Similarly, we wish to show that for a sequence $\{x_n\}$ of points in X, $\{x_n\}$ is convergent in a crisp metric d on X iff $\{x_n\}$ is convergent to x with respect to the fuzzy metric, M_d.

Proposition 5.14. Let $(X, M_d, *)$ be a fuzzy metric space, where M_d is the fuzzy metric induced by a crisp metric d. If $\{x_n\}$ is a sequence of points in X which converge to $x \in X$ with respect to M_d, then $\{x_n\}$ is convergent to x with respect to d.

Proof. Let $\{x_n\}$ be a sequence of points in X which converge to a point $x \in X$ with respect to a fuzzy metric M_d, induced by a crisp metric d.

By definition of convergence with respect to M_d, then the sequence of real numbers $M_d(x_n, x, t)$ converges to 1 for any t. Thus, let us choose $t = 1$, and suppose $\epsilon > 0$. We may choose $\delta = \frac{\epsilon}{1+\epsilon}$. Therefore, by the definition of fuzzy convergence, we may choose $N \in \mathbb{N}$ such that $\frac{1}{1+d(x_n,x)} > 1 - \delta$. Therefore, by choice of δ, we see that $\frac{1}{1+d(x_n,x)} > 1 - \frac{\epsilon}{1+\epsilon} = \frac{1}{1+\epsilon}$. Hence, $d(x_n, x) < \epsilon$. Therefore, $\{x_n\}$ converges to the point x with respect to d.

□

Remark 5.15. It is similar to prove that a sequence which is convergent with respect to a crisp metric is convergent to the same point with respect to the induced fuzzy metric.

Definition 5.16. (Grabiec, 1984) [14] A fuzzy metric space $(X, M, *)$ is called complete iff every Cauchy sequence in $(X, M, *)$ is convergent to a point in X.

From the above propositions and remarks, it follows that a fuzzy metric space $(X, M_d, *)$, where M_d is induced by crisp metric d, is complete iff (X, d) is complete in the crisp sense.

5.3 Fuzzy Topology

Using the absolute value metric on $\mathbb{R}$ for inspiration, mathematicians have derived a theory of analysis that includes continuous mappings and contraction mappings. Mathematicians generalized this to definitions using any arbitrary metric on any set; we now explore such a theory on fuzzy metric spaces.

First, we need define some basic topology of fuzzy metrics. Fuzzy closed sets are defined in the same fashion as in crisp metric theory. Fuzzy open balls are defined in similar a similar manner as in crisp metric theory; consequently, we define an open set in a fuzzy metric space in the same manner as in crisp metric theory. These definitions yield similar theorems to those of crisp metric theory. We state these relevant definitions below; we then show that these definitions lead to twin theorems of crisp metric spaces.

Definition 5.17. A fuzzy open ball centered at x, with radius, ϵ and bound t, is the crisp set $B[x, \epsilon, t] = \{y \in X \mid M(x, y, t) \geq 1 - \epsilon\}$

Definition 5.18. A subset A of X is said to be open iff $\forall x \in A$ and $t > 0$, there exists $\epsilon \in (0, 1)$ such that $B[x, \epsilon, t] \subseteq A$.

Definition 5.19. A subset A of X is said to be closed with respect to M iff for any sequence $\{a_n\}$ of points in A, if $\{a_n\}$ is convergent to a point a, then $a \in A$.

Proposition 5.20. A set which is open with respect to a crisp metric space is open in the induced fuzzy metric space.

Proof. Consider a crisp metric space (X, d), and $(X, M_d, \cdot)$ the induced fuzzy metric space. Let A be an open set with respect to d. Suppose $x \in A$, and fix $t \in (0, \infty)$. We wish to find an $\epsilon \in (0, 1)$ such that $B[x, \epsilon, t] \subseteq A$, in the crisp sense. We know that since A is open in (X, d), then there must exist $\delta > 0$ such that the crisp open ball $B[x, \delta] \subseteq A$.

Consider $\epsilon = \frac{\delta}{t+\delta}$, and suppose that $y \in B[x, \epsilon, t]$. We wish to show that $y \in A$. By definition of the fuzzy open ball, $M_d(x, y, t) > 1 - \epsilon$. Since M_d is induced by d, this implies $\frac{t}{t+d(x,y)} > 1 - \epsilon = 1 - \frac{\delta}{t+\delta} = \frac{t}{t+\delta}$. Therefore, $d(x, y) < \delta$. Thus, by choice of δ using the definition of the crisp open ball, $y \in A$. Therefore, the fuzzy open ball, $B[x, \epsilon, t] \subseteq A$, and hence, A is open in $(X, M, *)$.

$\square$

Example 5.21. From Proposition 5.20, any examples of open sets in crisp metric spaces are open sets in their induced fuzzy metric spaces.

1. Consider $a, b \in \mathbb{R}$ such that $a < b$. The interval (a, b) is open in $(\mathbb{R}, M_{|\cdot|}, \cdot)$, where $M_{|\cdot|}$ is the fuzzy metric induced by the absolute value metric.
2. For any $a \in (0, \infty)$, the set $\{\langle x, y, z\rangle \in \mathbb{R}^3 \mid 3 < \sqrt{x^2 + y^2 + z^2} < a\}$ is open with respect to the fuzzy metric space $(\mathbb{R}^3, M_E, \cdot)$, where M_E is the fuzzy metric induced by the Euclidean metric on $\mathbb{R}^3$.

Now, as in crisp logic, a set which is not open with respect to a fuzzy metric, may or may not be closed; a set which is not closed may or may not be open. However, we can show that in a fuzzy metric space, the compliment of an open set is closed, and the compliment of a closed set is open.

Proposition 5.22. (i) The complement of an open set is closed; (ii) the complement of a closed set is open in a fuzzy metric space.

Proof. (i) Let A be an open set in a fuzzy metric space $(X, M, *)$. We wish to show that A^c is closed. We do this by contradiction. Suppose that A^c is not closed.

Thus there exists $t > 0$ and a convergent sequence of points in A^c, $\{x_n\}$. Suppose $\{x_n\}$ converges to a point $x \notin A^c$; $x \in A$. Because A is open, there exists $\epsilon \in (0, 1)$ such that $B[x, \epsilon, t] \subseteq A$. Additionally, because $\{x_n\}$ converges to x, there exists $N \in \mathbb{N}$ such that $M(x_n, x, t) > 1 - \epsilon$, $\forall n \geq N$. Consequently, if we fix $n \geq N$ then $M(x_n, x, t) > 1 - \epsilon$. Therefore, $x_n \in B[x, \epsilon, t] \subseteq A$. Consequently, $x_n \in A$, which contradicts our hypothesis that $x_n \in A^c$. We thus conclude A^c is indeed closed. Therefore, in a fuzzy metric space, the complement of an open set is closed, paralleling crisp metric theory.

(ii) Let A be a closed subset of a fuzzy metric space $(X, M, *)$. We wish to show that A^c is open. We again prove this by contradiction. We do this by supposing A^c is not open, and constructing a sequence of points in A which converges to a point not in A.

Suppose A^c is not open and fix $t > 0$. Then, there must exist a point $x \in A^c$ such that $B[x, \epsilon, t] \nsubseteq A^c$ $\forall \epsilon > 0$. Therefore, $B[x, \epsilon, t] \cap A \neq \emptyset$ $\forall \epsilon > 0$.

Choose a countable sequence of values for ϵ, $\frac{1}{n+1}$. Therefore, we equivalently state that $B[x, \frac{1}{n+1}, t] \cap A \neq \emptyset$, for all $n \in \mathbb{N}$. Therefore, we may choose a point $x_1 \in A \cap B[x, \frac{1}{2}, t]$, and $x_2 \in A \cap B[x, \frac{1}{3}, t]$. Continuing, we choose a sequence of points $\{x_n\}$ such that $x_n \in A \cap B[x, \frac{1}{n+1}, t]$ $\forall n \in \mathbb{N}$. By the definition of the open ball B, this implies $\lim_{n \to \infty} x_n = x$. Thus, by definition of closed fuzzy metric space A, $x \in A$. This contradicts our hypothesis $x \in A^c$; so we conclude that A^c must be open. Thus the complement of any closed set in a fuzzy metric space is open with respect to the fuzzy metric.

$\square$

It becomes apparent at this point that many of the proof techniques of fuzzy metric theory are similar to those of crisp metric theory. This becomes intuitively clear when we recall that a crisp metric space models a generalized distance; a fuzzy metric space models an uncertainty measure of a generalized distance.

5.4 Fuzzy Contraction Mappings and Continuity

Lastly, we turn our attention to two related concepts of crisp metric theory; contraction mappings and continutity. As it turns out, there are a multitude of ways one may define contraction mappings on a fuzzy metric space. We consider one of the strongest forms of fuzzy contraction, GS-contraction. This is a fuzzification of the traditional, crisp contraction mapping inequality on a metric space (X, d), $d(f(x), f(y)) \leq kd(x, y)$ for all $x, y \in X$, for some $k \in [0, 1)$. Recall that we may use a fuzzy metric to induce a crisp semimetric. Using the same general construction of Definition 5.6, we rewrite the above inequality as $\frac{1}{M(f(x),f(y),t)} - 1 \leq k\big(\frac{1}{M(x,y,t)} - 1\big)$, for some $k \in (0, 1)$ and for any $t > 0$. This defines what is known as a GS-contraction mapping on the fuzzy metric space $(X, M, *)$(Gregori & Sapena, 2002)[16]. Henceforth, when we refer to a contraction mapping, or a contraction mapping, we refer to this GS-contraction definition.

Definition 5.23. Let $(X, M, *)$ be a fuzzy metric space. Let $f : X \to X$. We say that f is a contraction mapping (GS-contraction mapping) iff there exists $k \in (0, 1)$ such that,

$$\frac{1}{M(f(x), f(y), t)} - 1 \leq k\big(\frac{1}{M(x, y, t)} - 1\big) \text{for all } x, y \in X,\ \ t > 0$$

From the above definition, we extend to and define a contractive sequence of points. Let f be a contraction mapping on fuzzy metric space $(X, M, *)$. Fix any $x_0 \in X$. Define a sequence $\{x_n\}$ of points in X such that $x_n = f^{(n)}(x_0)$. That is, $x_1 = f(x_0)$, $x_2 = f(f(x_0))$, and so on. In this way, we define a sequence such that

$$\frac{1}{M(x_{n+1}, x_{n+2}, t)} - 1 \leq k\big(\frac{1}{M(x_n, x_{n+1}, t)} - 1\big)\ \forall n \in \mathbb{N}$$

We call a sequence a contractive sequence (GS-Contractive [16]) in the fuzzy metric space $(X, M, *)$ if it follows this rule for some $k \in (0, 1)$.

Remark 5.24. Suppose that $f : X \to X$ is a contraction mapping on a fuzzy metric space $(X, M_d, \cdot)$; M_d is the fuzzy metric induced by a crisp metric d on X. f is a crisp contraction mapping with respect to d.

This is shown by choosing $t = 1$. Consequently, there must exist $k \in (0,1)$ such that $\frac{1}{M(f(x),f(y),1)} - 1 \leq k\big(\frac{1}{M(x,y,1)} - 1\big)$. Therefore, $\frac{1}{\frac{1}{1+d(f(x),f(y)}} - 1 \leq k\big(\frac{1}{\frac{1}{1+d(x,y)}} - 1\big)$; hence,

$(1+d(f(x), f(y))-1 \leq k\big((1+d(x,y))-1\big)$, thus, $d(x,y) \leq kd(f(x), f(y))$. Therefore, f is a crisp contraction mapping on X.

Definition 5.25. (Gregori & Sapena, 2002) [16] Let $(X, M, *)$ be a fuzzy metric space. Let $f : X \to X$. We say that f is t-uniformly continuous iff for each ϵ such that $0 < \epsilon < 1$, there exists δ, $0 < \delta < 1$, such that, for any $x, y \in X$, $t > 0$, if $M(x,y,t) > 1 - \delta$ then $M(f(x), f(y), t) > 1 - \epsilon$.

Proposition 5.26. (Gregori & Sapena, 2002)[16]. Let $(X, M, *)$ be a fuzzy metric space. If $f : X \to X$ is a fuzzy contraction mapping on X, then f is t-uniformly continuous on X.

Proof. Let $x, y \in X$ and $\epsilon \in (0,1)$. Fix $t > 0$.
We need find $\delta \in (0,1)$ such that if $M(x,y,t) > 1-\delta$ then $M(f(x), f(y), t) > 1 - \epsilon$.

By hypothesis, f is a contraction mapping, thus we know there must exist
$k \in (0,1)$ such that $\frac{1}{M(f(x),f(y),t)} - 1 \leq k\big(\frac{1}{M(x,y,t)} - 1\big)$. Choose $\delta = \frac{\epsilon}{k(1-\epsilon)+\epsilon}$ and suppose that $M(x,y,t) > 1 - \delta$.

Therefore, $M(x,y,t) > \frac{k(1-\epsilon)}{k(1-\epsilon)+\epsilon}$. Consequently, $\frac{1}{M(x,y,t)} < \frac{k(1-\epsilon)+\epsilon}{k(1-\epsilon)} = 1 + \frac{\epsilon}{k(1-\epsilon)}$.

Thus, $\frac{1}{M(x,y,t)} - 1 < \frac{\epsilon}{k(1-\epsilon)}$; and hence $k\big(\frac{1}{M(x,y,t)} - 1\big) < \frac{\epsilon}{(1-\epsilon)}$.

Therefore, by definition of contraction mapping, $\frac{1}{M(f(x),f(y),t)} - 1 < \frac{\epsilon}{1-\epsilon} = \frac{1}{1-\epsilon} - 1$.

This implies that $M(f(x), f(y), t) > 1 - \epsilon$; f is $t-$uniformly continuous.

$\square$

Of course, with a contraction mapping on a fuzzy complete metric space, we wish to show that there exists a fixed point of the mapping, just as in crisp metric theory. That is, for a mapping f on a fuzzy metric space $(X, M, *)$, if X is a complete fuzzy metric space, then there exists $x \in X$ such that $f(x) = x$.

Proposition 5.27. Let $(X, M, *)$ be a complete metric space. If $f : X \to X$ is a contraction mapping on X, then there exists a unique fixed point $x \in X$ of f, such that $f(x) = x$.

Proof. Let $x_0 \in X$. We define a sequence of points $\{x_n\}$ such that $x_n = f^{(n)}(x_0)$. As in Definition 5.23, we see that $\{x_n\}$ is a contractive sequence of points in X. Thus, there must exist $k \in (0, 1)$ such that, for all $t > 0$ and all $n \in \mathbb{N}$,

$$\frac{1}{M(x_{n+2}, x_{n+1}, t)} - 1 \leq k\Big(\frac{1}{M(x_{n+1}, x_n, t)} - 1\Big)$$

Since this holds for all $n \in \mathbb{N}$, we see that

$$k\Big(\frac{1}{M(x_n, x_n, t)} - 1\Big) \leq k^2\Big(\frac{1}{M(x_n, x_{n-1}, t)} - 1\Big)..... \leq k^{n+1}\Big(\frac{1}{M(x_1, x_0, t)} - 1\Big)$$

It follows that $\frac{1}{M(x_{n+2}, x_{n+1}, t)} - 1 \leq k^{n+1}\big(\frac{1}{M(x_1, x_0, t)} - 1\big)$, where $\big(\frac{1}{M(x_1, x_0, t)} - 1\big)$ is fixed, and $k \in (0, 1)$. Thus, we see that,

$$\lim_{n\to\infty} \frac{1}{M(x_{n+2}, x_{n+1}, t)} - 1 \leq \lim_{n\to\infty} k^{n+1}\big(\frac{1}{M(x_1, x_0, t)} - 1\big) = 0.$$

Therefore, $\lim_{n\to\infty} M(x_{n+2}, x_{n+1}, t) = 1. \qquad (\star)$

Now, let $m, l \in \mathbb{N}$ such that $l \leq m$. By property (iv) of fuzzy metrics, and by the monotonicity of T-norm, $*$, we find that

$$M(x_m, x_l, t) \geq M(x_m, x_{m-1}, \frac{t}{2}) * M(x_{m-1}, x_{m-2}, \frac{t}{4}) * ...M(x_{l+1}, x_l, \frac{t}{2(m-l)}.$$

Continuing, by $\star$, we see that

$$\lim_{m,l\to\infty} M(x_m, x_{m-1}, \tfrac{t}{2}) * M(x_{m-1}, x_{m-2}, \tfrac{t}{4}) * ...M(x_{l+1}, x_l, \tfrac{t}{2(m-l)} = 1 * 1 * 1... * 1 = 1.$$

Therefore, $\lim_{m,l\to\infty} M(x_m, x_l, t) = 1$. Thus, the sequence $\{x_n\}$ is a Cauchy Sequence. As hypothesized, X is a complete fuzzy metric space, therefore, the sequence $\{x_n\}$ must have a limit point, which we denote x.

From this construction, it is natural to consider this x to be a possible fixed point of f. That is, we wish to show $f(x) = x$. This is equivalent to showing that $M(f(x), x, t) = 1 \; \forall t > 0$, by property (i) of fuzzy metrics. Notice though, by our proof that $\{x_n\}$ is convergent, that $\lim_{n\to\infty} M(x_n, x, t) = 1$. Recall $x_n = f(x_{n-1})$. Therefore, $1 = \lim_{n\to\infty} M(x_n, x, t) = \lim_{n\to\infty} M(f(x_{n-1}), x, t) = M(f(x), x, t)$. Therefore, $f(x) = x$.

Finally, we need show that x is unique. This is done by contradiction. Suppose there exist two unique fixed points, $x, y \in X$ of f. That is, $x \neq y$, $f(x) = x$, and $f(y) = y$; thus $f(x) \neq f(y)$.

From the definition of contraction mapping, there must exist $k \in (0, 1)$ such that

$$\frac{1}{M(x, y, t)} - 1 \leq k\big(\frac{1}{M(f(x), f(y), t)} - 1\big)$$

However, $M(f(x), f(y), t) = M(x, y, t) < 1$, and $k < 1$; thus

$$\frac{1}{M(x, y, t)} - 1 \leq k\big(\frac{1}{M(x, y, t)} - 1\big) < \frac{1}{M(x, y, t)} - 1$$

Therefore, we have a contradiction. We hence conclude that any contraction mapping on a complete fuzzy metric space has a unique fixed point.

□

5.5 Places to Go Next

Due to its twin nature with crisp metric spaces, many of the results of crisp metric theory are extended to fuzzy metric theory. Some of the most principle of these results have been included in this chapter. Research is currently being undertaken to address more in-depth theorems of fuzzy metric theory. A sample of papers of interest include:

(i) (Xia & Guo, 2004) [49] In their paper, "Fuzzy Metric Spaces", Xia and Guo construct a fuzzy metric theory different to that which is most commonly researched, that which is outlined above. They construct, rather, a crisp way of defining a metric on a fuzzy set. They do so by viewing the elements of the fuzzy sets as fuzzy points. These points are the same as those used in deriving a fuzzy subgroupoid in chapter 4 of this paper.

(ii) (Gregori et al., 2020) [15] In their paper, "Contractive Sequences in Fuzzy Metric Spaces", Gregori, Minana, and Miravet explore four types of contraction mappings, including the GS-contraction mapping that was discussed in the above section. Additionally, the authors prove a chain of strength of these types of contraction mappings, and their corresponding results.

(iii) (Paknazar & De La Sen, 2020) [38] In their paper, "Some New Approaches to Modular and Fuzzy Metric and Related Best Proximity Results", Paknazar and Sen explore new implications of fuzzy metric spaces when additional properties are assumed, similar to the addition of the non-Archimedean property discussed earlier in the chapter.

CHAPTER 6

Fuzzy Theory of Measure

We continue our exploration of fuzzy mathematics, now to a fuzzy theory of measures. In fuzzy research, measure theory incorporates fuzzy mathematics through many forms. We outline core principles of some of the most commonly used interpretations of a fuzzy theory of measure. These include a measure of the degree of crispness, otherwise known as specificity. As implied by its name, specificity measures assign a degree to how specific any fuzzy information is. In crisp logic, all information is considered to be completely specific, or specificity$=1$. However, as we have seen in chapter 1, fuzzy logic is used to model vague information, or information which is not specific. For example, consider a fuzzy set A to be defined to model the set of all values "close to a", for some $a \in \mathbb{R}$. Does A have high standards for what is considered to be 'close to a'? How high difficult is this standard to meet? Specificity measures answer these questions. Specificity measures assign a value to A which indicates the strength of the standard by which an element is considered close to a. We explore these specificity measures as an example of some of the many constructions which the fuzzy theory of measures may include.

Although these types of measure are the topic of modern research, and yield useful results, they do not naturally induce a fuzzy theory of measures twin to that of crisp measure theory. We wish to follow a twin pattern of crisp measure theory, as we have done in chapter 4 with fuzzy groups, and chapter 5 with fuzzy metrics, we first consider adopting the traditional definitions of a measure. For example, in crisp measure theory, we frequently use a prime example of measure, the Lebesque measure. This works well with crisp sets, but does not necessarily work well with fuzzy sets. To illustrate, let μ denote the Lebesque measure, and consider a proper fuzzy subset, A, of $\mathbb{R}$. How might we define $\mu(A)$? Traditionally, we would define $\mu(A)$ based on the infimum of the sum of lenghts of a collection of intervals, $\{(a_i, b_i)\}$, whose union contains A. That is, $\mu(A) = \inf \sum_{i=1}^{\infty}(b_i - a_i)$, where $A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i)$. If we use this definition, and A is a truly fuzzy set, then we must interpret the subsethood of A to the union of intervals, as being a fuzzy subsethood. However, as (a_i, b_i) are crisp intervals, this would imply that $A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i)$ iff $\bigcup_{i=1}^{\infty}(a_i, b_i)$ contains all points with non-zero membership in A. We call this crisp set of points of non-zero membership in A the support of A.

Definition 6.1. Let X be a universal set, and let $A \in [0,1]^X$. The support of A is denoted and defined to be $supp(A) = \{x \in X \mid A(x) > 0\}$.

Therefore, the Lebesque measure ignores all fuzziness of the set A, and $\mu(A) = \mu(supp(A))$. The vagueness over the edges is ignored, making any fuzziness a moot point. In fact, even in a generalized crisp measure theory, constructing a crisp measure on fuzzy sets is difficult, and nullifies many of the theorems of crisp measure theory. This is often as a result of the extremal laws of chapter 2; for a proper fuzzy set A of universal set X, $A \cap A^c \neq \emptyset$ and $A \cup A^c \neq X$.

To illustrate, let us suppose μ is a crisp measure on a set X. From our standard results in measure theory, we would expect that μ is monotone in relation to the order of subsethood. That is, if A and B are crisp measurable subsets of X, such that $A \subseteq B$, we show that $\mu(A) \leq \mu(B)$. This is shown using finite additivity of μ. Since A and B are crisp, we note $(B \cap A^c) \cap A = \emptyset$ and thus $\mu(B) = \mu(A) + \mu(B \cap A^c)$ by finite additivity. Therefore, since μ is nonnegative, we see $\mu(B) = \mu(A) + \mu(B \cap A^c) \geq \mu(A)$. Thus μ would be monotone. However, if A is a proper fuzzy set, then it is not true that $A \cap A^c = \emptyset$, and thus $(A \cap A^c) \cap B = A \cap (A^c \cap B) \neq \emptyset$. Thus, the proof of monotonicity fails.

We thus cannot apply the same definition of a crisp measure on to fuzzy sets and retain many of the implications of crisp measure theory. Hence, we explore a concept of measuring with fuzziness using weaker properties than those of crisp measure theory. By weakening the requirements of a crisp measure, we may develop a theory of "fuzzy measures", not to be confused with any measure which incorporates fuzziness. We then identify sufficient conditions for conclusions similar to crisp measure theory. These include theorems regarding convergence of measure, as well as some non-linear properties of integration. Similar to the fuzzy constructions of previous chapters, every crisp measure is a fuzzy measure. This becomes clear when a fuzzy measure is defined solely by weaker properties than those of a crisp measure; thus these weaker properties are implied by the properties of stronger, crisp properties. Therefore, it is this theory of fuzzy measure which is the generalization, and it is the crisp measure which is the special case.

6.1 Specificity Measure

Let us first consider specificity measures. These measures assign values to fuzzy sets, as they relate to crisp singletons. These can be used to determine the "trust" one might put in the accuracy of a decision made by use of the fuzzy set. This is not dissimilar to the use of the correlation coefficient in regression models.

For example, if we determine that a value is "about 5", we might construct a fuzzy set, $A_5 \in [0,1]^{\mathbb{R}}$, such that $A_5(x) = \max\{0, 1 - |x-5|\}$. This fuzzy set, as illustrated in chapter 1, leaves room for the ambiguity of x as being "about 5". However, we may construct another fuzzy set, $B_5 \in [0,1]^{\mathbb{R}}$, such that $B_5(x) = \max\{0, 1-2|x-5|\}$. By observing the graphs below, we see that, while both A_5 and B_5 illustrate "about 5", B_5 is much more restrictive on how "about 5" a value is perceived to be.

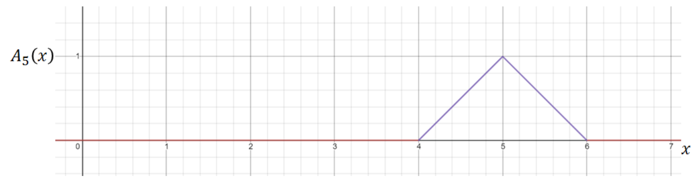

Figure 8: A_5

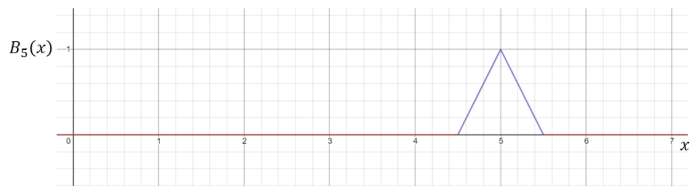

Figure 9: B_5

Therefore, using B_5 gains us more specificity, which we may use in our decision making. Specificity can be generally interpreted as a measure on fuzzy sets to determine their closeness to the nearest fuzzy singleton with $\lambda = 1$ (see ch.4; definition 4.3).

Definition 6.2. Specificity Measure: (Marin et al., 2020) [27] Let X be a finite universal set. We call the function $S_p : [0,1]^X \to [0,1]$ a specificity measure on X iff $\forall A, B \in [0,1]^X$,

(i) $S_p(A) = 1$ iff A is a crisp singleton (fuzzy singleton of $\lambda = 1$)

(ii) $S_p(\emptyset) = 0$

(iii) if $A \subseteq B$ and A, B are normalized, then $S_p(A) \geq S_p(B)$. (monotonically decreasing)

Example 6.3. Let X be a universal set. Consider $S_p : [0,1]^X \to [0,1]$ defined below for $A \in [0,1]^X$. If A is normalized, let a denote any element in X at which A attains value 1.

$$S_p(A) = \begin{cases} 0 & A \text{ is not normalized} \\ 1 - \bigvee_{x \in X - \{a\}} A(x) & A \text{ is normalized} \end{cases}$$

(i) To show this is a specificity measure, suppose first that for a fuzzy set A, A is a crisp singleton. Thus there exists a unique element of X, a such that $A(a) = 1$ and $A(x) = 0 \ \forall x \in X - \{a\}$. Therefore,

$$S_p(A) = 1 - \bigvee_{x \in X - \{a\}} A(x) = 1 - \bigvee_{x \in X - \{a\}} 0 = 1 - 0 = 1.$$

Conversely, suppose for a fuzzy set A, that $S_p(A) = 1$. Then by definition of S_p, A must be normalized, and there must exist $a \in X$ such that $A(a) = 1$. Thus, $S_p(A) = 1 = 1 - \bigvee_{x \in X - \{a\}} A(x)$. Thus $0 = \bigvee_{x \in X - \{a\}} A(x)$, and thus, since $A(x)$ is non-negative, we find that $A(x) = 0 \ \forall x \in X - \{a\}$. thus A is a crisp singleton; therefore, (i) is met.

(ii) Next, note that $\emptyset$ is not normalized, thus $S_p(\emptyset) = 0$;(ii) is met.

(iii) Finally, suppose there exist two normalized fuzzy sets $A, B \in [0,1]^X$, such that $A \subseteq B$. Since A is normalized, there must exist an element, $a \in X$ such that $A(a) = 1$. Thus by definition of fuzzy subsethood $B(a) = 1$. We may thus calculate

$$S_p(A) = 1 - \bigvee_{x \in X - \{a\}} A(x) \text{ and } S_p(B) = 1 - \bigvee_{x \in X - \{a\}} B(x)$$

We know $A(x) \leq B(x) \ \forall x \in X$ by fuzzy subsethood. Thus,

$$\bigvee_{x \in X - \{a\}} A(x) \leq \bigvee_{x \in X - \{a\}} B(x), \text{ and hence}$$

$$S_p(A) = 1 - \bigvee_{x \in X - \{a\}} A(x) \geq 1 - \bigvee_{x \in X - \{a\}} B(x) = S_p(B)$$

Thus (iii) holds. Therefore, S_p defines a specificity measure on X.

Specificity measures are, as stated previously, a way of distinguishing how close fuzzy sets are to a value that is certain. However, specificity measures need not be arbitrarily assigned, separate from any other fuzzy information. Rather, they can be generated by structures similar to ones already discussed in chapter 2. Let us consider a similarity measure, which is more closely related to a relation on $[0,1]^X$ than a true measure. Similarity measures are a means to determine the "similarity " of one fuzzy set to another. It is implied by their name that similarity measures include properties inspired by similarity relations (see ch. 2). Since the research is ongoing, the specific properties of the definition of a similarity measure are not uniformly recognized among fuzzy research; similarity measures are often defined having many different combinations of properties (see (Couso, 2013)[10]) . For the sake of this chapter, let us consider the following definition of a specificity measure, which is derived by the sufficient condition outlined in (Marin et al. , 2020) [27].

Definition 6.4. Let X be a finite universal set; consider $[0,1]^X$, the set of all fuzzy sets on X. A function $\beta : ([0,1]^X)^2 \to [0,\infty)$ is a similarity measure on X iff $\forall A, B, C \in [0,1]^X$

(i) $\beta(A,B) = 1$ if and only if $A = B$ (reflexive)

(ii) $\beta(A,B) = \beta(B,A)$ (symmetric)

(iii) if $A \subseteq B \subseteq C$ and $\max\limits_{x \in X}\{A(x)\} = \max\limits_{x \in X}\{B(x)\}$, then $\beta(A,C) \leq \beta(A,B)$

Example 6.5. Let X be a finite universal set. Consider β defined for $A, B \in [0,1]^X$ below. β is a similarity measure on X.

$$\beta(A,B) = \frac{1}{|X|}\sum_{x \in X}\{1 - |A(x) - B(x)|\}$$

(i) First suppose that for two fuzzy sets, $A, B \in [0,1]^X$, $\beta(A,B) = 1$. Therefore,

$$|X| = \sum_{x \in X} \{1 - |A(x) - B(x)|\}$$

Thus, $|X| = |X| - \sum_{x \in X} \{|A(x) - B(x)|\}$, and therefore, $0 = \sum_{x \in X} \{|A(x) - B(x)|\}$.

Therefore, $|A(x) - B(x)| = 0 \ \forall x$. Thus, $A(x) = B(x) \ \forall x$; we have fuzzy set equality $A = B$. The proof of the converse is similar. We thus conclude that β is reflexive. (ii) The symmetric property follows directly from the symmetric property of the absolute value metric. (iii) Suppose $A, B, C \in [0,1]^X$ such that $A \subseteq B \subseteq C$ and $\max_{x \in X}\{A(x)\} = \max_{x \in X}\{B(x)\}$. Therefore,

$$\beta(A,B) = \frac{1}{|X|} \sum_{x \in X} \{1 - |A(x) - B(x)|\} \text{ and } \beta(A,C) = \frac{1}{|X|} \sum_{x \in X} \{1 - |A(x) - C(x)|\}$$

Note however, that for any $x \in X$, $A(x) \leq B(x) \leq C(x)$ by definition of fuzzy subsets. Thus, $|A(x) - B(x)| = B(x) - A(x)$, and $|A(x) - C(x)| = C(x) - A(x)$. Therefore, $B(x) - A(x) \leq C(x) - A(x)$, and hence

$$1 - (B(x) - A(x)) \geq 1 - (C(x) - A(x))$$

Therefore, as this inequality holds for all $x \in X$,

$$\sum_{x \in X} \{1 - |A(x) - B(x)|\} \geq \sum_{x \in X} \{1 - |A(x) - C(x)|\}.$$

Thus, by multiplication by $\frac{1}{|X|}$, we find that

$$\beta(A,B) = \frac{1}{|X|} \sum_{x \in X} \{1 - |A(x) - B(x)|\} \geq \frac{1}{|X|} \sum_{x \in X} \{1 - |A(x) - C(x)|\} = \beta(A,C).$$

Therefore, condition (iii) is satisfied.

As stated previously, a specificity measure can be generated from a similarity measure. We provide the following proof, which is derived from the work in (Marin et al. , 2020) [27].

Proposition 6.6. Let X be a finite, nonempty universal set. Let β be a similarity measure on X. Let us define a function S_β as below. S_β is a specificity measure on X.

Let $A \in [0,1]^X$, such that $\max_{x \in X}(A(x)) = A(a)$ for some fixed point $a \in X$.

$$S_\beta(A) = \begin{cases} 0 & A = \emptyset \\ \beta(A, \{a\}) & A \neq \emptyset \end{cases}$$

Proof. (i) Let $A \in [0,1]^X$. We need show that $\beta(A) = 1$ if and only if A is a crisp singleton.

$\Longrightarrow$) Suppose A is a crisp singleton. That is, $A = \{a\}$ for some $a \in X$. $A = \{a\}$. Thus, a is the point of maximal membership in A, as $A(a) = 1$. Therefore by property (i) of similarity measures, $\beta(A, \{a\}) = \beta(A, A) = 1$, and thus, $S_\beta(A) = 1$.

$\Longleftarrow$) Suppose $S_\beta(A) = 1$. Then by definition of S_β, $A \neq \emptyset$, thus $S_\beta(A) = \beta(A, \{a\}) = 1$, where a is an element in X such that A attains maximal degree. Therefore, again by property (i) of similarity measures, $A = \{a\}$, a crisp singleton.

(ii) $S_\beta(\emptyset) = 0$ by definition. Thus, (ii) of specificity measure is met.

(iii) Let A and B be normalized fuzzy subsets of X such that $A \subseteq B$. Therefore, there must exist $a \in X$ such that $A(a) = 1 = B(a)$ for fixed element a. Consider the crisp singleton $\{a\}$. Note then that, $\{a\} \subseteq A \subseteq B$, and thus

$$\max_{x \in X}\{a\}(x) = \max_{x \in X} A(x) = \max_{x \in X} B(x) = 1$$

Hence, by property (iii) of similarity measures, $\beta(\{a\}, B) \leq \beta(\{a\}, A)$. Additionally, by the symmetric property of β, this implies $\beta(B, \{a\}) \leq \beta(A, \{a\})$. Therefore, $S_\beta(B) = \beta(B, \{a\}) \leq \beta(A, \{a\}) = S_\beta(A)$, and thus the third property of specificity measures is satisfied. Therefore, S_β is a specificity measure on X. $\square$

We have thus far discussed similarity measures, used to determine the reliability of fuzzy information as contrasted with a crisp singleton. There are a number of properties which may be proven of these measures; however, the theory of these measures does not induce a theory twin to that of crisp measure theory. In chapters 4 and 5, we explored definitions of fuzzy structures such that a theory twin to that of crisp logic was induced. To do this for the fuzzy theory of measure, we need define another one of the measures of fuzziness being researched; a fuzzy measure. The constructions related to these fuzzy measures closely follow the core theorems of crisp measure theory. This theory does include the ability of measuring fuzzy sets in parts of the construction; however, we attribute the marker, 'fuzzy', to the non-additive nature of these measures. A fuzzy measures can be interpreted as modelling the statement, "We are something more (or less) than the sum of our parts".

6.2 Fuzzy Measure

We now turn our attention to a measure known as a fuzzy measure; not to be confused with a measure of fuzzy sets. These fuzzy measures induce theorems of convergence with similar conclusions to those of crisp measure theory, but with different necessary and sufficient conditions. We also induce new integrals, namely the Sugeno integral, which is commonly used as an aggregation operator for large sets of fuzzy data. This form of integration helps us to make informed decisions based on subjective, fuzzy data; this is particularly applied to sets of polling data, which are frequently more representative when fuzzy than when demanding strict, "yes/no" answers from the responders. Therefore, we continue throughout this chapter to develop a generalized fuzzy measure theory. This enables us to explore a famous example of a fuzzy measure, the Sugeno measure. Finally, we explore the application of the corresponding Sugeno integral. To begin this process, we first need recall some definitions of crisp measure theory; namely a $\sigma-$algebra, and a crisp measure.

Definition 6.7. $\sigma-$algebra: Let X be a universal set, and Φ a nonempty collection of subsets of X. Φ is a $\sigma-$algebra on X iff (i) $X \in \Phi$

(ii) $A \in \Phi$ implies $A^c \in \Phi$

(ii) $\{A_i\}_{i=1}^{\infty} \subset \Phi$ implies $\bigcup_{i=1}^{\infty} A_i \in \Phi$

Definition 6.8. Crisp Measure on a $\sigma-$algebra: Let X be a universal set, and let Φ be a $\sigma-$algebra. We call a function $\mu : \Phi \to [0, \infty]$ a measure on X iff

(i) $\mu(\emptyset) = 0$

(ii) if $\{A_i\}_{i=1}^{\infty}$ is a pairwise disjoint sequence of sets in Φ, then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

We have illustrated that the Lebesque measure is not particularly well-equipped to measure fuzzy sets. It is also true that a general crisp measure does not work well with fuzzy sets, due to the extremal laws of fuzzy set theory. Therefore, in the theory of fuzzy measures, we loosen the condition of countable additivity in our definition, and instead, require the weaker condition of monotonicity. In some research, it may also be required in the definition that the fuzzy measure be upper and lower continuous. However, for the purposes of this chapter, we only consider the first two, universally required, properties of a fuzzy measure.

Definition 6.9. (Wang & Klir, 1992) [48] Fuzzy Measure: Let X be a universal set. Let Φ a set of subsets of X contain the empty set. Let $m : \Phi \to [0, \infty]$ be a non-negative set function. m is a fuzzy measure iff,

(i) $m(\emptyset) = 0$

(ii) for $A, B \in \Phi$, if $A \subseteq B$, then $m(A) \leq m(B)$ (m is monotonic)

When studying fuzzy measures, imagine the measure as acting on sets for which at least part of the set is malleable. In the process of unionization, components of sets either lose or gain meaning. These are the parts which we consider, malleable. This is similar to the process of mitosis of cells. A singular cell has only one nucleus. However, throughout the life of the cell, it splits into two daughter cells, each with their own nucleus. Consider the cells to be sets, and the components of the cells to be elements of the set. Every element of the daughter cells came from the original, parent cell. Thus consider the parent cell to be the union of the two daughter cells. Assign a value to each cell by the number of nuclei it has. However, note that the union of the daughter cells has a lesser value than the sum of the value of the daughter cells; the value we assign to cells by nuclei is nonadditive.

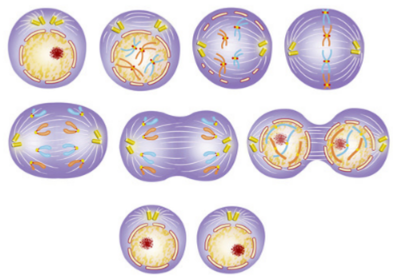

Figure 10: Mitosis: Fuzzy Union in Reverse [28]

For an alternative interpretation, consider two people, x and y, working for a company. Both x and y are excellent workers on their own, but lack the ability to work well with others. Suppose x can complete 5 tasks of equal difficulty in one hour; suppose y can complete 7 similar tasks in an hour. Thus, let A_x be the set of tasks which x can complete in an hour, and A_y the set of elements y can complete in an hour. Through this interpretation, we infer that no people can complete zero tasks in an hour. Thus $|A_0| = |\emptyset| = 0$. Also, even if some employees don not work well with others, we would still expect that two people can perform at least as many tasks as either of the original people on their own; that is $|A_x| \leq |A_x y|$, thus the proposed fuzzy measure would be considered monotonic. Of course, it is possible that y works so poorly with others that y in fact detracts from the efficiency of x, such that $|A_{xy}| \leq |A_x|$. We would hope and suppose that such toxic employees are not maintained in the company. Therefore, we interpret the proposed measure to indeed be monotonic. Even though these are crisp sets, the boundaries of these sets is malleable in that, through the fusion of these sets, some of the boundary is consumed and lost by the fusion process. Therefore, $A_{xy} = A_x \cup A_y$ may only consist of 10 tasks for this scenario. We lost two tasks in the fusion process of A_x and A_y, likely due to disagreements between x and y regarding how a task should be done. Alternatively, consider two other employees y and z, and let A_y and A_z be the sets of tasks y and z can complete in an hour respectively. If y and z do have the skills of working well together, then the fusion of sets A_y and A_z may result in a larger set; if $|A_y| = 3$ and $|A_z| = 4$, then through the efficient ability of the employees to optimize task delegation, $|A_{yz}| = 8$.

Example 6.10. Let Φ be a $\sigma-$algebra of a set X. Let f be a nonnegative, extended real valued function on X. Let μ be function $\mu : \Phi \to [0, \infty]$ defined for any $A \in \Phi$, below. μ is a fuzzy measure on X, but not a crisp measure.

$$\mu(A) = \bigvee_{x \in X} (A(x) \cdot f(x))$$

First, if $A = \emptyset$, then $A(x) = 0 \; \forall x \in X$. Therefore, $\mu(\emptyset) = 0$.

Next, let $A, B \in \Phi$, such that $A \subseteq B$. Therefore, $A(x) \leq B(x) \; \forall x \in X$ and hence, $A(x) \cdot f(x) \leq B(x) \cdot f(x) \; \forall x \in X$. Thus, $\mu(A) \leq \mu(B)$, and hence μ is monotone.

Therefore, μ is a fuzzy measure.

However, μ is not a crisp measure. Consider the simple counterexample of a constant function, $f : X \to [0, \infty)$, $f(x) = 3 \; \forall x \in X$. Let $a, b \in X$, $a \neq b$. Define fuzzy sets A, B such that $A(a) = 0.5$ for fixed $a \in X$, and $A(x) = 0 \; \forall x \neq a$. Similarly, $B(b) = 0.7$ for fixed $b \in X$, and $B(x) = 0 \; \forall x \neq b$. Thus $A \cap B = \emptyset$, $\mu(A) = 1.5$, $\mu(B) = 2.1$, and $\mu(A \cup B) = 2.1$. Therefore, $\mu(A \cup B) \neq \mu(A) + \mu(B)$, and therefore, as μ is not additive, μ is not a crisp measure.

Example 6.11. Consider the finite universal set $X = \{x_1, x_2 ... x_n\}$. Define $\mu : [0,1]^X \to [0, \infty)$ for $A \in [0,1]^X$,

$$\mu(A) = \sqrt[n]{\sum_{i=1}^{n} A(x_i)}$$

Here we see that for the empty set, $\emptyset(x_i) = 0 \ \forall x_i$ thus $\mu(\emptyset) = 0$.

Additionally, suppose that $A, B \in [0,1]^X$ such that $A \subseteq B$. Thus, $A(x_i) \leq B(x_i) \ \forall i$, and thus $\sqrt[n]{\sum_{i=1}^{n} A(x_i)} \leq \sqrt[n]{\sum_{i=1}^{n} B(x_i)}$. Therefore, μ is monotone.

However, μ is not countably additive. To show this, consider $n = 3$. Let A, B be fuzzy sets such that $A(x_1) = 0.5$, and $A(x_i) = 0 \ \forall i \geq 2$; and $B(x_3) = 0.7$ and $B(x_i) = 0 \ \forall i < 3$. Thus, $A \cap B = \emptyset$, but $\mu(A \cup B) = \sqrt[3]{1.2}$. However, $\mu(A) = \sqrt[3]{0.5}$ and $\mu(B) = \sqrt[3]{0.7}$. Thus we find that μ is not additive, thus μ is not a crisp measure.

6.3 λ Fuzzy Measures and Sugeno

One of the most commonly used fuzzy measures is the Sugeno λ measure. This measure yields an integral with some similar properties as that of the Lebesque integral. We have seen that a fuzzy measure of a set may not be equal to the sum of the measures of its parts. The Sugeno λ measure defines a rule to determine how much more than (or less than) the sum of our parts. This is known as the $\lambda-$rule. The $\lambda-$rule can be inductively defined for the case of a finite union of sets. However, as in crisp measure theory, we define a $\lambda-$fuzzy measure as a fuzzy measure meeting the $\lambda-$rule in the case of countably infinite unions; this is known as the $\sigma - \lambda-$rule. When a $\lambda-$fuzzy measure reaches a maximum of 1, we call this the Sugeno measure.

Definition 6.12. (Wang & Klir, 1992) [48] $\lambda-$rule: Let X be a universal set, Φ a collection of subsets of X, μ an extended, real valued function on Φ, $\mu : \Phi \to [0, \infty]$. μ is said to satisfy the $\lambda-$ rule iff there exists $\lambda \in \left(\frac{-1}{\vee\mu}, \infty\right) \cup \{0\}$ such that for any disjoint sets $A, B \in \Phi$,

$$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B).$$

If $\lambda = 0$, then the $\lambda-$ rule becomes crisp additivity;

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

If $\lambda \neq 0$, then by multiplying both sides by λ, adding 1, and factoring the expression on the right, we find that the $\lambda-$rule as is equivalent to the following; this allows us to more easily define the finite and $\sigma - \lambda-$rules.

$$1 + \lambda\mu(A \cup B) = [1 + \lambda\mu(A)] \cdot [1 + \lambda\mu(B)]$$

From this form, we may inductively infer a finite $\lambda-$rule. For a finite sequence of pairwise disjoint sets in Φ, whenever the union of such sets is in Φ, we define the finite $\lambda-$rule as follows.

Definition 6.13. (Wang & Klir, 1992) [48] Let Φ be a nonempty collection of subsets of a universal set X. Let μ be a nonnegative, real-valued function on Φ. μ meets the finite $\lambda-$ rule iff there exists $\lambda \in \left(\frac{-1}{\vee\mu}, \infty\right) \cup \{0\}$ such that for any finite sequence $\{A_i\}_{i=1}^n$ of sets in Φ, such that $\bigcup\limits_{i=1}^{n} A_i \in \Phi$, $A_i \cap A_j = \emptyset \ \forall i \neq j$,

$$\lambda = 0 \implies \mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$$

$$\lambda \neq 0 \implies 1 + \lambda\mu(\bigcup_{i=1}^{n} A_i) = \Pi_{i=1}^{n}[1 + \lambda\mu(A_i)]$$

However, in many constructions, we need this form of λ additivity to hold for the union of an infinite sequence of disjoint sets in Φ; particularly so when Φ is a $\sigma-$algebra, and closed under infinite union. If λ additivity holds for an infinite sequence, we say that μ meets the $\sigma - \lambda-$rule.

In general, finite does not imply infinite. That is, the finite $\lambda-$rule does not necessarily imply the $\sigma - \lambda-$rule. However, counterintuitively, the $\sigma - \lambda-$rule also does not necessarily imply the finite $\lambda-$rule. This is due to the fact that it is not required that $\emptyset \in \Phi$ or that $\mu(\emptyset) = 0$ for an arbitrary function μ. If it is the case that $\emptyset \in \Phi$ and $\mu(\emptyset) = 0$, then the $\sigma - \lambda-$rule implies the finite. We thus explore a sufficient condition from (Wang & Klir, 1992) [48] to establish this implication.

Lemma 6.14. Let X be a universal set, Φ a class of subsets of X, and $\mu : \Phi \to [0, \infty]$ such that μ meets the $\sigma - \lambda-$rule. If there exists $A \in \Phi$ such that $\mu(A) < \infty$, and $\emptyset \in \Phi$, then $\mu(\emptyset) = 0$

Proof. Let $A \subseteq X$ have finite measure. Then, let us construct the sequence of disjoint sets in Φ, $\{A_i\}_{i=1}^{\infty}$ such that $A_1 = A$ and $A_i = \emptyset \ \forall i \geq 2$.

Thus, by the $\sigma - \lambda-$rule, one of the following case holds.

Case 1) $\lambda = 0$. Thus $\mu(A) = \mu(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} \mu(A_i) = \mu(A) \sum\limits_{i=2}^{\infty} \mu(\emptyset)$.

Therefore, $0 = \sum\limits_{i=2}^{\infty} \mu(\emptyset)$. Thus, as μ is nonnegative, $\mu(\emptyset) = 0$.

Case 2) $\lambda \neq 0$. Thus,

$$1 + \lambda\mu(A) = \Pi_{i=1}^{\infty}[1 + \lambda\mu(A_i)] = [1 + \lambda\mu(A)]\Pi_{i=2}^{\infty}[1 + \lambda\mu(\emptyset)]$$

Therefore, as λ and $\mu(A)$ are finite, we find $1 = \Pi_{i=2}^{\infty}[1 + \lambda\mu(\emptyset)]$. Therefore, $1 = 1 + \lambda\mu(\emptyset)$ and hence, either $\lambda = 0$ or $\mu(\emptyset) = 0$.

We have supposed $\lambda \neq 0$, thus $\mu(\emptyset) = 0$. Therefore, in any case, $\mu(\emptyset) = 0$

□

Theorem 6.15. (Wang & Klir, 1992) [48] Let X be a universal set, and Φ a class of subsets containing the empty set. Let μ be a nonnegative function on Φ. μ meets the finite $\lambda-$rule if it meets the $\sigma - \lambda-$rule and there exists a set of finite measure in Φ.

Proof. Let $\{A_i\}_{i=1}^{n}$ be a sequence of pairwise disjoint sets in Φ, such that the union of these sets is in Φ. By the lemma above, $\mu(\emptyset) = 0$. Therefore, let us consider the infinite sequence $\{B_i\}_{i=1}^{\infty}$ such that $B_i = A_i \ \forall i \leq n$ and $B_i = \emptyset \ \forall i > n$. Thus we see that, $\bigcup\limits_{i=1}^{n} A_i = \bigcup\limits_{i=1}^{\infty} B_i$.

Case 1) $\lambda = 0$. By definition of the $\sigma - \lambda-$rule,

$$\mu(\bigcup_{i=1}^{n} A_i) = \mu(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{n} \mu(A_i) + \sum_{i=n+1}^{\infty} \mu(\emptyset)$$

Recall $\mu(\emptyset) = 0$; thus $\mu(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} \mu(A_i) + 0 = \sum\limits_{i=1}^{n} \mu(A_i)$. Therefore, the finite $\lambda-$rule holds.

Case 2) $\lambda \neq 0$. We need show that $1 + \lambda\mu(\bigcup\limits_{i=1}^{n} A_i) = \Pi_{i=1}^{n}[1 + \lambda\mu(A_i)]$. First note by the $\sigma - \lambda-$rule,

$$1 + \lambda(\bigcup_{i=1}^{\infty} B_i) = \Pi_{i=1}^{\infty}[1 + \lambda\mu(B_i)].$$

By set equality of the union of A_i and B_i respectively, we find that $1 + \lambda(\bigcup\limits_{i=1}^{n} A_i) = 1 + \lambda(\bigcup\limits_{i=1}^{\infty} B_i)$. Therefore,

$$1 + \lambda(\bigcup_{i=1}^{n} A_i) = \Pi_{i=1}^{\infty}[1 + \lambda\mu(B_i)].$$

Additionally, by definition of B_i,

$$\Pi_{i=1}^{\infty}[1 + \lambda\mu(B_i)] = (\Pi_{i=1}^{n}[1 + \lambda\mu(A_i)])(\Pi_{i=n+1}^{\infty}[1 + \lambda\mu(\emptyset)]).$$

Recall that $\mu(\emptyset) = 0$, thus $1 + \lambda\mu(\emptyset) = 1$. Therefore,

$$1 + \lambda(\bigcup_{i=1}^{n} A_i) = \Pi_{i=1}^{n}[1 + \lambda\mu(A_i)] \cdot 1 = \Pi_{i=1}^{n}[1 + \lambda\mu(A_i)]$$

Therefore, the finite $\lambda-$rule holds.

$\square$

Remark 6.16. Some of the conditions of the theorem above are met trivially if Φ is a $\sigma-$algebra; this is due to the fact any $\sigma-$algebra contains the empty set, and is closed under countable union. Thus, suppose the case that μ is a nonnegative, real-valued function on a $\sigma-$algebra, Φ; suppose μ meets the $\sigma-\lambda-$rule. μ meets the finite $\lambda-$rule if there exists any set $B \in \Phi$ such that $\mu(B) < \infty$. Hence, similar to crisp measure theory, it is common among fuzzy research to study constructions when Φ is a $\sigma-$algebra.

Definition 6.17. λ Fuzzy Measure: (Wang & Klir, 1992) [48] Let X be a universal set. Let Φ be a collection of subsets of X. Let $\mu : \Phi \to [0, \infty]$. μ is a $\lambda-$fuzzy measure iff μ is a fuzzy measure meeting the $\sigma-\lambda-$rule. If we additionally require that $\mu(X) = 1$, then we call this measure the Sugeno λ measure.

Example 6.18. Suppose $X = \{x_1, x_2, x_3\}$, and a $\sigma-$ algebra $\{0,1\}^X$. Let μ be a non-negative set function meeting the $\sigma-\lambda-$rule, and suppose that $\mu(\{x_1\}) = 0.4$, $\mu(\{x_2\}) = 0.3$, and $\mu(\{x_3\}) = 0.8$. We may calculate a λ such that μ is the Sugeno measure ($\mu(X) = 1$). Note $X = \{x_1\} \cup \{x_2\} \cup \{x_3\}$. Thus, using the $\lambda-$rule of μ, we find that if $A = \{x_1, x_2\}$, then

$$\mu(A) = \mu(\{x_1\}) + \mu(\{x_2\}) + \lambda\mu(\{x_1\})\mu(\{x_2\}) \text{ and}$$

$$\mu(X) = \mu(A) + \mu(\{x_3\}) + \lambda\mu(A)\mu(\{x_3\})$$

Thus, $\mu(X) = [0.4 + 0.3 + \lambda(0.3)(0.4)] + (0.8) + \lambda[0.4 + 0.3 + \lambda(0.3)(0.4)](0.8) = 1.5 + 0.68\lambda + 0.096\lambda^2$.

To define the Sugeno measure, we require that in addition to the $\lambda-$rule, that $\mu(X) = 1$. Thus, we solve $1 = 1.5 + 0.68\lambda + 0.096\lambda^2$, and find that $\lambda \in \{-\frac{5}{6}, -\frac{25}{4}\}$.

Recall that for the $\lambda-$rule to hold, we only consider $\lambda \in (\bigwedge_{A\in\Phi} -\frac{1}{\mu(A)}, \infty)$. However, since $X \in \Phi$, and $\mu(X) = 1$; thus for the Sugeno measure, we only consider $\lambda \in (-1, \infty)$. Therefore, we choose $\lambda = -\frac{5}{6}$.

We have thus defined the Sugeno λ measure on X, based on the given weights of the singleton elements. In fact, for any $X = \{x_1, x_2, ...x_n\}$, we may determine the value of λ for the Sugeno measure given only the weights of the singleton sets of elements in X. It has also been shown that for any such construction, with any weights given to the singleton elements of a finite set, solving for λ using the lambda rule yields up to n values for λ, but only one value which is in the interval, $(-1, \infty)$, that is applicable. (Tahani, Keller, 1990) [47].

In summary, the Sugeno λ measure assigns measures to subsets of a universal set, equivalent to a ranked value of the set to the system. Unlike crisp measures, this measure is not additive, but instead adheres to the $\sigma - \lambda$ rule; this rule implies that the Sugeno λ measure adheres to the finite λ rule, by the fact that any fuzzy measure assigns value 0 to the empty set. As stated previously, the Sugeno λ measure, and fuzzy measures generally, induce a fuzzy measure theory similar to that of the results of crisp measure theory. One such topic of fuzzy measures explores the necessary and sufficient conditions of function convergence implications. Of course, these conditions are sometimes different than those of crisp measure theory.

6.4 Convergence in Fuzzy Measure

Without countable additivity of a measure on a $\sigma-$algebra, implications of convergence from crisp measure theory frequently do not hold under the same necessary conditions. It is thus often a subject of research to determine minimal necessary conditions on convergence implications of fuzzy measures. We now examine which theorems of convergence with respect to fuzzy measure hold for measurable functions. First, we need recall some definitions of crisp measure theory, which we retain in fuzzy measure theory.

Definition 6.19. Measurable Function: Let (X, Φ) and (Y, Ψ) be measurable space. That is Φ and Ψ are $\sigma-$algebras on X and Y respectively. A function $f : X \to Y$ is said to be measurable iff $\forall B \in \Psi$, $f^{-1}(B) \in \Phi$.

For the purpose of this section, we focus our attention on the case that $Y = \mathbb{R}$

and Ψ is the power set of $\mathbb{R}$. We denote $F(X) = \{f : X \to \mathbb{R} \mid f \text{ is measurable}\}$. Additionally, we use the same definition for convergence in measure and convergence pointwise almost everywhere in the same manner as we would in crisp measure theory.

Definition 6.20. Convergence: Let (X, Φ, μ) be a fuzzy measure space, in that (X, Φ) is a measurable space, and μ is a fuzzy measure on Φ. Let $\{f_n\}_{n=1}^{\infty} \subset F(X)$, and $f \in F(X)$. We say that f_n converges to f in measure, denoted $f_n \xrightarrow{\mu} f$, iff $\lim\limits_{n\to\infty} \mu(\{x \in X \mid |f_n(x) - f(x)| > \epsilon\} = 0$ for all $\epsilon > 0$. Additionally, we say that f_n converges to f pointwise, denoted $f_n \xrightarrow{p.w.} f$, iff for all $x \in X$ and $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geq N$, then $|f_n(x) - f(x)| < \epsilon$. If there exists a crisp set of measure zero, $A \subseteq X$ such that f_n converges pointwise to f on A^c, then we say that f_n converges pointwise almost everywhere to f, denoted $f_n \xrightarrow{a.e.} f$.

For this section, we show a sufficient condition for convergence almost everywhere to imply convergence in measure. To determine this sufficient condition requires a definition and lemma from (Kawabe, 2019) [20] and (Li, 2003)[25].

Definition 6.21. Let X be a universal set, and $\{A_n\}$ be a sequence of subsets of X. Let $A \subseteq X$. We say that A_n decreases to A, denoted $A_n \downarrow A$, iff (i) $\forall n \in \mathbb{N}$, $A_n \supseteq A_{n+1}$ and (ii) $\forall x \in A^c$ there exists $N \in \mathbb{N}$, such that $x \notin A_n$ $\forall n \geq N$.

Definition 6.22. (Kawabe, 2019) [20] Let (X, Φ, μ), be a fuzzy measure space. μ is said to be strongly order continuous iff for any sequence $\{A_n\} \subset \Phi$, such that $\{A_n\}$ is decreasing to A ($A_n \downarrow A \in \Phi$) and $\mu(A) = 0$, then $\lim\limits_{n\to\infty} \mu(A_n) = 0$.

Theorem 6.23. (Li, 2003)[25] Let (X, Φ, μ) be a fuzzy measure space, and let μ be strongly order continuous. If a sequence of finite, real-valued functions converges in measure, it converges pointwise almost everywhere.

Proof. Let $\{f_n\} \subseteq F(X)$, and $f \in F(X)$ such that $f_n \xrightarrow{a.e.} f$. Let D be the set of values in X such that $f_n(x) \nrightarrow f(x)$. By supposition of convergence almost everywhere, $\mu(D) = 0$.

Note then that $D = \bigcup_{m=1}^{\infty} \big(\bigcap_{n=1}^{\infty} \big(\bigcup_{k=n}^{\infty} \{x \mid |f_k(x) - f(x)| \geq \frac{1}{m}\}\big)\big)$.

Let $A^m = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{x \mid f_k(x) - f(x)| \geq \frac{1}{m}\}$ for fixed $m \in \mathbb{N}$. Since A^m is a subset of a set of zero measure, D, we find $\mu(A^m) = 0 \ \forall m \in \mathbb{N}$.

Additionally, let $A_n^m = \bigcup_{k=n}^{\infty} \{x \mid |f_k(x) - f(x)| \geq \frac{1}{m}\}$.

Thus we find that for fixed m, $A_n^m \supseteq A_{n+1}^m \forall n$. Additionally, for any $x \in (A^m)^c$, there must exist some finite number of A_n^m such that $x \in A_n^m$. Therefore, choose N to be greater than the last such n such that $x \in A_n^m$. Therefore, $x \notin A_n^m \ \forall n \geq N$. Therefore, we find that $A_n^m \downarrow A^m$.

Since $A^m \subseteq D$, we find $\mu(A^m) = 0$. Therefore, by the strongly order continuous property of μ, $\lim_{n\to\infty} \mu(A_n^m) = 0$. Now let us note that for any fixed n and m,

$$\{x \mid |f_n(x) - f(x)| \geq \frac{1}{m}\} \subseteq \bigcup_{k=n}^{\infty} \{x \mid |f_k(x) - f(x)| \geq \frac{1}{m}\} = A_n^m.$$

Thus, by monotonicity of μ,

$$\mu\big(\{x \mid |f_n(x) - f(x)| \geq \frac{1}{m}\}\big) \leq \lim_{n\to\infty} \mu\big(A_n^m\big) = 0.$$

Therefore, the limit of the measure of all such sets has measure 0;

$$\lim_{n\to\infty} \mu\big(\{x \mid |f_n(x) - f(x)| \geq \frac{1}{m}\}\big) = 0.$$

Now let, $\epsilon > 0$. Fix m such that $\frac{1}{m} \leq \epsilon$. Therefore, $\{x \mid |f_n(x) - f(x)| \geq \epsilon\} \subseteq \{x \mid |f_n(x) - f(x)| \geq \frac{1}{m}\}$. Thus by monotonicity of μ, we find that

$$\lim_{n\to\infty} \mu(\{x \mid |f_n(x) - f(x)| \geq \epsilon\}) \leq \lim_{n\to\infty} \mu\big(\{x \mid |f_n(x) - f(x)| \geq \frac{1}{m}\}\big) = 0$$

Therefore, $\lim_{n\to\infty} \mu(\{x \mid |f_n(x) - f(x)| \geq \epsilon\}) = 0$; as this is for arbitrary $\epsilon > 0$, $f_n \xrightarrow{\mu} f$.

□

We thus see that some of the implications of crisp convergence are extended to fuzzy measure theory under certain conditions. Further implications of fuzzy convergence are the subject of current research in fuzzy measure theory [20]. Having illustrated a the existance of these implcations, we now turn our attention to the topic of fuzzy integration.

6.5 Sugeno Integration

Much of the research of fuzzy integration focuses on aggregating sets of data, such as the public opinion of a candidate, accounting for the vague factor of party loyalty. It is this fuzziness which inspires an operation analogous to integrating over a fuzzy set. We thus explore a fuzzy integral regularly the topic of fuzzy research, the Sugeno integral, as well as its application to these forms of aggregation problems.

Definition 6.24. Sugeno Integral: (Wang & Klir, 1992) [48] Let (X, Φ, μ) be a continuous measure space. Let $A \in \Phi$ and let $f : X \to [0, \infty)$. Denote for $\alpha \geq 0$, $f_\alpha = \{x \in X \mid f(x) \geq \alpha\}$. We define the Sugeno Integral of f to be

$$(S)\int_A f = \bigvee_{\alpha \geq 0} \alpha \wedge \mu(f_\alpha \cap A)$$

When using the Sugeno Integral, it is most common to use the Sugeno Measure, and require that f be a function into $[0, 1]$; equivalently, f defines a fuzzy subset of X. One of the most common applications of the Sugeno integral is as an aggregation operator of fuzzy-weighted data. For example, suppose we survey some n people, and ask, "On a scale of 0 to 1, how strongly would you approve of your local congressman?". This question is clearly vague and subjective, and thus responses will likely result in a list of n values, many neither 0 nor 1. Now, consider that we wish to not count all responses equally. The purpose of many polls is to predict results of the next election. To do so, we may not wish to count all responses equally, as we would if we merely were to average all responses. Many polls adjust in a binary fashion to this problem by only counting results of people that have voted in the previous two major elections. An alternative, fuzzy method, would be to ask an additional question in the survey. For example, let us ask the responders, "How open are you to consider voting for the opposing party?". Partisan loyalty plays a major roll in many elections, and many participants will vote for a party, regardless of how low they may approve of a particular candidate. This is also often due to the "lesser of two evils" reasoning. We may denote the openness of a person to vote for an opposing candidate as a fuzzy subset of our set of responses, A. We wish to give more weight to these responses from people with high membership in A, and a smaller weight to those with low membership in A a smaller weight. In this aggregation application of the Sugeno integral, the universal set is finite, and we thus yield an equivalent definition, common in fuzzy research.

Definition 6.25. Sugeno Integral: (Grabiec, 1988) [14] Let X be a finite set of data sources. Let μ denote a the Sugeno measure. Let h be a fuzzy subset of X. Let X be indexed by decreasing order of h, such that $h(x_1) \geq h(x_2) \geq ...h(x_n)$. Additionally, define sets $W_i = \{x_j\}_{j=1}^{i}$. We then calculate the Sugeno Integral of h to be

$$(S)\int_X h = \bigvee_{i=1}^{n} (h(x_i) \wedge \mu(W_i))$$

Example 6.26. We survey four people regarding their subjectively placed value of a video streaming company, FlixTube. Each of the participants gives a scaled ranking of FlixTube on a scale of 0 to 1. This value placed on FlixTube, h, is the function we would like to integrate. We do this to determine a meaningful interpretation of the value of FlixTube by the entire focus group. However, perhaps not all responders answers should hold equal weight. Not all opinions are equally valid. To illustrate, denote $X = \{a, b, c, d\}$, the set of all participants in the survey. If respondent $a \in X$

watches FlixTube every day, then we should likely consider a's response of importance; whereas if respondent b has only watched two videos on FlixTube, then perhaps the inexperienced opinion of b should not be held in as high regard as a's. It is for this reason we additionally ask our participants to tell us how often they watch FlixTube. The weight assigned to each of their opinions is determined accordingly through an increasing function into $[0, 1]$. We call this assignment of weight to each participant the fuzzy set W. If we held all opinions equally, we could just as well take the average of the responses as the value of FlixTube. However, by weighting each response as in W, we may determine a more natural, overall value of Flixtube, a fuzzy average.

Suppose $h = \{(a, 0.3), (b, 0.5), (c, 0.8), (d, 0.9)\}$, and the fuzzy set W is the value of each participants opinion. $W = \{(a, 0.8), (b, 0.6), (c, 0.2), (d, 0.3)\}$.

We wish to take the Sugeno integral of h over W. To do this, we apply the Sugeno measure to take on the values of W at the singletons. That is $\mu(\{a\} = 0.8$, $\mu(\{b\}) = 0.6$, $\mu(\{c\}) = 0.2$, and $\mu(\{d\}) = 0.9$. Before continuing with the integral, we must first determine the value of λ for the Sugeno measure.

Note that we have chosen to calculate a Sugeno measure on X, for $\Phi = \{0, 1\}^X$, the crisp power set of X. Additionally, as this is for the Sugeno measure, $\mu(X) = 1$. Because X is the union of the four singletons, we may calculate λ by using the finite λ rule.

Case 1) If $\lambda = 0$, then by definition of the finite $\lambda-$rule, we find that $1 = \mu(X) = \mu(\{a\}) + \mu(\{b\}) + (\mu(\{c\}) + \mu(\{d\}) = 0.8 + 0.6 + 0.2 + 0.3 = 1.9$. Thus we have a contradiction.

Case 2) Suppose that $\lambda \neq 0$. Therefore, $1 + \lambda\mu(X) =$ $(1 + \lambda(\{a\}))(1 + \lambda(\{b\}))(1 + \lambda(\{c\}))(1 + \lambda(\{d\}))$. Since $\mu(X) = 1$, $\mu(\{a\}) = 0.8$, $\mu(\{b\}) = 0.6$, $\mu(\{c\}) = 0.2$, and $\mu(\{d\}) = 0.9$, we find that

$$1 + \lambda = (1 + 0.8\lambda)(1 + 0.6\lambda)(1 + 0.2\lambda)(1 + 0.9\lambda).$$

By using technology, we solve for the solutions of this equation; we find that the equation holds true for $\lambda = 0$ and $\lambda \approx -0.93543$. We have shown in case 1 that $\lambda \neq 0$; thus we must conclude that $\lambda \approx -0.93543$.

Now, we define for any sets A and B in the power set of X, $1 + \lambda(\mu(A \cup B)) = (1 + \lambda\mu(A))(1 + \mu(B))$ for $\lambda \approx -0.93543$. Thus we have defined the Sugeno measure on this set X.

Now let us consider the function h. By definition of the Sugeno integral, we must index X to be in decreasing order of h. This helps for the integral to not hold outliers of high (or low) opinion in too high regard. Thus, we let $x_1 = d, x_2 = c, x_3 = b$, and $x_4 = a$; thus we also infer a collection of sets $W_1 = \{x_1\}$, $W_2 = \{x_1, x_2\}$, $W_3 = \{x_1, x_2, x_3\}$ and $W_4 = X$. Now we apply the definition of the integral of h over W to be

$$(S)\int_X h = \bigvee_{i=1}^{4} (h(x_i) \wedge \mu(W_i))$$

We calculate $(S)\int_X h$ as follows. For $i = 1$, we find $\mu(W_1) = \mu(\{d\}) = 0.3$. For $i = 2$, we use the finite $\lambda-$rule to determine;

$$\mu(W_2) = \mu(\{x_1, x_2\}) = \mu(\{d, c\}) = \mu(\{d\}) + \mu(\{c\}) + \lambda\mu(\{d\})\mu(\{c\}).$$

Therefore, by our previous calculation of $\lambda \approx -0.93543$;

$$\mu(W_2) = 0.3 + 0.2 + (-0.93543)(0.3)(0.2) \approx 0.4439.$$

Continuing, we find

$$\mu(W_3) = \mu(W_2 \cup \{b\}) = \mu(W_2) + \mu(\{b\}) + \lambda\mu(W_2)\mu(\{b\}) \approx 0.7948.$$

Finally, as $W_4 = X$, $\mu(W_4) = 1$. Therefore,

$h(x_1) \wedge \mu(W_1) = 0.9 \wedge 0.3 = 0.3$

$h(x_2) \wedge \mu(W_2) = 0.8 \wedge 0.4439 = 0.4439$

$h(x_3) \wedge \mu(W_3) = 0.5 \wedge 0.7948 = 0.5$

$h(x_4) \wedge \mu(W_4) = 0.3 \wedge 1 = 0.3$

We calculate the Sugeno integral to be the supremum of all such values;

$$(S) \int_X h = \bigvee_i h(x_i) \wedge \mu(W_i) = 0.5.$$

We thus conclude that the value of FlixTube to the responders can be represented as the fuzzy average of their responses, 0.5.

To help illustrate the nature of the Sugeno integral, we now ask what properties of the Lebesque integral may be applied to the Sugeno integral. We provide proofs of some of these properties, listed in (Ralescu & Adams 1980) [41], below. Our following proofs do not depend on Φ be a $\sigma-$algebra of crisp sets, and thus these properties hold for $\sigma-$algebras of fuzzy sets. We prove each of these for the generalized definition of the Sugeno integral;

$$(S) \int_A f = \bigvee_{\alpha \geq 0} \alpha \wedge \mu(f_\alpha \cap A).$$

Proposition 6.27. Key properties of the Fuzzy Sugeno Integral: Let (X, Φ, μ) be a fuzzy measure space, and let $f, g : X \to [0, \infty)$ and $A, B \in \Phi$. The following hold true of the Sugeno integral.

(i) $f(x) \leq g(x) \forall x \in A, \implies (S) \int_A f d\mu \leq (S) \int_A g d\mu$

(ii) $A \subseteq B \implies (S) \int_A f d\mu \leq (S) \int_B f d\mu$

(iii) $\mu(A) = 0 \implies (S) \int_A f d\mu = 0$

(iv) $c \in [0, \infty) \implies (S) \int_A c = \mu(A) \wedge c$

Proof. Let (X, Φ, μ) be a fuzzy measure space and $f, g : X \to [0, \infty)$; suppose $A, B \in \Phi$.

(i) Suppose $f(x) \leq g(x)\ \forall x \in A$. To show $\int_A f \leq \int_A g$ we first show that $(f_\alpha \cap A) \subseteq (g_\alpha \cap A)$ for any $\alpha \geq 0$. To do this, first let $\alpha \geq 0$, and note that if $x \in f_\alpha$, then $\alpha \leq f(x) \leq g(x)$. Thus $x \in g_\alpha$. Therefore, $f_\alpha \subseteq g$. Therefore, by intersection, $f_\alpha \cap A \subseteq g_\alpha \cap A$. Consequently, by monotonicity of fuzzy measure μ, we see $\mu(f_\alpha \cap A) \leq \mu(g_\alpha \cap A)$. Therefore, we find $\alpha \wedge \mu(f_\alpha \cap A) \leq \alpha \wedge \mu(g_\alpha \cap A)$. This inequality holds for any $\alpha \geq 0$, and thus it holds for the supremum over all $\alpha \geq 0$. Hence $\int_A f = \bigvee\limits_{\alpha \geq 0} \alpha \wedge \mu(f_\alpha \cap A) \leq \bigvee\limits_{\alpha \geq 0} \alpha \wedge \mu(g_\alpha \cap A) = \int_A g$. Therefore, property (i) holds.

(ii) Suppose $A \subseteq B$. Thus, similar to the proof of the previous property, we find that by intersection of f_α, $f_\alpha \cap A \subseteq f_\alpha \cap B$. Now, using the same principles of monotonicity of μ and supremum over all α; we find that
$\int_A f = \bigvee\limits_{\alpha \geq 0} \alpha \wedge \mu(f_\alpha \cap A) \leq \bigvee\limits_{\alpha \geq 0} \alpha \wedge \mu(f_\alpha \cap B) = \int_B f$. Therefore, property (ii) holds.

(iii) Suppose $\mu(A) = 0$. Note that $f_\alpha \cap A \subseteq A\ \forall \alpha \geq 0$. Thus by monotonicity of μ, we find that $\mu(f_\alpha \cap A) = 0\ \forall \alpha \geq 0$. Therefore,
$\int_A f = \bigvee\limits_{\alpha \geq 0} \alpha \wedge \mu(f_\alpha \cap A) = \bigvee\limits_{\alpha \geq 0} \alpha \wedge 0 = \bigvee\limits_{\alpha \geq 0} 0 = 0$; property (iii) holds.

(iv) Let $c \in [0, \infty)$. Then $\int_A c = \bigvee\limits_{\alpha \geq 0} \alpha \wedge \mu(c_\alpha \cap A)$. Note, for constant c, $c_\alpha \cap A = \emptyset$ if $\alpha > c$ and $c_\alpha \cap A = A$ if $\alpha \leq c$. Since $\mu(\emptyset) = 0$, we thus only consider the case that $\alpha \leq c$. Thus, we find that $\int_A f = \bigvee (\alpha \wedge \mu(A))\ :\ \alpha \leq c$. Therefore, $\int_A f = c \wedge \mu(A)$; property (iv) holds.

□

Therefore, the Sugeno integral holds some common behaviors similar to that of the traditional, Lebesque integral. Thus forming a fuzzy measure theory, yield some results analogous to those of crisp measure theory, and not others. It is a topic of current research to continue identifying these conditions for convergence implications, as well as formulating new integral inequalities, as in crisp measure theory.

6.6 Places to Go Next

In this chapter, we have explored base definitions and theorems of some of the most commonly used constructions measure theory incorporating fuzziness. It is our aim that using this chapter, one may integrate themselves in fuzzy research literature with greater ease. With this aim in mind, we recommend the following works to explore further principles of the fuzzy theory of measures.

(i) (Couso et al., 2013) [10] In this paper, Couso, Garrido, and Sanchez explore the consequences of the many possible properties of fuzzy similarity measures, dissimilarity measures, $f-$near degrees, and possibilistic similarity measures.

(ii) (Kawabe, 2020) [20] In "Convergence in measure theorems of nonlinear integrals of functions integrable to the pth power", Kawabe explores explores necessary conditions of implication theorems of convergence in fuzzy measure space. This includes applications using the Choquet and Sugeno integral.

(iii) (Boczek et al., 2020) [7] In this paper, Bockzek, Hovana, Hutnik, and Kaluska expore the extension of Holders inequality $\int fgd\mu \leq (\int f^p)^{1/p} \cdot (g^q)^{1/q})$ for p, q conjugates], as well as Minkoski's inequality, $(\int (f+g)^s)^{1/s} \leq (\int f^s d\mu)^{1/s} + (\int g^s d\mu)^{1/s}$ for $s \in (0, \infty)]$.

CHAPTER 7

Approximate Reasoning and Fuzzy Systems

We have thus far discussed fuzziness through redefining common mathematical structures to model vagueness. However, pure mathematics is not the only place fuzziness exists. As mentioned in chapter 1, there is a multitude of applications of fuzzy logic. These include topics in control theory [2], risk assessment [19], and medical diagnosis [34]. The benefit of using such systems is often derived from the ability of fuzziness to allow computers to think more similarly to a human.

We wish now in this chapter to illustrate one such common topic of fuzzy research, a fuzzy inference systems. These systems are a collection of fuzzy If-Then rules on fuzzy sets, in combination with some form of aggregation and defuzzification functions. The product of such as system is a crisp output. Additionally, a special type of such systems, fuzzy additive systems, are explored by Bart Kosko, in (Kosko, 1994), to show that these structures may be used to approximate any continuous function on a compact subset of $\mathbb{R}^n$. In this chapter, we first the basis of these systems, fuzzy if-Then rules, as well as commonly used defuzzifiers. We then illustrate a common example, the Mamdani fuzzy inference system, and its application in control. Finally, we define a topic of recent fuzzy research, the Fuzzy Additive System (FAS), or Standard Additive Model (SAM). This additive system may be used to approximate a continuous function on a compact subset of $\mathbb{R}^n$ (Fuzzy Approximation Theorem) [22].

7.1 Fuzzy If-Then Rules

Let us now define a fuzzification of the implication operation; fuzzy If-Then rules. These rules can be viewed as a type of connective, similar to those discussed in chapter 3. There are multiple ways to construct a fuzzy implication which lead to an equivalent form of the crisp implication definition (when limited to crisp truth values).

Definition 7.1. Q-implication: (Nguyen, Walker, & Walker, 2019) [35] Let T, C, and η be a T-norm, T-conorm, and negation, respectively, on complete lattice L. Suppose C is dual to T with respect to η. We define for elements $x, y \in L$, the truth value that x implies y as follows.

$$(x \implies y) = (x \; T \; y)C(x^{\eta})$$

Example 7.2. (Buckley, 2002) [9] Let us consider the T-norm, T-conorm, and negation, min, max, and $' : x \mapsto 1 - x$. We define the Q-implication using these norms for $x, y \in [0, 1]$ by

$$(x \implies y) = (x \wedge y) \vee (1 - x)$$

Thus, the statement, "If x is A then y is B " , may be modeled by

$$\big(A(x) \implies B(y)\big) = (A(x) \wedge B(y)) \vee (1 - A(x))$$

Example 7.3. Consider the T-norm of multiplication, negation $' : x \mapsto 1 - x$, and the corresponding dual T-conorm, $C : (x, y) \mapsto 1 - (1 - x)(1 - y) = x + y - xy$. Thus, if we write a rule, "If x then y", the truth value of such a statement is defined as

$$(x \implies y) = (x \cdot y)C(1 - x) = (xy) + (1 - x) - xy(1 - x) = 1 - x + x^2 y$$

Thus, the statement, "If x is A then y is B " , may be modeled by

$$\big(A(x) \implies B(y)\big) = 1 - A(x) + (A(x))^2 B(y)$$

Recall that fuzzy sets model vague information. Thus, fuzzy implications are a means of modelling the truth of an implication such as, "If you are smart, you will make a lot of money". How smart is smart? What constitutes a 'lot' of money? The truth of this statement is almost certainly not 100% true; there are many smart people living in poor to moderate conditions in the world. However, the intelligence of a person certainly plays some factor in determining the size of one's salary. The population generally understands this. Yet, we hear statements such as this implication, and we understand that the implication is true, to a partial degree. In the general consensus, implications such as this are considered, 'true enough'. However, other fuzzy statements are not considered true enough for common speech. For example, "If a cow has brown spots, its milk will taste bad", is not accepted to be true enough for common knowledge. Therefore, we mathematically calculate the fuzzy truth value of any vague implication statement through these fuzzy implication operations,.

Example 7.4. Let us consider the set of all days X, and let C be the set of all cold days. Additionally, let Y be a set of people, and let S be the set of people staying indoors, where the membership of y is S is determined to be a fuzzy value based on the amount of time one spends indoors throughout a day. We then define the fuzzy Q-implication, "If x is C then y in S", by using the $(\wedge, \vee, ')$ norms; $\wedge = \min$, $\vee = \max$, and $' : x \mapsto 1 - x$, as defined in chapter 3. Let $X = \{x_1, x_2, x_3, x_4\}$ and $Y = \{y_1, y_2, y_3, y_4, y_5\}$. Define C such that $C = \{(x_1, 0.4), (x_2, 0.7), (x_3, 0.2), (x_4, 0.8)\}$ and $S = \{(y_1, 0.4), (y_2, 0.7), (y_3, 0.9), (y_4, 0.6), (y_5, 0.2)\}$.

We now can model the fuzzy phrase, "If x is a cold day, then y stays indoors". In similar fashion to crisp logic, we may define the value of the implication using a table truth values. For a $Q-$implication, we define the the implication as the table of values, $(C \times S) \cup (C^c \times Y)$, taking into account all possible combinations of elements of X and Y. We use the min T-norm to define the membership of the point (x, y) in $C \times S$. We then see that $C \times S$ takes the values.

	y1	y2	y3	y4	y5
x1	0.4	0.4	0.4	0.4	0.2
x2	0.4	0.7	0.7	0.6	0.2
x3	0.2	0.2	0.2	0.2	0.2
x4	0.4	0.7	0.8	0.6	0.2

Table 4: $C \times S$

By using the negation $' : x \mapsto 1 - x$, $C^c = \{(x_1, 0.6), (x_2, 0.3), (x_3, 0.8), (x_4, 0.2)\}$. Thus, we calculate $C^c \times Y$ to be as follows.

	y1	y2	y3	y4	y5
x1	0.6	0.6	0.6	0.6	0.6
x2	0.3	0.3	0.3	0.3	0.3
x3	0.8	0.8	0.8	0.8	0.8
x4	0.2	0.2	0.2	0.2	0.2

Table 5: $C^c \times Y$

Finally, we calculate the truth value of the implication statement to be $(x \implies y) = (C \times S) \cup (C^c \times Y)$; we use the max T-conorm to calculate the membership of each point, (x_i, y_j) in the union of these sets.

	y1	y2	y3	y4	y5
x1	0.6	0.6	0.6	0.6	0.6
x2	0.4	0.7	0.7	0.6	0.3
x3	0.8	0.8	0.8	0.8	0.8
x4	0.4	0.7	0.8	0.6	0.2

Table 6: $(C \times S) \cup (C^c \times Y) = (x_i \implies y_j)$: Cold Implies Reclusiveness

We thus calculate the truth value of the implication for any collection of combinations of x_i and y_j; we do this by taking the minimum of the truth values for all such x_i and y_i. For the entire sets X and Y, the statement in this example $(x_i \implies y_j) = 0.2$. However, for X and $Y - \{y_1, y_5\}$, the implication $(x_i \implies y_j) = 0.6$.

We may alternatively define an implication operation using any T-norm, and the order of the set of truth values, L. We model this implication by closely following that of the crisp implication definition. As such, we strive for this implication to be completely true whenever the antecedent is less than or equal to the consequent ($(x \implies y) = 1 \; if \; x \leq y$) as it is in crisp logic. However, for a consequent strictly less than the antecedent $(y < x)$, we delve beyond the scope of crisp logic, and allow for a fuzzy implication to be partially true.

Definition 7.5. R-implication (Nguyen, Walker, & Walker, 2019) Let T be a T-norm on complete lattice L. Let $x, y \in L$. We define the truth value of a fuzzy R-implication as

$$(x \implies y) = \bigvee \{z \in X \mid x \; T \; z \; \leq y\}.$$

Equivalently, for the purpose of application in fuzzy set implications, consider $x, y \in X$, a universal set, and L a complete lattice. Let $A, B \in L^X$. We then may say that "If x is A, then y is B", with fuzzy truth value of the statement defined by

$$(A(x) \implies B(y)) = \bigvee \{z \in L \mid A(x) \; T \; z \leq B(y)\}.$$

Example 7.6. Consider the statement, "If person, x, is tall, then the number of date requests they receive, y, is high". Let A be the fuzzy set of tall people in the universal set; let B the fuzzy set of people who receive a high number of date requests. Let us consider a person, Eric, in the universal set, who has membership 0.8 in the fuzzy set T. Suppose Eric receives 4 date requests within the allotted time frame; this is determined by the membership function of B to give Eric a membership value of 0.5 in the set B. To determine the truth statement of the implication above, we need choose a T-norm.

If $T = \min$, then $(A(x) \implies B(y)) = \bigvee\{z \in [0,1] \mid 0.8 \wedge z \leq 0.5\} = 0.5$.
If $T = \cdot$ (real number multiplication), then
$(A(x) \implies B(y) = \bigvee\{z \in [0,1] \mid 0.8 \cdot z \leq 0.5\} = 0.625$. Thus this implication is shown to have a mediocre truth value for element, $x =$Eric and either T-norm.

Remark 7.7. For an $R-$ implication, whenever $L = [0,1]$, $(x \implies y) = 1$ if $x \leq y$. This is evident through the identity operation of T-norms, $x\ T\ 1 = x \leq y$, and thus $(x \implies y) = 1$. However, the converse may not necessarily hold. Consider the T-norm $\triangle$ from example 3.9, defined as

$$x \triangle y = \begin{cases} x \wedge y & x \vee y = 1 \\ 0 & x \vee y \neq 1 \end{cases}$$

Therefore, for $x = 0.5$ and $y = 0.4$ we see that $x \triangle z = 0 \leq y$, $\forall z \in [0,1)$. Thus, $(x \implies y) = \bigvee[0,1) = 1$, yet $x > y$.

Alternatively, if we require that T be continuous, then we may make the conclusion if, $(x \implies y) = 1$ then $x \leq y$. To show this, let $(xy) = 1$. Suppose for sake of contradiction that $x > y$, then $x\ T\ 1 > y$ by use of the identity property of T-norms. Let $\epsilon = x - y > 0$. Since T is continuous, there must exist a $\delta > 0$ such that if $1 - w < \delta$ then $x\ T\ w > y$. Thus choose $w = 1 - \frac{\delta}{2}$. Therefore, $x\ T\ w > y$; thus since T is order preserving,
$(x \implies y) = \bigvee\{z \in X \mid x\ T\ z \leq y\} \leq w = 1 - \frac{\delta}{2} < 1$. Hence we have a contradiction. We thus conclude that for continuous T, if $(x \implies y) = 1$ then $x \leq y$.

Lemma 7.8. A fuzzy implication (of either $Q-$implication or $R-$implication definition) is a crisp implication when limited to crisp truth values.

Proof. $R-$implication) First let us consider the R-implication on $\{0,1\}$ defined for any T-norm, T. Let $x, y \in \{0,1\}$.

Case 1) If $x = 0$, then $(x \implies y) = \bigvee\{z \in [0,1] \mid x\ T\ z\ \leq y\}$. Note however, by Lemma 3.10, that $0\ T\ z = 0 \leq y\ \forall z \in [0,1]$, and thus $(x \implies y) = \bigvee[0,1] = 1$.

Case 2) Next, suppose that $x = y = 1$. Then for fized $z = 1$, we find that $x\ T\ z\ = 1 = y$, and thus $(x \implies y) = 1$.

Case 3) Suppose $x = 1$ and $y = 0$. Recall that for any $z \in [0,1]$ that $1\ T\ z = z$, thus, $1\ T\ z \leq 0$ if and only if $z = 0$. Thus, $(x \implies y) = 0$.

Therefore, when limited to crisp truth values, the R-fuzzy implication is an equivalent definition to crisp implication.

$Q-$implication) Now let us consider the Q-implication, such that T is a dual T-norm to T-conorm, C, with respect to negation, η, on $[0,1]$. Let $x, y \in \{0,1\}$.

Case 1) If $x = 0$, then by Lemma 3.10, $xTy = 0$. Additionally, by the identity property of T-conorm C, $(x\ T\ y)\ C\ x^\eta = 0\ C\ 1 = 1$. Thus, $(x \implies y) = 1$.

Case 2) If however $x = y = 1$, then by identity of T-norm T, $x\ T\ y = 1$. Thus by the identity property of C, $(x\ T\ y)\ C\ x^\eta = 1\ C\ 0 = 1$. Thus $(x \implies y) = 1$.

Case 3) Lastly, if $x = 1$ and $y = 0$, then by identity of T-norms, $x\ T\ y = 0$; also, by definition of negation, $x^\eta = 0$. Thus $(x\ T\ y)\ C\ x^\eta = 0\ C\ 0 = 0$, by identity property of C; $(x \implies y) = 0$.

Therefore, a Q-fuzzy implication is a crisp implication when limited to truth values of $\{0,1\}$.

$\square$

Thus, we have exemplified that the extension of implications to model vagueness does not contradict any results of crisp logic involving implications. As we have done throughout thispaper, we conclude that fuzzy logic is not disjoint from crisp logic. Rather, it is an extension of crisp logic to be defined for fuzzy truth values, as well as crisp truth values.

7.2 Defuzzification Methods

Progressing towards the discussion of fuzzy systems, we need introduce a process known as defuzzification. This process is exactly what it "says on the label"; it is a means of converting a fuzzy value to a crisp value. Suppose there exists a fuzzy set B of a universal set Y. A defuzzification operation is a function, $F_c : [0,1]^Y \to Y$, which converts the fuzzy set B to a single, crisp element $y \in Y$. In the context of fuzzy systems, this is frequently used in converting a final, fuzzy output into the "best" crisp output when this is required for decision making. For example, consider the fuzzy set A_7, of real numbers that are "about 7". If we were required to convert this fuzzy information to a crisp, single output value, we would choose such a value which most closely models the fuzzy set. We could choose 6, or 5, but of course, the "best" deffuzification of this set is 7. This process is thus trivial in the case that A be defined as a set of values 'about' or 'around' a crisp value. However, as we have seen throughout this paper, many fuzzy sets are not so defined. As such, defuzzification is not always so trivial. There are multiple defuzzification methods used in the context of fuzzy systems, to defuzzify any fuzzy set into a single crisp value. Some of the most common include the Center of Gravity (or Centroid Defuzzifier), First of Maxima, Mean of Maxima, and Weighted Average.

Definition 7.9. Smallest/Largest/Mean of Maxima:(Alassar, Ahmed & Abuhadrous, 2010) [3] This method is commonly used by the Mamdani Fuzzy System. Suppose that there is a fuzzy set, B, of Y such that Y is a fully ordered set and the membership function $B(x) : Y \to [0,1]$ obtains a maxima. The smallest of maxima defuzzifier, (SOM) is the function $SOM(B) = y$, where y is the least element in Y for which B obtains maximum membership. Similarly, the largest of maxima, $LOM(B) = y$, such that y is the greatest element for which $B(y)$ is the maximum value of B. Lastly, Mean of Maxima, $MOM(B)$, for finite set Y, is the average of all elements of Y for which B obtains maximum membership.

Example 7.10. Consdider the fuzzy set B, of $\mathbb{Z}$, defined such that

$$B(x) = \begin{cases} 0 & x \leq 0 \\ 0.3 & 0 < x \leq 10 \\ 0.7 & 10 < x \leq 14 \\ 0 & 14 < x \end{cases}$$

Thus $SOM(B) = 10$, $LOM(B) = 14$ and $MOM(B) = 12$.

Definition 7.11. Center of Gravity/Centroid Defuzzifier: (Kosko, 1994) [22] Another commonly used defuzzification method is the Center of Gravity (Centroid defuzzifier). This operation calculates a crisp value based on the area under the curve of any integrable membership function of B, whenever the universal set Y is a subset of the real numbers.

Let $B \in [0,1]^Y$, where $Y \subseteq \mathbb{R}$. We define COG to be the function

$$y = COG(B) = \frac{\int_Y x \cdot B(x)}{\int_Y B(x)} \text{ ; or if Y is finite, } y = COG(B) = \frac{\sum_{i=1}^{n} x_i \cdot B(x_i)}{\sum_{i=1}^{n} B(x_i)}$$

Example 7.12. Consider the set $X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and the subset $B = \{(0, 0.8), (1, 0.7), (2, 0.43), (3, 0.7), (4, 0.8), (5, 0.78), (6, 0.23), (7, 0.39), (8, 0.64), (9, 0.44)\}$.

Since X is a finite set, we calculate the Center of Gravity to be

$$COG(B) =$$

$$\frac{(0\cdot 0.8+1\cdot 0.7+2\cdot 0.43+3\cdot 0.7+4\cdot 0.8+5\cdot 0.78+6\cdot 0.23+7\cdot 0.39+8\cdot 0.64+9\cdot 0.44)}{(0.8+0.7+0.43+0.7+0.8+0.78+0.23+0.39+0.64+0.44)}$$

$$= (23.95) \div (5.91) \approx 4.05$$

We therefore, conclude that for a system given fuzzy output B, the defuzzified crisp output used for decision making is 4.05.

Example 7.13. Consider $Y = [0, 10]$ and $B \in [0, 1]^{[0,10]}$, such that

$$B(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{2}x - 1 & 2 \leq x < 4 \\ 1 & 4 \leq x \leq 8 \\ -\frac{1}{2}x + 5 & 8 < x \end{cases}$$

Thus, $\int_X x \cdot B(x)dx = \int_2^4 x(\frac{1}{2}x - 1)dx + \int_4^8 x \cdot 1dx + \int_8^{10} x \cdot (-\frac{1}{2}x + 5)$

$= \frac{10}{3} + 24 + \frac{26}{3} = 36.$

Additionally, $\int_X B(x)dx = \int_2^4 (\frac{1}{2}x - 1)dx + \int_4^8 1dx + \int_8^{10} (-\frac{1}{2}x + 5)dx = 1 + 4 + 1 = 6.$

Therefore, for the defuzzified value of B, $COG(B) = \frac{36}{6} = 6.$

7.3 Mandani Fuzzy System

We now look at an example of one of the most commonly used fuzzy systems, the Mamdani Fuzzy Inference System. The Mamdani system uses a collection of rules of the R-implication form, as well as the min T-norm. We denote such rules as $\{R_i\}_{i=1}^n$ where R_i is the statement "If x is A_i, then y is B_i", where x and y come from universal sets X and Y respectively, and $A_i \in [0, 1]^X$ and $B_i \in [0, 1]^Y$. First consider the individual statement, "If x is A, then y is B". The Mamdani system desires that each rule have truth value 1; since min is continuous, by Remark 7.7, this requires that $A(x) \leq B(y)$.

Generally, the Mamdani system uses a set of rules of the form, " If x is in A_i then y is in B_i", where $B_i \neq B_j \; \forall i \neq j$. Some rules may be combined statements, or some rules of different input sets may have the same output set. However, rules of these forms may be considered as to be rules of the simple form by combining rules in the following ways.

Consider the case of a set of rules;

If x is A_1, then y is B,

If x is A_2, then y is B.

...

If x is A_n, then y is B.

We interpret the space between these implications as having an implied "or" between them; if x is in A_1 or A_2 or...A_n, then y is in B. 'Or' is modeled with union, just as in crisp logic; thus we use a T-conorm, generally the max operation, to find that we may combine these rules into one of the form, "If x is A, then y is B", where $A = \bigcup_{i=1}^{n} A_i$.

Similarly, consider the case of multiple input variables: "If x_1 is A_1 and x_2 is A_2, then y is in B". For this case, we then may interpret the multiple variables using the min operation to model 'and'. We thus construct a fuzzy cross product $(A_1 \times A_2)$ such that $(A_1 \times A_2)(x_1, x_2) = \min\{A_1(x_1), A_2(x_2)\}$. Thus the rule can be equivalently written as, "If $(x_1, x_2) \in (A_1 \times A_2)$, then $y \in B$.

Therefore, any set of implications of the above cases may be reduced to a collection of implications of the form; "If x is in A_i, then y is in B_i, where $B_i \neq B_j \forall i \neq j$. For the remainder of our discussion, we consider a system of rules of this form.

In this system of fuzzy implications, the input or output sets are frequently of the form of a symmetric triangular fuzzy number. These fuzzy numbers are thus defined to help clarify the upcoming examples of a Mamdani fuzzy system.

Definition 7.14. Symmetric Triangular Fuzzy Number: A fuzzy set, A, of universal set, $\mathbb{R}$, is said to be a triangular fuzzy number iff A attains a value of 1 at a unique element $a \in \mathbb{R}$, and is a piecewise, triangular-shaped membership function, symmetric on either side of the element a. That is, for $A \in [0,1]^{\mathbb{R}}$, A is a triangular fuzzy number if it is of the following form. Let $a \in \mathbb{R}$ be fixed, and suppose there exist lines $bx + y_1$ and $-bx + y_2$ for some $b \in (0, \infty)$ such that these lines intersect at the point $(a, 1)$. Then the fuzzy set A is defined as follows:

$$A(x) = \begin{cases} \max\{0, bx + y_1\} & x \leq a \\ \max\{0, -bx + y_2\} & a \leq x \end{cases}$$

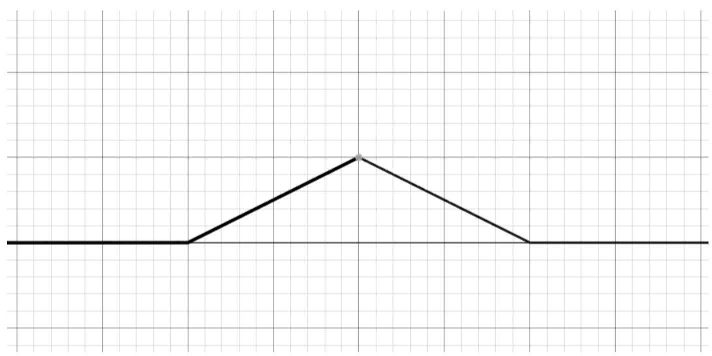

Figure 11: Symmetric Triangular Fuzzy Number

Example 7.15. Consider three fuzzy rules, $\{R_i\}_{i=1}^3$ of the form, "If x is A_i then y is B_i", where A_i and B_i are fuzzy subsets of $[0, 10]$. Define these sets as follows.

Let $\{A_i\}_{i=1}^3$ be a collection of symmetric triangular fuzzy numbers with $b = \frac{1}{2}$, such that $A_1(1) = A_2(2) = A_3(3) = 1$. Let $\{B_i\}_{i=1}^3$ be symmetric triangular fuzzy numbers with $b = \frac{1}{2}$, such that $B_1(3) = B_2(5) = B_3(7) = 1$.

Now let us consider the element 3. $A_1(3) = 0$, $A_2(3) = 0.5$, and $A_3(3) = 1$. As 3 has non-zero membership in A_2 and A_3, we thus consider rules R_2 and R_3 to be "fired"; we thus calculate an output fuzzy set, B_y, such that the truth values of these implication rules for y is 1.

Hence, we calculate output fuzzy sets B_2' such that $B_2'(y) = \min\{B_2(y), 0.5\}$, and $B_3'(y) = \min\{B_3(y), 1\} = B_3(y)$. We then aggregate the two output sets, B_2' and B_3' by the union of the two sets. That is $B'(y) = (B_2' \cup B_3')(y) = \max\{\min\{B_2(y), 0.5\}, B_3(y)\}\}$. Therefore,

$$B'(y) = \begin{cases} 0 & 0 \leq y < 3; 9 \leq y \\ 0.5y - 1.5 & 3 \leq y \leq 4 \\ 0.5 & 4 \leq y < 6 \\ 0.5y - 2.5 & 6 \leq y < 7 \\ -0.5y + 4.5 & 7 \leq y < 9 \end{cases}$$

B' represents the output of the Mamdani system. B' is a fuzzy set, which we need defuzzify using, in most cases, one of the defuzzification methods discussed in the previous section.

Using any of the maxima defuzzifiers, we find that $SOM(B') = LOM(B') = MOM(B') = 7$. However, by using the center of gravity deffuzfier, we find that

$\int_0^{10} y \cdot B'(x)dy =$

$\int_3^4 y(0.5y - 1.5)dy + \int_4^6 y(0.5)dy + \int_6^7 y(0.5y - 2.5)dy + \int_7^9 y(-0.5y + 4.5)dy$

$= \frac{11}{12} + 5 + \frac{59}{12} + \frac{23}{3} = 18.5$

Additionally, we find that $\int_0^{10} B'(y) =$

$\int_3^4 (0.5y - 1.5)dy + \int_4^6 (0.5)dy + \int_6^7 (0.5y - 2.5)dy + \int_7^9 (-0.5y + 4.5)dy$

$= \frac{1}{4} + 1 + \frac{3}{4} + 1 = 3$

Thus, we calculate the defuzzified output to be

$$COG(B') = \frac{\int_0^{10} yB'(y)dx}{\int_0^{10} B'(y)dy} = \frac{18.5}{3} = \frac{37}{6} \approx 6.17$$

Remark 7.16. (Buckley, 2002) [9] In the above example 7.15, we have used inputs as crisp elements, however the Mamdani system may also take fuzzy input values. In the above case of a crisp singleton, we may equivalently state that the input a, is the fuzzy singleton, $a_\lambda : \lambda = 1$ (see Ch.4). Therefore, we may equivalently define a set $A' = A \cap a_1$, such that $A'(x) = A(x)$ if $x = a$ and $A'(x) = 0$ otherwise. Consider the value $h(A')$, denoting the height of A'; the supremum of the membership function A. Thus $h(A') = A(a)$. Therefore, we may define our previously used output set B', for input x, as being $B'(y) = \min\{h(A'), B(y)\}$. By using this definition, we may replace a_1 by any fuzzy subset of X, and thus the system may accept any fuzzy input.

Example 7.17. (Izuierdo et al., 2015) [18] Suppose we wish to rank a house from 1 to 10, based on how suitable it would be for a possible buyer. To do this, let us consider two fuzzy variables of concern for the buyer; "Is the house expensive?", and "Is the house far from work?". While we could create an analysis of the system using crisp variables (price and distance) this would not model the fuzziness of the original variables. Thus, by scaling data from surveys asking what values are considered "far" and "expensive", we create fuzzy sets modelling these vague variables. We thus construct the following fuzzy If-Then rules.

"If the house is inexpensive or close to work, then suitability is good."
"If the house is expensive or far from work, then suitability is low."
"If the house is average-priced and close to 50 km from work, then suitability is regular."

We use the max T-conorm to model "or", the min T-conorm to model "and", and the max operation for fuzzy output set aggregation. This system obtains the form diagrammed below. Hence, for a particular house costing \$50,000, and distance 35 kilometers from work, we find the suitability of living at a given house with these parameters is 6.03845.

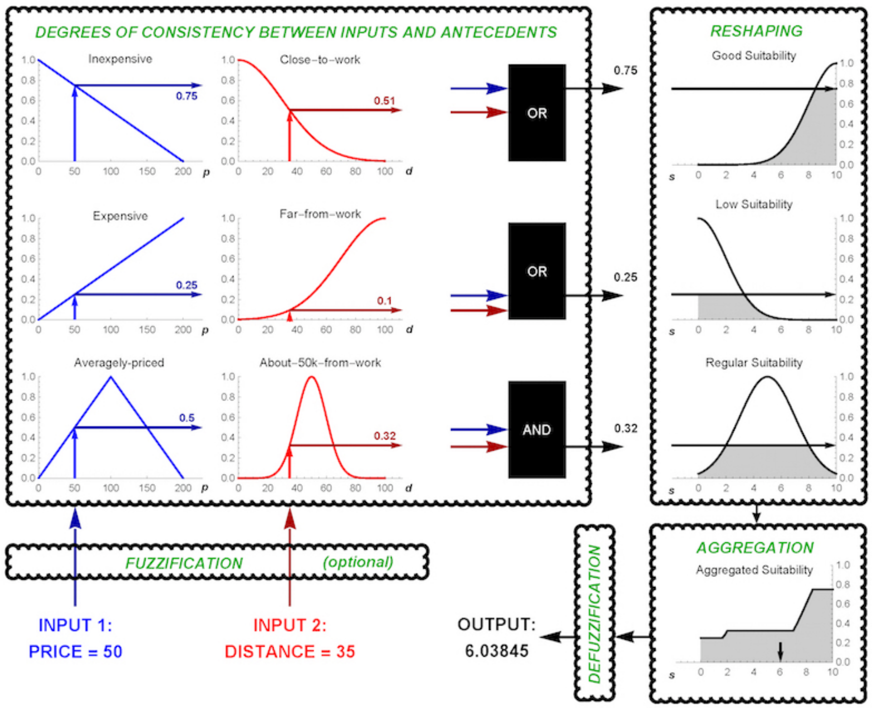

Figure 12: Mamdani System for House-Buyer Suitability [18]

Example 7.18. Consider the application of controlling an air conditioning system. Let $\{A_i\}_{i=1}^5$ be a set of fuzzy subsets of the interval $[50, 90]$, such that A_i are the fuzzy sets modeling 'cold', 'cool', 'just right' , 'warm', 'hot', respectively. We model 5 fuzzy output sets $\{B_i\}_{i=1}^5$ of $[0, 100]$, which represent fuzzy commands for the air conditioner fan speed, 'blast', 'fast', 'medium', 'slow', 'stop'.

With these fuzzy sets, we may construct a set of fuzzy rules such that we may control the strength of the air conditioner by temperature input; we thus account for the vague way we as humans quantify whether the temperature is hot or cold.

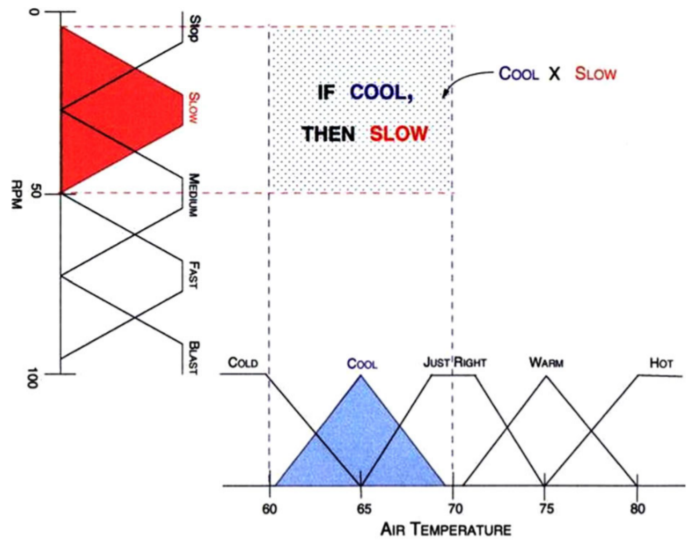

Figure 13: Fuzzy Air Conditioning Control [23]

Thus, we have a programmable system of rules of the form, "If air temperature is cool, then fan speed is slow". Using this system, the air conditioner controls the air temperature with logic allowing for the vagueness that humans use in communicating comfort or discomfort with their given surrounding.

7.4 Additive Fuzzy Systems

Finally, we now discuss a special form of a fuzzy system, an Additive Fuzzy System, or Standard Additive Model (SAM); thus named for the weighted averaging aggregation technique used. These systems have a broad array of applications, including medical diagnosis [34], and can be used to approximate continuous functions.

Definition 7.19. Fuzzy Additive System: (Kosko, 1994) [22] Let $\{R_i\}_{i=1}^{n}$ be a collection of fuzzy If-Then rules of the form, "If x is A_i, then y is B_i". We thus define the output of the Fuzzy Additive System, or Standard Additive Model to be

$$F(x) = centroid(B) = \frac{\int_{-\infty}^{\infty} y \sum_{i=1}^{m} B_i'(y)dy}{\int_{-\infty}^{\infty} \sum_{i=1}^{m} B_i'(y)dy}.$$

Here B is the fuzzy output of the system, determined using the T-norm $\cdot$ of real number multiplication, rather than min. That is, for fixed $x \in X$, $B(y) = \sum_{i=1}^{n} w_i A_i(x) B_i(y)$ where w_i represents the weight of the rule R_i, and $\sum_{i=1}^{n} w_i = 1$.

There is thus an equivalent form of this definition where w_i is the weight of rule R_i; the centroid of each of the fuzzy output set, B_i, is calculated by the centroid defuzzifier to be c_i; amd the area (volume) of each B_i is denoted V_i. We thus calculate the value of $F(x)$ to be:

$$F(x) = \frac{\sum_{i=1}^{n} w_i A_i(x) V_i c_i}{\sum_{i=1}^{n} w_i A_i(x) V_i}$$

Note if we weight all rules equally, we may ignore w_i completely. Thus, if we define a value $p_i(x)$ to be $p_i(x) = \frac{A_i(x)v_i}{\sum_{i=1}^{n} A_i(x) \cdot v_i}$, we find, $F(x) = \sum_{i=1}^{n} p_i(x) c_i$ (Nguyen & Peterson, 2008) [33]. Thus the Additive Fuzzy System may be viewed as a weighted average of the centroid of each output set.

Remark 7.20. (Pal et al., 2000) [39] With equal weights, the two definitions of 7.19 are the equivalent.

$$F(x) = centroid\left(\sum_{i=1}^{n} w_i A_i(x) B_i\right) = \frac{\sum_{i=1}^{n} A_i(x) V_i c_i}{\sum_{i=1}^{n} A_i(x) V_i} = \frac{\int_{\mathbb{R}} y \cdot B'(y)dy}{\int_{\mathbb{R}} B'(y)dy}$$

To show this, let $x \in \mathbb{R}^n$. B_i' is defined as a fuzzy set of real numbers such that $B_i'(y) = A_i(x) \cdot B_i(y)$. This is sometimes denoted $B_i'(y|x) = A_k(x) \cdot B_k(y)$ if fixed x is not clear. We then define the output of an element x of the fuzzy system, F to be

$$F(x) = \frac{\int_{-\infty}^{\infty} y \sum_{i=1}^{m} B_i'(y)dy}{\int_{-\infty}^{\infty} \sum_{i=1}^{m} B_i'(y)dy} = \frac{\sum_{i=1}^{m} A_i(x) \int_{-\infty}^{\infty} yB_i(y)dy}{\sum_{i=1}^{m} A_i(x) \int_{-\infty}^{\infty} B_i(y)dy}$$

$$= \frac{\sum_{i=1}^{m} A_i(x)V_i \frac{\int_{-\infty}^{\infty} yB_i(y)dy}{V_i}}{\sum_{i=1}^{m} A_i(x) \int_{-\infty}^{\infty} B_i(y)dy} = \frac{\sum_{i=1}^{m} A_i(x)V_ic_i}{\sum_{i=1}^{m} A_i(x)V_i}$$

The primary mathematical application of this system comes in the form of function approximation. Kosko [22] illustrates; by using a finite number of fuzzy If-Then rules, we can approximate a real valued function on any compact subset of $\mathbb{R}^n$. To show this, we use the Euclidean metric, E, on $\mathbb{R}^n$.

Theorem 7.21. Fuzzy Approximation Theorem (Kosko, 1994) [22] Let $X \subseteq \mathbb{R}^n$, such that X is compact (closed and bounded). Let $f : X \to \mathbb{R}$ be continuous. We may construct a fuzzy additive system F of finite rules which approximates f to any accuracy.

Proof. Let $\epsilon > 0$. We need show that we may construct a fuzzy additive system, F, such that $|F(x) - f(x)| < \epsilon \ \forall x \in X$. Since f is continuous on a compact X, there exists a $\delta > 0$ such that if $E(x, y) < \delta$, then $|f(x) - f(y)| < \frac{\epsilon}{4}$.

Since X is compact, we may construct a finite collection of crisp, open cubes of equal size, $\{A_i\}_{i=1}^m$, in $\mathbb{R}^n$, such that the diameter of A_i is less than δ for all i, and $\{A_i\}_{i=1}^m$ covers X. Consequently the length of A_i for any dimension is less than δ. Construct these such that for any A_i, the corner of A_i exists at the center of its neighbor.

Let $\{B_i\}_{i=1}^m$ be a collection of symmetric, integrable fuzzy output sets, centered at $f(m_i)$ where m_i represents the center of A_i, for each i. Thus, as B_i is symmetric, we find that the centroid of B_i is $f(m_i)$. We thus define a fuzzy additive system, F, of m rules of the form, "If x is A_i, then y is B_i".

Now let us show that this F approximates f within ϵ. Let $x \in X$. Thus, x exists in some collection of sets in $\{A_i\}$. Select y in a cube which overlaps with a cube in which x exists; say x exists in A_j and y exists in A_k. Thus, for any $z \in A_j \cap A_k$, we find that $|x - z| < \delta$ and $|y - z| < \delta$. Thus,

(i) $|f(x) - f(y)| \leq |f(x) - f(z)| + |f(z) - f(y)| < \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{\epsilon}{2}$.

Additionally, choosing $x = m_j$ and $y = m_k$, we find that (ii) $|f(m_j) - f(m_k)| < \frac{\epsilon}{2}$.

Now note that, if $x \in A_i$, then $|x - m_i| < \delta$. Thus (iii) $|f(x) - f(m_i)| < \frac{\epsilon}{4} < \frac{\epsilon}{2}$.

Also, since A_i are crisp sets, $F(x)$ is the weighted average of all $f(m_i)$ such that $x \in A_i$.

$$F(x) = \frac{\sum\limits_{i \ s.t. \ x \in A_i} w_i V_i f(m_i)}{\sum\limits_{i \ s.t. \ x \in A_i} w_i V_i}$$

Therefore, $F(x)$ lies between at least two $f(m_i)$, say $f(m_p)$ and $f(m_r)$, such that $x \in A_p \cap A_r$. Therefore, by (ii), $|F(x) - f(m_p)| \leq |f(m_p) - f(m_r)| < \frac{\epsilon}{2}$. Also, by (i), for any element , y, that is in the same A_i as x, $|f(y) - f(x)| < \frac{\epsilon}{2}$. Thus, choose $y = m_p$' by (iii), we find that

$$|F(x) - f(x)| \leq |F(x) - f(m_p)| + |f(m_p) - f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Thus, $F(x)$ approximates f using m rules to accuracy within ϵ. $\square$

7.5 Places to Go Next

We have thus far discussed the core constructions of fuzzy system, and their mathematical base of fuzzy If-Then rules. These systems, as we would expect, are particularly useful when modeling or controlling for vagueness and uncertainty. This has a broad array of applications, some of which explored in the papers listed below. These papers help to illustrate the diversity of the impact of fuzzy systems

(i) (Akisue et al., 2018) [2] In "Development of a fuzzy system for dissolved oxygen control in a recombinant Escherichia coli cultivation for heterologous protein expression", authors Rafael Akiue, Antonio Horta, and Ruy de Sousa Jr. address a possible improvement to the program SUPRSYS_HCDC. The Federal University of Sao Carlos's Chemical Engineering Department uses this program to regulate conditions, including oxygen injection, of a E. coli cultivation. The original program allows for no smooth transition within its decision tree, and thus Akisue et al. present the possible improvement of using a fuzzy system to regulate the oxygen in the cultivation.

(ii) (Karmipour et al., 2016) [19] In "New Fuzzy Model Based for Risk Assessment Based on Different Types of Consequences", K. Karimpour, R. Zarghami, M.A. Moosavian, and H. Bahmanyar develop a fuzzy system using the Mamdani Algorithm to evaluate risk level. This includes an application of evaluating the risk of personal injury with respect to a pipe carrying propane under different temperature conditions.

(iii) (Kosko, 2018) [23] In "Additive Fuzzy Systems: From Generalized Mixtures to Rule Continua", Bart Kosko shows that an additive fuzzy system, as discussed in this chapter, has an underlying structure of a probability mixture. Here, If-Then rules correspond to probability distribution densities.

(iv) (Nguyen et al., 2014) [34] In “Medical Diagnosis by Fuzzy Standard Additive Model with Wavelets”, Thanh Nguyen, Abbas Khosravi, Douglas Creighton and Saeid Nahavandi explore the application of combining standard additive models and wavelet features to reduce the data size, in order to diagnose breast cancer and heart disease.

CHAPTER 8

Closing Remarks

Inspired by a desire to codify and model vagueness, the core definitions and constructions of fuzzy set theory has set the standards for current fuzzy research. As set theory is the foundation of modern mathematics, the extension of such mathematics to be defined for fuzzy sets has been regularly the topic of interest to fuzzy mathematicians. This paper has aimed to provide a guided entry into the fuzzy world. As such, we have outlined the most standard structure of fuzzy set theory, as well as L-fuzzy set theory (ch.2 & ch.3). We have done this so that one may explore any fuzzy research with standard preliminaries.

Additionally, we have modeled selected techniques of fuzzifying mathematical structures. This has been done by redefining traditional definitions, to incorporate fuzzy set theory. We have consistently done so with the goal of maintaining equivalence to crisp definitions when limited to crisp truth values or crisp certainty values. In the realm of algebra (ch.4), we have fuzzified the concept of closure using an inequality to define fuzzy groupoids, groups, and quotient groups. We have shown that each of these are equivalent to their crisp counterparts when the membership of their elements is crisp. In the realm of analysis, we have appended a variable of certainty to a metric (ch.5). We have shown that this induces a crisp metric when such variable is fixed, particularly when certainty is 1 as it is in crisp metrics. For fuzzy measures (ch.6), we have weakened a singular condition of the core definition, additivity, and replaced it with a weaker, $\lambda-$additivity for $\lambda-$fuzzy measures. We have shown that these measures are equivalent to crisp measures when $\lambda = 0$. In each of these methods of fuzzyfication, we have provided paradigms to assist in developing intuition and mastery of the respective areas of fuzzy mathematics. Each of these methods may be generalized to techniques of fuzzifying traditional mathematics. We have consistently done so to expand the capacity of crisp structures to include fuzziness, without loosing the results of crispness. Thus, we maintain that

$$Crisp\ Mathematics\ \subset\ Fuzzy\ Mathematics.$$

Finally, in the world of programming, fuzzy logic is researched for its direct ability to model vagueness. We have illustrated that a human brain uses fuzzy reasoning in many of its classifications and decision making (ch.1). Inspired by this knowledge, we have explored a structure known as a fuzzy inference system, a system of fuzzy rules using fuzzy classifiers for decision making (ch.7).

This paper has given a survey of fuzzy logic and has illustrated some of its applications. We have done so with the intent that this paper help prepare a mathematician desiring to begin working on, or assist in, fuzzy research. There are many papers on the subject, for which this paper may serve as a supplementary material for. However, we have built the paper to prepare the reader for the recent works listed in the "Places to Go Next" sections of each chapter. These include further research in fuzzy set theory, fuzzy algebra, fuzzy metrics, fuzzy measures, and fuzzy additive systems.

BIBLIOGRAPHY

[1] M. Akgül, Some properties of fuzzy groups, Journal of Mathematical Analysis and Applications. 133 (1988) 93-100.

[2] R.A. Akisue, A.C.L. Horta, R. de Sousa, Development of a fuzzy system for dissolved oxygen control in a recombinant Escherichia coli cultivation for heterologous protein expression, Computer Aided Chemical Engineering. 43 (2018) 1129-1134.

[3] Alassar, Ahmed Abuhadrous, I.M. Elaydi, H.. Modeling and control of 5 DOF robot arm using supervisory control. 2010 The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010. 3. (2010) 351 - 355

[4] E.G. Antzoulatos, E.K. Miller, Differences between Neural Activity in Prefrontal Cortex and Striatum during Learning of Novel Abstract Categories, Neuron. 71 (2011) 243-249.

[5] Aufmann, R. N., Clegg, D. K., Nation, R. D., Lockwood, J. S., amp; SSM MATHEMATICAL EXCURSIONS (4th ed.). CENGAGE LEARNING CUSTOM P, (2017).

[6] K.T. Atanassov, Intuitionistic fuzzy sets, Fuzzy Sets Syst. 20 (1986) 87-96.

[7] M. Boczek, A. Hovana, O. Hutník, M. Kaluszka, Hölder-Minkowski type inequality for generalized Sugeno integral, Fuzzy Sets Syst. 396 (2020) 51-71.

[8] M.H. Bornstein, M.E. Arterberry, The Development of Object Categorization in Young Children, Developmental psychology. 46 (2010) 350-365.

[9] J.J. Buckley, E. Eslami, An Introduction to Fuzzy Logic and Fuzzy Sets, Physica-Verl, Heidelberg ; New York, 2002.

[10] I. Couso, L. Garrido, L. Sánchez, Similarity and dissimilarity measures between fuzzy sets: A formal relational study, Inf. Sci. 229 (2013) 122-141.

[11] Einstein, A. Geometry and experience: An address delivered before the Prussian Academy of Sciences in Berlin, January 27, 1921. St. John's: St. John's Press, (1971).

[12] A. Emniyet, M. Şahin, Fuzzy Normed Rings, Symmetry (Basel). 10 (2018) 515.

[13] A. George, P. Veeramani, On some results in fuzzy metric spaces, Fuzzy Sets Syst. 64 (1994) 395-399.

[14] M. Grabiec, Fixed points in fuzzy metric spaces, Fuzzy Sets Syst. 27 (1988) 385-389.

[15] V. Gregori, J. Miñana, D. Miravet, Contractive sequences in fuzzy metric spaces, Fuzzy Sets Syst. 379 (2020) 125-133.

[16] V. Gregori, A. Sapena, On fixed-point theorems in fuzzy metric spaces, Fuzzy Sets Syst. 125 (2002) 245-252.

[17] J.A. Goguen, L-fuzzy sets, Journal of Mathematical Analysis and Applications. 18 (1967) 145-174.

[18] L.R. Izquierdo, D. Olaru, S.S. Izquierdo, S. Purchase, G.N. Soutar, Fuzzy Logic for Social Simulation Using NetLogo, Journal of Artificial Societies and Social Simulation. 18 (2015). http://jasss.soc.surrey.ac.uk/18/4/1.html#novak2012

[19] K. Karimpour, R. Zarghami, M.A. Moosavian, H. Bahmanyar. New Fuzzy Model for Risk Assessment Based on Different Types of Consequences. Oil Gas Science and Technology - Revue d'IFP Energies nouvelles, Institut Français du Pétrole, 2016, 71 (1), pp.17.

[20] J. Kawabe, Convergence in measure theorems of nonlinear integrals of functions integrable to the pth power, Fuzzy Sets Syst. 396 (2020) 29-41.

[21] Kosko, B. Fuzzy thinking: the new science of fuzzy logic, 1. ed. ed., Hyperion, New York, 1993.

[22] B. Kosko, Fuzzy systems as universal approximators, TC. 43 (1994) 1329-1333.

[23] Kosko, Bart. Additive Fuzzy Systems: From Generalized Mixtures to Rule Continua: Additive Fuzzy Systems. International Journal of Intelligent Systems. 33. (2018).

[24] What are common causes of blurry vision?, Larson Eye Care. (August 8th, 2019) https://larsoneyecare.com/what-are-common-causes-of-blurry-vision/

[25] J. Li, Order continuous of monotone set function and convergence of measurable functions sequence, Applied Mathematics and Computation. 135 (2003) 211-218.

[26] W. Liu, Fuzzy invariant subgroups and fuzzy ideals, Fuzzy Sets Syst. 8 (1982) 133-139.

[27] N. Marín, G. Rivas-Gervilla, D. Sánchez, R.R. Yager, Specificity measures based on fuzzy set similarity, Fuzzy Sets Syst. 401 (2020) 189-199.

[28] Wilkin, D., Mitosis. (2019, November 20). https://www.ck12.org/biology/mitosis/lesson/Mitosis-Advanced-BIO-ADV/

[29] K. Mittal, A. Jain, K.S. Vaisla, O. Castillo, J. Kacprzyk, A comprehensive review on type 2 fuzzy logic applications: Past, present and future, Eng Appl Artif Intell. 95 (2020) 103916

[30] S. Mohinta, T. Samanta, An introduction to fuzzy soft graph, Mathematica Moravica. 19 (2015) 35-48.

[31] J.A. Momoh, X.W. Ma, K. Tomsovic, Overview and literature survey of fuzzy set theory in power systems, TPWRS. 10 (1995) 1676-1690.

[32] N.N. Morsi, S.E. Yehia, Fuzzy-quotient groups, Inf. Sci. 81 (1994) 177-191.

[33] Thi Thanh Nguyen, J. Peterson, Enhancing the Performance of the Fuzzy System Approach to Prediction, FSKD. 1 (Oct 2008) 259-265.

[34] Thanh Nguyen, A. Khosravi, D. Creighton, S. Nahavandi, Medical diagnosis by fuzzy standard additive model with wavelets, FUZZ-IEEE. (Jul 2014) 1937-1944.

[35] H.T. Nguyen, C.L. Walker, E. Walker, A First Course in Fuzzy Logic, Fourth edition ed., CRC Press/Taylor Francis Group, Boca Raton ; London ; New York, 2019.

[36] V. Novák, S. Lehmke, Logical structure of fuzzy IF-THEN rules, Fuzzy Sets Syst. 157 (2006) 2003-2029.

[37] S. Ovchinnikov, Similarity relations, fuzzy partitions, and fuzzy orderings, Fuzzy Sets Syst. 40 (1991) 107-126.

[38] M. Paknazar, M. De La Sen, Some new approaches to modular and fuzzy metric and related best proximity results, Fuzzy Sets Syst. 390 (2020) 138-159.

[39] S.K. Pal, A. Ghosh, M.K. Kundu, Soft Computing for Image Processing, (2000).

[40] C. Phillips, C.L. Karr, G. Walker, Helicopter flight control with fuzzy logic and genetic algorithms, Engineering applications of artificial intelligence. 9 (1996) 175-184.

[41] D. Ralescu, G. Adams, The fuzzy integral, Journal of Mathematical Analysis and Applications. 75 (1980) 562-570.

[42] Rosch, E. H. . On the internal structure of perceptual and semantic categories. Cognitive development and acquisition of language, Elsevier Inc. (1973) 111-144.

[43] A. Rosenfeld, Fuzzy groups, Journal of Mathematical Analysis and Applications. 35 (1971) 512-517.

[44] B. Šešelja, A. Tepavčević, $\Omega-$groups in the language of $\Omega-$groupoids, Fuzzy Sets Syst. 397 (2020) 152-167.

[45] Rainbow Clouds Crown Belukha Mountain, Siberia's Highest Peak, The Siberian Times, (Dec. 23rd, 2019) https://siberiantimes.com/other/others/news/rainbow-clouds-crown-belukha-mountain-siberias-highest-peak/

[46] Shakhatreh, Mohammad Shatanawi, Wasfi Jaradat, M. Fora, Ali. On fuzzy partitions. International Journal of Pure and Applied Mathematics. 30 (2006).

[47] H. Tahani, J.M. Keller, Information fusion in computer vision using the fuzzy integral, T-SMC. 20 (1990) 733-741.

[48] Z. Wang, G.J. Klir, Fuzzy Measure Theory, Plenum Press, New York u.a, 1992

[49] Xia, ZQ., Guo, FF. Fuzzy metric spaces. Journal of Applied Mathematics and Computing 16, (2004) 371–381

[50] Yeh, R.T., and Bang, S.Y. . Fuzzy relations, fuzzy graphs and their applications to clustering analysis. In Zadeh et al. (1975) 125-150.

[51] L.A. Zadeh, Fuzzy sets, Information and Control. 8 (1965) 338-353.

[52] L.A. Zadeh, Similarity relations and fuzzy orderings, Inf. Sci. 3 (1971) 177-200.

[53] H.-. Zimmermann, Fuzzy Set Theory — and its Applications, Second, Revised Edition. ed., Springer Netherlands, Dordrecht, (1991).