Shiyou Lian

Principles of Imprecise-Information Processing

A New Theoretical and Technological System

Springer

Principles of Imprecise-Information Processing

Shiyou Lian

Principles of Imprecise-Information Processing

A New Theoretical and Technological System

Springer

Shiyou Lian
Xi'an Shiyou University
Xi'an
China

ISBN 978-981-10-1547-2 ISBN 978-981-10-1549-6 (eBook)
DOI 10.1007/978-981-10-1549-6

Library of Congress Control Number: 2016940820

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

Preface

With the development of information and intelligence sciences and technologies as well as the rise in social requirements, imprecise-information processing about flexible linguistic values is becoming more and more important and urgent, and it will play an indispensable role in intelligent systems, especially in the anthropomorphic intelligent systems.

Imprecision, which is different from uncertainty, is another independent attribute of information, and the now so-called fuzziness is actually a kind of imprecision. Therefore, the author proposed explicitly the concepts of imprecise information and imprecise-information processing, and the book, just as the title shows, is a monograph on imprecise-information processing.

Actually, on imprecise-information processing, many scholars have been doing research with some results, among which the fuzzy set theory introduced by American Professor Loft Zadeh is the most famous. In fact, since Zadeh proposed the concept of fuzzy sets in 1965, the fuzzy-information processing technology based on fuzzy set theory has developed rapidly and made some achievements. However, so far, some important theoretical and technical problems in fuzzy-information processing have not been solved very well. For this reason, not a few scholars worked to improve and develop fuzzy set theory, and presented many new ideas, theories, and methods, which all have their respective angles of view and characteristics. But on the whole, people have not yet reached a common view, and the existed problems are neither solved really. Making a general survey of the decades of imprecise-information processing, although people presented many theories and methods, a theoretical and technological system has not yet been formed, that is, widely approved and has solid foundation of mathematics and logic like that for uncertain-information processing. In particular, some scholars still put imprecision of information into the category of uncertainty of information or mix the two together to do research. Therefore, imprecise-information processing is still a significant subject necessitating careful research.

After years of concentrated study, the author discovers that the imprecision of information originates from the phenomenon of “continuous distribution or change”

of magnitudes of a feature of things (or in other words, "uniform chain similarity" of things) and the treating way of "flexible clustering" of human brain. Thus, based on this and combined with the ways of human brain dealing with imprecise information in daily language, I have examined and explored the principles and methods of imprecise-information processing in an all-round way. As a result, a series of new theories and methods different from fuzzy technology were obtained, which forms a new theoretical and technological system of imprecise-information processing. This book is just a summation of these research results. Of course, viewed from the relationship between flexible sets and fuzzy sets, this book can also be viewed as an "amendment" to fuzzy-information processing technology; however, it does not follow traditional thinking to make modifications and supplementations in the existing framework of fuzzy set theory. Rather, tracing to the source and opening a new path, this book researches and explores the imprecise-information processing with new perspectives and ideas.

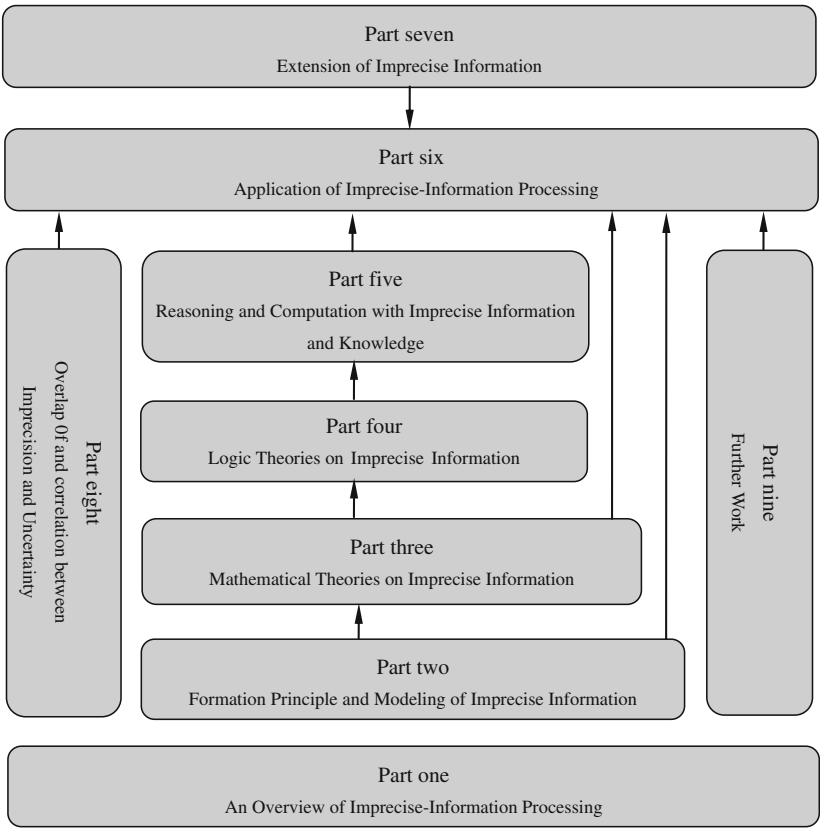
As early in the start of the 1990s, while building expert systems, from the doubt of Zadeh's CRI (compositional rule of inference) fuzzy reasoning, the author began to think about the problem of "fuzzy." In the period, I analyzed the objective cause that brings about the "fuzziness" of information and proposed some terminologies, concepts and methods such as "flexible linguistic values," "flexible concepts," "degreed logic,"and "reasoning with degrees". In 2000, a book *Degree theory* was published in which I summed up the research results at that time. After that, I continued to explore in this direction, further examined the formation principles and mathematical models of flexible concepts, and realized that fuzzy set is somewhat too general in describing a "fuzzy concept." Accordingly, I introduced the terminology and concept of "flexible sets," further examined the flexible linguistic values, and then founded the corresponding theories of mathematics and logic and meanwhile also found the geometric models and practical models of flexible concepts, the logical semantics of propositions, and the mathematical essence of flexible linguistic rules. The series of new discoveries and new progress made me more confident and determined to continue the cultivation in this field. During further researches, I realized gradually that the more essential characteristic of those so-called fuzzy (vague) concepts modeled by using fuzzy sets should be "flexible" rather than "fuzzy or vague," and the information containing flexible linguistic values is actually a kind of imprecise information. Therefore, I took specifically "imprecise-information processing" as a direction and objective and carried out an all-round research. In September 2009, the results obtained were gathered as a book and formally published with the name of *Principles of Imprecise- Information Processing* (Chinese version).

After this book was published, I also had some further understanding and thinking. For instance, we can also research approximate reasoning and computing at the level of linguistic functions and can extend flexible linguistic value to more general quantifiable linguistic value. Meantime, some problems and deficiencies in this book are also found such as the exposition about "uniform chain similarity relation" of things, the wording of "real number space," the discrimination between flexible concept and vague (fuzzy) concept, the analyses of inference in

truth-degreed logic, and the logical semantics of propositions, which all need improvement, and there are some redundancies in Chap. 12. In particular, the comparison is not made in this book between the principles and methods of approximate reasoning and computation we present and those in fuzzy set theory. In addition, some contents in this book are not so closely related to the theme of this book. Thus, I continued again the work nonstop. I did research further and at the same time also made revisions, corrections, and extensions to the original work: deleted or reduced some contents, extended some contents, changed some formulations and especially added many new contents (such as "flexible linguistic functions" and "quantifiable linguistic values" as well as the logical and mathematical principles of approximate reasoning). Thus, some original chapters and sections were deleted, but some new chapters and sections were added, and most of original chapters and sections were rewritten or adapted; correspondingly, the structure of text was also made a large modification—changed from the original 8 parts and 21 chapters to 9 parts and 26 chapters. Thus, a new work about "Principles of Imprecise-Information Processing," that is, the second edition of original book, has been formed. At the beginning of 2015, the manuscript of the new work had been basically completed, and then, some polishing was done. The new work has two versions: one in Chinese and one in English, the latter being this book.

Compared with the original edition, the new edition made much new progress both in depth and in extent—not only the quantity is increased but also the quality is raised, and the whole theoretical and technological system is more compact and coherent. This book has nine parts. The first part gives an outline of imprecise-information processing; the second part reveals the formation principle of imprecise information and establishes its mathematical models; the third part is the mathematical theory on imprecise information; the fourth part is the logic theory on imprecise information; the fifth part expatiates on the principles and methods of reasoning and computation with imprecise information and knowledge; the sixth part is the application and acquiring techniques of imprecise knowledge; the seventh part is the extension of imprecise information; the eighth part expatiates on and deals with the overlap of and the correlation between imprecision and uncertainty; and finally, the ninth part is further work. In terms of structure, Part I is the introduction, Part II the origin, Parts III and IV the basis, Part V the main body, Part VI the application (interface), Part VII the extension, Part VIII the cross, and Part IX the frontier. Their logical relationships and the hierarchy of this book are shown in the following diagram.

This book researches imprecise-information processing by using mathematical and logical methods, but meanwhile, it also develops the corresponding theories of mathematics and logic. The whole book presents over 100 important concepts, derives over 40 theorems and more than 100 formulas, functions, and rules, and gives over 70 specific methods and algorithms. Besides, there are also brief commentaries of some existing viewpoints and methods (which are mainly of fuzzy set theory) in this book.



Architecture of *Principles of Imprecise-Information Processing*

This book also has a feature; that is, there are many symmetrical, antithetical, or corresponding concepts and terminologies such as "flexible linguistic value" and "flexible set," "membership function" and "consistency function," "geometric model," and "algebraic model," "combined linguistic value" and "synthetic linguistic value," "form of possession" and "form of membership," "logical composition" and "algebraical composition," "conjunction-type rule" and "disjunction-type rule," "complementary flexible partition" and "exclusive flexible partition," "flexible linguistic value" and "rigid linguistic value," "medium value" and "neutral value," "L-N function" and "N-L function," "certain rule" and "uncertain rule," "natural logical semantics" and "extended logical semantics," "reasoning with truth-degrees" and "reasoning with believability-degrees," "degree-true inference" and "near-true inference," "numerical ××" and "linguistic ××," "conceptual ××" and "practical ××," "×× of single conclusion" and "×× of multiple conclusions," "×× on the same space" and "×× from distinct spaces," "one-dimensional ××" and "multi-dimensional ××," and "typical ××" and

"non-typical ××," thus forming many pairs of parallel or complementary theories and methods—they are arranged in a crisscross pattern and together constitute a multidimensional system of theories and technologies.

This book makes an all-round and systematic research of imprecise-information processing, but the focus is on clarifying concepts, straightening out relationships, revealing principles, and presenting methods to lay a theoretical foundation and build a technological platform for further research and application. In fact, on the basis of this book, we can directly develop related applications and also carry out further researches.

Imprecise-information processing is a big subject; in addition, the vision and the level of the author are limited, so the deficiencies and defects in both content and expression in this book are unavoidable although great effort was made. Therefore, the author sincerely invites experts and scholars to grant instructions and the readers to comment and make suggestions!

Xi'an, China Shiyou Lian
January 2016

Acknowledgements

On the formal publication of this book, I owe a debt of gratitude to my daughter Xiaoya Lian for her great contributions in the translation of the manuscript. Without her help, it is hard to know when this book could ever meet the readers!

I also want to thank my elder daughter's family for their support in spirit!

I would also like to thank all persons who gave me help and support in the completion of this book!

Finally, I still want to thank the people who helped me in the publication of the original work, because without yesterday's original work, there would not be today's new work.

Contents

Part I
Introduction

Chapter 1
Overview of Imprecise-Information Processing

Abstract This chapter introduces firstly what is imprecise information and then examines the origin of imprecise information, thus revealing the formation principle of imprecise information, and then, it discusses the distinction and correlation between imprecision and uncertainty of information, the research issues of imprecise-information processing, and significance of studying imprecise-information processing and the related disciplines and fields; finally, it outlines the work of the book. Besides, a survey of researches on imprecise-information processing is given in the chapter.

Keywords Imprecise information · Flexible linguistic values · Uncertain information · Artificial intelligence

1.1 What Is Imprecise Information?

Imprecise information here refers mainly to the information that is expressed by words with imprecise meanings. For example, "tall" is a word with imprecise meaning in that there is no strict and rigid standard for a certain height to be considered "tall." Therefore, the word "tall" expresses imprecise information [1].

Words with imprecise meanings can be found everywhere in our daily communication and written materials. Here are some examples: "morning" and "evening" characterizing time, "nearby" characterizing location, "far" and "near" characterizing distance, "much" and "little" characterizing quantity, "big" and "small" characterizing volume or space, and "slight," "a little," "very," and "extremely" characterizing strength. Other words such as "fast," "slow," "hot," "cold," "good," "bad," "diligent," "hardworking," "serious," "friendly," "beautiful," "kindhearted," "brave," "ardently love," and "very likely" are all words with imprecise meanings. Thus, it is clear that there is imprecision almost everywhere in our communication (Look, this statement itself contains imprecision: What is "almost"?).

Imprecise information also includes the information expressed by words with precise meanings but which can be replaced by numerical values, because,

S. Lian, *Principles of Imprecise-Information Processing*,
DOI 10.1007/978-981-10-1549-6_1

compared to numerical values, words appear to be less precise. For example, school records can be represented by "word grades," such as "good" or "excellent," but also can be represented by corresponding numerical scores, such as 85 and 98. Here, the former is not as precise as the latter. Additionally, the occurrence of a random event can be described by the word "likely," but can also be described by a numerical probability; similarly, the former is also not as precise as the latter. Therefore, the information expressed by such words as "good," "excellent," and "likely," is also imprecise.

Note: In addition to the above-mentioned causes, imprecision can also be caused by inappropriate words or inappropriate measuring units (the imprecision caused by measuring units is usually called inexactness). For example, when describing a person's place of residence, the name of country is not as precise as that of the province or city. Also, "ton" is not as precise as "kilogram" and "kilogram" is not as precise as "gram" in describing weights. This book will not cover the study of these two kinds of imprecision.

1.2 Origin of Imprecise Information

As stated above, imprecise information is caused by words with imprecise meanings. Then, why are the meanings of these words not precise? We know that words are actually the linguistic symbolic representations of corresponding concepts in human brain. The reason these words' meanings are imprecise is that the concepts represented by these words have no strict definitions. That is, these concepts' connotations have no rigid standards or conditions and their denotations have no rigid boundaries. In other words, their connotative conditions and denotative boundaries have a certain softness or flexibleness. For example, for the word "tall," heights over 1.75 m are all "tall" to a certain degree, and for "young," ages under 40 years are all "young" to a certain degree. For another example, the boundary between "hot" and "cold" weather is actually a "flexible boundary." That is, "hot" gradually transitions to "cold," and in turn, "cold" also gradually transitions to "hot." Therefore, the concepts expressed by words with imprecise meanings are actually "flexible concepts." That is to say, imprecise information turns out to be caused by flexible concepts in our brain. Then, how are these flexible concepts formed?

We know that everything has some attributes or states, and there are some relationships between things. To facilitate the narration, we call the attributes, states, and relations of things collectively to be the features of things.

Observing and examining the boundless universe we live in, it can be found that for one and the same feature, each relevant object has its specific magnitude and these magnitudes are not exactly the same, but assume continuous distribution or continuous change, thus forming a continuous range. Examples:

- Human heights continuously distribute or change from about 0.3 to 2.5 m, forming a range of [0.3, 2.5], which is a continuous real interval.

- Air temperatures continuously change from about −45 to 45 °C, forming a range of [−45, 45], which is also a continuous real interval.
- Ages of humans continuously distribute or change from about 1 to 120, forming the range of {1, 2, …, 120}, and this is a continuous set of integers.

Note that the "continuous" here includes the "continuous" of real numbers, the "continuous" of rational numbers (i.e., "dense"), and the "consecutive" of integers (i.e., order) and to be the same later.

It can be seen from the above examples that after a certain measure being introduced, the magnitudes of a feature of things are explicitly shown as concrete numbers.

We call the numbers representing magnitudes of a certain feature of things to be the **numerical feature values**, or simply, the **numerical values**, of things. Thus, we will treat the magnitudes and numerical feature values, i.e., numerical values, of the things as synonym later.

It is not hard to see that the continuity of magnitudes of a feature of things makes corresponding things show as the uniform chain similarity relation. Then, facing with one and another things being uniformly chained similar, how should the human brain save and deal with relevant information? Of course, numerical feature values can directly describe things precisely, but if they are used all the time and everywhere, then the human brain would be unable to bear the enormous amount of data and humans would be unable to tolerate the trouble.

It can be seen that in the continuous numerical feature values, the relation between adjacent numbers is the approximation relation. Thus, we can cluster these numerical values according to approximation relation and then express the numbers in one and the same class by using one and the same word. This kind of word summarizing a batch of numerical values is a big-granule value—**linguistic value**. The linguistic values are also a kind of values representing magnitudes of features of things, namely **linguistic feature values**. Clearly, the number of linguistic values of one feature is very finite. Thus, by clustering and summarizing numerical values, we can use finite number of linguistic values to represent an infinite number of numerical values. Thus, the amount of information can be greatly reduced and the complexity lowered.

Actually, according to the law of "quantitative change to qualitative change," there is an obvious difference between things' properties represented by numerical values far apart in range. So, speaking from this point, numerical feature values of things must be clustered and partitioned. Another benefit of clustering and partitioning numerical values and then expressing them in linguistic values is that we can understand and grasp things at a higher level.

However, unfortunately, it is difficult to do the usual natural and objective clustering and partitioning of these continuous numerical values (for detailed analysis, see Sect. 2.1). For this reason, the human brain adopts the clever strategy of flexible clustering and flexible partitioning to obtain corresponding "flexible classes" (for concrete principle and method, see Chaps. 2, 3) and afterward summarize the thing's properties stood for by flexible classes, thus obtaining "flexible concepts"

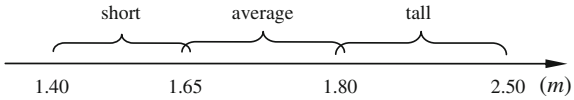


Fig. 1.1 An example of flexible clustering and flexible classes in a range of heights of adults

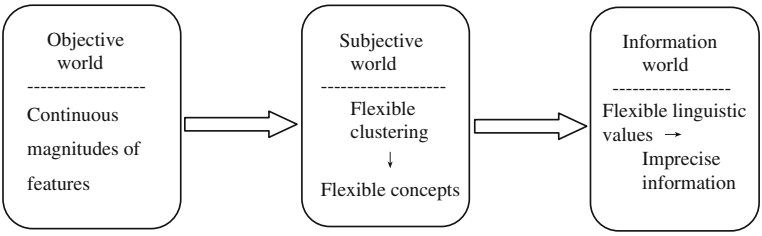


Fig. 1.2 The diagram of the origin of imprecise information

and "flexible linguistic values" representing flexible concepts. For example, as shown in Fig. 1.1, through flexible partitioning of the range [1.4, 2.5] of heights of adults by flexible clustering, we obtain the corresponding flexible classes, flexible concepts, and flexible linguistic values: "short," "average," and "tall."

From stated above, we see that it is just the phenomenon of "continuous distribution or change" of magnitudes of a feature of things (or in other words, "uniform chain similarity" of things) in the objective world and the treating way of "flexible clustering" of the human brain that result in the flexible concepts in human brain, and then, there occur flexible linguistic values and corresponding imprecise information. Thus, the origin of imprecise information can be diagramed as follows (see Fig. 1.2).

Now, there exists yet another question: When does the human brain flexible treating with respect to the continuous magnitudes of features? In other words, are those flexible concepts in human brain obtained independently by each individual's own flexible clustering of corresponding feature's magnitudes? We will discuss the problem in Sect. 19.1.

There might be readers who think, "Aren't the flexible concepts talked about here same as 'fuzzy concepts' in some other literatures"? Right, the flexible concepts we talk about here are just the fuzzy concepts called in some literatures. Then, why do we call them flexible concepts but not fuzzy concepts? Section 19.3 will give the answer.

1.3 Distinction and Correlation Between Imprecision and Uncertainty of Information

Besides imprecise information, there also exists uncertain information in usual information processing and communication. For example:

It might rain tomorrow.

This sentence carries a piece of uncertain information—it rain tomorrow.

Note that the uncertainty we talk about here refers only to the uncertainty (of information) that is caused from randomness of things or people's lack of knowledge of things. It is not that kind of uncertainty, said in the literature [2], including fuzziness (i.e., imprecision), vagueness, unknownness, non-specificity, strife, discord, conflict, and ignorance. Of course, the uncertainty we talk about here also does not include fallibility, instability, inaccuracy, incompletion, and ambiguity.

1. **Distinction between imprecision and uncertainty**

Imprecise information is the information that describes the features and relations of things not specifically, strictly, or exactly enough. Uncertain information is the information of which the authenticity cannot be determined, that is, the event, or the properties, relationships, or behaviors of things expressed by which is not certain or not sure.

From the last section, imprecise information originates from the continuous distribution or change of numerical feature values of relevant things (or the uniform chain similarity of things) and the flexible treating mechanism of human brain. Uncertain information originates then from the feature of "partial share" of relevant sets and the relations of "partial correspondence" or "partial inclusion" between relevant sets (see Sect. 25.3).

Although the imprecise information has an objective basis, it is a "man-made" product, so it has a certain subjectivity. The uncertain information is the objective expression of properties or behaviors of things that people can't be sure, but in the description of the degree of uncertainty, there may be subjective factors.

Imprecise information is directly expressed by the relevant statements (of which the uncertainty is shown in the linguistic value(s) of the relevant statements). Uncertainty information, in general, cannot be expressed directly, but it is expressed indirectly by the aid of a main-clause-structured compound sentence (we call it the possibly type modal proposition, see Sect. 25.1). For example, the above uncertain information "It rain tomorrow" is expressed by "It might rain tomorrow," that is, "'It rain tomorrow' is possible."

From the above, we can see that uncertainty and imprecision are two mutually independent attributes of information. Uncertain-information processing solves the possibility problem of the truth or falsity of information, while imprecise-information processing solves the strength problem of the truth or falsity of information.

2. **Correlation between imprecision and uncertainty**

Now that uncertainty and imprecision are two mutually independent attributes of information, the correlation between them would be not the relation of subordination. But, we find that there are some links between the two of them.