Czech Technical University in Prague
Faculty of Transportation Sciences

Department of Applied Mathematics
Branch of Study: Engineering Informatics in Transportation and Telecommunications
Study Programme: Engineering Informatics

Fuzzy Sets in Stochastic Modelling

Fuzzy množiny ve stochastickém modelování

DOCTORAL THESIS

Author:	Pavel Provinský
Supervisor:	doc. Ing. Ivan Nagy, CSc.
Year:	2021

Acknowledgements

<div align="right">Pavel Provinský</div>

Title:
Fuzzy Sets in Stochastic Modelling

Author:	Pavel Provinský
Branch of Study:	Engineering Informatics in Transportation and Telecommunications
Type of Document:	Doctoral Thesis
Supervisor:	doc. Ing. Ivan Nagy, CSc. Department of Applied Mathematics, Faculty of Transportation Sciences, Czech Technical University in Prague

Abstract: The text introduces floppy logic, which is a new multi-valued logic. Floppy logic consistently links fuzzy sets to probability theory.

The most important results of this work include proof that all statements equivalent in standard two-valued logic are also equivalent in floppy logic. It follows that floppy logic retains all the properties of standard two-valued logic which can be expressed as an equivalence.

Another important result is the proof that floppy logic is a model of Kolmogorov probability theory. We can therefore apply all the concepts and tools of probability theory in floppy logic.

Much focus was given to practical examples of work with floppy logic.

Floppy logic is compared to several other theories and also presented in historical context.

Keywords:	floppy logic, fuzzy logic, multi-valued logic, non-truth-functional logic

Název práce:

Fuzzy množiny ve stochastickém modelování

Autor: Pavel Provinský

Obor: Inženýrská informatika v dopravě a spojích

Druh práce: Dizertační práce

Vedoucí práce: doc. Ing. Ivan Nagy, CSc.
                  Ústav aplikované matematiky, Fakulta dopravní, České vysoké
                  učení technické v Praze

Abstrakt: Tento text zavádí floppy logiku, což je nová mnohohodnotová logika.
Floppy logika konzistentně propojuje fuzzy množiny s teorií pravděpodobnosti.
Mezi nejdůležitější výsledky této práce patří důkaz, že všechny výrazy ekvivalentní
ve standardní dvouhodnotové logice jsou ekvivalentní i ve floppy logice. Z toho
plyne, že floppy logika zachovává všechny vlastnosti standardní dvouhodnotové
logiky, které lze vyjádřit jako ekvivalence.
Dalším důležitým výsledkem je důkaz, že floppy logika je modelem Kolmogorovy
teorie pravděpodobnosti. Můžeme tedy ve floppy logice používat všechny pojmy a
nástroje teorie pravděpodobnosti.
Velká pozornost je věnována praktické práci s floppy logikou a příkladům.
Floppy logika je také porovnána s několika dalšími teoriemi a je zasazena do his-
torického kontextu.

Klíčová slova: floppy logika, fuzzy logika, mnohohodnotová logika, nefunkcionální
                   logika

Contents

Introduction

The primary aim of this work was to create a theory that consistently linked fuzzy sets to probability theory. The theory had to work with the phenomena described by fuzzy sets, probability distributions and exact numbers. It had to apply the tools of probability theory to all these phenomena. The whole required mathematical consistency.

The new multi-value logic described in this text is named *floppy logic*.

During the work, floppy logic revealed itself to be not only the bridge searched between fuzzy sets and probability theory but it was also fascinating in itself.

One of the beautiful and remarkable results of this theory is that floppy logic retains all the properties of standard two-valued logic which can be formulated as equivalences. Floppy logic, therefore, preserves distributivity, the law of contradiction, the law of excluded middle, De Morgan laws, the law of contraposition, and so on. Although floppy logic is similar to fuzzy logic, no fuzzy logic possesses this property. The fact that such a multi-valued generalisation of standard two-valued logic is possible at all is certainly a great surprise.

Another beautiful feature of floppy logic is that it is a model of Kolmogorov probability theory. It is therefore possible to use the full wealth of relationships and knowledge of standard probability theory within floppy logic. For example, it is feasible to define the mean value of a floppy set, use Bayes' theorem or apply the law of total probability.

Floppy logic also retains the simplicity and practicality of fuzzy logic. This will surely be appreciated by anyone who wants to apply this new logic in solving specific problems.

The work is divided into five chapters.

The first chapter is "Floppy Logic in a Historical Context". I mention here some important notions in logic and probability formed in the twentieth century. These ideas are immediately associated with floppy logic and demonstrate whether and to

what extent they contributed to the creation of floppy logic. This chapter is very brief yet contains many references to other literature.

The second chapter "Mathematical Foundations of Floppy Logic" is the mathematical core of this work. It introduces the basic concepts and relationships of floppy logic. The most important results of this chapter are three theorems: the first two link floppy logic with probability theory and the third links with Boolean logic. The proofs of these theorems are given in the appendices.

The third chapter "Other Interesting Mathematical Outcomes" describes three other interesting topics. It first introduces the mean value of a floppy set as an example of the link between fuzzy sets and probability theory.

The chapter then provides many interesting results concerning floppy implication and introduces several generalisations of the inference rules modus ponens and modus tollens.

Finally, the chapter describes an interesting relationship between the floppy membership function of a floppy set and t-norms and t-conorms used in fuzzy logic. From this relationship, it is possible to derive the quantitative measurement of the dependence of two statements.

The fourth chapter "Using Floppy Logic to Describe a System" examines in detail everything required to describe and control a system using floppy logic. It starts with the selection of suitable primary fuzzy sets, continues with the fuzzification of input data, implementation of system rules and defuzzification of output data, and ends with optimal control of the system.

The fifth chapter is "Comparison of Floppy Logic to Other Theories". The chapter compares floppy logic with fuzzy logic and Adams' and Stalnaker's probability logic. For example, comparison of floppy implication with Adams' Thesis is a fascinating study.

The basis of this work is articles [51, 52, 53]. However, the text has been reworked and greatly expanded.

I am confident that discussion of this remarkable theory will not only be instructive but also a pleasant experience!

Chapter 1

Floppy Logic in a Historical Context

In the first chapter, I attempt to describe the basic ideas which contributed to or hindered the formulation of floppy logic. In floppy logic, several streams of thought combine. One of them is certainly multi-valued logic.

1.1 Multi-Valued Logic

The first article discussing multi-valued logic, written by Jan Łukasiewicz, appeared in 1920 [42]. The author wrote [41] that he was inspired to create a three-valued logic according to Aristotle's reasoning in the ninth volume of Hermeneutics, where a future of naval battle is considered. Before this uncertain future event takes place, one cannot consider statements about it to be true or false.

Łukasiewicz assigned a logical value of $\frac{1}{2}$ (or 2 in the original version) to such uncertain statements and supplemented the logical table with rules for the third value.

Łukasiewicz attached great importance to the creation of this new logic [41]: "Each such logic can be the basis of slightly different mathematics, and each such mathematics can be the basis of slightly different physics."

1.2 Truth-Functionality

When Łukasiewicz discussed logic, he assumed that multi-valued logic must be truth-functional, just like classical logic [41]. This means that the truth-value of a composite proposition depends only on the truth-values of its atom propositions.

This assumption precludes a link between logic and probability theory which is not truth-functional.

Floppy logic is not truth-functional either.

1.3 Other Multi-Valued Logic

Other three-valued systems of logic were independently introduced by Dmitry Anatolievich Bochvar and Stephen Cole Kleene in 1938 and in 1945 [7, 33].

Many types of multi-valued logic are now known, such as four-valued logic [5] or 16-valued logic [59].

Let us focus on that logic which has truth values from the interval [0, 1]. The best-known types of logic are probabilistic logic and fuzzy logic.

1.4 Probabilistic Logic

Probabilistic logic is not a unified theory, rather more of a school of thought which assumes that probability can be considered a generalisation of a truth value.

The notion that a close link exists between probability and logic can be found in the works of many thinkers of the nineteenth and twentieth centuries. We may cite, for example, Augustus De Morgan [14], George Boole [8], Frank Plumpton Ramsey [54], and Rudolf Carnap [13].

However, Łukasiewicz wrote [41]: "Previous attempts to combine multi-valued logic with probability have encountered great difficulties."

Floppy logic also interprets probability or conditional probability as a possible generalisation of a truth value.

1.5 Adams' and Stalnaker's Logic

Well-known works on probabilistic logic were written by Robert C. Stalnaker in 1970 [60] and Ernest Wilcox Adams in 1975 [2].

In these theories, probability is a generalisation of the truth value. The main hypothesis behind these theories is *Adams' Thesis*, which states that the probability of implication $P(A \rightarrow B)$ is equal to relevant conditional probability $P(B|A)$.

David Lewis showed [40] that this probability logic has one major problem: it cannot be used to work with sentences wherein an implication contains some other implication. For example, the sentence $(A \wedge (A \Rightarrow B)) \Rightarrow B$ departs from the possibilities of this type of probabilistic logic. This outcome is known as the *Lewis' triviality result*.

Floppy logic also obtains an exact relationship between $P(A \rightarrow B)$ and $P(B|A)$, although this is different from Adams' thesis. In floppy logic, Lewis' triviality result does not apply.

1.6 Fuzzy Sets

The first article exploring fuzzy sets was written by Lotfali Askar Zadeh in 1965 [68]. In that work, the author introduced fuzzy sets and their membership functions. The membership function of a fuzzy set can obtain any value from the interval [0,1]. Therefore, an element can belong to a fuzzy set, for example, only to 20%.

Floppy logic adopts the idea of fuzzy sets and membership functions and applies them.

1.7 Operations with Fuzzy Sets

Fuzzy logic generalises the intersection and union of fuzzy sets in several ways.

In his first work on fuzzy logic [68], Zadeh already suggested two methods of operating with fuzzy sets. From these two methods, Gödel and product fuzzy logics were devised.

The following years saw the invention of Łukasiewicz[1], drastic, and many other types of fuzzy logic.

However, floppy logic does not use fuzzy operations. Instead, it uses an ordinary set intersection and union.

---

[1]This logic, named after a Polish mathematician, is a different type of logic to the one published in 1920 by Jan Łukasiewicz. This logic introduced Robin Giles in a paper [24] in 1976.

1.8   Problems in Fuzzy Logic

Although fuzzy logic has become very popular and the most widely used multi-value logic, it has certain problems.

The different types of fuzzy logic provide different results. It is not clear which fuzzy logic should be used in certain situations.

The second problem is that no fuzzy logic satisfies all the laws of standard two-valued logic. Proof of this can be found, for example, in [47]. Details on which laws are satisfied in the most commonly applied types of fuzzy logic are given in Table 1.1.

Table 1.1:   Logical laws in the most commonly used types of fuzzy logic. G = Gödel logic, Ł = Łukasiewicz logic, P = product logic, all with standard negation. Data are drawn from [47].

		G	Ł	P
involution	$\neg(\neg A) = A$	OK	OK	OK
commutativity	$A \vee B = B \vee A$	OK	OK	OK
	$A \wedge B = B \wedge A$			
associativity	$(A \vee B) \vee C = A \vee (B \vee C)$	OK	OK	OK
	$(A \wedge B) \wedge C = A \wedge (B \wedge C)$			
distributivity	$A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$	OK	$\times$	$\times$
	$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$			
idempotence	$A \vee A = A$	OK	$\times$	$\times$
	$A \wedge A = A$			
absorption	$A \vee (A \wedge B) = A$	OK	$\times$	$\times$
	$A \wedge (A \vee B) = A$			
absorption with 1 and 0	$A \vee 1 = 1$	OK	OK	OK
	$A \wedge 0 = 0$			
neutral elements	$A \vee 0 = A$	OK	OK	OK
	$A \wedge 1 = A$			
law of contradiction	$A \wedge \neg A = 0$	$\times$	OK	$\times$
law of excluded middle	$A \vee \neg A = 1$	$\times$	OK	$\times$
De Morgan's laws	$\neg(A \vee B) = \neg A \wedge \neg B$	OK	OK	OK
	$\neg(A \wedge B) = \neg A \vee \neg B$			

Floppy logic satisfies all the properties listed in Table 1.1 and all properties of standard Boolean logic which can be formulated as equivalences.

1.9 Probability Theory

Another stream of thinking significant to floppy logic is the evolution of probability theory.

Probability theory emerged in the mid-seventeenth century and is associated with the names of Blaise Pascal, Pierre de Fermat and Christian Huygens.

In 1933, Andrey Nikolajevic Kolmogorov performed an axiomatisation of the theory [35]. This axiomatisation is very important to us.

First, floppy logic is a model of the Kolmogorov probability theory.

Second, in this axiomatisation, Kolmogorov attributed probability not to the individual possible results of a random experiment but their sets.

Applied to fuzzy sets, this step is the main difference between floppy logic and fuzzy logic.

We can also find other axiomatisations of probability theory. The axiomatisation performed by Alfred Renyi [57] is compelling since the main concept in his theory is not probability but conditional probability.

This is similar to floppy logic, where probability is also derived from conditional probability.

1.10 Attempts to Link Probability Theory and Fuzzy Logic

The notion to link probability theory and fuzzy set theory is not new.

Both approaches can be combined in many ways. For example, Lotfi Zadeh introduced classical probability of fuzzy phenomena [70].

In another article, the same author examined the probability of fuzzy events not given by real numbers but fuzzy numbers [67].

In paper [48], the authors discuss the conditional probability of fuzzy objects.

An interesting relationship can be found in an article [21] by Brian R. Gaines, who introduces "uncertainty logic". After adding the law of excluded middle, this concept leads to Rescher's probabilistic logic, which is introduced in the book [55]. Conversely, an addition to the requirement of functionality in logic leads to Łukasiewicz's

fuzzy logic. The addition of a combination of both requirements leads to contradiction.

Another relationship between probability theory and fuzzy sets can be found in the similar properties of conditional probability and fuzzy relative cardinality. This similarity was highlighted by Bart Kosko [37].

Possibility theory, which falls between probability theory and the theory of fuzzy sets, was developed by Lotfi Zadeh [69], Didier Dubois and Henri M. Prade [17] and is important to mention.

I would also like to mention one more work concerning the relationship between probability theory and the theory of fuzzy sets. The authors in the paper [45] studied the question of which t-norms and t-conorms, in conjunction with Zadeh's definition of probability of fuzzy events [70], satisfied Kolmogorov's axioms.

A very similar question arose in the formulation of floppy logic.

Despite these attempts, unifying probability theory and fuzzy set theory into a single consistent theory failed.

The main reasons of the incompatibility between fuzzy sets and probability were as follows:

(1) Fuzzy logic and probability theory work with different forms of uncertainty [31].

(2) Fuzzy logic is truth-functional whereas probability theory is not [31].

However, we can find a non-truth-functional fuzzy logic system combined with probability [23].

Floppy logic is not truth-functional and operates with the probabilities of floppy sets whereas a floppy set is a crisp set of fuzzy sets.

1.11   Floppy Logic

In September 2014, Ivan Nagy, my colleague, posed an interesting question of whether it would be possible somehow to link statistics and fuzzy logic consistently. It would be very useful. Frequently, more complex systems described with fuzzy logic simultaneously require statistical data on this system to be processed.

Ivan Nagy immediately suggested a solution: if a structure which satisfied all the Kolmogorov Axioms of probability theory was found in the world of fuzzy sets, then all probability tools could be consistently applied when working with fuzzy sets. This idea was refreshing, and after two weeks, the first results were obtained.

The first article on this new logic was published in 2017 [51]. The paper presented a theoretical foundation, and the new logic was called *floppy logic*.

Published in 2018, the second article [52] provided many examples of working with floppy logic.

The third paper [53] demonstrated that floppy logic could be considered a multi-valued generalisation of standard two-valued logic since it preserves all the properties of Boolean logic which can be formulated as equivalences. It also contains some compelling results on floppy implication.

Chapter 2

Mathematical Foundations of Floppy Logic

2.1 Basic Floppy Logic

2.1.1 Assumptions and Definitions

Let A_1, A_2, ... be fuzzy sets whose membership functions are defined in the same domain. We will call these sets *primary fuzzy sets*.

Let $\mu_{A_1}(x)$, $\mu_{A_2}(x)$, ... be the membership functions of primary fuzzy sets A_1, A_2, ...

Let X be the domain of the membership functions $\mu_{A_1}(x)$, $\mu_{A_2}(x)$, ...

Let $\boldsymbol{S}$ be the set of all primary fuzzy sets A_1, A_2, ...

Let $\mathcal{P}(\boldsymbol{S})$ be the power set of $\boldsymbol{S}$.

Let $\boldsymbol{S_x}$ be a set of all primary fuzzy sets A_i whose membership functions are greater than 0 for given x.

Let the following assumptions be satisfied:

Assumption 2.1.1. For all $x \in X$, $\boldsymbol{S_x}$ is a finite or countable set.

Assumption 2.1.2. The membership functions of fuzzy sets $A_i \in \boldsymbol{S}$ assume values from the interval $[0, 1]$.

Assumption 2.1.3.

$$\forall x \in X : \sum_{A_i \in \boldsymbol{S}} \mu_{A_i}(x) = 1. \tag{2.1}$$

Assumption 2.1.4. A measure space $(X, \mathcal{A}, P)$ is defined in the set X, where $\mathcal{A}$ is a σ-algebra on X, and P is a probability measure.

Assumption 2.1.5. All membership functions $\mu_{A_i}(x)$ of fuzzy sets $A_i \in \boldsymbol{S}$ are measurable in the sets $X_i \in \mathcal{A}$ corresponding to measure P.

Definition 2.1.1. All subsets of $\boldsymbol{S}$ are called *floppy sets*. Floppy sets are denoted in bold capital letters.

Definition 2.1.2. Each floppy set $\boldsymbol{B} \subseteq \boldsymbol{S}$ is associated with a function $\boldsymbol{\mu_B}(x)$. The function $\boldsymbol{\mu_B}(x)$ is defined as:

$$\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x). \tag{2.2}$$

The function $\boldsymbol{\mu_B}(x)$ is called the *floppy membership function of the floppy set* and is denoted by the bold Greek letter $\boldsymbol{\mu}$.

Definition 2.1.3. Each floppy set $\boldsymbol{B} \subseteq \boldsymbol{S}$ is associated with a number $R(\boldsymbol{B})$. The number $R(\boldsymbol{B})$ is defined as:

$$R(\boldsymbol{B}) = \int_X \boldsymbol{\mu_B}(x)\, \mathrm{d}P, \tag{2.3}$$

where the integral is a Lebesgue integral and P is the probability measure given in Assumption 2.1.4. The number $R(\boldsymbol{B})$ is called the *probability of the floppy set* and is denoted by the capital letter R.

The following theorem shows that a number selected in this way genuinely has probability properties.

2.1.2 Theorem of Basic Floppy Probability Space

Definition 2.1.4. The space $(\boldsymbol{S}, \mathcal{P}(\boldsymbol{S}), R)$ is called a *basic floppy probability space.*

In floppy logic, the following theorem applies:

Theorem 2.1.1. Each basic floppy probability space satisfies all Kolmogorov Axioms.

Proof of this theorem is given in Appendix A.

2.1.3 Remarks

Remark 2.1.1. By proving Theorem 2.1.1, the objective set by Ivan Nagy is achieved. A structure which satisfies all Kolmogorov Axioms is found in the world of fuzzy sets. All standard probabilistic tools and concepts in floppy logic can therefore be used. For example, Bayes' theorem can be applied, or the mean value or median of the floppy set can be introduced consistently.

Remark 2.1.2. The result of Theorem 2.1.1 could be thought of as trivial since the existence of a probability space is assumed in Assumption 2.1.4 and the existence of a probability space is proved in Theorem 2.1.1. This is not true because $(X, \mathcal{A}, P)$ and $(\boldsymbol{S}, \mathcal{P}(\boldsymbol{S}), R)$ are completely different probability spaces. The probability measures P and R were distinguished for this reason.

Remark 2.1.3. The floppy set is crisp set of the primary fuzzy sets. Its floppy membership function is the sum of membership functions of its elements.

Remark 2.1.4. In articles [51] and [52], the function $\boldsymbol{\mu_B}(x)$ was called "the membership function of the floppy set". This denomination was selected to indicate that this function has a very similar role in floppy logic to the membership function of the fuzzy set in fuzzy logic. However, this denomination led to misunderstandings, and it was therefore changed to "the floppy membership function of the floppy set".

Remark 2.1.5. The intersection of two different single-element floppy sets is equivalent to an empty set. When a certain reality is modelled, the primary fuzzy sets which respect this property must therefore be selected.

For example, the temperature of water can be described with the following three primary fuzzy sets: unpleasantly cold, pleasant, and unpleasantly warm. It would never be suggested, for example, that water is unpleasantly cold and pleasant simultaneously. Therefore, it is acceptable that the intersection of these two properties (single-element floppy set) is an empty set.

However, the following primary fuzzy sets cannot be used: cold, pleasant, and warm, since water can be, for example, warm and pleasant simultaneously.

Remark 2.1.6. In articles [51] and [52], Assumption 2.1.1 was stronger: "$\boldsymbol{S}$ is a finite or countable set." However, this stronger version is not necessary for proving Theorem 2.1.1, therefore this assumption was changed.

Remark 2.1.7. Floppy logic is not truth-functional. This means that $\boldsymbol{\mu_{A \cap B}}(x)$ cannot be computed from $\boldsymbol{\mu_A}(x)$ and $\boldsymbol{\mu_B}(x)$. The common elements (primary fuzzy sets) of floppy sets $\boldsymbol{A}$ and $\boldsymbol{B}$ must be known.

Similarly, in probability theory, the probability of an event $A \cap B$ cannot be calculated from the probabilities of events A and B.

However, standard binary logic and fuzzy logic are truth-functional.

Remark 2.1.8. Many examples of working with floppy sets in practice are presented in Chapter 4.

2.2 Generalised Floppy Logic

2.2.1 Definitions

We have two probabilistic spaces. The space $(X, \mathcal{A}, P)$ was assumed whereas the space $(\boldsymbol{S}, \mathcal{P}(\boldsymbol{S}), R)$ was defined. Let us create a joint probability space.

A sample space, sufficiently σ-algebra rich, and a probabilistic measure are needed. Let us assume satisfaction of Assumptions 2.1.1 to 2.1.5. The sample space of this joint probability space is the Cartesian product $\boldsymbol{S} \times X$.

Definition 2.2.1. The smallest σ-algebra over $\mathcal{P}(\boldsymbol{S}) \times \mathcal{A}$ is denoted $\mathcal{C}$.

Definition 2.2.2. Elements of $\mathcal{C}$ are called *generalised floppy sets*.

The generalised floppy sets are denoted by bold capital letters with an upper index G.

Definition 2.2.3. The *floppy membership function of a generalised floppy set* $\boldsymbol{C^G} \in \mathcal{C}$ is defined by the rule:

$$\boldsymbol{\mu_{C^G}}(x) = \sum_{A_i \in \boldsymbol{S}:\,[A_i, x] \in \boldsymbol{C^G}} \mu_{A_i}(x). \tag{2.4}$$

The floppy membership function of floppy sets is denoted by the bold letter $\boldsymbol{\mu}$.

Definition 2.2.4. The *probability R of generalised floppy set* $\boldsymbol{C^G} \in \mathcal{C}$ is defined by the rule:

$$R\left(\boldsymbol{C^G}\right) = \int_X \boldsymbol{\mu_{C^G}}(x)\,\mathrm{d}P. \tag{2.5}$$

The floppy membership function of the generalised floppy set is a generalisation of the floppy membership function of the floppy set because:

$$\boldsymbol{\mu_{B \times X}}(x) = \sum_{A_i \in \boldsymbol{S}:\,[A_i, x] \in \boldsymbol{B} \times X} \mu_{A_i}(x) = \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x) = \boldsymbol{\mu_B}(x). \tag{2.6}$$

Similarly, the probability of the generalised floppy set is a generalisation of the probability of the floppy set because:

$$R\left(\boldsymbol{B} \times X\right) = \int_X \boldsymbol{\mu_{B \times X}}(x)\,\mathrm{d}P = \int_X \boldsymbol{\mu_B}(x)\,\mathrm{d}P = R\left(\boldsymbol{B}\right). \tag{2.7}$$

The probabilistic measure R for generalised floppy sets is a generalisation of the probabilistic measure P:

Let Y be a subset of σ-algebra $\mathcal{A}$ assumed in Assumption 2.1.4. Therefore:

$$R\left(\boldsymbol{S} \times Y\right) = \int_X \sum_{A_i \in \boldsymbol{S}:\,[A_i, x] \in \boldsymbol{S} \times Y} \mu_{A_i}(x)\,\mathrm{d}P = \int_Y 1\,\mathrm{d}P = P\left(Y\right). \tag{2.8}$$

2.2.2 Theorem of Generalised Floppy Probability Space

Definition 2.2.5. The space $(\boldsymbol{S} \times X, \mathcal{C}, R)$ is called *a generalised floppy probability space*.

The names "probability of generalised floppy set" in the Definition 2.2.4 and "generalised floppy probability space" in the Definition 2.2.5 are justified in the following theorem:

29

Theorem 2.2.1. Each generalised floppy probability space satisfies all Kolmogorov axioms.

The proof of Theorem 2.2.1 is given in Appendix B.

Generalised and basic floppy sets can thus be seen as probabilistic events, and it is possible to work with them in this manner.

Consequence 2.2.1. Theorem 2.2.1 allows us to work with $R\left(\boldsymbol{C^G}|x\right)$ as with a conditional probability. For discrete cases, we can write:

$$R\left(\boldsymbol{C^G}\right) = \sum_{i=1}^{n} R\left(\boldsymbol{C^G}|x_i\right) \cdot P\left(x_i\right). \qquad (2.9)$$

For continuous cases, we can can write:

$$R\left(\boldsymbol{C^G}\right) = \int_X R\left(\boldsymbol{C^G}|x\right) \cdot f\left(x\right) \mathrm{d}x, \qquad (2.10)$$

where $f\left(x\right)$ is a probability density function and the integral is a Riemann integral. Generally, we can write:

$$R\left(\boldsymbol{C^G}\right) = \int_X R\left(\boldsymbol{C^G}|x\right) \mathrm{d}P, \qquad (2.11)$$

where the integral is a Lebesgue integral.

This equation can be compared to Definition 2.2.4:

$$R\left(\boldsymbol{C^G}\right) = \int_X \boldsymbol{\mu_{C^G}}\left(x\right) \mathrm{d}P. \qquad (2.12)$$

The equation

$$\int_X \boldsymbol{\mu_{C^G}}\left(x\right) \mathrm{d}P = \int_X R\left(\boldsymbol{C^G}|x\right) \mathrm{d}P \qquad (2.13)$$

applies to all possible sets $\boldsymbol{C^G}$. It also applies to those $\boldsymbol{C^G}$ whose membership function has non-zero values only in an arbitrarily narrow interval from X.

Therefore, the equation

$$\boldsymbol{\mu_{C^G}}\left(x\right) = R\left(\boldsymbol{C^G}|x\right) \qquad (2.14)$$

must apply for all x, except for a null set. In this work, I assume that it applies for all $x \in X$.

This important relationship between conditional probability $R\left(\boldsymbol{C^G}|x\right)$ and the floppy membership function $\boldsymbol{\mu}_{\boldsymbol{C^G}}(x)$ can be interpreted in the following manner: the floppy membership function $\boldsymbol{\mu}_{\boldsymbol{C^G}}(x)$ can be calculated as the conditional probability of event $\boldsymbol{C^G}$ given an exact x.

Example 2.2.1. 60% of experts decide that water is hot when the temperature reaches 50 °C. The floppy membership function of floppy set "hot" is therefore 0.6 for a temperature of 50 °C.

2.3 Isomorphism Theorem

2.3.1 Formulation

I now introduce a very important theorem in floppy logic. A similar statement appeared in [51], although no detailed mathematical proof has yet been provided.

Theorem 2.3.1. Let $\varepsilon = (E, \wedge, \vee, \neg, \bot, \top, \equiv)$ be a finite Boolean algebra of sentences.

Let $E_A = \{U_1, U_2, U_3, \ldots U_n\}$ be a set of all atoms of this Boolean algebra and n number of these atoms.

Let $\boldsymbol{S} = \{A_1, A_2, A_3, \ldots A_n\}$ be a set of n primary fuzzy sets and X be their domain.

Let $\delta = (\mathcal{P}(\boldsymbol{S}), \cap, \cup, ', \emptyset, \boldsymbol{S}, =)$ be a Boolean algebra of subsets of $\boldsymbol{S}$.

Let f be a relation which satisfies the following:

a.	$\forall k \in \{1, 2, 3, \ldots n\} : f(U_k) = \{A_k\},$	(2.15)
b.	$\forall V, W \in E : f(V \wedge W) = f(V) \cap f(W),$	(2.16)
c.	$\forall V, W \in E : f(V \vee W) = f(V) \cup f(W),$	(2.17)
d.	$\forall V \in E : f(\neg V) = f(V)'.$	(2.18)

Then:

1.	f is an isomorphism,	
2.	ε and δ are isomorphic,	
3.	$\forall V, W \in E : V \equiv W \Leftrightarrow f(V) = f(W),$	(2.19)
4.	$\forall V, W \in E, \forall x \in X : V \equiv W \Rightarrow \boldsymbol{\mu}_{f(V)}(x) = \boldsymbol{\mu}_{f(W)}(x),$	(2.20)
5.	$\forall V, W \in E : V \equiv W \Rightarrow R(f(V)) = R(f(W)).$	(2.21)

The terms Boolean algebra, atom, and isomorphism are explained in Appendix C. Also, Theorem 2.3.1 is proved in Appendix C.

2.3.2   Consequences

Consequence 2.3.1. Theorem 2.3.1 states that sentences which are equivalent in standard Boolean logic (consisting of a finite set of atomic statements) are assigned the same floppy set, and hence, the same floppy membership function and thus the same probability. This means that statements which are equivalent in standard Boolean logic are also equivalent in floppy logic.

Consequence 2.3.2. This means that floppy logic satisfies all the standard binary logic properties which are formulated as an equivalence of two sentences. All the properties listed in Table 1.1 are thus especially satisfied.

Consequence 2.3.3. Floppy logic can hence be considered a generalisation of standard Boolean logic.

Consequence 2.3.4. Standard Boolean logic often models statements using sets such as Venn diagrams. In view of Theorem 2.3.1, these sets can be interpreted as floppy sets.

Consequence 2.3.5. Theorem 2.3.1 states that the Boolean algebras ε and δ are isomorphic. If it is appropriate, the logical and set operations can therefore be used in combination or exchanged.

Consequence 2.3.6. The probability of implication, for example, can be expressed as:

$$R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) = R\left(\neg \boldsymbol{A} \vee \left(\boldsymbol{A} \wedge \boldsymbol{B}\right)\right) = R\left(\boldsymbol{A}' \cup \left(\boldsymbol{A} \cap \boldsymbol{B}\right)\right). \tag{2.22}$$

$\boldsymbol{A}'$ and $\boldsymbol{A} \cap \boldsymbol{B}$ are disjoint sets, therefore:

$$R\left(\boldsymbol{A}' \cup \left(\boldsymbol{A} \cap \boldsymbol{B}\right)\right) = R\left(\boldsymbol{A}'\right) + R\left(\boldsymbol{A} \cap \boldsymbol{B}\right). \tag{2.23}$$

Theorem 2.1.1 declares that R satisfies all the properties of probability. Standard equations for $R\left(\boldsymbol{A} \cap \boldsymbol{B}\right)$ and $R\left(\boldsymbol{A}'\right)$ can therefore be used, obtaining:

$$R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) = 1 - R\left(\boldsymbol{A}\right) + R\left(\boldsymbol{B}|\boldsymbol{A}\right) \cdot R\left(\boldsymbol{A}\right). \tag{2.24}$$

It was found that Adams' Thesis in floppy logic does not apply.

Consequence 2.3.7. Similarly, transitivity of implication can be expressed as:

$$R\left[\left[(\boldsymbol{A} \Rightarrow \boldsymbol{B}) \wedge (\boldsymbol{B} \Rightarrow \boldsymbol{C})\right] \Rightarrow \left[\boldsymbol{A} \Rightarrow \boldsymbol{C}\right]\right] = 1. \tag{2.25}$$

Floppy logic is not limited by Lewis' triviality result.

Consequence 2.3.8. In floppy logic, both probabilities (as in probabilistic logic) and floppy membership functions (as in fuzzy logic) can be worked with. Sentences equivalent in standard binary logic possess the same floppy membership functions and probabilities. It is very surprising that both of these different generalisations of the truth-value encounter each other in a single theory.

As shown in Section 2.2.2, the relationship between probability and the floppy membership function is, in floppy logic, given by the equation:

$$\boldsymbol{\mu_A}(x) = R(\boldsymbol{A}|x). \tag{2.26}$$

Consequence 2.3.9. Let us accept that

$$\forall x \in Y : \boldsymbol{\mu_A}(x) = 1 \qquad \text{is the same as} \qquad \forall x \in Y : A(x), \tag{2.27}$$

$$\exists x \in Y : \boldsymbol{\mu_A}(x) = 1 \qquad \text{is the same as} \qquad \exists x \in Y : A(x), \tag{2.28}$$

then floppy logic is also a generalisation of predicate logic.

However, care must be taken. The notation which uses the membership functions of floppy sets is more general than the notation which uses $\forall$ and $\exists$ quantifiers. Therefore, for example, the basic relationship of predicate logic

$$\neg\left(\forall x \in Y : A(x)\right) \equiv \exists x \in Y : \neg\left(A(x)\right) \tag{2.29}$$

does not apply in floppy logic.

Chapter 3

Other Interesting Mathematical Outcomes

3.1 Mean Value of Floppy Sets

Theorems 2.1.1 and 2.2.1 guarantee that we can introduce probabilistic concepts for floppy sets. Let us demonstrate this for the mean value of a floppy set.

Definition 3.1.1. Let us define the mean value of a floppy set as follows:

$$\langle \boldsymbol{B} \rangle = \mathrm{E}\left(x | \boldsymbol{B}\right). \tag{3.1}$$

For discrete cases, the mean value of a floppy set can be expressed as:

$$
\begin{aligned}
\langle \boldsymbol{B} \rangle = \mathrm{E}\left(x_i | \boldsymbol{B}\right) &= \sum_{x_i \in X} x_i \cdot R\left(x_i | \boldsymbol{B}\right) = \\
&= \frac{1}{R\left(\boldsymbol{B}\right)} \sum_{x_i \in X} x_i \cdot R\left(\boldsymbol{B} | x_i\right) \cdot R\left(x_i\right) = \\
&= \frac{\sum_{x_i \in X} x_i \cdot \boldsymbol{\mu_B}\left(x_i\right) \cdot P\left(x_i\right)}{\sum_{x_i \in X} \boldsymbol{\mu_B}\left(x_i\right) \cdot P\left(x_i\right)} = \\
&= \frac{\int_X x_i \cdot \boldsymbol{\mu_B}\left(x_i\right) \mathrm{d}P}{\int_X \boldsymbol{\mu_B}\left(x_i\right) \mathrm{d}P}
\end{aligned}
\tag{3.2}
$$

where the integrals are Lebesgue integrals.

35

For continuous cases, we can write:

$$
\begin{aligned}
\langle \boldsymbol{B} \rangle = \mathrm{E}\left(x|\boldsymbol{B}\right) &= \int_X x \cdot f\left(x|\boldsymbol{B}\right) \mathrm{d}x = \\
&= \frac{1}{R\left(\boldsymbol{B}\right)} \int_X x \cdot R\left(\boldsymbol{B}|x\right) \cdot f\left(x\right) \mathrm{d}x = \\
&= \frac{\int_X x \cdot \boldsymbol{\mu_B}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}{\int_X \boldsymbol{\mu_B}\left(x\right) \cdot f\left(x\right) \mathrm{d}x} = \\
&= \frac{\int_X x \cdot \boldsymbol{\mu_B}\left(x\right) \mathrm{d}P}{\int_X \boldsymbol{\mu_B}\left(x\right) \mathrm{d}P}
\end{aligned}
\tag{3.3}
$$

where the integrals on the final row are Lebesgue integrals and the other integrals are Riemann integrals.

Let suppose that $\{\boldsymbol{B_1}, \boldsymbol{B_2}, \dots \boldsymbol{B_n}\}$ is a finite set of pairwise disjoint floppy sets whose union is whole set $\boldsymbol{S}$.

We can then write the following interesting relationship:

$$
\begin{aligned}
\mathrm{E}\left(\langle \boldsymbol{B_i} \rangle\right) &= \sum_{i=1}^{n} \langle \boldsymbol{B_i} \rangle \cdot R\left(\boldsymbol{B_i}\right) = \\
&= \sum_{i=1}^{n} \frac{1}{R\left(\boldsymbol{B_i}\right)} \cdot \int_X x \cdot R\left(\boldsymbol{B_i}|x\right) \mathrm{d}P \cdot R\left(\boldsymbol{B_i}\right) = \\
&= \int_X x \cdot \left(\sum_{i=1}^{n} R\left(\boldsymbol{B_i}|x\right)\right) \mathrm{d}P = \\
&= \int_X x \cdot 1 \cdot \mathrm{d}P.
\end{aligned}
\tag{3.4}
$$

Therefore:

$$
\mathrm{E}\left(\langle \boldsymbol{B_i} \rangle\right) = \mathrm{E}\left(x\right).
\tag{3.5}
$$

This means that to calculate or estimate the mean value of a random variable, we only need to know the mean values of the floppy sets and their probabilities. Knowledge of the probability distribution of this random variable is not required.

Example 3.1.1. The performance of a football team is described by a random variable. The domain of this random variable is a set of real numbers. We do not know the course of the probability density.

The random variable is described by three floppy sets: "win", "draw", "lose". Their courses are also unknown. Suppose that three, one, and zero points for a win, draw, and loss are the mean values of the respective floppy sets.

Our team won 18 times, drew 5 times, and lost 12 times. We can estimate the mean value of its performance as

$$
\begin{aligned}
\mathrm{E}\left(x\right) \;=\; & \sum_{i=1}^{n} \langle \boldsymbol{B_i} \rangle \cdot R\left(\boldsymbol{B_i}\right) = \\
=\; & 3 \cdot \frac{18}{35} + 1 \cdot \frac{5}{35} + 0 \cdot \frac{12}{35} = 1.69.
\end{aligned} \tag{3.6}
$$

It is a fascinating idea that the points allocated to individual variants can be understood as the mean values of relevant floppy sets.

3.2 A Noteworthy Implication

One of the important consequences of the isomorphism theorem is the equation for floppy implication (see Consequence 2.3.6):

$$
R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) = 1 - R\left(\boldsymbol{A}\right) + R\left(\boldsymbol{B}|\boldsymbol{A}\right) \cdot R\left(\boldsymbol{A}\right). \tag{3.7}
$$

This equation, along with the law of total probability and Bayes' theorem, allows us to infer a range of interesting relationships for floppy implications. These relationships not only hold for floppy logic but also standard Boolean logic, assuming that we substitute the probabilities of sentences (floppy sets) with truth values 0 or 1.

3.2.1 Bayes' Theorem

$$
\begin{aligned}
R\left(\boldsymbol{B}|\boldsymbol{A}\right) \;=\;& \frac{R\left(\boldsymbol{A}|\boldsymbol{B}\right) \cdot R\left(\boldsymbol{B}\right)}{R\left(\boldsymbol{A}\right)}, & (3.8)\\
R\left(\boldsymbol{B}|\boldsymbol{A}\right) \cdot R\left(\boldsymbol{A}\right) \;=\;& R\left(\boldsymbol{A}|\boldsymbol{B}\right) \cdot R\left(\boldsymbol{B}\right), & (3.9)\\
R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) - 1 + R\left(\boldsymbol{A}\right) \;=\;& R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right) - 1 + R\left(\boldsymbol{B}\right), & (3.10)\\
R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right) - R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) \;=\;& R\left(\boldsymbol{A}\right) - R\left(\boldsymbol{B}\right), & (3.11)\\
R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) \;=\;& R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right) + R\left(\boldsymbol{B}\right) - R\left(\boldsymbol{A}\right). & (3.12)
\end{aligned}
$$

The following is a variant of the floppy membership function:

$$
\boldsymbol{\mu_{A \Rightarrow B}}\left(x\right) = \boldsymbol{\mu_{B \Rightarrow A}}\left(x\right) + \boldsymbol{\mu_B}\left(x\right) - \boldsymbol{\mu_A}\left(x\right). \tag{3.13}
$$

3.2.2 Contraposition

A contraposition of implication is a logical law which states that implication $A \Rightarrow B$ is logically equivalent to implication $\neg B \Rightarrow \neg A$.

In floppy logic, the law of contraposition follows directly from Theorem 2.3.1. The variants for probabilities and floppy membership functions are provided below:

$$R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) \;=\; R\left(\neg \boldsymbol{B} \Rightarrow \neg \boldsymbol{A}\right), \tag{3.14}$$

$$\boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \boldsymbol{B}}\left(x\right) \;=\; \boldsymbol{\mu}_{\neg \boldsymbol{B} \Rightarrow \neg \boldsymbol{A}}\left(x\right). \tag{3.15}$$

3.2.3   Modus Ponens

Modus ponens is a logical rule which states: If A holds and $A \Rightarrow B$ holds, then B holds.

A generalisation of modus ponens in floppy logic might appear as follows: We know the probability of $\boldsymbol{A}$ and the probability of $\boldsymbol{A} \Rightarrow \boldsymbol{B}$. What is the probability of $\boldsymbol{B}$?

The correct answer is that we do not have enough information to answer the question.

Let us now consider the case where we are given more information. For instance, we may know the probability of $\neg \boldsymbol{A} \Rightarrow \boldsymbol{B}$. We can then write the following:

$$
\begin{aligned}
R\left(\boldsymbol{B}\right) \;&=\; R\left(\boldsymbol{A} \cap \boldsymbol{B}\right) + R\left(\neg \boldsymbol{A} \cap \boldsymbol{B}\right) = & (3.16)\\
&=\; R\left(\boldsymbol{B}|\boldsymbol{A}\right) \cdot R\left(\boldsymbol{A}\right) + R\left(\boldsymbol{B}|\neg \boldsymbol{A}\right) \cdot R\left(\neg \boldsymbol{A}\right) = & (3.17)\\
&=\; R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) - 1 + R\left(\boldsymbol{A}\right) + R\left(\neg \boldsymbol{A} \Rightarrow \boldsymbol{B}\right) - 1 + R\left(\neg \boldsymbol{A}\right). & (3.18)
\end{aligned}
$$

Therefore:

$$R\left(\boldsymbol{B}\right) = R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) + R\left(\neg \boldsymbol{A} \Rightarrow \boldsymbol{B}\right) - 1. \tag{3.19}$$

This is a very interesting generalisation of modus ponens. It is surprising that we do not need $R\left(\boldsymbol{A}\right)$ to compute $R\left(\boldsymbol{B}\right)$. All the required information about $\boldsymbol{A}$ is given in the probabilities $R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right)$ and $R\left(\neg \boldsymbol{A} \Rightarrow \boldsymbol{B}\right)$.

Let us now show that equation (3.19) is indeed a generalisation of modus ponens: If $R\left(\boldsymbol{A}\right) = 1$ and $R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) = 1$, then:

$$R\left(\neg\left(\neg \boldsymbol{A}\right) \vee \left(\neg \boldsymbol{A} \wedge \boldsymbol{B}\right)\right) = 1, \qquad R\left(\neg \boldsymbol{A} \Rightarrow \boldsymbol{B}\right) = 1, \qquad R\left(\boldsymbol{B}\right) = 1. \tag{3.20}$$

Other generalisations of modus ponens can be similarly derived:

$$
\begin{aligned}
R\left(\boldsymbol{B}\right) &= R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) + R\left(\boldsymbol{A}\right) - R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right), & (3.21)\\
R\left(\boldsymbol{B}\right) &= 2 - R\left(\boldsymbol{B} \Rightarrow \neg\boldsymbol{A}\right) - R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right), & (3.22)\\
R\left(\boldsymbol{B}\right) &= R\left(\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) + R\left(\neg\boldsymbol{A}\right) - R\left(\boldsymbol{B} \Rightarrow \neg\boldsymbol{A}\right). & (3.23)
\end{aligned}
$$

The following are variants for floppy membership functions:

$$
\begin{aligned}
\boldsymbol{\mu_B}\left(x\right) &= \boldsymbol{\mu_{A \Rightarrow B}}\left(x\right) + \boldsymbol{\mu_{\neg A \Rightarrow B}}\left(x\right) - 1, & (3.24)\\
\boldsymbol{\mu_B}\left(x\right) &= \boldsymbol{\mu_{A \Rightarrow B}}\left(x\right) + \boldsymbol{\mu_A}\left(x\right) - \boldsymbol{\mu_{B \Rightarrow A}}\left(x\right), & (3.25)\\
\boldsymbol{\mu_B}\left(x\right) &= 2 - \boldsymbol{\mu_{B \Rightarrow \neg A}}\left(x\right) - \boldsymbol{\mu_{B \Rightarrow A}}\left(x\right), & (3.26)\\
\boldsymbol{\mu_B}\left(x\right) &= \boldsymbol{\mu_{\neg A \Rightarrow B}}\left(x\right) + \boldsymbol{\mu_{\neg A}}\left(x\right) - \boldsymbol{\mu_{B \Rightarrow \neg A}}\left(x\right). & (3.27)
\end{aligned}
$$

3.2.4 Modus Ponens for a Finite Partition of the Sample Space

Let $\{\boldsymbol{A}_i\}$ be a finite set of n pairwise disjoint floppy sets whose union is the whole sample space. We can then generalise modus ponens as follows:

$$
\begin{aligned}
R\left(\boldsymbol{B}\right) &= \sum_{i=1}^{n} R\left(\boldsymbol{A}_i \Rightarrow \boldsymbol{B}\right) - n + 1, & (3.28)\\
R\left(\boldsymbol{B}\right) &= \frac{n - \sum_{i=1}^{n} R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}_i\right)}{n-1}, & (3.29)\\
R\left(\boldsymbol{B}\right) &= n - \sum_{i=1}^{n} R\left(\boldsymbol{B} \Rightarrow \neg\boldsymbol{A}_i\right), & (3.30)\\
R\left(\boldsymbol{B}\right) &= \frac{\sum_{i=1}^{n} R\left(\neg\boldsymbol{A}_i \Rightarrow \boldsymbol{B}\right) - 1}{n-1}, & (3.31)
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\mu_B}\left(x\right) &= \sum_{i=1}^{n} \boldsymbol{\mu_{A_i \Rightarrow B}}\left(x\right) - n + 1, & (3.32)\\
\boldsymbol{\mu_B}\left(x\right) &= \frac{n - \sum_{i=1}^{n} \boldsymbol{\mu_{B \Rightarrow A_i}}\left(x\right)}{n-1}, & (3.33)\\
\boldsymbol{\mu_B}\left(x\right) &= n - \sum_{i=1}^{n} \boldsymbol{\mu_{B \Rightarrow \neg A_i}}\left(x\right), & (3.34)\\
\boldsymbol{\mu_B}\left(x\right) &= \frac{\sum_{i=1}^{n} \boldsymbol{\mu_{\neg A_i \Rightarrow B}}\left(x\right) - 1}{n-1}. & (3.35)
\end{aligned}
$$

Example 3.2.1. Let us show that in standard two-valued logic, the relation 3.29 holds for four sentences $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$, and $\boldsymbol{B}$:

Table 3.1: Demonstration of the validity of equation 3.29

$R\left(\boldsymbol{A_1}\right)$	$R\left(\boldsymbol{A_2}\right)$	$R\left(\boldsymbol{A_3}\right)$	$R\left(\boldsymbol{B}\right)$	$R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A_1}\right)$	$R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A_2}\right)$	$R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A_3}\right)$	$\frac{n-\sum_{i=1}^{n} R(B \Rightarrow A_i)}{n-1}$
1	0	0	**1**	1	0	0	**1**
1	0	0	**0**	1	1	1	**0**
0	1	0	**1**	0	1	0	**1**
0	1	0	**0**	1	1	1	**0**
0	0	1	**1**	0	0	1	**1**
0	0	1	**0**	1	1	1	**0**

We see that the fourth and eighth columns of Table 3.1 are the same.

3.2.5 Modus Tollens

Modus tollens is a logical rule which states that if $A \Rightarrow B$ holds and B does not hold, then A does not hold.

In floppy logic, we can describe these generalisations of modus tollens as follows:

$$R\left(\neg\boldsymbol{A}\right) = R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) + R\left(\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}\right) - 1, \tag{3.36}$$

$$R\left(\neg\boldsymbol{A}\right) = R\left(\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) + R\left(\neg\boldsymbol{B}\right) - R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right), \tag{3.37}$$

$$R\left(\neg\boldsymbol{A}\right) = 2 - R\left(\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}\right) - R\left(\boldsymbol{B} \Rightarrow \boldsymbol{A}\right), \tag{3.38}$$

$$R\left(\neg\boldsymbol{A}\right) = R\left(\boldsymbol{B} \Rightarrow \neg\boldsymbol{A}\right) + R\left(\boldsymbol{B}\right) - R\left(\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}\right). \tag{3.39}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}\left(x\right) = \boldsymbol{\mu}_{\boldsymbol{A}\Rightarrow\boldsymbol{B}}\left(x\right) + \boldsymbol{\mu}_{\boldsymbol{A}\Rightarrow\neg\boldsymbol{B}}\left(x\right) - 1, \tag{3.40}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}\left(x\right) = \boldsymbol{\mu}_{\boldsymbol{A}\Rightarrow\boldsymbol{B}}\left(x\right) + \boldsymbol{\mu}_{\neg\boldsymbol{B}}\left(x\right) - \boldsymbol{\mu}_{\boldsymbol{B}\Rightarrow\boldsymbol{A}}\left(x\right), \tag{3.41}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}\left(x\right) = 2 - \boldsymbol{\mu}_{\neg\boldsymbol{A}\Rightarrow\boldsymbol{B}}\left(x\right) - \boldsymbol{\mu}_{\boldsymbol{B}\Rightarrow\boldsymbol{A}}\left(x\right), \tag{3.42}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}\left(x\right) = \boldsymbol{\mu}_{\boldsymbol{B}\Rightarrow\neg\boldsymbol{A}}\left(x\right) + \boldsymbol{\mu}_{\boldsymbol{B}}\left(x\right) - \boldsymbol{\mu}_{\neg\boldsymbol{A}\Rightarrow\boldsymbol{B}}\left(x\right). \tag{3.43}$$

3.2.6 Modus Tollens for a Finite Partition of the Sample Space

Let $\{\boldsymbol{B}_i\}$ be a finite set of n pairwise disjoint floppy sets whose union is the entire sample space. We can then generalise modus tollens:

$$R(\neg\boldsymbol{A}) = \frac{\sum_{i=1}^{n} R(\boldsymbol{A} \Rightarrow \boldsymbol{B}_i) - 1}{n - 1}, \tag{3.44}$$

$$R(\neg\boldsymbol{A}) = n - \sum_{i=1}^{n} R(\boldsymbol{B}_i \Rightarrow \boldsymbol{A}), \tag{3.45}$$

$$R(\neg\boldsymbol{A}) = \sum_{i=1}^{n} R(\boldsymbol{B}_i \Rightarrow \neg\boldsymbol{A}) - n + 1, \tag{3.46}$$

$$R(\neg\boldsymbol{A}) = \frac{n - \sum_{i=1}^{n} R(\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}_i)}{n - 1}, \tag{3.47}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}(x) = \frac{\sum_{i=1}^{n} \boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \boldsymbol{B}_i}(x) - 1}{n - 1}, \tag{3.48}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}(x) = n - \sum_{i=1}^{n} \boldsymbol{\mu}_{\boldsymbol{B}_i \Rightarrow \boldsymbol{A}}(x), \tag{3.49}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}(x) = \sum_{i=1}^{n} \boldsymbol{\mu}_{\boldsymbol{B}_i \Rightarrow \neg\boldsymbol{A}}(x) - n + 1, \tag{3.50}$$

$$\boldsymbol{\mu}_{\neg\boldsymbol{A}}(x) = \frac{n - \sum_{i=1}^{n} \boldsymbol{\mu}_{\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}_i}(x)}{n - 1}. \tag{3.51}$$

3.2.7 Relationships between Implications

Implications can be substituted according to the following rules:

$$R(\boldsymbol{A} \Rightarrow \boldsymbol{B}) = R(\boldsymbol{B} \Rightarrow \boldsymbol{A}) - R(\boldsymbol{A}) + R(\boldsymbol{B}), \tag{3.52}$$

$$R(\boldsymbol{A} \Rightarrow \boldsymbol{B}) = R(\boldsymbol{B}) + 1 - R(\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}), \tag{3.53}$$

$$R(\boldsymbol{A} \Rightarrow \boldsymbol{B}) = R(\neg\boldsymbol{A}) + 1 - R(\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}), \tag{3.54}$$

$$\boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \boldsymbol{B}}(x) = \boldsymbol{\mu}_{\boldsymbol{B} \Rightarrow \boldsymbol{A}}(x) - \boldsymbol{\mu}_{\boldsymbol{A}}(x) + \boldsymbol{\mu}_{\boldsymbol{B}}(x), \tag{3.55}$$

$$\boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \boldsymbol{B}}(x) = \boldsymbol{\mu}_{\boldsymbol{B}}(x) + 1 - \boldsymbol{\mu}_{\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}}(x), \tag{3.56}$$

$$\boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \boldsymbol{B}}(x) = \boldsymbol{\mu}_{\neg\boldsymbol{A}}(x) + 1 - \boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}}(x). \tag{3.57}$$

$$R(\boldsymbol{A} \Rightarrow \boldsymbol{B}) + R(\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}) + R(\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}) + R(\neg\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}) = 3, \tag{3.58}$$

$$\boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \boldsymbol{B}}(x) + \boldsymbol{\mu}_{\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}}(x) + \boldsymbol{\mu}_{\neg\boldsymbol{A} \Rightarrow \boldsymbol{B}}(x) + \boldsymbol{\mu}_{\neg\boldsymbol{A} \Rightarrow \neg\boldsymbol{B}}(x) = 3. \tag{3.59}$$

3.3 Dependence Measurement

It is possible to compare the (floppy) membership functions of intersection and union in both floppy logic and fuzzy logic:

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two floppy sets with membership functions $\boldsymbol{\mu_A}(x)$ and $\boldsymbol{\mu_B}(x)$. The following applies to the union of floppy sets:

$$\max\left\{\boldsymbol{\mu_A}(x),\, \boldsymbol{\mu_B}(x)\right\} \leq \boldsymbol{\mu_{A\cup B}}(x) \leq \min\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x),\, 1\right\}. \qquad (3.60)$$

Therefore, $\boldsymbol{\mu_{A\cup B}}(x)$ is bounded by Gödel and Łukasiewicz t-conorms.

The following applies to the intersection of floppy sets:

$$\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\, 0\right\} \leq \boldsymbol{\mu_{A\cap B}}(x) \leq \min\left\{\boldsymbol{\mu_A}(x),\, \boldsymbol{\mu_B}(x)\right\}. \qquad (3.61)$$

$\boldsymbol{\mu_{A\cap B}}(x)$ must be greater than or equal to $\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\, 0\right\}$, because if $\boldsymbol{\mu_{A\cap B}}(x)$ were less than $\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\, 0\right\}$, then $\boldsymbol{\mu_{A\cup B}}(x) > 1$ would apply.

Therefore, $\boldsymbol{\mu_{A\cap B}}(x)$ is bounded by Łukasiewicz and Gödel t-norms.

The intersection is equal to the Łukasiewicz t-norm, and the union is equal to the Łukasiewicz t-conorm simultaneously. In this case, $\boldsymbol{\mu_{A\Leftrightarrow B}}(x)$ is as minimal as possible.

Similarly, the intersection is equal to the Gödel t-norm, and the union is equal to the Gödel t-conorm simultaneously. In this case, $\boldsymbol{\mu_{A\Leftrightarrow B}}(x)$ is as maximal as possible.

Interestingly, a coefficient which measures dependence can be derived from this comparison of (floppy) membership functions in floppy logic and fuzzy logic:

If the floppy membership functions of intersection and equivalence are as maximal as possible and the floppy membership function of union is as minimal as possible, then the dependence of $\boldsymbol{A}$ and $\boldsymbol{B}$ is as maximal possible (for a given $\boldsymbol{\mu_A}(x)$ and $\boldsymbol{\mu_B}(x)$), and vice versa.

The floppy membership function of equivalence (or intersection or union) can be used to measure dependency, either directly or after normalisation to the interval $[0, 1]$.

Of course, normalisation is possible only if a spectrum of possible values of equivalence (or intersection or union) exists in the floppy membership function. If some of the floppy membership functions $\boldsymbol{\mu_A}(x)$ or $\boldsymbol{\mu_B}(x)$ are equal to zero or one, then the membership function of equivalence (and intersection and union) is given unambiguously and normalisation is impossible.

The normalised measure of dependency can be calculated in the following manner:

Intersection:

$$\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\} \leq \boldsymbol{\mu_{A \cap B}}(x) \leq \min\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\}, \quad (3.62)$$

$$0 \leq \frac{\boldsymbol{\mu_{A \cap B}}(x) - \max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}}{\min\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\} - \max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}} \leq 1, \quad (3.63)$$

$$K_{\cap}(x) = \frac{\boldsymbol{\mu_{A \cap B}}(x) - \max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}}{\min\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\} - \max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}}. \quad (3.64)$$

Union:

$$\max\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\} \leq \boldsymbol{\mu_{A \cup B}}(x) \leq \min\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x),\ 1\right\}, \quad (3.65)$$

$$0 \leq \frac{\boldsymbol{\mu_{A \cup B}}(x) - \max\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\}}{\min\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x),\ 1\right\} - \max\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\}} \leq 1, \quad (3.66)$$

$$K_{\cup}(x) = \frac{\boldsymbol{\mu_{A \cup B}}(x) - \max\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\}}{\min\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x),\ 1\right\} - \max\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\}}. \quad (3.67)$$

Equivalence:

$$\boldsymbol{\mu_{A \Leftrightarrow B}}(x) = \boldsymbol{\mu_{A \cap B}}(x) + \boldsymbol{\mu_{A' \cap B'}}(x) = \quad (3.68)$$

$$= 1 - \boldsymbol{\mu_A}(x) - \boldsymbol{\mu_B}(x) + 2\boldsymbol{\mu_{A \cap B}}(x), \quad (3.69)$$

$$\dots$$

$$\boldsymbol{\mu_{A \Leftrightarrow B}}(x) \geq 1 - \boldsymbol{\mu_A}(x) - \boldsymbol{\mu_B}(x) + 2\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}, (3.70)$$

$$\boldsymbol{\mu_{A \Leftrightarrow B}}(x) \leq 1 - \boldsymbol{\mu_A}(x) - \boldsymbol{\mu_B}(x) + 2\min\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\}, \quad (3.71)$$

$$0 \leq \frac{\boldsymbol{\mu_{A \Leftrightarrow B}}(x) - 1 + \boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 2\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}}{2\min\left\{\boldsymbol{\mu_A}(x),\ \boldsymbol{\mu_B}(x)\right\} - 2\max\left\{\boldsymbol{\mu_A}(x) + \boldsymbol{\mu_B}(x) - 1,\ 0\right\}} \leq 1, (3.72)$$

$$K_{\Leftrightarrow}(x) =$$

$$= \frac{\boldsymbol{\mu}_{\boldsymbol{A} \Leftrightarrow \boldsymbol{B}}(x) - 1 + \boldsymbol{\mu}_{\boldsymbol{A}}(x) + \boldsymbol{\mu}_{\boldsymbol{B}}(x) - 2 \max\{\boldsymbol{\mu}_{\boldsymbol{A}}(x) + \boldsymbol{\mu}_{\boldsymbol{B}}(x) - 1, \, 0\}}{2 \min\{\boldsymbol{\mu}_{\boldsymbol{A}}(x), \, \boldsymbol{\mu}_{\boldsymbol{B}}(x)\} - 2 \max\{\boldsymbol{\mu}_{\boldsymbol{A}}(x) + \boldsymbol{\mu}_{\boldsymbol{B}}(x) - 1, \, 0\}}. \quad (3.73)$$

It is curious that

$$K_{\cap}(x) = K_{\Leftrightarrow}(x) = 1 - K_{\cup}(x) = K(x), \quad (3.74)$$

where $K(x)$ is the *dependence coefficient*.

For independent events $\boldsymbol{A}$ and $\boldsymbol{B}$, for a given x, the following is satisfied:

$$R(\boldsymbol{A} \cap \boldsymbol{B}|x) = R(\boldsymbol{A}|x) \cdot R(\boldsymbol{B}|x), \quad (3.75)$$

$$\boldsymbol{\mu}_{\boldsymbol{A} \cap \boldsymbol{B}}(x) = \boldsymbol{\mu}_{\boldsymbol{A}}(x) \cdot \boldsymbol{\mu}_{\boldsymbol{B}}(x). \quad (3.76)$$

In this case, for all x, the following simple relationship applies:

$$K(x) = \boldsymbol{\mu}_{\boldsymbol{X}}(x), \quad (3.77)$$

where $\boldsymbol{X}$ is a *friend* of $\boldsymbol{Y} \in \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{A}', \boldsymbol{B}'\}$ and

$$\boldsymbol{\mu}_{\boldsymbol{Y}}(x) = \min\{\boldsymbol{\mu}_{\boldsymbol{A}}(x), \, \boldsymbol{\mu}_{\boldsymbol{B}}(x), \, \boldsymbol{\mu}_{\boldsymbol{A}'}(x), \, \boldsymbol{\mu}_{\boldsymbol{B}'}(x)\}. \quad (3.78)$$

$\boldsymbol{A}$ and $\boldsymbol{B}$ are friends and $\boldsymbol{A}'$ and $\boldsymbol{B}'$ are friends.

A similar dependency coefficient can be created if we repeat the entire procedure with probabilities instead of floppy membership functions.

Chapter 4

Using Floppy Logic to Describe a System

4.1 Introductory Example

This chapter illustrates how to work with a system using floppy logic. We start with the choice of primary fuzzy sets and end with optimal control. First, however, the most important concepts and relationships of floppy logic are demonstrated in the following example.

We want to describe a quantity, such as the speed of a car, using primary fuzzy sets, such as A_1 – "too slow", A_2 – "satisfactory", A_3 – "too fast". The membership functions of these primary fuzzy sets can be represented as in Figure 4.1.
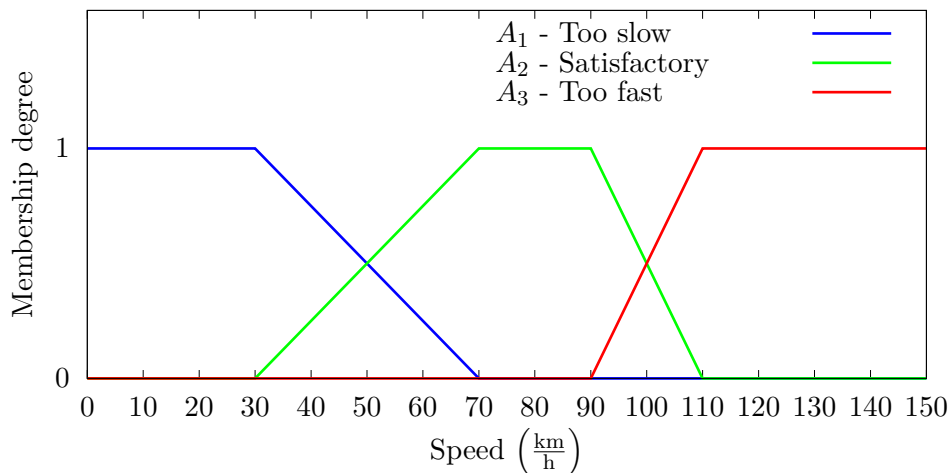
Figure 4.1: Primary fuzzy sets

Floppy sets are crisp sets of these primary fuzzy sets. For example, floppy set $\boldsymbol{B_{011}}$ – "Not too slow" obtains two primary fuzzy sets A_2 and A_3:

$$\boldsymbol{B_{011}} = \{A_2, A_3\}. \tag{4.1}$$

The floppy membership function of a floppy set is the sum of the membership functions of its elements:

$$\boldsymbol{\mu_{B_{011}}}(x) = \mu_{A_2}(x) + \mu_{A_3}(x). \tag{4.2}$$

We can write (see Section 2.2.2):

$$\boldsymbol{\mu_{B_{011}}}(x) = R\left(\boldsymbol{B_{011}}|x\right). \tag{4.3}$$

This conditional probability can be interpreted as the probability that somebody (e.g. an expert) decides that speed x "is not too slow".

Besides floppy sets, speed can also be described by the probability density function $f(x)$. This function can be, for example, estimated from measured data.

Let the speed of the car at this point and time have a normal distribution with a mean value of $90\frac{\text{km}}{\text{h}}$ and standard deviation of $15\frac{\text{km}}{\text{h}}$. This probability density function is shown in Figure 4.2.
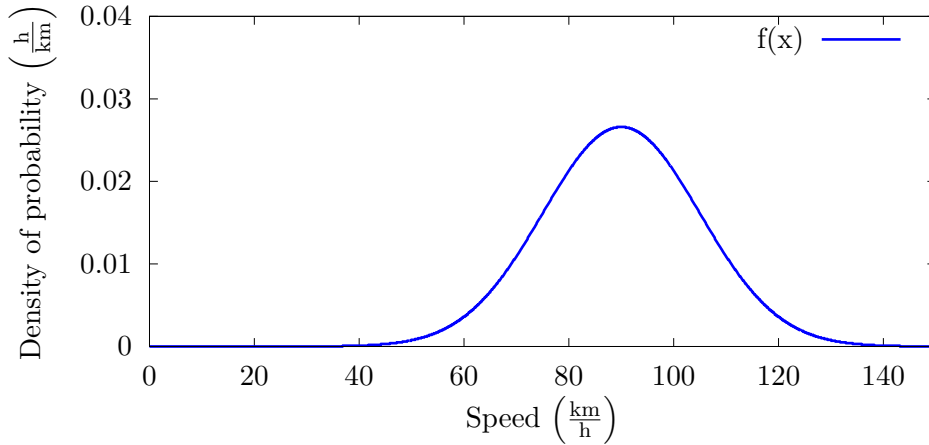

Figure 4.2: Probability density function

When we know the membership functions and the probability density, we can calculate, for example, the probability or the mean value of the floppy sets. For example:

$$R\left(\boldsymbol{B_{011}}\right) = \int_X \boldsymbol{\mu_{B_{011}}}(x) \cdot f(x)\,\mathrm{d}x = 0.9841, \tag{4.4}$$

$$\langle \boldsymbol{B_{011}} \rangle = \frac{\int_X x \cdot \boldsymbol{\mu_{B_{011}}}(x) \cdot f(x)\,\mathrm{d}x}{\int_X \boldsymbol{\mu_{B_{011}}}(x) \cdot f(x)\,\mathrm{d}x} = \frac{89.08}{0.9841} = 90.52. \tag{4.5}$$

46

The functions from Figures 4.1 and 4.2 were inserted.

When we model statements using floppy sets, conjunction and disjunction are modelled by ordinary intersection and union. For example, the statement "the speed is too slow or too fast" is modelled as follows:

$$\boldsymbol{B_{100}} \vee \boldsymbol{B_{001}} = \{A_1\} \cup \{A_3\} = \{A_1, A_3\}. \tag{4.6}$$

The relevant floppy membership function can be written in the following manner:

$$\boldsymbol{\mu_{B_{100} \vee B_{001}}}(x) = \mu_{A_1}(x) + \mu_{A_3}(x). \tag{4.7}$$

Other logical connectives must first be converted into conjunctions, disjunctions, and negations. It does not matter how it is done (see Isomorphism Theorem 2.3). For example, the sentence "If the speed is not too slow, then it is satisfactory," can be modelled as follows:

$$\neg\boldsymbol{B_{100}} \Rightarrow \boldsymbol{B_{010}} = \boldsymbol{B_{100}} \vee \boldsymbol{B_{010}} = \{A_1\} \cup \{A_2\} = \{A_1, A_2\}. \tag{4.8}$$

Another way might be

$$\begin{aligned}
\neg\boldsymbol{B_{100}} \Rightarrow \boldsymbol{B_{010}} &= \neg\boldsymbol{B_{010}} \Rightarrow \boldsymbol{B_{100}} = \\
&= \boldsymbol{B_{010}} \vee (\neg\boldsymbol{B_{010}} \wedge \boldsymbol{B_{100}}) = \\
&= \{A_2\} \cup \left(\{A_2\}' \cap \{A_1\}\right) = \\
&= \{A_2\} \cup (\{A_1, A_3\} \cap \{A_1\}) = \\
&= \{A_2\} \cup \{A_1\} = \{A_1, A_2\}. \tag{4.9}
\end{aligned}$$

The relevant floppy membership function can be written as

$$\boldsymbol{\mu_{\neg B_{100} \Rightarrow B_{010}}}(x) = \mu_{A_1}(x) + \mu_{A_2}(x). \tag{4.10}$$

In this example, the basic properties of floppy logic are shown.

The following sections will show in detail the requirements for describing and controlling a system using floppy logic:

- How to select primary fuzzy sets and their membership functions.

- How to create input and output floppy sets from primary fuzzy sets.

- How to calculate the probabilities of input floppy sets.

- How to implement system rules and calculate the probabilities of output floppy sets.

- How to determine probability distributions and the mean values of output quantities from the probabilities of output floppy sets.

- How to control a system described by floppy logic.

4.2   Selection of Primary Fuzzy Sets

We want to describe a discrete or continuous quantity using several primary fuzzy sets. What primary fuzzy sets should we select?

1. We select the primary fuzzy sets which correspond to the floppy sets used in the system rules.

2. Each point of the domain of the input and output quantities must be covered by a primary fuzzy set.

3. Primary fuzzy sets describing the same variable must be mutually incompatible.[1]

Example 4.2.1. Let us have a system where we switch on heating according to the water temperature. We consider the following system rules:

If the water is cold, then . . .

If the water is pleasant, then . . .

Our first choice of primary fuzzy set is "cold" and "pleasant" since these properties appeared in the system rules.

However, this is not completely right since these properties are compatible. Water can be cold and pleasant simultaneously. Our second choice therefore is "unpleasantly cold", "cold and pleasant", "pleasant but not cold".

This is still not satisfactory, because we do not consider hot water. We add the primary fuzzy set "warm".

This is also not right, because water can be warm and pleasant simultaneously. Our final choice therefore looks like this: "unpleasantly cold", "pleasantly cold", "pleasant but not cold and not warm", "pleasantly warm", "unpleasantly warm".

Then we add a system rule to consider warm water. For example:

If water is warm, then . . . '

---

[1]Primary fuzzy sets describing the same variable must be mutually incompatible because the intersection of two different single-element floppy sets is the empty set.

The floppy sets, "cold", "pleasant", and "warm", which appeared in the system rules, contain two, three, and two elements (primary fuzzy sets), respectively.

4.3 Determination of the Membership Functions of Primary Fuzzy Sets

In the previous section, we selected primary fuzzy sets. Membership functions must now be assigned to these primary fuzzy sets. When we search for membership functions, we follow these three rules:

1. The sum of the membership functions of all primary fuzzy sets must be 1 everywhere:

$$\sum_{A_i \in \boldsymbol{S}} \mu_{A_i}(x) = 1. \tag{4.11}$$

2. The floppy membership function of floppy set $\boldsymbol{B}$ is the sum of the membership functions of its elements:

$$\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x). \tag{4.12}$$

3. The floppy membership function of floppy set $\boldsymbol{B}$ can be understood as a conditional probability:

$$\boldsymbol{\mu_B}(x) = R(\boldsymbol{B}|x). \tag{4.13}$$

Now, let us study some examples of how to determine the membership functions of primary fuzzy sets in different situations.

4.3.1 Determination of Membership Functions from Experts' Opinions

We start from the relation:

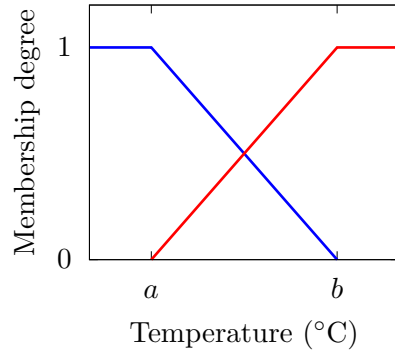$$\boldsymbol{\mu_B}(x) = R(\boldsymbol{B}|x). \tag{4.14}$$

Figure 4.3: Membership functions for cold and warm water

Conditional probability $R\left(\boldsymbol{B}|x\right)$ can be understood as the probability that an expert (or someone else) will state that for a given x, the system possesses property $\boldsymbol{B}$.

Of course, sometimes it is not necessary to ask experts. We can often estimate membership functions ourselves. We can estimate conditional probability $R\left(\boldsymbol{B}|x\right)$ for several x and determine the course of the membership function.

This method can be applied to the floppy membership function of floppy sets and to the membership function of primary fuzzy sets.

Example 4.3.1. We want to describe water temperature using two primary fuzzy sets: cold (C) and warm (W).

For membership functions $\mu_C\left(x\right)$ and $\mu_W\left(x\right)$, we want to use piecewise linear functions, as shown in Figure 4.3.

We want to determine the membership functions to best suit experts' opinions. We must therefore estimate the coefficients a and b as best we can.

We ask fifty people (experts) to determine whether the water is cold or warm. The results are shown in Table 4.1.

Table 4.1: Responses from participants to whether the water is cold or warm

Temperature (°C)	Cold	Warm
10	50	0
15	48	2
20	42	8
25	27	23
30	4	46
35	0	50

For $\mu_C(x)$ and $\mu_W(x)$, it applies:

$$\mu_C(x) = 1 - \mu_W(x).\tag{4.15}$$

It is therefore sufficient to determine, for example, only $\mu_C(x)$.

We use the least-squares method to estimate the coefficients a and b and minimise the expression:

$$
\begin{aligned}
50 \cdot (1 - \mu_C(10))^2 + 0 \cdot (0 - \mu_C(10))^2 + \\
+ 48 \cdot (1 - \mu_C(15))^2 + 2 \cdot (0 - \mu_C(15))^2 + \\
\cdots \\
+ 0 \cdot (1 - \mu_C(35))^2 + 50 \cdot (0 - \mu_C(35))^2.
\end{aligned}
\tag{4.16}
$$

We obtain the parameters:

$$a = 18.246,\tag{4.17}$$
$$b = 31.404.\tag{4.18}$$

4.3.2 Determination of the Membership Functions of Primary Fuzzy Sets from the Applied Floppy Sets

The following example shows how to proceed if the membership functions of primary fuzzy sets are required and some floppy membership functions have already been specified.

Example 4.3.2. We now describe the water temperature again and begin with three floppy sets: A_1 – Unhealthily cold, A_3 – Pleasant, A_4 – Warm. We assume that these properties appear in the system rules. The membership functions are presented in Figure 4.4(a).

First, we add floppy set A_2 – Healthily cold – so that the sum is not less than one anywhere (Fig. 4.4(b)).

If the sum for x is greater than one and two nonzero floppy sets A_i and A_j exist for this x, we substitute floppy sets A_i and A_j with the three primary fuzzy sets "Only A_i", "Only A_j" and "A_i and A_j".

The primary fuzzy sets "Only A_i" are denoted A with binary index containing 0 and 1 at the i-th position. The fuzzy sets "A_i and A_j" will contain 1 at the i-th and j-th positions.
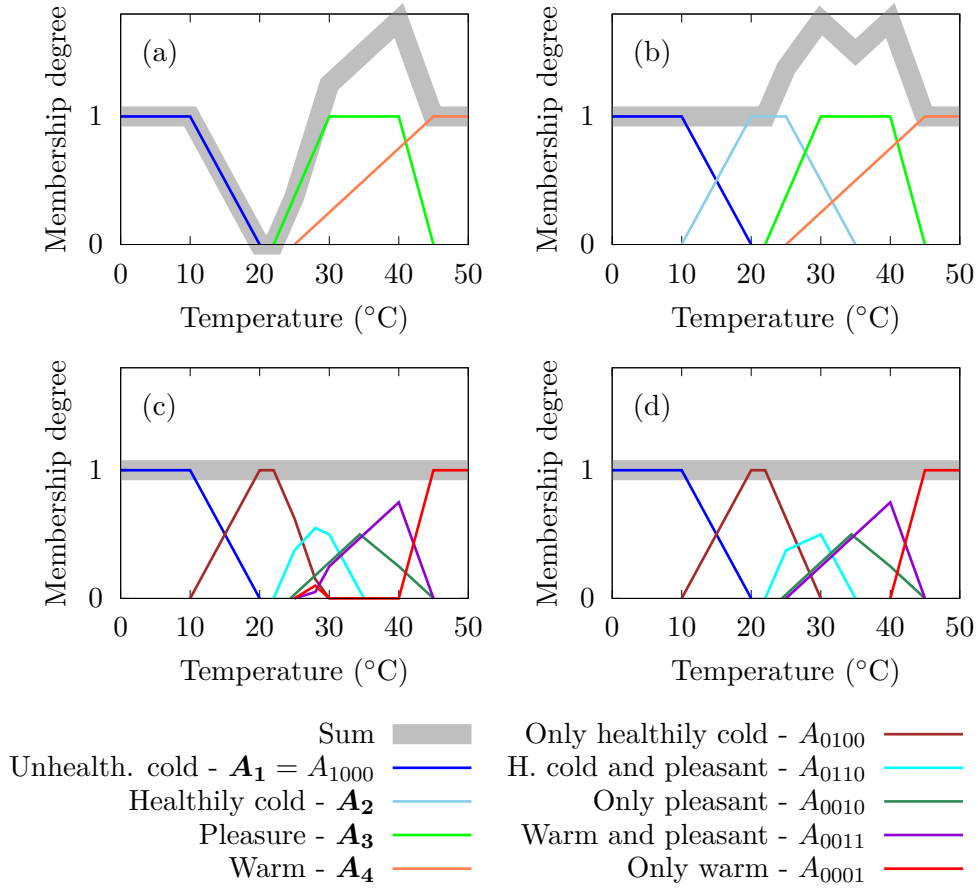
Figure 4.4: Applied membership functions

For simplicity, the (floppy) membership functions in the following equations are denoted in the same manner as the floppy sets and primary fuzzy sets:

For example, for floppy sets $\boldsymbol{A_3}$ and $\boldsymbol{A_4}$, the following must be satisfied:

$$A_{0010} + A_{0001} + A_{0011} = 1, \tag{4.19}$$

$$A_{0010} + A_{0011} = \boldsymbol{A_3}, \tag{4.20}$$

$$A_{0001} + A_{0011} = \boldsymbol{A_4}, \tag{4.21}$$

such that

$$A_{0011} = \boldsymbol{A_3} + \boldsymbol{A_4} - 1, \tag{4.22}$$

$$A_{0010} = \boldsymbol{A_3} - A_{0011}, \tag{4.23}$$

$$A_{0001} = \boldsymbol{A_4} - A_{0011}. \tag{4.24}$$

The membership degree of floppy set "$\boldsymbol{A_i}$ and $\boldsymbol{A_j}$" is the sum minus 1, and we must

subtract the same value from A_i and A_j to obtain the membership functions of "Only A_i", "Only A_j".

A problem arises if for some x more non-zero membership functions exist. This case occurs for $x \in (25, 35)$.

Floppy sets A_2, A_3, and A_4 can be substituted with primary fuzzy sets A_{0100}, A_{0010}, A_{0001}, A_{0110}, A_{0101}, A_{0011}, and A_{0111}. We suppose that no expert states that the water is cold and warm simultaneously. Therefore, we can eliminate sets A_{0101} and A_{0111}.

The following must be satisfied:

$$A_{0100} + A_{0010} + A_{0001} + A_{0110} + A_{0011} = 1, \tag{4.25}$$

$$A_{0100} + A_{0110} = A_2, \tag{4.26}$$

$$A_{0010} + A_{0110} + A_{0011} = A_3, \tag{4.27}$$

$$A_{0001} + A_{0011} = A_4. \tag{4.28}$$

For $x \in [30, 35)$ is the A_3 floppy membership function 1. The membership functions of primary fuzzy sets can therefore be calculated unambiguously.

Many solutions exist at interval $(25, 30)$. For example, let us decide that the best choice in our mathematical model is $A_{0110}(28) = 0.55$ and $A_{0011}(28) = 0.05$. This solution is illustrated in Figure 4.4(c).

This solution is rather strange, however. The primary fuzzy set "Only warm" is non-zero at interval $(25, 30)$ and then at interval $(40, 50)$. If we want a more elegant solution, we can divide the fuzzy set "Only warm" into two fuzzy sets, one for each interval.

Another option here is to select the values on interval $(25, 30)$ to be zero on the fuzzy set "Only warm". This solution is illustrated in Figure 4.4(d). These membership functions are listed in Appendix D.

4.3.3 Two-Dimensional Case – Marginal Floppy Sets

The following example shows how to proceed to create joint (floppy) membership functions from marginal (floppy) membership functions.

Example 4.3.3. Variable T is the air temperature. We select the following floppy sets to describe this property: T_1 – Cold, T_2 – Tepid, T_3 – Warm.

Variable P is the air pressure. We select the following floppy sets to describe this property: P_1 – Low, P_2 – Normal, P_3 – High.

These floppy sets satisfy all their required rules.

We want to describe the conditions in the air according to both temperature and pressure. The Cartesian product of temperature and pressure is therefore the domain of the floppy membership functions of the new floppy sets. Floppy set $\boldsymbol{T_1}$ is understood as "it is cold and pressure is arbitrary". Other floppy sets may be similarly understood. These marginal floppy sets are listed in Appendix E and shown in Figure 4.5(a) and (b).

We can observe that the sum of all floppy membership functions of $\boldsymbol{T_i}$ and $\boldsymbol{P_j}$ is 2 everywhere. We therefore introduce the primary fuzzy sets "Only $\boldsymbol{T_i}$", "Only $\boldsymbol{P_j}$" and "$\boldsymbol{T_i}$ and $\boldsymbol{P_j}$".

The primary fuzzy sets "Only $\boldsymbol{T_i}$" and "Only $\boldsymbol{P_j}$" are denoted $A_{i,0}$ and $A_{0,j}$. The primary fuzzy sets "$\boldsymbol{T_i}$ and $\boldsymbol{P_j}$" are denoted $A_{i,j}$.

As in the previous example, instead of the (floppy) membership functions, we write directly the relevant floppy or primary fuzzy sets.

We solve the following system of equations:

$$\sum_i A_{i,0} + \sum_j A_{0,j} + \sum_i \sum_j A_{i,j} = 1, \tag{4.29}$$

$$A_{i,0} + \sum_j A_{i,j} = \boldsymbol{T_i}, \tag{4.30}$$

$$A_{0,j} + \sum_i A_{i,j} = \boldsymbol{P_j}, \tag{4.31}$$

$$\sum_i \boldsymbol{T_i} = 1, \tag{4.32}$$

$$\sum_j \boldsymbol{P_j} = 1. \tag{4.33}$$

In solving this system, we find that the membership functions of all fuzzy sets "Only $\boldsymbol{T_i}$" and "Only $\boldsymbol{P_j}$" must be zero.

The system has an unambiguous solution only for a small number of non-zero fuzzy sets. Larger systems have more solutions.

However, one solution, always valid, is very simple:

$$A_{i,j} = \boldsymbol{T_i} \cdot \boldsymbol{P_j}. \tag{4.34}$$

The primary fuzzy set "Tepid and normal" obtained in this way is shown in Figure 4.5(c).

Multiplication of the floppy membership functions of marginal floppy sets is, of course, possible even in cases which are more than two-dimensional.
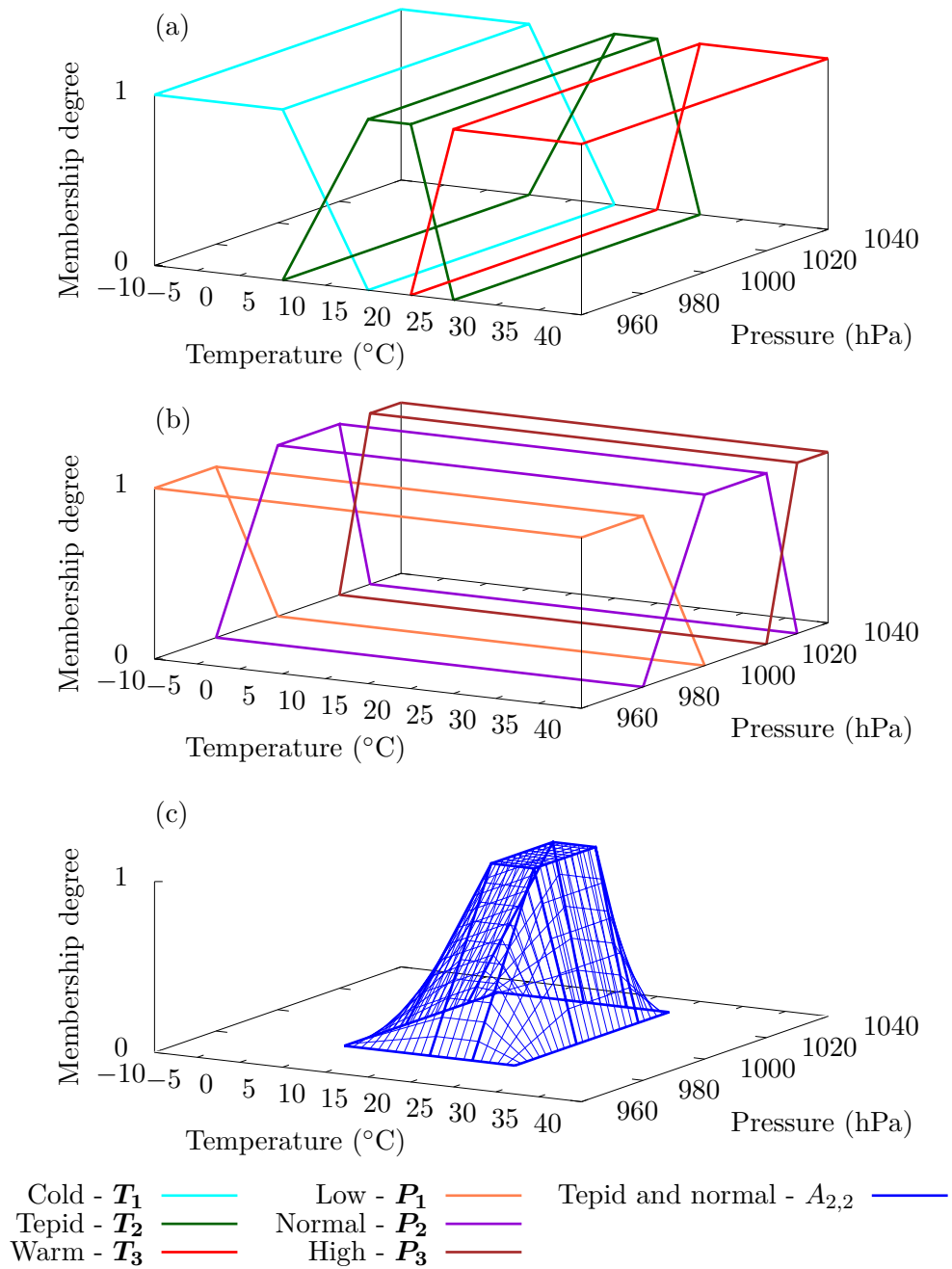
Figure 4.5: Marginal and joint (floppy) membership functions

4.3.4 Two-Dimensional Case – Clusters

This example demonstrates that clusters can be understood as primary fuzzy sets (or floppy sets) and how their (floppy) membership functions can be calculated.

The example was inspired by article [9], in which the authors described a fuzzy C-means clustering algorithm. This method was introduced in [6].

Example 4.3.4. We need to divide points in the plane into three clusters. These clusters are understood as the primary fuzzy sets $A_1, \ldots A_3$. The cluster centres are at points $C_1, \ldots C_3$. How do we obtain the cluster membership functions which satisfy all the required rules?

The method is very simple: Let us assign a non-negative function $g_i(X)$ to each cluster.[2] For example, function $g_i(X) = \frac{1}{d_i}$, where d_i is the distance between points X and C_i.[3] This is a standard choice in a C-means method.

Now, we calculate cluster membership functions according to the equation:

$$\mu_{A_i}(X) = \frac{g_i(X)}{\sum\limits_j g_j(X)}. \tag{4.35}$$

Many functions g_i are permissible. However, which functions best match the described reality?

We require such membership functions so that their ratio in X corresponds to the ratio in which the points at X belong to individual clusters. The best choice of $g_i(X)$ is therefore the density of intensity. This may often be calculated as the probability density function of the cluster multiplied by the intensity of the cluster.

For example, our three clusters have three different multivariate normal distributions (see [38] and Appendix E) with mean vectors (centres of clusters):

$$\mathbf{m_1} = \begin{pmatrix} -100 \\ 30 \end{pmatrix}, \qquad \mathbf{m_2} = \begin{pmatrix} 20 \\ -5 \end{pmatrix}, \qquad \mathbf{m_3} = \begin{pmatrix} 80 \\ 0 \end{pmatrix}, \tag{4.36}$$

covariance matrices:

$$\mathbf{M_1} = \begin{pmatrix} 1600 & -800 \\ -800 & 1600 \end{pmatrix}, \quad \mathbf{M_2} = \begin{pmatrix} 200 & 0 \\ 0 & 200 \end{pmatrix}, \quad \mathbf{M_3} = \begin{pmatrix} 300 & 100 \\ 100 & 300 \end{pmatrix}, \tag{4.37}$$

and cluster intensities of 70, 50 and 80 points per minute.

---

[2]At least one positive function must be everywhere.
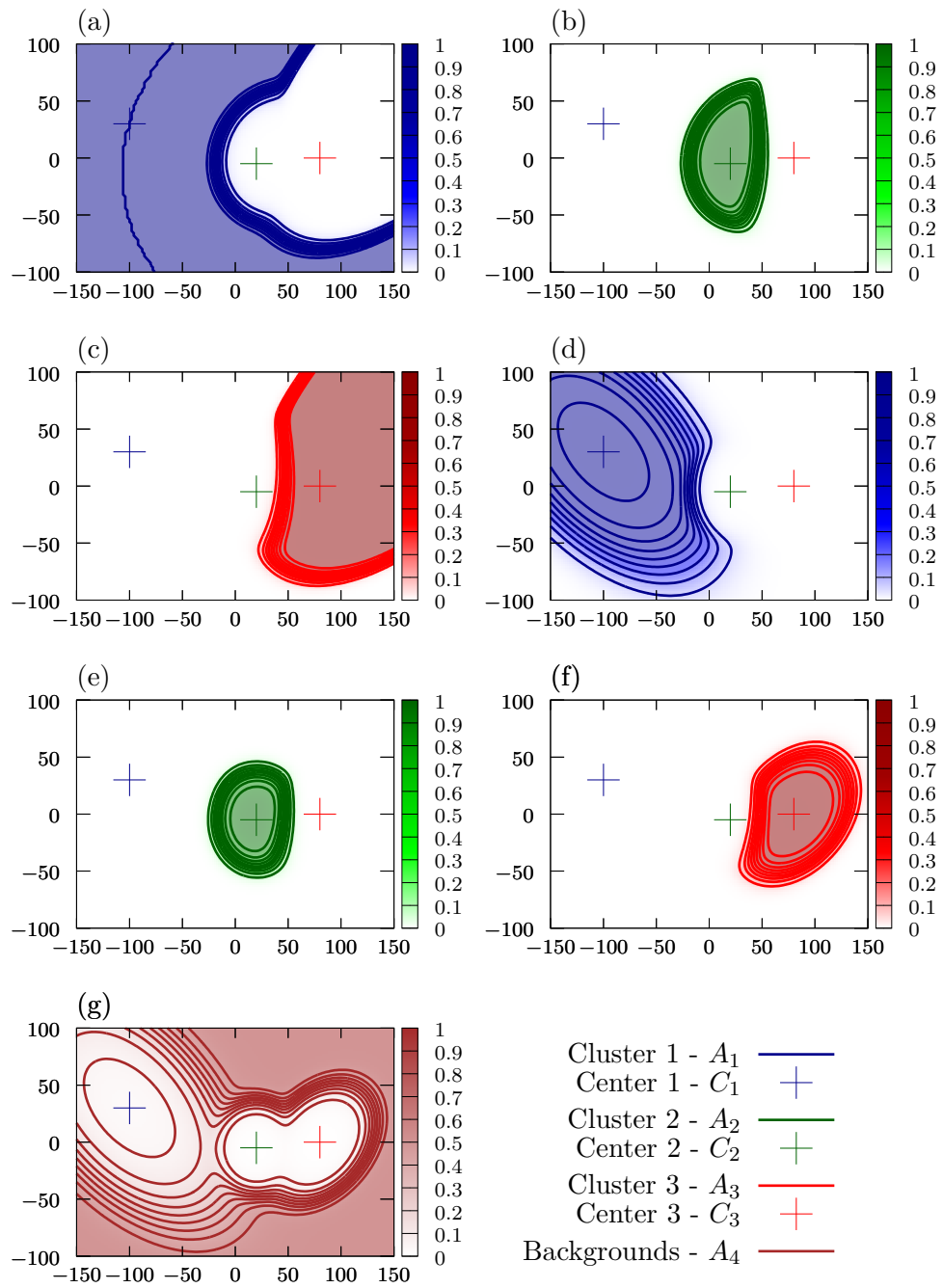
[3]Of course, the singularities must be solved.

Figure 4.6: Cluster membership functions

We multiply the relevant multivariate normal distributions by the intensity of the clusters and substitute into equation 4.35. These three membership functions are shown in Figure 4.6(a), (b), (c).

But these membership functions are somewhat strange. A large membership degree is also evident in places very distant from the centres.

How can this be improved? Article [64] can offer insight. The authors considered not only clusters but also backgrounds. We add the new fuzzy set A_4 – "Backgrounds". Its function $g_4(X)$ is constant everywhere. For example, the intensity of this fuzzy set is 30 points per minute in rectangle 300×200. Therefore, $g_4(X) = \frac{30}{300 \cdot 200}$.

We now recompute the membership functions of all A_i as shown above. The results are presented in Figure 4.6(d), (e), (f), (g). These functions better typify the degree of cluster membership.

4.4   Fuzzification

The previous two sections show how to determine primary fuzzy sets and their membership functions.

From these primary fuzzy sets, we construct floppy sets which occur in the rules defining the system. A floppy set is a set of respective primary fuzzy sets. The floppy membership function of the floppy set is the sum of the membership functions of its elements.

Some floppy sets appearing in the system rules are input whereas others are output.

To implement the system rules, we must first determine the probabilities of the input floppy sets.

Let us demonstrate how to calculate the probabilities of input floppy sets in several different situations.

4.4.1   An Exact Number Is Given

We know that the exact value of variable A is x_0.

The probability of floppy set $\boldsymbol{B}$ is calculated according to the equation:

$$R(\boldsymbol{B}|x_0) = \boldsymbol{\mu_B}(x_0). \tag{4.38}$$

Example 4.4.1. Let us continue in Example 4.3.4. What is the probability that a point with coordinates $(10, 40)$ belongs to cluster A_2?

$$R\left(\{A_2\} \mid (10, 40)\right) = \boldsymbol{\mu}_{\{A_2\}}(10, 40) = \frac{g_2(10, 40)}{\sum\limits_{j=1}^{4} g_j(10, 40)}, \qquad (4.39)$$

where

$$\boldsymbol{\mu}_{\{A_2\}}(x) = \mu_{A_2}(x). \qquad (4.40)$$

After applying the functions from Example 4.3.4, we obtain:

$$R\left(\{A_2\} \mid (10, 40)\right) = \frac{0.0001961}{0.0007277} = 0.270. \qquad (4.41)$$

We found that the ratio of the probabilities that the point belong to the cluster A_1 to A_4 is equal to the ratio of the intensity densities. It is the right result. This is why we chose intensity densities as g_i functions.

4.4.2   A Probability Distribution Is Given

We know the probability distribution of variable A. It might be the probability distribution given in Assumptions 2.1.4 and 2.1.5.

The probability of floppy set $\boldsymbol{B}$ can be calculated according to the equation:

$$R(\boldsymbol{B}) = \int_X \boldsymbol{\mu_B}(x)\, \mathrm{d}P, \qquad (4.42)$$

where the integral is the Lebesgue integral.

If A is a discrete random variable, then we can use a simpler equation:

$$R(\boldsymbol{B}) = \sum_{x_i \in X} \boldsymbol{\mu_B}(x_i) \cdot P(x_i), \qquad (4.43)$$

where $P(x_i)$ is the known probability function.

If A is a continuous random variable, then we can use the equation:

$$R(\boldsymbol{B}) = \int_X \boldsymbol{\mu_B}(x) \cdot f(x)\, \mathrm{d}x, \qquad (4.44)$$

where the integral is a Riemann integral and $f(x)$ is the known probability density.

Example 4.4.2. Let us continue in Example 4.3.3. We know the joint probability distribution of air temperature t and air pressure p. What is the probability that air pressure is high?

$$R\left(\{A_{1,3}, A_{2,3}, A_{3,3}\}\right) = \int_{p=0}^{\infty} \int_{t=-\infty}^{\infty} \boldsymbol{\mu}_{\{A_{1,3}, A_{2,3}, A_{3,3}\}}(t, p) \cdot f(t, p) \ \mathrm{d}t \ \mathrm{d}p, \qquad (4.45)$$

where

$$\boldsymbol{\mu}_{\{A_{1,3}, A_{2,3}, A_{3,3}\}}(t, p) = \mu_{A_{1,3}}(t, p) + \mu_{A_{2,3}}(t, p) + \mu_{A_{3,3}}(t, p). \qquad (4.46)$$

After applying the functions from Appendix E, we obtain:

$$R\left(\{A_{1,3}, A_{2,3}, A_{3,3}\}\right) = 0.2290. \qquad (4.47)$$

4.4.3 A Floppy Set and an Exact Value Are Given

We know that variable A is described by floppy set $\boldsymbol{C}$ and that the exact value of variable A is x_0.

The probability of floppy set $\boldsymbol{B}$ is calculated according to the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}, \boldsymbol{x_0}\right) = \frac{R\left(\boldsymbol{B} \cap \boldsymbol{C}|\boldsymbol{x_0}\right)}{R\left(\boldsymbol{C}|\boldsymbol{x_0}\right)} = \frac{\boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x_0\right)}{\boldsymbol{\mu}_{\boldsymbol{C}}\left(x_0\right)}. \qquad (4.48)$$

Example 4.4.3. We continue in Example 4.3.4. A point has coordinates $[50, 0]$. We know that the point belongs to either cluster A_2 or cluster A_3. What is the probability that the point belongs to cluster A_3?

$$R\left(\{A_3\} \mid \{A_2, A_3\}, [50, 0]\right) =$$

$$= \frac{\boldsymbol{\mu}_{\{A_3\} \cap \{A_2, A_3\}}\left([50, 0]\right)}{\boldsymbol{\mu}_{\{A_2, A_3\}}\left([50, 0]\right)} = \frac{\boldsymbol{\mu}_{\{A_3\}}\left([50, 0]\right)}{\boldsymbol{\mu}_{\{A_2, A_3\}}\left([50, 0]\right)} =$$

$$= \frac{\frac{g_3([50,0])}{\sum_{j=1}^{4} g_j([50,0])}}{\frac{g_2([50,0]) + g_3([50,0])}{\sum_{j=1}^{4} g_j([50,0])}} = \frac{g_3\left([50, 0]\right)}{g_2\left([50, 0]\right) + g_3\left([50, 0]\right)}. \qquad (4.49)$$

After applying the functions from Example 4.3.4, we obtain:

$$R\left(\{A_3\} \mid \{A_2, A_3\}, [50, 0]\right) = \frac{0.008327}{0.01227} = 0.679. \qquad (4.50)$$

4.4.4 A Floppy Set and a Probability Distribution Are Given

We know that variable A is described by floppy set $\boldsymbol{C}$, and we also know the probability distribution of variable A.

The probability of floppy set $\boldsymbol{B}$ can be calculated according to the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}\right) = \frac{R\left(\boldsymbol{B} \cap \boldsymbol{C}\right)}{R\left(\boldsymbol{C}\right)} = \frac{\int_X \boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x\right) \mathrm{d}P}{\int_X \boldsymbol{\mu}_{\boldsymbol{C}}\left(x\right) \mathrm{d}P}, \tag{4.51}$$

where the integrals are Lebesgue integrals.

If A is a discrete random variable, then we use the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}\right) = \frac{\sum_{x_i \in X} \boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x_i\right) \cdot P\left(x_i\right)}{\sum_{x_i \in X} \boldsymbol{\mu}_{\boldsymbol{C}}\left(x_i\right) \cdot P\left(x_i\right)}. \tag{4.52}$$

If A is a continuous random variable, then we use the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}\right) = \frac{\int_X \boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}{\int_X \boldsymbol{\mu}_{\boldsymbol{C}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}, \tag{4.53}$$

where the integrals are Riemann integrals.

Example 4.4.4. We continue in Example 4.3.2. We know that the water is cold. The probability density function of temperature is $f\left(x\right)$. What is the probability that the water is pleasant?

$$R\left(\{A_{0110}, A_{0010}, A_{0011}\} \,|\, \{A_{1000}, A_{0100}, A_{0110}\}\right) =$$

$$= \frac{\int_X \boldsymbol{\mu}_{\{A_{0110}, A_{0010}, A_{0011}\} \cap \{A_{1000}, A_{0100}, A_{0110}\}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}{\int_X \boldsymbol{\mu}_{\{A_{1000}, A_{0100}, A_{0110}\}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x} =$$

$$= \frac{\int_X \boldsymbol{\mu}_{\{A_{0110}\}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}{\int_X \boldsymbol{\mu}_{\{A_{1000}, A_{0100}, A_{0110}\}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}. \tag{4.54}$$

After applying the functions from Appendix D, we obtain:

$$R\left(\{A_{0110}, A_{0010}, A_{0011}\} \,|\, \{A_{1000}, A_{0100}, A_{0110}\}\right) = \frac{0.1456}{0.6844} = 0.2127. \tag{4.55}$$

4.4.5 An Interval and a Probability Density Function Are Given

We know that the value of continuous random variable A lies in the interval $[a, b]$ and that the probability density function is $f\left(x\right)$.

The probability of floppy set $\boldsymbol{B}$ can be calculated according to the equation:

$$R\left(\boldsymbol{B}|\,[a, b]\right) = \frac{R\left(\boldsymbol{B} \cap [a, b]\right)}{R\left([a, b]\right)} = \frac{\int_a^b \boldsymbol{\mu}_{\boldsymbol{B}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}{\int_a^b f\left(x\right) \mathrm{d}x}. \tag{4.56}$$

Example 4.4.5. We continue in Example 4.3.2. The water temperature is in the interval $[20, 40]$ °C. What is the probability that the water is pleasant?

$$R\left(\{A_{0110}, A_{0010}, A_{0011}|\,[20, 40]\}\right) = \frac{\int_{20}^{40} \boldsymbol{\mu}_{\{A_{0110}, A_{0010}, A_{0011}\}}(x) \cdot f(x)\,\mathrm{d}x}{\int_{20}^{40} f(x)\,\mathrm{d}x}. \qquad (4.57)$$

After applying the functions from Appendix D, we obtain:

$$R\left(\{A_{0110}, A_{0010}, A_{0011}|\,[20, 40]\}\right) = \frac{0.3944}{0.6247} = 0.6314. \qquad (4.58)$$

4.4.6 A Floppy Set, an Interval and a Probability Density Function Are Given

We know that the continuous random variable A is described by floppy set $\boldsymbol{C}$, its exact value lies within the interval $[a, b]$ and the probability density function is $f(x)$.

The probability of floppy set $\boldsymbol{B}$ is calculated according to the equation:

$$
\begin{aligned}
R\left(\boldsymbol{B}|\boldsymbol{C}, [a, b]\right) &= \frac{R\left(\boldsymbol{B} \cap \boldsymbol{C}|\,[a, b]\right)}{R\left(\boldsymbol{C}|\,[a, b]\right)} = \\
&= \frac{\frac{\int_a^b \boldsymbol{\mu}_{B \cap C}(x) \cdot f(x)\,\mathrm{d}x}{\int_a^b f(x)\,\mathrm{d}x}}{\frac{\int_a^b \boldsymbol{\mu}_{C}(x) \cdot f(x)\,\mathrm{d}x}{\int_a^b f(x)\,\mathrm{d}x}} = \\
&= \frac{\int_a^b \boldsymbol{\mu}_{B \cap C}(x) \cdot f(x)\,\mathrm{d}x}{\int_a^b \boldsymbol{\mu}_{C}(x) \cdot f(x)\,\mathrm{d}x}.
\end{aligned}
\qquad (4.59)
$$

Example 4.4.6. We continue in Example 4.3.2. The water is not warm and the water temperature is over 15 °C. The probability density function is $f(x)$. What is the probability that the water is pleasant?

$$
\begin{aligned}
&R\left(\{A_{0110}, A_{0010}, A_{0011}\} \,|\, \{A_{1000}, A_{0100}, A_{0110}, A_{0010}\}, [15, \infty)\right) = \\
&= \frac{\int_{15}^{\infty} \boldsymbol{\mu}_{\{A_{0110}, A_{0010}, A_{0011}\} \cap \{A_{1000}, A_{0100}, A_{0110}, A_{0010}\}}(x) \cdot f(x)\,\mathrm{d}x}{\int_{15}^{\infty} \boldsymbol{\mu}_{\{A_{1000}, A_{0100}, A_{0110}, A_{0010}\}}(x) \cdot f(x)\,\mathrm{d}x} = \\
&= \frac{\int_{15}^{\infty} \boldsymbol{\mu}_{\{A_{0110}, A_{0010}\}}(x) \cdot f(x)\,\mathrm{d}x}{\int_{15}^{\infty} \boldsymbol{\mu}_{\{A_{1000}, A_{0100}, A_{0110}, A_{0010}\}}(x) \cdot f(x)\,\mathrm{d}x}.
\end{aligned}
\qquad (4.60)
$$

After applying the functions from Appendix D, we obtain:

$$R\left(\{A_{0110}, A_{0010}, A_{0011}\} \,|\, \{A_{1000}, A_{0100}, A_{0110}, A_{0010}\}, [15, \infty)\right) = \frac{0.2660}{0.6461} = 0.4117. \qquad (4.61)$$

4.4.7 A Floppy Set, a Subset of X and a Probability Distribution Are Given

We know that variable A is described by floppy set $\boldsymbol{C}$, and we also know the probability distribution of variable A. The exact value of A lies in set Y, which is a subset of X and $Y \in \mathcal{A}$.

The probability of floppy set $\boldsymbol{B}$ can be calculated according to the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}, Y\right) = \frac{\int_Y \boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x\right) \mathrm{d}P}{\int_Y \boldsymbol{\mu}_{\boldsymbol{C}}\left(x\right) \mathrm{d}P}, \tag{4.62}$$

where the integrals are Lebesgue integrals.

If A is a discrete random variable, then we apply the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}, Y\right) = \frac{\sum_{x_i \in Y} \boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x_i\right) \cdot P\left(x_i\right)}{\sum_{x_i \in Y} \boldsymbol{\mu}_{\boldsymbol{C}}\left(x_i\right) \cdot P\left(x_i\right)}. \tag{4.63}$$

If A is a continuous random variable, then we apply the equation:

$$R\left(\boldsymbol{B}|\boldsymbol{C}, Y\right) = \frac{\int_Y \boldsymbol{\mu}_{\boldsymbol{B} \cap \boldsymbol{C}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}{\int_Y \boldsymbol{\mu}_{\boldsymbol{C}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x}, \tag{4.64}$$

where the integrals are Riemann integrals.

Example 4.4.7. Let us continue in Example 4.3.3. It is warm. The joint probability density function is $f\left(t, p\right)$. The set Y is given:

$$t < 32\,°\mathrm{C}, \qquad p > 952\,\mathrm{hPa}, \qquad \frac{p}{t} \le 34\,\frac{\mathrm{hPa}}{°\mathrm{C}}. \tag{4.65}$$

What is the probability that the air pressure is low?

$$
\begin{aligned}
R\left(\{A_{1,1}, A_{2,1}, A_{3,1}\} \,|\, \{A_{3,1}, A_{3,2}, A_{3,3}\}, Y\right) &= \\
= \frac{\int_{t=\frac{952}{34}}^{32} \int_{p=952}^{34 \cdot t} \boldsymbol{\mu}_{\{A_{1,1}, A_{2,1}, A_{3,1}\} \cap \{A_{3,1}, A_{3,2}, A_{3,3}\}}\left(t, p\right) \cdot f\left(t, p\right) \,\mathrm{d}p\,\mathrm{d}t}{\int_{t=\frac{952}{34}}^{32} \int_{p=952}^{34 \cdot t} \boldsymbol{\mu}_{\{A_{3,1}, A_{3,2}, A_{3,3}\}}\left(t, p\right) \cdot f\left(t, p\right) \,\mathrm{d}p\,\mathrm{d}t} &= \\
= \frac{\int_{t=28}^{32} \int_{p=952}^{34 \cdot t} \boldsymbol{\mu}_{\{A_{3,1}\}}\left(t, p\right) \cdot f\left(t, p\right) \,\mathrm{d}p\,\mathrm{d}t}{\int_{t=28}^{32} \int_{p=952}^{34 \cdot t} \boldsymbol{\mu}_{\{A_{3,1}, A_{3,2}, A_{3,3}\}}\left(t, p\right) \cdot f\left(t, p\right) \,\mathrm{d}p\,\mathrm{d}t}. &
\end{aligned}
\tag{4.66}
$$

After applying the functions from Appendix E, we obtain:

$$R\left(\{A_{1,1}, A_{2,1}, A_{3,1}\} \,|\, \{A_{3,1}, A_{3,2}, A_{3,3}\}, Y\right) = \frac{0.002085}{0.02828} = 0.0737. \tag{4.67}$$

4.5 System Rules

4.5.1 System Rules as a Conditional Probability Distribution

In the previous section, we calculated the probabilities of input floppy sets. Using system rules, we now calculate the probabilities of output floppy sets. The system rules will therefore take the form of a conditional probability distribution. This method enables more options than strict IF-THEN rules.

The probabilities of output floppy sets can be calculated as follows:

$$R^P\left(\boldsymbol{B_i}\right) = \sum_j R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right) \cdot R\left(\boldsymbol{A_j}\right), \tag{4.68}$$

where $\boldsymbol{A_j}$ are input floppy sets and $\boldsymbol{B_i}$ are output floppy sets. $R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right)$ is the conditional probability distribution which describes the system rules. Probabilities $R\left(\boldsymbol{A_j}\right)$ are the probabilities calculated in the previous section 4.4.

Note that now we have two different probabilities of output floppy sets. A priori probability is given by Definition 2.1.3.

In Equation 4.68, an a posteriori probability which depends on the probabilities of the input floppy sets was introduced. We denote the a posteriori probability by the upper index P.

Example 4.5.1. Let us continue in Example 4.3.4. We suppose that points in the plane are patients and that individual clusters are different types of patients.

A system can be described according to the following rules:

1. If we give medicine to a patient from "cluster 1", their condition will be good.

2. If we give medicine to a patient from "cluster 2", their condition will be good in 70% or fair in 30% of cases.

3. If we give medicine to a patient from "cluster 3", their condition will be serious.

4. If we give medicine to a patient from "background", their condition will be good in 50%, fair in 30%, or serious in 20% of cases.

We thus have four input floppy sets ($\boldsymbol{A_1}$ = cluster 1, $\boldsymbol{A_2}$ = cluster 2, $\boldsymbol{A_3}$ = cluster 3, $\boldsymbol{A_4}$ = background) and three output floppy sets ($\boldsymbol{B_1}$ = good, $\boldsymbol{B_2}$ = fair, $\boldsymbol{B_3}$ = serious). The matrix of the conditional probability distribution is:

$$R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right) = \begin{pmatrix} 1 & 0.7 & 0 & 0.5 \\ 0 & 0.3 & 0 & 0.3 \\ 0 & 0 & 1 & 0.2 \end{pmatrix}. \tag{4.69}$$

Note that the sum in each column is equal to one.

We have a patient. The probabilities that he or she belongs to individual clusters are as follows:

$$R\left(\boldsymbol{A_1}\right) = 0.2, \quad R\left(\boldsymbol{A_2}\right) = 0.5, \quad R\left(\boldsymbol{A_3}\right) = 0.2, \quad R\left(\boldsymbol{A_4}\right) = 0.1. \tag{4.70}$$

We calculate the a posteriori probabilities of the output floppy sets:

$$R^P\left(\boldsymbol{B_i}\right) = \sum_{j=1}^{4} R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right) \cdot R\left(\boldsymbol{A_j}\right) =$$

$$= \begin{pmatrix} 1 & 0.7 & 0 & 0.5 \\ 0 & 0.3 & 0 & 0.3 \\ 0 & 0 & 1 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} 0.2 \\ 0.5 \\ 0.2 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 0.60 \\ 0.18 \\ 0.22 \end{pmatrix}. \tag{4.71}$$

4.5.2 Estimation of System Rules Parameters

Values $R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right)$ are sometimes known, but often it is necessary to determine them from data. How should we proceed?

Our model, which consists of one family of input floppy sets, one family of output floppy sets, and the conditional probability distribution, is equivalent to a discrete model with one input and one output variable. The values $R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right)$ can therefore be estimated as follows [46, p. 14]:

$$R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right) = \frac{\kappa_{i,j}\left(n\right)}{\sum_{\forall k} \kappa_{k,j}\left(n\right)}, \tag{4.72}$$

where $\kappa_{i,j}\left(n\right)$ are statistics obtained from n data vectors. These statistics describe how many times a particular combination of input and output floppy sets has occurred.

Example 4.5.2. We continue in Example 4.3.4. Again, we suppose that points in the plane are patients and that individual clusters are different types of patients, yet we do not know the values $R(\boldsymbol{B}_i|\boldsymbol{A}_j)$.

Table 4.2 presents data on five patients.

Table 4.2: Patients data

Patient	Input floppy sets				Output floppy sets		
	Cluster 1	Cluster 2	Cluster 3	Backgr.	Good	Fair	Serious
1		1			1		
2	1					1	
3			1			0.5	0.5
4	0.8			0.2	1		
5		0.6		0.4		0.7	0.3

The first data vector states that input floppy set $\boldsymbol{A_2}$ occurred with output floppy set $\boldsymbol{B_1}$. We can therefore determine $\kappa_{i,j}(1)$:

$$\kappa_{i,j}(1) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4.73}$$

Then, we add the second patient's data:

$$\kappa_{i,j}(2) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4.74}$$

The third patient's data indicate that the output floppy sets "Fair" and "Serious" have a membership degree of 0.5. We add these data:

$$\kappa_{i,j}(3) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 \end{pmatrix}. \tag{4.75}$$

Similarly, for the fourth patient, we see non-integer degrees of membership in the input floppy sets:

$$\kappa_{i,j}(4) = \begin{pmatrix} 0.8 & 1 & 0 & 0.2 \\ 1 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 \end{pmatrix}. \tag{4.76}$$

A problem arises when we discover non-integer membership degrees in the input and output floppy sets simultaneously. We do not know exactly how we should increase individual statistics $\kappa_{i,j}$.

We may exclude such data or proceed as if the input and output data were independent. We selected this (not entirely correct) method in this example:

$$\kappa_{i,j}(5) = \begin{pmatrix} 0.8 & 1 & 0 & 0.2 \\ 1 & 0.42 & 0.5 & 0.28 \\ 0 & 0.18 & 0.5 & 0.12 \end{pmatrix}. \tag{4.77}$$

Now we calculate $R(\boldsymbol{B_i}|\boldsymbol{A_j})$ according to equation 4.72:

$$R(\boldsymbol{B_i}|\boldsymbol{A_j}) = \begin{pmatrix} 0.4444 & 0.625 & 0 & 0.3333 \\ 0.5556 & 0.2625 & 0.5 & 0.4667 \\ 0 & 0.1125 & 0.5 & 0.2 \end{pmatrix}. \tag{4.78}$$

Note that the sum in each column is equal to one.

The conditional probability distribution $R(\boldsymbol{B_i}|\boldsymbol{A_j})$ is now determined only by the data for the five patients in Table 4.2. However, we may have some other information in the form of an expert estimate or some other research. If we want to include this information in our estimate of $R(\boldsymbol{B_i}|\boldsymbol{A_j})$, we create an a priori statistic $\kappa_{i,j}(0)$. Then we add the data from Table 4.2 and proceed according to equation 4.72.

The sum in each column $\kappa_{i,j}(0)$ indicates how many fictitious patients the estimate corresponds to.

For example, an expert's estimate might appear as follows:

$$\kappa_{i,j}(0) = \begin{pmatrix} 10 & 700 & 0 & 0 \\ 0 & 300 & 0 & 0 \\ 0 & 0 & 10 & 0 \end{pmatrix}. \tag{4.79}$$

The sum in the first and third columns is ten. This means that the expert is not very confident with the patients from clusters 1 and 3. The sum in the second column is 1000. The expert is very confident with the patients from cluster 2. The sum in the fourth column is zero. This means that the expert does not estimate at all the conditions with patients from the "background".

If we include the expert estimate $\kappa_{i,j}(0)$, we obtain an estimate of $R(\boldsymbol{B_i}|\boldsymbol{A_j})$ as follows:

$$R(\boldsymbol{B_i}|\boldsymbol{A_j}) = \begin{pmatrix} 0.9153 & 0.6999 & 0 & 0.3333 \\ 0.0847 & 0.2999 & 0.0455 & 0.4667 \\ 0 & 0.0002 & 0.9545 & 0.2 \end{pmatrix}. \tag{4.80}$$

4.6 Defuzzification

4.6.1 Calculation of A Posteriori Probability Distributions

In Section 4.5.1 we calculated the a posteriori probabilities of output floppy sets. This form of output data can be absolutely satisfactory in some situations. For example: It will be 80% sunny, 15% cloudy, and 5% rainy.

Sometimes we want to know which floppy set has the highest probability. For example: It will probably be sunny.

Sometimes we are interested in point estimation. For example: It will be 28 °C. These point estimations as either a mean or median can be computed from an a posteriori probability distribution.

The aim in this section is to calculate an a posteriori probability distribution from the a posteriori probabilities of the output floppy sets. The point or interval estimations can then be calculated as usual.

Let $\{\boldsymbol{B_j}\}$ be a family of pairwise disjoint floppy sets whose union is the entire sample space.

Then, the a posteriori probability distribution can be computed with the law of total probability and Bayes' theorem.

Variant for discrete case:

$$
\begin{aligned}
R^P(x_i) &= \sum_j R\left(x_i|\boldsymbol{B_j}\right) \cdot R^P(\boldsymbol{B_j}) = \\
&= \sum_j \frac{R\left(\boldsymbol{B_j}|x_i\right) \cdot R\left(x_i\right)}{R\left(\boldsymbol{B_j}\right)} \cdot R^P(\boldsymbol{B_j}) = \\
&= \sum_j \frac{\boldsymbol{\mu_{B_j}}\left(x_i\right) \cdot P\left(x_i\right)}{\sum_{x_i \in X} \boldsymbol{\mu_{B_j}}\left(x_i\right) \cdot P\left(x_i\right)} \cdot R^P(\boldsymbol{B_j}).
\end{aligned}
\tag{4.81}
$$

Variant for continuous case:

$$
\begin{aligned}
f^P(x) &= \sum_j f\left(x|\boldsymbol{B_j}\right) \cdot R^P(\boldsymbol{B_j}) = \\
&= \sum_j \frac{R\left(\boldsymbol{B_j}|x\right) \cdot f\left(x\right)}{R\left(\boldsymbol{B_j}\right)} \cdot R^P(\boldsymbol{B_j}) = \\
&= \sum_j \frac{\boldsymbol{\mu_{B_j}}\left(x\right) \cdot f\left(x\right)}{\int_{x \in X} \boldsymbol{\mu_{B_j}}\left(x\right) \cdot f\left(x\right) \mathrm{d}x} \cdot R^P(\boldsymbol{B_j}).
\end{aligned}
\tag{4.82}
$$

Example 4.6.1. Let us have a system which heats water with the aid of the sun.

The weather is described according to three primary fuzzy sets: "sunny"", "cloudy", "rainy". Water temperature is described as in Example 4.3.2. The system is defined by the following rules:

1. If it is sunny, then the water is cold in 10% or warm in 50% of cases.

2. If it is cloudy, then the water is cold in 20% or warm in 30% of cases.

3. If it is rainy, then the water is cold in 40% or warm in 10% of cases.

Today it is 50% sunny and 50% cloudy. Calculate the probability distribution for the water temperature.

We determine three output floppy sets:

- cold $-$ $\boldsymbol{B_1} = \{A_{1000}, A_{0100}, A_{0110}\}$,

- not cold and not warm $-$ $\boldsymbol{B_2} = \{A_{0010}\}$,

- warm $-$ $\boldsymbol{B_3} = \{A_{0011}, A_{0001}\}$.

We create matrix $R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right)$:

$$R\left(\boldsymbol{B_i}|\boldsymbol{A_j}\right) = \begin{pmatrix} 0.1 & 0.2 & 0.4 \\ 0.4 & 0.5 & 0.5 \\ 0.5 & 0.3 & 0.1 \end{pmatrix}. \tag{4.83}$$

We compute the a posteriori probabilities of output floppy sets $\boldsymbol{B_i}$:

$$R^P\left(\boldsymbol{B_i}\right) = \begin{pmatrix} 0.1 & 0.2 & 0.4 \\ 0.4 & 0.5 & 0.5 \\ 0.5 & 0.3 & 0.1 \end{pmatrix} \cdot \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.15 \\ 0.45 \\ 0.40 \end{pmatrix}. \tag{4.84}$$

We calculate the a posteriori probability density function of temperature:

$$\begin{aligned} f^P(x) &= \frac{\boldsymbol{\mu}_{B_1}(x) \cdot f(x)}{\int_{x \in X} \boldsymbol{\mu}_{B_1}(x) \cdot f(x)\,\mathrm{d}x} \cdot 0.15 + \\ &+ \frac{\boldsymbol{\mu}_{B_2}(x) \cdot f(x)}{\int_{x \in X} \boldsymbol{\mu}_{B_2}(x) \cdot f(x)\,\mathrm{d}x} \cdot 0.45 + \\ &+ \frac{\boldsymbol{\mu}_{B_3}(x) \cdot f(x)}{\int_{x \in X} \boldsymbol{\mu}_{B_3}(x) \cdot f(x)\,\mathrm{d}x} \cdot 0.40. \end{aligned} \tag{4.85}$$
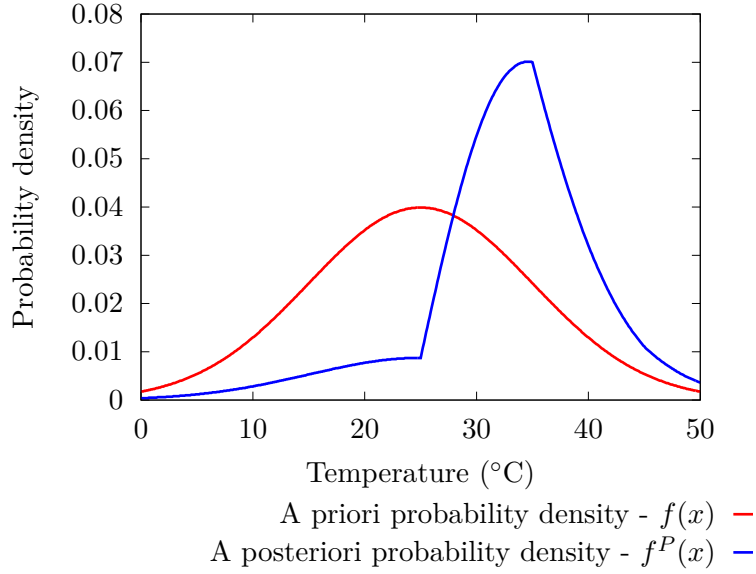
Figure 4.7: Comparison of a priori and a posteriori probability densities

After applying the functions from Appendix D, we obtain:

$$f^P(x) = [0.2192 \cdot \boldsymbol{\mu_{B_1}}(x) + 3.7375 \cdot \boldsymbol{\mu_{B_2}}(x) + 2.0489 \cdot \boldsymbol{\mu_{B_3}}(x)] \cdot f(x). \quad (4.86)$$

A comparison of a priori and a posteriori probability densities of water temperature is presented in Figure 4.7.

4.6.2 Calculation of Mean Values

The mean value can be calculated from the a posteriori probability distribution in the standard manner.

For discrete cases, we use the equation:

$$\mathrm{E}(x) = \sum_{x_i \in X} x_i \cdot R^P(x_i) \quad (4.87)$$

For continuous cases, we use the equation:

$$\mathrm{E}(x) = \int_{x \in X} x \cdot f^P(x)\,\mathrm{d}x. \quad (4.88)$$

Sometimes, the a posteriori probability distribution is barely determinable. To compute it, we must know the membership functions of all output floppy sets and the a priori probability distribution. We do not know these required properties, especially regarding subjective variables such as "the need for something".

In these cases, we cannot compute the a posteriori probability distribution, but we can calculate the mean value using equation 3.5. In that equation, we need to know or estimate only the mean values of all individual output floppy sets.[4]

Example 4.6.2. We continue with Example 4.6.1. Calculate the mean value to make a point estimate of the water temperature.

We compute the mean value of x from the a posteriori probability density function:

$$\mathrm{E}\left(x\right) = \int_{x \in X} x \cdot f^P(x)\,\mathrm{d}x. \tag{4.89}$$

After applying the functions from Appendix D, we obtain:

$$\mathrm{E}\left(x\right) = 32.98. \tag{4.90}$$

If we do not know the a posteriori probability density, we calculate the mean value using the mean values of the individual output floppy sets.

Assume that the mean value of cold water is 20.01 °C, the mean value of not cold and not warm water is 33.53 °C, and the mean value of warm water is 37.22 °C. (We calculated these values by applying the functions from Appendix D in equation 3.3.)

We proceed according to formula 3.5:

$$\mathrm{E}\left(x\right) = \begin{pmatrix} 20.01 \\ 33.53 \\ 37.22 \end{pmatrix}^T \cdot \begin{pmatrix} 0.15 \\ 0.45 \\ 0.40 \end{pmatrix} = 32.98. \tag{4.91}$$

The resulting mean values obtained in both methods were, of course, the same.

4.7 Control of Floppy Systems

4.7.1 General Approach

To control floppy systems, we can successfully use the strategies developed to control systems in general. Much quality literature is available on this topic [4, 28, 39].

---

[4]Definition 3.1.1.

This section presents a procedure based on penalising individual states of the system.

A penalty is assigned to each combination of control and output variables.

Floppy systems are stochastic. We therefore do not know the exact output which will be achieved. The resulting penalty is therefore also unknown.

However, we know the probability distribution for the individual outputs. This allows us to calculate the mean value of the penalty for each variant of control.

We then select a control variant which has the smallest mean value for the penalty.

4.7.2 Static Discrete Systems

A system specified by one family of input floppy sets $\{A_j\}$, one family of output floppy sets $\{B_i\}$, and the conditional probability distribution matrix $\{B_i|A_j\}$ corresponds to a discrete static model.

If we add a discrete control, we can use the standard procedure for controlling discrete models.

Example 4.7.1. We continue with Example 4.5.1.

We add the control variables: $u_1 = $ "We administer the medicine." $u_2 = $ "We do not administer the medicine."

We add the following system rules:

5. If we do not administer the medicine to a patient from "cluster 1", their condition will be good in 50% of cases or fair in 50%.

6. If we do not administer the medicine to a patient from "cluster 2", their condition will be good in 20% of cases, fair in 40%, or serious in 40%.

7. If we do not administer the medicine to a patient from "cluster 3", their condition will be good in 10% of cases, fair in 70%, or serious in 20%.

8. If we do not administer the medicine to a patient from "background", their condition will be good in 30% of cases, fair in 40%, or serious in 30%.

We write all rules into matrices:

$$R\left(\boldsymbol{B_i}|\boldsymbol{A_j}, u_1\right) = \begin{pmatrix} 1 & 0.7 & 0 & 0.5 \\ 0 & 0.3 & 0 & 0.3 \\ 0 & 0 & 1 & 0.2 \end{pmatrix}, \tag{4.92}$$

$$R\left(\boldsymbol{B_i}|\boldsymbol{A_j}, u_2\right) = \begin{pmatrix} 0.5 & 0.2 & 0.1 & 0.3 \\ 0.5 & 0.4 & 0.7 & 0.4 \\ 0 & 0.4 & 0.2 & 0.3 \end{pmatrix}. \tag{4.93}$$

The probabilities that our patient belongs to individual clusters are as follows:

$$R\left(\boldsymbol{A_1}\right) = 0.2, \qquad R\left(\boldsymbol{A_2}\right) = 0.5, \qquad R\left(\boldsymbol{A_3}\right) = 0.2, \qquad R\left(\boldsymbol{A_4}\right) = 0.1. \tag{4.94}$$

We calculate the a posteriori probabilities of output floppy sets for individual control variants:

$$R^P\left(\boldsymbol{B_i}, u_1\right) = \begin{pmatrix} 1 & 0.7 & 0 & 0.5 \\ 0 & 0.3 & 0 & 0.3 \\ 0 & 0 & 1 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} 0.2 \\ 0.5 \\ 0.2 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 0.60 \\ 0.18 \\ 0.22 \end{pmatrix}, \tag{4.95}$$

$$R^P\left(\boldsymbol{B_i}, u_2\right) = \begin{pmatrix} 0.5 & 0.2 & 0.1 & 0.3 \\ 0.5 & 0.4 & 0.7 & 0.4 \\ 0 & 0.4 & 0.2 & 0.3 \end{pmatrix} \cdot \begin{pmatrix} 0.2 \\ 0.5 \\ 0.2 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0.48 \\ 0.27 \end{pmatrix}. \tag{4.96}$$

We determine the penalty for individual output floppy sets and control variables.

Good condition in a patient is not penalised, fair condition is penalised with 10 points, and serious condition with 50 points. If we administer the medicine, we add 3 points (price, side effects).

We therefore create the following penalty vectors for output floppy sets and control variants:

$$k\left(u_1\right) = \begin{pmatrix} 3 \\ 13 \\ 53 \end{pmatrix}, \qquad\qquad k\left(u_2\right) = \begin{pmatrix} 0 \\ 10 \\ 50 \end{pmatrix}. \tag{4.97}$$

We calculate the mean value of the penalty for both control variants:

$$K\left(u_1\right) = k\left(u_1\right)^T \cdot R^P\left(\boldsymbol{B_i}, u_1\right) = \begin{pmatrix} 3 \\ 13 \\ 53 \end{pmatrix}^T \cdot \begin{pmatrix} 0.60 \\ 0.18 \\ 0.22 \end{pmatrix} = 15.8, \qquad (4.98)$$

$$K\left(u_2\right) = k\left(u_2\right)^T \cdot R^P\left(\boldsymbol{B_i}, u_2\right) = \begin{pmatrix} 0 \\ 10 \\ 50 \end{pmatrix}^T \cdot \begin{pmatrix} 0.25 \\ 0.48 \\ 0.27 \end{pmatrix} = 18.3. \qquad (4.99)$$

If we administer the medicine, the mean value of the penalty is lower. We select this variant.

4.7.3 Dynamic Discrete Systems

The previous example was static. In this example, the current output variable depends on the previous output variable and the previous discrete control.

Because the outputs in the individual steps depend on each other, the control in one step also affects the other steps. Therefore, if we want to penalise individual control variants, we must penalise the consequences of control not only in the current steps but also any future steps.

The n-step penalisation of individual control variants is therefore:

$$K\left(u_{v_0, v_1, v_2, \dots v_n}\right) = \sum_{m=1}^{n} k\left(u_{v_{m-1}}\right)^T \cdot R_m^P\left(\boldsymbol{B_i}\right), \qquad (4.100)$$

where $u_{v_{m-1}}$ is v-th variant of the control in $(m-1)$-th step, $u_{v_0, v_1, v_2, \dots v_n}$ is the control sequence u_{v_0}, u_{v_1}, u_{v_2}, $\dots u_{v_n}$, $R_m^P\left(\boldsymbol{B_i}\right)$ is the vector of the probabilities of output floppy sets in m-th step, and $k\left(u_{v_{m-1}}\right)$ is the penalty vector dependent on the control $u_{v_{m-1}}$.

Assume that the probability of the output floppy sets in one round depends on the probability of output floppy sets in the previous round, as follows:

$$R_m^P\left(\boldsymbol{B_i}\right) = R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_{m-1}}\right) \cdot R_{m-1}^P\left(\boldsymbol{B_j}\right), \qquad (4.101)$$

where $R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_{m-1}}\right)$ is a matrix dependent on the $m-1$-th control.

We then substitute into equation 4.100:

$$
\begin{aligned}
K\left(u_{v_0,v_1,v_2,\ldots v_n}\right) &= k\left(u_{v_0}\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_0}\right) \cdot R_0^P\left(\boldsymbol{B_j}\right) + \\
&+ k\left(u_{v_1}\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_1}\right) \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_0}\right) \cdot R_0^P\left(\boldsymbol{B_j}\right) + \\
&+ \ldots = \\
&= \left[k\left(u_{v_0}\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_0}\right) + \right. \\
&+ k\left(u_{v_1}\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_1}\right) \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_0}\right) + \\
&+ \left. \ldots \right] \cdot R_0^P\left(\boldsymbol{B_j}\right).
\end{aligned}
\tag{4.102}
$$

The expression in square brackets is denoted $L\left(u_{v_0,v_1,v_2,\ldots v_n}\right)$:

$$
K\left(u_{v_0,v_1,v_2,\ldots v_n}\right) = L\left(u_{v_0,v_1,v_2,\ldots v_n}\right) \cdot R_0^P\left(\boldsymbol{B_j}\right).
\tag{4.103}
$$

We calculate $K\left(u_{v_0,v_1,v_2,\ldots v_n}\right)$ for all combinations of control and select the variant with the least penalty. From this variant, we use only the first control u_{v_0}.

The disadvantage of this control method is that we must calculate the penalty for all control variants. However, the expressions $L\left(u_{v_0,v_1,v_2,\ldots v_n}\right)$ need only be computed once. If we save the results, we can use them without change in any subsequent control steps.

Example 4.7.2. We can use one of two medicines to treat a patient: $u_1 =$ "the first medicine", $u_2 =$ "the second medicine."

The patient's condition is described according to three floppy sets: $\boldsymbol{B_1} =$ "good", $\boldsymbol{B_2} =$ "fair", $\boldsymbol{B_3} =$ "serious".

Each day, we determine the patient's condition by determining the probabilities of individual floppy sets $\boldsymbol{B_i}$.

We model the progress of the patient's condition in the following manner:

$$
R_m^P\left(\boldsymbol{B_i}\right) = R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_{v_{m-1}}\right) \cdot R_{m-1}^P\left(\boldsymbol{B_j}\right)
\tag{4.104}
$$

The system rules are given by the following matrices:

$$
R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_1\right) = \begin{pmatrix} 0.9 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0 & 0.1 & 0.8 \end{pmatrix}, \quad R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_2\right) = \begin{pmatrix} 0.8 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.5 \\ 0 & 0.1 & 0.4 \end{pmatrix}.
\tag{4.105}
$$

The penalty is given by the following vectors:

$$
k\left(u_1\right) = \begin{pmatrix} 3 \\ 13 \\ 53 \end{pmatrix}, \qquad k\left(u_2\right) = \begin{pmatrix} 5 \\ 15 \\ 55 \end{pmatrix}.
\tag{4.106}
$$

We perform a three-step control and decide which medicine to administer to a patient whose condition was rated today by two doctors as serious and one doctor as fair.

We know the current vector $R_0^P (\boldsymbol{B_j})$:

$$R_0^P (\boldsymbol{B_j}) = \begin{pmatrix} 0 \\ 1/3 \\ 2/3 \end{pmatrix}. \tag{4.107}$$

Now, we calculate $L\left(u_{v_0,v_1,v_2}\right)$ and $K\left(u_{v_0,v_1,v_2}\right)$ for all control variants. For example, for a three-step control $u = (u_{1,2,2})$, we calculate:

$$
\begin{aligned}
L\left(u_{1,2,2}\right) &= \\
&= k\left(u_0\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_0\right) + \\
&\quad + k\left(u_1\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_1\right) \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_0\right) + \\
&\quad + k\left(u_2\right)^T \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_2\right) \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_1\right) \cdot R\left(\boldsymbol{B_i}|\boldsymbol{B_j}, u_0\right) =
\end{aligned}
$$

$$
\begin{aligned}
&= \begin{pmatrix} 3 \\ 13 \\ 53 \end{pmatrix}^T \cdot \begin{pmatrix} 0.9 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0 & 0.1 & 0.8 \end{pmatrix} + \\
&\quad + \begin{pmatrix} 5 \\ 15 \\ 55 \end{pmatrix}^T \cdot \begin{pmatrix} 0.8 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.5 \\ 0 & 0.1 & 0.4 \end{pmatrix} \cdot \begin{pmatrix} 0.9 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0 & 0.1 & 0.8 \end{pmatrix} = \\
&\quad + \begin{pmatrix} 5 \\ 15 \\ 55 \end{pmatrix}^T \cdot \begin{pmatrix} 0.8 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.5 \\ 0 & 0.1 & 0.4 \end{pmatrix} \cdot \begin{pmatrix} 0.8 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.5 \\ 0 & 0.1 & 0.4 \end{pmatrix} \cdot \begin{pmatrix} 0.9 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0 & 0.1 & 0.8 \end{pmatrix} = \\
&= \begin{pmatrix} 20.87 & 31.60 & 88.55 \end{pmatrix}.
\end{aligned}
\tag{4.108}
$$

$$K\left(u_{1,2,2}\right) = \begin{pmatrix} 20.87 & 31.60 & 88.55 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1/3 \\ 2/3 \end{pmatrix} = 69.57. \tag{4.109}$$

The results are given in Table 4.3.

We obtained the smallest mean value of the penalty for the control sequence $u_{2,2,1}$. Today, therefore, we administer the second medicine.

When we check the patient's condition tomorrow, we perform the entire procedure again. We already have the calculated values $L\left(u_{v_0,v_1,v_2}\right)$. It is therefore enough

Table 4.3: Calculation of penalties for individual control sequences

Control sequence	$L\left(u_{v_0,v_1,v_2}\right)$	$K\left(u_{v_0,v_1,v_2}\right)$
$u_{1,1,1}$	(13.66 27.92 111.26)	83.48
$u_{1,1,2}$	(16.71 29.46 103.43)	78.77
$u_{1,2,1}$	(17.58 29.04 90.22)	69.83
$u_{1,2,2}$	(20.87 31.60 88.55)	69.57
$u_{2,1,1}$	(17.72 36.10 72.70)	60.50
$u_{2,1,2}$	(20.62 37.19 69.71)	58.87
$u_{2,2,1}$	(21.96 38.18 68.42)	56.34
$u_{2,2,2}$	(25.14 40.41 65.79)	57.33

to calculate the values $K\left(u_{v_0,v_1,v_2}\right)$. For example, if all doctors agree that a patient's condition is fair, then we determine the smallest penalty for variant $u_{1,1,1}$ and administer the first medicine to the patient.

Chapter 5

Comparison of Floppy Logic to Other Theories

5.1 Floppy Logic and Kolmogorov Probability Theory

The relationship between floppy logic and probability theory is very simple. Theorems 2.1.1 and 2.2.1 state that both basic floppy logic and generalised floppy logic are models of probability theory.

We can therefore use all the constructions and tools of Kolmogorov probability theory in floppy logic. An example is the floppy set's mean value. This was introduced in section 3.1 and used in section 4.6.2.

We can translate the concepts of floppy logic into the language of probability theory.

The set of all primary fuzzy sets is the sample space.

A floppy set is a probability event.

The floppy membership function of floppy set $\boldsymbol{B}$ is a conditional probability:

$$\boldsymbol{\mu_B}\left(x\right) = R\left(\boldsymbol{B}|x\right), \tag{5.1}$$

where x is an element of the domain.

From the assumptions of floppy logic (Section 2.1.1), this relation follows for all $x \in X$, except for a null set. (Section 2.2.2.) The natural choice is to accept this relation for all $x \in X$.

The case is similar to a probability density function in probability theory. This function must be non-negative throughout the domain, except for a null set. However, it is usually selected non-negatively everywhere.

The number

$$R\left(\boldsymbol{B}\right) = \int_X \boldsymbol{\mu_B}\left(x\right) \mathrm{d}P \tag{5.2}$$

is the probability of event $\boldsymbol{B}$.

Interestingly, in floppy logic, the probability is derived from the floppy membership function, which is interpreted as the conditional probability. In floppy logic, therefore, the fundamental concept is conditional probability, not probability, as in Kolmogorov probability theory. It would therefore be interesting to compare floppy logic with alternative probability theories (e.g. [50, 57]) in which conditional probability is the central concept.

5.2 Floppy Logic and Standard Bivalent Logic

The relationship between floppy logic and standard Boolean logic is discussed in Theorem 2.3.1. This theorem states that every two statements which are equivalent in standard Boolean logic are also equivalent in floppy logic.

Therefore, distributivity, idempotence, the law of excluded middle and all the properties listed in Table 1.1 are also preserved in floppy logic.

It is very surprising that some multi-valued logic exists with these properties.

For the reasons above, floppy logic can be considered a multi-valued generalisation of standard propositional logic.

If we accept that floppy membership functions are the generalisations of quantifiers (Consequence 2.3.9), we may consider floppy logic as a generalisation of standard predicate logic.

5.3 Floppy Logic and Fuzzy Logic

5.3.1 The Basic Differences

Floppy logic and fuzzy logic have in common that they both work with fuzzy sets and their membership functions.

The basic difference between floppy logic and fuzzy logic is that events are modelled according to fuzzy sets in fuzzy logic whereas in floppy logic, they are modelled according to floppy sets. A floppy set is a set of primary fuzzy sets. Floppy sets are introduced in Sections 2.1.1 and 2.2.1.

This shift from elements to sets is not a new concept. A similar step was made by Andrey Nikolaevich Kolmogorov in his probability theory [35], where he assigned probabilities not to elements but to subsets of the sample space.

The second difference is that fuzzy logic is truth-functional, while floppy logic is not. More about this is examined in Section 5.3.4.

Another important difference is that floppy logic, unlike fuzzy logic, preserves the equivalence of statements. More about this is written in detail in Section 2.3.

The great advantage of floppy logic is that it is fully compatible with Kolmogorov probability theory. See Sections 2.1.2 to 2.2.2.

Another advantage of floppy logic is that it consistently works with data in the form of exact values, probability distributions, crisp sets and floppy sets simultaneously. See Chapter 4.

5.3.2   Intersections and Unions in Fuzzy Logic

In his first article on fuzzy logic [68], Zadeh already discusses two methods of generalising the intersection and union of sets. The first pair of these functions is the minimum and maximum of membership functions (Gödel t-norm and t-conorm):

$$
\begin{aligned}
\mu_{A\cap B} &= \min\left\{\mu_A, \mu_B\right\}, \\
\mu_{A\cup B} &= \max\left\{\mu_A, \mu_B\right\}.
\end{aligned}
$$

The second pair is now called the algebraic product (= product t-norm) and algebraic sum (= product t-conorm):

$$
\begin{aligned}
\mu_{A\cap B} &= \mu_A \cdot \mu_B, \\
\mu_{A\cup B} &= \mu_A + \mu_B - \mu_A \cdot \mu_B.
\end{aligned}
$$

The third pair of these functions was introduced by Robin Giles in 1976 [24] (bounded product = Łukasiewict t-norm) and Lotfi Zadeh in 1975 [66] (bounded sum = Łukasiewicz t-conorm):

$$
\begin{aligned}
\mu_{A\cap B} &= \max\left\{0, \mu_A + \mu_B - 1\right\}, \\
\mu_{A\cup B} &= \min\left\{1, \mu_A + \mu_B\right\}.
\end{aligned}
$$

The next pair, drastic sum and drastic product, was introduced by Didier Dubois in 1979 [16, 44]:

$$\mu_{A \cap B} = \begin{cases} \mu_A & \text{if } \mu_B = 1, \\ \mu_B & \text{if } \mu_A = 1, \\ 0 & \text{if } \mu_A, \mu_B < 1. \end{cases}$$

$$\mu_{A \cup B} = \begin{cases} \mu_A & \text{if } \mu_B = 0, \\ \mu_B & \text{if } \mu_A = 0, \\ 1 & \text{if } \mu_A, \mu_B > 0. \end{cases}$$

Over time, the notion arose that any t-norm and t-conorm could be used as a generalisation of intersection and union. This concept first appeared in the study [18].

T-norms are binary operations which first appeared in 1942 in Menger's article [43]. They were studied before the advent of fuzzy logic. For example, in paper [58] from 1960, all four basic t-norms above can be found. Many families of t-norms and t-conorms are now known; for example Schweizer-Sklar, Hamacher, Frank, Yager, Dombi, Sugeno-Weber, Aczel-Alsina, and Mayor-Torrens [34].

All t-norms (and t-conorms) can be arranged by size [18]. The smallest is the drastic t-norm, the largest is the Gödel t-norm. Similarly, the smallest t-conorm is the Gödel t-conorm, and the largest is the drastic t-conorm.

Negation, as with the complement, can also be modelled in several ways; for example, standard (= Łukasiewicz) negation, Gödel negation [30] and the Sugeno class of negations [22] are all known.

5.3.3   Intersections and Unions in Floppy Logic

In floppy logic, we generally apply intersection, union and complement.

Interestingly, it is possible to compare the (floppy) membership functions of intersection and union in both floppy logic and fuzzy logic.

The floppy membership function of the intersection of two floppy sets in floppy logic is bounded from above by a Gödel t-norm and from below by a Łukasiewicz t-norm. The floppy membership function of the union of two floppy sets is bounded from above by a Łukasiewicz t-conorm and from below by a Gödel t-conorm. More about this is described in Section 3.3.

5.3.4 Truth-Functionality Problem

The previous result is curiously related to the truth-functionality problem. Fuzzy logic is truth-functional, while floppy logic is not.

This means that in fuzzy logic, we can calculate the membership function of $\mu_{A \cap B}(x)$ or $\mu_{A \cup B}(x)$ from the membership functions of $\mu_A(x)$ and $\mu_B(x)$. By contrast, this is not possible in floppy logic. In floppy logic, the elements of individual floppy sets and their membership functions must be known.

Similarly, in probability theory, we must know the conditional probability, which then determines the degree of dependence of both events.

If we compare floppy logic and fuzzy logic, we can generally state that in fuzzy logic, the degree of dependence is determined by the choice of the relevant t-norm and t-conorm. The product t-norm and t-conorm correspond to independent events. The t-(co)norms between product and Gödel t-(co)norms correspond to positive dependent events. The t-(co)norms between product and Łukasiewicz t-(co)norms correspond to negative dependent events.

5.3.5 Implications in Floppy Logic and Fuzzy Logic

Implication is another important operation. In floppy logic, the case is simple. The implication is calculated according to the equation:

$$\boldsymbol{\mu_{A \Rightarrow B} = 1 - \mu_A + \mu_{A \cap B}}. \tag{5.3}$$

Some interesting results from floppy implication are examined in section 3.2.

Fuzzy logic has three main methods of defining an implication [3, 47]. The first (e.g [62]) uses the equation:

$$\mu_{A \Rightarrow B} = \mu_{\dot{\neg} A \dot{\vee} B}, \tag{5.4}$$

where $\dot{\neg}$ is a fuzzy negation and $\dot{\vee}$ is a t-conorm.

For standard negation and the Łukasiewicz, Gödel and product t-conorms, we obtain:

- Łukasiewicz implication: $\mu_{A \Rightarrow B} = \begin{cases} 1 - \mu_A + \mu_B & \text{if } \mu_A > \mu_B, \\ 1 & \text{if } \mu_A \leq \mu_B. \end{cases}$

- Kleene-Dienes implication: $\mu_{A \Rightarrow B} = \max\{1 - \mu_A, \mu_B\}$.

- Reichenbach implication: $\mu_{A \Rightarrow B} = 1 - \mu_A + \mu_A \cdot \mu_B$.

A second method of defining fuzzy implication is explored in [19, 30]:

$$\mu_{A \Rightarrow B} = \max \{ \mu_C : \ \mu_{A \dot\wedge C} \leq \mu_B \}, \tag{5.5}$$

where $\dot\wedge$ is a continuous t-norm.

For Łukasiewicz, Gödel and product t-norms, we obtain the following implications:

- Łukasiewicz implication.

- Gödel implication: $\mu_{A \Rightarrow B} = \begin{cases} \mu_B & \text{if } \mu_A > \mu_B, \\ 1 & \text{if } \mu_A \leq \mu_B. \end{cases}$

- Goguen implication: $\mu_{A \Rightarrow B} = \begin{cases} \mu_B / \mu_A & \text{if } \mu_A > \mu_B, \\ 1 & \text{if } \mu_A \leq \mu_B. \end{cases}$

The third method of obtaining a fuzzy implication is the equation:

$$\mu_{A \Rightarrow B} = 1 - \mu_A + \mu_{A \dot\wedge B}, \tag{5.6}$$

which is similar to the corresponding relationship from floppy logic. For Łukasiewicz, Gödel and product t-norms, we obtain Kleene-Dienes, Łukasiewicz and Reichenbach implications (in this order).

5.3.6   Probability in Fuzzy Logic and Floppy Logic

Fuzzy logic contains Zadeh's definition of the probability of a fuzzy set [70]:

$$P(A) = \int_{\Omega} \mu_A(x) \ \mathrm{d}P, \tag{5.7}$$

where the integral is a Lebesgue Stieltjes integral and P is a probability measure. A is a fuzzy set and $\mu_A(x)$ is its membership function.

In floppy logic, the probability of a floppy set is defined as:

$$R(\boldsymbol{B}) = \int_{\Omega} \boldsymbol{\mu_B}(x) \ \mathrm{d}P, \tag{5.8}$$

where the integral is a Lebesgue (or Lebesgue Stieltjes) integral and P is the probability measure presumed in Assumption 2.1.4. $\boldsymbol{B}$ is a floppy set and $\boldsymbol{\mu_B}(x)$ is its membership function.

These two equations are very similar, although they yield the same results only for single-element floppy sets.

For example, an event "warm or tepid water" in fuzzy logic is modelled according to the fuzzy union of the fuzzy sets "warm water" and "tepid water": $A_W \dot{\vee} A_T$.

In floppy logic, this event can be modelled according to the two-element floppy set: $\{A_W, A_T\}$. In both these cases we obtain, generally, different (floppy) membership functions and probabilities.

5.3.7 Assumptions of Floppy Logic

For floppy logic to be a model of Kolmogorov probability theory, Assumptions 2.1.1 to 2.1.5 must be satisfied. Let us now examine the validity of these assumptions in fuzzy logic:

In floppy logic, Assumption 2.1.1 is required so that the floppy membership functions of floppy sets, which are the sum of the membership functions of its elements, are well defined. In fuzzy logic, this assumption is usually not required.

In fuzzy logic, Assumption 2.1.2 is often accepted (e.g. [30, 47, 68]). Sometimes it is generalised (e.g. [26, 65]).

Assumption 2.1.3 is not a component of fuzzy logic and is sometimes highlighted as such (e.g. [27]). However, especially in practical applications, this assumption is often added (e.g. [10, 12, 20, 32, 49, 56]).

If we want to use Zadeh's definition of the probability of a fuzzy set [70], we must accept Assumptions 2.1.4 and 2.1.5.

One more assumption must be satisfied to model reality well: "Two properties which are described by two different primary fuzzy sets cannot simultaneously describe reality." This assumption is explained in Remark 2.1.5.

This assumption is not required in fuzzy logic.

The unpleasant consequence of this assumption is that in floppy logic, we must work with a larger set of primary fuzzy sets than in fuzzy logic. If the number of these primary fuzzy sets is great, we can use standard fuzzy operations to automatically estimate their membership functions. If the events are independent, we use the product t-norm. If the events are heavily positively dependent, we can use the Gödel t-norm. If the events are heavily negatively dependent, we can use the Łukasiewicz t-norm.

5.4 Floppy Logic and Adams' and Stalnaker's Probability Logic

Except for the membership function, the probability of a floppy set $R(\boldsymbol{A})$ is the second possible generalisation of the truth value in floppy logic. In this aspect, floppy logic is similar to the probability logic of Ernest Adams and Robert Stalnaker.

The main idea of these theories is Adams' Thesis or the PCCP hypothesis [29]: The probabilities of conditionals are conditional probabilities [2, 60]:

$$P(A \Rightarrow B) = P(B|A), \qquad \text{if } P(A) > 0. \qquad (5.9)$$

Adams adds that:

$$P(A \Rightarrow B) = 1, \qquad \text{if } P(A) = 0. \qquad (5.10)$$

This idea, best known from Adams' and Stalnaker's works, can already be found in some older studies, for example [54] from 1931.

Two important arguments against this idea are presented by Lewis [40]. However, Adams [1] and Douven and Verbrugge [15] defend the PCCP hypothesis.

In floppy logic, the relationship between the probability of conditionals and conditional probabilities is rather more complicated. We can derive[1] the equation:

$$R(\boldsymbol{A} \Rightarrow \boldsymbol{B}) = 1 - R(\boldsymbol{A}) + R(\boldsymbol{B}|\boldsymbol{A}) \cdot R(\boldsymbol{A}). \qquad (5.11)$$

A comparison of the probability of implication in probabilistic logic and floppy logic is presented in Figure 5.1.

The benefits of floppy implication include, for example, the fact that the implication is logically equivalent to its contrapositive (as in standard logic):

$$R(\boldsymbol{A} \Rightarrow \boldsymbol{B}) = R(\neg\boldsymbol{B} \Rightarrow \neg\boldsymbol{A}). \qquad (5.12)$$

The PCCP implication does not possess this property.

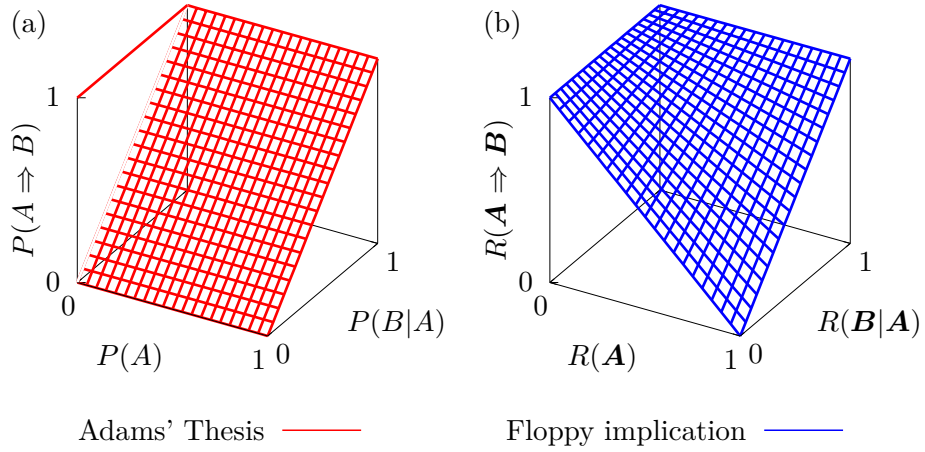[1]For the derivation of this equation, see Consequence 2.3.6.

Figure 5.1: Comparison of the probability of implication in probabilistic logic according to Adams' Thesis (a) and floppy logic (b)

Let us show one more advantage of floppy logic versus probability logic. In floppy logic, Lewis' triviality result [40] presents no problems, and implications may be used in the antecedent and consequent of other implications. For example, in floppy logic, the distributivity and transitivity of an implication apply such that:

$$R\left[\left[\boldsymbol{A} \Rightarrow (\boldsymbol{B} \Rightarrow \boldsymbol{C})\right] \Rightarrow \left[(\boldsymbol{A} \Rightarrow \boldsymbol{B}) \Rightarrow (\boldsymbol{A} \Rightarrow \boldsymbol{C})\right]\right] = 1, \quad (5.13)$$

$$R\left[\left[(\boldsymbol{A} \Rightarrow \boldsymbol{B}) \wedge (\boldsymbol{B} \Rightarrow \boldsymbol{C})\right] \Rightarrow \left[\boldsymbol{A} \Rightarrow \boldsymbol{C}\right]\right] = 1. \quad (5.14)$$

The truth of these statements derives directly from Theorem 2.3.1 and the distributivity and transitivity of implications in standard bivalent logic.

Conclusion

The main objective of this research was to find a consistent link between fuzzy sets and probability theory. The objective was attained and resulted in a new multi-valued logic, which was named floppy logic.

Floppy logic is a non-truth-functional logic similar to fuzzy logic. It works with floppy sets, which are crisp families of primary fuzzy sets.

Three important theorems concerning floppy logic were proved. The first two theorems claimed that floppy logic is a model of Kolmogorov probability theory. Therefore, it is possible to apply all the concepts and tools of this theory in floppy logic. This principle was demonstrated with an example which defined the mean value of a floppy set. Bayes' theorem and the law of total probability were also frequently applied.

The third theorem linked floppy logic with Boolean logic. The theorem demonstrated that each two statements which are logically equivalent in standard Boolean logic are also equivalent in floppy logic. It followed that floppy logic retains all the properties of standard two-valued logic which can be expressed as equivalences. I am confident that this remarkable feature lets us consider floppy logic as a multi-valued generalisation of standard two-valued logic.

Floppy logic also links two streams of thinking which generalise the truth value in different ways. It is a probabilistic logic which generalises truth values by applying probability and fuzzy logic which generalises truth values using membership functions. Both these concepts have their place in floppy logic.

Several other interesting results concern the mean value of the floppy set, floppy implication, and measurement of statement dependence.

In addition to the theoretical results, the practical aspects of floppy logic were explored. The thesis described detailed work with a system, including the selection of suitable primary fuzzy sets, fuzzification, application of system rules, defuzzification and optimal control. The thesis attempted to demonstrate that floppy logic is a relatively simple, intuitive, practical and elegant theory.

This research is certainly only the foundation of a new theory. Many options are available for the application or theoretical development of floppy logic. I cordially encourage the reader to engage in this research.

For example, it may be important to elaborate the concept of floppy logic as a generalisation of predicate logic.

Reduction of the computational difficulty of the multi-step control algorithm would be very useful. It would also be interesting to further develop the notion of measurement of the dependence of statements or to derive additional relationships for floppy implication or other floppy operations.

I am confident that this text contains everything essential for researchers to conduct theoretical or practical work with floppy logic.

References

[1]   ADAMS E.W. *A primer of probability logic.* Stanford, Calif.: Center for the Study of Language and Information, 1998. ISBN 157586066X.

[2]   ADAMS E.W. *The logic of conditionals: An application of probability to deductive logic.* Dordrecht: Springer Science & Business Media, 1975. ISBN 9789048183432.

[3]   BACZYŃSKI M., JAYARAM B. (S, N)- and R-implications: A state-of-the-art survey. *Fuzzy Sets and Systems.* 2008, 159(14), pp. 1836 –1859. ISSN 0165-0114.

[4]   BELLMAN R.E. *Adaptive control processes: a guided tour.* Princeton: Princeton university press, 2015. ISBN 9781400874668.

[5]   BELNAP N. A Useful Four-Valued Logic. In: J.M. DUNN, G. EPSTEIN, eds. *Modern Uses of Multiple-Valued Logic.* Dordrecht: Springer, 1977, pp. 5–37. ISBN 978-94-010-1161-7.

[6]   BEZDEK J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms.* Boston, MA: Springer US, 1981. ISBN 978-1-4757-0450-1.

[7]   BOCHVAR D.A. On a three-valued logical calculus and its application to the analysis of contradictions. *Rec. Math. (Mat. Sbornik) N. S.* 1938, 4(46)(2), pp. 287–308, doi: `10.14445/22312803/IJCTT-V10P119`. ISSN 23490829.

[8]   BOOLE G. *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities.* London: Walton and Maberly, 1854.

[9]   BORA D.J., GUPTA A.K. A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm. *International Journal of Computer Trends and Technology.* 2014, vol. 10(issue 2), pp. 108–113, doi: `10.14445/22312803/IJCTT-V10P119`. ISSN 23490829.

[10]   BROWN D.G. Classification and boundary vagueness in mapping presettlement forest types. *International Journal of Geographical Information Science.* 1998, 12(2), pp. 105–129, doi: `10.1080/136588198241914`. ISSN 1365-8816.

[11]   BURKILL J.C. *The Lebesgue integral.* London: Cambridge University Press, 1963.

[12]   CALCEV G. Some remarks on the stability of Mamdani fuzzy control systems. *IEEE Transactions on Fuzzy Systems.* 1998, 6(3), pp. 436–442, doi: `10.1109/91.705511`.

[13] CARNAP R. *Logical Foundations of Probability.* Chicago: University of Chicago Press, 1950.

[14] DE MORGAN A. *Formal Logic.* London: Taylor and Walton, 1847.

[15] DOUVEN I., VERBRUGGE S. The probabilities of conditionals revisited. *Cognitive Science.* 2013, 37(4), pp. 711–730, doi: `10.1111/cogs.12025`. ISSN 03640213.

[16] DUBOIS D. Quelques classes d'opérateurs remarquables pour combiner des ensembles flous (Some notable operator classes to combine fuzzy sets). *Busefal, Automne.* 1979, (1), pp. 29–35.

[17] DUBOIS D., PRADE H. Possibility Theory. *Wiley Encyclopedia of Electrical and Electronics Engineering.* 2001-08-21, doi: `10.1002/047134608X.W3502`.

[18] DUBOIS D., PRADE H.M. *Fuzzy sets and systems: theory and applications.* New York: Academic Press, 1980. ISBN 01-222-2750-6.

[19] ESTEVA F., GODO L., HÁJEK P., NAVARA M. Residuated fuzzy logics with an involutive negation. *Archive for mathematical logic.* 2000, 39(2), pp. 103–124, doi: `10.1007/s001530050006`. ISSN 1432-0665.

[20] FRANSSEN H.H., VAN EIJNSBERGEN A., STEIN A. Use of spatial prediction techniques and fuzzy classification for mapping soil pollutants. *Geoderma.* 1997, 77(2-4), pp. 243–262, doi: `10.1016/S0016-7061(97)00024-4`. ISSN 0016-7061.

[21] GAINES B.R. Fuzzy and probability uncertainty logics. *Information and Control.* 1978, vol. 38(issue 2), pp. 154–169, doi: `10.1016/S0019-9958(78)90165-1`. ISSN 00199958.

[22] GALAR M, BUSTINCE H, FERNANDEZ J, SANZ J, BELIAKOV G. Fuzzy entropy from weak fuzzy subsethood measures. *Neural network world.* 2010, 20(1), pp. 139–158. ISSN 1210-0552.

[23] GERLA G. Inferences in probability logic. *Artificial Intelligence.* 1994, 70(1-2), pp. 33–52, doi: `10.1016/0004-3702(94)90102-3`. ISSN 00043702.

[24] GILES R. Łukasiewicz logic and fuzzy set theory. *International Journal of Man-Machine Studies.* 1976, 8(3), pp. 313–327, doi: `10.1016/S0020-7373(76)80003-X`. ISSN 00207373.

[25] GIVANT S., HALMOS P. *Introduction to Boolean Algebras.* New York: Springer Science+Business Media, 2009. ISBN 978-0-387-68436-9.

[26] GOGUEN J.A. L-Fuzzy Sets. *Journal of Mathematical Analysis and Applications.* 1967, 18(1), pp. 145–174, doi: `10.1016/0022-247X(67)90189-8`. ISSN 0022247X.

[27] GONZÁLEZ M.A.S. Developing mass appraisal models with fuzzy systems. In: T. KAUKO, M. D'AMATO, eds. *Mass Appraisal Methods: An International Perspective for Property Valuers.* Oxford, UK: Wiley-Blackwell, 2008, pp. 181–202. ISBN 9781444301021.

[28] GOODWIN G.C., GRAEBE S.F., SALGADO M.E. *Control System Design.* Upper Saddle River: Prentice Hall, 2001. ISBN 0139586539.

[29] HÁJEK A. Probability, logic, and probability logic. In: L. GOBLE, ed. *The Blackwell guide to philosophical logic.* 2001, pp. 362–384. ISBN 9780631206927.

[30] HÁJEK P. *Metamathematics of Fuzzy Logic.* Boston: Kluwer Academic Publishers, 1998. ISBN 0-7923-5238-6.

[31] HÁJEK P., GODO L., ESTEVA F. *Fuzzy Logic and Probability.* 2013. Available also from: arXiv: 1302.4953 (cs.AI).

[32] KAREMORE G., MULLICK J.B., SUJATHA R, NIELSEN M., SANTHOSH C. Classification of protein profiles using fuzzy clustering techniques: An application in early diagnosis of oral, cervical and ovarian cancer. In: *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE.* Buenos Aires: Curran Associates, 2010, pp. 6361–6364. ISBN 978-1-4244-4123-5.

[33] KLEENE S.C. On the Interpretation of Intuitionistic Number Theory. *The Journal of Symbolic Logic.* 1945, 10(4), pp. 109–124, doi: 10.2307/2269016. ISSN 0022-4812.

[34] KLEMENT E.P., MESIAR R., PAP E. *Triangular Norms.* Dordrecht: Springer Netherlands, 2000. ISBN 978-94-015-9540-7.

[35] KOLMOGOROV A.N. *Grundbegriffe der Wahrscheinlichkeitsrechnung (Foundations of the Theory of Probability).* Berlin: J. Springer, 1933.

[36] KOPÁČEK J. *Matematická analýza nejen pro fyziky (III) (Mathematical Analysis Not Only for Physicists (III)).* 3., upr. vyd. Praha: Matfyzpress, 2007. ISBN 978-80-7378-020-3.

[37] KOSKO B. Fuzziness vs. Probability. *International Journal of General Systems.* 1990, vol. 17(2-3), pp. 211–240, doi: 10.1080/03081079008935108. ISSN 0308-1079.

[38] KOTZ S., READ C.B., BALAKRISHNAN N., VIDAKOVIC B., JOHNSON N.L. *Encyclopedia of Statistical Sciences.* Hoboken, NJ, USA: John Wiley & Sons, 2004. ISBN 9780471667193.

[39] KUČERA V. *Analysis and design of discrete linear control systems.* Upper Saddle River: Prentice-Hall, Inc., 1991. ISBN 978-0-13-033085-7.

[40] LEWIS D. Probabilities of conditionals and conditional probabilities. In: W.L. HARPER, R. STALNAKER, G. PEARCE, eds. *Ifs.* Dordrecht: Springer, 1976, pp. 129–147. ISBN 978-94-009-9117-0.

[41] ŁUKASIEWICZ J. Geneza logiki trójwartościowej (The genesis of three-valued logic). *Nauka Polska.* 1939, 24, pp. 215–223.

[42] ŁUKASIEWICZ J. O logice trójwartościowej (On three-valued logic). *Ruch Filozoficzny.* 1920, 6, pp. 170–171.

[43]   MENGER K. Statistical Metrics. *Proceedings of the National Academy of Sciences of the United States of America.* 1942, 28(12), pp. 535–537, doi: `10.1073/pnas.28.12.535`.

[44]   MIZUMOTO M. Fuzzy sets and their operations, II. *Information and Control.* 1981, 50(2), pp. 160 –174, doi: `https://doi.org/10.1016/S0019-9958(81)90191-1`. ISSN 0019-9958.

[45]   MONTES I., HERNÁNDEZ J., MARTINETTI D., MONTES S. Characterization of continuous t-norms compatible with Zadeh's probability of fuzzy events. *Fuzzy Sets and Systems.* 2013, vol. 228(October, 2013), pp. 29–43, doi: `10.1016/j.fss.2012.11.020`. ISSN 01650114.

[46]   NAGY I., SUZDALEVA E. *Algorithms and Programs of Dynamic Mixture Estimation: Unified Approach to Different Types of Components.* Cham: Springer, 2017. ISBN 978-3-319-64670-1.

[47]   NAVARA M., OLŠÁK P. *Základy fuzzy množin (Basics of Fuzzy Sets).* Praha: Vydavatelství ČVUT, 2002. ISBN 80-01-02585-3.

[48]   OKUDA T., TANAKA H., ASAI K. A formulation of fuzzy decision problems with fuzzy information using probability measures of fuzzy events. *Information and Control.* 1978, vol. 38(issue 2), pp. 135–147, doi: `10.1016/S0019-9958(78)90151-1`. ISSN 00199958.

[49]   ÖZDEMIR D., AKARUN L. A fuzzy algorithm for color quantization of images. *Pattern Recognition.* 2002, 35(8), pp. 1785–1791, doi: `10.1016/S0031-3203(01)00170-4`. ISSN 00313203.

[50]   POPPER K. *The logic of scientific discovery.* New York: Basic Books, 1959.

[51]   PROVINSKÝ P. Floppy Logic - a Younger Sister of Fuzzy Logic. *Neural Network World.* 2017, vol. 27(issue 5), pp. 479–497, doi: `10.14311/NNW.2017.27.025`. ISSN 2336-4335.

[52]   PROVINSKÝ P. Floppy Logic - Instructions for Use. *Neural Network World.* 2018, 28(5), pp. 473–494, doi: `10.14311/NNW.2018.28.026`. ISSN 2336-4335.

[53]   PROVINSKÝ P. Floppy logic as a generalization of standard Boolean logic. *Neural Network World.* 2020, 30(3), pp. 193–209, doi: `10.14311/NNW.2020.30.014`. ISSN 2336-4335.

[54]   RAMSEY F.P. Truth and Probability. In: R. BRAITHWAITE, ed. *The Foundations of Mathematics and Other Logical Essays.* London: Routledge a Kegan Paul, 1931, pp. 156–198.

[55]   RESCHER N. *Many-valued logic.* New York: McGraw-Hill, 1969. ISBN 0070518939.

[56]   RIVEROL C, COONEY J. Fault Detection Using Fuzzy Similarities. *International Journal of Soft Computing.* 2006, 1(3), pp. 193–198. ISSN 1816-9503.

[57] RÉNYI A. On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae.* 1955, vol. 6(3-4), pp. 285–335, doi: `10.1007/BF02024393`. ISSN 0001-5954.

[58] SCHWEIZER B., SKLAR A., et al. Statistical metric spaces. *Pacific J. Math.* 1960, 10(1), pp. 313–334, doi: `10.1109/TSMC.1977.4309687`.

[59] SHRAMKO Y., WANSING H. Some useful 16-valued logics: How a computer network should think. *Journal of Philosophical Logic.* 2005, 34(2), pp. 121–153, doi: `10.1007/s10992-005-0556-5`. ISSN 1573-0433.

[60] STALNAKER R.C. Probability and conditionals. *Philosophy of science.* 1970, 37(1), pp. 64–80, doi: `10.1086/288280`. ISSN 00318248.

[61] ŠTĚPÁN J. *Teorie pravděpodobnosti: Matematické základy (Probability Theory: Mathematical Foundations.* Praha: Academia, 1987.

[62] TRILLAS E., VALVERDE L. On implication and indistinguishability in the setting of fuzzy logic. In: J. KACPRZYK, R.R. YAGER, eds. *Management decision support systems using fuzzy sets and possibility theory.* Köln: Verlag TÜV Rheinland, 1985, pp. 198–212. ISBN 3885851431.

[63] VESELÝ J. *Základy matematické analýzy: První díl (Fundamentals of Mathematical Analysis: First Part).* Praha: Matfyzpress, 2004. ISBN 978-80-7378-063-02.

[64] WEN Z.L., HAN J.L., LIU F.S. Galaxy clusters identified from the SDSS DR6 and their properties. *The Astrophysical Journal Supplement Series.* 2009, 183(2), pp. 197–213, doi: `10.1088/0067-0049/183/2/197`. ISSN 0067-0049.

[65] YAGER R.R. Connectives and quantifiers in fuzzy sets. *Fuzzy Sets and Systems.* 1991, 40(1), pp. 39–75, doi: `10.1016/0165-0114(91)90046-S`. ISSN 01650114.

[66] ZADEH L.A. Calculus of fuzzy restrictions. In: L.A. ZADEH, K.-S. FU, K. TANAKA, M. SHIMURA, eds. *Fuzzy Sets and their Applications to Cognitive and Decision Processes.* 1975, pp. 1 –39. Available also from: `http://www.sciencedirect.com/science/article/pii/B9780127752600500062`. ISBN 978-0-12-775260-0.

[67] ZADEH L.A. Fuzzy probabilities. *Information Processing.* 1984, vol. 20(issue 3), pp. 363–372, doi: `10.1016/0306-4573(84)90067-0`. ISSN 03064573.

[68] ZADEH L.A. Fuzzy sets. *Information and Control.* 1965, 8(3), pp. 338–353, doi: `10.1016/S0019-9958(65)90241-X`. ISSN 00199958.

[69] ZADEH L.A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems.* 1978, vol. 1(issue 1), pp. 3–28, doi: `10.1016/0165-0114(78)90029-5`. ISSN 01650114.

[70] ZADEH L.A. Probability Measures of Fuzzy Events. *Journal of Mathematical Analysis and Applications.* 1968, 23(2), pp. 421–427, doi: `10.1016/0022-247X(68)90078-4`. ISSN 0022247x.

Appendices

Appendix A

Proof of Theorem on Basic Floppy Probability Space

A.1 Kolmogorov Axioms

Kolmogorov axioms are taken from the studies [35, 61] in the following form:

Let us have a set of values which can be adopted by a random variable. This set is denoted Ω and known as the *sample space.*

Let us also have the σ-algebra $\mathcal{B}$ of subsets of Ω. We therefore assume:

Axiom A.1.1. $\emptyset \in \mathcal{B}$.

Axiom A.1.2. If $A \in \mathcal{B}$, then $A' \in \mathcal{B}$. [1]

Axiom A.1.3. If $A_1,\ A_2,\ \ldots,\ A_n,\ \ldots \in \mathcal{B}$, then

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}. \tag{A.1}$$

The elements of $\mathcal{B}$ are called *events.*

A real number $P(A)$ is assigned to all events $A \in \mathcal{B}$. The number $P(A)$ is known as the *the probability of event A.*

We assume that the probability satisfies:

Axiom A.1.4. $P(\Omega) = 1$.

---

[1] A' is the designation for the set $\Omega - A$.

Axiom A.1.5. $\forall A \in \mathcal{B} : P(A) \geq 0$.

Axiom A.1.6. If $M = \{A_n\}$ is a finite or countable sequence of pairwise disjoint sets from $\mathcal{B}$, then

$$P\left(\bigcup_{A_n \in M} A_n \right) = \sum_{A_n \in M} P(A_n). \tag{A.2}$$

A.2 The Floppy Membership Function $\boldsymbol{\mu_B}(x)$ and the Probability $R(\boldsymbol{B})$ of Floppy Set $\boldsymbol{B}$ Are Well Defined

Lemma A.2.1. For all $x \in X$ and for each set $\boldsymbol{B} \subseteq \boldsymbol{S}$, the sum $\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ converges absolutely to a number from the interval $[0, 1]$.

First, we prove absolute convergence:

$\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x) = \sum_{A_i \in \boldsymbol{B} \cap \boldsymbol{S}_x} \mu_{A_i}(x)$, which is a finite or countable sum. (Assumption 2.1.1.)

Let n be a natural number.

A sequence of partial sums $\sum_{A_i \in \boldsymbol{B} \cap \boldsymbol{S}_x \wedge i < n} \mu_{A_i}(x)$ is non-decreasing since we add only non-negative functions. (Assumption 2.1.2.)

The sequence of partial sums $\sum_{A_i \in \boldsymbol{B} \wedge i < n} \mu_{A_i}(x)$ is bounded above by the number 1:

$$\sum_{A_i \in \boldsymbol{B} \wedge i < n} \mu_{A_i}(x) \leq \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x) \leq \sum_{A_i \in \boldsymbol{S}} \mu_{A_i}(x) = 1, \tag{A.3}$$

where Assumption 2.1.3 was applied in the previous equation.

A finite limit of a non-decreasing, bounded from above sequence always exists, and therefore the sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ converges (e.g. [63, p. 55]).

The sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ converges absolutely since we add only non-negative numbers.

Let us show that the sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ converges to a number in the interval $[0, 1]$:

The sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ is equal to or greater than zero since we add only non-negative numbers. (Assumption 2.1.2.)

The sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ is equal to or less than one because:

$$\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x) \leq \sum_{A_i \in \boldsymbol{S}} \mu_{A_i}(x) = 1, \tag{A.4}$$

where Assumption 2.1.3 is applied.

Lemma A.2.2. The sum $\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ does not depend on the order in which the addends are added together.

The sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x) = \sum_{A_i \in \boldsymbol{B} \cap \boldsymbol{S}_x} \mu_{A_i}(x)$ is a finite or countable sum according to Assumption 2.1.1.

A finite sum does not depend on the order of the addends.

If the sum is countable, then the expression $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x)$ converges absolutely according to Lemma A.2.1. The sum of an absolutely convergent sequence also does not depend on the order of addends (e.g. [63, p. 99]).

Lemma A.2.3. The function $\boldsymbol{\mu_B}(x)$ is measurable on the set X with respect to the measure P.

The sum $\sum_{A_i \in \boldsymbol{B}} \mu_{A_i}(x) = \sum_{A_i \in \boldsymbol{B} \cap \boldsymbol{S}_x} \mu_{A_i}(x)$ is a finite or countable sum (Assumption 2.1.1.)

All functions μ_{A_k} are measurable on the set X with respect to the measure P. (Assumption 2.1.5.)

Finite series:

The sum of two measurable functions is a measurable function (e.g. [11, p. 24]). A finite sum of measurable functions is thus a measurable function. Therefore, if the sum $\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B} \cap \boldsymbol{S}_x} \mu_{A_i}(x)$ is a finite sum, then it is measurable.

Countable series:

If the sum $\boldsymbol{\mu_B}(x) = \sum_{A_i \in \boldsymbol{B} \cap \boldsymbol{S}_x} \mu_{A_i}(x)$ is countable, then its partial sums are measurable functions (according to the previous paragraph). A sequence of partial sums is non-decreasing. (Assumption 2.1.2.) and a limit of a monotonous sequence of measurable functions is a measurable function (e.g. [11, p. 25]).

In both cases, the function $\boldsymbol{\mu_B}(x)$ is measurable.

Lemma A.2.4. The Lebesgue integral $R(\boldsymbol{B}) = \int_X \boldsymbol{\mu_B}(x)\,\mathrm{d}P$ always exists.

The function $\boldsymbol{\mu_B}(x)$ is bounded (according to Lemma A.2.1) and measurable on the set X with respect to the measure P (according to Lemma A.2.3).

$(X, \mathcal{A}, P)$ is a space with a probability measure. (Assumption 2.1.4.) Therefore, it satisfies all Kolmogorov axioms, especially Axiom A.1.4. So $P(X) = 1$. Therefore, the set X is measurable with respect to the measure P, and its measure is finite.

A Lebesgue integral of a bounded measurable function over a measurable set with a finite measure always exists (e.g. [11, pp. 29 - 31]).

A.3 The Basic Floppy Probability Space Satisfies All Kolmogorov Axioms

Proposition A.3.1. $\mathcal{P}(\boldsymbol{S})$ satisfies Axiom A.1.1.

The power set $\mathcal{P}(\boldsymbol{S})$ contains all subsets of $\boldsymbol{S}$ and therefore specifically $\emptyset$. Axiom A.1.1 is therefore satisfied.

Proposition A.3.2. $\mathcal{P}(\boldsymbol{S})$ satisfies Axiom A.1.2.

A set of elements of $\boldsymbol{S}$ that does not belong to $\boldsymbol{B}$ is a subset of $\boldsymbol{S}$. The power set $\mathcal{P}(\boldsymbol{S})$ contains all subsets of $\boldsymbol{S}$ and therefore specifically $\boldsymbol{B}'$. Axiom A.1.2 is therefore satisfied.

Proposition A.3.3. $\mathcal{P}(\boldsymbol{S})$ satisfies Axiom A.1.3.

The union of any system of subsets $\boldsymbol{S}$ is a subset of $\boldsymbol{S}$. The power set $\mathcal{P}(\boldsymbol{S})$ contains all subsets of $\boldsymbol{S}$ and therefore specifically $\bigcup_{i=1}^{\infty} \boldsymbol{B_i}$ where $\boldsymbol{B_i} \in \mathcal{P}(\boldsymbol{S})$. Axiom A.1.3 is therefore satisfied.

Proposition A.3.4. $(\boldsymbol{S}, \mathcal{P}(\boldsymbol{S}), R)$ satisfies Axiom A.1.4.

Axiom A.1.4 is applicable because:

$$R(\boldsymbol{S}) = \int_X \boldsymbol{\mu_S}(x)\,\mathrm{d}P = \int_X \sum_{A_i \in \boldsymbol{S}} \mu_{A_i}(x)\,\mathrm{d}P = \int_X 1\,\mathrm{d}P = P(X) = 1, \qquad \text{(A.5)}$$

where Assumption 2.1.3 was first applied followed by Axiom A.1.4 for the probability space $(X, \mathcal{A}, P)$ (according to Assumption 2.1.4).

Proposition A.3.5. $(\boldsymbol{S}, \mathcal{P}(\boldsymbol{S}), R)$ satisfies Axiom A.1.5.

A number $R(\boldsymbol{B})$ exists for every set $\boldsymbol{B}$. (Lemma A.2.4.) The number $R(\boldsymbol{B}) = \int_X \boldsymbol{\mu_B}(x)\,\mathrm{d}P$ is non-negative since we integrate a non-negative function (according to Lemma A.2.1.)

The space $(\boldsymbol{S}, \mathcal{P}(\boldsymbol{S}), R)$ therefore satisfies Axiom A.1.5.

Proposition A.3.6. $(\boldsymbol{S}, \mathcal{P}\left(\boldsymbol{S}\right), R)$ satisfies Axiom A.1.6.

Let $M = \{\boldsymbol{B_i}\}$ be a finite or countable sequence of pairwise disjoint sets from $\mathcal{P}\left(\boldsymbol{S}\right)$. Each fuzzy set $A_k \in \bigcup_{\boldsymbol{B_i} \in M} \boldsymbol{B_i}$ is therefore an element of just one floppy set $\boldsymbol{B_i}$. We can therefore write:

$$
\begin{aligned}
R\left(\bigcup_{\boldsymbol{B_i} \in M} \boldsymbol{B_i}\right) &= \int_X \boldsymbol{\mu}_{\bigcup_{\boldsymbol{B_i} \in M} \boldsymbol{B_i}}(x)\, \mathrm{d}P = \\
&= \int_X \sum_{A_k \in \bigcup \boldsymbol{B_i}} \mu_{A_k}(x)\, \mathrm{d}P = \\
&= \int_X \sum_{\boldsymbol{B_i} \in M} \left[\sum_{A_k \in \boldsymbol{B_i}} \mu_{A_k}(x) \right] \mathrm{d}P = \\
&= \int_X \sum_{\boldsymbol{B_i} \in M} \boldsymbol{\mu}_{\boldsymbol{B_i}}(x)\, \mathrm{d}P.
\end{aligned}
\tag{A.6}
$$

The sum and the integral can be exchanged since $\boldsymbol{\mu}_{\boldsymbol{B_n}}(x)$ are non-negative and measurable functions according to Lemmas A.2.1 and A.2.3 (e.g. [36, p. 106]).

Therefore:

$$
R\left(\bigcup_{\boldsymbol{B_i} \in M} \boldsymbol{B_i}\right) = \sum_{\boldsymbol{B_i} \in M} \int_X \boldsymbol{\mu}_{\boldsymbol{B_i}}(x)\, \mathrm{d}P = \sum_{\boldsymbol{B_i} \in M} R\left(\boldsymbol{B_i}\right).
\tag{A.7}
$$

Therefore, the space $(\boldsymbol{S}, \mathcal{P}\left(\boldsymbol{S}\right), R)$ satisfies Axiom A.1.6.

Appendix B

Proof of Theorem on Generalised Floppy Probability Space

Proposition B.0.1. $\mathcal{C}$ satisfies Axioms A.1.1, A.1.2 and A.1.3.

$\mathcal{C}$ is a σ-algebra. (Definition 2.2.1.)

Proposition B.0.2. $(\boldsymbol{S} \times X, \mathcal{C}, R)$ satisfies Axiom A.1.4.

$$R\left(\boldsymbol{S} \times X\right) = \int_X \boldsymbol{\mu_{S \times X}}\left(x\right) \mathrm{d}P = \int_X \sum_{A_i \in \boldsymbol{S}:[A_i,x] \in \boldsymbol{S} \times X} \mu_{A_i}\left(x\right) \mathrm{d}P = \int_X 1 \, \mathrm{d}P = 1. \quad \text{(B.1)}$$

Assumption 2.1.3 followed by Axiom A.1.4 for the probability space $(X, \mathcal{A}, P)$ (according to Assumption 2.1.4) were applied.

Proposition B.0.3. $(\boldsymbol{S} \times X, \mathcal{C}, R)$ satisfies Axiom A.1.5.

We can write:

$$R\left(\boldsymbol{C^G}\right) = \int_X \boldsymbol{\mu_{C^G}}\left(x\right) \mathrm{d}P = \int_X \sum_{A_i \in \boldsymbol{S}:[A_i,x] \in \boldsymbol{C^G}} \mu_{A_i}\left(x\right) \mathrm{d}P, \quad \text{(B.2)}$$

where $\mu_{A_i}\left(x\right)$ is measurable on the set $\left\{x \in X : [A_i, x] \in \boldsymbol{C^G}\right\}$ (Assumption 2.1.5) and non-negative (Assumption 2.1.2). Thus, the sum $\sum_{A_i \in \boldsymbol{S}:[A_i,x] \in \boldsymbol{C^G}} \mu_{A_i}\left(x\right)$ is measurable and non-negative. The integral $R\left(\boldsymbol{C^G}\right) = \int_X \sum_{A_i \in \boldsymbol{S}:[A_i,x] \in \boldsymbol{C^G}} \mu_{A_i}\left(x\right) \mathrm{d}P$ therefore exists and is non-negative.

Proposition B.0.4. $(\boldsymbol{S} \times X, \mathcal{C}, R)$ satisfies Axiom A.1.6.

Let $M = \left\{ \boldsymbol{C}_n^{\boldsymbol{G}} \right\}$ be a finite or countable sequence of pairwise disjoint sets from $\mathcal{C}$. Therefore, each point $[A_i, x] \in \bigcup\limits_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} \boldsymbol{C}_n^{\boldsymbol{G}}$ is an element of just one floppy set $\boldsymbol{C}_n^{\boldsymbol{G}}$.

We can therefore write:

$$
\begin{aligned}
R\left(\bigcup_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} \boldsymbol{C}_n^{\boldsymbol{G}} \right) &= \int_X \boldsymbol{\mu}_{\bigcup_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} C_n^{\boldsymbol{G}}} (x)\, \mathrm{d}P = \\
&= \int_X \sum_{A_i \in \boldsymbol{S} : [A_i, x] \in \bigcup_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} C_n^{\boldsymbol{G}}} \mu_{A_i}(x)\, \mathrm{d}P = \\
&= \int_X \sum_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} \left[\sum_{A_i \in \boldsymbol{S} : [A_i, x] \in \boldsymbol{C}_n^{\boldsymbol{G}}} \mu_{A_i}(x) \right] \mathrm{d}P = \\
&= \int_X \sum_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} \boldsymbol{\mu}_{C_n^{\boldsymbol{G}}}(x)\, \mathrm{d}P. \qquad \text{(B.3)}
\end{aligned}
$$

The sum and the integral can be exchanged since $\boldsymbol{\mu}_{C_n^{\boldsymbol{G}}}(x)$ are non-negative and measurable functions.

Therefore:

$$
R\left(\bigcup_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} \boldsymbol{C}_n^{\boldsymbol{G}} \right) = \sum_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} \int_X \boldsymbol{\mu}_{C_n^{\boldsymbol{G}}}(x)\, \mathrm{d}P = \sum_{\boldsymbol{C}_n^{\boldsymbol{G}} \in M} R\left(\boldsymbol{C}_n^{\boldsymbol{G}} \right). \qquad \text{(B.4)}
$$

The space $(\boldsymbol{S} \times X, \mathcal{C}, R)$ therefore satisfies Axiom A.1.6.

Appendix C

Proof of Theorem on Isomorphism

C.1 Explanation of Concepts

C.1.1 Boolean Algebra

A *Boolean algebra* is a non-empty set B, together with two binary operations $\wedge$ and $\vee$ (on B), a unary operation $'$, and two distinguished elements 0 and 1, satisfying the following axioms [25, p. 10]:

$$0' = 1, \qquad\qquad\qquad 1' = 0, \qquad\qquad\qquad (C.1)$$

$$p \wedge 0 = 0, \qquad\qquad\qquad p \vee 1 = 1, \qquad\qquad\qquad (C.2)$$

$$p \wedge 1 = p, \qquad\qquad\qquad p \vee 0 = p, \qquad\qquad\qquad (C.3)$$

$$p \wedge p' = 0, \qquad\qquad\qquad p \vee p' = 1, \qquad\qquad\qquad (C.4)$$

$$(p') = p, \qquad\qquad\qquad\qquad\qquad\qquad\qquad (C.5)$$

$$p \wedge p = p, \qquad\qquad\qquad p \vee p = p, \qquad\qquad\qquad (C.6)$$

$$(p \wedge q)' = p' \vee q', \qquad\qquad (p \vee q)' = p' \wedge q', \qquad\qquad (C.7)$$

$$p \wedge q = q \wedge p, \qquad\qquad\qquad p \vee q = q \vee p, \qquad\qquad\qquad (C.8)$$

$$p \wedge (q \wedge r) = (p \wedge q) \wedge r, \qquad p \vee (q \vee r) = (p \vee q) \vee r, \qquad (C.9)$$

$$p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r), \qquad p \vee (q \wedge r) = (p \vee q) \wedge (p \vee r). \qquad (C.10)$$

These axioms are not independent. Often, a smaller group of axioms is selected.

Boolean algebra is often denoted as follows: $(B, \wedge, \vee, ', 0, 1, =)$. Operations $\wedge$, $\vee$, and $'$ are often called *meet, join,* and *complement.* Elements 0 and 1 are called *zero* and *one.*

C.1.2 Atoms

An *atom q* of a Boolean algebra $\mathcal{B}$ is an element that satisfies the following propositions [25, p. 117]:

 a. $q \neq 0$,
 b. for every element $p \in \mathcal{B}$, either $p \wedge q = q$ or $p \wedge q = 0$, but not both.

Example 1: Atoms of a Boolean algebra of sentences given by sentence A_1, A_2, A_3, $\dots A_n$ are sentences $B_1 \wedge B_2 \wedge B_3 \wedge \dots \wedge B_n$, where either $B_i = A_i$ or $B_i = \neg A_i$.

Example 2: Atoms of a Boolean algebra of subsets of $\boldsymbol{S}$ are one-element subsets of $\boldsymbol{S}$.

C.1.3 Isomorphism

The Isomorphism f of two Boolean algebras $\mathcal{B}_1$ and $\mathcal{B}_2$ is a bijection which preserves the operations meet, join, and complement. Thus:

 a. $\forall p, q \in \mathcal{B}_1 : f(p \wedge q) = f(p) \wedge f(q),$ (C.11)
 b. $\forall p, q \in \mathcal{B}_1 : f(p \vee q) = f(p) \vee f(q),$ (C.12)
 c. $\forall p \in \mathcal{B}_1 : f(p') = f(p)'.$ (C.13)

C.2 Proof of Theorem 2.3.1

The proof begins with the statement: "Each finite Boolean algebra $\mathcal{B}$ is isomorphic to the field $\mathcal{P}(n)$, or, equivalently, to the Boolean algebra 2^n, for a non-negative integer n. In fact, n is the number of atoms in $\mathcal{B}$" [25, p. 127].

ε and δ are finite algebras, $\mathcal{P}(\boldsymbol{S})$ is the field $\mathcal{P}(n)$, n is the number of atoms in ε and δ. Therefore, ε is isomorphic to δ.

An isomorphism between ε and δ therefore must exist. Let us find it:

Let g be a binary relation over E and a set B of binary indices of length n.

The indices to the statements are as follows: If a sentence V satisfies $V \wedge U_k = U_k$, then 1 is in the k-th place.

If V satisfies $V \wedge U_k = \bot$, then 0 is in the k-th place. No other possibility exists since U_k are atoms.

Let us show that g is a bijection:

a. Each statement is assigned a binary index. Therefore, g is a *left-total relation*.

b. Each statement is assigned exactly one binary index.

   Let us suppose that two equivalent sentences $V \equiv W$ are assigned two different binary indices. Let these indices be different in the k-th place. This means that $V \wedge U_k = U_k$ and $W \wedge U_k = \bot$ (or vice versa).

   V and W are equivalent and can therefore be exchanged: $W \wedge U_k = U_k$. Thus, $U_k = \bot$, which is a contradiction since U_k is an atom and atoms are non-zero elements.

   Therefore, g is a *left-total function*.

c. Each binary index is assigned a statement. The statement $U_i \vee U_j \vee U_k \vee \ldots$ can be assigned to the binary index with 1 in the i-th, j-th, k-th, ...places.

   Therefore, g is a *surjection*.

d. E and B are finite sets with the same number of elements. The number of elements is 2^n [25, p. 82].

   Therefore, g is a *bijection*.

Let h be a binary relation $\mathcal{P}(\boldsymbol{S})$ to the set B of binary indices of length n.

The indices are assigned to the floppy sets as follows:

If a floppy set $\boldsymbol{B}$ satisfies $\boldsymbol{B} \cap \{A_k\} = \{A_k\}$, then 1 is in the k-th place.

If $\boldsymbol{B} \cap \{A_k\} = \emptyset$, then 0 is in the k-th place.

h is also a bijection. It can be proven in the same manner as bijection g.

Therefore, $h^{-1} \circ g$ is a bijection. Let us show that $h^{-1} \circ g$ is an isomorphism:

a. Preservation of the binary operation *meet*:

   Sentences V and W have the same binary indices as floppy sets $h^{-1} \circ g\,(V)$ and $h^{-1} \circ g\,(W)$.

   $V \wedge W$ has a binary index, where 1 is in the places where 1 is in both binary indices of V and W simultaneously.

   $h^{-1} \circ g\,(V) \cap h^{-1} \circ g\,(W)$ has a binary index where 1 is in the places where 1 is in both binary indices of $h^{-1} \circ g\,(V)$ and $h^{-1} \circ g\,(W)$ simultaneously.

   Therefore, $V \wedge W$ and $h^{-1} \circ g\,(V) \cap h^{-1} \circ g\,(W)$ have the same binary index.

   Therefore: $\forall V, W \in E : h^{-1} \circ g\,(V \wedge W) = h^{-1} \circ g\,(V) \cap h^{-1} \circ g\,(W)$.

b. Preservation of binary operation *join* can be proved in the same manner.

c. Preservation of unary operation *complementation*:

A binary index of sentence $\perp$ has 0 in all positions since $\forall k : \perp \wedge U_k = \perp$.

A binary index of sentence $\top$ has 1 in all positions since $\forall k : \top \wedge U_k = U_k$.

Applying the axiom of Boolean Algebra: $V \wedge \neg V = \perp$. Therefore, in the index of $\neg V$, 1 must not be in the positions where 1 is in the index of V.

Now, applying the axiom: $V \vee \neg V = \top$. Therefore, in the index of $\neg V$, 1 must be in all positions where 1 is not in the index of V.

Therefore, the index of $\neg V$ has 1 exactly in these positions where V has 0.

Similarly, it can be shown that $(h^{-1} \circ g\,(V))'$ has 1 exactly in these positions where $h^{-1} \circ g\,(V)$ has 0.

Therefore, $\neg V$ and $(h^{-1} \circ g\,(V))'$ have the same binary indices.

Therefore: $\forall V \in E : h^{-1} \circ g\,(\neg V) = (h^{-1} \circ g\,(V))'$.

d. Preservation of neutral elements:

Sentence $\perp$ and empty floppy set $\emptyset$ have a binary index with all zeros. Therefore:

$h^{-1} \circ g\,(\perp) = \emptyset$.

Similarly, it can be shown that $h^{-1} \circ g\,(\top) = \boldsymbol{S}$.

Bijection $h^{-1} \circ g$ preserves all operations, therefore it is an *isomorphism.*

Let us show that $h^{-1} \circ g$ satisfies all properties of relation f:

Properties 2.16, 2.17, and 2.18 were proved above. Now, we prove property 2.15:

First, $U_i \wedge U_j = \perp$, if $i \neq j$ is applied. The second possibility for atoms $U_i \wedge U_j = U_i$ cannot apply since U_j would not be an atom, which is a contradiction.

Therefore, all atoms have only one 1 in their binary indices. Sentence U_k has 1 in the k-th position and the floppy set $\{A_k\}$.

Therefore: $\forall k \in \{1, 2, 3, \ldots n\} : h^{-1} \circ g\,(U_k) = \{A_k\}$.

Now, let us show that $h^{-1} \circ g = f$.

First, the set of all sentences whose binary indices contain no more than k 1 is denoted M_k.

It was proved above that $h^{-1} \circ g\,(U_k) = f\,(U_k)$.

Now, let us show that $h^{-1} \circ g\left(\bot\right) = f\left(\bot\right)$. It is simple, because $h^{-1} \circ g\left(\bot\right) = \emptyset$ and $f\left(\bot\right) = f\left(V \wedge \neg V\right) = f\left(V\right) \cap f\left(\neg V\right) = f\left(V\right) \cap f\left(V\right)' = \emptyset$.

The equivalence $h^{-1} \circ g = f$ for all sentences from M_1 is thus proved.

All sentences from M_2 can be expressed as a disjunction of two sentences from M_1:

$$\forall Z \in M_2 : \exists V, W \in M_1 : Z = V \vee W. \tag{C.14}$$

Therefore:

$$\begin{aligned} h^{-1} \circ g\left(Z\right) &= h^{-1} \circ g\left(V \vee W\right) = \\ &= h^{-1} \circ g\left(V\right) \cup h^{-1} \circ g\left(W\right) = \\ &= f\left(V\right) \cup f\left(W\right) = f\left(Z\right). \end{aligned} \tag{C.15}$$

The equivalence $h^{-1} \circ g = f$ for all sentences from M_2 is thus proved.

This procedure can be repeated to prove the equivalence $h^{-1} \circ g = f$ for M_3, $M_4, \ldots M_n$.

Since $M_n = E$, the equivalence $h^{-1} \circ g = f$ is proved for all sentences from E.

Therefore, f is an *isomorphism*.

It represents the first assertion of Theorem 2.3.1. The following assertions are implied immediately.

Appendix D

Entry of Functions Used in Examples 4.3.2, 4.4.4, 4.4.5, 4.4.6, 4.6.1 and 4.6.2

The membership functions of the primary fuzzy sets are presented as they are shown in Figure 4.4(d).

Unhealthily cold – A_{1000}:

$$
\begin{aligned}
t \in (-\infty, 10] && \mu_{A_{1000}}(t) &= 1, \\
t \in (10, 20] && \mu_{A_{1000}}(t) &= 2 - 0.1 \cdot t, \\
t \in (20, \infty) && \mu_{A_{1000}}(t) &= 0. && \text{(D.1)}
\end{aligned}
$$

Only healthily cold – A_{0100}:

$$
\begin{aligned}
t \in (-\infty, 10] && \mu_{A_{0100}}(t) &= 0, \\
t \in (10, 20] && \mu_{A_{0100}}(t) &= -1 + 0.1 \cdot t, \\
t \in (20, 22] && \mu_{A_{0100}}(t) &= 1, \\
t \in (22, 30] && \mu_{A_{0100}}(t) &= \frac{30}{8} - \frac{t}{8}, \\
t \in (30, \infty) && \mu_{A_{0100}}(t) &= 0. && \text{(D.2)}
\end{aligned}
$$

Healthily cold and pleasant – A_{0110}:

$$
\begin{aligned}
t \in (-\infty, 22] && \mu_{A_{0110}}(t) &= 0, \\
t \in (22, 25] && \mu_{A_{0110}}(t) &= -\frac{22}{8} + \frac{t}{8}, \\
t \in (25, 30] && \mu_{A_{0110}}(t) &= -\frac{10}{40} + \frac{t}{40}, \\
t \in (30, 35] && \mu_{A_{0110}}(t) &= \frac{35}{10} - \frac{t}{10}, \\
t \in (35, \infty) && \mu_{A_{0110}}(t) &= 0.
\end{aligned}
\tag{D.3}
$$

Only pleasant – A_{0010}:

$$
\begin{aligned}
t \in (-\infty, 25] && \mu_{A_{0010}}(t) &= 0, \\
t \in (25, 35] && \mu_{A_{0010}}(t) &= -\frac{25}{20} + \frac{t}{20}, \\
t \in (35, 45] && \mu_{A_{0010}}(t) &= \frac{45}{20} - \frac{t}{20}, \\
t \in (45, \infty) && \mu_{A_{0010}}(t) &= 0.
\end{aligned}
\tag{D.4}
$$

Warm and pleasant – A_{0011}:

$$
\begin{aligned}
t \in (-\infty, 25] && \mu_{A_{0011}}(t) &= 0, \\
t \in (25, 40] && \mu_{A_{0011}}(t) &= -\frac{25}{20} + \frac{t}{20}, \\
t \in (40, 45] && \mu_{A_{0011}}(t) &= \frac{135}{20} - \frac{3 \cdot t}{20}, \\
t \in (45, \infty) && \mu_{A_{0011}}(t) &= 0.
\end{aligned}
\tag{D.5}
$$

Only warm – A_{0001}:

$$
\begin{aligned}
t \in (-\infty, 40] && \mu_{A_{0001}}(t) &= 0, \\
t \in (40, 45] && \mu_{A_{0001}}(t) &= -\frac{40}{5} + \frac{t}{5}, \\
t \in (45, \infty) && \mu_{A_{0001}}(t) &= 1.
\end{aligned}
\tag{D.6}
$$

Let us select the probability density as follows:

$$
f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}},
\tag{D.7}
$$

where

$$
\mu = 25,
\tag{D.8}
$$

$$
\sigma = 10.
\tag{D.9}
$$

Appendix E

Entry of Functions Used in Examples 4.3.3, 4.4.2 and 4.4.7

The (floppy) membership functions are presented as they are shown in Figure 4.5.

Cold – $\boldsymbol{T_1}$:

$$
\begin{array}{llll}
t \in (-\infty, 10] & p \in \mathbb{R} & \boldsymbol{\mu_{T_1}}(t, p) = 1, & \\
t \in (10, 20] & p \in \mathbb{R} & \boldsymbol{\mu_{T_1}}(t, p) = 2 - \dfrac{t}{10}, & \\
t \in (20, \infty) & p \in \mathbb{R} & \boldsymbol{\mu_{T_1}}(t, p) = 0. & \text{(E.1)}
\end{array}
$$

Tepid – $\boldsymbol{T_2}$:

$$
\begin{array}{llll}
t \in (-\infty, 10] & p \in \mathbb{R} & \boldsymbol{\mu_{T_2}}(t, p) = 0, & \\
t \in (10, 20] & p \in \mathbb{R} & \boldsymbol{\mu_{T_2}}(t, p) = -1 + \dfrac{t}{10}, & \\
t \in (20, 25] & p \in \mathbb{R} & \boldsymbol{\mu_{T_2}}(t, p) = 1, & \\
t \in (25, 30] & p \in \mathbb{R} & \boldsymbol{\mu_{T_2}}(t, p) = 6 - \dfrac{t}{5}, & \\
t \in (30, \infty) & p \in \mathbb{R} & \boldsymbol{\mu_{T_2}}(t, p) = 0. & \text{(E.2)}
\end{array}
$$

Warm – $\boldsymbol{T_3}$:

$$
\begin{array}{llll}
t \in (-\infty, 25] & p \in \mathbb{R} & \boldsymbol{\mu_{T_3}}(t, p) = 0, & \\
t \in (25, 30] & p \in \mathbb{R} & \boldsymbol{\mu_{T_3}}(t, p) = -5 + \dfrac{t}{5}, & \\
t \in (30, \infty) & p \in \mathbb{R} & \boldsymbol{\mu_{T_3}}(t, p) = 1. & \text{(E.3)}
\end{array}
$$

Low – $\boldsymbol{P_1}$:

$$t \in \mathbb{R} \qquad p \in (-\infty, 960] \qquad \boldsymbol{\mu}_{\boldsymbol{P_1}}(t,p) = 1,$$

$$t \in \mathbb{R} \qquad p \in (960, 980] \qquad \boldsymbol{\mu}_{\boldsymbol{P_1}}(t,p) = 49 - \frac{p}{20},$$

$$t \in \mathbb{R} \qquad p \in (980, \infty) \qquad \boldsymbol{\mu}_{\boldsymbol{P_1}}(t,p) = 0. \qquad (\text{E.4})$$

Normal – $\boldsymbol{P_2}$:

$$t \in \mathbb{R} \qquad p \in (-\infty, 960] \qquad \boldsymbol{\mu}_{\boldsymbol{P_2}}(t,p) = 0,$$

$$t \in \mathbb{R} \qquad p \in (960, 980] \qquad \boldsymbol{\mu}_{\boldsymbol{P_2}}(t,p) = -48 + \frac{p}{20},$$

$$t \in \mathbb{R} \qquad p \in (980, 1010] \qquad \boldsymbol{\mu}_{\boldsymbol{P_2}}(t,p) = 1,$$

$$t \in \mathbb{R} \qquad p \in (1010, 1020] \qquad \boldsymbol{\mu}_{\boldsymbol{P_2}}(t,p) = 102 - \frac{p}{10},$$

$$t \in \mathbb{R} \qquad p \in (1020, \infty) \qquad \boldsymbol{\mu}_{\boldsymbol{P_2}}(t,p) = 0. \qquad (\text{E.5})$$

High – $\boldsymbol{P_3}$:

$$t \in \mathbb{R} \qquad p \in (-\infty, 1010] \qquad \boldsymbol{\mu}_{\boldsymbol{P_3}}(t,p) = 0,$$

$$t \in \mathbb{R} \qquad p \in (1010, 1020] \qquad \boldsymbol{\mu}_{\boldsymbol{P_3}}(t,p) = -101 + \frac{p}{10},$$

$$t \in \mathbb{R} \qquad p \in (1020, \infty) \qquad \boldsymbol{\mu}_{\boldsymbol{P_3}}(t,p) = 1. \qquad (\text{E.6})$$

The membership functions of the primary fuzzy sets are obtained from the equation:

$$\mu_{A_{i,j}}(t,p) = \boldsymbol{\mu}_{\boldsymbol{T_i}}(t,p) \cdot \boldsymbol{\mu}_{\boldsymbol{P_j}}(t,p). \qquad (\text{E.7})$$

Let us select, as a joint probability density, a multivariate normal probability density with mean vector $m = \begin{pmatrix} 15 \\ 1000 \end{pmatrix}$ and covariance matrix $S = \begin{pmatrix} 100 & 20 \\ 20 & 400 \end{pmatrix}$.

The two-dimensional normal probability density is given by the equation:

$$f(t,p) = \frac{1}{2\pi\sqrt{|S|}} \cdot e^{-\frac{1}{2} \cdot (\mathrm{x-m})^{\mathrm{T}} \cdot \mathrm{S}^{-1} \cdot (\mathrm{x-m})}, \qquad (\text{E.8})$$

where $x = \begin{pmatrix} t \\ p \end{pmatrix}$.