Part I

Fundamentals

Jane M. Booker

I.1 Chapters 1–6

To bridge the gap between probability and fuzzy theories, the first step is to examine and understand the two sides of the gap. The first part of this book consists of six chapters that lay the foundations of both theories and provide the fundamental principles for constructing the bridge.

We (the editors) begin (in Chapter 1) with an introduction to the history of both theories and the stories describing the formulation of the gap between them. It is our intent to represent both "sides," but with the tone of reconciliation. There are cases where applications support one theory more than the other, and these are brought forth in the application chapters in Part II. There are also cases where either probability or fuzzy theory is useful or a hybrid approach combining the two is best, particularly when characterizing different kinds of uncertainties in a complex problem.

Following the philosophical discussion in Chapter 1, Chapters 2 and 3 provide the foundations of fuzzy theory and probability theory, respectively.

Chapter 4 is devoted to Bayesian probability theory. Bayes' theorem provides a powerful structure for bridging the gap between fuzzy and probability theory and lays some of the groundwork for the bridging mechanism. Chapter 5 then completes the building of the bridge in a mathematical and philosophical sense.

Because data and information in today's world often reside within the experience and knowledge of the human mind, Chapter 6 examines the formal use of eliciting and analyzing expert judgment. That topic comprises an entire book in its own right (Meyer and Booker, 2001). The contents of this chapter not only covers both theories but provides applications, making it the perfect transition chapter to the applications chapters in Part II.

I.2 Suggested reading

The reader who has little familiarity with either fuzzy or probability theory should review Chapters 2 and 3, respectively. We highly recommend Chapter 4 for an introductory understanding of how to bridge the gap between the two theories and recommend Chapter 5 to complete this understanding. Because many uncertainties in complex problem solving and mass communication are imbedded in the cognitive processes of the human mind, we suggest a review of the formal methods for handling expert knowledge.

References

M. A. Meyer and J. M. Booker (2001), *Eliciting and Analyzing Expert Judgment: A Practical Guide*, ASA–SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA.

Chapter 1
Introduction

Timothy J. Ross, Jane M. Booker, and W. Jerry Parkinson

1.1 Some history and initial thoughts

The history of humankind and its relationship to technology is well documented. One such technology—information technology—has been impacting human thought over at least the last 400 years.

For example, the 17th century was described by its technology historians as the Age of Experience, in which many developments from direct observation were realized in fundamental astronomy, physics, chemistry, and mathematics despite remaining mysteries in other fundamental concepts (e.g., sphericity of the planet). The 18th century was perceived as the Age of Reason—the Renaissance following the Dark Ages in Europe at the end of the 17th century. From the beginning of the 19th century until about 1950, we had what we could call the Age of Mathematics—here the arithmetical formalisms of the previous century were advanced into more formal calculus-based theories. From 1950 to 1980, we had the Age of Computing—a time where many computational models of physical processes were developed. This period paralleled the developments in digital computing from the ILLIAC through the IBM® 360 to the IBM PC. From 1980 to 1995, we had the Age of Knowledge—a time when, as evidenced in its literature, a great deal of effort was focused on the acquisition and appropriate use of knowledge. Again, we see a parallel development, this time in hardware, between the areas of symbolic computing (LISP machines, Mathematica) and parallel computing (CRAY, Connection-Machine, IBM SP). Finally, 1995 to the present is the Age of Cyberspace—an age of ubiquitous networks and multiple, rapid forms of communication.

Humans cannot be expected to reasonably adjust to today's rapid technological advancements. Today humans are in information overload. The engine behind this can be traced to the individual forces of computer hardware and software; the nascent field of network computing; integration of various engineering fields; new technologies such as GIS (geographic information systems) and GPS (global positioning systems) producing gigabytes of information each hour; intelligent databases; complex simulation models; collaborative engineering; inventive engineering; and new methods to present, distill, and deliver technological education to distant areas—so-called distance education (Arciszewski, 1999).

Superposed with this development in information technology we had a parallel development in the theoretical frameworks for assessing uncertainty in the information. Probability concepts date back to the 1500s, the time of Cardano when gamblers recognized the rules of probability in games of chance and, more important, that avoiding these rules resulted in a sure loss (i.e., the classic coin toss example of "heads you lose, tails I win," referred to as the "Dutch book"). The concepts were still very much in the limelight in 1685, when the Bishop of Wells wrote a paper that discussed a problem in determining the truth of statements made by two witnesses who were both known to be unreliable to the extent that they only tell the truth with probabilities p_1 and p_2, respectively. The Bishop's answer to this was based on his assumption that the two witnesses were independent sources of information (Lindley (1987b)).

Probability theory was initially developed in the 18th century in such landmark treatises as Jacob Bernoulli's *Ars Conjectandi* (1713) and Abraham DeMoiver's *Doctrine of Chances* (1718; 2nd ed., 1738). Later in that century a small number of articles appeared in the periodical literature that would have a profound effect on the field. Most notable of these were Thomas Bayes' *An Essay Towards Solving a Problem in the Doctrine of Chances* (1763) and Pierre Simon Laplace's formulation of the axioms relating to games of chance, *Mémoire sur la probabilité des causes par les evenemens* (1774). Laplace, only 25 years old at the time he began his work in 1772, wrote the first substantial article in mathematical statistics prior to the 19th century. Despite the fact that Laplace, at the same time, was heavily engaged in mathematical astronomy, his memoir was an explosion of ideas that provided the roots for modern decision theory, Bayesian inference with nuisance parameters (historians claim that Laplace did not know of Bayes' earlier work), and the asymptotic approximations of posterior distributions (Stigler (1986)).

By the time of Newton, physicists and mathematicians were formulating different theories of probability (see Chapter 3). The most popular ones remaining today are the relative frequency theory and the subjectivist, or personalistic, theory. The latter theory was initiated by Thomas Bayes (1763), who articulated his very powerful theorem for the assessment of subjective probabilities. The theorem specified that a human's degree of belief could be subjected to an objective, coherent, and measurable mathematical framework within the subjective probability theory. In the early days of the 20th century, Rescher developed a formal framework for a conditional probability theory, and Jan Lukasiewicz developed a multivalued, discrete logic (circa 1930). In the 1960s, Arthur Dempster developed a theory of evidence which, for the first time, included an assessment of ignorance, or the absence of information. In 1965, Zadeh introduced his seminal idea in a continuous-valued logic called fuzzy set theory. In the 1970s, Glenn Shafer extended Dempster's work[1] to produce a complete theory of evidence dealing with information from more than one source, and Lotfi Zadeh illustrated a possibility theory resulting from special cases of fuzzy sets. Later, in the 1980s other investigators showed a strong relationship between evidence theory, probability theory, and possibility theory with the use of what have been called fuzzy measures (Klir and Folger (1988)).

In the over three decades since its inception by Zadeh, fuzzy set theory (and its logical counterpart, fuzzy logic) has undergone tremendous growth. Over ten thousand papers, hundreds of books, almost a dozen journals, and several national and international societies bear witness to this growth. Table 1.1 shows a count of papers containing the word "fuzzy" in the title, as cited by INSPEC and MathSciNet databases. (Data for 2001 are not complete.) To this day, perhaps because the theory is chronologically one of the newest, fuzzy sets and

[1] We refer to this extension in the text as the Dempster–Shafer theory.

Table 1.1. *Number of papers with the word "fuzzy" in the title.**

Period	INSPEC	MathSciNet
1970–1979	570	441
1980–1989	2,383	2,463
1990–1999	23,121	5,459
2000–2001	5,940	1,670
Totals	32,014	10,033

*Compiled by Camille Wanat, Head, Engineering Library, University of California at Berkeley, June 21, 2002.

fuzzy logic remain steeped in controversy and debate for a variety of reasons. Although the philosophical and mathematical foundations of fuzzy sets are intuitive, they run counter to the thousands of years of dependence on binary set theory on which our entire Western cultural logic resides (first espoused by Aristotle in ancient Greece). In addition, some have seen fuzzy sets as a competitor to probability theory (the new-kid-on-the-block syndrome) in a variety of settings, such as in competition for precious page space in journals, for classes on campuses, for students in graduate classes, and even for consulting opportunities in industry, to name a few. The statistical societies have even sponsored debates in their own journals (e.g., *Statist. Sci.*, 1 (1986), pp. 335–358) and conferences on topics ranging from the mathematical (e.g., the axioms of subjective probability) to the linguistic (e.g., communicating better with engineers and scientists (Hoadley and Kettering (1990)) in an effort to win back "market share," i.e., to get science and engineering students to take classes in statistics and probability instead of evidence theory, fuzzy logic, soft computing, and other new uncertainty technologies.

However, the debate extends far beyond "market share" competitiveness. The core issues involve the philosophical and theoretical differences between these theories and how these theories are useful for application in today's complex, information-based society. Later in this chapter, there is a lengthy discussion on the running debate between advocates of the two theories concerning the merits of each theory in terms of modeling uncertainty and variability.[2]

It is the premise of this book that this perceived tension between probability theory and fuzzy set theory is precisely the mechanism necessary for scientific advancement. Within this tension are the searches, trials and errors, and shortcomings that all play a part in the evolution of any theory and its applications. In this debate between advocates of these two theories are the iterations necessary to reach a common ground that will one day seem so intuitive and plausible that it will be difficult to reflect or to remember that there ever was a debate at all! Our goal in writing this book is to illustrate how naturally compatible and complementary the two theories are and to help the reader see the power in combining the two theories to address the various forms of uncertainty that plague most complex problems. Some of this compatibility can be reached by examining how probability is interpreted. As will be demonstrated in this book, it is much easier to *bridge the gap* when a subjective or personalistic (e.g., Bayesian-based) interpretation of probability is used (see Chapter 3 for a discussion).

The contributors to this book were chosen for their experience and expertise in both theories but also for their efforts in and understanding of the compatibility of both fuzzy and probability theories. Even though some chapters are more fuzzy oriented and some are more

[2]It should be noted that within the probability community there is an equally intense debate over the interpretation of probability (i.e., frequentist versus subjective). To our knowledge, the only attempt ever made in resolving that debate is addressed in Chapter 3.

probability oriented,[3] the goal for each chapter is to provide some insight into establishing the common ground, i.e., *bridging the gap*.

In addition to the motivations for this book as explained above, there is ample evidence in the literature about the need for more information sharing among groups using various theories to assess and quantify uncertainty. For example, Laviolette et al. (1995) claim that "so few statisticians and probabilists have considered the efficacy of fuzzy methods"! Another observation (Bement (1996)) is that "the reason engineers (e.g., Zadeh) had to devise their own theory for handling different kinds of uncertainty was because statisticians failed to respond to the engineering needs." Thus far, most of the debate about FST (fuzzy set theory) and probability has appeared in "fuzzy" journals that "are not frequently read by statisticians." This statement rings true in many other works in the literature.

1.2 The great debate

1.2.1 The debate literature

There have been several organized debates at fuzzy conferences, including a recent one at the annual Joint Statistical Meeting (Anaheim, CA, August 1997); an oral faceoff between two individuals (G. Klir and P. Cheeseman) at the 8th Maximum Entropy Workshop (August 1–5, 1988, St. John's College, Cambridge, U.K.; see Klir (1989)); and there exist several archival collections of written debates. For example, the following citations give an idea of the activity produced in very well written arguments espousing the virtues of various uncertainty methods, although most of the debate centers on Bayesian probability or fuzzy logic:

- *Statist. Sci.*, 2 (1987), pp. 3–44;
- *Comput. Intell.*, 4 (1988), pp. 57–142;
- *IEEE Trans. Fuzzy Systems*, 2 (1994), pp. 1–45;
- *Technometrics*, 37 (1995), pp. 249–292.

In addition, numerous articles have been written outside of organized debates that have been critical of one or the other theory by protagonists of each. These articles include the following:

- Lindley, *Internat. Statist. Rev.*, 50 (1982), pp. 1–26 (with seven commentaries);
- Cheeseman, in the edited volume *Uncertainty in Artificial Intelligence*, North-Holland, Amsterdam, 1986, pp. 85–102;
- Cheeseman, *Comput. Intell.*, 4 (1988), pp. 58–66 (with 22 commentaries);
- Hisdal, *Fuzzy Sets Systems*, 25 (1988), pp. 325–356;
- Kosko, *Internat. J. Gen. Systems*, 17 (1990), pp. 211–240;
- Laviolette and Seaman, *Math. Sci.*, 17 (1992), pp. 26–41;

[3] It is not inconceivable that some problems are more fuzzy oriented and some are more probability oriented than others.

- Elkan, in *Proceedings of the American Association for Artificial Intelligence*, MIT Press, Menlo Park, CA, 1993, pp. 698–703 (with numerous commentaries in the subsequent AAAI magazine);

- Zadeh, *IEEE Trans. Circuits Systems*, 45 (1999), pp. 105–119.

The next section takes many of the points made in the various historical debates and organizes them into a few classic paradigms that seem to be at the heart of the philosophical differences between the two theories. While this is merely an attempt to summarize the debates for the purpose of spawning some thinking about the subsequent chapters of this book, it can in no way represent a complete montage of all the fine points made by so many competent scientists. Such a work would be another manuscript in itself.

1.2.2 The issues and controversy

Any attempt to summarize the nearly four decades of debate between the probability and fuzzy communities is a daunting task, and one that is in danger of further criticism for at least two reasons: First, we could not possibly include everyone's arguments (see the brief review of the many organized debates in the previous section); second, we could be precariously close to misrepresenting the opinions of those arguments that we do include here if we have inadvertently taken some arguments out of their originally intended contexts. Therefore, we apologize in advance for any omissions or possible misrepresentations included herein. It is important to mention that many of the opinions and arguments presented in this brief review of the "great debate" should be taken in their historical perspectives. Some of the individuals whose quotes are provided here—quotes perhaps made very early in the evolution of the process—may have changed their minds or changed their perspectives on some of the issues raised over the past 35 years. In addition, in changing views or perspectives, some of these individuals, and more not mentioned here, unwittingly have advanced the knowledge in both fields—fuzzy and probability theories—because it is the debate process, by its very nature, that has forced a positive evolution in bridging the gap between these two very powerful and very useful models of uncertainty and variability.

In what follows, we have organized our review of the great debate into some rather useful, although perhaps arbitrary, epistemological paradigms. We begin with a review of some of the polemical statements that fueled the fires of many of the original debates. Although strident in their character, in retrospect these polemics were well timed in terms of forcing people to look more closely at the side of the debate they were defending. The polemics tended to be one-sided, opposing fuzzy set theory, this being the *new sibling* in the family and requesting equal consideration as a viable theory for assessing uncertainty alongside its more mature and metaphorically larger *brother*, probability theory. Very few of the polemics were aimed in the other direction—probably more as a defensive reaction—and we will mention some of these.

Next, we organize the debate into some paradigms that seemed to characterize much of the disagreement. These are

- philosophical issues of chance, ambiguity, crispness, and vagueness;

- membership functions versus probability density functions;

- Bayes' rule;

- the so-called conjunction fallacy;

- the so-called disjunction contradiction;
- the excluded-middle laws;
- the infusion of fuzziness into probability theory.

We conclude our summary of the great debate with some rather pithy but positive statements.

The early polemics

We give first billing to the father of fuzzy logic, Professor Lotfi Zadeh, whose seminal paper in 1965 obviously started the debate. In a recent paper (Zadeh (1999)), where he discusses the need for developing methods to compute with words, he recounts a few remarks made in the early 1970s by two of his colleagues. These remarks revealed, in a very terse and crude way, a deep-seated proclivity of hard scientists to seriously consider only those things that are numerical in nature. To preface these remarks, Professor Zadeh cited a statement attributed to Lord Kelvin in 1883 about the prevailing 19th century respect for numbers and the utter disrespect for words:

> "I often say that when you can measure what you are speaking about and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind: it may be the beginning of knowledge but you have scarcely, in your thoughts, advanced to the state of science, whatever the matter may be."

Zadeh goes on to recall that, in 1972, Rudolph Kalman, remarking on Zadeh's first exposition on a linguistic variable, had this to say:

> "Is Professor Zadeh presenting important ideas or is he indulging in wishful thinking? No doubt Professor Zadeh's enthusiasm for fuzziness has been reinforced by the prevailing climate in the U.S.—one of unprecedented permissiveness. 'Fuzzification' is a kind of scientific permissiveness; it tends to result in socially appealing slogans unaccompanied by the discipline of hard scientific work and patient observation."

In a similar vein, in 1975 Zadeh's colleague at Berkeley, Professor William Kahan, offered his assessment:

> "Fuzzy theory is wrong, wrong, and pernicious. I cannot think of any problem that could not be solved better by ordinary logic. What we need is more logical thinking, not less. The danger of fuzzy theory is that it will encourage the sort of imprecise thinking that brought us so much trouble."

These statements, and others similar to them, set the stage for the great debate that, although continuing today, has calmed in its rhetoric in recent years.

In 1988, Peter Cheeseman excited numerous investigators in the fields of fuzzy set theory, logic, Dempster–Shafer evidence theory, and other theories with his work "An Inquiry into Computer Understanding." In this paper, he made the claim that both fuzzy set theory and Dempster–Shafer evidence theory violated *context dependency*, a required property of any method-assessing beliefs. This statement, and many others included in the paper, such as his misstatements about possibility distributions (Ruspini (1988)), incited such a debate that 22 commentaries followed and were published.

In a series of papers dealing with the expression of uncertainty within artificial intelligence, Dennis Lindley (1982, 1987a, b) perhaps expressed the most vociferous challenge to fuzzy set theory—or any other non-Bayesian theory—with his comments about the inevitability of probability:

> "*The only satisfactory description of uncertainty is probability*. By this is meant that every uncertainty statement must be in the form of a probability; that several uncertainties must be combined using the rules of probability, and that the calculus of probabilities is adequate to handle *all* situations involving uncertainty. In particular, alternative descriptions of uncertainty are unnecessary. These include the procedures of classical statistics; rules of combination... possibility statements in fuzzy logic... use of upper and lower probabilities... and belief functions."

In a single paragraph, Lindley's proclamations were sufficient to not only extend indefinitely the debate between fuzzy and probability, but also to prolong numerous other debates, such as those continuing for the past 100 years between frequentists and Bayesians.

Other statements were made that, although inflammatory, appeared to have less substance. Hisdal (1988a) stated "the fuzzy set group is... in the position of having a solution for which it has not yet found a problem" and "a theory whose formulas must be replaced by other ad hoc ones whenever it does not agree with experiment is not a finished theory." A rather unexplained quote from her, "Fuzziness $\neq$ randomness... this is a very strong, and... also a very surprising assertion," was followed by an even more enigmatic statement: "Fuzzy set theory has mostly assumed that some mystic agent is at work, making its fuzzy decisions according to some undefined procedure." Moreover, in (Hisdal (1988b)), she states of fuzzy set theory that "This seems to imply the belief, that human thinking is based on inexact, fuzzily-defined concepts. As I have heard a colleague express it, the theory of fuzzy sets is no theory at all, it is more like a collection of cooking recipes."

Laviolette and Seaman (1992) remarked that "fuzzy set theory represents a higher level of abstraction relative to probability theory" and questioned whether laws exist to govern the combination of membership values. In a rather curious metaphor explaining the relationship of the axiomatic differences in the two theories, specifically the fuzzy property of supersubsethood as articulated by Bart Kosko in 1990, they begin with "The foundation of probability is both operationally and axiomatically sound. Supersethood alone need not make one theory more broadly applicable than another in any practical sense. If we wish to design and construct a building, we need only Newtonian mechanics and its consequences. The fact that quantum mechanics includes Newtonian mechanics as a subset theory is of no practical consequence." Then, in speaking about the operational effectiveness of the two theories, they further suggest "We have proposed that fuzzy methods be judged by their sensitivity to changes in an associated probabilistic model. We take for granted that fuzzy methods are sub-optimal with respect to probabilistic methods. It is important that a method for judging the efficacy of fuzzy methods be developed and employed, given the widespread interest in FST." Finally, they state "... in our opinion, this operational deficiency remains the chief disadvantage of fuzzy representations of uncertainty."

From the other side of the argument, Kosko (1994) states that "probability is a very special case of fuzziness." In referring to the excluded-middle laws, he states that it "forces us to draw hard lines between things and non-things. We cannot do that in the real world. Zoom in close enough and that breaks down." (See section 2.1.2 in Chapter 2 on excluded-middle laws for an expansion of these ideas.) Kosko further states that "Our math and world view might be different today if modern math had taken root in the A-AND-not-A views of Eastern culture instead of the A-OR-not-A view of ancient Greece." To dismiss this as *unfortunate deconstructionism* is just to name call and to ignore historical fact. For a long time the probability view had a monopoly on uncertainty, but now "fuzzy theory challenges the probability monopoly... the probability monopoly is over."

In his 1990 paper "Fuzziness vs. Probability," Kosko addresses everything from randomness, to conditional probability, to Bayesian subjective probability. He very eloquently shows that his subsethood theorem is derived from first principles and, in commenting on the lack of a derivable expression for conditional probability, remarks that this is the "difference between showing and telling." He remarked that his "subsethood theorem suggests that randomness is a working fiction akin to the luminiferous ether of nineteenth-century physics—the phlogiston of thought." Noting that his derivation of the subsethood theorem had nothing to do with randomness, he states "The identification of relative frequency with probability is cultural, not logical. That may take some getting used to after hundreds of years of casting gambling intuitions as matters of probability and a century of building probability into the description of the universe. It is ironic that to date every assumption of probability—at least in the relative frequency sense of science, engineering, gambling, and daily life—has actually been an invocation of fuzziness." Giving equal attention to the subjective probability groups, he claims that "Bayesianism is a polemical doctrine." To justify this, he shows that Bayes' rule also stems from the subsethood theorem.[4] Additionally, in commenting on Lindley's argument (Lindley (1987a)) that only probability theory is a coherent theory and all other characterizations of uncertainty are incoherent, Kosko states "this polemic evaporates in the face of" the subsethood theorem and "ironically, rather than establish the primacy of axiomatic probability, Lindley seems to argue that it is fuzziness in disguise." Kosko finishes his own series of polemics by euphemistically pointing out that perhaps 100 years from now, "no one... will believe that there was a time when a concept as simple, as intuitive, as expressive as a fuzzy set met with such impassioned denial."

Kosko's prediction reflects the nature of change in human cognition and perhaps even philosophy. Whether that process can be best described as fuzzy or probabilistic in nature we leave as a mental exercise.

Philosophical issues: Chance, ambiguity, and crispness versus vagueness

It is sometimes useful to first sit back and ask some fundamental questions about what it is we are trying to do when we attempt to first characterize various forms of uncertainty, and then to posit mathematical forms for quantifying them. Philosophers excel at these questions. The history of humankind's pondering of this matter is rich, and won't be replicated here, but the ideas of a few individuals who have thought about such notions as chance, ambiguity, crispness, and vagueness are cited here. While we could look at a time in Western culture as far back as that of Socrates, Descartes, or Aristotle, or revisit the writings of Zen Buddhism, we shall skip all that and move directly to the 20th century, where a few individuals spoke profoundly on these matters of uncertainty.

Max Black, in writing his 1937 essay "Vagueness: An exercise in logical analysis," first cites remarks made by the ancient philosopher Plato about uncertainty in geometry, then embellishes on the writings of Bertrand Russell (1923) who emphasized that "all traditional logic habitually assumes that precise symbols are being employed." With these great thoughts prefacing his own arguments, he proceeded to his own, now famous quote:

[4]While subsethood by itself may not guarantee performance enhancement in practical applications, enhancements to existing solutions from the use of fuzzy systems is illustrated in the applications chapters of Part II of this book.

> "It is a paradox, whose importance familiarity fails to diminish, that the most highly developed and useful scientific theories are ostensibly expressed in terms of objects never encountered in experience. The line traced by a draftsman, no matter how accurate, is seen beneath the microscope as a kind of corrugated trench, far removed from the ideal line of pure geometry. And the 'point-planet' of astronomy, the 'perfect gas' of thermodynamics, or the 'pure-species' of genetics are equally remote from exact realization. Indeed the unintelligibility at the atomic or subatomic level of the notion of a rigidly demarcated boundary shows that such objects not merely are not but could not be encountered. While the mathematician constructs a theory in terms of 'perfect' objects, the experimental scientist observes objects of which the properties demanded by theory are and can, in the very nature of measurement, be only approximately true."

More recently, in support of Black's work, Quine (1981) states

> "Diminish a table, conceptually, molecule by molecule: when is a table not a table? No stipulations will avail us here, however arbitrary.... If the term 'table' is to be reconciled with bivalence, we must posit an exact demarcation, exact to the last molecule, even though we cannot specify it. We must hold that there are physical objects, coincident except for one molecule, such that one is a table and the other is not."[5]

Bruno de Finetti, in his landmark 1974 book *Theory of Probability*, quickly gets his readers attention by proclaiming "Probability does not exist; it is a subjective description of a person's uncertainty. We should be normative about uncertainty and not descriptive." He further emphasizes that the frequentist view of probability (objectivist view) "requires individual trials to be equally probable and stochastically independent." In discussing the difference between possibility and probability he states "The logic of certainty furnishes us with the range of possibility (and the possible has no gradations); probability is an additional notion that one applies within the range of possibility, thus giving rise to gradations ('more or less' probable) that are meaningless in the logic of uncertainty." In his book, de Finetti warns us that: "The calculus of probability can say absolutely nothing about reality"; and in referring to the dangers implicit in attempts to confuse certainty with high probability, he states, "We have to stress this point because these attempts assume many forms and are always dangerous. In one sentence: to make a mistake of this kind leaves one inevitably faced with all sorts of fallacious arguments and contradictions whenever an attempt is made to state, on the basis of probabilistic considerations, that something must occur, or that its occurrence confirms or disproves some probabilistic assumptions."

In a discussion about the use of such vague terms as "very probable," "practically certain," or "almost impossible," de Finetti states

> "The field of probability and statistics is then transformed into a Tower of Babel, in which only the most naïve amateur claims to understand what he says and hears, and this because, in a language devoid of convention, the fundamental distinctions between what is certain and what is not, and between what is impossible and what is not, are abolished. Certainty and impossibility then become confused with high or low degrees of a subjective probability, which is itself denied precisely by this falsification of the language. On the contrary,

[5]Historically, a sandpile paradox analogous to Quine's table was discussed much earlier, perhaps in ancient Greek culture.

the preservation of a clear, terse distinction between certainty and uncertainty, impossibility and possibility, is the unique and essential precondition for making meaningful statements (which could be either right or wrong), whereas the alternative transforms every sentence into a nonsense.''

Probability density functions versus membership functions

The early debate in the literature surrounding the meaning of probability and fuzzy logic also dealt with the functions in each theory; i.e., probability density functions (PDFs) and membership functions (MFs). The early claim by probabilists was that membership functions were simply probability density functions couched in a different context, especially in the modeling of subjective probability. The confusion perhaps stemmed from both the similarity between the functions and the similarity between their applications—to address questions about uncertainty. However, this confusion still prevails to a limited extent even today. For example, in the paper (Laviolette et al. (1995)), the authors provide a quote from the literature that intimated that membership functions are "probabilities in disguise.'' When one looks at the ontological basis for both these functions, it is seen that the only common feature of an MF and a PDF is that both are nonnegative functions. Moreover, PDFs do not represent probabilities and, while the maximum membership value allowed by convention is unity, there is no maximum value for a PDF. On the other hand, a cumulative probability distribution function (CDF), the integral of the PDF, is a nonnegative functions defined on the unit interval [0, 1] just as an MF. A CDF is a monotonically increasing function, and an MF is not. The integral of a PDF (the derivative of the CDF) must equal unity, and the area under an MF need not equal unity.[6] Based on these fundamental differences, it is not so obvious how either function could be a disguise for the other, as claimed by some. However, we recognize that the attempts to overcome, ignore, or mask the differences may be the result of wishful thinking, either from wanting to show that one theory is superfluous because its features can be captured by the other, or wanting to bridge the gap between the two.

One area in which PDFs and MFs have been equated is that of Bayesian probability analysis (Laviolette (1995), Cheeseman (1988)). In this case, the likelihood function in Bayes' rule, the function containing new information and used to update the prior probabilities, need not be a PDF. In Bayes' rule, there is a normalization taking place that ensures the posterior distribution (updated prior) will be a PDF, even if the likelihood function is not a PDF. In many applications of Bayes' rule, this likelihood function measures a degree of belief and, by convention, attains a value of unity. Geometrically, the likelihood function can be used in the same context as an MF. In fact, Laviolette uses this similarity to point out that fuzzy control can be approached in the same way as probabilistic control (see Chapter 8 for a more definitive discussion of this claim), and Cheeseman uses this similarity to mistakenly imply that membership functions are simply likelihood functions. But the bottom line is that subjective probabilities—those arising from beliefs rather than from relative frequencies—still must adhere to the properties mentioned above for PDFs and CDFs. They cannot be equated with MFs without making extreme assumptions to account for the differences. Another fundamental difference is that MFs measure degrees of belongingness. They do not measure likelihood (except in a similarity sense, i.e., likeliness) or degrees of belief or frequencies, or perceptions of chance, and the area under their curve is not constrained to a value of unity. They measure something less measurable—set membership in ambiguous

[6]While the area under an MF is not a constraint on these functions, it is a useful metric in some applications involving defuzzification methods (see Chapter 2).

or vague sets. Although both membership values and probabilities map to the unit interval [0, 1], they are axiomatically different (see the discussion in Chapter 5).

James Bezdek gives some excellent examples of the differences between membership values and probabilities (Bezdek (1993, 1994a, 1994b)). He summarizes much of the debate by pointing out that the probability-only advocates base their arguments on one of two philosophical themes: (i) nonstatistical uncertainty does not exist, or (ii) maybe nonrandom uncertainty does exist, but a probability model still is the best choice for modeling all uncertainties. Theme (i) engenders a debate even within the statistical community because there is a nice variety of choices on which to base a probabilistic approach: relative frequency, subjectivity, or the axiomatic approach (von Mises (1957)). In theme (ii), it has been argued by many (e.g., Klir (1989, 1994)) that nonstatistical forms of uncertainty do exist and that numerous theories address them in very useful ways.

Bezdek (1993) gives the following example that, in turn, has spawned numerous other examples. Suppose you are very thirsty and have the chance to drink the liquid from one of two different bottles, A and B. You must decide from which bottle to drink using the following information: bottle A has 0.91 membership in the set of potable liquids, whereas bottle B has a probability of 0.91 of being potable. For the sake of illustration, we will define this probability as a frequentist interpretation. Most people would choose bottle A since its contents are at least "reasonably similar" to a potable liquid and because bottle B has a 9% chance of being unsavory, or even deadly. Moreover, if both bottles could be tested for potability, the membership value for bottle A would remain unchanged, while the probability for bottle B being potable would become either 0 or 1—it is either potable or not potable after a test.[7] These two different kinds of information inherently possess philosophically different notions: fuzzy memberships represent similarities of objects to imprecisely defined properties, and probabilities convey assessments of relative frequencies.

Some probabilists have claimed that natural language is precise, indeed, and that there is no need to use a method like fuzzy logic to model linguistic imprecision. An example dealing with imprecision in natural language is also given by Bezdek (1993). Suppose that as you approach a red light, you must advise a driving student when to apply the brakes. Would you say "begin braking 74 feet from the crosswalk," or would your advice be for the student to "apply the brakes pretty soon"? As an exercise, think of how you would address this uncertainty with probability theory.[8] In this case, precision in the instruction is of little value. The latter instruction would be used by humans since the first command is too precise to be implemented in real time. Sometimes, striving for precision can be expensive, or adds little or no useful information, or both. Many times, such uncertainty cannot be presumed to be, or adequately assessed as, an element of chance.

Bayes' rule

Cheeseman (1988) gives an eloquent discussion of the utility of Bayesian probability in reasoning with subjective knowledge. He and Laviolette (1985) suggest that there is no need to use fuzzy set theory to model subjective knowledge since Bayesian methods are more than adequate for this. However, even the subjective interpretation of probability still has the restrictions of that theory. For example, there is no apparent difference in the way one would model ignorance that arises from lack of knowledge or ignorance that arises from

[7]We could argue that the subjective probability would also remain unchanged; however, this is more of a frequentist interpretation.

[8]Using probabilities, we could estimate the required distance and then instruct the student to brake where he/she maximized his/her probability assessment of stopping in time.

only knowing something about a random process (McDermott (1988)). Moreover, others still question the utility of Bayesian models in epistemology (Dempster (1988), Dubois and Prade (1988)).

Shafer (1987) makes the following points about Bayes' formula $P(A|U_0) = P(U_0|A)P(A)/P(U_0)$:

> "[It] can serve as the symbolic expression of a rule. This rule, the Bayesian rule of conditioning, says that when new knowledge or evidence tells us that the correct answer to the question considered by U is in the subset U_0, we should change our probability for another subset A from $P(A)$ to the new probability given by this quotient, above. Strictly speaking, this rule is valid only when the receipt of the information, U_0, is itself a part of our probability model—i.e., it was foreseen by the protocol. In practice, however, the rule is used more broadly than this in subjective probability judgment. This broader use is reasonable, but it must be emphasized that the calculation is only an argument—an argument that is imperfect because it involves a comparison of the actual evidential situation with a situation where U_0 was specified by a protocol, as one of the possibilities for the set of still possible outcomes at some point in the game of chance. The protocol is the specification, at each step, of what may happen next, i.e., the rules of the game of chance."

Alleged conjunction fallacy

Conjunction in fuzzy logic is modeled with the minimum operator as the smaller of the membership values between two sets. Woodall (1997) and Osherson and Smith (1981) cite the conjunction fallacy as one of the frailties of fuzzy logic. Others, such as McNeil and Freiberger (1993), have called this an intriguing cognitive puzzle that is not a frailty of fuzzy logic but rather a strength, as it exposes with human reasoning the reality of conjunctive processes. Basically, the cognitive puzzle can be illustrated with the following example: The intersection between the set "pet" and the set "fish" produces a membership value of the conjunction "pet fish" that is the smaller of the membership values for "pet" and for "fish" for a particular species of fish in the universe of vertebrates. Osherson and Smith considered that guppy, as a specific species, should have a higher value in the conjunctive set "pet fish" than it does in either "pet" or "fish"; it is not a common pet, nor a common fish, they argued, but it is certainly a strong prototype of the concept "pet fish." In explaining the minimum operator as a preferred operator for the conjunction of fuzzy sets, Zadeh (and others, e.g., Oden (1984) and Rosch (1973)) pointed out that humans *normalize* in their thinking of concepts, whereas strict logical operations don't involve such normalization. Being in the intersection of set "pet" and set "fish," i.e., in the set "pets and fish," is different from being in the set "pet fish." A linguistic modification (pet modifies fish) does *not* equate to a logical conjunction of concepts (sets). Other, more easily understood examples are "red hair" or "ice cube," where the terms "red" and "cube" are altered significantly from their usual meaning. In thinking about *pet–fish* people normalize; they compare a guppy to other pets and other fish and the guppy has a low membership, but when they compare it only to other *pet–fish*, it grades higher. Such cognitive puzzles, which lead to faulty human reasoning, have long been discussed in the literature (Tversky and Kahneman (1983)).

Normalization is typical in many uncertainty operations. A conditional probability is a normalization and, as we have seen in the previous section, Bayes' rule also involves normalization. For example, the probability of getting a 3 on one toss of a die is $\frac{1}{6}$, but the probability changes to $\frac{1}{3}$ if the result is conditioned (i.e., normalized) on the outcome being an odd number (i.e., 1, 3, or 5). In their arguments about "pet fish," Osherson and Smith implicitly normalized for their example of guppies.

Alleged disjunction contradiction

Another allegation about fuzzy logic suggested by some probabilists in the literature (Hisdal (1988b), Laviolette et al. (1985), Lindley (1987a)) has been termed the "disjunctive contradiction." In describing a fuzzy set modeling the heights of people, Hisdal (1988b) complains that the logical union between the fuzzy sets *medium height* and *tall height* should *not* be less than unity between the prototypical values of medium and tall. She calls this the "sharp depression" in the grade of membership curve for two adjacent linguistic labels (see Figure 1.1).

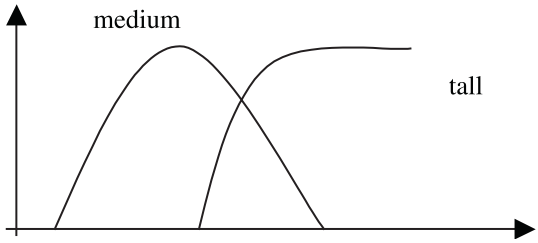

Figure 1.1.

In other words, Hisdal's question is, how could any person's membership in the set "medium OR tall" (which, for fuzzy operations, involves the maximum operator) be less than the largest membership of either if that person's height was between the prototype heights representing medium height and tall? The answer is quite simple. As the figure shows, a "tall" person can have a low membership in the set "medium." The set "tall" is not a subset of the set "medium," but there is an overlap—there exists a common ground in which members of both sets coexist with lower membership. Hisdal confused "medium or tall" with the set "at least medium." Hence there is no inconsistency in the use of the maximum operator here for disjunction, as stated by Hisdal, but rather a misinterpretation of the meaning of disjunction—or logical union. Finally, the argument of Osherson and Smith (1981) on this issue confuses the operations of "sum" and "max" in their arguments on the union operator; the former is an operator on functions and the latter is an operator on sets.

Those controversial excluded-middle laws

For crisp sets, Lindley (1971) showed that the excluded-middle laws result from the three basic laws of probability theory (convexity, addition, and multiplication) when he surmised that an event, E, and its negation, not-E, are mutually exclusive and their union exhaustive. He implicitly assumes that event occurrence is binary—it does or does not occur. Gaines (1978) showed that a probability theory results only after the law of the excluded-middle is added as an axiom to a general uncertainty logic. However, in either case, the critical assumption is that there is no provision for the common ground; as termed by Osherson and Smith, "there can be no apple that is not an apple." Lindley is convinced—he claims that "uncertainties are constrained by the laws of probability" and thereby ignores vagueness, fuzziness, ambiguity, and linguistic imprecision. The laws of probability theory work wonderfully with sets whose boundaries are clear and unambiguous, but a theory is also needed for sets with fuzzy boundaries.

On this same matter, Osherson and Smith (1981), and certainly many others at that time, argued that the "apple that is not an apple" concept is logically empty. Hisdal (1988b) erroneously termed this the *complementation paradox*. In fuzzy set theory, locutions of the type "tomatoes are both fruit and not fruit" can be supported, but in classical Aristotelian logic they cannot. Basically, how can the complement of a fuzzy set, where the set itself does not enjoy full membership over its support, be null throughout that support? The only way to accept the excluded-middle is to accept only "all or nothing" memberships. For example, if the concept "tall" is vague, then so is the concept "not tall." In this case, why would the conjunction "tall *and* not tall" still not be ambiguous to some degree (i.e., not null everywhere)? Alternatively, why wouldn't the disjunction "tall *or* not tall" have less than full membership everywhere?

Classical sets deal with sets whose boundaries are crisp, but very few concepts yield crisp boundaries—in reality, they don't exist (see Black (1937)). As espoused by Howard Gardner (1985), the greatest proponents of crispness—Socrates, Locke, Wittgenstein—all had labored under the deepest illusion!

There are at least two reasons why the law of the excluded-middle might be inappropriate for some problems. First, people may have a high degree of belief about a number of possibilities in a problem. A proposition should not be "crowded out" just because it has a large number of competing possibilities. The difficulties people have in expressing beliefs consistent with the axioms of probability logic are sometimes manifested in the rigidity of the law of the excluded-middle (Wallsten and Budescu (1983)). Second, the law of the excluded-middle results in an inverse relationship between the informational content of a proposition and it probability. For example, in a universe of n singletons, as more and more evidence becomes available on each of the singletons, the relative amount of evidence on any one diminishes (Blockley (1983)). This makes the excluded-middle laws inappropriate as a measure of modeling uncertainty in many situations.

Since the law of the excluded-middle (and its dual, the law of contradiction) is responsible for spawning a theory of probability from a general uncertainty theory (see Chapter 5), it is interesting to consider examples where it proves cumbersome and counterintuitive. David Blockley, in civil engineering, and Wallsen and Budescu, both in business and management, all argued in 1983 the points made in the previous paragraph.

Consider a hypothetical example involving several alternatives. A patient goes to two doctors to get opinions on persistent head pain. Both doctors consider four possible causes: stress headache (H), concussion (Co), meningitis (M), or a tumor (T). Prior to asking the patient to undergo more extensive and expensive tests, both doctors conduct a routine office examination, at different times and locations, and come to the same conclusions: they give little weight to meningitis but considerable weight to the other three possible causes. When pressed to give numerical values for their assessments (their degrees of belief) they give the following numbers: $w(T) = 0.8$, $w(H) = 0.8$, $w(Co) = 0.8$, $w(M) = 0.1$. Of course, a probabilist would normalize these to obtain probabilities and to get their sum to equal unity; i.e., $p(T) = 0.32$, $p(H) = 0.32$, $p(Co) = 0.32$, $p(M) = 0.04$. The problem arises when the probabilist informs the doctor that the probability that the patient's pain is due to something "other than a tumor" is $0.68 = 1 - p(T)$, which is a number higher than the normalized value for any of the other three possible causes! This runs counter to the beliefs of the doctors, who feel that a concussion and a headache are also strong possibilities based on their "limited" diagnosis. Such is the requirement imposed by the excluded-middle laws. Based on this information, should the patient spend more money on tests? If the doctors' original weights were assigned as possibilities, these weights would not need any normalization.

On the other hand, Bayesianists would claim that the original weights were simple likelihood values obtained by diagnosis to be used to modify prior probabilities of the historical evidence of the four possible causes of the patient's head pain. For example, suppose that the prior probabilities ϕ_i are $\phi(H) = 0.8$, $\phi(T) = 0.05$, $\phi(M) = 0.01$, and $\phi(Co) = 0.30$; then from Bayes' formula, we can predict for each potential cause of the patient's head pain (HP), C_i,

$$p(C_i|HP) = \frac{w(HP|C_i)\phi(C_i)}{\sum_{i=1}^{4} w(HP|C_i)\phi(C_i)},$$

where $w(HP|C_i)$ is the likelihood value for the ith possible cause of the head pain, as given in the previous paragraph. The resulting posterior probabilities (updated priors) would be calculated as

$$\begin{aligned} w(C_1|HP) &= w(H|HP) = 0.695, \\ w(C_2|HP) &= w(T|HP) = 0.043, \\ w(C_3|HP) &= w(C|HP) = 0.261, \\ w(C_4|HP) &= w(M|HP) = 0.001. \end{aligned}$$

Now the probability that the head pain is due to a tumor is 0.043, slightly higher than the historical evidence due to the doctors' rather high weight assigned after a rudimentary diagnostic exam. However, it remains that the probability (now updated) that it is a cause other than a tumor is 0.957, a number much higher than the rest of the potential causes. Therefore, whether one takes the philosophy of a Bayesianist or a frequentist, the excluded-middle law can be counterintuitive to people's initial judgments; in this case, the doctors' original assessments that three of the four possible causes all have a high likelihood hold. Since the evidence (or degrees of belief, or likelihood values, etc.) comes from humans, it is *not* a justification for using probability theory, as is often claimed by Lindley and others, to suggest that humans must be forced to think in "coherent" terms (the term "coherent" borrowed by probabilists to describe the additivity axiom of probability theory). Meyer and Booker (2001) in their formal elicitation methods advocate just the opposite approach. The highest quality of expert information comes from allowing the experts to express their beliefs free of such probabilistic constraints. Of course, we expect (or perhaps hope) that humans exhibit coherence as a linguistic feature of reasoning. However, studies (Meyer and Booker (2001)) have too often shown this is not the case, and we should not force them to obey the coherence property of probability theory. More to the point, we also should not force them to obey a set of axiomatic structures that run counter to their thinking or problem solving because such restrictions induce cognitive and motivational biases that both degrades the quality and changes the nature and content of the information given.

Probability theory needs an infusion of fuzzy logic

In a 1998 email communication to colleagues at the University of California at Berkeley, Lotfi Zadeh stated "probability theory needs an infusion of fuzzy logic to enhance its ability to deal with real-world problems." Earlier, in their 1988 series of commentaries on Peter Cheeseman's essay on his controversial *computer understanding*, Enrique Ruspini, Ron

Yager, and Lotfi Zadeh all made suggestions (*Comput. Intell.*, 4 (1988), pp. 57–142) about how powerful probability theory would be if it were able to consider notions of overlapping sets—however, then it wouldn't be the probability theory that we all know (or think we know) today. In a 1996 paper, Zadeh alludes to the fact that the historically main contribution—but not the only one—of fuzzy logic will be its ability to help us develop a new field, computing with words (CW). Here he states that a frequently asked question is, "What can be done with fuzzy logic that cannot be done equally well with other methodologies, e.g., predicate logic, probability theory, neural network theory, Bayesian network, and classical control?" One such answer involves the nascent field of CW.

In 1999, Zadeh stated that information granulation is associated with an increasing order of precision: from interval to fuzzy to evidence to random, the granulation goes from coarse to fine. He cites the successes of precision on monumental scientific tasks involving the atomic bomb, moon landings, supercomputers, telescopes, and carbon dating of rock. He also cites failures on more simple tasks: we cannot build robots that can move with the agility of animals; we cannot automate driving tasks in heavy traffic; we cannot automate language translation; we cannot summarize nontrivial stories automatically; we cannot model the behavior of economic systems; and we cannot construct machines that can compete with children in the performance of simple physical and cognitive tasks. Hoadley and Kettering (1990) point out that follow-up studies to the Space Shuttle *Challenger* disaster revealed that "relevant statistical data had been looked at incorrectly… probabilistic-risk assessment methods were needed to support NASA, and… its staff lacked specialists and engineers trained in the statistical sciences." They admit, however, that statistical science does not seem to deal well with the probabilistic risk assessments of complex engineered systems, because "there are too little data in the traditional sense."

The subject of specifically how fuzzy set theory can actually enhance the effectiveness of probability theory when dealing with some problems has been articulated recently by Zadeh (1995, 1996). In these works, Zadeh points out that probability theory is very useful in dealing with the uncertainty inherent in measurements or objects that can be measured; however, it is not very useful in dealing with the uncertainty imbedded in perceptions by humans. The former involves crisp sets, while the latter involves fuzzy sets. Examples of the latter are the following:

1. What is the probability that your tax return will be audited?
2. What is the probability that tomorrow will be a warm day?
3. Team A and Team B played each other 10 times. Team A won the first seven times, and Team B won the last three times. What is the probability that Team A will win the next game between the two teams?
4. Most young women are healthy. Mary is young. What is the probability that Mary is healthy?

In question 1, the difficulty arises from the basic property of conditional probability—namely, given $p(x)$, all that can be said is that the value of $p(x|y)$ is between 0 and 1. Thus if we know that 1% of all tax returns are audited, this tells us little about the probability that your tax return will be audited, except in a bounding sense that your return is one of the many from which the audit sample is taken. You may have an estimate of how often you have been audited, which could be used in a Bayesian probability model, and you can always provide your subjective belief as a probability for an estimate. If we have more detailed information about you, e.g., income, age, residence, occupation, etc., a better estimate might be possible from IRS data based upon the fraction of returns that are audited in a certain category, but this detail will never reach the level of a single individual. Probability theory alone cannot directly handle this question at the appropriate level.

In question 2, the difficulty is that the warmth is a matter of degree. The event "a warm day" cannot be defined precisely; it is a fuzzy event based on personal preference.

In question 3, the difficulty is that we are dealing with a time series drawn from a nonstationary process. In such cases, probabilities do not exist. In fact, that statement is the familiar quote attributed to de Finetti (1974) in his discussion about the origins of subjective probabilities, as enunciated earlier in this debate summary.

In question 4, we have a point of common sense reasoning, and probability theory is inadequate to handle the uncertainty in concepts such as "most young women" or "healthy."

To one degree or another, all four questions involve the same theme: the answers to these questions are not numbers; they are linguistic descriptions of fuzzy perceptions of probabilities. As pointed out by Zadeh (1999), these kinds of questions can be addressed by an enhanced version of probability theory—a theory enhanced by a generalization that accounts for both fuzzification and granulation. Here fuzzy granulation reflects the finite cognitive ability of humans to resolve detail and store information. Classical probability theory is much less effective in those fields in which dependencies between variables are not well defined; the knowledge of probabilities is imprecise and/or incomplete; the systems are not mechanistic; and human reasoning, perception, and emotion play an important role. This is the case in many fields ranging from economics to weather forecasting.

Our summary of the debate

In 1997, in discussing the choices available to analysts who are modeling uncertainty in their systems, Nguyen had this to say:

> "Of course, we are free to model mathematically the way we wish, but unless the modeling is useful for applications, the modeling problem may be simply a mathematical game. In the axiomatic theory of probability we do not allow all possible subsets of a sample space to be qualified as events. Instead we take a sigma-field of subsets of the sample space to be the collection of events of interest. It is not possible to extend the sigma-additive set-function to the whole power set of the reals. Acknowledging that randomness and fuzziness are different types of uncertainties, there is no compelling reason to believe that their associated calculi are the same."

The arguments of Lindley (1982) have been countered on mathematical grounds by Goodman, Nguyen, and Rogers (1991). "Saying that fuzzy sets theory is complementary rather than competitive does not presume deficiencies in probability."

Finally, rather than debate what is the correct set of axioms to use (i.e., which logic structure) for a given problem involving uncertainty, one should look closely at the problem, determine which propositions are vague or imprecise and which ones are statistically independent or mutually exclusive, determine which ones are consistent with human cognition, and use these considerations to apply a proper uncertainty logic, with or without the law of the excluded-middle. By examining a problem so closely as to determine these relationships, one finds out more about the structure of the problem in the first place. For example the assumption of a strong truth-functionality (for fuzzy logic) could be viewed as a computational device that simplifies calculations, and the resulting solutions would be presented as ranges of values that most certainly bound the true answer if the assumption is not reasonable. A choice of whether fuzzy logic is appropriate is, after all, a question of balancing the model with the nature of the uncertainty contained within it. Problems without an underlying physical model, problems involving a complicated weave of technical, social, political, and economic factors, and problems with incomplete, ill-defined, and inconsistent information where conditional probabilities cannot be supplied or rationally formulated perhaps are candidates for fuzzy set applications. It has become apparent that the engineering profession is overwhelmed by these sorts of problems.

The argument, then, appears to be focused on the question, "Which theory is more practical in assessing the various forms of uncertainty that accompany any problem?" The major intent of this book, therefore, is to shed light on this question by comparing just two of the theories, fuzzy sets and probability, for various applications in different fields. While we know that this book will not end the philosophical debate (which could be waged for many more decades), will not convert either the protagonists or antagonists of a particular theory, and will not lessen the importance or usefulness of any of the theories, it is our hope that its contents will serve to educate the various audiences about the possibility that the complementary use of these two theories in addressing uncertainties in many different kinds of problems can be very powerful indeed. Hence we hope that this book will help "bridge the gap" between probability theory and other nonprobabilistic alternatives in general, and fuzzy set theory in particular.

1.3 Fuzzy logic and probability: The best of both worlds

In contrast to the literature cited in section 1.2, there is a growing list of recent papers illustrating the power of combining the theories. What follows are very brief, very incomplete descriptions of a few of the works readily available to the editors of this book:

- *Cooper and Ross* (1998) *on subjective knowledge in system safety*. The authors describe a series of mathematical developments combining probability operations and fuzzy operations in the area of system reliability. They provide sort of a shopping cart of potential ideas in combining the two methods to assess both modeling and parametric uncertainty within the context of assessing the safety and reliability of manmade systems.

- *Ross et al.* (1999) *on system reliability*. In this work, the authors discuss a way to fuse probabilities, using fuzzy norms, to address the fact that real systems contain components whose interactions are between the extremes of completely independent and completely dependent. The fuzzy norms are based on Frank's t-norms, which can be shown to contain, as a family of norms, the probabilistic and fuzzy norms. The paper further shows that memberships and probability density functions can be combined using the same norms, thus paving the way for an algebraic approach to fuse both (random and fuzzy) kinds of information.

- *Smith et al.* (1997, 1998) *on uncertainty estimation*. In a series of two papers, Smith et al. used the graphical methods typical in fuzzy control applications to develop probabilistic density functions of the uncertainties inherent in the working parts of a system to estimate the systems reliability. In this case, the output MFs in the inference of a fuzzy control method were used to develop the density functions.

- *Zadeh* (1999) *on computing with words*. In this paper, Zadeh shows that a question of the type "What is the probability of drawing a small ball from an urn of balls of various sizes?" can be addressed with a combination of probability and fuzzy set theory. In this work there is a distinction drawn between being able to measure and manipulate numbers (as done with probabilities) and being able to measure and manipulate perceptions (linguistic data).

- *Nikolaidis* (1997–1999) *on designing under uncertainty*. In a sequence of papers (Nikolaidis et al. (1999) and Maglaras et al. (1997)), a contrast is established between probability theory and possibility theory in the area of structural design. In this case, structural failure and survival (safety) usually are taken to be complementary states, even though the transition between the two is gradual for most systems. In designs under uncertainty, it is important to understand how each of the two methods maximizes the expression of safety. In terms of design optimization against failure, the probabilistic optimization tries to reduce the probabilities of failure of the modes that are easiest to control in order to minimize the system failure probability, whereas fuzzy set optimization simply tries to equalize the possibilities of failure of all failure modes in minimizing the system failure possibility. They show that fuzzy set methods are not necessarily more conservative than probabilistic methods when assessing system failure. They also show that fuzzy set methods yield safer designs than probabilistic designs when there is limited data.

- *Rousseeuw* (1995) *on fuzzy and probabilistic clustering*. This work is important because it shows an application in classification where a fuzzy clustering method is actually a collaboration of two approaches: fuzzy theory and statistics. The usefulness of fuzzy clustering is apparent in that it helps the classification process avoid convergence only to local minima. In a completely fuzzy approach, each cluster center is influenced also by the many objects that essentially belong to other clusters. This bias may keep the fuzzy method from finding the true clusters. This problem is overcome by the use of objective functions, which are hybrids of purely fuzzy and purely probabilistic formalisms and which produce a high contrast clustering method capable of finding the true clusters in data.

- *Singpurwalla et al.* (2000) *on membership and likelihood functions*. In this published laboratory report, Singpurwalla et al. explore how probability theory and fuzzy MFs can be made to work in concert so that uncertainty of both outcomes and imprecision can be treated in a unified and coherent manner. In this work, the authors show that the MF can be interpreted as a likelihood if the fuzzy set takes the role of an observation and the elements of the fuzzy set take the role of the hypothesis. They show further that the early work of Zadeh (1968) in defining the probability of a fuzzy event as being equal to the expected value of the event's membership function is proportional to the measure they develop, which is a function of a prior probability of the classification of the event.

As indicated earlier, technology is moving so fast, in fact, that the systems on which we rely for our support and our discourse are becoming more, not less, complex. Decision-making has become so much guesswork under time constraints that analysts rarely have time to develop well-conceived models of today's systems. Complexity in the real world generally arises from uncertainty in various forms. Complexity and uncertainty are features of problems that have been addressed since humans began thinking abstractly; they are ubiquitous features that imbue most social, technical, and economic problems faced by the human race. Why is it then that computers, which have been designed by humans, after all, are not capable of addressing complex issues, that is, issues characterized by vagueness, ambiguity, imprecision, and other forms of uncertainty? How can humans reason about real systems when the complete description of a real system often requires more detailed data than a human could ever hope to recognize simultaneously and assimilate with understanding? It is because humans have the capacity to reason approximately, a capability that computers currently do not have. In reasoning about such systems, humans simply approximate behavior, thereby maintaining only a generic understanding about the problem. Fortunately, for humans' ability to understand complex systems, this generality and ambiguity is sufficient.

When we learn more and more about a system, its complexity decreases and our understanding increases. As complexity decreases, the precision afforded by computational methods become more useful in modeling the system. For systems with little complexity, and hence little uncertainty, closed-form mathematical expressions provide precise descriptions of the systems. For systems that are a little more complex, but for which significant data exists, model-free methods, such as artificial neural networks, provide a powerful and robust means to reduce some uncertainty through learning based on patterns in the available data. Finally, for the most complex systems, where little numerical data exists and where only ambiguous or imprecise information may be available, probabilistic or fuzzy reasoning can provide a way to understand system behavior by allowing us to interpolate approximately between observed input and output situations and thereby measuring, in some way, the uncertainty and variability. The imprecision in fuzzy models and the variability in probabilistic predictions generally can be quite high. The point, however, is to match the model type with the character of the uncertainty exhibited in the problem. In situations where precision is apparent, for example, fuzzy systems are not as efficient as more precise algorithms in providing us with the best understanding of the problem. On the other hand, fuzzy systems can focus on modeling problems with imprecise or ambiguous information. In systems where the model is understood but where variability in parameters of the model can be quite high, a probabilistic model can enrich understanding. Knowledge in using the tools illustrated in the subsequent chapters of this book will allow readers to address the vast majority of problems that are characterized either by their complexity or by their lack of a requirement for precision.

Our philosophy in preparing this book for a world of complex problems is to set a tone of conciliation between fuzzy theory and probability. We admit that there are differences between them. We do not propose that one is better than the other. We agree that one is more mature than the other. We admit that some applications are better served by one or the other. However, we contend that some applications can benefit from a hybrid approach of the two. Our intent is to help probabilists understand fuzzy set theory and fuzzy logic and to help fuzzy advocates see that probability theory, in conjunction with fuzzy logic, can be very powerful for some types of problems. Another objective of this book is to help the user select the best tool for the problem at hand. Our intended audience is statisticians, probabilists, and engineers.

1.4 Organization of the book

The book is organized into two major divisions. The chapters in Part I cover fundamental topics of probability, fuzzy set theory, uncertainty, and expert judgment elicitation. In Part I, we move from specific explanations of probability theory to more general descriptions of various forms of uncertainty, culminating in the elicitation of human evaluations and assessments of these uncertainties. Part II includes applications and case studies that illustrate the uses of fuzzy theory and probability and hybrid approaches.

Chapter 2 covers fuzzy set theory, fuzzy logic, and fuzzy systems. The origins of membership sets and functions are included, along with a discussion on the law of the excluded-middle. Fuzzy relationships and operations are illustrated with examples. Fuzzy logic and classical logic are compared.

Chapter 3 covers the foundations of probability theory and includes a discussion on the different interpretations of probability. Topics relating to probability distribution functions are presented as they relate to uncertainty and to hybrid fuzzy/probability approaches. The last section of this chapter provides a transition from probability to fuzzy theory by discussing the concepts of data, knowledge, and information.

One of the probability theories from Chapter 3 has Bayes' theorem at its core. Details regarding methods based on this approach to probability are presented in Chapter 4. Interpretations, applications, and issues associated with the Bayesian-based methods are presented. One reason for the special emphasis on these methods is the link it provides to fuzzy MFs, as described in the last section of the chapter.

Chapter 5 considers the uses of fuzzy set theory and probability theory. As such, it covers and compares the topics of vagueness, imprecision, chance, uncertainty, and ambiguity as they relate to both theories. Historical development of these topics is discussed, leading up to recent research work in possibility theory.

The process of defining and understanding the various kinds of uncertainty crosses into another subjective arena: the elicitation of expert knowledge. Chapter 6 provides guidelines for eliciting expert judgment in both realms of probability and fuzzy logic. While this chapter provides many of the elicitation fundamentals, it does so using two different applications for illustration.

Part II of the book, on applications and case studies, begins with Chapter 7 on image enhancement processing. This chapter compares image enhancement via the modification of the PDF of the gray levels with the new techniques that involve the use of knowledge-based (fuzzy expert) systems capable of mimicking the behavior of a human expert. This chapter includes some color images.

Chapter 8 provides a brief, basic introduction to engineering process control. It begins with classical fuzzy process control and introduces the relatively new idea of probabilistic process control. Two different scenarios of a problem are examined using classical process control, proportional-integral-derivative (PID) control, and probabilistic control techniques.

Chapter 9 presents a structural safety analysis study for multistory framed structures using a combined fuzzy and probability approach. The current probabilistic approach in treating system and human uncertainties and its inadequacy is discussed. The alternative approach of using fuzzy sets to model and analyze these uncertainties is proposed and illustrated with examples. Fuzzy set models for the treatment of some uncertainties in reliability assessment complement probabilistic models and are readily incorporated into the current analysis procedure for the safety assessment of structures.

Reliability is usually considered a probabilistic concept. Chapter 10 presents a case study of the reliability of aircraft structural integrity. The usefulness of fuzzy methods is demonstrated when insufficient test data are available and when reliance is necessary on the use of expert judgment.

Chapter 11 presents another case study in reliability. It begins with a complex reliability problem in the automotive industry, carries it through using Bayesian methods, and finally covers how certain aspects of the problem would benefit from the use of MFs. As in Chapter 10, here the use of expert judgment is prevalent when it is not feasible or practical to obtain sufficient test data for traditional reliability calculations.

Statistical process control (SPC) is a long-standing set of probability-based techniques for monitoring the behavior of a manufacturing process. Chapter 12 describes an application of fuzzy set theory for exposure control in beryllium part manufacturing by adapting and comparing fuzzy methods to the classical SPC. This very lengthy chapter considers the topic of fuzzy SPC as sufficiently important and novel (as introduced in the third author's Ph.D. dissertation); the chapter is essentially a tutorial on standard SPC extended to incorporate fuzzy information and fuzzy metrics.

Combining, propagating, and accounting for uncertainties is considered for fault tree logic models in Chapter 13. This scheme is based on classical (Boolean) logic. Now this method is expanded to include other logics. Three of the many proposed logics in the literature considered here are Lukasiewicz, Boolean, and fuzzy. This is a step towards answering the questions "Which is the most appropriate logic for the given situation?" and "How can uncertainty and imprecision be addressed?"

Chapter 14 investigates the problem of using fuzzy techniques to quantify information from experts when test data or a model of the system is unavailable. MFs are useful when the uncertainty associated with this expertise is expressed as rules. Two examples illustrate how a fuzzy-probability hybrid technique can be used to develop uncertainty distributions. One example is for cutting tool wear and the other is for reliability of a simple series system.

Chapter 15 introduces the use of Bayesian belief networks as an effective tool to show flow of information and to represent and propagate uncertainty based on a mathematically sound platform. Several illustrations are presented in detecting, isolating, and accommodating sensor faults. A probabilistic representation of sensor errors and faults is used for the construction of the Bayesian network. Fuzzy logic forms the basis for a second approach to sensor networks.

These chapters represent the collective experience of the authors' and editors' efforts and research to bridge the gap between the theories of fuzzy logic and probability. As stated above, our goal is to continue to develop, support, and promote these techniques for solving the complex problems of modern information technology.

References

T. Arciszewski (1999), *History of Information Technology*, discussion at the American Society of Civil Engineers Committee on Artificial Intelligence and Expert Systems Workshop, George Mason University, Fairfax, VA.

T. Bayes (1763), An essay towards solving a problem in the doctrine of chances, *Philos. Trans. Roy. Soc.*, 53, pp. 370–418.

T. Bement (1996), private communication.

J. Bezdek (1993), Fuzzy models—what are they and why? *IEEE Trans. Fuzzy Systems*, 1, pp. 1–6.

J. Bezdek (1994a), Fuzziness vs. probability: Again(!?), *IEEE Trans. Fuzzy Systems*, 2, pp. 1–3.

J. Bezdek (1994b), The thirsty traveler visits Gamont: A rejoinder to "Comments on fuzzy sets—what are they and why?" *IEEE Trans. Fuzzy Systems*, 2, pp. 43–45.

M. Black (1937), Vagueness: An exercise in logical analysis, *Internat. J. Gen. Systems*, 17, pp. 107–128.

D. Blockley (1983), Comments on "Model uncertainty in structural reliability," by Ove Ditlevsen, *J. Structural Safety*, 1, pp. 233–235.

P. Cheeseman (1986), Probabilistic versus fuzzy reasoning, in *Uncertainty in Artificial Intelligence*, L. Kanal and J. Lemmer, eds., North-Holland, Amsterdam, pp. 85–102.

P. Cheeseman (1988), An inquiry into computer understanding, *Comput. Intell.*, 4, pp. 58–66.

J. A. Cooper and T. Ross (1998), The application of new mathematical structures to safety analysis, in *Proceedings of the IEEE Conference on Fuzzy Systems* (*FUZZ-IEEE* '98), Vol. 1, IEEE, Piscataway, NJ, pp. 692–697.

B. de Finetti (1974), *Theory of Probability*, John Wiley, New York.

A. Dempster (1988), Comments on "An inquiry into computer understanding" by Peter Cheeseman, *Comput. Intell.*, 4, pp. 72–73.

D. Dubois and H. Prade (1988), Comments on "An inquiry into computer understanding" by Peter Cheeseman, *Comput. Intell.*, 4, pp. 73–76.

C. Elkan (1993), The paradoxical success of fuzzy logic, in *Proceedings of the American Association for Artificial Intelligence*, MIT Press, Menlo Park, CA, pp. 698–703.

B. Gaines (1978), Fuzzy and probability uncertainty logics, *Inform. and Control*, 38, pp. 154–169.

H. Gardner (1985), *The Mind's New Science: A History of the Cognitive Revolution*, Basic Books, New York.

I. Goodman, H. Nguyen, and G. Rogers (1991), On scoring approach to the admissibility of uncertainty measures in expert systems, *J. Math. Anal. Appl.*, 159, pp. 550–594.

E. Hisdal (1988a), Are grades of membership probabilities? *Fuzzy Sets Systems*, 25, pp. 325–348.

E. Hisdal (1988b), The philosophical issues raised by fuzzy set theory, *Fuzzy Sets Systems*, 25, pp. 349–356.

A. Hoadley and J. Kettering (1990), Communication between statisticians and engineer/physical scientists, *Technometrics*, 32, pp. 243–274 (includes 11 commentaries).

G. Klir (1989), Is there more to uncertainty than some probability theorists might have us believe? *Internat. J. Gen. Systems*, 15, pp. 347–378.

G. Klir (1994), On the alleged superiority of probabilistic representation of uncertainty, *IEEE Trans. Fuzzy Systems*, 2, pp. 27–31.

G. Klir and T. Folger (1988), *Fuzzy Sets, Uncertainty, and Information*, Prentice–Hall, Englewood Cliffs, NJ.

B. Kosko (1990), Fuzziness vs. probability, *Internat. J. Gen. Systems*, 17, pp. 211–240.

B. Kosko (1994), The probability monopoly, *IEEE Trans. Fuzzy Systems*, 2, pp. 32–33.

M. LAVIOLETTE AND J. SEAMAN (1992), Evaluating fuzzy representations of uncertainty, *Math. Sci.*, 17, pp. 26–41.

M. LAVIOLETTE, J. SEAMAN, J. BARRETT, AND W. WOODALL (1995), A probabilistic and statistical view of fuzzy methods, *Technometrics*, 37, pp. 249–292 (with seven commentaries).

D. LINDLEY (1971), *Making Decisions*, 2nd ed., John Wiley, New York.

D. LINDLEY (1982), Scoring rules and the inevitability of probability, *Internat. Statist. Rev.*, 50, pp. 1–26 (with seven commentaries).

D. LINDLEY (1987A), The probability approach to the treatment of uncertainty in artificial intelligence and expert systems, *Statist. Sci.*, 2, pp. 17–24.

D. LINDLEY (1987B), Comment: A tale of two wells, *Statist. Sci.*, 2, pp. 38–40.

G. MAGLARAS, E. NIKOLAIDIS, R. HAFTKA, AND H. CUDNEY (1997), Analytical-experimental comparison of probabilistic methods and fuzzy set based methods for designing under uncertainty, *Structural Optim.*, 13, pp. 69–80.

D. MCDERMOTT (1988), Comments on Cheeseman: Why plausible reasoning? Bayesian inference, *Comput. Intell.*, 4, pp. 91–92.

D. MCNEIL AND P. FRIEBERGER (1993), *Fuzzy Logic*, Simon and Schuster, New York.

M. A. MEYER AND J. M. BOOKER (2001), *Eliciting and Analyzing Expert Judgment: A Practical Guide*, ASA–SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA.

H. NGUYEN (1997), Fuzzy sets and probability, *Fuzzy Sets Systems*, 90, pp. 129–132.

E. NIKOLAIDIS, H. CUDNEY, R. ROSCA, AND R. HAFTKA (1999), Comparison of probabilistic and fuzzy set-based methods for designing under uncertainty, in *Proceedings of the 40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, St. Louis, MO*, Vol. 4, American Institute of Aeronautics and Astronautics, Reston, VA, pp. 2860–2874.

G. ODEN (1994), Integration of fuzzy linguistic information in language comprehension, *Fuzzy Sets Systems*, 14, pp. 29–41.

D. OSHERSON AND E. SMITH (1981), On the adequacy of prototype theory as a theory of concepts, *Cognition*, 9, pp. 35–58.

W. QUINE (1981), *Theories and Things*, Harvard University Press, Cambridge, MA.

E. ROSCH (1973), Natural categories, *Cognitive Psychology*, 4, pp. 328–350.

T. ROSS, C. FERREGUT, R. OSEGUEDA, AND V. KREINOVICH (1999), System reliability: A case when fuzzy logic enhances probability theory's ability to deal with real-world problems, in *Proceedings of the 18th International Conference of the North American Fuzzy Information Society*, R. Dave and T. Sudkamp, eds., IEEE, Piscataway, NJ, pp. 81–84.

P. ROUSSEEUW (1995), Discussion: Fuzzy clustering at the intersection, *Technometrics*, 37, pp. 283–286.

E. RUSPINI (1988), Intelligent understanding and wishful thinking: The dangers of epistemological alchemy, *Comput. Intell.*, 4, pp. 105–117.

B. RUSSELL (1923), Vagueness, *Australasian J. Philos.*, 1, pp. 85–88.

G. SHAFER (1987), Belief functions and possibility measures, in *Analysis of Fuzzy Information*, Vol. 1, J. Bezdek, ed., CRC Press, Boca Raton, FL, pp. 51–84.

N. SINGPURWALLA, T. BEMENT, AND J. BOOKER (2000), *Membership Functions and Probability Measures of Fuzzy Sets*, Report LA-UR 00-3660, Los Alamos National Laboratory, Los Alamos, NM.

R. SMITH, T. BEMENT, W. PARKINSON, F. MORTENSEN, S. BECKER, AND M. MEYER (1997), The use of fuzzy control system techniques to develop uncertainty distributions, in *Proceedings of the Joint Statistical Meetings, Anaheim, CA*, American Statistical Association, Alexandria, VA.

R. E. SMITH, J. M. BOOKER, T. R. BEMENT, M. A. MEYER, W. J. PARKINSON, AND M. JAMSHIDI (1998), The use of fuzzy control system methods for characterizing expert judgment uncertainty distributions, in *Proceedings of the 4th International Conference on Probabilistic Safety Assessment and Management* (*PSAM* 4), Vol. 1: Probabilistic Safety Assessment and Management, A. Mosleh and R. A. Bari, eds., Springer, London, pp. 499–502.

S. STIGLER (1986), Laplace's 1774 memoir on inverse probability, *Statist. Sci.*, 1, pp. 359–378.

A. TVERSKY AND D. KAHNEMAN (1983), Probability, representativeness, and the conjunction fallacy, *Psychological Rev.*, 90, pp. 293–315.

R. VON MISES (1957), *Probability, Statistics, and Truth*, Dover, New York.

T. WALLSTEN AND D. BUDESCU (1983), Encoding subjective probabilities: A psychological and psychometric review, *Management Sci.*, 29, pp. 151–173.

W. WOODALL (1997), An overview of comparisons between fuzzy and statistical methods, (extended abstract 154) in *Proceedings of the Joint Statistical Meetings, Anaheim, CA*, American Statistical Association, Alexandria, VA, p. 174.

L. ZADEH (1965), Fuzzy sets, *Inform. and Control*, 8, pp. 338–353.

L. ZADEH (1995), Discussion: Probability theory and fuzzy logic are complementary rather than competitive, *Technometrics*, 37, pp. 271–276.

L. ZADEH (1996), Fuzzy logic = computing with words, *IEEE Trans. Fuzzy Systems*, 4, pp. 103–111.

L. ZADEH (1998), *PT Needs FL* (*Part* 2), email communication to the BISC (Berkeley Initiative for Soft Computing) Group, Berkeley, CA, November 20.

L. ZADEH (1999), From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions, *IEEE Trans. Circuits Systems*, 45, pp. 105–119.