1

A Brief Introduction and History

THINGS TO LOOK FOR...

- Early views on reality, learning, logic, and reasoning.
- The early classic laws of thought.
- Foundations of fuzzy logic.
- A learning and reasoning taxonomy.
- The mathematics underlying crisp and fuzzy logic.
- Similarities and differences between crisp and fuzzy logic.
- Fuzzy logic and approximate reasoning.
- Fuzzy sets and membership functions.

1.1 Introduction

We open this text with a challenge and a foundation. Whether crisp or fuzzy, whether involving animals, humans, or machines, philosophers, scientists, and educators have studied, debated, and analyzed terms such as *think*, *ponder*, *logic*, *reason*, *philosophize*, or *learn* for centuries. Yet today, our understanding of these processes still has the opportunity to grow. Given such a history, what do we know?

Let us start with learning. Learning is a process that starts (at least) immediately after birth and continues, often unobtrusively, through the remaining years of life. Recent research, however, has found that learning may actually begin months earlier. Nevertheless, the term itself generally evokes childhood memories of old books, pedagogical teachers, and stuffy classrooms on warm spring afternoons when we would rather be outside playing. If we pause and reflect for a moment, we realize that learning is not limited to that proffered by the pendants of previous days but is a natural part of our daily existence. Each time that we encounter a fresh idea, make a new discovery, or solve one of life's many challenges, we are learning; we are growing and enriching our perceived model of our world or potentially what lies in space beyond.

Understanding the concepts of learning and reasoning is playing an increasingly significant role in the modern high-tech design, development, and implementation of perceptrons, neural networks, artificial intelligence, machine learning, and the primary topic of this text, i.e., fuzzy systems. We mentioned two terms: "crisp" and "fuzzy." We now introduce and explore thought and reasoning in such systems.

Introduction to Fuzzy Logic, First Edition. James K. Peckol.

1.2 Models of Human Reasoning

We now move from animals and humans to raise the question that has perplexed for eons. Can a computing machine be designed to think, to reason, to learn, to create, that is, to self-modify? Can such machines learn and operate like a human being? To be able to design and implement such tools and machines, we must first fully understand their intended applications and what these terms actually mean in such contexts.

Throughout history, as an outgrowth of the work of researchers in both the information processing and the epistemological schools, many different models of human reasoning have been explored, proposed, and tried. In each instance, the hypothetical model tries to capture the dynamic nature, inexactness, or intuitive nature of the underlying process.

Frequently, the heuristic character of human reasoning is quantified numerically. This is seen in Lotfi Zadeh's characterization of notions such as *young* or *old* on a mathematical scale or Ted Shortliffe's measures of *belief* and *disbelief*. Herb Simon has countered that people do not reason numerically. Perhaps they do not. However, mathematics is a reasonable first-order approach when attempting to capture the essence of the intuitive inexactness humans so readily accept but which computers have difficulty accommodating. Three or four simple words illustrate the essence of the two philosophies:

|(*yes, no*) → (*maybe*) → (maybe not)| . . . |(crisp) → (fuzzy)|

Let's begin our study by looking at some of the early works. This work is rooted mainly in the studies, writings, and teachings of early Greek philosophers including Socrates, Plato, Aristotle, Parmenides, and Heraclitus. Philosophy in the early days often included mathematics and related reasoning.

1.2.1 The Early Foundation

Socrates was one of the early classic Greek philosophers and is often regarded as the founding father of Western philosophy. He is particularly noted for his creation of the *Socratic Method* in which the teacher repeatedly poses clarifying questions until the student grasps and understands the concept(s) being taught.

Plato, a student of Socrates, formed the first institution of higher learning in the Western world. Reflecting the *Socratic Method*, he wrote dialogues in which the participants discuss, analyze, and dissect a topic from various perspectives. He is also considered the developer of the concept of *forms* in which an ideal world of *forms* exists in contrast to a false world of phenomena.

Aristotle, mentored by Plato, brought symbolic logic and the notion of scientific thinking to Western philosophy. In doing so, he contradicted Plato's ideal world of forms with a more pragmatic view and contributed advances in the branch of philosophy known as *metaphysics*. Among Aristotle's fundamental assertions was that it is impossible to both be something and not be the same thing at the same time. Today, that assertion falls apart in the field of fuzzy logic.

Parmenides was a pre-Socratic philosopher known as the *Philosopher of Changeless Being*. He believed that there are two views of reality. One is the way of truth or fact and the other is that change is impossible and existence is unchanging. Over the course of his career, he also had a significant influence on a young Plato.

Heraclitus views contradicted those of Parmenides with his insistence on ever-present change as fundamental to the universe. Such a view was reflected in his saying: "*No man ever steps in the same river twice*." His beliefs continued to identify him as one of the founders of the branch of metaphysics referred to as *ontology,* which deals with the nature of being.

From these early philosophers and their different views on reality, learning, and reasoning, we have the following classic laws of thought:

1.2.1.1 Three Laws of Thought

The *laws of thought* are stipulated to be the rules by which rational discourse can be considered to be based. Thus, they are rules that apply, without exception, to any subject matter of thought. The three laws are as follows:

- The Law of Identity
- The Law of Excluded Middle
- The Law of Non-contradiction

- Law of Identity
  The *law of identity* simply states that each entity or thing is identical to itself.
- Law of Non-contradiction
  The *law of non-contradiction* is given by Aristotle. Among his other assertions, he contends that when trying to determine the nature of reality, the following principle applies. A substance cannot have a quality and yet simultaneously not have that same quality.
- Law of Excluded Middle
  The fundamental *law of excluded middle*, which originated with Plato, states that for any proposition, that proposition is either true or its negation is true.

  Such a statement is an essential component of Boolean algebra and can be written as the classic exclusive OR: $A = (M \vee \sim M)$. A is equal to M or not M. The symbol $\vee$ is OR and the symbol ~ is the term "not."

1.3 Building on the Past – From Those Who Laid the Foundation

As the centuries have rolled forward and technology has advanced, with *thinking*, *reasoning*, and *learning* yet at its roots, the works and ideas of the early philosophers laid the foundation for contemporary concepts and ideas. Yet, as we move forward, if we pause and reflect for a moment, we realize that learning is not limited to that proffered by the pendants of previous days but is a natural part of our daily existence. Each time that we encounter a fresh idea, make a new discovery, or solve one of life's many problems, we are revisiting our store of knowledge, we are reasoning, and, hopefully, we are learning and collecting new knowledge to add to that store. We are growing and enriching our perceived model of the world.

As we have seen, educators, scientists, and philosophers have debated, studied, and analyzed learning for centuries. Our deep understanding of this process remains embryonic, yet our learning continues. Following in the footsteps of early thinkers are Feigenbaum, Hegle, Newell, Minsky, Papert, Winston, McCulloch, Pitts, Rosenblatt, Hebb, Lukasiewicz, Hopfield, Knuth, and Zadeh. We will examine the roles that each has played and how they have contributed. The next few paragraphs present a small portion of their works.

1.4 A Learning and Reasoning Taxonomy

One can distinguish a number of different forms or methods of learning ranging from the most elementary rote learning to more complex processes of learning by analogy and discovery. Variations on this taxonomy have appeared in much of the recent literature. In the following paragraphs, the term “student” is frequently used. Consider that a student or learner could be a human person or, potentially, a machine.

In some of his earlier works, Edward Feigenbaum proposed a five-phase learning process:

1) Request information.
2) Interpret the information.
3) Convert the information into a useable form.
4) Integrate the information into the existing knowledge store.
5) Apply the knowledge and evaluate the results.

The learning situation is composed of two parties, the *learner* and the *teacher* or *environment*, and a body of knowledge to be transferred from the environment to the learner. Based on the five criteria, a six-level learning taxonomy was proposed. The taxonomy considers two extremes: no active learner participation and complete active learner participation. Examining the taxonomy, one can easily see the influence of Socrates.

1) *Rote learning* – A memorization process that requires little thought of meaning by the learner.
2) *Learning with a teacher* – Most of the information is provided by the teacher. Missing details must be inferred by the learner.
3) *Learning by example* – Specific conceptual instances are given; however, generalization must be achieved by the learner.
4) *Learning by analogy* or *metaphor* – Related conceptual instances are given. The learner must recognize the relation and apply it to the task at hand.
5) *Learning by problem solving* – Knowledge embedded in the problem may be gained by the learner through solving the problem.
6) *Learning by discovery* – Knowledge exists but must be hypothesized by the learner through theory formation and extracted by experiment.

1.4.1 Rote Learning

Habit or *rote learning* is the simplest learning process. The environment supplies all of the knowledge, and the learner merely accepts and stores it with no thought to meaning or content. Despite its elementary nature, the rote acquisition of knowledge is essential to all higher forms of learning. The learner must retain base information to be able to apply it to future problems.

B.F. Skinner took a somewhat different view. He suggested that no clear connection had been demonstrated in education between ends and means. He contended that the educational process should be reduced to defining goals or acts that the learner was able to perform. Based on the "present" state of knowledge, a sequence of acts could be created to move the learner from the present to the desired state. Often, the teacher would not be necessary.

One of the most familiar and perhaps best early instances of mechanized rote learning is Arthur Samuel's program designed to play the game of checkers. The program was initially equipped with a number of suggested procedures for playing the game correctly. The intent was to have the program learn by memorizing successful (deemed significant) board positions as it encountered them and then to use them properly and effectively in future games. Ultimately, the program progressed to the level of skilled novice.

1.4.2 Learning with a Teacher

Learning with a teacher is the first level of increased complexity in the learning hierarchy. Here, the learner is beginning to take an active role in several phases of the process, specifically she or he may request information from the teacher. In this situation, abstracted or general information is presented in an integrated manner by the teacher. The learner must accept the information and then complete the store of knowledge by inferring the missing details.

Many successful programs have been written using such a paradigm. In these, the program played the role of the student or learner. Several programs, including Mostow's FOO program for playing the card game "hearts" and Waterman's poker player, were oriented toward game playing. Davis's TEIRESIAS program presented an interesting variation on this scheme.

Rather than being autonomous, TEIRESIAS was designed to sit in front of the MYCIN program written earlier by E.H. Shortliffe. MYCIN was a large rule-based system designed to assist physicians in the diagnosis and therapy of infectious diseases. The design and development process for any such large-scale system is both iterative and refining. If the system makes a misdiagnosis or offers advice contrary to the physician's diagnosis, the knowledge base must be modified. Under such a condition, TEIRESIAS would interact with the user to correct the difficulty. Such a situation reduces to a two-part task: first, explaining to the user the line of reasoning that led to the conclusion and then second, asking what additional or different information is needed to alter the result.

1.4.3 Learning by Example

Learning by example or induction increases the level of participation by the learner in the learning process. Unlike the previous example in which the teacher abstracted and then presented the material, here the student must assume the responsibility for the task. In such a context, specific conceptual instances are presented, and the student must recognize the significant or key features of the examples and then form the desired generalizations.

An early classic example of such an approach is Patrick Winston's work on "*Learning Structural Descriptions from Examples*." The goal of Winston's program was to learn elementary geometrical constructs such as those one might build using toy blocks. The program was presented with training instances from which it evolved an internal description of the concept it was to be learning. The knowledge acquired was incorporated into a semantic network where all of the interrelationships among the constituent elements were described.

Critical to the effective use of Winston's algorithm are the ideas that positive training instances are evolutionary rather than revolutionary. In Winston's algorithm, negative training instances are those that reflect only minimal differences from the concept being investigated; thus, no learning occurs.

1.4.4 Analogical or Metaphorical Learning

Analogical or metaphorical learning is probably one of the more common methods by which human beings acquire new knowledge. As Winston points out in his work *Learning and Reasoning by Analogy*, with such an approach, once again, the learner's contribution to the process is increased. When *learning by example*, the learner is presented with positive and negative instances of the concept to be learned. With an *analogy*, the student has only closely related instances from which to extract the desired concept.

Jaime Carbonell identifies *transformational* and *derivational* as the two principal methods of reasoning by analogy. When learning by transformational analogy, the line of reasoning proceeds incrementally from some old or known solution to the new or desired solution through a series of mappings *means-ends* called *transform operators*. The operators are applied using a *means-ends paradigm* until the desired transformation is achieved.

Knowledge acquisition by *derivational analogy* achieves learning by recreating the line of reasoning that resulted in the solution to the problem. The reconstruction includes both decision sequences and attendant justifications.

1.4.5 Learning by Problem Solving

Learning by problem solving can easily be viewed as subsuming all other forms of learning discussed. However, such a technique has sufficient merit in its own right that it deserves individual consideration. With this approach, the knowledge to be imparted is embedded in a problem or sequence of problems. The objective is for the learner to acquire that knowledge by solving the problem. The most serious difficulty with such an approach is the intolerance of individual method.

Consider the question: "What is the sum of ¼ and ¼?" Although a response of 2/4 would be completely correct, that answer may be considered wrong since it did not match the "correct" answer of ½.

1.4.6 Learning by Discovery

Learning by discovery is the antithesis of rote learning. In this paradigm, the learner is the initiator in all five phases of learning discussed earlier. There is no new knowledge in the world since all knowledge already exists and is merely waiting some clever individual to discover it.

According to Carbonell, two basic methods of acquiring knowledge by discovery are available: *observation* and *experimentation*. Observation is considered to be a passive approach because the learner collects information by watching a particular event and then later forms a theory to explain the phenomenon. In contrast, experimentation is viewed as active. Here, the process generally involves the learner postulating a new theory about the existence of a particular piece of knowledge and verifying that theory by experiment. In neither case is the possibility of serendipitous discovery excluded.

Discovery, Carbonell proposes, is a three-step process begun by hypothesis formation. The hypothesis may be either data driven, as is the case for observation, or theory driven as with experimentation. The initial step is followed by a refinement process in which partial theories are merged and boundary conditions are established. Finally, what has been learned and created is extended to new instances.

Clearly, any theory or model of human (or machine) learning must include aspects of each member of the established taxonomy. Today, no comprehensive and unified theory of human learning exists. Only partial theories that attempt to explain portions of the whole of human learning have been developed. Looking back over each of the ideas, we can take away the notion that learning and problem solving are effectively used interchangeably.

1.5 Crisp and Fuzzy Logic

As we move forward, we will explore, study, and learn two forms of logic and reasoning called *crisp logic* or reasoning and *fuzzy logic* or reasoning. The two graphical diagrams shown in Figure 1.1 suggest the difference between these two forms.

On the left is a crisp, precise circle and on the right is one in which the shape is less precise but can still be viewed in many contexts as a kind of circle.

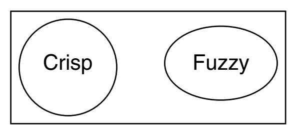

Figure 1.1 Crisp and fuzzy circles.

1.6 Starting to Think Fuzzy

Over the years, fuzzy logic has been found to be extremely beneficial and useful to people involved in research and development in numerous fields including engineers, computer software developers, mathematicians, medical researchers, and natural scientists. As we begin, with all those people involved, we raise the question: What is fuzzy?

Originally, the word fuzz described the soft feathers that cover baby chicks. In English, the word means indistinct, imprecise, blurred, not focused, or not sharp. In French, the word is flou and in Japanese, it is pronounced "aimai." In academic or technical worlds, the word fuzz or fuzzy is used in an attempt to describe the sense of ambiguity, imprecision, or vagueness often associated with human concepts.

Revisiting an earlier example, trying to teach someone to drive a car is a typical example of real-life fuzzy teaching and fuzzy learning. As the student approaches a red light or intersection, what do you tell him or her? Do you say, "Begin to brake 25 m or 75 ft from the intersection?" Probably not. More likely, we would say something more like "Apply the brakes soon" or "Start to slow down in a little bit." The first case is clearly too precise to be implemented or executed by the driver. How can one determine exactly when one is 25 m or 75 ft from an intersection? Streets and roads generally do not have clearly visible and accurate millimeter- or inch-embedded gradations. The second vague instruction is the kind of expression that is common in everyday language.

Children learn to understand and to manipulate fuzzy instructions at an early age. They quite easily understand phrases such as "Go to bed about 10:00." Perhaps with children, they understand too well. They are adept at turning such a fuzzy expression into one that is very precise. At 9:56, determined to stay up longer, they declare, "It's not 10:00 yet."

In daily life, we find that there are two kinds of imprecision: statistical and nonstatistical. Statistical imprecision is that which arises from such events as the outcome of a coin toss or card game. Nonstatistical imprecision, on the other hand, is that which we find in instructions such as "Begin to apply the brakes soon." This latter type of imprecision is called fuzzy, and qualifiers such as *very*, *quickly*, *slowly* or others on such expressions are called *hedges* in the fuzzy world.

Another important concept to grasp is the linguistic form of variables. *Linguistic variables* are variables with more qualitative rather than numerical values, comprising words or phrases in a natural or potentially an artificial language. That is, whether simple or complex, such variables are *linguistic* rather than *numeric*. Simple examples of such variables are *very*, *slightly*, *quickly*, and *slowly*. Other examples may be generated from a set of primary terms such as *young* or its antonym *old* or *tall* with its antonym *short*.

The first practical noteworthy applications of fuzzy logic and fuzzy set theory began to appear in the 1970s and 1980s. To effectively design modern everyday systems, one must be able to recognize, represent, interpret, and manipulate statistical and nonstatistical uncertainties. One should also learn to work with *hedges* and *linguistic variables*. One should use statistical models to capture and quantify random imprecision and fuzzy models to capture and to quantify nonrandom imprecision.

1.7 History Revisited – Early Mathematics

Fuzzy logic, with roots in early Greek philosophy, finds a wide variety of contemporary applications ranging from the manufacture of cement to the control of high-speed trains, auto focus cameras, and potentially self-driving automobiles. Yet, early mathematics began by emphasizing precision. The central theme in the philosophy of Aristotle and many others was the search for perfect numbers or golden ratios. Pythagoras and his followers kept the discovery of irrational numbers a secret. Their mere existence was also counter to many fundamental religious teachings of the time.

Later mathematicians continued the search for precision and were driven toward the goal of developing a concise theory of mathematics. One such effort was *The Laws of Thought* published by Stephan Korner in 1967 in the *Encyclopedia of Philosophy*. Korner's work included a contemporary version of *The Law of Excluded Middle* which stated that every proposition could only be TRUE or FALSE – there could be no in between. An earlier version of this law, proposed by Parmenides in approximately 400 BC, met with immediate and strong objections. Heraclitus, a fellow philosopher, countered that propositions could simultaneously be both TRUE and NOT TRUE. Plato, the student, made the same arguments to his teacher Socrates.

1.7.1 Foundations of Fuzzy Logic

Plato was among the first to attempt to quantify an alternative possible state of existence. He proposed the existence of a third region, beyond TRUE and FALSE, in which "opposites tumbled about." Many modern philosophers such as Bertrand Russell, Kurt Gödel, G.W. Leibniz, and Hermann Lotze have supported Plato's early ideas.

The first formal steps away from classical logic were taken by the Polish mathematician Lukasiewicz (also the inventor of Reverse Polish Notation, RPN). He proposed a three-valued logic in which the third value, called *possible*, was to be assigned a numeric value somewhere between TRUE and FALSE. Lukasiewicz also developed an entire set of notations and an axiomatic system for his logic. His intention was to derive modern mathematics.

In later works, he also explored four- and five-valued logics before declaring that there was nothing to prevent the development of infinite-valued logics. Donald Knuth proposed a similar three-valued logic and suggested using the values of −1, 0, 1. The idea never received much support.

1.7.2 Fuzzy Logic and Approximate Reasoning

The birth of modern fuzzy logic is usually traced to the seminal paper *Fuzzy Sets* published in 1965 by Lotfi A. Zadeh. In his paper, Zadeh described the mathematics of fuzzy subsets and, by extension, the mathematics of fuzzy logic. The concept of the fuzzy event was introduced by Zadeh (1968) and has been used in various ways since early attempts to model inexact concepts were prevalent in human reasoning. The initial work led to the development of the branch of mathematics called *fuzzy logic*. This logic, actually a superset of classical binary-valued logic, does not restrict set membership to absolutes (Yes or No) but tolerates varying degrees of membership.

Using these criteria, an element is assigned a grade of membership in a parent set. The domain of this attribute is the closed interval [0, 1]. If the grade of membership values is restricted to the two extrema, then fuzzy logic reduces to two-valued or crisp logic.

In his work, Zadeh proposed that people often base their thinking and decisions on imprecise or nonnumerical information. He further believed that the membership of an element in a set need not be restricted to the values 0 and 1 (corresponding to FALSE and TRUE) but could easily be extended to include all real numbers in the interval 0.0–1.0 including the endpoints. He further felt that such a concept should not be considered in isolation but rather viewed as a methodology that moves from a discrete world to a continuous one. To augment such thinking, he proposed a collection of operations supporting his new logic.

Zadeh introduced his ideas as a new way of representing the vagueness common in everyday thinking and language. His fuzzy sets are a natural generalization or superset of classical sets or Boolean logic that are one of the basic structures underlying contemporary mathematics. Under Boolean algebra, a proposition takes a narrow view that a value is either completely true or completely false. In contrast, fuzzy logic introduces the concept of partial truth under which values are expressed anywhere within, and including, the two extremes of TRUE and FALSE.

Based on the idea of the fuzzy variable, Zadeh (1979) further proposed a theory of *approximate reasoning*. This theory postulates the notion of a possibility distribution on a linguistic variable. Using this concept, he was able to reason using vague concepts such as *young, old, tall*, or *short*. Zadeh also introduced the ideas of *semantic equivalence* and *semantic entailment* on the possibility distributions of linguistic propositions. Using these concepts, he was able to determine that a *statement* and its *double negative* are equivalent and that *very small* is more restrictive than *small*. Such conclusions derive from either the *equality* or *containment* of corresponding distributions.

Zadeh's theory is generally effective in reconciling ambiguous natural language expressions. The scope of the work was initially limited to laboratory sentences comparing hair color, age, or height between various people. Zadeh's work provided a good tool for future efforts, particularly in combination with or to enhance other forms of reasoning.

As often follows the introduction of a new concept or idea, questions arise: Why does that thing do this? Why doesn't it do that? Can you make it do another thing? An early criticism of Zadeh's fuzzy sets was: "Why can't your fuzzy set members have an uncertainty associated with them?" Zadeh eventually dealt with the issue by proposing more sophisticated kinds of fuzzy sets. New criteria evolved the original concept into numbered types of fuzzy subsets. His initial work became type-1 fuzzy sets. Additional concepts grew from type-2 fuzzy sets to ultimately type-*n* in a 1976 paper to incorporate greater uncertainty into set membership. Naturally, if a type-2 or higher set has no uncertainty in its members, it reduces to a type-1. In this text, we will work primarily with type-1 fuzzy subsets.

1.7.3 Non-monotonic Reasoning

Non-monotonic reasoning is an attempt to duplicate the human ability to reason with incomplete knowledge and to make default assumptions when insufficient evidence exists to empirically support a hypothesis. This proposed method of reasoning may be contrasted with *monotonic* reasoning in the following way.

A *monotonic logic* states that if a conclusion can be derived from a set of premises X, and if X is a subset of some larger set of premises Y, then x, a member of X, may also be derived from Y. This does not hold true for non-monotonic logic since Y may contain statements that may prevent the earlier conclusion from being derived.

Consider the following example scenario: The objective is to cross a river, and at the edge of the river there is a row boat and a set of oars. Using monotonic logic, one can conclude that it is possible to cross the river by rowing the boat across. If the new information that the boat is painted red is added, this will not alter the conclusion. On the other hand, using non-monotonic logic, the same initial conclusion may be drawn; however, if the new information that the boat has a hole in it arises, the original conclusion can no longer be drawn.

An English philosopher, William of Ockham (or Occam) (1280–1348) held a number of beliefs that foreshadowed the development of non-monotonic logic. In particular, there is an element of his philosophy called *Occam's Razor* that provides a succinct description of this form of logic. Occam stated ". . .that for the purposes of explaining, things not known to exist should not, unless it is absolutely necessary, be postulated as existing" (1280–1348); this has also been called the *Law of Parsimony*.

This belief may be reformulated slightly as "in the absence of any information to the contrary, assume. . .." This kind of reasoning may be defined as a *plausible inference* and is applied when conclusions must be drawn despite the absence of total knowledge about a world. These consequences then become a belief that might be modified with subsequent evidence. In a closed world, what is not known to be true must be false. Therefore, one can infer negation if proving the affirmative is not possible. Inferring negation becomes more difficult, of course, in an open world.

A first-order theory implies a monotonic logic; however, a real-world situation is non-monotonic because of gaps or incompleteness in the knowledge base. The default inference can then be used to fill in these gaps, which is very similar to some of Piaget's arguments.

McCarthy (1980) presents an idea that he calls *circumscription*. Circumscription is a rule of conjecture that argues when deriving a conclusion, that the only relevant entities are the facts on hand, and those whose existence follows from these facts. The correctness of the conclusion depends upon all of the relevant facts having been taken into account. Rephrased, if A is a collection of facts, conclusions derived from circumscription are conjectures that A includes all the relevant facts and that the objects whose existence follows from A are all relevant objects.

Reiter (1980), on the other hand, argues for default inferences from a closed-world perspective. Under such an assumption, he asserts that if R is some relation, then one can assume not R (the opposite of R or R does not exist) if assuming not R is consistent to do so. This consistence is based on not being able to prove R from the information on hand. If such a proof cannot be done, then the proof must not be true, or, similarly, if an object cannot be proven to exist in the current world, then the object does not exist.

Looking at the relationship between fuzzy logic and Reiter's form of non-monotonic logic, Reiter asserts that a default inference provides a representation for (almost all) the fuzzy subsets (and with most in terms of defaults). Reiter's assertion is not strictly correct because the inference is either true or not true, whereas a fuzzy grade of membership expresses a degree of belief in the entity.

A fundamental difference between these two theories is that Reiter's theory appears to require a global domain, whereas McCarthy's theory does not. McDermott and Doyle (1980) argue that this may not be a weakness in Reiter's approach. In either case, the intention is to extend a given set of facts (beliefs) by inferring new beliefs from the existing ones. These new beliefs are held until the evidence is introduced to contradict them. When such counterevidence occurs, a reorganization of the belief system is required.

In the discussion of his TMS (Truth Maintenance System) system, Doyle (1979) suggests that such a reorganization may take either of the two forms: *world model reorganization* or *routine revision*. Routine revision requires maintaining a body of facts that are expressed as universally true but may have some exceptions. Such a need usually occurs as a result of inferences, default assumptions, or observations. World model reorganization involves more wholesale restructuring of the model when something goes wrong. The aforementioned world model reorganization is usually quite complex and is typically the result of induction hypothesis, testimony, analogy, and intuition.

Note that these two (monotonic and non-monotonic reasoning) are very similar to Piaget's concepts of assimilation and accommodation. From Doyle's point of view, non-monotonic logic is reasoning with revision and that if a default election is made from a number of possible alternatives based on the alternatives not being believed, then the concept or argument under debate or consideration is not extensible.

1.8 Sets and Logic

1.8.1 Classical Sets

Classical sets are considered crisp because their members satisfy precise properties. For example, for illustration, let H be the set of integer real numbers from 6 to 8. Using set notation, one can express H as:

$$H = \{r \varepsilon R \mid 6 \le r \le 8\}$$
$$H = \{6, 7, 8\} \tag{1.1}$$

One can also define a function $\mu_H(r)$ called a *membership function* to specify the membership of r in the set H,

$$\begin{aligned} \mu_H(r) &= 1 \quad 6 \le r \le 8 \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{1.2}$$

The expression states:

> r is a member of the set H (membership in $H = 1$) if its value is 6, 7, or 8. Otherwise, it is not a member of the set (membership in $H = 0$).

One can also present the same information graphically as in Figure 1.2.

Whichever representation is chosen, it remains clear that every real number, r, is surrounded by crisp boundaries and is either in the set H or not in the set H.

Moreover, because the membership function μ maps the associated *universe of discourse* of every classical set onto the set {0, 1}, it should be evident that crisp sets correspond to a two-valued logic. An element is either in the set or it is not in the set, and it is either TRUE or it is FALSE.

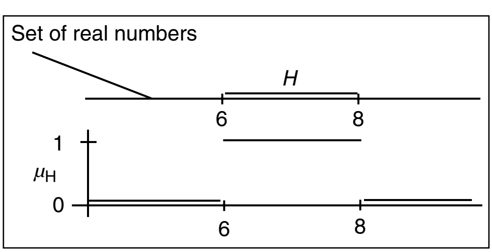

Figure 1.2 Membership in subset *H*.

1.8.2 Fuzzy Subsets

In relation to crisp sets, as we noted, fuzzy sets are supersets (of crisp sets) whose members are composed of collections of objects that satisfy imprecise properties to varying degrees. As an example, we can write the statement that X is a real number close to 7 as:

F = set of real numbers close to 7

But what do "set" and "close to" mean and how do we represent such a statement in mathematically correct terms?

Zadeh suggests that F is a fuzzy subset of the set of real numbers and proposes that it can be represented by its membership function, mF. The value of mF is the extent or grade of membership of each real number r in the subset of numbers close to 7. With such a construct, it is evident that fuzzy subsets correspond to a continuously valued logic and that any element can have various degrees of membership in the subset.

Let's look at another example. Consider that a car might be traveling on a freeway at a velocity between 20 and 90 mph. In the fuzzy world, we identify or define such a range as the *universe of discourse*. Within that range, we might also say that the range of 50–60 is the *average* velocity.

In the fuzzy context, the term *average* would be classed as a *linguistic* variable. A velocity below 50 or above 60 would not be considered a member of the *average* range. However, values within and equal to the two extrema would be considered members.

1.8.3 Fuzzy Membership Functions

The following paragraphs are partially reused in Chapter 4.

The fuzzy property "close to 7" can be represented in several different ways. Who decides what that representation should be? That task falls upon the person doing the design.

To formulate a membership function for the fuzzy concept "close to 7," one might hypothesize several desirable properties. These might include the following properties:

- *Normality* – It is desirable that the value of the membership function (grade of membership for 7 in the set F) for 7 be equal to 1, that is, $\mu_F(7) = 1$.

  We are working with membership values 0 or 1.

- *Monotonicity* – The membership function should be monotonic. The closer r is to 7, the closer $\mu_F(r)$ should be to 1.0 and vice versa.

  We are working with membership values in the range 0.0–1.0.

- *Symmetry* – The membership function should be one such that numbers equally distant to the left and right of 7 have equal membership values.

  We are working with membership values in the range 0.0–1.0.

It is important that one realize that these criteria are relevant only to the fuzzy property "close to 7" and that other such concepts will have appropriate criteria for designing their membership functions.

Based on the criteria given, graphic expressions of the several possible membership functions may be designed. Three possible alternatives are given in Figure 1.3. Depending upon whether one is working in a crisp or fuzzy domain, the range of grade of membership (vertical) axis should be labeled either {0–1} or {0.0–1.0}.

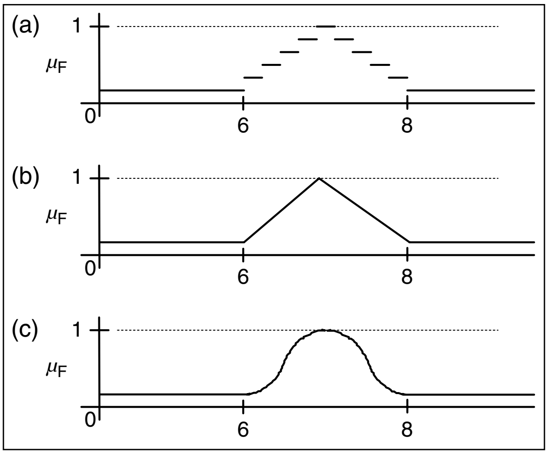

Figure 1.3 Membership functions for "close to 7."

Note that in graph c, every real number has some degree of membership in F although the numbers far from 7 have a much smaller degree. At this point, one might ask if representing the property "close to 7" in such a way makes sense.

Example 1.1
Consider a shopping trip with a friend in Paris who poses the question:

How much does that cost?

To which you answer:

About 7 euros.

which certainly can be represented graphically.

Example 1.2
As a further illustration, let the crisp set H and the fuzzy subset F represent the heights of players on a basketball team. If, for an arbitrary player p, we know that the membership in the set H is given as $\mu_H(p) = 1$, then all we know is that the player's height is somewhere between 6 and 8 ft. On the other hand, if we know that the membership in the set P is given as $\mu_F(p) = 0.85$, we know that the player's height is close to 7 ft. Which information is more useful?

Example 1.3
Consider the phrase:

Etienne is old.

The phrase could also be expressed as:

Etienne is a member of the set of old people.

If Etienne is 75, one could assign a fuzzy truth value of 0.8 to the statement. As a fuzzy set, this would be expressed as:

$$\mu_{\text{old}}\,(\text{Etienne}) = 0.8$$

From what we have seen so far, membership in a fuzzy subset appears to be very much like probability. Both the degree of membership in a fuzzy subset and a probability value have the same numeric range: 0–1. Both have similar values: 0.0 indicating (complete) nonmembership for a fuzzy subset and FALSE for a probability and 1.0 indicating (complete) membership in a fuzzy subset and TRUE for a probability. What, then, is the distinction?

Consider a natural language interpretation of the results of the previous example. If a probabilistic interpretation is taken, the value 0.8 suggests that there is an 80% chance that Etienne is old. Such an interpretation supposes that Etienne is or is not old and that we have an 80% chance of knowing it. On the other hand, if a fuzzy interpretation is taken, the value of 0.8 suggests that Etienne is more or less old (or some other term corresponding to 0.8).

To further emphasize the difference between probability and fuzzy logic, let's look at the following example.

Example 1.4
Let

L = the set of all liquids
P = the set of all potable liquids

Suppose that you have been in the desert for a week with nothing to drink, and you find two bottles (Figure 1.4).

Which bottle do you choose?

Bottle A could contain wash water. Bottle A could not contain sulfuric acid.

> A *membership value* of 0.9 means that the contents of A are very similar to perfectly potable liquids, namely, water.
> A *probability* of 0.9 means that over many experiments that 90% yield B to be perfectly safe and that 10% yield B to be potentially deadly. There is 1 chance in 10 that B contains a deadly liquid.

Observation:

What does the information on the back of the bottles reveal?

> After examination, the membership value for A remains unchanged, whereas the probability value for B containing a potable liquid drops from 0.9 to 0.0 (Figure 1.5).

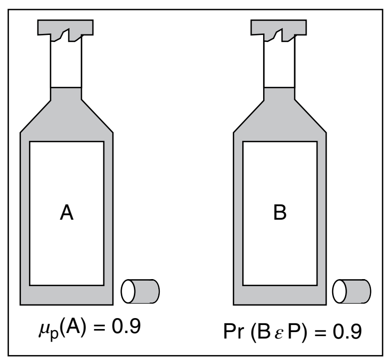

Figure 1.4 Bottles of liquids – front side.

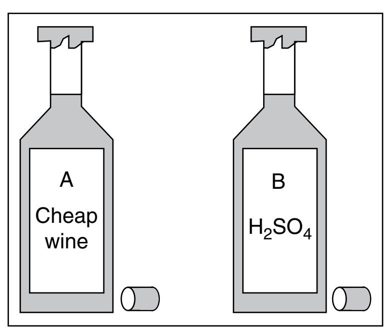

Figure 1.5 Bottles of liquids – back side.

As the previous example illustrates, fuzzy memberships represent similarities to objects that are imprecisely defined, while probabilities convey information about relative frequencies.

1.9 Expert Systems

Expert systems are an outgrowth of postproduction systems augmented with various elements of probability theory and fuzzy logic to emulate human reasoning within constrained domains. The general structure of such systems is a decision-making portion known as an *inference engine* associated with a hierarchical knowledge structure or *knowledge base*.

The knowledge base typically consists of the domain-specific knowledge and at least one level of abstraction. This second level contains the knowledge about knowledge or *meta-knowledge* for the domain. As such, this level of abstraction provides the inference engine with the criteria for making decisions or reasoning within the specific domain of application.

In some cases, these systems have performed with remarkable success. Most notables are *Dendral,* Feigenbaum (1969), *Meta-Dendral*, Buchanan (1978), *Rule-Based Expert Systems: MYCIN Experiments,* Shortliffe (1984), and *Prospector,* Duda (1978). In each case, there have been tens and perhaps hundreds of man-years devoted to tailoring each to a specific task. Nonetheless, those could be integrated, with very little effort, as a brute-force solution to reasoning. These systems have little ability to learn from previous experience and are extremely fragile at the boundaries of their knowledge.

It is clear that human understanding of human knowledge is in its infancy. Sundry schemes have been suggested and tried while attempting to explain and to emulate or simulate the human thought and reasoning processes we so casually take for granted. None has demonstrated, nor even proposed, a universal answer; they merely chip at small corners and suggest that the elephant is a thin line like a rope, round like a cylinder, or flat like a wall.

1.10 Summary

We began this chapter with a look at some of the early works in learning and reasoning. We have seen that vagueness and imprecision are common in everyday life. Very often, the kind of information we encounter may be placed into two major categories: statistical and nonstatistical. The former we model using probabilistic methods, and the latter we model using fuzzy methods.

We introduced and examined the concepts of monotonic, non-monotonic, and approximate reasoning and several models of human reasoning and learning.

We examined the concepts of crisp and fuzzy subsets and crisp and fuzzy subset membership. We learned that the possible degree of membership $\mu_x F()$ (membership of the variable x in the set F) of a variable x in a fuzzy subset spans the range [0.0–1.0] and that when we restrict the membership function so as to admit only two values, 0 and 1, fuzzy subsets reduce to classical sets. We also introduced the membership graph as a tool for expressing membership functions.

Review Questions

1.1 In the chapter opening, we introduced the term learning. Describe what it means.

1.2 Identify the three laws of thought and briefly describe each of them.

1.3 The chapter identified what is termed Feigenbaum's five-phase learning process. Identify and describe each of those phases.

1.4 Based on Feigenbaum's learning process, identify and describe each member of the corresponding six-level taxonomy.

1.5 In the context of this chapter, discuss what the term fuzzy means.

1.6 Identify and describe the difference between a classical set and a fuzzy set.

1.7 What is a fuzzy membership function? What information does such a function give you?

1.8 Formulate and graph a membership function for the concept very tall.

1.9 Formulate and graph membership functions for Examples 1.2 and 1.3.

1.10 What kind of information are we expressing with a crisp membership graph? What kind of information does such a graph give you?

1.11 What kind of information are we expressing with a fuzzy membership graph? What kind of information does such a graph give you?

1.12 Why do you think that membership graphs have different shapes?

1.13 Identify and describe how crisp and fuzzy membership graphs differ.

1.14 The chapter identified two kinds of imprecision. What are they and describe each and the differences.

1.15 To effectively design modern real-world systems, the chapters identified several important things. Identify and describe them.

1.16 Draw and annotate a membership graph for members of the crisp set old people.

1.17 Draw and annotate a membership graph for members of the fuzzy set old people.