Early Developments in Fuzzy Logic
Selected Issues

iciu

Prof. George J. Klir

Literature

- The Enigma of Cox's Proof
- G. J. Klir, Is There More to Uncertainty Than Some Probability Theorists Might Have us Believe?
- M. Colyvan, The Philosophical Significance of Cox's Theorem
- R. T. Cox, Probability, Frequency and Reasonable Expectation
- Some Criticism of Fuzzy Logic
- The Logic That Dares Not Speak Its Name
- Ch. Elkan, The Paradoxical Success of Fuzzy Logic
- R. Belohlavek, G. J. Klir, On Elkan's Theorems: Clarifying Their Meanning via Simple Proofs

All materials are available at
`www.inf.upol.cz/downloads/early-developments-in-fuzzy-logic.zip`

esf european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
INVESTMENTS IN EDUCATION DEVELOPMENT

Appendix A. The Enigma of Cox's Proof

The claim that the only way to deal with uncertainty is to use the rules of probability theory in some of the debates described in Section 2.4 is often justified by referring to a proof by Cox (1946)[1]. According to Cox, the aim of the proof is "to show that by employing the algebra of symbolic logic it is possible to derive the rules of probability from two quite primitive notions". Although his derivation is actually not presented as a proof of a theorem, but rather as a sequence of intertwined formal and intuitive arguments, it has been routinely referred to as "Cox's proof" or "Cox's theorem". In order to describe this derivation, we consider it appropriate to adhere to the original Cox's notation: Letters **a**, **b**, **c** ... denote propositions; ~ **a** denotes the negation of proposition **a**; **ab** and $\mathbf{a} \vee \mathbf{b}$ denote, respectively, the conjunction and disjunction of propositions **a** and **b**; and $\mathbf{b} \mid \mathbf{a}$ denotes "some measure of reasonable credibility of the proposition **b** when the proposition **a** is known to be true".

Using this notation and employing the Boolean algebra of classical propositional logic (or classical set theory), Cox aims at proving that the only sensible way to combine reasonable credibilities is to use the rules of probability theory. He begins with the assumption (axiom) that

$$\mathbf{cb} \mid \mathbf{a} = F(\mathbf{c} \mid \mathbf{ba}, \mathbf{b} \mid \mathbf{a}), \tag{1}$$

where F is some function of two variables to be determined. Employing the associative law of conjunctions of propositions, he derives the equation

$$F(F(\mathbf{d} \mid \mathbf{cba}, \mathbf{c} \mid \mathbf{ba}), \mathbf{b} \mid \mathbf{a}) = F(\mathbf{d} \mid \mathbf{cba}, F(\mathbf{c} \mid \mathbf{ba}, \mathbf{b} \mid \mathbf{a})), \tag{2}$$

where **a**, **b**, **c**, **d** are any propositions. Letting $\mathbf{d}|\mathbf{cba} = x$, $\mathbf{c}|\mathbf{ba} = y$, and $\mathbf{b}|\mathbf{a} = z$, Eq. (2) becomes

$$F(F(x,y),z) = F(x,F(y,z)). \tag{3}$$

Function F must satisfy this functional equation, known as associativity equation (Aczel 1966, 253), for arbitrary values x, y, and z.

Once Cox converted his original problem into this purely mathematical problem, his challenge was to solve this equation. In a long and tedious derivation, described in detail in a large Appendix to his paper, Cox managed to show the following: If F has continuous second-order derivatives, then

$$Cf(F(p,q)) = f(p)f(q) \tag{4}$$

[1] Also covered later without any substantial change in his book (Cox 1961).

is the solution of (3), where, as stated by Cox, "*f* is an arbitrary function of a single variable and *C* is an arbitrary constant". This derivation is correct and it was a significant contribution to the theory of functional equations when Cox's paper was published.

In order to determine a relation between **b**|**a** and ~ **b**|**a**, Cox furthermore assumes (employs as an axiom) that

$$\sim \mathbf{b} \mid \mathbf{a} = S(\mathbf{b} \mid \mathbf{a}), \tag{5}$$

where *S* is some function to be determined. Employing the law of double negation and one of the De Morgan's laws of classical logic, he derives the functional equation

$$xS(S(y)/x) = yS(S(x)/y), \tag{6}$$

where x = **c**|**a** and $y = S($**cd**|**a**$)$. He then shows (again by a tedious but correct derivation in the Appendix of his paper) the following: If *S* is twice differentiable, then

$$S(p) = (1 - p^m)^{1/m} \tag{7}$$

is the solution of (6), where *m* is an arbitrary constant. Cox considers the value of *m* purely conventional and chooses $m = 1$ to obtain the desired formula of probability theory.

The results obtained by Cox became a subject of controversy in some of the debates described in Section 2.4. The controversy was triggered by two closely related claims, a rather extreme claim that "the only satisfactory description of uncertainty is probability" (Lindley[1]) and an associated claim that "the strongest argument for the use of standard probability theory is a proof by Cox" (Cheeseman[2]). Since the Cox's proof was published more than 40 years before the debates and in a journal specializing on physics, it is understandable that many people participating in the debates were initially not aware of it. However, some of them[3] quickly recognized that the proof was contingent upon the assumption that function *F* in the associativity functional equation (3) has a continuous second derivative. This, they argued, excludes possibility and necessity measures since they are based on max and min functions, which are clearly not satisfy this assumption. Next, some advocates of the original meaning of Cox's proof responded by defending it via reference to another method for solving equation (3), which was developed by Aczél (1966). For example, Smith and Erickson (1989) wrote (p. 38):

[1] *Statistical Science*, 2(1), 1987, p.17.
[2] *Computational Intelligence* (Canadian), 4(1), 1988,
[3] For example, (Dubois and Prade 1988).

> By assuming that F is twice differentiable in both variables, Cox derived from Eq. (3)[1] a differential equation which he then solved. Some fuzzy set advocates have pounced upon this assumption as invalidating Cox's theory, in evident ignorance of the work of Aczél (1966, 1987), who derived the same general solution without assuming differentiability.

This was more than a decade later still echoed by Jaynes (2003, 668):

> The issue of nondifferentiable functions arises from time to time in probability theory. In particular, when one solves a functional equation such as those studied in Chapter 2,[2] to assume differentiability is to have a horde of compulsive mathematical nitpickers descend upon one, with claim that we are excluding a large class of potentially important solutions. However, we noted that this is not the case; Aczél demonstrated that Cox's functional equations can all be solved without assuming differentiability (at the cost of much longer derivation) and with just the same solution that we found above.

It is correct that Aczél was able to solve (3) without the assumption that function F has a continuous second derivative, but assuming instead that F is reducible on both sides, which he defines (Aczél 1966, 255) as "$F(t, u) = F(t, v)$ or $F(u, w) = F(v, w)$ only if $u = v$." The authors of the above statements seem to tacitly assume that the requirement that F be reducible on both sides is weaker than the requirement that F has a continuous second derivative and, hence, the controversy regarding functions max and min is resolved. However such an assumption is wrong. Functions max and min do not have continuous second derivatives, but they are non-reducible on both sides as well. The two requirements are actually not comparable. In addition to the product function, which clearly satisfies both of them, and the max and min functions, which satisfy neither of them, there also exist associative functions that satisfy only one of them. For example, function $F(x,y) = \sqrt{x^2 + y^2}$ is associative and twice differentiable, but nor reducible. On the other hand, function $F(x,y) = f^{-1}(f(x), f(y))$, where $f(x) = x/2$ for $x \in [0,0.5]$ and $f(x) = 1.5x - 0.5$ for $x \in (0.5,0]$ (and $f(y)$ is defined in the same way), is clearly associative and not differentiable, but it is reducible from both sides. To show its reducibility, let $F(x,y) = F(x,z)$. Then,

$$f^{-1}(f(x)f(y)) = f^{-1}(x)f(z)),$$

[1] Eq. (29) in (Smith and Eriscon 1989).

[2] In Chapter 2 of this book, Jaynes basically outlines the Cox's proof.

$$f(x)f(y) = f(x)f(z),$$

$$f(y) = f(z).$$

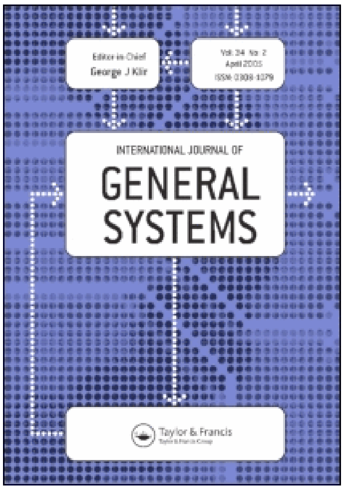
Since f is a bijective function, we obtain that $y = z$, so F is reducible from one side. Moreover, since F is a symmetric function, it is also reducible from the other side.

In this Appendix, we set to examine, strictly on mathematical grounds, why the Cox's proof does not justify claims made by some advocates of probability theory (Lindley, Cheeseman, and others) that *probability is the only sensible description of uncertainty*, that *every uncertainty statement must be in the form of probability*, that *the calculus of probabilities is adequate to handle all situations involving uncertainty*, and the like. However, the Cox's proof, as a justification for such extreme claims, is also vulnerable on philosophical grounds. This is beyond the aim of this Appendix, but we consider it worth to refer to a paper by Colyvan (2004), where these aspects are thoroughly discussed.

Fortunately, the debates outlined in Section 2.4 (involving not only fuzzy logic, but also some other theories of uncertainty different from probability theory) has ended on a positive side in the sense that they helped to curtail the extreme claims of some radical probability advocates to claims that are more reasonable, as adequately expressed by one of the strong advocates of probability theory, Kevin Van Horn[1]:

> Although there is not a completely compelling case for Cox's axioms, and thus one cannot claim that probability theory is the only workable logic of uncertain reasoning, there are strong grounds for a weaker, but still interesting claim: probability is the *simplest* workable logic of uncertain reasoning one could hope to construct.

[1]Conclusion in a discussion paper by Van Horn (*Intern J. of Approximate Reasoning*, 35, 2004, 109-110).



International Journal of General Systems
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713642931

IS THERE MORE TO UNCERTAINTY THAN SOME PROBABILITY THEORISTS MIGHT HAVE US BELIEVE?*

George J. Klir[a]
[a] Department of Systems Science, Thomas J. Watson School, State University of New York, Binghamton, New York, USA

To cite this Article Klir, George J.(1989) 'IS THERE MORE TO UNCERTAINTY THAN SOME PROBABILITY THEORISTS MIGHT HAVE US BELIEVE?*', International Journal of General Systems, 15: 4, 347 — 378
To link to this Article: DOI: 10.1080/03081078908935057
URL: http://dx.doi.org/10.1080/03081078908935057

Int. J. General Systems, Vol. 15, pp. 347–378
Reprints available directly from the publisher
Photocopying permitted by license only

IS THERE MORE TO UNCERTAINTY THAN SOME PROBABILITY THEORISTS MIGHT HAVE US BELIEVE?*

GEORGE J. KLIR

Department of Systems Science, Thomas J. Watson School, State University of New York, Binghamton, New York 13901, USA

(Received 1 August 1988; in final form 14 July 1989)

The aim of the paper is to challenge the claims (as described by Lindley[38]), "that probability is the only sensible description of uncertainty and is adequate for all problems involving uncertainty. All other methods are inadequate". The paper concentrates primarily on the codification of the concept of uncertainty and the discussion of adequate principles of maximum and minimum uncertainty.

INDEX TERMS: Uncertainty, probability theory, fuzzy set theory, possibility theory, Dempster–Shafer theory, fuzziness, nonspecificity, dissonance, confusion.

1. INTRODUCTION: SETTING THE STAGE

My position in this debate is rather modest: I intend to challenge certain extreme claims regarding the concept of uncertainty that are maintained by some probability theorists. The claims I have in mind are perhaps most explicitly (forcefully) expressed by Lindley in the following quote from his recent paper[38] (italics added by me):

The only satisfactory description of uncertainty is probability. By this I mean 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 ... probability is *the only* sensible description of uncertainty and is adequate for *all* problems involving uncertainty. *All* other methods are inadequate. ... *Anything* that can be done with fuzzy logic, belief functions, upper and lower probabilities, or *any other* alternative to probability can *better* be done with probability.

Similar claims can also be found in another paper by Lindley[37] and in several debate-like papers by Cheeseman.[3–6] For example, Cheeseman[4] declares as the theme of one of his papers "that all reasoning under uncertainty can be fully captured, and captured correctly, by probability". In another paper,[3] he writes:

The numerous schemes for representing and reasoning about uncertainty that have appeared in the AI literature are unnecessary—probability is all that is needed.

*This paper, which expresses my view about the concept of uncertainty, is based upon an official debate organized at the Eighth Maximum Entropy Workshop at St. John's College, Cambridge, U.K., August 1–5, 1988. My opponent, whose position was that uncertainty can be fully and correctly captured by probability theory, was Peter Cheeseman. The paper was distributed among participants of the Workshop, but it is not included in the Proceedings of the Workshop (*Maximum Entropy and Bayesian Methods*, edited by J. Skilling, Kluwer, Dordrecht, 1989).

Although other probability theorists express similar views, the chosen quotes are sufficient for my purpose, as they capture well the essence of the claims I propose to challenge in this debate. That is, I intend to argue that probability theory is not capable of capturing the full scope of uncertainty and, consequently, various broader or alternative formalisms are desirable for dealing with certain situations involving uncertainty.

Since there is a great variety of views on the nature of the concept of probability,[15] we must be clear about what we mean by this concept. I assume that our debate is based upon the concept of probability as it is understood and employed in the Bayesian/Maximum Entropy methodology.[22] In this context, only the concept of a conditional probability is recognized: it is viewed as a measure of a reasonable credibility of (belief in) one proposition when another proposition (evidence) is known to be true. As is well known, this concept of conditional probability was formally derived from some reasonable requirements by Cox.[10] Mathematically, this concept is equivalent to the concept of conditional probability in the standard quantitative (numerical) probability theory,[15] as axiomatized, for example, by Kolmogorov.[32]

Hence, when arguing the adequacy of probability theory for describing uncertainty in this debate, I assume that the subject under discussion is the standard quantitative probability à la Kolmogorov,[32] particularly its Bayesian interpretation employed in the derivation by Cox.[10] That is, I assume that the term "probability" is not used in this debate in the broad sense, covering, for example, varieties of classificatory and comparative notions of probability.[50]

To make sure that my position in this debate is properly understood, let me emphasize that I am critical of neither the Bayesian/Maximum Entropy methodology, as well described by Jaynes,[22] nor related methodologies, such as the one based on the principle of minimum information (which seems to be a generalization of the Bayesian rule[51]) or the entropy minimax methodology, developed and applied with great skill and success in the area of predictive modelling by Christensen.[7–9] On the contrary, I believe that there is now enough evidence to demonstrate that these related methodologies are superior to orthodox statistics in terms of foundational and pragmatic grounds.[8,23,45] In fact, I have been using the various entropy principles for many years in my work in systems science, as documented in some of my publications.[26,27,31]

In summary, when uncertainty is conceptualized in terms of probability theory, I believe that the right way to deal with situations involving uncertainty is to use the Bayesian/Maximum Entropy methodology or some of the related methodologies. Hence, I am in full agreement with the Bayesian/Maximum Entropy position within the domain of probability theory. I also believe, however, that probabilistic conceptualization is restrictive in the sense that it does not capture the full scope of uncertainty. In particular, I believe, and intend to argue, that uncertainty is a multidimensional concept and that probability theory allows us to capture only one of its dimensions. To also capture the other dimensions of uncertainty, mathematical frameworks that are either complementary to or broader than probability theory are needed.

I hope that the principal issue of this debate is now clear:

Claims to be challenged—probability theory is the only satisfactory mathematical framework to describe uncertainty; it is adequate for dealing with all situations involving uncertainty (Lindley, Cheeseman).

My counterclaims—probability theory is capable of conceptualizing only one type of uncertainty; to capture the full scope of uncertainty, one has to go beyond probability theory (Klir).

To develop arguments supporting my counterclaims, I deem it essential to start with a general (common sense) discussion of the concept of uncertainty.

2. WHAT IS UNCERTAINTY?

As a starting point, let us consult a standard dictionary about the term "uncertainty". We find that it has a broad semantic content. For example, *Webster's New Twentieth Century Dictionary* defines *uncertainty* as the quality or state of being uncertain. It gives six clusters of meanings of the term *uncertain*:

1. not certainly known, questionable, problematical;
2. vague, not definite or determined;
3. doubtful, not having certain knowledge, not sure;
4. ambiguous;
5. not steady or constant, varying;
6. liable to change or vary; not dependable or reliable.

When we use the dictionary again to examine these various meanings, two major types of uncertainty emerge quite naturally. They are quite well captured by the terms "vagueness" and "ambiguity".

In general, *vagueness* is associated with the lack of precise or sharp distinctions or boundaries. *Ambiguity*, on the other hand, is associated with one-to-many relations, that is, situations in which several alternatives are left unspecified or a desired categorization of an element is left undecided due to ignorance.

Each of the two major types of uncertainty—vagueness and ambiguity—is connected with a fairly large set of kindred concepts. Some of the concepts connected with vagueness are: fuzziness, haziness, cloudiness, unclearness, indistinctiveness, sharplessness, indefiniteness; some of the concepts connected with ambiguity are: nonspecificity, variety, generality, diversity, divergence, equivocation, incongruity, discrepancy, dissonance, disagreement. There are also concepts that are connected with both vagueness and ambiguity. An example is the concept of imprecision. When it is used to express that a distinction made in some context is not sharp, imprecision relates to vagueness; when, on the other hand, it is used to describe the lack of specificity in a situation, imprecision relates to ambiguity.

Further inspection of the concept of ambiguity reveals that two distinct types of ambiguity can readily be distinguished. One is connected with the variety of alternatives that in a given situation are left unspecified; this type of ambiguity is well described by the term "*nonspecificity*". The other type of ambiguity is connected with the disagreement resulting from the attempt to classify an element of a given universal set into two or more disjoint subsets of interest under total or partial ignorance regarding relevant characteristics of the element. The term "*dissonance*" seems to be sufficiently suggestive of this type of ambiguity.

The concept of *uncertainty* is closely connected with the concept of *information*. When our uncertainty in some situation is reduced by an action (such as an observation, performing an experiment, receiving a message, or finding an historical record), the action may be viewed as a source of information pertaining

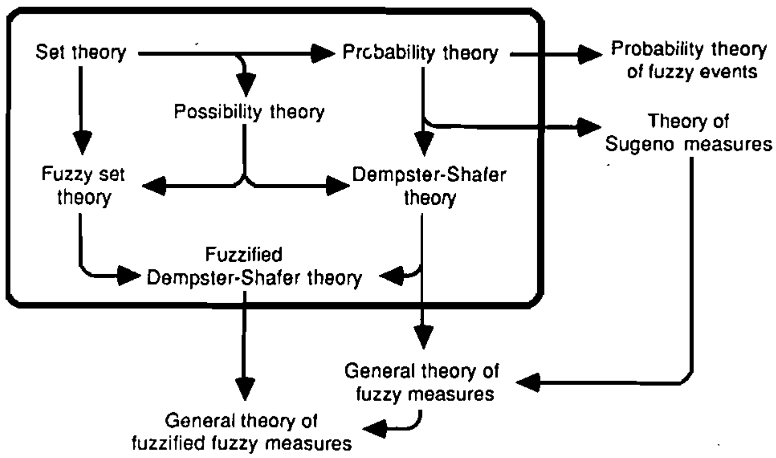


Figure 1 Some mathematical theories capable of conceptualizing situations under uncertainty.

to the situation under consideration. The amount of information obtained by the action may be measured by the reduction of uncertainty that results from the action.

Information in the sense just described does not cover semantic and pragmatic aspects. Consequently, it is not what we usually mean by information in human communication. The notion of information used here is defined strictly in terms of uncertainty reduction within given syntactic and semantic frameworks, which are assumed to be fixed in each particular application. This restricted concept of information may be well described by the term "*uncertainty-based information*".

3. MATHEMATICAL FRAMEWORKS FOR CONCEPTUALIZING UNCERTAINTY

The classical mathematical frameworks for characterizing situations under uncertainty have been set theory and probability theory. Since the mid-1960s, a number of generalizations of these classical theories became available for conceptualizing uncertainty. Names of some of these generalized theories are given in Figure 1; each arrow in the figure indicates some sort of generalization.

In this debate, my aim is to compare the six theories identified in Figure 1 by the shaded area (fuzzy set theory, possibility theory, basic and fuzzified Dempster–Shafer theory, and the two classical theories: set theory and probability theory) by their capabilities of conceptualizing the three main types of uncertainty: vagueness, nonspecifity, and dissonance. These theories are chosen because they are currently the most developed theories for dealing with situations under uncertainty. Moreover, appropriate measures of uncertainty are now well justified in each of the theories. For the remaining theories listed in Figure 1, let me cover them only by appropriate references: probability theory of fuzzy events,[55] theory of Sugeno measures,[1,46] general theory of fuzzy measures,[46] general theory of fuzzified fuzzy measures (not developed as yet).

To facilitate the discussion, let me overview relevant basic properties of the four non-classical theories; knowledge of set theory and probability theory is assumed.

Let X denote a universal set under consideration, assumed here to be finite for the sake of simplicity, and let $P(X)$ denote the power set of X. Then, the Dempster–Shafer theory is based upon a function

$$m: P(X) \rightarrow [0, 1]$$

such that

$$m(\phi) = 0 \quad \text{and} \quad \sum_{A \subseteq X} m(A) = 1.$$

This function is called a *basic assignment*; the value $m(A)$ represents the degree of belief (based on relevant evidence) that a specific element of X belongs to set A, but not to any particular subset of A. Every set $A \in P(X)$ for which $m(A) \neq 0$ is called a *focal element*. The pair (F, m), where F denotes the set of all focal elements of m, is called a *body of evidence*.

Associated with each basic assignment m is a pair of measures, a *belief measure*, Bel, and a *plausibility measure*, Pl, which are determined for all sets $A \in P(X)$ by the equations

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B), \tag{1}$$

$$\text{Pl}(A) = \sum_{B \cap A \neq \phi} m(B). \tag{2}$$

These equations and the definition of the basic assignment form the core of the Dempster–Shafer theory. This theory is best described by Shafer.[43]

Belief and plausibility measures are connected by the equation

$$\text{Pl}(A) = 1 - \text{Bel}(\bar{A}) \tag{3}$$

for all $A \in P(X)$, where $\bar{A}$ denotes the complement of A. Furthermore,

$$\text{Bel}(A) \leqq \text{Pl}(A) \tag{4}$$

for all $A \in P(X)$.

A belief measure (or a plausibility measure) becomes a *probability measure* when all focal elements are *singletons*. In this case, $\text{Bel}(A) = \text{Pl}(A)$ for all $A \in P(X)$, which follows immediately from Eqs. (1) and (2). When some focal elements are not singletons, the additivity of probability theory bifurcates into the more general properties of superadditivity for belief measures,

$$\text{Bel}(A \cup B) \geqq \text{Bel}(A) + \text{Bel}(B) - \text{Bel}(A \cap B), \tag{5}$$

and subadditivity for plausibility measures,

$$\text{Pl}(A \cap B) \leqq \text{Pl}(A) + \text{Pl}(B) - \text{Pl}(A \cup B). \tag{6}$$

When all focal elements are nested (ordered by set inclusion), the body of evidence is called *consonant*. In this case, we obtain special plausibility measures, which are called *possibility measures* (or *consonant plausibility measures*), and the corresponding special belief measures, which are called *necessity measures*. A possibility measure, π, is conveniently (and uniquely) determined by a *possibility distribution function*

$$r: X \to [0, 1]$$

via the formula

$$\pi(A) = \max_{x \in A} r(x) \tag{7}$$

for all $A \in P(X)$. The corresponding necessity measure, η, is then determined for all $A \in P(X)$ by a formula equivalent to Eq. (3),

$$\eta(A) = 1 - \pi(\bar{A}). \tag{8}$$

A theory that deals with consonant bodies of evidence in terms of possibility and necessity measures is usually called a *possibility theory*. The properties of superadditivity and subadditivity of the Dempster–Shafer theory, expressed by Eqs. (5) and (6), assume in possibility theory for all A, $B \in P(X)$ the forms

$$\eta(A \cap B) = \min [\eta(A), \eta(B)], \tag{9}$$

$$\pi(A \cup B) = \max [\pi(A), \pi(B)], \tag{10}$$

respectively.[14,29] Furthermore, given a consonant body of evidence

$$F = \{A_1, A_2, \ldots, A_n\} \tag{11}$$

such that $A_1 \subset A_2 \subset \cdots \subset A_n$, the basic assignment in possibility theory is connected with the possibility distribution via the formula

$$m(A_i) = r(x_i) - r(x_{i+1}) \tag{12}$$

for some $x_i \in A_i$, some $x_{i+1} \in A_{i+1}$, and $i = 1, 2, \ldots, n$, where $r(x_{n+1}) = 0$ by convention.[29] Similarly, given a basic assignment of a consonant body of evidence of the form (11), the corresponding possibility distribution is calculated by the formula

$$r(x_i) = \sum_{k=i}^{n} m(A_k) \tag{13}$$

for each $x_i \in A_i$.[29]

Possibility theory can be formulated not only in terms of consonant bodies of evidence within the Dempster–Shafer theory, but also in terms of fuzzy sets.[54] It was introduced in this latter manner by Zadeh.[56] A *fuzzy set* is a set whose boundary is not sharp. That is, the change from nonmembership to membership in a fuzzy set is gradual rather than abrupt. This gradual change is expressed by a *membership grade function*, μ_A, of the form

$$\mu_A: X \rightarrow [0, 1],$$

where A is a label of the fuzzy set defined by this function within the universal set X. The value $\mu_A(x)$ expresses the grade of membership of element x of X in the fuzzy set A or, in other words, the degree of compatibility of x with the concept represented by the fuzzy set. A fuzzy set A is called *normalized* when $\max_{x \in X} \mu_A(x) = 1$. If $\mu_A(x) \leqq \mu_B(x)$ for all $x \in X$, then A is called a *fuzzy subset* of B.

An important concept associated with fuzzy sets is an α-cut. Given a fuzzy set A and a specific number $\alpha \in [0, 1]$, the *α-cut*, A_α, is a crisp (nonfuzzy) set

$$A_\alpha = \{x \in X \mid \mu_A(x) \geqq \alpha\}.$$

The set of all elements of X for which $\mu_A(x) > 0$ is called a *support* of the fuzzy set A; it is usually denoted by $\mathrm{supp}(A)$.

For some applications, the concept of the fuzzy set can be extended in various ways. An important extension is to allow a more general form of the membership grade function,

$$\mu_A: X \rightarrow L,$$

where L denotes a partially ordered set (usually a lattice). Fuzzy sets defined by this more general function are called *L-fuzzy sets*.[16] Simple but useful examples of L-fuzzy sets are *interval-valued fuzzy sets*. They are defined by membership grade functions of the form

$$\mu_A: X \rightarrow P([0, 1]),$$

where for each $x \in X$, $\mu_A(x)$ is a closed interval in $[0, 1]$.

Given a regular fuzzy set A with membership grade function μ_A (the range of μ_A is $[0, 1]$), Zadeh[56] defines a possibility distribution function, r_A, associated with A as numerically equal to μ_A, i.e.,

$$r_A(x) = \mu_A(x) \tag{14}$$

for all $x \in X$; then, he defines the corresponding possibility measure π_A by the equation

$$\pi_A(B) = \max_{x \in B} r_A(x) \tag{15}$$

for all $B \in P(X)$. In this interpretation of possibility theory, focal elements correspond to distinct α-cuts A_α of the fuzzy set A. This follows from the property that $A_\alpha \subseteq A_\beta$ when $\alpha > \beta$.

As observed by Yager,[53] the Dempster–Shafer theory can be fuzzified. In its fuzzified form, the basic assignment is a function

$$\tilde{m}: \tilde{P}(X) \rightarrow [0, 1],$$

where $\tilde{P}(X)$ denotes the set of all fuzzy subsets of X. This function must satisfy the same requirements for the extended domain $\tilde{P}(X)$ as function m does for the domain $P(X)$. Plausibility and belief measures based upon $\tilde{m}$ are expressed by the following generalized counterparts of Eqs. (1) and (2),

$$\mathrm{Bel}(A) = \sum_{B \in F} \tilde{m}(B) \left[1 - \max_{x \in X} \min (1 - \mu_A(x), \mu_B(x)) \right], \tag{1'}$$

$$\mathrm{Pl}(A) = \sum_{B \in F} \tilde{m}(B) \left[\max_{x \in X} \min (\mu_A(x), \mu_B(x)) \right], \tag{2'}$$

where $\mu_A(x)$ and $\mu_B(x)$ are degrees of membership of element x in fuzzy sets A and B, respectively, and F is the set of all focal elements (fuzzy sets) associated with $\tilde{m}$.

4. MEASURES OF UNCERTAINTY

Let me explore now the conceptualization and measurement of the various types of uncertainty within the six mathematical frameworks I chose to cover in this debate. Let me start with classical set theory.

When a situation is expressed in terms of a set of alternatives that are left undecided, it is clear that the uncertainty assumes the form of nonspecificity. The more alternatives, the less specific the situation is; when only one alternative is possible, the situation is fully specific. A measure of this sort of uncertainty was introduced by Hartley in 1928,[18], even though Hartley called it a measure of information.

Given a finite crisp set A of possible alternatives, Hartley derived a simple function

$$I(A) = \log_2 |A|, \tag{16}$$

where $|A|$ denotes the cardinality of set A, as the only meaningful measure (except for a multiplication constant) of the amount of information needed for characterizing one element of the set. Later, the uniqueness of the measure was proven axiomatically by Rényi.[41] For our purpose, it is important to realize that *nonspecificity* expressed by the *Hartley measure* is the only type of uncertainty that can be conceptualized within *classical set theory.*

A measure of *probabilistic uncertainty* (and the associated information) was established by Shannon in 1948.[44] This measure, whose basic form is

$$H(p(x) \mid x \in X) = -\sum_{x \in X} p(x) \log_2 p(x), \tag{17}$$

where $(p(x) \mid x \in X)$ denotes a probability distribution on X, is usually called the *Shannon entropy.* It is well justified, in a number of alternative ways, as a unique measure of uncertainty conceptualized in terms of probability theory.[23,29,41,45]

What type of uncertainty does the Shannon entropy measure? We may easily conclude that it does not measure nonspecificity: the probabilities $p(x)$ are required to be real numbers in $[0, 1]$, and thus fully specific, and each of them focuses on a single (i.e., specific) alternative x. We observe, however, that each probability $p(x)$ in Eq. (17) expresses the degree of belief (based upon some evidence) that x is the true alternative in a given context. In probability theory, X (when finite) is always viewed as a set of exhaustive and mutually exclusive alternatives (outcomes, states, elementary events, basic propositions). Hence, the beliefs expressed by the probabilities in the distribution $(p(x)\,|\,x \in X)$ conflict with each other. The greater the lack of discrimination among the beliefs, the greater the conflict. This is precisely how the Shannon entropy behaves. Hence, the Shannon entropy is a *measure of dissonance* in a probability distribution (using the terminology introduced in Section 2).

Although we understand now that the Hartley measure and the Shannon entropy quantify different types of uncertainty, this fact was obscured by two prevailing views of experts in probabilistic information theory about the Hartley measure. According to one of the views, the Hartley measure is a special case of the Shannon entropy that distinguishes only between zero and nonzero probabilities; otherwise, it is totally insensitive to the actual values of the probabilities. That is, probabilities are withdrawn according to this view and only possibilities of elements of X are retained. There is certainly no need to introduce probabilities and, then, withdraw them to obtain the Hartley measure.

According to the second view, the Hartley measure is a special case of the Shannon entropy that emerges from the uniform probability distribution. This view is ill-conceived since the Hartley measure is totally independent of any probabilistic assumptions, as correctly recognized by Kolmogorov[33] and Rényi.[41] Strictly speaking, the Hartley measure is based upon one concept only—the concept of a finite set of possible alternatives, which can be interpreted as experimental outcomes, states of a system, events, messages, and the like, or as sequences of these. In order to use this measure, possible alternatives must be distinguished, within a given universal set, from those that are not possible. It is thus the *possibility* of each relevant alternative that matters in the Hartley measure. Hence, the Hartley measure can be meaningfully generalized only through broadening the notion of possibility. This avenue is now available in terms of possibility theory.

A natural generalization of the Hartley measure in possibility theory was discovered by Higashi and Klir in 1983.[20] They coined for it the name *U-uncertainty*. When a possibility distribution is expressed in terms of a normalized fuzzy set A, the U-uncertainty has the form

$$U(A) = \int_0^1 \log_2 |A_\alpha|\, d\alpha, \tag{18}$$

where $|A_\alpha|$ denotes the cardinality of the α-cut of the fuzzy set A. The uniqueness of this function as a *possibilistic measure of nonspecificity* under appropriate requirements was proved by Klir and Mariano.[30] For fuzzy sets that are not normalized and for which $\max_{x \in X} \mu_A(x) = a$, each value $U(A)$, given by Eq. (18), must be divided by a.

When possibility theory is interpreted in terms of the Dempster–Shafer theory, the possibilistic measure of nonspecificity is expressed by a function V that for any given nested body of evidence (F, m) assumes the form

$$V(m) = \sum_{A \in F} m(A) \log_2 |A|. \tag{19}$$

Clearly, when m corresponds to a possibility distribution represented by a fuzzy set A via Eq. (12), then $V(m) = U(A)$. Furthermore, it is now well established that function V defined by Eq. (19) is applicable as a measure of nonspecificity to any arbitrary body of evidence, not only those that are nested.[29] Its uniqueness in this general setting was proved by Ramer.[40]

Since focal elements of probability measures are singletons, $V(m) = 0$ for every probability measure. That is, there is no nonspecificity in probability measures; all probability measures are fully specific, as already anticipated on intuitive grounds. Hence, *probability theory is not capable of conceptualizing nonspecificity*, one of the basic types of uncertainty.

Let me discuss now the meaning of the Shannon entropy within the Dempster–Shafer theory. Since the Dempster–Shafer theory is a generalization of probability theory, obtained by relaxing the additivity requirement, is it reasonable to expect that some appropriately generalized form of the Shannon entropy exists that is universally applicable within the broader theory? This question has already been answered in the affirmative by establishing that the Shannon entropy in fact bifurcates within the Dempster–Shafer theory into the following two generalized forms:

$$E(m) = -\sum_{A \in F} m(A) \log_2 \mathrm{Pl}(A), \tag{20}$$

$$C(m) = -\sum_{A \in F} m(A) \log_2 \mathrm{Bel}(A). \tag{21}$$

Function E defined by Eq. (20) is usually called a *measure of dissonance* and function C given by Eq. (21) is called a *measure of confusion*. Since both E and C collapse into the Shannon entropy when m represents a probability distribution, they are sometimes referred to as *entropy-like uncertainty measures*.

What do functions E and C actually measure? From Eq. (2) and the general property of basic assignments (satisfied for every $A \in P(X)$),

$$\sum_{B \cap A = \phi} m(B) + \sum_{B \cap A \neq \phi} m(B) = 1,$$

we obtain

$$E(m) = -\sum_{A \in F} m(A) \log_2 \left[1 - \sum_{B \cap A = \phi} m(B) \right]. \tag{22}$$

The term

$$K = \sum_{B \cap A = \phi} m(B)$$

in this expression for $E(m)$ clearly represents the total conflict between the belief in A and other beliefs within a given body of evidence. It is obtained simply by adding the basic assignment values of all focal elements that are disjoint with A and, consequently, the beliefs allocated to them are in conflict with the belief focusing on A. The value of K ranges from 0 to 1. The function

$$-\log_2[1-K],$$

which is employed in Eq. (22), is monotonic increasing with K; it extends the range from $[0,1]$ to $[0,\infty)$. The choice of the logarithmic function is based on the same motivation as the choice of the logarithmic function in the Shannon entropy. Now, we can readily see that $E(m)$ defines the mean (expected) value of the conflict in beliefs associated with a given body of evidence (F, m); the name "measure of dissonance" is thus quite appropriate. This observation reinforces my previous argument that the Shannon entropy measures the degree of conflict (dissonance) among beliefs expressed by a probability distribution.

Let me now explain the meaning of function C given by Eq. (21). From Eq. (1) and the general property of basic assignments (satisfied for every $A \in P(X)$),

$$\sum_{B \subseteq A} m(B) + \sum_{B \nsubseteq A} m(B) = 1,$$

we get

$$C(m) = -\sum_{A \in F} m(A) \log_2 \left[1 - \sum_{B \nsubseteq A} m(B) \right]. \tag{23}$$

The term

$$L = \sum_{B \nsubseteq A} m(B)$$

in this expression of $C(m)$ stands for the sum of all focal elements that either do not overlap with set A or overlap with it only partially. Since beliefs in these focal elements B are in actual or potential conflict with the belief in A (since $B \nsubseteq A$ by definition), L represents the total real and potential conflict with the belief in A. The reasons for using

$$-\log_2[1-L]$$

instead of L in Eq. (23) are the same as already discussed in the context of function E. The conclusion is that $C(m)$ defines the mean (expected) value of not only the real conflict (as function E does), but also of the potential conflict associated with a given body of evidence. This multitude of partially or totally conflicting focal elements is a source of confusion; hence the name "measure of confusion".

Since focal elements of possibility measures are nested, the plausibility of each focal element must be 1 (by Eq. (2)) and, consequently, $E(m)=0$ when m defines a possibility measure. That is, consonant bodies of evidence (and the associated possibility and necessity measures) are free of dissonance (real conflict in beliefs).

However, they are not free of potential conflict since $C(m)\neq 0$ in general; $C(m)=0$ if and only if the body of evidence contains only one focal element.

The question of how to measure the degree of vagueness or fuzziness of a fuzzy set has been one of the basic issues of fuzzy set theory. Various measures of vagueness, more often called *measures of fuzziness*, have been proposed.[29] One way of measuring fuzziness, which was suggested by Yager[52] and further investigated by Higashi and Klir,[19] is to view it as the lack of distinction between the fuzzy set and its complement. Clearly, the less a fuzzy set differs from its complement, the fuzzier it is. Using this approach, the measure of fuzziness depends on the complementation operator employed (which is not unique[19]) and on the distance function by which the distinction between the set and its complement is expressed. Let F_c denote a measure of fuzziness based upon the complementation operator c and the Hamming distance. Then, assuming again a finite universal set X, F_c is given by the formula

$$F_c(A)=|X|-\sum_{x\in X}|\mu_A(x)-c(\mu_A(x))|. \tag{24}$$

Observe that the concept of the amount of fuzziness, measured by function F_c, has no applicability in probability theory: no vagueness is allowed in defining probabilities.

When the Dempster–Shafer theory is fuzzified, measures of the three relevant types of uncertainty, given by Eqs. (19), (20), (21), are still applicable provided that the entries in their formulas are properly interpreted: since focal elements A are now fuzzy sets, $|A|$ in Eq. (19) must be calculated by the formula

$$|A|=\sum_{x\in X}\mu_A(x), \tag{25}$$

which defines a *simple* (*scalar*) *cardinality of fuzzy set* A; in Eqs. (20) and (21), values of $\mathrm{Pl}(A)$ and $\mathrm{Bel}(A)$ must be calculated by Eqs. (2′) and (1′), respectively. In addition, it becomes also meaningful to measure fuzziness of a given fuzzified body of evidence. We may, for example, use the formula

$$\tilde{F}_c(\tilde{m})=\sum_{A\in F}\tilde{m}(A)F_c(A), \tag{26}$$

where $\tilde{F}_c(\tilde{m})$ denotes the fuzziness of $(F,\tilde{m})$ and $F_c(A)$, which is determined by Eq. (24), denotes the fuzziness of the focal element (a fuzzy set) A.

5. LIMITATIONS OF PROBABILITY THEORY

The purpose of this section is to discuss some limitations or inadequacies of probability theory for dealing with various situations under uncertainty. Comparisons are restricted only to the theories identified in Figure 1 by the shaded area. The section is organized in the form of a collection of independent or loosely related remarks regarding various conceptual, formal, computational, methodological, and application-related issues. The remarks are general; specific examples are discussed in Section 6.

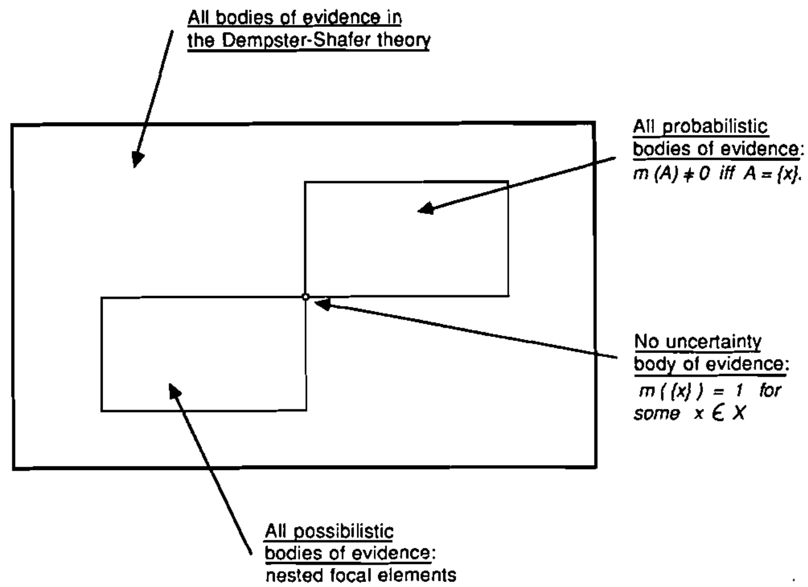


Figure 2 Subset relationship of probabilistic and possibilistic bodies of evidence within the Dempster–Shafer theory.

1. First, it is essential for our purpose to compare mathematical structures of the theories of our concern, using probability theory as a reference. When comparing probability theory with the Dempster–Shafer theory, it is clear that the latter is more general. This follows from the fact that the additivity axiom of probability theory is replaced with the weaker superadditivity and subadditivity axioms. As a consequence, it is not required in the Dempster–Shafer theory that degrees of belief be allocated to singletons of the universal set. When such allocations are possible and make sense in a particular application, the Dempster–Shafer theory assumes automatically the mathematical structure of probability theory. Hence, probability theory is a subset of the Dempster–Shafer theory in terms of the bodies of evidence that can be conceptualized by either of the theories. Another subset of the Dempster–Shafer theory is possibility theory, which is virtually disjoint with probability theory. The only body of evidence they share consists of one focal element that is a singleton (the case of total certainty). Since the additivity axiom of probability theory is replaced with the maximum axiom (Eq. (7)) of possibility theory, which guarantees the nested structure of focal elements in possibility theory, the two theories are complementary (neither is contained in the other). The subset relationship among the three theories is summarized in Figure 2.

As far as fuzzy set theory is concerned, it is clearly a generalization of classical set theory: the range $\{0, 1\}$ of characteristic functions extends to the range $[0, 1]$ of the membership grade functions. This results in the violation of some properties of the Boolean lattice of classical set theory. Which properties are violated depends on the chosen operators for fuzzy set union, intersection, and complement, which are not unique and whose choice depends on the context of each application. For the standard fuzzy set operators, only the law of excluded middle and the law of contradiction are violated. A direct comparison of fuzzy set theory with probability theory is not possible since fuzzy set theory does not involve the power set $P(X)$ or a σ-field of selected subsets of X, which are essential for the formalization of probability theory. As already mentioned, however, probability theory can be compared with the fuzzy set interpretation of possibility theory. In this comparison, the two theories emerge as totally distinct except in the one (trivial) body of evidence. Probability theory can be effectively combined with fuzzy set theory by defining, for example, probabilities of fuzzy events,[55] or by characterizing probabilities in terms of fuzzy numbers.[13,24] These combinations are, however, generalizations (enrichments) of probability theory and, consequently, they should not be considered as part of probability theory in this debate.

2. It follows from the brief exposition in Section 4 that uncertainty and uncertainty-based information are, in general, multidimensional concepts. Depending on the mathematical framework employed, uncertainty is manifested by one or more of the following four types we now recognize: *fuzziness* (vagueness), *nonspecificity* (lack of informativeness), *dissonance* (pure conflict), and *confusion* (pure and potential conflict). The multidimensionality of uncertainty was obscured when uncertainty was investigated solely in terms of classical set theory and probability theory, in each of which uncertainty is manifested only by one of its dimensions: by nonspecificity in classical set theory (expressed by the Hartley measure), and by dissonance in probability theory (expressed by the Shannon entropy). The following is a summary of the applicability of the individual measures of the four types of uncertainty in the six mathematical theories of our concern:

a) *classical (crisp) set theory*—nonspecificity (Hartley measure), expressed by Eq. (16);

b) *probability theory*—dissonance and confusion collapse into one measure (Shannon entropy) expressed by Eq. (17);

c) *fuzzy set theory*—fussiness, exemplified by Eq. (24), and nonspecificity, expressed by Eq. (18);

d) *possibility theory*—confusion, expressed by Eq. (21), and nonspecificity, expressed either by Eq. (18) or by Eq. (19), depending on the interpretation employed;

e) *Dempster–Shafer theory*—nonspecificity, dissonance, and confusion, given by Eqs. (19), (20), (21), respectively;

f) *fuzzified Dempster–Shafer theory*—nonspecificity, dissonance, and confusion, expressed by modified Eqs. (19), (20), (21), respectively, in which m is replaced with $\tilde{m}$, $|A|$ is obtained by Eq. (25), and $\mathrm{Pl}(A)$, $\mathrm{Bel}(A)$ are calculated by Eqs. (2′), (1′), respectively.

The undeniable fact that probability is capable of expressing only one of the four distinct types of uncertainty we now recognize and are able to measure makes the extreme probabilistic claim that "every uncertainty statement must be in the form of a probability" unattainable.

3. It is clear from the foregoing that the one-dimensional probabilistic information theory will have to be extended into a multidimensional information

theory, far better equipped to capture the semantic richness of the concepts of uncertainty and uncertainty-based information. Such a research program involves many challenging philosophical, mathematical, and computational issues. In the four non-classical mathematical frameworks (c)–(f), the maximum and minimum uncertainty principles emerge as multiple objective criteria optimization problems or, alternatively, single objective optimization problems with objective functions that are justifiable aggregates of relevant measures of uncertainty. Considering the Dempster–Shafer theory as an example, the uncertainty principles would be expressed either in terms of three objective functions, V, E, and C, or, alternatively, in terms of an appropriate aggregate of these functions. For instance, we may take the total uncertainty, expressed by the sum

$$S(m) = V(m) + E(m) + C(m), \tag{27}$$

as the objective function. In this latter case, for example, the principle of maximum uncertainty for a problem within a finite universal set X would be formulated as follows: determine a basic assignment $m(A)$, for all $A \in P(X)$, that maximizes the function

$$S(m) = \sum_{A \subseteq X} m(A) \log_2 \frac{|A|}{\mathrm{Pl}(A) \cdot \mathrm{Bel}(A)} \tag{28}$$

subject to the given contraints $c_1, c_2, \ldots$, which represent the available information relevant to the matter of concern, as well as the general constraints of the Dempster–Shafer theory.

4. In probability theory, total ignorance is expressed (employing the maximum entropy principle) by the uniform probability distribution on X. This choice is well justified (on several different grounds) providing we require that the situation be characterized by a single probability distribution (the usual requirement of probability theory). This requirement, however, is too strong to allow us to obtain an honest characterization of total ignorance. Indeed, if no information is available about the situation under consideration, then every probability distribution on the given universal set is equally possible (or equally probable, if you like). Hence, an honest characterization of total ignorance should be expressed in terms of the full set of possible probability distributions on X, allowing thus nonspecificity in the formulation. Such a formulation, however, is foreign to probability theory. In the broader framework of the Dempster–Shafer theory (as well as in the narrower framework of possibility theory), where uncertainty in the form of nonspecificity is acceptable, total ignorance is expressed by $m(X) = 1$ and $m(A) = 0$ for all $A \neq X$. This is certainly an honest expression of total ignorance that perfectly agrees with our common sense: we know that the element is in the universal set X, but we have no evidence about its location in any subset of X. Using Eq. (1) or Eq. (2), we can also express total uncertainty in the form

$$\mathrm{Bel}(X) = 1 \quad \text{and} \quad \mathrm{Bel}(A) = 0 \quad \text{for all} \quad A \neq X,$$

or in the form

$$\mathrm{Pl}(\phi)=0 \quad \text{and} \quad \mathrm{Pl}(A)=1 \quad \text{for all} \quad A\neq\phi,$$

respectively. These equivalent forms are also perfectly agreeable to our common sense: the degree of our belief in every set except the universal set X is zero while, at the same time, every set except the empty set is considered fully plausible ($\mathrm{Pl}(A)=1$). When we measure the three applicable types of uncertainty, we obtain $V(m)=\log_2|X|$, $E(m)=0$, and $C(m)=0$. This, again, is perfectly compatible with our intuition: total ignorance is expressed fully in terms of nonspecificity (lack of informativeness). Indeed, having no information, we have no basis to form any meaningful beliefs and, consequently, there cannot be any conflict (be it real or potential) among beliefs that do not exist. In probability theory, total ignorance involves the same amount of uncertainty ($H(p)=\log_2|X|$), but this uncertainty is expressed solely in terms of the maximized conflict among beliefs (measured by the Shannon entropy); these beliefs are not derived from any evidence (which is nonexistent), but are dictated by the restrictive formalism of probability theory.

5. As already mentioned, probability theory captures totally different bodies of evidence (with the one trivial exception) than possibility theory. These theories also involve different types of uncertainty. In probability theory, our ignorance is expressed solely in terms of dissonance (measured by the Shannon entropy). In possibility theory, it is expressed predominantly in terms of nonspecificity (measured by function V), even though some amount of potential conflict (measured by function C) may also be included. Furthermore, since possibility distributions are in a one-to-one correspondence with fuzzy sets, it is also meaningful to characterize possibility distributions by their degrees of fuzziness. All these observations indicate that the two formalisms are complementary in their applicability. For example, probability is suitable for characterizing the number of persons that are expected to ride in a particular car each day. Possibility theory, on the other hand, is suitable for characterizing the number of persons that can ride in that car at any one time. Since the physical characteristics of a person (such as size or weight) are intrinsically vague, it is not realistic to describe the situation by a sharp distinction between possible and impossible instances. A more realistic possibility distribution (defined here on the set of positive integers) might be, for example, $r(x)=1$ for $x\leqq 5$, $r(6)=0.9$, $r(7)=0.5$, $r(8)=0.1$, and $r(x)=0$ for $x\geqq 9$. The need for possibilistic (consonant) characterization of uncertainty in economics has been argued for many years by the British economist Shackle.[42]

6. Possibility theory is computationally less sensitive to errors in the assessment of possibility degrees than probability theory is in the assessment of probability degrees. This follows from the type of operations employed in the two theories. In possibility theory, where the basic operations are the maximum and minimum, the error does not accumulate when we operate on possibility distributions. That is, the error cannot exceed the largest error in the assessment of possibilities no matter how many times we operate with them. In probability theory, where the basic operations are the sum and product, the error increases with the number of operations performed.

7. An interesting and conceptually useful notion of a *geometry of fuzzy sets* was recently introduced by Kosko.[35] He interprets fuzzy subsets of a finite universal set X with n elements as points in the n-dimensional unit cube $[0,1]^n$. That is, the entire cube represents the fuzzy power set $\tilde{P}(X)$, its vertices represent the crisp

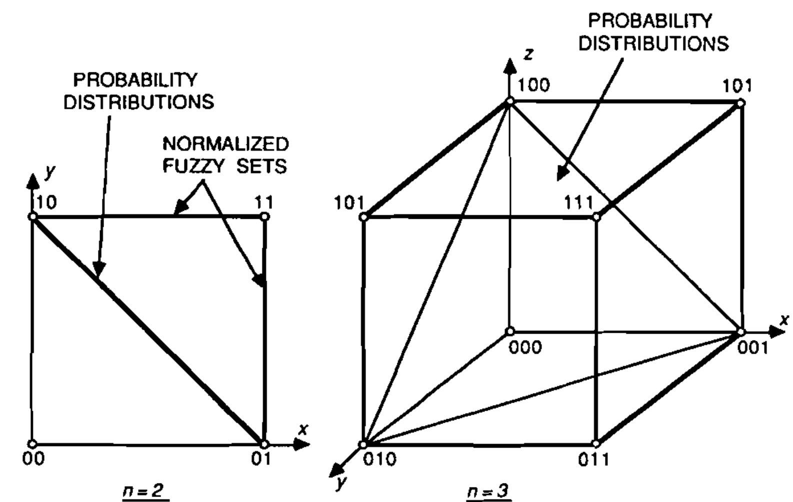


Figure 3 Examples illustrating the geometry of fuzzy sets.

power set $P(X)$. This interpretation suggests that a suitable distance $d(A,B)$ be defined between fuzzy sets. Using, for example, the concept of the Hamming distance, we have

$$d(A,B)=\sum_{x\in X}|\mu_A(x)-\mu_B(x)|. \tag{29}$$

The cardinality $|A|$ of a fuzzy set A, given by Eq. (25), can be then viewed as the distance $d(A,\phi)$ of A from the empty set. Observe that probability distributions are represented by sets whose cardinality is 1. Hence, the set of all probability distributions definable on X is represented by an $(n-1)$-dimensional simplex of the n-dimensional unit cube. Examples of this simplex for $n=2,3$ are shown in Figure 3.

8. The geometry of fuzzy sets introduced in the previous remark suggests that, perhaps, theorems of probability theory can be derived, in a unified fashion, from some more fundamental laws that hold for the whole n-dimensional unit cube and not only for the probabilistic $(n-1)$-dimensional simplex. Kosko demonstrates that, indeed, this can be done.[35] First, he defines for every pair of fuzzy sets $A,B\in P(X)$ the degree of subsethood, $S(A,B)$, of A in B by the formula

$$S(A,B)=\left(|A|-\sum_{x\in X}\max(0,\mu_A(x)-\mu_B(x))\right)\Big/|A|. \tag{30}$$

The $\sum$ term in this formula describes the aggregated (summed) violations of the subset inequality $\mu_A(x)\leq\mu_B(x)$, the difference in the numerator describes the lack of these violations, and the cardinality $|A|$ in the denominator is a normalization factor to obtain the range

$$0\leq S(A,B)\leq 1. \tag{31}$$

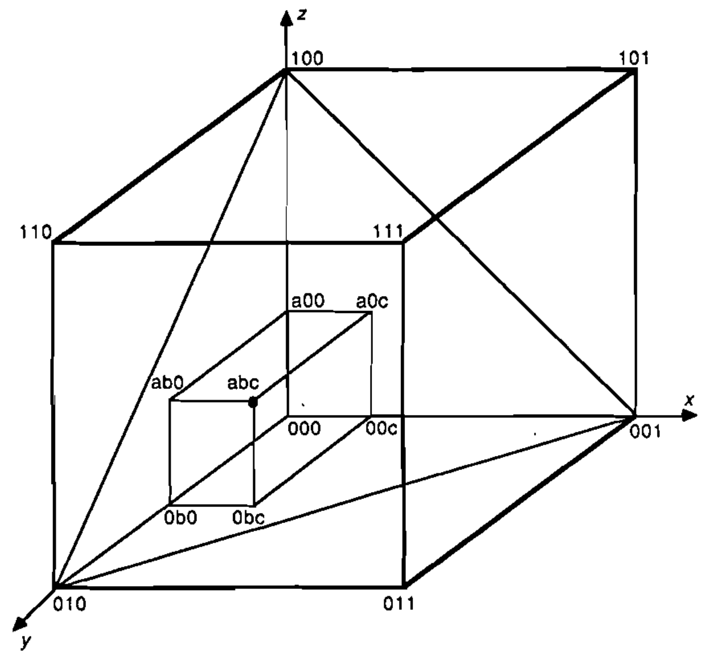


Figure 4 Illustration of the meaning of a probabilistic Boolean lattice in terms of the geometry of fuzzy sets.

Kosko shows[34] that $S(A, B)$ can also be expressed, in a more convenient form, as

$$S(A, B) = \frac{|A \cap B|}{|A|}. \tag{32}$$

Using this formula, it is easy to derive the following properties of $S(A, B)$:

$$S(A, B) = 1 \quad \text{iff} \quad A \subseteq B; \tag{33}$$

$$S(A, B \cup C) = S(A, B) + S(A, C) - S(A, B \cap C); \tag{34}$$

$$S(A, B \cap C) = S(A, B) \cdot S(A \cap B, C); \tag{35}$$

$$S(A, B) = \frac{S(B, A) \cdot |B|}{|A|}; \tag{36}$$

$$\frac{S(A \cap B, C)}{S(A \cap B, \bar{C})} = \frac{S(A \cap C, B)}{S(A \cap \bar{C}, B)} \frac{S(A, C)}{S(A, \bar{C})}. \tag{37}$$

Let me introduce now a probabilistic interpretation of the cardinality and the subsethood function $S(A, B)$. In order to easily visualize the interpretation, let me utilize the fuzzy set geometry within the 3-dimensional unit cube; a generalization to $n > 3$ is trivial. Figure 4 is used as a guide. Let $X = \{\alpha, \beta, \gamma\}$ denote the universal set in our case and let membership grades of the various fuzzy sets be represented in the unit cube $[0,1]^3$ by values of the coordinates z, y, x, respectively. Consider a point (a, b, c) in the 2-dimensional probabilistic simplex (shaded in Figure 4), which represents a fuzzy set Q defined by

$$\mu_Q(\alpha)=a, \quad \mu_Q(\beta)=b, \quad \mu_Q(\gamma)=c.$$

Since Q is located on the probabilistic simplex, $|Q|=a+b+c=1$ by definition, and we may interpret the triple (a, b, c) as a probability distribution. Consider now the fuzzy power set $\tilde{P}(Q)$. Its geometrical representation is a rectangular parallelepiped (shown in Figure 4), whose vertices are of our interest. One of them is the original point (a, b, c) and the other vertices are obtained by replacing one or more entries in (a, b, c) with a zero. The family of fuzzy sets represented by these vertices forms a Boolean lattice under standard fuzzy set operators. Let $\mathscr{B}(a, b, c)$ denote this lattice. Observe that the following one-to-one correspondence as well as numerical equality holds between cardinalities of the fuzzy sets in $\mathscr{B}(a, b, c)$ and probabilities $p(A)$ obtained from the probability distribution (a, b, c) for all $A \in P(X)$:

$$\begin{aligned}
|a,b,c| &= a+b+c = p(\{\alpha,\beta,\gamma\}),\\
|a,b,0| &= a+b \quad = p(\{\alpha,\beta\}),\\
|a,0,c| &= a+c \quad = p(\{\alpha,\gamma\}),\\
|0,b,c| &= b+c \quad = p(\{\beta,\gamma\}),\\
|a,0,0| &= a \quad = p(\{\alpha\}),\\
|0,b,0| &= b \quad = p(\{\beta\}),\\
|0,0,c| &= c \quad = p(\{\gamma\}),\\
|0,0,0| &= 0 \quad = p(\phi).
\end{aligned}$$

Hence, given a probability distribution (a, b, c), we have

$$p(A)=|A'|, \tag{38}$$

where $A \in P(X)$, $A' \in \mathscr{B}(a, b, c)$, and $\mathrm{supp}(A)=\mathrm{supp}(A')$. Now, using Eqs. (32) and (38), we readily obtain

$$p(B \mid A)=S(A', B'), \tag{39}$$

where $A, B \in P(X)$, $A', B' \in \mathscr{B}(a, b, c)$, $\mathrm{supp}(A)=\mathrm{supp}(A')$, and $\mathrm{supp}(B)=\mathrm{supp}(B')$. Applying Eqs. (38) and (39) to Eqs. (31) and (33)–(37), we can convert all the latter equations to their probabilistic form. Observe that the obtained equations all conform to laws of probability theory: Eqs. (31), (34), and (35) become identical with the properties described by Lindley as convexity, addition, and multiplication

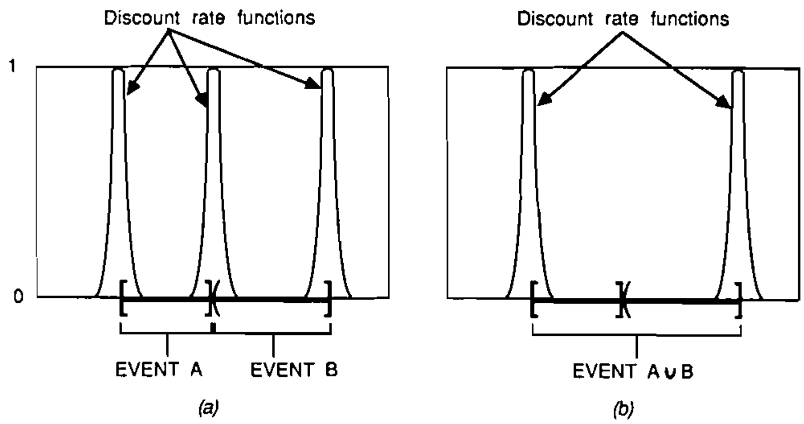


Figure 5 An example illustrating the violation of the additivity axiom of probability theory.

rules, respectively;[38] Eq. (36) becomes Bayes theorem; and Eq. (37) becomes Bayes theorem in odds form.[38] Although the concept of fuzzy subsethood $S(A, B)$ is capable of capturing all conditional probabilities within any Boolean lattice $\mathscr{B}(a_1, a_2, \ldots, a_n)$ embedded in $P(X)$, it is more general and fundamental than the concept of conditional probability: $S(A, B)$ is applicable to any pair of fuzzy sets in $\tilde{P}(X)$ while conditional probabilities are meaningful only within specific subsets of $\tilde{P}(X)$, each of which is a Boolean lattice $\mathscr{B}(a_1, a_2, \ldots, a_n)$, $a_1 + a_2 + \cdots + a_n = 1$. Moreover,

$$S(A, B) + S(A, \bar{B}) \geqq 1 \tag{40}$$

in general, with equality obtained within any Boolean lattice $\mathscr{B}(a_1, a_2, \ldots, a_n)$.

9. Numerous arguments have been or can be raised against the necessity and adequacy of the various axioms of probability theory. Let me mention just a few of them. One argument questioning the adequacy of the additivity axiom was presented by Viertl.[49] It is based on the fact that measurements are often inherently fuzzy due to unavoidable measurement errors. Consider, for example, two disjoint events, A and B, defined in terms of adjoining intervals of real numbers, as shown in Figure 5a. Observations in close neighborhoods (within a measurement error) of the endpoint of each event are unreliable and should be properly discounted, for example, according to the discount rate functions shown in Figure 5a. That is, observations in the neighborhood of the endpoints should carry less evidence than those outside them. When measurements are taken for the union of the two events, as shown in Figure 5b, one of the discount rate functions is not applicable. Hence, the same observations produce more evidence for the single event $A \cup B$ than for the two disjoint events A and B. This implies that the degree of belief in $A \cup B$ (probability of $A \cup B$) should be greater than the sum of the degrees of belief in A and B (probabilities of A and B). The additivity axiom is thus violated.

Some of the axioms that Cox employed for proving his theorem,[10] which are explicitly stated by Horwitz *et al.*[21] and also by Cheeseman,[5] can be questioned as well. For example, it is not obvious why, according to the complementarity axiom, belief in the complement of a proposition should be "a monotonically decreasing function of the belief in the proposition itself".[5,21] This is too strong to capture realistic situations characterized by the lack of belief in both a proposition and its negation. The axioms of completeness and scalar continuity[5,21] are also questionable. They require that a specific real number representing a degree of belief must be assigned to each proposition that is meaningful in a given context. This requirement rules out imprecision of any kind in expressing beliefs, be it imprecision in the form of vagueness or nonspecificity. It does not allow us, for example, to express our beliefs in terms of intervals of real numbers and deal with them by interval analysis,[39] or to express them by fuzzy numbers and manipulate them by fuzzy arithmetic.[13,24] It is questionable whether such a strict and rigid requirement of precision is realistic and desirable in modelling complex humanistic phenomena such as national economies, where human psychology, reasoning, and communication play essential roles.

Cox's proof is also based on one hidden axiom, which is usually not listed explicitly: in solving his functional equations, he assumes that both of the unknown functions have continuous second-order derivatives. It is likely that alternative solutions of the functional equations might be obtained if this assumption is withdrawn. For example, minimum and maximum functions do not satisfy this hidden axiom. This question, for which we have no answer at this time, is a difficult one, but certainly worthy of investigation.

6. EXAMPLES

Several very simple examples involving uncertainty are discussed in this section to illustrate the four types of uncertainty introduced in Section 4. In each case, I use a mathematical theory that I consider most appropriate for dealing with the situation. Probabilistic conceptualizations of the examples are left to my opponent, Peter Cheeseman.

1. Bogler[2] describes the following example of multiple sensor target identification in which intelligence reports are also employed as a source of information. It is assumed, based on an intelligence report, that there are 100 possible target types. Let $X=\{x_1, x_2, \ldots, x_{100}\}$ denote the set of these target types. It is also known, from another intelligence report, that only target type x_1 entered the relevant tactical area, but the reporting agent had access only to records pertaining to 40% of the targets entering the tactical area. What is the probability that an aircraft drawn at random is any particular target type?

The formulation of this problem in the Dempster–Shafer theory is very simple and noncontroversial. Let $A=\{x_1\}$. Then, using the given (initial) evidence, e_1, we have $m_1(A)=0.4$ and $m_1(X)=0.6$. By Eqs. (1) and (2), we readily obtain:

$$\mathrm{Bel}_1(A)=0.4, \quad \mathrm{Pl}_1(A)=1;$$

$$\mathrm{Bel}_1(\bar{A})=0, \quad \mathrm{Pl}_1(\bar{A})=0.6.$$

Interpreting the belief and plausibility functions as lower and upper probabilities, respectively, we obtain the following nonspecific determination of the relevant probabilities: $p_1(A)\in[0.4, 1]$, $p_1(\bar{A})\in[0, 0.6]$, $p_1(X)=1$. Consider now a later report from a sensor (evidence e_2) indicating that not only targets of type x_1 might be in the population of incoming targets, but also ten other target types, say $x_2, x_3, \ldots, x_{11}$. Let $B = \{x_1, x_2, \ldots, x_{11}\}$. Then, $m_2(B) = 1$, $\mathrm{Bel}_2(B) = \mathrm{Pl}_2(B) = 1$, $\mathrm{Bel}(\bar{B}) = \mathrm{Pl}(\bar{B}) = 0$. Hence, $p_2(B) = 1$ and $p_2(\bar{B}) = 0$. Using now the Dempster rule for combining evidence e_1 and e_2, we obtain:

	m_{12}	Bel_{12}	Pl_{12}	p_{12}
A	0.4	0.4	1	[0.4, 1]
$\bar{A}$	0	0	0.6	[0, 0.6]
B	0.6	1	1	1
$\bar{B}$	0	0	0	0
X	0	1	1	1

Calculating amounts of the three types of uncertainty in the three bodies of evidence, we obtain:

	$V(m)$	$E(m)$	$C(m)$
m_1	4	0	0.5
m_2	3.5	0	0
m_{12}	2.1	0	0.5

We can see that by combining evidence e_1 with evidence e_2, nonspecificity is substantially reduced. Observe that the three bodies of evidence are nested in this example ($E(m) = 0$ for each of them). Hence, we can also use the possibilistic rule of combining evidence (minimum of possibility distributions) to obtain exactly the same result.

2. Let me discuss another example in the area of multiple sensor target identification described by Bogler.[2] Assume that the universal set is again the set of 100 possible target types, but only two of the target types are involved in this example, a fighter and a bomber, which I denote by f and b, respectively. Evidence came in this case from two sensors. A short range sensor provides a support of 0.6 that the target is a fighter, while the radar warning receiver gives a support of 0.95 that the detected target is a bomber. What degree of support does the combined evidence provide? Using again the Dempster–Shafer theory, we obtain:

	m_1	m_2	m_{12}	Bel_{12}	Pl_{12}	p_{12}
$\{f\}$	0.6	0	0.07	0.07	0.12	[0.07, 0.12]
$\{b\}$	0	0.95	0.88	0.88	0.93	[0.88, 0.93]
X	0.4	0.05	0.05	1	1	1

The three types of uncertainties have the values:

	$V(m)$	$E(m)$	$C(m)$
m_1	2.7	0	0.4
m_2	0.3	0	7
m_{12}	0.3	0.3	0.4

We can see that by combining the two conflict-free bodies of evidence ($E(m_1)=E(m_2)=0$), which conflict with each other, we obtain a conflicting body of evidence ($E(m_{12})=0.3$). By combining bodies of evidence based upon the two sensors, nonspecificity of the first body of evidence and confusion (potential conflict) of the second body of evidence are substantially reduced at the cost of a small amount of dissonance.

3. This example has the same context as the two previous examples, but involves a different problem. Since the purpose is to illustrate a general principle, the example is discussed first in general terms and, then, some numerical instances of the results are examined. Let X denote a set of possible target types and let A and B denote particular subsets of X that can be identified by two sensors in the incoming population of targets with some degrees of belief, say a and b, respectively. Assume that $A \cap B \neq \phi$. Knowing that a and b are the total degrees of beliefs in A and B, respectively, what degree of belief should be allocated to the set $A \cap B$? Conceptualizing the problem again in terms of the Dempster–Shafer theory, I propose to deal with it in terms of a relevant principle of maximum uncertainty. The principle may be employed as an optimization problem with either three objective functions (V, E, and C, given by Eqs. (19), (20), and (21), respectively) or with a single aggregate of these objective functions (e.g., their sum S, given by Eq. (28)). Regardless of the optimization alternative chosen, the constraints of the optimization problem are expressed in this example by the equations and inequalities

$$\begin{aligned} &m(X)+m(A)+m(B)+m(A\cap B) &&= 1,\\ &m(A)+m(A\cap B) &&= a,\\ &m(B)+m(A\cap B) &&= b,\\ &m(X), m(A), m(B), m(A\cap B) &&\geqq 0, \end{aligned}$$

where $a, b \in [0, 1]$ are given numbers (total beliefs focusing on A, B, respectively). The equations are consistent, independent and underdetermined, with one degree of freedom. Selecting, for example, $m(A\cap B)$ as the free variable, we readily obtain

$$\begin{aligned} m(A) &= a - m(A\cap B),\\ m(B) &= b - m(A\cap B),\\ m(X) &= 1-a-b+m(A\cap B). \end{aligned} \tag{41}$$

Since all the unknowns must be nonnegative, the first two equations set the upper bound of $m(A\cap B)$, whereas the third equation specifies its lower bound; we obtain

$$\max(0, a+b-1) \leqq m(A \cap B) \leqq \min(a, b). \tag{42}$$

Let $R = [\max(0, a+b-1), \min(a, b)]$ denote this range of values of $m(A \cap B)$ that satisfy the given constraints.

Using Eqs. (41), the objective functions can be expressed solely in terms of the free variable $m(A \cap B)$. After a simple rearrangement of terms, we obtain

$$V(m) = m(A \cap B)\log_2(K_1) + K_2,$$

$$E(m) = 0,$$

$$C(m) = -m(A \cap B)\log_2 m(A \cap B) + m(A \cap B)\log_2(a \cdot b) + K_3,$$

where

$$K_1 = \frac{|X| \cdot |A \cap B|}{|A| \cdot |B|},$$

$$K_2 = (1-a-b)\log_2|X| + a\log_2|A| + b\log_2|B|,$$

$$K_3 = -a\log_2 a - b\log_2 b.$$

Let M_V, M_E, and M_C denote the values or sets of values of $m(A \cap B)$ for which functions V, E, and C, respectively, reach their maxima. Then, we can determine by simple considerations that

$$M_V = \begin{cases} \max(0, a+b-1) & \text{when} \quad K_1 < 1 \\ \min(a+b) & \text{when} \quad K_1 > 1 \\ R & \text{when} \quad K_1 = 1, \end{cases}$$

$$M_E = R,$$

$$M_C = \begin{cases} \max(0, a+b-1) & \text{when} \quad ab/e < \max(0, a+b-1) \\ \min(a, b) & \text{when} \quad ab/e > \min(a, b) \\ ab/e & \text{when} \quad ab/e \in R, \end{cases}$$

where e is the base of natural logarithms ($e \doteq 2.7$).

Let R_a denote the set of admissible (nondominated or noninferior) solutions of our optimization problem. There are nine possible combinations of M_V and M_C, each of which determines R_a. These combinations are specified in Table 1. We can

Table 1 Admissible solutions in the example discussed

R_a	$K_1 < 1$	$K_1 > 1$	$K_1 = 1$
$ab/e < \max(a+b-1)$	$\max(0, a+b-1)$	R	$\max(0, a+b-1)$
$ab/e > \min(a, b)$	R	$\min(a, b)$	$\min(a, b)$
$ab/e \in R$	$[\max(0, a+b-1), ab/e]$	$[ab/e, \min(a, b)]$	ab/e

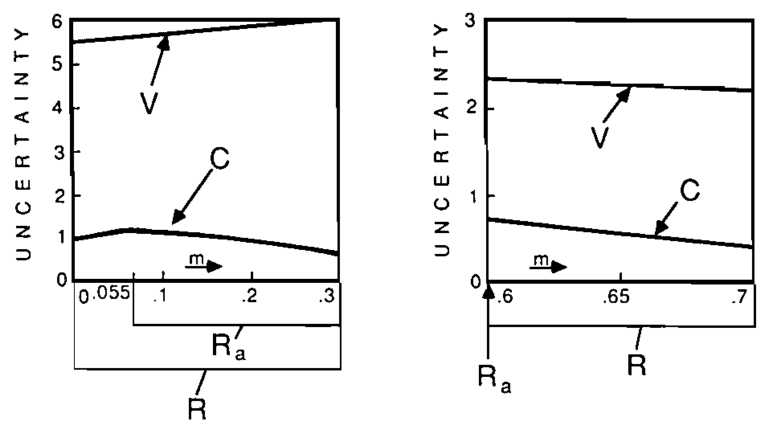


Figure 6 Example of the use of the maximum uncertainty principle in the Dempster–Shafer theory.

see that there are actually only six different types of R_a, three of which are represented by unique numbers and the remaining three by intervals of real numbers.

When employing the total uncertainty $S(m)$, given by Eq. (28), as the objective function in our example, we obtain

$$S(m) = m(A \cap B)[\log_2 K_1 - \log_2 m(A \cap B) + \log_2 ab] + K_2 + K_3$$

We can readily determine the maximum, M_S, of this objective function within the restricted domain R:

$$M_S = \begin{cases} \max(0, a+b-1) & \text{when} \quad K_1 ab/e < \max(0, a+b-1) \\ \min(a,b) & \text{when} \quad K_1 ab/e > \min(a,b) \\ K_1 ab/e & \text{when} \quad K_1 ab/e \in R. \end{cases}$$

Let us examine now three numerical instances of this example.

i) Let $|X| = 150$, $|A| = 40$, $|B| = 50$, $|A \cap B| = 20$, $a = 0.5$, and $b = 0.3$. Then, $R = [0, 0.3]$, $K_1 = 1.5$, $ab/e = 0.055$, $M_V = 0.3$, $R_a = [0.055, 0.3]$, $M_S = 0.083$ (Figure 6a). Hence, the solution is interval-valued when we view the problem as a multiple-objective criteria optimization problem: $m(A \cap B) \in [0.055, 0.3]$. Depending on the selected value of $m(A \cap B)$, using possibly some additional objective criteria, the values of $m(A)$, $m(B)$, and $m(X)$ are determined by Eqs. (41). When we use the total uncertainty $S(m)$ as the only objective function, we obtain

$$
\begin{array}{llll}
m(A \cap B) = 0.083, & \mathrm{Bel}(A \cap B) = 0.083, & \mathrm{Pl}(A \cap B) = 1, \\
m(A) \quad = 0.417, & \mathrm{Bel}(A) \quad = 0.5, & \mathrm{Pl}(A) \quad = 1, \\
m(B) \quad = 0.217, & \mathrm{Bel}(B) \quad = 0.3, & \mathrm{Pl}(B) \quad = 1, \\
m(X) \quad = 0.283, & \mathrm{Bel}(X) \quad = 1, & \mathrm{Pl}(X) \quad = 1.
\end{array}
$$

ii) Let $|X| = 20$, $|A| = 8$, $|B| = 9$, $|A \cap B| = 3$, $a = 0.7$, and $b = 0.9$. Then, $R = [0.6, 0.7]$, $K_1 = 0.833$, $ab/e = 0.232$, $M_V = 0.6$, $R_a = 0.6$, $M_S = 0.6$ (Figure 6b). Hence, the solution is unique, $m(A \cap B) = 0.6$, independent of the approach employed. We obtain,

$$
\begin{array}{llll}
m(A \cap B) = 0.6, & \mathrm{Bel}(A \cap B) = 0.6, & \mathrm{Pl}(A \cap B) = 1, \\
m(A) \quad = 0.1, & \mathrm{Bel}(A) \quad = 0.7, & \mathrm{Pl}(A) \quad = 1, \\
m(B) \quad = 0.3, & \mathrm{Bel}(B) \quad = 0.9, & \mathrm{Pl}(B) \quad = 1, \\
m(X) \quad = 0, & \mathrm{Bel}(X) \quad = 1, & \mathrm{Pl}(X) \quad = 1.
\end{array}
$$

iii) Let $|X| = 80$, $|A| = 40$, $|B| = 50$, $|A \cap B| = 20$, $a = 0.5$, and $b = 0.6$. Then, $R = [0.1, 0.5]$, $K_1 = 0.8$, $ab/e = 0.11$, $M_V = 0.1$, $R_a = [0.1, 0.11]$, $M_S = 0.1$. Hence, $m(A \cap B) \in [0.1, 0.11]$ for the multiple-objective criteria approach and $m(A \cap B) = 0.1$ for the single-objective criteria approach.

4. Consider two variables v_1, v_2, each of which has two possible states, say 0 and 1. For convenience, let the joint states of the variables be labelled by an index i in the following way:

v_1	v_2	i
0	0	0
0	1	1
1	0	2
1	1	3

Assume that we have a record of 1,000 observations of the variables; some of the observations contain values of both variables, some of them contain a value of only one variable due to some measurement or communication constraints (not essential for our discussion). Observing a value of one variable only may be interpreted in the Dempster–Shafer theory as observing a set of two joint states. For example, observing that $v_1 = 1$ (and not knowing the state of v_2) may be viewed as observing the subset $\{2, 3\}$ of the four joint states. Numbers of observations N of the eight relevant sets of states (defined by their characteristic functions) are given in Table 2, which also contains values of the estimated basic assignment m, based on frequency interpretation, as well as the corresponding degrees of belief Bel and plausibility Pl.

Belief and plausibility degrees can readily be calculated for any of the eight remaining subsets of states. For example, $\mathrm{Bel}(\{1, 2, 3\}) = 0.721$ and

Table 2 Illustration to Example 4

i: 0	*1*	*2*	*3*	*N*(*A*)	*m*(*A*)	*Bel*(*A*)	*Pl*(*A*)
A: 1	1	0	0	212	0.212	0.373	0.839
0	0	1	1	128	0.128	0.161	0.627
1	0	1	0	315	0.315	0.446	0.786
0	1	0	1	151	0.151	0.214	0.554
1	0	0	0	106	0.106	0.106	0.633
0	1	0	0	55	0.055	0.055	0.418
0	0	1	0	25	0.025	0.025	0.468
0	0	0	1	8	0.008	0.008	0.287

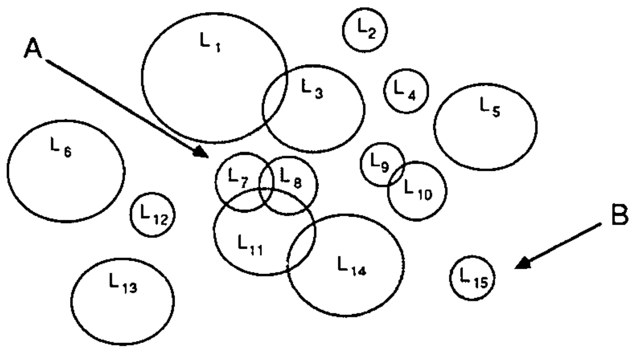


Figure 7 Estimates of 15 experts regarding the location of the epicenter of an earthquake.

$\mathrm{Pl}=(\{1,2,3\})=0.992$. We can also calculate how the total amount of uncertainty of this body of evidence is distributed among the three types of uncertainty: $V(m)=0.806$, $E(m)=0.559$, $C(m)=2.104$. A comparison of these results with a probabilistic treatment of this example should be illuminating.

5. Dong and Wong[11] describe an example in which a group of experts give their estimates of possible location of the epicenter of an earthquake. Suppose that 15 estimates are given as shown in Figure 7. Observe that these estimates are both nonspecific and conflicting with each other. Using the evidence on hand, what is the likelihood that the epicenter is inside of any particular area of interest (e.g., densely populated areas A and B)? Each of the estimates has a weight of evidence 1/15, provided that we consider all reports as equally reliable and otherwise equivalent in their value. Then, degrees of belief and plausibility can readily be calculated: $\mathrm{Bel}(A)=2/15=0.13$, $\mathrm{Pl}(A)=5/15=0.33$; $\mathrm{Bel}(B)=1/15=0.07$, $\mathrm{Pl}(B)=3/15=0.2$. Hence, we obtain the following interval-valued estimates of probabilities $p(A)$ and $p(B)$ that the epicenter is in area A or B, respectively: $p(A)\in[0.13,0.33]$, $p(B)\in[0.07,0.2]$.

6. A convenient and computationally efficient way of solving integer optimization problems with variables $v_k(k=1,2,\ldots,n)$ is to replace the constraints "v_k is an integer" with less specific and fuzzy constraints "v_k is almost an integer". The latter constraints are fuzzy sets (fuzzy numbers[13,24]). For each variable, these fuzzy sets can be expressed, for example, by the membership grade functions

$$\mu_{A_i}(x)=\frac{1}{1+c(x-i)^2},$$

where i is a nonnegative integer ($i \in I$), x denotes values of the variable (usually nonnegative real numbers), and c is some positive constant (usually $c \geqq 1$), which can be properly adjusted to achieve a good performance. The total constraint is the fuzzy set of all approximate integers, say set A, which is obtained by taking union of sets A_i for all $i \in I$:

$$\mu_A(x)=\max_{i \in I} \frac{1}{1+c(x-i)^2}=\frac{1}{1+c[x-\mathrm{tr}(x+0.5)]^2},$$

where tr denotes the truncation function (the largest integer not greater than x). By fuzzifying the constraints, an integer optimization problem becomes an ordinary optimization problem and, according to our experience (with $c=10$), computational efficiency increases.

The area of fuzzy optimization and decision making (as illustrated by this example) is now well established and very successful.[57] The use of fuzzy constraints is motivated not only by the desire to reduce computational complexity, but also by the desire to allow the decision maker to express his problem (if preferred) in fuzzy terms of natural language.

7. There are many applications described in the literature (and some successfully implemented) in which it is essential or, at least, advantageous to use descriptions in natural language (with all its imprecisions) and common sense reasoning. In medicine, for example, hepatitis can be well described by the inherently fuzzy and nonspecific proposition "total proteins are usually normal, albumin is decreased, α-globulins are slightly decreased, β-globulins are slightly decreased, and γ globulins are increased", but any attempt to reduce the nonspecificity or fuzziness of this description is unwarranted. In this way, the medical knowledge can be expressed in terms of a fuzzy relation M by which diseases in set A are related to symptoms in set B. Then, given M and a fuzzy set S of symptoms observed in a patient, the fuzzy set D of possible diseases can be inferred by means of the compositional rule of inference

$$D=S \circ M,$$

where $\circ$ denotes the max/min composition, i.e.,

$$\mu_D(d)=\max_{s \in S}\,[\min(\mu_S(s), \mu_S(s,d))]$$

for each $d \in D$. Further details regarding this application, including the process of compiling medical knowledge (constructing the fuzzy relation M), are discussed in my book.[29]

Linguistic descriptions have already proved very successful in the design of control systems.[47] Here, we deal with a set of fuzzy rules, such as "if the temperature is very high and the pressure is decreasing rapidly, then reduce the temperature significantly", which can often be elicited from experienced human operators.

It is not clear to me how probability theory could effectively describe and manipulate the great variety of descriptions or rules that are possible in natural language.

7. CONCLUSIONS

The principal conclusion emerging from my arguments should be that probability theory is capable of conceptualizing only one type of uncertainty. This is by design: axioms of probability theory do not allow any imprecision in characterizing situations under uncertainty, be it imprecision in the form of nonspecificity or vagueness. The only type of uncertainty that remains is conflict. Probability theory conceptualizes uncertainty strictly in terms of conflict among degrees of belief allocated to mutually exclusive alternatives; it can be well characterized as a conflict-oriented mathematical theory of uncertainty. This restriction of probability theory to one type of uncertainty—conflict—is a virtue in some application contexts and a vice in others.

Few would disagree with the following remarkable principles of wisdom, expressed by the ancient Chinese philosopher Lao Tsu:[36]

> Knowing ignorance is strength.
> Ignoring knowledge is sickness.

At first sight, it might seem that these principles are perfectly operationalized by the principles of maximum and minimum entropy. This, unfortunately, is not the case, not because of any defect in the maximum and minimum entropy principles, but because of the principal limitations of probability theory. Due to the nature of probability theory, ignorance in the form of nonspecificity or vagueness cannot be recognized. This violates the first principle of Lao Tsu: we should fully recognize our ignorance. Moreover, probability theory is not capable of conceptualizing knowledge that involves nonspecificity or vagueness and, consequently, such knowledge must be ignored. This violates the second principle of Lao Tsu: utilize all knowledge available, regardless of its form. In order to fully operationalize the Lao Tsu principles of wisdom, we need much broader principles of maximum and minimum uncertainty, principles that involve all types of uncertainty.

The relationship among uncertainty, complexity and credibility of systems models, which is of utmost importance to systems modelling, is not well understood as yet. We only know that uncertainty is a valuable commodity, which can be traded for a reduction of complexity or an increase of credibility of models in the modelling business. Since well-justified measures of the various types of uncertainty are now available for several mathematical frameworks in which uncertainty can be conceptualized (Section 4), this trading can be made operational at a scale previously unsuspected. It is undeniable that major research must yet be undertaken not only to develop sound multidimensional principles of uncertainty in the novel mathematical theories (Section 3), but also to learn how to use these theories in various application areas. After all, probability theory has been with us for over three centuries,[17] while the new theories for conceptualizing uncertainty are a phenomenon of less than three decades.

A turning point in our understanding of the concept of uncertainty was reached when it became clear that more than one type of uncertainty must be recognized within the Dempster–Shafer theory, and even within the restricted domain of possibility theory. This new insight into the concept of uncertainty was obtained by examining uncertainty within mathematical frameworks more general than the two classical theories employed for characterizing uncertainty (classical set theory and probability theory).

The emergence of generalizations of existing mathematical theories is a significant current trend in mathematics, as exemplified by the change in emphasis from quantitative to qualitative, from functions to relations, from graphs to hypergraphs, from ordinary geometry (Euclidean as well as non-Euclidean) to fractal geometry, from ordinary automata to dynamic cellular automata, from classical analysis to a study of singularities (catastrophe theory), from ordinary artificial languages to developmental languages, from precise analysis to interval analysis, from classical logic to logic of inconsistency, from two-valued logic to multiple-valued logics, from single objective to multiple objective criteria optimization, and, as most relevant to the subject of this debate, from probability measures to fuzzy measures and from classical set theory to fuzzy set theory. These generalizations, stimulated primarily by advances in computer technology and modern systems thinking, have enriched not only our insights but also our capabilities for modelling the intricacies of the real world.

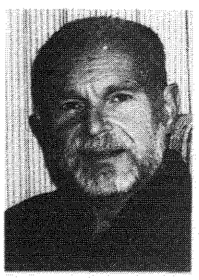
The issue of our next debate should not be whether probability theory and the other theories are right or wrong. It should rather be the question of which of the theories are relevant and appropriate for conceptualizing and dealing with uncertainty in each particular context. Furthermore, probability theory and the other theories should not be viewed as necessarily conflicting with each other. Some may usefully complement each other (as probability and possibility theories do), others may supplement and reinforce each other (as fuzzy events imported into probability theory do). Therefore, let us cooperate rather than compete.

REFERENCES

1. G. Banon, "Distinction between several subsets of fuzzy measures." *Fuzzy Sets and Systems*, **5**, No. 3, 1981, pp. 291–305.
2. P. Bogler, "Shafer–Dempster reasoning with applications to multisensor target identification systems." *IEEE Trans. on Systems, Man. and Cybernetics*, **SMC-17**, No. 6, 1987, pp. 968–977.
3. P. Cheeseman, "In defense of probability." *Proc. 9th Intern. Joint Conf. on Artificial Intelligence*, Los Angeles, 1985, pp. 1002–1009.
4. P. Cheeseman, "Probabilistic versus fuzzy reasoning." In: *Uncertainty in Artificial Intelligence*, edited by L. N. Kanal and J. F. Lemmer, North-Holland, Amsterdam and New York, 1986, pp. 85–102.
5. P. Cheeseman, "An inquiry into computer understanding." *Computational Intelligence*, **4**, No. 1, 1988, pp. 58–66.
6. P. Cheeseman, "In defense of an inquiry into computer understanding." *Computational Intelligence*, **4**, No. 1, 1988, pp. 129–142.
7. R. Christensen, *Entropy Minimax Sourcebook* (4 volumes). Entropy Limited, Lincoln, Mass., 1980–81.
8. R. Christensen, *Multivariate Statistical Modeling*. Entropy Limited, Lincoln, Mass., 1983.
9. R. Christensen, "Entropy minimax multivariate statistical modeling." *Intern. J. of General Systems*, **11**, No. 3, 1985, pp. 231–277 (I. Theory); **12**, No. 3, 1986, pp. 227–305 (II. Applications).
10. R. T. Cox, "Probability, frequency and reasonable expectation." *American Journal of Physics*, **14**, No. 1, 1946, pp. 1–13.
11. W. Dong and F. S. Wong, "From uncertainty to approximate reasoning: Part 1: Conceptual models and engineering interpretations. *Civil Eng. Systems*, **3**, 1986, pp. 143–202.

12. D. Dubois and H. Prade, "Properties of measures of information in evidence and possibility theories." *Fuzzy Sets and Systems*, **25**, No. 2, 1987, pp. 161–182.
13. D. Dubois and H. Prade, "Fuzzy numbers: An overview." In: *Analysis of Fuzzy Information*, edited by J. C. Bezdek, CRC Press, Boca Raton, FL., 1987, pp. 3–39.
14. D. Dubois and H. Prade, *Possibility Theory*. Plenum Press, New York, 1988.
15. T. L. Fine, *Theories of Probability: An Examination of Foundations*. Academic Press, New York, 1973.
16. J. A. Goguen, "L-fuzzy sets." *J. of Math. Analysis and Applications*, **18**, 1967, pp. 145–174.
17. I. Hacking, *The Emergence of Probability*. Cambridge University Press, Cambridge, 1975.
18. R. V. L. Hartley, "Transmission of information." *The Bell Systems Technical J.*, **7**, 1928, pp. 535–563.
19. M. Higashi and G. J. Klir, "On measures of fuzziness and fuzzy complements." *Intern. J. of General Systems*, **8**, No. 3, 1982, pp. 169–180.
20. M. Higashi and G. J. Klir, "Measures of uncertainty and information based on possibility distributions." *Intern. J. of General Systems*, **9**, No. 1, 1983, pp. 43–58.
21. E. J. Horvitz, D. E. Heckerman and C. P. Langlotz, "A framework for comparing alternative formalisms for plausible reasoning." *Proc. Fifth National Conf. on Artificial Intelligence*, Philadelphia, 1986, pp. 210–214.
22. E. T. Jaynes, "Bayesian methods: General background." In: *Maximum Entropy and Bayesian Methods in Applied Statistics*, edited by J. H. Justice, Cambridge University Press, Cambridge, 1986, pp. 1–25.
23. E. T. Jaynes, *Papers on Probability, Statistics and Statistical Physics* (edited by R. D. Rosenkrantz). D. Reidel, Boston, 1983.
24. A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic: Theory and Applications*. Van Nostrand Reinhold, New York, 1985.
25. W. J. M. Kickert, *Fuzzy Theories on Decision-Making*. Martinus-Nijhoff, Boston, 1978.
26. G. J. Klir, "Pattern discovery in activity arrays." In: *Systems: Approaches, Theories, Applications*, edited by W. E. Hartnett. D. Reidel, Dordrecht, Holland, 1977, pp. 121–158.
27. G. J. Klir, *Architecture of Systems Problem Solving*. Plenum Press, New York, 1985.
28. G. J. Klir, "Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like?" *Fuzzy Sets and Systems*, **24**, No. 2, 1987, pp. 141–160.
29. G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*. Prentice-Hall, Englewood Cliffs, NJ, 1988.
30. G. J. Klir and M. Mariano, "On the uniqueness of possibilistic measure of uncertainty and information." *Fuzzy Sets and Systems*, **24**, No. 2, 1987, pp. 197–219.
31. G. J. Klir and E. C. Way, "Reconstructability analysis: Aims, results, open problems." *Systems Research*, **2**, 1985, pp. 141–163.
32. A. N. Kolmogorov, *Foundations of the Theory of Probability*. Chelsea, New York, 1950.
33. A. N. Kolmogorov, "Three approaches to the quantitative definition of information." *Problems of Information Transmission*, **1**, No. 1, 1965, pp. 1–7.
34. B. Kosko, "Fuzzy entropy and conditioning." *Information Sciences*, **40**, No. 1, 1987, pp. 1–10.
35. B. Kosko, "Fuzziness vs. probability." *Unpublished manuscript*, E.E. Department, University of Southern California, 1987.
36. Lao Tsu, *Tao Te Ching*. Vintage Books, New York (Sec. 71), 1972.
37. D. V. Lindley, "Scoring rules and the inevitability of probability." *Intern. Statistical Review*, **50**, 1982, pp. 1–26.
38. D. V. Lindley, "The probability approach to the treatment of uncertainty in artificial intelligence and expert systems." *Statistical Science*, **2**, No. 1, 1987, pp. 17–24.
39. R. E. Moore, *Methods and Applications of Interval Analysis*. SIAM, Philadelphia, 1979.
40. A. Ramer, "Uniqueness of information measure in the theory of evidence." *Fuzzy Sets and Systems*, **24**, No. 2, 1981, pp. 183–196.
41. A. Rényi, *Probability Theory*. North-Holland, Amsterdam. (Chapter IX, Introduction to Information Theory, 540–616), 1970.
42. G. L. S. Shackle, *Decision, Order and Time in Human Affairs*. Cambridge University Press, Cambridge, 1961.
43. G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
44. C. E. Shannon, "The mathematical theory of communication." *The Bell System Technical Journal*, **27**, 1948, pp. 379–423, 623–656.
45. J. E. Shore and R. W. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy." *IEEE Trans. on Information Theory*, **26**, 1980, pp. 26–37.

46. M. Sugeno, "Fuzzy measures and fuzzy integrals: a survey." In: *Fuzzy Automata and Decision Processes*, edited by M. M. Gupta, G. N. Saridis and B. R. Gaines, North-Holland, Amsterdam and New York, 1977, pp. 89–102.
47. M. Sugeno, (editor), *Industrial Applications of Fuzzy Control*. North-Holland, New York, 1985.
48. M. Tribus, *Rational Descriptions, Decisions, and Designs*. Pergamon Press, Oxford, 1969.
49. R. Viertl, "Is it necessary to develop a fuzzy Bayesian inference?" In: *Probabilty and Bayesian Statistics*, edited by R. Viertl, Plenum Press, New York, 1987, pp. 471–475.
50. P. Walley and T. L. Fine, "Varieties of modal (classificatory) and comparative probability." *Synthese*, **41**, 1979, pp. 321–374.
51. P. M. Williams, "Bayesian conditionalisation and the principle of minimum information." *British J. for the Philosophy of Science*, **31**, No. 2, 1980, pp. 131–144.
52. R. R. Yager, "On the measure of fuzziness and negation. Part I: Membership in the unit interval." *Intern. J. of General Systems*, **5**, No. 4, 1979, pp. 221–229.
53. R. R. Yager, "Toward general theory of reasoning with uncertainty: nonspecificity and fuzziness." *Intern. J. of Intelligent Systems*, **1**, No. 1, 1986, pp. 45–67.
54. L. A. Zadeh, "Fuzzy sets." *Information and Control*, **8**, No. 3, 1965, pp. 338–353.
55. L. A. Zadeh, "Probability measures of fuzzy events." *J. Math. Analysis and Applications*, **23**, 1968, pp. 421–427.
56. L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility." *Fuzzy Sets and Systems*, **1**, No. 1, 1978, pp. 3–28.
57. H. J. Zimmermann, *Fuzzy Sets, Decision Making and Expert Systems*. Kluwer, Boston, 1987.



George J. Klir is a Distinguished Professor of Systems Science and Chairman of the Department of Systems Science, Thomas J. Watson School of Engineering, Applied Science and Technology, State University of New York at Binghamton. He received the M.S. degree in electrical engineering from the Prague Institute of Technology in 1957, and the Ph.D. degree in Computer Science from the Czechoslovak Academy of Sciences in 1964. He is also a graduate of the IBM Systems Research Institute in New York. Before joining the State University of New York, Dr. Klir had been with the Computer Research Institute and Charles University in Prague, the University of Baghdad, the University of California at Los Angeles and Fairleigh Dickinson University in New Jersey; he has also worked part-time for IBM and Bell Laboratories, and taught summer courses at the University of Colorado, Portland State University in Oregon and Rutgers University in New Jersey. During the academic years 1975–76 and 1982–83, he was a Fellow at the Netherlands Institute for Advanced Studies in Wassenaar, Holland, and in 1980 he was a fellow of the Japan Society for the Promotion of Science. Dr. Klir's main research activities have been in general systems methodology, logic design and computer architecture, switching and automata theory, information theory, knowledge-based systems, fuzzy set theory, and the philosophy of science. He is the author of over two hundred articles published in various professional journals and holds a number of patents. He is an author or editor of fourteen books, among them *Cybernetic Modelling* (Iliffe, London, 1967), *An Approach to General Systems Theory* (Van Nostrand Reinhold, New York, 1969), *Trends in General Systems Theory* (John Wiley, New York, 1972), *Methodology of Switching Circuits* (Van Nostrand, New York, 1972), *Architecture of Systems Problem Solving* (Plenum Press, New York, 1985) and *Fuzzy Sets, Uncertainty, and Information* (Prentice-Hall, Englewood Cliffs, 1988). Dr. Klir has been Editor-in-Chief of the *International Journal of General Systems* since 1974. He was President of the Society for General Systems Research in 1981–82 and President of the International Federation for Systems Research (IFSR) in 1980–84. He is currently Editor of the *IFSR Book Series on Systems Science and Engineering*, President of the North American Fuzzy Information Processing Society (NAFIPS), and a member of Editorial Boards of 13 journals. He is a senior member of IEEE and a member of ISSS, AAAS, ACM, and PSA.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER International Journal of Approximate Reasoning 37 (2004) 71–85
INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

The philosophical significance of Cox's theorem

Mark Colyvan *

Department of Philosophy, University of Queensland, Brisbane, Qld. 4072, Australia

Received 1 June 2002; received in revised form 1 July 2003; accepted 1 November 2003
Available online 31 December 2003

Abstract

Cox's theorem states that, under certain assumptions, any measure of belief is isomorphic to a probability measure. This theorem, although intended as a justification of the subjectivist interpretation of probability theory, is sometimes presented as an argument for more controversial theses. Of particular interest is the thesis that the only coherent means of representing uncertainty is via the probability calculus. In this paper I examine the logical assumptions of Cox's theorem and I show how these impinge on the philosophical conclusions thought to be supported by the theorem. I show that the more controversial thesis is not supported by Cox's theorem.

Keywords: Cox; Belief; Uncertainty; Non-classical logic; Probability; Excluded middle

1. Introduction

Benacerraf [6] once warned that when philosophical conclusions are argued from formal mathematical results, one should look very carefully at the assumptions of the arguments in question. For any such argument cannot rest on the formal result alone; there must be some philosophical premise, and this is often illicitly smuggled through the back door. Benacerraf is not suggesting that one can never draw philosophical conclusions from formal results, or that

* Tel.: +61-7-3365-2613; fax: +61-7-3365-1968.
E-mail address: mcolyvan@uq.edu.au (M. Colyvan).

0888-613X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2003.11.001

all such arguments are flawed—just that it is important to identify the often suppressed philosophical premises and to assess their plausibility. I think this is very good advice and with this advice in mind I wish to examine the formal result known as Cox's theorem. This theorem states that, under the assumptions of the theorem, any measure of belief is isomorphic to a probability measure [9,10]. The theorem has been used to support a variety of philosophical conclusions, ranging from a justification of the Bayesian approach to probability, to a more radical thesis that probability is the only coherent representation of uncertainty. In particular, I will examine the logical underpinnings of the theorem—classical propositional calculus—and show that, in certain contexts at least, these logical assumptions are hard to defend. This, in turn, undermines the more radical philosophical theses that the theorem might be thought to support. I begin by discussing a kind of uncertainty for which classical logic is inappropriate.

2. Belief and non-epistemic uncertainty

Agents typically do not believe propositions to degree one or zero. Belief comes in degrees. This is because there is typically uncertainty about the truth value of the proposition in question. Good epistemic agent recognise this and set about quantifying the extent of the uncertainty and/or their degree of certainty. Providing the details of a representation of reasoning carried out by human (or more commonly, ideal) agents operating under uncertainty is often referred to as the project of delivering the logic of plausible inference. [1] It is usually assumed that uncertainty arises because of incomplete information—it is simply an epistemic matter. I will argue that this is not the case. Some uncertainty may remain even when the agent is in possession of all the relevant data. This is bad news for classical logic and classical probability theory.

There are two ways in which an agent can be uncertain about the state of a system. The first is familiar. This is where there is uncertainty about some underlying fact of the matter: System S is either in state σ or it is not, but agent A does not know which. A might be in possession of some probabilistic information about the state of S—either numerical ("the probability that S is in state σ is x") or non-numerical ("it is more likely that S is in state σ than not"). Call this *epistemic uncertainty*. Now compare this with a second, quite different kind of uncertainty; uncertainty where there is no fact of the matter about

[1] I share Shafer's [48] concerns about the use of the term 'plausible' here, but this term is well entrenched in the literature, and since I can think of no better term, I'll continue to use it. I stress however, that I am using the term more broadly than is usual. I include any formal account of belief and reasoning under uncertainty.

whether system S' is in state σ' or not. Indeed, here the uncertainty arises *because* there is no underlying fact of the matter. Call this second kind of uncertainty *non-epistemic uncertainty*. The idea here is that, for reasons I'll discuss shortly, the system S' is neither in state σ' nor not in state σ'—S' is not in a determinate state with respect to σ'. It follows that even an agent in possession of all the relevant data will be uncertain as to the truth value of the proposition 'S' is in state σ''. (Or, equivalently, the agent will not know the answer to the question 'Is S' in state σ'?'.)

It follows that if there are any instances of non-epistemic uncertainty, an agent could not be in possession of probabilistic information in such cases. After all, what would it mean to say that the probability that system S' is in state σ' is x when there is no fact of the matter about the state of S'? Classical probability theory presupposes that there is an underlying fact of the matter. To see this we need only consider one of the axioms of classical probability theory:

$$\Pr(Q \vee \neg Q) = 1.$$

This implies that the proposition $Q \vee \neg Q$ is certain (because it is a logical truth). This axiom of probability theory is the probabilistic analogue of the logical principle of excluded middle. It would thus seem that in any domain where excluded middle fails, (classical) probability theory is an inappropriate tool for representing uncertainty. [2]

Now there are several candidates for such domains, none of which, admittedly, are entirely uncontroversial. To start with, consider fictional discourse. In a work of fiction such as H.G. Wells' The Time Machine there is nothing more to the story than what is written (and perhaps the logical and natural implications of what is written). There is no fact of the matter about details not in the story. So, for example, in the 1960 movie of the novel the time traveller sets off for the future taking with him three books. What were the three books? Well that's (quite deliberately) not part of the story so (plausibly) there's no fact of the matter about what the three books were. It seems that classical logic—in particular excluded middle—fails here. It is not true that either the time traveller took or did not take Descartes' meditations with him. Moreover, the probability of this disjunction is not one (as standard probability theory insists). Indeed, it seems quite misguided to talk of probabilities at all in such cases. [3] I should add that this example is not as irrelevant to science as it might

[2] See [7,13,14] for more on this issue.

[3] Some might insist that the question about what the books were is meaningless, but this is very hard to sustain. There is nothing ungrammatical about the sentence and the meaning is perfectly clear. On what grounds is the case for the sentence's meaningless to be based? I can think of none. Indeed, the reason that some are inclined to call such questions meaningless is because they *do* understand the meaning, see what the implications are, and only then deny that it has meaning.

at first seem. Science makes wide use of fictional entities (like incompressible fluids, and Turing machines) and others that turn out to be fictional (like the planet Vulcan, which was supposed to have an orbit inside Mercury's). And it is clear that there are true propositions about such fictions—'the halting problem is unsolvable', for instance. So it would not do to dismiss fictional discourse as a mere philosophical curiosity. [4]

Another example of a domain in which excluded middle might be thought to fail is mathematics. Consider the status of Goldbach's conjecture: all even numbers greater than two can be written as the sum of two primes. At present this conjecture has not been proven nor has its negation been proven. Now let us suppose that you are a constructivist about mathematics. That is, you believe that 'P is true' is just to say that P has a constructively respectable proof from some constructively respectable set of axioms. It is well known that such constructivists embrace intuitionistic logic where both double negation elimination and excluded middle fail [26]. But even non-constructivists may accept that there are no-fact-of-the-matter propositions in mathematics. Take, for example, an independent question of set theory such as the continuum hypothesis. Neither this nor its negation is provable from the standard ZFC axioms—it is provably independent of those axioms. Many non-constructivists (for example, mathematical fictionalists like Field [12]) also believe that at least some independent statements are neither true nor false (and so it is not true that such an independent statement or its negation holds).

A third example of where excluded middle might be thought to fail is in domains where vague predicates are employed. Let us suppose that we wish to know how many young people there are in a crowd. We might be uncertain about this because there are some borderline cases. Take, for example, someone who is in their late 1920s. Do we count such a person as young or not? There seems no definitive way to answer this question. The problem is that the word 'young' is vague (in the sense that it permits borderline cases). [5] There are some well-known approaches to vagueness according to which excluded middle holds—for example, Williamson's [55] epistemic account of vagueness, the

[4] See [20, pp. 70–73], [21, Chapter 7] and [40, pp. 128–131] for more on the logic and semantics of non-denoting fictional terms. Fictional discourse also raises problems at the level of predicate logic. In classical predicate logic all names refer—even names like 'Vulcan'. Another deviation from classical logic motivated by such considerations is *free logic* where "empty" names are permitted (see [17,18,35,43]).

[5] It is also context sensitive. But let us put that aside; let us assume that the context is fixed. I should also mention that vagueness is rather widespread in both natural language and in science so it is unreasonable to dismiss it as another philosophical curiosity. See [44,45] for some of the problems arising from vagueness in ecology and conservation biology.

supervaluational account [15,51] and the paraconsistent approach [4,28,29]. Still, rejecting excluded middle remains a very plausible strategy. [6]

Indeed, those who would like to apply probability theory to domains with vague predicates should take little comfort from the above excluded-middle-preserving approaches. For example, on what is generally thought to be the leading contender among these approaches—the supervaluational account—probabilities are still out of place. Although $P \vee \neg P$ is a theorem, if P is borderline, P is usually thought to be neither true nor false. In either case, it seems to make little sense to speak of the probability of P being true (when P is borderline). On the paraconsistent approach, excluded middle is preserved at the expense of (one sense of) the law of non-contradiction. That is, borderline statements (such as 'a 28 year old is young') are seen as both true and false. That is, we have some true instances of $P \wedge \neg P$. Those who find giving up excluded middle objectionable are unlikely to be happy with this. Williamson's epistemic approach (according to which there is an unknowable fact of the matter concerning borderline cases) is the only option that would seem palatable to defenders of the view that probability theory is appropriate in such domains. The problem is that Williamson's view is deeply unintuitive and finds few supporters because of this. It would be inappropriate to try to settle the matter of the correct account of vagueness here; I simply mention vagueness as another very plausible source of non-epistemic uncertainty.

It is worth pausing for a moment to emphasise how vagueness gives rise to uncertainty. Consider a scientific question such as 'how many species are there in a given eco-system?' Obviously there will be epistemic uncertainty associated with this question but let us suppose that an agent is in possession of all the relevant data. It turns out that even in possession of all the data, the answer to the question may remain out of reach because of the vagueness of the scientific terms 'eco-system' and 'species'. The boundary of a eco-system will always admit borderline cases. Less obvious, perhaps, is that the term 'species' is vague. Consider the possibility of a speciation event occurring at the moment that the question is asked. Do we count the species in question as one or two? It is also worth stressing that no further information can be brought to light that will settle the matter. Perhaps we must settle for upper and lower bounds as the answer to the question. In the example of a speciation event taking place, we might give the interval $[n, n+1]$ as our answer. So we see that vagueness can give rise to this peculiar kind of uncertainty—an uncertainty that cannot be eliminated by gathering further data.

Now it might be argued that since the uncertainty in question here is uncertainty about the truth value of a vague proposition, we can state the problem classically in the meta-language. We can say that we do not know

[6] See, for example, [19,22,34,42] for some of the approaches that abandon excluded middle.

whether P (for some vague proposition P) is true. Let v be the valuation function (which maps from the domain of discourse D to the truth value set TV), then the problem is that of determining whether $v(P) = a$, where a is a particular truth value in TV. But, so the argument goes, '$v(P) = a$' is either true or false and so we have forged a link between non-epistemic uncertainty at the object-language level and epistemic uncertainty at the meta-language level. Indeed some do opt for a classical metalogic, but there is a case to be made for non-classicality all the way up. A non-classical metalogic would be called for, for instance, if there is higher-order vagueness. An adequate discussion of this would take us too far a field; I mention it merely to make the point that non-epistemic uncertainty does not reduce to epistemic uncertainty in any straightforward fashion. [7]

So far I have argued that in domains where excluded middle fails, the applicability of probability theory is highly questionable. The claim that classical probability theory is the only coherent representation of uncertainty suggests (among other things) that there are no domains about which we reason with uncertainty, where excluded middle fails. On the face of it at least, this is false: there are many such domains: there are fictional domains, constructive domains and domains with vague predicates. Thus any defender of classical logic needs to convince us that classical logic can, despite appearances, cope with these problematic domains. This is a large (if not impossible) task, for it involves, among other things, providing a classical account of fictional discourse, a defence of certain philosophical views about the philosophy of mathematics (perhaps defending platonism) and a defence of something like Williamson's epistemic approach to vagueness. [8]

Before I move on to a discussion of Cox's theorem, let us consider a couple of objections to my conclusion that probability theory is not appropriate for non-epistemic uncertainty. The first objection comes from quantum mechanics. According to the Copenhagen interpretation of quantum mechanics, there is no fact of the matter about the state of certain quantum systems before a measurement is made. But quantum theory itself provides us with probabilities about the state of the system in question (see [27]), so it seems that we have a

[7] See [55] for a discussion of higher-order vagueness.

[8] Worse still, there would seem to be inconsistent domains about which we reason. I have in mind here inconsistent mathematical theories (such as the early calculus and naive set theory) and inconsistent scientific theories (such as the conjunction of general relativity and quantum mechanics). Classical logic and classical probability theory are inappropriate in such domains since in classical logic everything follows from a contradiction and in classical probability theory, all contradictions have probability zero and all conditional probabilities conditional on a contradiction are undefined. Again if we reason about such domains, as we surely do, then it is clear that the classical theories are inadequate. See [38] for some recent papers on inconsistency in science and [37] for an account of a paraconsistent belief revision theory.

counterexample to the claim that we cannot use probability theory unless there is an underlying fact of the matter. The problem with this objection, however, is that it confuses what the quantum mechanical probabilities are about. The quantum mechanical probabilities are *not* about the state of the quantum system in question *before measurement*; rather, the probabilities are usually construed to be about the state of the system *were it to be measured* (or, if you prefer, they might be construed to be about the measurements themselves—the probability of the measurement turning out a particular way). Either way, the probabilities are not construed as being about systems in indeterminate states.

The next objection concerns denotational failure. According to some (e.g., [49]), when there is failure of denotation, there is no fact of the matter about the truth of the offending sentence (i.e., the offending sentence is truth-valueless). Let us, for the sake of argument, accept this view. (Indeed, I have already entertained fictional discourse—which is one special kind of denotational failure—as a source of non-epistemic uncertainty.) Suppose you see a male colleague, whom rumour would have it was supposed to be having marital problems, looking rather depressed and you speculate that his wife has left him. You might even believe that this is the most likely explanation for his depressed state. That is, you assign a subjective probability of greater than 0.5 to the truth of the proposition 'My colleague's wife has left him'. Now, as it turns out, your colleague is not, nor has he ever been, married. We thus have a case of denotational failure and so, by hypothesis, the sentence in question does not take a truth value. But, it still seems sensible to attribute a probability of truth to the sentence in question. [9] I agree that it seems sensible to entertain a probability of truth for the sentence in question, but it is not clear that it is sensible to do this *on the view under consideration*. After all, it also seems sensible to say that the sentence in question is false (this was Russell's [46] view), and on this view it does make sense to talk of the probability of such sentences being true. The issue is not whether it seems sensible to attribute truth-value gaps to sentences that have non-referring terms and whether it seems sensible to speak in terms of probability about these same sentences; the issue is whether the latter is sensible *given a commitment to the former*. That is, is it sensible to say, for instance, that some sentence is neither true nor false but it is probably true? It would seem not, for this would commit one to a kind of Moore's paradox. [10]

Now if it still seems sensible, on the view under consideration (i.e., the truth-value-gap view), to talk about the probability of truth for sentences with non-referring terms, it is because there is an implicit assumption that there is no denotational failure. So, for example, the probability that your colleague's wife

[9] I thank Daniel Nolan for raising this objection.

[10] This is the paradox of an agent asserting '*P* but I don't believe it'.

has left him is something like the probability that his wife has left him, given that he is married. If it turns out that he is not married, the probability in question is the probability that his wife has left him, given that he is both married and not married. This probability is undefined. So even if it may *seem* sensible on this view to talk about the probability of truth for sentences with non-referring terms, it is not.

3. Coxs theorem and its assumptions

Thus far I've outlined two quite distinct sorts of uncertainty and argued that only epistemic uncertainty is amenable to probabilistic treatment. Now I turn to Cox' theorem [11] and how the lessons of the last section impact on the philosophical significance of this theorem. Cox's theorem can be stated as follows:

Theorem 1 (Cox). *Any measure of belief is isomorphic to a probability measure.*

The theorem is explicitly premised on the following assumptions: (i) belief is a real-valued function (ii) an agent's belief in $\neg P$ is a function of his/her belief in P and (iii) an agent's belief in $P \wedge Q$ is a function of the agent's belief in P given Q and the agent's belief in Q.

There has been a great deal of discussion on the assumptions of Cox's theorem and alternatives to these assumptions [7,11,23,24,48,52], but there has been little if any discussion of the logical assumptions (or alternatives to these), and yet these are crucial to understanding the significance of the theorem. It is those logical assumptions I now wish to examine.

The logical assumptions of the theorem are, of course, none other than classical propositional logic, though very few state this explicitly and unambiguously. For example, Cox invokes "*the* algebra of symbolic logic" (italics added). But even in Cox's day there was more than one such logic. Jaynes [30] tells us that the logic is "deductive logic". But again 'deductive logic' is ambiguous between the many logics deserving of this title. Elsewhere [31, p. 9] Jaynes suggests that the logic is "two-valued logic or Aristotelian logic", obviously thinking that classical (two-valued) propositional calculus and Aristotelian logic are the same (which they are not). Others such as van Horn [52] refer to '*the* propositional calculus' (italics added), again as though there were only one such logic. [12] But what they all have in mind is quite clearly

[11] There are, in fact, a number of theorems along similar lines (e.g., [1,2,8,16,25,36,39,52]). Cox's theorem [9,10], however, is undoubtedly the most well known and so I'll be content to focus on it, although I'll often use the phrase 'Cox's theorem' to apply to the more general class of results.

[12] Though in footnote 1 on [52, p. 5] van Horn suggests that we may also consider numerical identity statements. This suggests that full (classical) first-order logic is what is needed.

classical propositional calculus. The problem is that none of them calls the logic in question by name and so it is (at least initially) unclear what logic they have in mind. Worse still, some suggest (e.g., by the use of the definite article 'the') that there is only one choice here. Those (like Jaynes [31]) who do point out that there are other logics to choose from do not bother to defend the choice of classical logic in any systematic fashion. [13]

Combine this unclarity about the logic in question (or the number of candidate logics) with a very commonly held view that logic is domain independent. [14] According to this view, the choice of logic does not depend on the domain of application. [15] So if we combine this commonly held view about logic with the view that "logic" is classical propositional logic (or first-order classical predicate calculus), then we are led to a view that classical logic is all we need for deductive inferences on any domain. Again it is clear that some commentators on Cox's theorem hold such a view. Indeed, Van Horn states this quite explicitly: "the propositional calculus is applicable to *any* problem domain for which we can formulate useful propositions" (p. 11, italics in original). Of course there is a sense in which Van Horn is right—classical propositional logic is *applicable* to any domain, but that is not the issue. The issue is whether classical propositional logic can be applied to any domain *and get the right answers*. It is clear that it cannot. One needs only consider arguments involving modality to see the inadequacy of classical propositional logic. [16] Van Horn, of course, is not alone in holding such a view of logic, though I have never seen anyone suggest that classical propositional calculus is the universal logic—the usual candidates are classical first-order logic or a extension of it such as S5 modal logic. But what I'm arguing here is that no classical

[13] Jaynes does make some rather obscure comments by way of defence of classical propositional logic. For instance, in a section of his book [31, p. 23] called 'Nitpicking' Jaynes raises the possibility of alternative logics and suggests that "[multiple-valued logics] can have no useful content that is not already in two-valued logic; that is, that an n-valued logic applied to one set of propositions is either equivalent to a two-valued logic applied to an enlarged set, or else it contains internal inconsistencies." It is not clear what he means by this, and the appendix where the argument for this claim is supposed to be found is of no help. In any case, Jaynes seems to be thinking of multi-valued logics as the only non-classical logics. As we have already seen, there are others—for example, free logics.

[14] This widely held view found a powerful advocate in Tarski [50].

[15] In essence, this is a monist or one-size-fits-all view of logic, as opposed to a more pluralist horses-for-courses view. See [5,41] for discussion on the monism–pluralism debate.

[16] Consider the argument from 'there is uncertainty' to 'possibly there is uncertainty'. This argument is clearly valid and yet the validity cannot be demonstrated by classical propositional calculus, because the only way to formalise this argument in this logic is as P therefore Q which is invalid. To demonstrate the validity of such arguments, modal logic is required. See [18] for a good introduction to modal logics and their applications.

logic is up to this task. Classical logic simply fails in some domains in which we routinely perform logical inferences.

Cox's theorem, if it is to demonstrate the adequacy of probability theory for plausible reasoning across *all* domains, it must be derivable from assumptions that are not domain specific. But as I've already argued, classical logic *is* domain specific. Or at least, we have been offered no argument to the effect that it is not. All we are typically given are rather casual acceptances of classical propositional calculus as though there were no other, or, at least, no other worthy of serious consideration. So what is delivered is not a logic of plausible reasoning, *simpliciter*, instead we have a logic of plausible reasoning that is defensible only when there is no referential failure, vagueness or the like. Now perhaps this is all some commentators have in mind—a limited-scope logic of plausible reasoning. If this is the case, then this limitation needs to be stressed. But it is clear that not all contributors to the literature on Cox's theorem have such a modest project in mind. [17] Again Van Horn states this point of view very clearly: "recall the purpose of this enterprise: to construct a *universal* system or logic of plausible reasoning" (p. 11, again emphasis in original). My point is simply that if the enterprise is, as Van Horn suggests, that of constructing a universal system, it had better not rest on classical logic. On the other hand, if the enterprise is the more modest one suggested above, this needs to be made clear.

Now let us turn briefly to the question of whether the proof of the theorem requires any of the contentious features of classical logic? What if, for instance, the proof only relied on inferences and logical equivalences that are not controversial in the context of the representation of belief—inferences such as *modus ponens* and equivalences such as de Morgan laws? There is no need to ponder such questions too long, for the standard proofs of Cox-style results quite clearly rely on disputed logical principles. First, an example from Cox's original proof and then another example from a more recent proof. In Cox's proof that the belief in $Q \vee \neg Q$ is maximal, Cox quite explicitly assumes the classical principle of double negation elimination: $\neg\neg Q \equiv Q$. If we limit our attention to epistemic uncertainty and exclude all forms of non-epistemic uncertainty, then the assumption seems harmless. On the other hand, if we are interested in uncertainty in the broadest sense (including constructive domains, vague domains and so on) the assumption is highly controversial. [18] For a more recent example I once again turn to Van Horn [52] who also uses double negation elimination (in the proof of Proposition 2 on p. 11) and assumes, in

[17] And I stress that this *is* a modest project, because vagueness is ubiquitous in both scientific and everyday discourse. The limited-scope logic of plausible reasoning will thus be rarely applicable outside pure mathematics.

[18] Indeed, intuitionists deny this principle.

the proof of Lemma 11 (on p. 20), that $\neg B \equiv ((A \vee \neg B) \wedge (\neg A \vee \neg B))$. This last assumption is very closely related to excluded middle. With the usual classical assumptions in place about the distribution of $\vee$ over $\wedge$, [19] it amounts to the assumption that $A \wedge \neg A$ is false, which under further (classical) assumptions is equivalent to excluded middle. So at the end of the day, controversial features of classical logic are assumed in the original proof of Cox's theorem and these assumptions remain in modern presentations.

The assumption of classical logic is particularly troublesome if Cox's theorem is to be wielded as a weapon against non-classical systems of belief representation. And, I should add, that some commentators do put Cox's theorem to such a purpose. For example, Lindley [36] draws the following conclusion (from a similar theorem): "The message is essentially that only probabilistic descriptions of uncertainty are reasonable" (p. 1) and Jaynes [31] suggests that "the mathematical rules of probability theory [...] are [...] the unique consistent rules for conducting inference (i.e., plausible reasoning) of any kind" (p. xxii). [20] But Cox's result is simply a representation theorem demonstrating that if belief has the structure assumed for the proof of the theorem, classical probability theory is a legitimate calculus for representing degrees of belief. But as it stands it certainly does not legitimate *only* classical probability theory as a means of representing belief, nor does it prove that such a representation is adequate for all domains.

What are the alternatives then? What would these alternate belief theories look like? If we want a probability theory for non-epistemic uncertainty, we may wish to base it on a logic in which excluded middle fails. This means that propositions of the form $P \vee \neg P$ would not automatically receive maximal probability. There are a couple of ways of doing this. One approach would be to allow tautologies to take probability assignments less than one. The other approach is to underwrite the probability theory with a non-classical logic. In this latter case, the tautologies of the non-classical logic will receive maximal probability—it is just that classical tautologies such as $P \vee \neg P$ would not, in general, get assigned the maximal value. [21] In some of the logics in contention, there may be no tautologies (as is the case with Kleene's three valued system K3 and the most popular fuzzy logics [40]). If we use one of these as the underlying logic, there would not be any logical truths and so there would not be any propositions automatically assigned the maximal probability. Some work has been carried out in these directions but there is much more to do.

---

[19] Interestingly, distribution fails in quantum logics. So quantum logicians may contest the logical equivalence that Van Horn relies on, but for slightly different reasons. See [3] for an early presentation of quantum logic.

[20] Shafer [48] also notes (disapprovingly) this use of Cox's theorem to rule against anything other than standard probability theory.

[21] See [54] for a constructive probability theory.

4. Conclusion

Let me finish by noting a few points of contact between this paper and Glen Shafer's recent discussion of Cox's theorem in this journal [48]. Shafer notes that Cox's theorem relies not only on its stated, explicit assumptions, but it also relies on implicit assumptions—such as the assumption that belief should be represented by a real-valued function. I note one other implicit (or at least undefended) assumption—the use of classical propositional logic. Shafer's work on belief functions [47] casts doubt over the plausibility of the assumption that belief is adequately represented by a real-valued function. [22] I've pointed out that work in logic in the latter part of the 20th century casts doubt over the plausibility of the assumption (used by Cox and others) that classical logic is the appropriate logic to underwrite a formal theory of plausible reasoning.

The connection between this paper and Shafer's runs even deeper. Not only are both papers questioning implicit assumptions of Cox's theorem. It turns out that our concerns may well be two sides of the one coin. Although our starting points are apparently quite different—mine being logic, Shafer's being the representation of imprecise belief. It turns out that starting with concerns such as mine (i.e., concerns about vagueness and other forms of non-epistemic uncertainty), one very natural way of responding to these issues is to give up the classical logical principle of excluded middle. This in turn naturally leads to a non-classical belief theory that is very similar to Shafer's. [23] In essence we are both reject the unrealistic precision assumed by standard belief theory. Shafer rejects the assumption that belief functions are real valued; I reject the logical assumption of excluded middle.

Another point of contact is that Shafer stresses that the assumptions of Cox's theorem need to be more than merely plausible, they need to be self-evident. He points out that both the explicit assumptions and the implicit assumption that belief functions are real-valued fail in this regard. I concur and I add one further assumption to this list of non-self-evident assumptions. In the context of the representation of uncertainty classical logic is not self-evidently the appropriate logic. Indeed, I think it is demonstrably *not* the appropriate logic, but even if you disagree with me on this stronger claim, the fact remains that classical logic is not self-evident. So those who would employ Cox-style results for the purpose of providing a logic of plausible inference, need to first mount a defence of classical logic.

[22] And I find myself in full agreement with Shafer on this issue. See [32,53] and for other approaches to abandoning the assumption that a single real number is adequate for characterising belief.

[23] See [13] for details.

The final point of contact between my discussion here and Shafer's is that we are both interested in widening the historical focus of the discussion of Cox's theorem. Shafer wants to draw to the attention of commentators on Cox's theorem the earlier work (by continental probability theorists) on the logical interpretation of probability—frequentism and subjectivism are not, nor were they in 1946 (when Cox wrote his paper), the only games in town. I wish to bring to the discussion the issue of the underlying logic—classical logic is not, nor was it in 1946, the only game in town. [24] I think a lot is to be gained by considering these broader historical and, I might add, interdisciplinary perspectives. Once one does this, one sees that Cox's theorem is an interesting representation theorem that has prompted some fruitful debate, but ultimately the theorem rests on some rather questionable assumptions about the structure of human belief.

Acknowledgements

I would like to thank Phil Dowe, Scott Ferson, Hartry Field, Dominic Hyde, Daniel Nolan, Graham Priest and Helen Regan for useful discussion or comments on this paper. A version of this paper was presented at the Australasian Association of Philosophy Conference at the University of Queensland in July 2000. I am grateful to the participants in the subsequent discussion for their valuable contributions. I would also like to thank an anonymous referee of this journal for many extremely helpful suggestions and comments. These led to improvements throughout the paper.

References

[1] J. Aczél, Lectures on Functional Equations and their Applications, Academic Press, New York, 1966.
[2] R. Aleliunas, A summary of a new normative theory of probabilistic logic, in: R. Shachter, T. Levitt, L. Kanal, J. Lemmer (Eds.), Uncertainty in Artificial Intelligence, vol. 4, North-Holland, New York, 1990, pp. 8–14.
[3] G. Birkhoff, J. von Neumann, The logic of quantum mechanics, Annals of Mathematics 37 (1936) 823–843.
[4] J.C. Beall, M. Colyvan, Heaps of gluts and Hyde-ing the sorites, Mind 110 (2001) 401–408.
[5] J.C. Beall, G. Restall, Logical pluralism, Australasian Journal of Philosophy 78 (2000) 475–493.

[24] By 1946 a variety of non-classical logics had been developed—various many-valued logics and intuitionist logic, in which excluded middle fails. Quantum logic and various modal logics were also known by this time. Moreover, many of these logics had found interesting and fruitful applications. See [40] for a good introduction to non-classical logics and [33] for some of the relevant history.

[6] P. Benacerraf, What mathematical truth could not be, in: A. Morton, S.P. Stich (Eds.), Benacerraf and his Critics, Blackwell, 1996, pp. 9–59.
[7] M. Colyvan, Is probability the only coherent approach to uncertainty?, Risk Analysis, forthcoming.
[8] P. Cheeseman, An inquiry into computer understanding, Computational Intelligence 4 (1988) 58–66.
[9] R.T. Cox, Probability frequency and reasonable expectation, American Journal of Physics 14 (1946) 1–13.
[10] R.T. Cox, The Algebra of Probable Inference, Johns Hopkins Press, Baltimore, 1961.
[11] D. Dubois, H. Prade, The logical view of conditioning and its application to possibility and evidence theories, International Journal of Approximate Reasoning 4 (1990) 23–46.
[12] H. Field, Mathematical objectivity and mathematical objects, in: S. Laurence, C. Macdonald (Eds.), Contemporary Readings in the Foundations of Metaphysics, Blackwell, Oxford, 1998, pp. 387–403.
[13] H. Field, Indeterminacy degree of belief and excluded middle, Noûs 34 (2000) 1–30.
[14] H. Field, Truth and the Absence of Fact, Oxford University Press, Oxford, 2001.
[15] K. Fine, Vagueness truth and logic, Synthese 30 (1975) 265–300.
[16] T.L. Fine, Theories of Probability, Academic Press, New York, 1973.
[17] R.A. Girle, Logics for knowledge, possibility and existence, Notre Dame Journal of Formal Logic 19 (1978) 200–214.
[18] R.A. Girle, Modal Logics and Philosophy, Acumen, Teddington, UK, 2000.
[19] J. Goguen, The logic of inexact concepts, Synthese 19 (1968–69) 325–373.
[20] S. Haack, Philosophy of Logics, Cambridge University Press, Cambridge, 1978.
[21] S. Haack, Deviant Logic Fuzzy Logic: Beyond the Formalism, University of Chicago Press, Chicago, 1996.
[22] S. Halldén, The Logic of Nonsense, Uppsala Universitets Arsskrift, Uppsala, 1949.
[23] J.Y. Halpern, A counterexample to theorems of Cox and Fine, Journal of Artificial Intelligence Research 10 (1999) 67–85.
[24] J.Y. Halpern, Cox's theorem revisited, Journal of Artificial Intelligence Research 11 (1999) 429–435.
[25] D. Heckerman, An axiomatic framework for belief updates, in: J.F. Lemmer, L. Kanal (Eds.), Uncertainty in Artificial Intelligence, vol. 2, North-Holland, Amsterdam, 1988, pp. 11–22.
[26] A. Heyting, Intuitionism: An Introduction, third revised ed., North-Holland, Amsterdam, 1971.
[27] R.I.G. Hughes, The Structure and Interpretation of Quantum Mechanics, Harvard University Press, Cambridge, MA, 1989.
[28] D. Hyde, From heaps and gaps to heaps of gluts, Mind 106 (1997) 641–660.
[29] D. Hyde, A reply to Beall and Colyvan, Mind 110 (2001) 409–411.
[30] E.T. Jaynes, How does the brain do plausible reasoning?, in: G.J. Erickson, C.R. Smith (Eds.), Maximum Entropy and Bayesian Methods in Science and Engineering, Kluwer, Dordrecht, 1988.
[31] E.T. Jaynes, Probability Theory: The Logic of Science, Cambridge University Press, Cambridge, 2003.
[32] R. Jeffrey, Bayesianism with a human face, in: Probability and the Art of Judgment, Cambridge University Press, 1992 (originally published in 1983), pp. 77–107.
[33] W. Kneale, M. Kneale, The Development of Logic, Clarendon Press, Oxford, 1962.
[34] S. Körner, Conceptual Thinking, Cambridge University Press, Cambridge, 1955.
[35] K. Lambert (Ed.), Philosophical Applications of Free Logic, Oxford University Press, Oxford, 1991.
[36] D.V. Lindley, Scoring rules and the inevitability of probability, International Statistical Review 50 (1982) 1–26.

[37] E. Mares, A paraconsistent theory of belief revision, Erkenntnis, forthcoming.
[38] J. Meheus (Ed.), Inconsistency in Science, Kluwer, Dordrecht, 2002.
[39] J.B. Paris, The Uncertain Reasoner's Companion: A Mathematical Perspective, Cambridge University Press, Cambridge, 1994.
[40] G. Priest, An Introduction to Non-Classical Logic, Cambridge University Press, Cambridge, 2001.
[41] G. Priest, Logic: one or many?, in: B. Brown, J. Woods (Eds.), Logical Consequences, Kluwer, Dordrecht, forthcoming.
[42] H. Putnam, Vagueness and alternative logic, in: Realism and Reason: Philosophical Papers, vol. 3, Cambridge University Press, Cambridge, 1983, pp. 271–286.
[43] R. Routley, Some things do not exist, Notre Dame Journal of Formal Logic 7 (1966) 251–276.
[44] H.M. Regan, M. Colyvan, M.A. Burgman, A proposal for fuzzy IUCN categories and criteria, Biological Conservation 92 (2000) 101–108.
[45] H.M. Regan, M. Colyvan, M.A. Burgman, A taxonomy and treatment of uncertainty for ecology and conservation biology, Ecological Applications 12 (2002) 618–628.
[46] B. Russell, On denoting, Mind 14 (1905) 479–493.
[47] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[48] G. Shafer, Comments on "Constructing a logic of plausible inference: a guide to Cox's theorem" by Kevin S. Van Horn, International Journal of Approximate Reasoning, 35 (2004) 97–105.
[49] P.F. Strawson, Identiying reference and truth values, Theoria 30 (1964) 96–118.
[50] A. Tarski, On the concept of logical consequence, in: Logic, Semantics, Metamathematics: Papers from 1923 to 1938, second ed., Hackett, Indianapolis, 1983 (originally published in 1936), pp. 409–420.
[51] B.C. Van Fraassen, Singular terms truth value gaps and free logic, Journal of Philosophy 63 (1966) 481–495.
[52] K.S. Van Horn, Constructing a logic of plausible inference: a guide to Cox's theorem, International Journal of Approximate Reasoning 34 (2003) 3–24.
[53] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.
[54] B. Weatherson, From classical to constructive probabiliy, Notre Dame Journal of Formal Logic, forthcoming.
[55] T. Williamson, Vagueness, Routledge, London, 1994.

AMERICAN JOURNAL of PHYSICS

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 14, NUMBER 1 JANUARY–FEBRUARY, 1946

Probability, Frequency and Reasonable Expectation

R. T. Cox
The Johns Hopkins University, Baltimore 18, Maryland

Frequency or Reasonable Expectation as the Primary Concept

THE concept of probability has from the beginning of the theory involved two ideas: the idea of frequency in an ensemble and the idea of reasonable expectation. The choice of one or the other as the primary meaning of probability has distinguished the two main schools of thought in the theory.[1]

If a box contains two white balls and one black ball, indistinguishable except by color, both schools agree that the probability that a blindfolded man will draw a white ball on a single trial is $\frac{2}{3}$ and the probability that he will draw a black ball is $\frac{1}{3}$. On the frequency theory, the primary meaning of these probabilities is in terms of the ensemble. The ensemble may be an indefinitely large number of such boxes having the same contents, or it may be an indefinitely large number of drawings from the same box, the ball drawn being replaced each time. The significant point is that the initial circumstances are assumed to be capable of indefinite repetition, these repetitions constituting the ensemble. That the probability of a white ball is $\frac{2}{3}$ means simply that the number of trials giving a white ball as result is $\frac{2}{3}$ the number of trials in the whole ensemble. According to the frequency theory, this is not a prediction of the theory of probability but the definition of the probability. Probability in that theory is a characteristic of the ensemble and, without the ensemble, cannot be said to exist.

Again, according to both schools, the probability of a white ball in two successive drawings, when the first ball drawn is not replaced, is $\frac{2}{3}\times\frac{1}{2}$, or $\frac{1}{3}$. According to the frequency theory, this implies that two balls are drawn successively from each of an ensemble of boxes containing originally two white balls and one black ball. On $\frac{2}{3}$ of the trials a white ball is drawn first and one white and one black ball are left in the box. Then, in $\frac{1}{2}$ of the trials which give this result, a white ball is drawn next, so that $\frac{1}{3}$ of the whole number of trials give white balls on both drawings. These examples illustrate the general fact that, when probability is identified with frequency in an ensemble, the probabilities are calculated by arithmetic in particular examples and, in more general cases, the rules of probability are found by ordinary algebra.[2]

[1] If minor differences are counted, the number of schools seems to be somewhere between two and the number of authors, and probably nearer the latter number. But the clearest line of division is the one mentioned

[2] An exposition of the frequency theory, with some comment on other theories, has been given by G. Bergmann, *Am. J Physics* 9, 263 (1941) Readers with a wider knowledge of philosophy than mine will be better able to compare his views with those of this paper.

Probability is recognized also as providing a measure of the reasonable expectation of an event in a single trial. That the probability of drawing a white ball is $\frac{2}{3}$ and of drawing a black ball is $\frac{1}{3}$ means that a white ball is a more likely result of a trial than a black ball, and the numbers $\frac{2}{3}$ and $\frac{1}{3}$ serve to compare the likelihoods of the two results. According to the second main school of probability, this measure of reasonable expectation, rather than the frequency in an ensemble, is the primary meaning of probability.

If it could be shown that every measure of reasonable expectation is also a frequency in some ensemble and that every frequency in an ensemble measures a reasonable expectation, then the choice of one or the other as the primary meaning of probability would not be very important. I shall not attempt to discuss whether there are frequencies in an ensemble that are not measures of reasonable expectation. It is enough for my present purpose to show that the two interpretations are not always identical. For this it will suffice to point out that there are probabilities in the sense of reasonable expectations for which no ensemble exists and for which, if one is conceived, it is clearly no more than a convenient mental artifice. Thus, when the probability is calculated that more than one planetary system exists in the universe, it is barely tenable even as an artifice that this refers to the number of universes having more than one planetary system among an indefinitely large number of universes, all resembling in some way *the* universe, which by definition is all-inclusive

Moreover, there is so gradual a transition from the cases in which there is a discoverable ensemble and those in which there is none that a theory which requires a sharp distinction between them offers serious difficulties. A few examples will illustrate this point. Let us consider the probability that the number of heads thrown in a certain number of tosses of an unbiased coin shall lie within certain limits, and let us compare with this the probability, often considered, that the true value of a physical constant lies within certain limits. The two probabilities have something in common, but there is a difference between them. The difference lies in the causes that oblige us to deal with probabilities rather than certainties in discussing the score in tossing a coin and the value of a physical constant In discussing the score in a given number of tosses of a coin, we have to use probabilities because the score will vary from one trial to another. The true value of a physical constant, on the other hand, is unique. We have to speak of the probability that it lies within certain limits only because our knowledge is incomplete.

Sometimes, it is true, the probability that the value of a physical constant lies within certain limits is equivalent to another probability, that the error in the average of a number of measurements lies within these limits. If there are no systematic sources of error, we may imagine an ensemble of measurements and treat the measurements made as a random sample of this ensemble. The probability in question may then be found in a manner similar to that used in dealing with the coin. For example, the probability of certain limits for the true value of the Joule equivalent may perhaps be considered in this way.

The case is somewhat different with the reciprocal fine-structure constant that appears in quantum mechanics. For here, in addition to the values derived from measurements, there is evidence of another sort in the argument adduced by Eddington[3] that this constant may be expected to be an integer, having the value 137 If it should be estimated from the measurements alone that there is an equal probability that the constant lies inside or outside of certain limits which include 137, then Eddington's argument will increase the probability that it lies inside these limits and correspondingly decrease the probability that it lies outside.

As a final example, we may consider the case of a purely mathematical constant, of which the existence has been proved but the value determined only within certain limits. A problem of the theory of numbers, discussed by Hardy[4] among others, provides a good example. It concerns the equivalence of an integer to a sum

[3] A. S. Eddington, *Relativity theory of protons and electrons* (Macmillan, New York, and Univ. Press, Cambridge, 1936)

[4] G H. Hardy, *Some famous problems of the theory of numbers and in particular Waring's problem* (Clarendon Press, Oxford, 1920)

of cubes of smaller integers. It has been proved that any integer is given by the sum of not more than 9 cubes, and that any integer beyond a certain one is given by the sum of 8 or fewer. It is expected that large enough integers may all be expressed by the sum of some still smaller number of cubes, and the problem is to find the minimum number required for all integers above a certain value. It has been proved that, if this value is taken high enough, the number in question is 4, 5, 6, 7 or 8. This is as far as rigorous proof has gone, but the evidence of computation makes some of these numbers less likely than others. In very large samples—all of the first 40,000 integers and the 2000 ending at one million—the integers requiring 8 cubes are found to drop out early in the progress to higher integers and those requiring 7 disappear somewhat farther on, while those requiring 6 occur more and more rarely until there are only two among the 2000 integers next below one million. Hardy concludes that the minimum number for large enough integers is almost certainly neither 8 nor 7 and probably not 6. There remain 5 and 4 as the likely numbers, and he seems to favor 4 as the more probable.

Let us consider now these four examples. the probability of certain limits for (i) the score in a number of throws of a coin, (ii) the value of the Joule equivalent, (iii) the value of the reciprocal fine-structure constant and (iv) the value of the least number of cubes for the expression of large integers. It will most likely be granted that other examples can be interpolated among these, so that the differences will be very slight between each example and those next before and after. We shall have then a graded series of examples of probability. At one end of the series the interpretation of probability in terms of frequency will be valid, at the other end it will be impossible. For it is certainly impossible to discuss the statistical spread of the determinations of a number which has never in fact been determined and of which the determination, when it is made, will give a single and logically inevitable value.

Nevertheless, it must be admitted that there is a kind of reasoning common to all these examples. The gambler in the first example, the physicist in the second and third, and the mathematician in the fourth are all using similar processes of inference.

In this connection it is worth while to observe how much of the theory of probability deals with relations between probabilities: between the probability that an event will not occur and the probability that it will occur, between the probability of both of two events and their separate probabilities, between these probabilities and the probability that at least one of the two events will occur. In the case of probabilities that can be identified with frequencies in an ensemble, these relations are readily obtained by ordinary algebra, as was mentioned earlier. But the same or at least similar relations are involved in inference concerned with reasonable expectation even when no ensemble is discoverable. Thus, under any definition of probability, or even without an attempt to define it precisely, there will still be agreement that the less likely an event is to occur the more likely it is not to occur. The occurrence of both of two events will not be more likely and will generally be less likely than the occurrence of the less likely of the two. But the occurrence of at least one of the events is not less likely and is generally more likely than the occurrence of either.

For example, if it were a question of the credibility of a certain hypothesis for the origin of terrestrial life or of human language, one would hold it as a point against the hypothesis that it postulated the occurrence of two events, of which neither was considered very probable, but the hypothesis would gain in credibility if it could be justified by postulating merely that one or the other of these events had occurred. Generally speaking, a simple hypothesis is preferred to a complex one. If this preference is founded on a reasonable belief rather than being a mere convention, its justification would seem to be that two or more postulates are less likely to be true than a single one of about the same likelihood.

This difficulty of the frequency theory of probability may now be summarized. There is a field of probable inference which lies outside the range of that theory. The derivation of the rules of probability by ordinary algebra from the characteristics of the ensemble cannot justify the use of these rules in this outside field.

Nevertheless, the use of these rules in this field is universal and appears to be a fundamental part of our reasoning. Thus the frequency theory is inadequate in the sense that it fails to justify what is conceived to be a legitimate use of its own rules.

From a purely rational point of view, the extent of this field of inference outside the range of the frequency theory is irrelevant to the point in question. Even if the valid instances of reasoning in this field were rare and of little consequence, it would still be logically necessary to maintain the inadequacy of the frequency theory. As a practical matter, however, if these instances were few or trivial, we should probably be content to ignore them. But actually, as I have tried to suggest by the examples given, it is rather the cases in which a strictly definable ensemble exists that are exceptional. This is not to say that they are numerically few. There are many of them, and they have a particular interest, but they still do not appear to comprise in our ordinary practice the greater part of the uses of probable inference Nor are the other uses by any means trivial. Kemble,[5] in an interesting paper which covers, among other things, some of the ground thus far traversed here, has made the point that the frequency definition of probability does not suffice to establish the connection between statistical mechanics and thermodynamics, which is certainly crucial in physical theory.

A very original and thoroughgoing development of the theory of probability, which does not depend on the concept of frequency in an ensemble, has been given by Keynes.[6] In his view, the theory of probability is an extended logic, the logic of probable inference. Probability is a relation between a hypothesis and a conclusion, corresponding to the degree of rational belief and limited by the extreme relations of certainty and impossibility. Classical deductive logic, which deals with these limiting relations only, is a special case in this more general development. Hence it follows in general that the theory of probability cannot be based entirely on concepts of classical logic. In particular, the relation of probability cannot be defined in terms of certainty, since certainty itself is a special case of probability. The frequency definition of probability is therefore invalid, since it depends on the relations of certainty involved in the knowledge of numbers of instances. Probability is taken as a primary concept, like distance or time in mechanics, not reducible to any more elementary terms

Merely to describe Keynes' position, as I have done, without giving the reasoning by which he is led to it, does his work very poor justice. The reasoning is, to me at least, very convincing, and is the original source of a large part of the opinions given here, though the arguments I have used are not the same as his. Nevertheless, it must be conceded that his work does not bring us very far in the solution of the problem mentioned earlier, that of justifying the few basic rules of probable inference necessary for the development of the theory. These rules, in Keynes' theory, are simply taken as axiomatic. Now some primary assumptions will have to be made by anyone who accepts, as I am strongly inclined to do, his general point of view as to the nature of probability, because some rational starting point is needed to replace the frequency definition, once that has been abandoned. But Keynes' axioms seem to me, as they have doubtless seemed to others, including Kemble, somewhat too arbitrary and too sophisticated to be entirely suitable as axioms. They do not appeal very directly to common sense, and it is hard to see how they would have been formulated without considering colored balls in a box, dice, coins, or some of the other devices associated with the concept of the ensemble. It is rather as if Euclid had placed the Pythagorean theorem among the axioms of plane geometry.

Russell[7] makes a criticism somewhat different in form, but which may have the same ground as this. After conceding the strength of Keynes' argument against the frequency theory, he nevertheless prefers that theory, if it can be logically established, because of its explicit definition of probability. It is this definition that makes it possible to avoid the assumption of axioms such as characterize Keynes' theory.

[5] E C Kemble, *Am J. Physics* **10**, 6 (1942).
[6] J. M Keynes, *A treatise on probability* (Macmillan, London, 1929).

[7] B. Russell, *Philosophy* (Norton, New York, 1927).

Other authors, who, like Keynes, present an axiomatic development, choose somewhat different sets of postulates, but those I have seen still show some of the tool marks of their original derivation from the study of games of chance, with the consequent implication of an ensemble. I think this is true even of the carefully chosen postulates of Jeffreys and Wrinch, whether in their original form or as revised by Jeffreys.[8]

Relations of Reasonable Expectation Consistent with Symbolic Logic

In what follows next, I shall try to show that by employing the algebra of symbolic logic it is possible to derive the rules of probability from two quite primitive notions, which are independent of the concept of the ensemble and which, as I think, appeal rather immediately to common sense. This algebra has been applied to probability by a number of writers, including Boole,[9] who originated it. Still, its possibilities in this respect do not seem to have been fully realized. It may be well here to give a brief introduction to the Booleian algebra, at least to as much of it as the later argument will require.

Letters, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\cdots$, will denote propositions. There is an advantage in speaking of the probabilities of propositions rather than of events, partly for the sake of greater generality but mainly because speaking of events easily invokes the notion of sequence in time, and this may become a source of confusion. A proposition may, of course, assert the occurrence of an event, but it may just as well assert something else, for example, something about a physical constant. The proposition not-$\mathbf{a}$ will be denoted by $\sim\mathbf{a}$, the proposition $\mathbf{a}$-and-$\mathbf{b}$ by $\mathbf{a}\cdot\mathbf{b}$, and the proposition $\mathbf{a}$-or-$\mathbf{b}$ by $\mathbf{a}\vee\mathbf{b}$.

It is to be borne in mind that the proposition $\sim\mathbf{a}$ is not the particular proposition which in some sense is the opposite of $\mathbf{a}$. Thus if $\mathbf{a}$ is the proposition, "The stranger was a short, fat old man without coat or hat," $\sim\mathbf{a}$ is not the proposition, "The stranger was a tall, thin young woman with coat and hat." To assert $\sim\mathbf{a}$ means nothing more than to answer "no" to the question, "Is $\mathbf{a}$ wholly true?" If $\mathbf{a}$ is in several parts, $\mathbf{a}_1$, $\mathbf{a}_2$, $\cdots$, to assert $\sim\mathbf{a}$ is not to affirm that $\mathbf{a}_1$, $\mathbf{a}_2$, $\cdots$ are all false but only to say that at least one of them is false.

Since the letters $\mathbf{a}$, $\mathbf{b}$ denote propositions and not events, the order in which they appear in the symbols $\mathbf{a}\cdot\mathbf{b}$ and $\mathbf{a}\vee\mathbf{b}$ is only the order in which two propositions are stated, not the order in time in which two events occur. Also the form $\mathbf{a}\cdot\mathbf{a}$ indicates only that a proposition is twice stated, not that an event has twice occurred.

It is also to be understood that $\mathbf{a}\vee\mathbf{b}$ means $\mathbf{a}$-or-$\mathbf{b}$ in the sense of the child who asks, "May I have a nickel or a dime?" without meaning to exclude the possibility of both a nickel and a dime, not in the sense of the orator saying "Sink or swim, survive or perish." Thus $\mathbf{a}\vee\mathbf{b}$ has the sense for which the form $\mathbf{a}$ and/or $\mathbf{b}$ is often employed.

Finally it may be noted that if the proposition, "It is raining," is true, then the proposition, "It is raining or snowing," is also true. To assert a proposition $\mathbf{a}$ is to imply every proposition $\mathbf{a}\vee\mathbf{b}$ of which $\mathbf{a}$ is one term.

With the meaning of the symbols thus understood, the rules for their combination may be set down as follows:

$$\sim\sim\mathbf{a}=\mathbf{a}, \tag{1}$$

$$\mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{a}, \quad (2) \qquad \mathbf{a}\vee\mathbf{b}=\mathbf{b}\vee\mathbf{a}, \quad (2')$$

$$\mathbf{a}\cdot\mathbf{a}=\mathbf{a}, \quad (3) \qquad \mathbf{a}\vee\mathbf{a}=\mathbf{a}, \quad (3')$$

$$\mathbf{a}\cdot(\mathbf{b}\cdot\mathbf{c})=(\mathbf{a}\cdot\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\ \mathbf{b}\cdot\mathbf{c}, \tag{4}$$

$$\mathbf{a}\vee(\mathbf{b}\vee\mathbf{c})=(\mathbf{a}\vee\mathbf{b})\vee\mathbf{c}=\mathbf{a}\vee\mathbf{b}\vee\mathbf{c}, \tag{4'}$$

$$\sim(\mathbf{a}\cdot\mathbf{b})=\sim\mathbf{a}\vee\sim\mathbf{b}, \tag{5}$$

$$\sim(\mathbf{a}\vee\mathbf{b})=\sim\mathbf{a}\cdot\sim\mathbf{b}, \tag{5'}$$

$$\mathbf{a}\cdot(\mathbf{a}\vee\mathbf{b})=\mathbf{a}, \quad (6) \qquad \mathbf{a}\vee(\mathbf{a}\cdot\mathbf{b})=\mathbf{a}. \quad (6')$$

These eleven rules are not all independent. From six of them it is possible to prove the remaining five, and the set of six may be chosen in various ways. It is necessary only to include the first and one from each similarly numbered pair of the others. Thus, for example, Eq. (5′) is derived as follows.

$$\begin{aligned}\sim(\mathbf{a}\vee\mathbf{b}) &= [\text{by Eq. (1)}]\sim(\sim\sim\mathbf{a}\vee\sim\sim\mathbf{b})\\ &= [\text{by Eq. (5)}]\sim\sim(\sim\mathbf{a}\cdot\sim\mathbf{b})\\ &= [\text{by Eq. (1)}]\sim\mathbf{a}\cdot\sim\mathbf{b}.\end{aligned}$$

[8] H. Jeffreys, *Theory of probability* (Clarendon Press, Oxford, 1939).

[9] G. Boole, *An investigation of the laws of thought* (Macmillan, London, 1854).

Now let the symbol $\mathbf{b}|\mathbf{a}$ denote some measure of the reasonable credibility of the proposition **b** when the proposition **a** is known to be true.[10] The term is indefinite at this point, because, if there is one such measure, there will be any number of others. If $\mathbf{b}|\mathbf{a}$ is one such measure, then an arbitrary function $f(\mathbf{b}|\mathbf{a})$ will also be a measure. Consequently, the symbol is not now to be identified with the conventional probability. To avoid that implication, I shall call $\mathbf{b}|\mathbf{a}$ the *likelihood* of the proposition **b** on the hypothesis **a**, taking advantage of a suggestion made by Margenau,[11] but at the same time taking the liberty of giving the term a more inclusive meaning than the one he proposed.

It is not to be supposed that a relation of likelihood exists between any two propositions. If **a** is the proposition "Caesar invaded Britain" and **b** is "Tomorrow will be warmer than today," there is no likelihood $\mathbf{b}|\mathbf{a}$, because there is no reasonable connection between the two propositions.

It is now time to make the first of the two assumptions mentioned earlier as providing a basis for the principles of probable inference. We assume, whatever measure be chosen, that the likelihood $\mathbf{c}\cdot\mathbf{b}|\mathbf{a}$ is determined in some way by the two likelihoods $\mathbf{b}|\mathbf{a}$ and $\mathbf{c}|\mathbf{b}\cdot\mathbf{a}$, or

$$\mathbf{c}\cdot\mathbf{b}|\mathbf{a}=F(\mathbf{c}|\mathbf{b}\cdot\mathbf{a},\ \mathbf{b}|\mathbf{a}), \tag{7}$$

where F is some function of two variables.

Written in symbolic form, this assumption may not appear very axiomatic. Actually it is a familiar enough rule of common sense, as an example will show. Let **b** denote the proposition that an athlete can run from one given place to another, and let **c** denote the proposition that he can run back without stopping. The physical condition of the runner and the topography of the course are described in the hypothesis **a**. Then $\mathbf{b}|\mathbf{a}$ is the likelihood that he can run to the distant place, estimated on the information given in **a**, and $\mathbf{c}|\mathbf{b}\cdot\mathbf{a}$ is the likelihood that he can run back, estimated on the initial information and the further assumption that he has just run one way. These are just the likelihoods that would have to be considered in estimating the likelihood, $\mathbf{c}\cdot\mathbf{b}|\mathbf{a}$, that he can run the complete course without stopping. In postulating only that the last-named likelihood is some function of the other two, we are making the least restrictive assumption possible.

The form of the function F is partly conventional because of the indefiniteness of the measure to be used for likelihood. But it is not wholly so, for it must be consistent with the algebra of propositions. Accordingly we make use of Eq. (4) to derive a functional equation involving F, as follows:

$$\begin{aligned}\mathbf{d}\cdot\mathbf{c}\cdot\mathbf{b}|\mathbf{a}&=[\text{by Eq. (4)}](\mathbf{d}\cdot\mathbf{c})\cdot\mathbf{b}|\mathbf{a}\\&=[\text{by Eq. (7)}]F(\mathbf{d}\cdot\mathbf{c}|\mathbf{b}\cdot\mathbf{a},\ \mathbf{b}|\mathbf{a}).\end{aligned}$$

But

$$\begin{aligned}\mathbf{d}\cdot\mathbf{c}|\mathbf{b}\cdot\mathbf{a}&=[\text{by Eq. (7)}]F[\mathbf{d}|\mathbf{c}\cdot(\mathbf{b}\cdot\mathbf{a}),\ \mathbf{c}|\mathbf{b}\cdot\mathbf{a}]\\&=[\text{by Eq. (4)}]F(\mathbf{d}|\mathbf{c}\cdot\mathbf{b}\cdot\mathbf{a},\ \mathbf{c}|\mathbf{b}\cdot\mathbf{a}).\end{aligned}$$

Hence

$$\mathbf{d}\cdot\mathbf{c}\cdot\mathbf{b}|\mathbf{a}=F[F(\mathbf{d}|\mathbf{c}\cdot\mathbf{b}\cdot\mathbf{a},\ \mathbf{c}|\mathbf{b}\cdot\mathbf{a}),\ \mathbf{b}|\mathbf{a}].$$

Also

$$\begin{aligned}\mathbf{d}\cdot\mathbf{c}\cdot\mathbf{b}|\mathbf{a}&=[\text{by Eq. (4)}]\mathbf{d}\cdot(\mathbf{c}\cdot\mathbf{b})|\mathbf{a}\\&=[\text{by Eq. (7)}]F[\mathbf{d}|(\mathbf{c}\cdot\mathbf{b})\cdot\mathbf{a},\ \mathbf{c}\cdot\mathbf{b}|\mathbf{a}]\\&=[\text{by Eqs. (4) and (7)}]\\&\qquad F[\mathbf{d}|\mathbf{c}\cdot\mathbf{b}\cdot\mathbf{a},\ F(\mathbf{c}|\mathbf{b}\cdot\mathbf{a},\ \mathbf{b}|\mathbf{a})].\end{aligned}$$

Equating these two expressions for $\mathbf{d}\cdot\mathbf{c}\cdot\mathbf{b}|\mathbf{a}$ and, for simplicity, letting $\mathbf{d}|\mathbf{c}\cdot\mathbf{b}\ \mathbf{a}=x$, $\mathbf{c}|\mathbf{b}\cdot\mathbf{a}=y$, and $\mathbf{b}|\mathbf{a}=z$, we have

$$F[F(x,y),z]=F[x,F(y,z)]. \tag{8}$$

The function F must be such as to satisfy Eq. (8) for arbitrary values of x, y and z. It is easily shown by substitution that this equation is satisfied if

$$Cf[F(p,q)]=f(p)f(q),$$

where f is an arbitrary function of a single variable, and C is an arbitrary constant. It is shown in the appendix that this is also the general solution, provided F has continuous second derivatives We have then

$$Cf(\mathbf{c}\cdot\mathbf{b}|\mathbf{a})=f(\mathbf{c}|\mathbf{b}\cdot\mathbf{a})f(\mathbf{b}|\mathbf{a}).$$

The choice of the function f is purely a matter of convention. For it has already been pointed out that, if $\mathbf{b}|\mathbf{a}$ is a measure of the credibility of

---

[10] Keynes has traced the use of such a symbol to H. McColl, *Proc. Lond Math. Soc.* **11**, 113 (1880). McColl uses the symbol x_a for the probability of the proposition x on the hypothesis a.

[11] H. Margenau, *Am. J. Physics* **10**, 224 (1942). R. A. Fisher has used the term in a quite different sense.

$\mathbf{b}$ on the hypothesis $\mathbf{a}$, then so also is $f(\mathbf{b}|\mathbf{a})$. We might then continue the discussion with $f(\mathbf{b}|\mathbf{a})$ as the symbol of likelihood in place of $\mathbf{b}|\mathbf{a}$ and never have to specify the function f. But this would give two symbols where one would be enough. As a matter of convenience, therefore, we write

$$C\mathbf{c}\cdot\mathbf{b}|\mathbf{a}=\mathbf{c}|\mathbf{b}\cdot\mathbf{a}\ \mathbf{b}|\mathbf{a}. \qquad (9)$$

This is, of course, the same as choosing the function f to make $f(\mathbf{b}|\mathbf{a})=\mathbf{b}|\mathbf{a}$. Since the choice was conventional, it follows that another choice could have been made. We might, for example, have let $f(\mathbf{b}|\mathbf{a})=\exp(\mathbf{b}|\mathbf{a})$, whence it would have followed that the likelihood of $\mathbf{c}\cdot\mathbf{b}$ was, except for an arbitrary additive constant, equal to the sum of the likelihoods which determine it. This would have given us a likelihood related to the one we have as entropy is related to thermodynamic probability in statistical mechanics. It would have been an allowable choice, but a less convenient one than that which was made.

If in Eq. (9) we let $\mathbf{c}=\mathbf{b}$, and note that $\mathbf{b}\cdot\mathbf{b}=\mathbf{b}$, by Eq. (3), we obtain, after dividing by $\mathbf{b}|\mathbf{a}$,

$$C=\mathbf{b}|\mathbf{b}\cdot\mathbf{a}.$$

Thus we see that when the hypothesis includes the conclusion the likelihood has the constant value C, whatever the propositions may be. This is what we should expect, since $\mathbf{b}$ is certain on the hypothesis $\mathbf{b}\cdot\mathbf{a}$, and we do not recognize degrees of certainty.

The value to be assigned to C, the likelihood of certainty, is purely conventional. If it is desired to make the likelihoods with which we are dealing correspond as nearly as possible to ordinary probabilities, then C will be given the value 1. Other choices are often made, especially in conversation. The phrase "one chance in a hundred" may be taken to mean unit likelihood on a scale in which certainty is represented by 100. Statements that have the form of assertions about numbers in an ensemble may be merely convenient ways of stating likelihoods on a scale chosen for its aptness to the question considered. In a general discussion the most convenient value for C is unity, and we therefore write Eq. (9) in the form

$$\mathbf{c}\cdot\mathbf{b}|\mathbf{a}=\mathbf{c}|\mathbf{b}\cdot\mathbf{a}\ \mathbf{b}|\mathbf{a}. \qquad (10)$$

This has the same form as the ordinary rule for the probability of two events. However, it does not make our likelihood correspond uniquely to the ordinary probability. For Eq. (10) raised to any power m is

$$(\mathbf{c}\cdot\mathbf{b}|\mathbf{a})^m=(\mathbf{c}|\mathbf{b}\cdot\mathbf{a})^m(\mathbf{b}|\mathbf{a})^m.$$

Thus any power of our likelihood satisfies an equation of the same form as Eq. (10) and corresponds equally well to the ordinary probability.

Next to be sought is a second assumption of probable inference, which is to provide a relation between the likelihoods of the propositions $\mathbf{b}$ and $\sim\mathbf{b}$ on the same hypothesis $\mathbf{a}$. Since $\sim\mathbf{b}$ is determined when $\mathbf{b}$ is specified, a reasonable assumption, and the least restrictive possible, appears to be that $\sim\mathbf{b}|\mathbf{a}$ is determined by $\mathbf{b}|\mathbf{a}$, or

$$\sim\mathbf{b}|\mathbf{a}=S(\mathbf{b}|\mathbf{a}), \qquad (11)$$

where S is some function of a single variable.

By Eq. (1), $\sim\sim\mathbf{b}|\mathbf{a}=\mathbf{b}|\mathbf{a}$, and therefore $S[S(\mathbf{b}|\mathbf{a})]=\mathbf{b}|\mathbf{a}$. Thus S must be such a function that

$$S[S(x)]=x, \qquad (12)$$

where x may have any possible value of a likelihood between those of certainty and impossibility. This does not impose enough restriction on S to be of much use by itself. Another functional equation may be obtained by considering $S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})$; thus,

$$S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})=\sim(\mathbf{c}\vee\mathbf{b})|\mathbf{a}$$
$$=[\text{by Eq. } (5')]\sim\mathbf{c}\cdot\sim\mathbf{b}|\mathbf{a}.$$

We wish to eliminate the propositions $\sim\mathbf{c}$ and $\sim\mathbf{b}$, so as to obtain an equation in the propositions $\mathbf{c}$, $\mathbf{b}$ and $\mathbf{a}$ and the function S. First we eliminate $\sim\mathbf{c}$.

$$\sim\mathbf{c}\cdot\sim\mathbf{b}|\mathbf{a}=[\text{by Eq. } (10)]\sim\mathbf{c}|\sim\mathbf{b}\cdot\mathbf{a}\ \sim\mathbf{b}|\mathbf{a}$$
$$=[\text{by Eq. } (11)]S(\mathbf{c}|\sim\mathbf{b}\cdot\mathbf{a})S(\mathbf{b}|\mathbf{a}).$$

Thus we have

$$S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})=S(\mathbf{c}|\sim\mathbf{b}\cdot\mathbf{a})S(\mathbf{b}|\mathbf{a}),$$

or

$$S(\mathbf{c}|\sim\mathbf{b}\cdot\mathbf{a})=S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})/S(\mathbf{b}|\mathbf{a}).$$

Taking the function S of both sides of this equation and using Eq. (12), we obtain

$$\mathbf{c}|\sim\mathbf{b}\cdot\mathbf{a}=S[S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})/S(\mathbf{b}|\mathbf{a})]. \qquad (13)$$

Next we eliminate $\sim\mathbf{b}$:

$$\begin{aligned}\mathbf{c}|\sim\mathbf{b}\cdot\mathbf{a} &= [\text{by Eq. (10)}]\mathbf{c}\cdot\sim\mathbf{b}|\mathbf{a}/\sim\mathbf{b}|\mathbf{a}\\ &= [\text{by Eq. (2)}]\sim\mathbf{b}\cdot\mathbf{c}|\mathbf{a}/\sim\mathbf{b}|\mathbf{a}\\ &= [\text{by Eq. (10)}]\sim\mathbf{b}|\mathbf{c}\cdot\mathbf{a}\,\mathbf{c}|\mathbf{a}/\sim\mathbf{b}|\mathbf{a}\\ &= [\text{by Eq. (11)}]S(\mathbf{b}|\mathbf{c}\cdot\mathbf{a})\mathbf{c}|\mathbf{a}/S(\mathbf{b}|\mathbf{a}).\end{aligned}$$

Therefore we may write in place of Eq. (13),

$$S(\mathbf{b}|\mathbf{c}\cdot\mathbf{a})\mathbf{c}|\mathbf{a}/S(\mathbf{b}|\mathbf{a}) = S[S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})/S(\mathbf{b}|\mathbf{a})].$$

It is convenient now to have $\mathbf{a}$ as the common hypothesis in all the likelihoods. We note that

$$\begin{aligned}\mathbf{b}|\mathbf{c}\cdot\mathbf{a} &= [\text{by Eq. (10)}]\mathbf{b}\cdot\mathbf{c}|\mathbf{a}/\mathbf{c}|\mathbf{a}\\ &= [\text{by Eq. (2)}]\mathbf{c}\cdot\mathbf{b}|\mathbf{a}/\mathbf{c}|\mathbf{a}.\end{aligned}$$

Substituting this expression in the preceding equation and multiplying both sides by $S(\mathbf{b}|\mathbf{a})$, we obtain

$$\begin{aligned}&S(\mathbf{c}\cdot\mathbf{b}|\mathbf{a}/\mathbf{c}|\mathbf{a})\mathbf{c}|\mathbf{a}\\ &\quad = S[S(\mathbf{c}\vee\mathbf{b}|\mathbf{a})/S(\mathbf{b}|\mathbf{a})]S(\mathbf{b}|\mathbf{a}). \qquad (14)\end{aligned}$$

This equation must hold for arbitrary meanings of the propositions $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. Let $\mathbf{b}=\mathbf{c}\cdot\mathbf{d}$. Then

$$\mathbf{c}\vee\mathbf{b} = \mathbf{c}\vee(\mathbf{c}\cdot\mathbf{d}) = [\text{by Eq. (6')}]\mathbf{c},$$

and

$$\begin{aligned}\mathbf{c}\cdot\mathbf{b} &= \mathbf{c}\cdot(\mathbf{c}\cdot\mathbf{d})\\ &= [\text{by Eq. (4)}](\mathbf{c}\cdot\mathbf{c})\cdot\mathbf{d} = [\text{by Eq. (3)}]\mathbf{c}\cdot\mathbf{d}.\end{aligned}$$

Making these substitutions in Eq. (14), we obtain

$$S(\mathbf{c}\cdot\mathbf{d}|\mathbf{a}/\mathbf{c}|\mathbf{a})\mathbf{c}|\mathbf{a} = S[S(\mathbf{c}|\mathbf{a})/S(\mathbf{c}\cdot\mathbf{d}|\mathbf{a})]S(\mathbf{c}\cdot\mathbf{d}|\mathbf{a}).$$

This may be written in a highly symmetric form if we let $\mathbf{c}|\mathbf{a}=x$ and $S(\mathbf{c}\cdot\mathbf{d}|\mathbf{a})=y$, and make use of the fact that $\mathbf{c}\cdot\mathbf{d}|\mathbf{a}=[\text{by Eq. (12)}]\, S[S(\mathbf{c}\cdot\mathbf{d}|\mathbf{a})]=S(y)$. In these terms we have

$$xS[S(y)/x] = yS[S(x)/y]. \qquad (15)$$

This equation must be satisfied by the function S for all of the values of x and y obtainable by arbitrarily varying the propositions $\mathbf{c}$, $\mathbf{d}$ and $\mathbf{a}$. If the function S is twice differentiable, the solution of Eq. (15) together with Eq. (12) is, as shown in the appendix,

$$S(p) = (1-p^m)^{1/m},$$

where m is an arbitrary constant. Hence by Eq. (11),

$$(\mathbf{b}|\mathbf{a})^m + (\sim\mathbf{b}|\mathbf{a})^m = 1.$$

Now, whatever the value of m, if $\mathbf{b}|\mathbf{a}$ measures the credibility of $\mathbf{b}$ on the hypothesis $\mathbf{a}$, then so also will $(\mathbf{b}|\mathbf{a})^m$. It has already been pointed out that $(\mathbf{b}|\mathbf{a})^m$ may replace $\mathbf{b}|\mathbf{a}$ in Eq. (10). Therefore we may take $(\mathbf{b}|\mathbf{a})^m$ as the symbol of likelihood without being under any necessity of assigning a value to m. This is the same as to say that the choice of a value for m is purely conventional. For simplicity of notation we let $m=1$ and write

$$\mathbf{b}|\mathbf{a} + \sim\mathbf{b}|\mathbf{a} = 1. \qquad (16)$$

This has the same form as the ordinary rule relating the probability of $\sim\mathbf{b}$ to that of $\mathbf{b}$, or, as it is usually said, the rule for the probability that an event will not occur, given the probability that it will occur.

If in Eq. (16) we let $\mathbf{b}=\mathbf{a}$, then

$$\mathbf{a}|\mathbf{a} + \sim\mathbf{a}|\mathbf{a} = 1.$$

The two likelihoods are now those of certainty and impossibility. Since certainty has been given the likelihood 1, it now follows that impossibility has the likelihood zero.

Two other useful theorems are easily obtained. By Eq. (10),

$$\mathbf{c}\cdot\mathbf{b}|\mathbf{a} + \sim\mathbf{c}\cdot\mathbf{b}|\mathbf{a} = (\mathbf{c}|\mathbf{b}\cdot\mathbf{a} + \sim\mathbf{c}|\mathbf{b}\cdot\mathbf{a})\mathbf{b}|\mathbf{a}.$$

By Eq. (16),

$$\mathbf{c}|\mathbf{b}\cdot\mathbf{a} + \sim\mathbf{c}|\mathbf{b}\cdot\mathbf{a} = 1.$$

Therefore,

$$\mathbf{c}\cdot\mathbf{b}|\mathbf{a} + \sim\mathbf{c}\cdot\mathbf{b}|\mathbf{a} = \mathbf{b}|\mathbf{a}, \qquad (17)$$

This is one of the theorems. The other is obtained as follows.

$$\begin{aligned}\mathbf{c}\vee\mathbf{b}|\mathbf{a} &= [\text{by Eq. (16)}]1-\sim(\mathbf{c}\vee\mathbf{b})|\mathbf{a}\\ &= [\text{by Eq. (5')}]1-\sim\mathbf{c}\cdot\sim\mathbf{b}|\mathbf{a}\\ &= [\text{by Eq. (17)}]1-\sim\mathbf{b}|\mathbf{a}+\mathbf{c}\cdot\sim\mathbf{b}|\mathbf{a}.\end{aligned}$$

Now, by Eq. (16), $1-\sim\mathbf{b}|\mathbf{a}=\mathbf{b}|\mathbf{a}$. Also,

$$\begin{aligned}\mathbf{c}\cdot\sim\mathbf{b}|\mathbf{a} &= [\text{by Eq. (2)}]\sim\mathbf{b}\cdot\mathbf{c}|\mathbf{a}\\ &= [\text{by Eq. (17)}]\mathbf{c}|\mathbf{a}-\mathbf{b}\cdot\mathbf{c}|\mathbf{a}\\ &= [\text{by Eq. (2)}]\mathbf{c}|\mathbf{a}-\mathbf{c}\cdot\mathbf{b}|\mathbf{a}.\end{aligned}$$

Therefore,

$$\mathbf{c}\vee\mathbf{b}|\mathbf{a} = \mathbf{c}|\mathbf{a}+\mathbf{b}|\mathbf{a}-\mathbf{c}\cdot\mathbf{b}|\mathbf{a}, \qquad (18)$$

which has the same form as the ordinary rule for the probability that at least one of two events will occur. When, as is often done, the rule is stated for mutually exclusive events, the last term in the right-hand member does not appear. This conceals the rather interesting symmetry of the equation between the propositions $\mathbf{c} \vee \mathbf{b}$ and $\mathbf{c} \cdot \mathbf{b}$.

The indefiniteness of the concept of likelihood, defined only as a measure of reasonable credibility, has been removed by the conventions which have been adopted. The symbol $\mathbf{b} | \mathbf{a}$ stands now for a particular measure of credibility. Since this measure has been shown to be subject to the ordinary rules of probability, it is appropriate to call it the probability of the proposition **b** on the hypothesis **a**, discarding the term likelihood, which was less definitely defined.

The rules obtained, being only relations between probabilities, do not of themselves assign numerical values to all the probabilities arising in specific problems. The only numerical values thus far obtained are those corresponding to certainty and impossibility, and these were assigned by convention rather than required by the rules of symbolic logic. It is hardly to be supposed that every reasonable expectation should have a precise numerical value. In a number of cases, however, the familiar rule of insufficient reason may be employed. If there are n propositions of which, with respect to a given hypothesis, one and no more than one can be true, and if the hypothesis gives no reason for considering any one of them more likely than another, then, by the rules obtained, each of them has the probability $1/n$.

Probability and Frequency

The whole discussion thus far has consisted of two parts. The first part was intended to show that the rules of probable inference are credited by common sense with a wider validity than can be established by deducing them from the frequency definition of probability. In the second part they were derived without reference to this definition, from rather elementary postulates. It remains now to see what is the connection between probability, as here understood, and the frequency of an event.

Let us suppose that two capsules contain equal masses of radon, but that the contents are of different ages, one having been produced by the very recent decay of radium, and the other having been drawn from a vessel in which radon has been accumulating for a long time over radium in solution. Suppose there are two identical ion counters, each receiving radiation from one capsule, and each placed with respect to its capsule in the same relative position as the other. One of the two capsules will be the first to cause 1000 discharges in its ion counter. More than one hypothesis will ascribe to each capsule the same probability of being first. A physicist will estimate equal probabilities on the ground of the many observations which have been made on rates of radioactive transformation, with some additional evidence from quantum mechanics that the stability of such an aggregate of elementary particles as an atomic nucleus is independent of its age. Another person, quite unfamiliar with all this, will estimate equal probabilities merely on the ground that he does not know which capsule contains the older radon and has therefore no possible reason to suppose that one sample rather than the other will be the first to cause 1000 discharges.

These are the extreme cases, the first estimate being highly significant and the second quite trivial, but each is right on the hypothesis given. Kemble calls the estimate of the physicist one of *objective* probability and that of the other person one of *subjective*, or *primary*, probability. The latter term seems preferable to me, as it does to him. It is true that the estimate of the nonphysicist is subjective in the sense that it is relative to his limited information, but it is objective in the sense that another person with the same information would reasonably make the same estimate. It seems questionable whether there is a real difference in the *kind* of judgment made by the nonphysicist and the physicist. The nonphysicist bases his estimate on the fact that the capsules are indistinguishable. The physicist bases *his* estimate on the accumulated evidence that the atoms of radon themselves are indistinguishable. The difference seems to be not so much a difference in the nature of the evidence as in its amount and relevance or, to use Keynes' suggestive term, its weight.

Now let the experiment with the radon capsules be tried a number of times, with the old and new samples identified in advance of each trial. Even a long run of instances in which the older sample is first will not change the probabilities as estimated by the physicist. The evidence on which he made his first estimate had so much weight that no additional number of instances, not enormously large, could require a new estimate. The probabilities have for practical purposes become stable. Strictly speaking, since probability is relative to an experience that is never complete, it is always subject to change by new experience. A stable probability is a limit that is not strictly attainable, but that can in certain cases be approximated as nearly as necessary for practical use. It is to be expected that a stable probability will give a better basis for prediction than will an unstable one.

Let **a** be a hypothesis of which a number of instances may be examined. Let $\mathbf{b}_r$ mean that a certain proposition **b** is valid in the rth instance of **a**. Unless the hypothesis **a** itself assigns a stable probability to **b**, then $\mathbf{b}_s|\mathbf{a}$, $\mathbf{b}_s|\mathbf{a}\cdot\mathbf{b}_r$ and $\mathbf{b}_s|\mathbf{a}\cdot\sim\mathbf{b}_r$ will generally all be different; the knowledge that **b** is valid or that it is not valid in one instance will affect the reasonable expectation of its validity in another instance. But now let there be included in the hypothesis a proposition **p**, which asserts that the probability is stable and equal to p, some number between 0 and 1. This means that the probability of $\mathbf{b}_s$ is the same whether $\mathbf{b}_r$ or $\sim\mathbf{b}_r$ or neither is included in the hypothesis. Thus

$$\mathbf{b}_s|\mathbf{a}\cdot\mathbf{p}\cdot\mathbf{b}_r=\mathbf{b}_s|\mathbf{a}\cdot\mathbf{p}\cdot\sim\mathbf{b}_r=\mathbf{b}_s|\mathbf{a}\cdot\mathbf{p}=p.$$

Then by Eqs. (10) and (16) we obtain

$$\mathbf{b}_s\cdot\mathbf{b}_r|\mathbf{a}\cdot\mathbf{p}=p^2,$$
$$\mathbf{b}_s\cdot\sim\mathbf{b}_r|\mathbf{a}\cdot\mathbf{p}=p(1-p),$$
$$\sim\mathbf{b}_s\cdot\sim\mathbf{b}_r|\mathbf{a}\cdot\mathbf{p}=(1-p)^2.$$

Let $\mathbf{n}_N$ mean that the number of instances of **b** in N instances of **a** is exactly n. Then by Eqs. (10), (16) and (18) it is possible to derive the well-known result of Bernoulli, that

$$\mathbf{n}_N|\mathbf{a}\cdot\mathbf{p}=p^n(1-p)^{N-n}N!/n!(N-n)!.$$

This is a maximum when $p=n/N$, and the maximum becomes sharper as N is increased. Thus, when there is a stable probability, the frequency may confidently be expected to approach it as a limit.

There will sometimes be questions in which the existence of a stable probability is known but its value is undetermined. As a rather artificial but simple example, let it be supposed that there are two dice, both dynamically symmetric, but one of them defectively marked, having two faces instead of one stamped with four dots. Then for either of these dice there is a stable probability of throwing a four, equal to $\frac{1}{6}$ if it is the true die and to $\frac{1}{3}$ if it is the defective one. Suppose one die of the pair is picked up at random and, without being examined, is tossed N times. If a four turns up on n of these throws, what is the probability of a four on the next throw?

The problem may be generalized as follows. Let it be supposed that in the ensemble of instances of a proposition **a**, another proposition **b** is known to have a stable probability, but the value of this stable probability is unknown. As before, **p** will denote the proposition that the probability is stable and equal to a number p, but in the present case **p** is not a part of the hypothesis. Instead, the hypothesis contains a weaker proposition which only assigns a probability to the proposition **p** corresponding to every value of p.

We may let the single symbol **a** represent the entire initial hypothesis, including this proposition. Thus, in the example of the dice, **a** will describe the two dice and will also assert that one is chosen at random and tossed without being identified. There are then only two possible stable probabilities, $\frac{1}{6}$ and $\frac{1}{3}$, and they are equally probable at the beginning. Hence in this example $\mathbf{p}|\mathbf{a}$ has the value $\frac{1}{2}$ if p is either $\frac{1}{6}$ or $\frac{1}{3}$ and the value zero if p is any other number.

Returning to the general problem, we suppose that N instances of **a** are observed, and **b** is found valid in n of them and invalid in the rest. What is now the probability of **b** in the $N+1$th instance of **a**? As before, let $\mathbf{n}_N$ denote the proposition that n is the number of instances of **b** in N instances of **a**. The problem is to find $\mathbf{b}_{N+1}|\mathbf{a}\cdot\mathbf{n}_N$, given n and N, and also $\mathbf{p}|\mathbf{a}$ for every value of p.

The theorems available are enough to give the result

$$\mathbf{b}_{N+1}|\mathbf{a}\cdot\mathbf{n}_N=\frac{\sum p^{n+1}(1-p)^{N-n}\mathbf{p}|\mathbf{a}}{\sum p^n(1-p)^{N-n}\mathbf{p}|\mathbf{a}},$$

where the summations are over-all values of p.

If, on the hypothesis **a**, the stable probability has a continuous range of possible values from 0 to 1, and if $\mathrm{f}(p)\mathrm{d}p$ denotes the probability of a value between p and $p+\mathrm{d}p$, the summations are replaced by integrals, and we have

$$\mathbf{b}_{N+1}|\mathbf{a}\cdot\mathbf{n}_N=\frac{\int_0^1 p^{n+1}(1-p)^{N-n}\mathrm{f}(p)\mathrm{d}p}{\int_0^1 p^n(1-p)^{N-n}\mathrm{f}(p)\mathrm{d}p}.$$

It was assumed by Laplace that an unknown probability is equally likely to have any value from 0 to 1. On this assumption, $\mathrm{f}(p)$ is constant in the last equation. The integrals in this case are known; and the result, sometimes called the rule of succession, is simply

$$\mathbf{b}_{N+1}|\mathbf{a}\cdot\mathbf{n}_N=(n+1)/(N+2),$$

or approximately, for large values of n and N,

$$\mathbf{b}_{N+1}|\mathbf{a}\cdot n_N=n/N.$$

Several authors have pointed out that an unknown stable probability is not necessarily one for which all values from 0 to 1 are equally likely, and the rule of succession has been shown to lead to some absurd results. Nevertheless, we should expect that for large numbers it will generally be right. For we know, by the theorem of Bernoulli given earlier, that, when there is a stable probability, the ratio n/N is very likely to be nearly equal to it when N is large. Also we have understood a stable probability to be the limit that the probability approaches as the weight of the evidence is increased, and usually the surest way to increase the weight of evidence is to increase the number of observed instances. If the ratio n/N and the probability approach a common limit, then certainly they must approach each other.

If the absurdities to which Laplace's rule has led are examined, they are found to fall into three classes: those in which N is not a very large number, those in which $n/N=1$, and those in which $n/N=0$. (The last two are really one class, since to say that **b** is valid in N out of N instances of **a** is the same as to say that $\sim$**b** is valid in none of N instances.) If these conditions are excluded, Laplace's rule may be derived from a much less drastic assumption than the assumption that all values of the stable probability are equally likely.

If from the general equation for $\mathbf{b}_{N+1}|\mathbf{a}\cdot\mathbf{n}_N$ we eliminate n by letting $n/N=\nu$, we obtain

$$\mathbf{b}_{N+1}|\mathbf{a}\cdot\mathbf{n}_N=\frac{\int_0^1 p[p^\nu(1-p)^{1-\nu}]^N\mathrm{f}(p)\mathrm{d}p}{\int_0^1 [p^\nu(1-p)^{1-\nu}]^N\mathrm{f}(p)\mathrm{d}p}.$$

If $0<\nu<1$, then $p^\nu(1-p)^{1-\nu}$ has a maximum value when $p=\nu$. The Nth power of this expression, when N is large enough, will have so pronounced a maximum that its values when p is more than slightly different from ν will be relatively negligible. Hence, unless $\mathrm{f}(p)$ is extremely small when $p=\nu$, its only values of importance in the integrals will be those for which p and ν are very nearly equal. Therefore, unless $\mathrm{f}(p)$ is rapidly varying around this point, it may be replaced in the integrals by the constant $\mathrm{f}(\nu)$.

Thus we arrive again at Laplace's rule. Its generality is much less than Laplace supposed. But it serves to show how a probability approaches stability as the number of instances is increased, and this is all we should expect of it.[12]

* * *

Professor K. O. Friedrichs, of New York University, read a preliminary draft of this paper. I wish to thank him for this kindness and for his help in correcting some mathematical inaccuracies. He is not responsible, of course, for any errors that remain or for the opinion expressed as to the nature of probability.

Appendix: The Solution of the Functional Equations

The first equation to be solved is

$$\mathrm{F}[\mathrm{F}(x,y),z]=\mathrm{F}[x,\mathrm{F}(y,z)]. \tag{8}$$

Let $\mathrm{F}(x,y)=u$, and let $\mathrm{F}(y,z)=v$. Then Eq. (8) becomes $\mathrm{F}(u,z)=\mathrm{F}(x,v)$. Differentiating this with respect to x, y and z in turn, and writing $\mathrm{F}_1(p,q)$ for $\partial\mathrm{F}(p,q)/\partial p$ and $\mathrm{F}_2(p,q)$ for $\partial\mathrm{F}(p,q)/\partial q$, we obtain

$$\mathrm{F}_1(u,z)\partial u/\partial x=\mathrm{F}_1(x,v), \tag{19}$$

$$\mathrm{F}_1(u,z)\partial u/\partial y=\mathrm{F}_2(x,v)\partial v/\partial y, \tag{20}$$

$$\mathrm{F}_2(u,z)=\mathrm{F}_2(x,v)\partial v/\partial z. \tag{21}$$

Differentiating Eq. (20) with respect to x, y and z in turn,

[12] The problem of inverse probability when $n/N=1$ (or 0), which is important in the application of probability to inductive reasoning, is discussed at length by Jeffreys in reference 8.

writing $F_{11}(p, q)$ for $\partial F_1(p, q)/\partial p$, and similarly representing the other second derivatives, we obtain

$$F_{11}(u, z)(\partial u/\partial x)(\partial u/\partial y)+F_1(u, z)\partial^2 u/\partial x\partial y = F_{12}(x, v)\partial v/\partial y, \quad (22)$$

$$F_{11}(u, z)(\partial u/\partial y)^2+F_1(u, z)\partial^2 u/\partial y^2 = F_{22}(x, v)(\partial v/\partial y)^2+F_2(x, v)\partial^2 v/\partial y^2, \quad (23)$$

$$F_{12}(u, z)\partial u/\partial y = F_{22}(x, v)(\partial v/\partial y)(\partial v/\partial z) + F_2(x, v)\partial^2 v/\partial y\partial z. \quad (24)$$

Differentiating Eq. (19) with respect to z, or Eq. (21) with respect to x, we obtain

$$F_{12}(u, z)\partial u/\partial x = F_{12}(x, v)\partial v/\partial z. \quad (25)$$

Among Eqs. (20), (22), $\cdots$, (25) we can now eliminate the functions of u and v other than their derivatives. Thus, eliminating $F_{12}(u, z)$ and $F_{12}(x, v)$ among Eqs. (22), (24) and (25), we find

$$[F_{11}(u, z)(\partial u/\partial y)^2 - F_{22}(x, v)(\partial v/\partial y)^2](\partial u/\partial x)(\partial v/\partial z) = F_2(x, v)(\partial^2 v/\partial y\partial z)(\partial v/\partial y)(\partial u/\partial x) - F_1(u, z)(\partial^2 u/\partial x\partial y)(\partial u/\partial y)(\partial v/\partial z)$$

Combining this with Eq. (23), we can eliminate $F_{11}(u, z)$ and $F_{22}(x, v)$ together, obtaining

$$F_1(u, z)(\partial v/\partial z)[(\partial^2 u/\partial y^2)(\partial u/\partial x) - (\partial^2 u/\partial x\partial y)(\partial u/\partial y)] = F_2(x, v)(\partial u/\partial x)[(\partial^2 v/\partial y^2)(\partial v/\partial z) - (\partial^2 v/\partial y\partial z)(\partial v/\partial y)].$$

Combining this with Eq. (20), we eliminate $F_1(u, z)$ and $F_2(x, v)$ together and obtain

$$\frac{\partial^2 u/\partial x\partial y}{\partial u/\partial x} - \frac{\partial^2 u/\partial y^2}{\partial u/\partial y} = \frac{\partial^2 v/\partial y\partial z}{\partial v/\partial z} - \frac{\partial^2 v/\partial y^2}{\partial v/\partial y}.$$

This may be written in the form,

$$\frac{\partial}{\partial y}\ln\left(\frac{\partial u/\partial x}{\partial u/\partial y}\right) = -\frac{\partial}{\partial y}\ln\left(\frac{\partial v/\partial y}{\partial v/\partial z}\right).$$

Now $u = F(x, y)$ and $v = F(y, z)$, so that

$$\frac{\partial u/\partial x}{\partial u/\partial y} = \frac{F_1(x, y)}{F_2(x, y)}, \quad \text{and} \quad \frac{\partial v/\partial y}{\partial v/\partial z} = \frac{F_1(y, z)}{F_2(y, z)}.$$

We have then,

$$\frac{\partial}{\partial y}\ln\left[\frac{F_1(x, y)}{F_2(x, y)}\right] = -\frac{\partial}{\partial y}\ln\left[\frac{F_1(y, z)}{F_2(y, z)}\right].$$

Since x appears only in the left-hand member and z only in the right-hand member of this equation, it follows that each member is a function only of the remaining variable y. It will be convenient to denote the integral of this function by $\ln \Phi(y)$, so that we have

$$\frac{\partial}{\partial y}\ln\left[\frac{F_1(x, y)}{F_2(x, y)}\right] = \frac{d}{dy}\ln \Phi(y), \quad (26)$$

and

$$\frac{\partial}{\partial y}\ln\left[\frac{F_1(y, z)}{F_2(y, z)}\right] = -\frac{d}{dy}\ln \Phi(y). \quad (27)$$

Permuting x, y and z in Eq. (27), we obtain

$$\frac{\partial}{\partial x}\ln\left[\frac{F_1(x, y)}{F_2(x, y)}\right] = -\frac{d}{dx}\ln \Phi(x). \quad (28)$$

Multiplying Eq. (28) by dx and Eq. (26) by dy and adding, we obtain

$$\frac{\partial}{\partial x}\ln\left[\frac{F_1(x, y)}{F_2(x, y)}\right]dx + \frac{\partial}{\partial y}\ln\left[\frac{F_1(x, y)}{F_2(x, y)}\right]dy = -d\ln \Phi(x) + d\ln \Phi(y).$$

The left-hand member being now a complete differential, we may integrate and so find

$$F_1(x, y)/F_2(x, y) = h\Phi(y)/\Phi(x), \quad (29)$$

where h is a constant of integration.

To make use of this result, we divide Eq. (20) by Eq. (21), obtaining

$$\frac{F_1(u, z)}{F_2(u, z)}\frac{\partial u}{\partial y} = \frac{\partial v/\partial y}{\partial v/\partial z}.$$

The right-hand member is simply $F_1(y, z)/F_2(y, z)$, and, with the aid of Eq. (29), the equation may be written as

$$\frac{\Phi(z)}{\Phi(u)}\frac{\partial u}{\partial y} = \frac{\Phi(z)}{\Phi(y)}.$$

Replacing in this equation u by its value $F(x, y)$ we have

$$\partial F(x, y)/\partial y = \Phi[F(x, y)]/\Phi(y). \quad (30)$$

Similarly, from Eqs. (19) and (20) we obtain

$$\partial F(y, z)/\partial y = \Phi[F(y, z)]/\Phi(y),$$

which becomes, when x and y are written for y and z,

$$\partial F(x, y)/\partial x = \Phi[F(x, y)]/\Phi(x). \quad (31)$$

Combining Eqs. (30) and (31) to obtain the differential dF (the variables being understood as x and y) we find

$$dF/\Phi(F) = dx/\Phi(x) + dy/\Phi(y).$$

If we denote $\int[dp/\Phi(p)]$ by $\ln f(p)$, we obtain, by integrating and taking the exponentials of both members of this equation,

$$Cf(F) = f(x)f(y),$$

where C is a constant of integration. This then is the solution of Eq. (8).

The solution of the equation

$$xS[S(y)/x] = yS[S(x)/y] \quad (15)$$

is obtained in a similar manner, but more quickly. This equation and the three derived from it by differentiation with respect to x, to y, and to x and y may be written as follows, when $S(y)/x$ is denoted by u and $S(x)/y$ by v:

$$xS(u) = yS(v), \quad (32)$$

$$uS'(u) - S(u) = -S'(v)S'(x), \quad (33)$$

$$S'(u)S'(y) = -vS'(v) + S(v), \quad (34)$$

$$uS''(u)S'(y)/x = vS''(v)S'(x)/y. \quad (35)$$

Multiplying Eq. (32) by Eq. (35), we eliminate x and y simultaneously, obtaining

$$uS''(u)S(u)S'(y) = vS''(v)S(v)S'(x).$$

With this equation, together with Eqs. (33) and (34), it is possible to eliminate $S'(x)$ and $S'(y)$. The result is the equation

$$\frac{uS''(u)S(u)}{[uS'(u)-S(u)]S'(u)} = \frac{vS''(v)S(v)}{[vS'(v)-S(v)]S'(v)}.$$

Since each member of the foregoing equation is the same function of a different variable, this function must be equal to a constant. Calling this constant k, we have

$$uS''(u)S(u)=k[uS'(u)-S(u)]S'(u).$$

This may be put in the form

$$dS'/S'=k(dS/S-du/u),$$

whence, by integration,

$$S'=A(S/u)^k,$$

where A is a constant.

The variables being separable, another integration gives

$$S^m=Au^m+B,$$

where m has been written for $1-k$, and B is a constant of integration. It is now found by substitution that Eq. (15) can be satisfied for arbitrary values of x and y only if $B=A^2$. Finally, if the solution of Eq. (15) is also to satisfy the equation $S[S(x)]=x$, it is necessary that $A=-1$. Thus we obtain $[S(u)]^m+u^m=1$.

Application of Group Theory to the Calculation of Vibrational Frequencies of Polyatomic Molecules[1]

Arnold G. Meister and Forrest F. Cleveland
Illinois Institute of Technology, Chicago, Illinois

WILSON[2] has devised a method for obtaining the vibrational frequencies of polyatomic molecules in which group theory is used to simplify the calculations. The method is especially good for molecules having considerable symmetry and several equivalent atoms, that is, atoms with identical nuclei that transform into one another for all operations of the point group of the molecule. A further advantage of the method is that it requires no coordinate system, but only bond distances, interbond angles and unit vectors directed along the bonds.

Since one who is beginning calculations of vibrational frequencies may find the symbolism of Wilson's papers difficult, and since other papers involving the method omit many of the details, it seems worth while to give an elementary treatment of a few typical molecules for those desiring to start work in this field. The H_2O molecule is considered first because it has only a small number of atoms, has no degenerate frequencies, and permits the reader to concentrate on the method without being confused by the complexity of the molecule.[3] Then the CH_3Cl, CH_4 and CD_4 molecules are treated to show how the method is applied when doubly or triply degenerate frequencies are present.[4]

THE H_2O MOLECULE

Symmetry Coordinates

The methods given in a previous paper[5] are used to determine the point group of the molecule as well as the number of fundamental vibrations of each type. It is found that the H_2O molecule belongs to the point group $\mathbf{C}_{2v}$ and that there are two vibrations of type A_1 and one of type B_2. Since a nonlinear molecule containing N atoms has $3N-6$ vibrational degrees of freedom, $3N-6$ coordinates are necessary to describe the vibrations of the molecule. To attain the simplification made possible by the use of group theory, it is

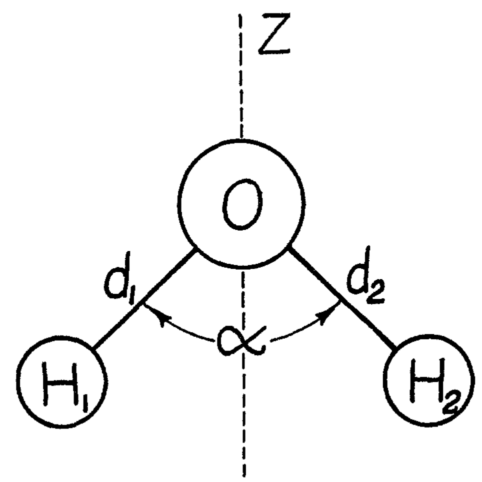


Fig. 1. Bond distances d_1 and d_2, interbond angle α, and principal symmetry axis Z for the H_2O molecule.

[1] Communication No. 43 from the *Spectroscopy Laboratory*.

[2] E. B. Wilson, Jr., *J. Chem. Physics* **7**, 1047 (1939); **9**, 76 (1941).

[3] A more complicated molecule, CH_2Cl_2, involving only nondegenerate frequencies has been discussed by G. Glockler, *Rev. Mod. Physics* **15**, 125 (1943).

[4] A treatment in outline form of the CH_3Cl molecule is given at the end of Wilson's second paper, reference 2.

[5] A. G. Meister, F. F. Cleveland and M. J. Murray, *Am. J. Physics* **11**, 239 (1943).

SOME CRITICISM OF FUZZY LOGIC

Quote from Zadeh's Obituary for Joe Goguen in *Fuzzy Sets and Systems*, 2007, 158(8), pp. 809-810:

My first paper on fuzzy sets was published in June 1965, while I held the position of the Chair of the Department of Electrical Engineering at UC Berkeley. Shortly after the publication of my paper, a student walked into my office and identified himself as Joe Goguen, a graduate student in mathematics. He told me that he did read my paper and was interested in developing the concept of a fuzzy set within the framework of category theory. We discussed his ideas for a while. At the end of our discussion, he asked me to become his research supervisor. I responded affirmatively, since it was quite obvious that Joe Goguen was not an average graduate student — he was a superior intellect. This meeting was the beginning of my lifelong relationship with Goguen. We met frequently to discuss various issues of fuzzy set theory. Nobody in the mathematics department took interest in his work.

Criticism of fuzzy sets by Rudolf Kalman at the International Conference on Man and Computer in Bordeaux, France, in 1972, after Zadeh introduced the concept of a linguistic variable:

I would like to comment briefly on Professor Zadeh's presentation. His proposal could be severely, ferociously, even brutally criticized from a technical point of view. This would be out of place here. But a blunt question remains: 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 political climate in the US — 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. I must confess that I cannot conceive of "fuzzification" as a viable alternative for the scientific method.

Another criticism of fuzzy sets made in 1975 by William Kahan, a mathematician at University of California-Berkeley:

Fuzzy theory is wrong, wrong, and pernicious. I cannot think of any problem that could not be solved by ordinary logic. ... What Zadeh is saying is the same sort of thing as, "Technology got us into this mess and now it can't get us out." Well, technology did not get us into this mess. Greed and weakness and ambivalence got us into this mess. 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 has brought us so much trouble.

Typical criticism expressing extreme claims of some probability theorists, especially those supporting the Bayesian methodology:

Our thesis is simply stated: *the only satisfactory description of uncertainty is probability*. By this is meant that every uncertainty statement must be in the form of 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. We speak of "the inevitability of probability." ... Our argument may be summarized by saying that probability is the only sensible description of uncertainty. All other methods are inadequate. ... My challenge that anything that can be done with fuzzy logic, belief functions, upper and lower probabilities, or any other alternative to probability, can better be done with probability, remains.

A peculiar criticism of fuzzy logic by Charles Elkan:

In July 1993, Charles Elkan presented a paper entitled "The paradoxical success of fuzzy logic" at the Eleventh National Conference on Artificial Intelligence in Washington, D.C., sponsored by the American Association for Artificial Intelligence (AAAI). The paper is also included in the Conference Proceedings (Elkan 1993). In this paper, Elkan claims that the apparent success of fuzzy logic in many practical applications is paradoxical since fuzzy logic collapses upon closer scrutiny into the classical, bivalent logic. To support this claim, he uses one definition and one theorem expressed in terms of the following notation: A,B denote assertions; $t(A), t(B) \in [0,1]$ denote the degrees of truth in A,B, respectively; and $\wedge, \vee, \neg$ denote logical connectives of conjunction, disjunction, and negation, respectively. The following is the definition (exactly as it appears in the paper), which is supposed to define a particular system of fuzzy logic:

Definition 1:

$$t(A \wedge B) = \min\{t(A), t(B)\}$$

$$t(A \vee B) = \max\{t(A), t(B)\}$$

$$t(\neg A) = 1 - t(A)$$

$t(A) = t(B)$ if A and B are logically equivalent.

The definition is followed by the following clarification of the term "logically equivalent": "In the last case of this definition, let "logically equivalent" mean equivalent according to the rules of classical two-valued propositional calculus". Next is the statement of Elkan's theorem, preceded by his short remark "Fuzzy logic is intended to allow an indefinite variety of numerical truth values. The result proved here is that only two different truth values are in fact possible in the formal system of Definition 1."

Theorem 1: For any two assertions A and B, either $t(A) = t(B)$ or $t(A) = 1 - t(B)$.

WELL-DOCUMENTED DEBATES

1. Probability theory versus evidence theory and other ways of dealing with uncertainty in AI: *Statistical Science* 2(1), 1987, 1-44.
2. Bayesian probabilistic approach to managing uncertainty in AI versus other approaches: *Computational Intelligence* (Canadian) 4(1), 1988, 57-142.
3. An AI view of the treatment of uncertainty: *The Knowledge Engineering Review* 2(1), 1988, 59-91
4. Cambridge debate: Bayesian approach to dealing with uncertainty versus other approaches: *Intern. J. of General Systems* 15(4), 1989, 347-378.
5. Fuzziness versus probability: *IEEE Trans on Fuzzy Systems* 2(1), 1994, 1-41.
6. The paradoxical success of fuzzy logic (Charles Elkan): *IEEE Expert* 9(4), August 1994, 2-49.
7. Probabilistic and statistical view of fuzzy methods: *Technometrics* 37(3), 1995, 249-292.

Title:The logic that dares not speak its name
Source: *The Economist (US).* 331.7859 (Apr. 16, 1994): p89.
Document Type:Article

Full Text:

A HOVERING helicopter is one of the world's more improbable sights. It teeters forever on the edge of instability; only continuous adjustments by the pilot keep it airborne. The pilot is able to do this because, during his training, he has learnt to "feel" what the machine is up to. He knows how to respond to each little pitch and yaw. Automating such a process is a nightmare, but Michio Sugeno, of the Tokyo Institute of Technology, has done it. Better still, to fly his remote-controlled helicopter you simply talk to it.

Putting the human voice in charge of what is already a difficult task may sound like unnecessary complication. In fact, it makes perfect sense. Dr Sugeno's machine is controlled by a system that was originally devised to mimic the imprecision of language: fuzzy logic.

Information technology is classically binary. On or off. There or not there. It would be nice to report that there are two reasons for this, but there are more. Discrete variables are easy to handle. They are not confusing. And a lot of mathematics and philosophy tells you how to handle them. Ever since Aristotle, a dominant theme in Western thinking has been the idea of the law of the excluded middle: that everything is, in principle at least, A or not-A. In the jargon of logicians, logic is two-valued: each statement has a "truth value" of one or zero.

Fuzzy is different. Fuzzy logicians recognise that in reality things are often a bit A and a bit non-A: a half-eaten apple is still, to a degree, an apple. They have developed new mathematical theory to handle the idea. Other logicians disagree. For industrialists, to whom an excluded middle is something you sweat for at a health farm, the debate is beside the point. Whether or not fuzziness is mathematically respectable, they see it as the tool they need for controlling complex processes. From washing machines to video cameras, from making markets to making cement, commerce is going fuzzy at the edges.

Fuzzy logic came, as you would expect, out of the University of California, Berkeley, in the 1960s. Lotfi Zadeh, an electrical engineer, observed that a lot of words are used as if they were mathematical concepts. Words such as "many" and "few" are like numbers. "Frequently" and "rarely" are like probabilities. "Very" and "somewhat" are like mathematical operators such as multiplication and division. Dr Zadeh thought these words were like mathematical concepts, but not the same as them, because they were fuzzy (as is the idea of being "like"). Philosophers such as Bertrand Russell and Max Black had tinkered with ideas of this sort, but Dr Zadeh put them in a form familiar to mathematicians by incorporating them into set theory.

Game, set and match?

A set is a collection of things, such as the socks in your chest of drawers, or the numbers between one and eight. Something can belong to more than one set; tennis socks, for instance, also belong to the set of sports clothes. But they are either in or out. Your tennis socks are not part of the set of your shirts, not even a little bit. That, at least, is classical set theory. And thanks to Russell and others, set theory is taken as the basis of mathematics in general, so that by meddling with it, you can change the way a lot of other mathematics, from logic onwards, is done.

Fuzzy set theory allows something to be partly in one set and partly in another. It answers a paradox first propounded by Zeno, an ancient Greek philosopher: if you remove the grains from a pile of sand one at a time, at what point does it cease to be a pile? The fuzzy answer is that it leaves the set of piles of sand as smoothly as the individual grains are taken away from it.

This sort of set theory is apt for describing things that vary continuously. Suppose you have two sets: hot and cold. Is air at 20 degrees C hot? Not really. Is it cold? Up to a point. Inventing the categories "warm" and "cool" helps, but not much. 20 degrees C might be right in the middle of "warm", but what about 27 degrees C? In a fuzzy world such questions, and the need for ever smaller sets, do not arise. 27 degrees C might be 10% warm and 40% hot. 0 degrees C, however, is 100% cold--so is -5 degrees C. Fuzzy sets allow the qualitative question "how hot?" to be answered mathematically.

At first, fuzzy set theory was an embarrassment to researchers. According to Dr Zadeh's adherents, "ordinary" set theory, and therefore all the mathematics that derive from it--ie, all mathematics to date--is actually just a special case of fuzzy set theory: one with the fuzziness turned down to zero so that the "membership value" of each element is either one or zero for any given set. This was widely denied by mathematicians. Many claimed that partial membership of a number of sets was simply a restatement of probability theory (it is 10% probable that 27 degrees C is warm; 40% probable that it is hot). And many natural scientists objected to the implication of "fuzziness" when it came to their results.

In fact, the scientists' objections are as much semantic as real. Philosophers have often been amazed that mathematics can describe the world. But the mathematical models themselves have to be tweaked and rounded to do their jobs. The resulting equations may give an adequate representation of reality, but precise they are not.

Fuzzy logicians say that the probability argument is based on a misapprehension. Probability measures the likelihood of something happening. Fuzziness measures the degree to which it is happening. Knowing how likely it is to rain tomorrow is not the same as knowing how heavily it is raining today. To emphasise the difference and reinforce fuzzy logic's claim to mathematical fundamentalism, Bart Kosko, of the University of Southern California, has recently derived some of the basic principles of probability theory from fuzzy sets.

Nevertheless there is still resistance from some quarters. A theorem devised by Charles Elkan, of the University of California, San Diego, suggests that fuzzy logic may not be all that it claims--that it collapses to a more traditional sort of two-valued logic under close scrutiny.

Such arguments might be the stuff of Nobel prizes, were they to be awarded for mathematics. What has caused non-academics to sit up and take notice of fuzzy logic is that, in the physical world, systems that incorporate it seem to work rather well.

Soft sell

Think again about the temperature of the air. Air conditioners are designed to keep it constant. Nature varies it continuously. Devices like air conditioners are generally run by mathematical models. Information is gathered (in this case from temperature sensors), processed in some way, and the results used to control something (such as the speed of a fan-motor). The maths may be "wired in", by being embedded in the components of a traditional, analogue, electrical circuit. Or the calculations may be performed by a digital computer program stored in a chip. Either way, the engineer who designs a machine has first to devise a model of what it is supposed to do.

This is difficult to do precisely, but lack of precision can exacerbate problems such as "hunting"--a tendency to overshoot or undershoot that is caused by the lag between a change in the motor speed and a change in the air temperature. It has to be damped by anticipating the change, which needs more complex maths, and lots of it. Yet cooling a room is a relatively simple task--there is only one input (the temperature) and one output (the motor speed). Most control processes have several inputs, and may have more than one output. Turning the relationships between them into explicit, accurate maths can be tricky.

Fuzzy systems do not need to be precise. The underlying model is not a string of differential equations, but a set of simple rules such as "if the air is hot, run the motor fast". "Hot" and "fast" are fuzzy sets. The fuzzy controller takes the temperature in the same way as a traditional air conditioner, but then "fuzzifies" it by working out how much it belongs to each fuzzy input set (see diagram). By applying the relevant rules to each input set, the membership of each fuzzy output set can be calculated. After this, the result is "defuzzified" into a single value that the motor can understand, using a process known as centroid averaging. And presto, the air conditioner responds.

The advantage of this imprecision is that it reacts well to tinkering. To build a mathematical model requires a good understanding of what is being modelled. Fuzzy logic does not. As long as the underlying rules are sensible, all that is needed is to tune the device by trial and error--changing the boundaries of the fuzzy sets until the desired outcome is achieved. It is also quicker than traditional model-building, cutting down costly development time.

Fuzzy logic was first picked up as an industrial tool by Hitachi in Japan. Seiji Yasunobu (the engineer who invented centroid defuzzification) used it in an automatic control system for the city of Sendai's subway system. Even in a country famed for the precision of its underground railways, Sendai's is impressive. Each train stops to within 7cm (3 inches) of the right spot on the platform. In addition, the trains travel more smoothly and use about 10% less energy than their human-controlled equivalents. The person in the driver's compartment is there for little more than reassurance.

Hitachi's success at Sendai set off a scramble to apply fuzzy logic to other equipment. At the head of the race was a company called Omron, in Kyoto, which acquired a series of patents registered by Takeshi Yamakawa of the Kyushu Institute of Technology in Fukuoka. In 1987 (the year the Sendai subway opened) Dr Yamakawa demonstrated a set of chips that could cope with large numbers of rules and defuzzify the results 100 times faster than conventional chips--fast enough to keep an inverted pendulum upright. (Sounds easy? Try balancing a ruler on your finger.)

Dr Yamakawa achieved this trick by eschewing digital technology in the chips themselves. Once he had worked out the rules, he designed the chips around them. Instead of wasting time turning them into ones and zeroes, his analogue chips worked directly with the signals from the system's sensors. Without time spent translating back and forth between analogue and digital signals, the processing was much quicker and control of unstable, rapidly changing systems became possible.

Since 1987, the market for machines and factories with fuzzy controllers has exploded. Japan's Ministry of International Trade and Industry estimates that, worldwide, it was worth $2 billion in 1991. Omron alone expects its 1995 sales of fuzzy devices to top YEN 100 billion (about $1 billion). The company now makes around 80 products, from counterfeit banknote detectors to blood-pressure monitors, that depend on fuzzy-logic control.

After a slow start, fuzzy control is taking off outside Japan too. Cement manufacturers worldwide are spending about $15m a year to equip their plants, following the example of the Danish cement works that was one of the first places to use fuzzy production control. The car industry is pretty fuzzy everywhere. American car makers expect to use fuzzy-logic gear worth about $6 billion by 1997. Plenty of American products, from petrol pumps to computer spreadsheets, embody fuzzy logic, but often masquerading as "human- like reasoning" or a "knowledge-based system". Marketing departments worry that customers are not yet ready to be told that fuzzy is good.

As the tasks to which fuzzy controllers are put become more sophisticated, so do the ways of working out the rules. Early controllers were tuned by hand. Now they are often linked to neural nets--computer programs that mimic nervous systems. These "learn" by reinforcing successful behaviour, much as a rat in a laboratory learns to associate performing a trick with receiving food.

Other fuzzy controllers are "expert systems"--attempts to embody the knowledge of a human expert in computer software. The idea of "downloading" a human lifetime's experience into a machine appeals to computer programmers. In some areas, such as assisting medical diagnosis, expert systems using conventional logic have been successfully designed. But expertise itself is intractable stuff. Often people cannot quantify just how they go about things. However, they can, if pushed, usually describe it. Such descriptions are the stuff of fuzzy logic.

Expertise can be as specialised as flying a helicopter, or it can be quite mundane. Driving, for instance. Car makers have already fuzzified their transmissions, brakes and suspensions. But now there are expert systems that can control the vehicle itself. In 1987 Nissan demonstrated the first fuzzy driver. It steered a car using the white lines on a road for guidance. In 1989 Dr Kosko and Seong-Gon Kong devised an expert system that could park an articulated lorry. The race is now on to build something that can drive itself in traffic, and take its owner safely home after a night on the town.

Beyond control

Negotiating rush-hour traffic and landing a helicopter may sound tricky, but true believers in fuzzy logic think that it is destined for greater things. They reckon that a system of logic that draws its inspiration from human language should be ideal for modelling the intelligence behind language.

Artificial intelligence (AI) is a tantalising prize. Many have grasped for it using traditional computing methods but it has remained beyond reach. Toshiro Terano, the director of the Laboratory for International Fuzzy Engineering Research (LIFE), in Yokohama, hopes that his team will do better. LIFE has dropped the study of fuzzy control in order to concentrate on fuzzy information processing and fuzzy computing. Anca Ralescu, a professor from the University of Cincinnati who is working as the institute's assistant director, is running a project that explores the relationship between images and language. One of its applications is face recognition, a classic example of a task that people find easy and machines find hard. AI researchers usually solve it using neural nets, which have proved good at recognising pictures. Dr Ralescu's team, though, has written programs which can recognise verbal descriptions as well. Tell one of them what a particular face looks like (big nose, long hair, round face) and it will analyse the images and pick out your choice from a couple of dozen possibilities. One of the programs is so sensitive that it can recognise moods from facial expressions.

Dr Ralescu's work is largely theoretical. A more practical example of LIFE's work is Satoru Fukami's attempt to build a foreign- exchange dealing model. Fuzzy expert systems for financial dealing already exist. Yamaichi Securities has developed one for stockmarket trading. So far, however, the models have relied solely on price information. The LIFE forex model will use numbers but will also use the sort of non-numerical information that markets thrive on, such as feelings about politics, to refine the outcome.

Perhaps the biggest prize that LIFE is pursuing is general-purpose fuzzy computing. Along with other computer research groups, it is seeking a way of overcoming that bane of programmers, the bug. Computer programs are fragile creatures, apt to fall apart if the slightest piece is awry. Being able to write programs that do not trip on bugs but that can gloss over errors in the code as easily as a person can make sense of a mis-spelt word or a distorted voice would be a triumph. Such "soft" computing is the target of several projects around the world, including the Berkeley Initiative in Soft Computing and Japan's Real World Computing project. Not all use fuzzy logic. Another iconoclastic system of logic known as Bayesian reasoning, which deals in estimates of the probability of events about which there is little or no information, is also popular. But, on the face of it, using fuzzy logic at the points in the program where decisions are made (and bugs are often found) looks sensible.

Some people, however, doubt that fuzzy logic will work better than other methods. Dr Elkan, for instance, thinks that the success of fuzzy controllers will be hard to repeat in more sophisticated applications. Fuzzy control is "structurally shallow": it does not involve long chains of reasoning. The rules that combine the inputs control the output more or less directly. Computer programs are not like this. They usually involve many layers of processing, with the results of one layer being fed into the next.

Shallow thinking has two big advantages. It makes it easy to tune the system, because the effects of different changes in input do not interfere with each other. And it reduces the likelihood of getting mutually contradictory results from parallel chains of reasoning. If it involved long chains of reasoning, fuzzy AI would probably still suffer from the problems that have scuppered other attempts at AI, doubly so if Dr Elkan is right that fuzzy logic collapses into a more traditional two-valued logic.

Dr Kosko agrees that increasing complexity causes problems, but argues that fuzzy logic can be kept simple, even for complex jobs. This means limiting the number of rules. That, in turn, means concentrating on what is critical--lots of rules for areas where small changes in input can have big effects on output; broader strokes for the stable regions in between. If this approach works, technology will be fuzzy in future--but only fairly fuzzy. Such is fuzziness.

Source Citation (MLA 7th Edition)
"The logic that dares not speak its name." *The Economist* [US] 16 Apr. 1994: 89+. *Infotrac Newsstand*. Web. 7 Mar. 2013.
Document URL
http://go.galegroup.com/ps/i.do?id=GALE%7CA15301149&v=2.1&u=bingul&it=r&p=STND&sw=w

Gale Document Number: GALE|A15301149

The Paradoxical Success of Fuzzy Logic

Charles Elkan, University of California, San Diego

Fuzzy logic methods have been used successfully in many real-world applications, but the foundations of fuzzy logic remain under attack. Taken together, these two facts constitute a paradox. A second paradox is that almost all of the successful fuzzy logic applications are embedded controllers, while most of the theoretical papers on fuzzy methods deal with knowledge representation and reasoning. I hope here to resolve these paradoxes by identifying which aspects of fuzzy logic render it useful in practice, and which aspects are inessential. My conclusions are based on a mathematical result, on a survey of literature on the use of fuzzy logic in heuristic control and in expert systems, and on practical experience developing expert systems.

An apparent paradox

As is natural in a research area as active as fuzzy logic, theoreticians have investigated many formal systems, and a variety of systems have been used in applications. Nevertheless, the basic intuitions have remained relatively constant. At its simplest, fuzzy logic is a generalization of standard propositional logic from two truth values, *false* and *true*, to degrees of truth between 0 and 1.

Formally, let A denote an assertion. In fuzzy logic, A is assigned a numerical value $t(A)$, called the *degree of truth* of A, such that $0 \leq t(A) \leq 1$. For a sentence composed from simple assertions and the logical connectives "and" ($\wedge$), "or" ($\vee$), and "not" ($\neg$) degree of truth is defined as follows:

Definition 1: Let A and B be arbitrary assertions. Then

$t(A \wedge B) = \min\{t(A), t(B)\}$
$t(A \vee B) = \max\{t(A), t(B)\}$
$t(\neg A) = 1 - t(A)$
$t(A) = t(B)$ if A and B are logically equivalent. ■

Depending how the phrase "logically equivalent" is understood, Definition 1 yields different formal systems. A fuzzy logic system is intended to allow an indefinite variety of numerical truth values. However, for many notions of logical equivalence, only two different truth values are possible given the postulates of Definition 1.

Theorem 1: Given the formal system of Definition 1, if $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are logically equivalent, then for any two assertions A and B, either $t(B) = t(A)$ or $t(B) = 1 - t(A)$. ■

A direct proof of Theorem 1 appears in the sidebar, but it can also be proved using similar results couched in more abstract form:[1,2]

Proposition: Let P be a finite Boolean algebra of propositions and let τ be a truth-assignment function $P \rightarrow [0,1]$, supposedly truth-functional via continuous connectives. Then for all $p \in P$, $\tau(p) \in \{0, 1\}$ ■

The link between Theorem 1 and this proposition is that $\neg(A \wedge \neg B) \equiv B \vee (\neg A \wedge \neg B)$ is a valid equivalence of Boolean algebra. Theorem 1 is stronger in that it relies on only one particular equivalence, while the proposition is stronger because it applies to any connectives that are truth-functional and continuous (as defined in its authors' paper).

The equivalence used in Theorem 1 is rather complicated, but it is plausible intuitively, and it is natural to apply it in reasoning about a set of fuzzy rules, since $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are both reexpressions of the classical implication $A \rightarrow B$. It was chosen for this reason, but the same result can also be proved using many other ostensibly reasonable logical equivalences.

It is important to be clear on what exactly Theorem 1 says, and what it does not say. On the one hand, the theorem applies to any more general formal system that includes the four postulates listed in Definition 1. Any extension of fuzzy logic to accommodate first-order sentences, for example, collapses to two truth values if it admits the propositional fuzzy logic of Definition 1 and the equivalence used in the statement of Theorem 1 as a special case. The theorem also applies to fuzzy set theory given the equation $(A \cap B^c)^c = B \cup (A^c \cap B^c)$, because Definition 1 can be understood as axiomatizing degrees of membership for fuzzy set intersections, unions, and complements.

On the other hand, the theorem does not necessarily apply to versions of fuzzy logic that modify or reject any of the postulates of Definition 1 or the equivalence used in Theorem 1. However, it is possible to carry through the proof of the theorem in many variant fuzzy logic systems. In particular, the theorem remains true when negation is modeled by any operator in the Sugeno class,[3] and when disjunction or conjunction are modeled by operators in the Yager classes.[4] The theorem also does not depend on any particular definition of implication in fuzzy logic. New definitions of fuzzy implication are still being proposed as new applications of fuzzy logic are investigated.[5]

Of course, the last postulate of Definition 1 is the most controversial one. To preserve

An earlier version with the same title appeared in *Proceedings of the Eleventh National Conference on Artificial Intelligence* (AAAI '93), MIT Press, 1993, pp. 698–703.

Proof of Theorem 1

Theorem 1: Given the formal system of Definition 1, if $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are logically equivalent, then for any two assertions A and B, either $t(B) = t(A)$ or $t(B) = 1 - t(A)$.

Proof: Given the assumed equivalence, $t(\neg(A \wedge \neg B)) = t(B \vee (\neg A \wedge \neg B))$. Now

$$\begin{aligned} t(\neg(A \wedge \neg B)) &= 1 - \min\{t(A), 1 - t(B)\} \\ &= 1 + \max\{-t(A), -1 + t(B)\} \\ &= \max\{1 - t(A), t(B)\} \end{aligned}$$

and

$$t(B \vee (\neg A \wedge \neg B)) = \max\{t(B), \min\{1 - t(A), 1 - t(B)\}\}.$$

The numerical expressions above are different if

$$t(B) < 1 - t(B) < 1 - t(A),$$

that is if $t(B) < 1 - t(B)$ and $t(A) < t(B)$, which happens if $t(A) < t(B) < 0.5$. So it cannot be true that $t(A) < t(B) < 0.5$.

Now note that the sentences $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are both reexpressions of the material implication $A \to B$. One by one, consider the seven other material implication sentences involving A and B, which are

$$\begin{gathered} \neg A \to B \\ A \to \neg B \\ \neg A \to \neg B \\ B \to A \\ \neg B \to A \\ B \to \neg A \\ \neg B \to \neg A \end{gathered}$$

By the same reasoning as before, none of the following can be true:

$$\begin{gathered} 1 - t(A) < t(B) < 0.5 \\ t(A) < 1 - t(B) < 0.5 \\ 1 - t(A) < 1 - t(B) < 0.5 \\ t(B) < t(A) < 0.5 \\ 1 - t(B) < t(A) < 0.5 \\ t(B) < 1 - t(A) < 0.5 \\ 1 - t(B) < 1 - t(A) < 0.5 \end{gathered}$$

Now let $x = \min\{t(A), 1 - t(A)\}$ and let $y = \min\{t(B), 1 - t(B)\}$. Clearly $x \le 0.5$ and $y \le 0.5$ so if $x \ne y$, then one of the eight inequalities derived must be satisfied. Thus $t(B) = t(A)$ or $t(B) = 1 - t(A)$. ■

a continuum of degrees of truth, one naturally wants to restrict the notion of logical equivalence. In intuitive descriptions, fuzzy logic is often characterized as arising from the rejection of the law of excluded middle: the assertion $A \vee \neg A$. Unfortunately, rejecting this law is not sufficient to avoid collapse to just two truth values. Intuitionistic logic rejects the law of excluded middle, but the formal system of Definition 1 still collapses when logical equivalence means intuitionistic equivalence.[6] (The Gödel translations of classically equivalent sentences are intuitionistically equivalent.[6] For any sentence, the first three postulates of Definition 1 make its degree of truth and the degree of truth of its Gödel translation equal. Thus the proof in the sidebar can be carried over directly.) Dubois and Prade note that if all the properties of a Boolean algebra are preserved except for the law of excluded middle, their proposition no longer holds.[2] This observation is compatible with a collapse assuming only the equivalences of intuitionistic logic, because although intuitionistic logic rejects the law of excluded middle, it admits a doubly negated version of the law, namely $\neg\neg(\neg\neg A \vee \neg A)$. Of course, collapse to two truth values is avoided if we admit only the equivalences generated by the operators minimum, maximum, and complement to one. However, these equivalences are essentially the axioms of de Morgan, which allow only restricted reasoning about collections of fuzzy assertions.

Fuzzy logic in expert systems

The basic motivation for fuzzy logic is clear: While many ideas resemble traditional assertions, they are not naturally either true or false; uncertainty of some sort is attached to them. Fuzzy logic is an attempt to capture valid reasoning patterns about uncertainty. The notion is now well accepted that there are many different types of uncertainty, vagueness, and ignorance.[7] However, there is still debate as to what types of uncertainty are captured by fuzzy logic. Many papers have discussed (at a high level of mathematical abstraction) the question of whether fuzzy logic provides suitable laws of thought for reasoning about uncertainty — and if so, which varieties of uncertainty. The question of interest here is more empirical: whether or not fuzzy logic is in practice an adequate formalism for uncertain reasoning in knowledge-based systems.

I conducted a thorough search of the technical literature using the Inspec and Computer Articles databases of more than 1.3 million papers published since 1988. Using the abstracts as a guide, I found no published report of a deployed expert system that uses fuzzy logic as its primary formalism for reasoning under uncertainty. While many theoretical papers on fuzzy logic in expert systems have been published, and several prototype systems have been described, it is hard to find reports of fielded systems doing knowledge-intensive tasks such as diagnosis, scheduling, or design.

Recent conferences give a representative view of the extent of fuzzy logic application in current commercial and industrial knowledge-based systems. All the systems in actual use described at the 1992 IEEE International Conference on Fuzzy Systems are controllers, as opposed to reasoning systems. At the 1993 IEEE Conference on AI for Applications, no applications of fuzzy logic in knowledge-based systems were reported. Of the 16 deployed systems described at the 1993 AAAI Conference on Innovative Applications of AI, three — the CAPE,[8] Dodger,[9] and DYCE[10] systems — used fuzzy logic in some way. However, none of these systems uses fuzzy logic operators for reasoning about uncertainty. Input observations are assigned degrees of membership in fuzzy sets, but inference with these degrees of membership uses other formalisms.

In addition to DYCE, a team at IBM has developed and fielded several knowledge-based systems over the past five years. Some of these systems are used for software and hardware diagnosis, for data analysis, and for operator training.[11,12] The systems have varying architectures and cope with different varieties of uncertainty. Experience with them suggests that fuzzy logic is rarely suitable in practice for reasoning about uncertainty. The basic problem is that items of uncertain knowledge must be combined carefully to avoid incorrect inferences. Fixed domain-independent operators like those of fuzzy logic do not work.

The correct propagation of certainty degrees must account for the content of the uncertain propositions being combined.

This is necessary whether the uncertain propositions constitute deep or shallow knowledge. In the case of shallow knowledge, which may be defined as knowledge that is valid only in a limited context (for example, a correlation between a symptom and a fault), how degrees of uncertainty are combined must be adjusted to account for unstated background knowledge.

A simple example illustrates the difficulty. Consider a system that reasons in a shallow way using a notion of "strength of evidence," and assume that, as in many expert systems, this notion is left primitive and not analyzed more deeply. (Certainly "strength of evidence" is an intuitively meaningful concept that may or may not be probabilistic, but it is definitely different from "degree of truth.") For concreteness, suppose the context of discourse is a collection of melons, and in this context by definition $watermelon(x) \leftrightarrow redinside(x) \wedge greenoutside(x)$. For some melon m, suppose that $t(redinside(m)) = 0.5$ and $t(greenoutside(m)) = 0.8$, meaning that the evidence that m is red internally has strength 0.5, and that m is green externally with strength of evidence 0.8. Are the rules of fuzzy logic adequate for reasoning about this particular type of uncertainty? They say that the strength of evidence that m is a watermelon is $t(watermelon(m)) = \min\{0.5, 0.8\} = 0.5$. However, implicit background knowledge in this context says that being red inside and green outside are mutually reinforcing pieces of evidence toward being a watermelon, so m is a watermelon with strength of evidence over 0.5.

Deep knowledge can be defined as knowledge that is detailed and explicit enough to be valid in multiple contexts. Deep knowledge is general purpose and usable in complex chains of reasoning. However, Theorem 1 says that if more than two different truth values are assigned to the input propositions of long inference chains using fuzzy logic rules and one plausible equivalence, then it is possible to arrive at inconsistent conclusions. Fuzzy logic cannot be used for general reasoning under uncertainty with deep knowledge.

The fundamental issue here is that a conjunction's degree of uncertainty is not in general determined uniquely by the degree of uncertainty of the assertions entering into the conjunction. There does not exist a function f such that the rule $t(A \wedge B) = f(t(A), t(B))$ is always valid, whatever the type of uncertainty represented by $t(\cdot)$. For example, in the case of probabilistic uncertainty, the rule $t(A \wedge B) = t(A) \cdot t(B)$ is valid if and only if A and B represent independent events. In general, for probabilistic uncertainty all one knows is that $\max\{0, t(A) + t(B) - 1\} \leq t(A \wedge B) \leq \min\{t(A), t(B)\}$.

Methods for reasoning about uncertain evidence are an active research area in AI, and the conclusions here are not new. However, our practical experience independently confirms previous arguments about the inadequacy of systems for reasoning about uncertainty that propagate numerical factors according only to which connectives appear in assertions.[13]

Fuzzy logic in heuristic control

The application of fuzzy logic has been most successful in heuristic control, where there is wide consensus that traditional techniques of mathematical control theory are often inadequate. The reasons for this include the reliance of traditional methods on linear models of systems to be controlled, their propensity to produce "bang-bang" control regimes, and their focus on worst-case convergence and stability rather than typical-case efficiency. Heuristic control techniques give up mathematical simplicity and performance guarantees in exchange for increased realism and better performance in practice. For example, a heuristic controller using fuzzy logic has been shown to have less overshoot and quicker settling.[14]

The first demonstrations that fuzzy logic could be used in heuristic controllers were published in the 1970s.[15,16] Work continued through the 1980s, and recently there has been an explosion of industrial interest in the area.[17,18] One reason for this recent interest in fuzzy controllers is that they can be implemented by embedded specialized microprocessors.[19]

Despite industry interest, and consumer interest in Japan, fuzzy logic technology continues to meet resistance. For example, at IJCAI '91, Takeo Kanade gave a talk on computer vision, describing at length Matsushita's camcorder image stabilizing system without mentioning its use of fuzzy logic. Also, while a fuzzy logic controller is embedded in the 1994 Honda Accord's automatic transmission, the advertising brochures describe it as "grade logic."

Almost all currently deployed heuristic controllers using fuzzy logic are similar in five important aspects (a good example of this standard architecture appears in a paper by Sugeno and his colleagues[21]):

(1) The typical fuzzy controller knowledge base consists of fewer than 100 rules; often fewer than 20 rules are used. Fuzzy controllers are orders of magnitude smaller than systems built using traditional AI formalisms.

(2) The knowledge entering into fuzzy controllers is structurally shallow, both statically and dynamically. Conclusions produced by rules are not used as premises in other rules; statically rules are organized in a flat list, and dynamically there is no runtime chaining of inferences.

(3) The knowledge recorded in a fuzzy controller typically reflects immediate correlations between the inputs and outputs to be controlled, as opposed to a deep, causal model of the system. The premises of rules refer to sensor observations, and rule conclusions refer to actuator settings. (Rule premises refer to qualitative or "linguistic" sensor observations, and rule conclusions refer to qualitative actuator settings, whereas outputs and inputs of sensors and actuators are typically real-valued. This means that normally two controller components map between numerical values and qualitative values. In fuzzy logic terminology, these components are said to defuzzify outputs and implement membership functions.)

(4) In deployed fuzzy controllers, the numerical parameters of their rules and of their qualitative input and output modules are tuned in a learning process. The tuning can be done by human engineers or by learning algorithms; neural network methods have been especially successful.[22] What the tuning algorithms themselves have in common is that they are gradient-descent "hill-climbing" algorithms that learn by local optimization.[14]
(5) By definition, fuzzy controllers use fuzzy logic operators. Typically, minimum and maximum are used, as are explicit possibility distributions (usually trapezoidal) and some fuzzy implication operator.

The question that naturally arises is, Which of these five features are essential to the success of fuzzy controllers? It appears that the first four are vital to practical success, because they make the celebrated credit assignment problem solvable, while the use of fuzzy logic is not essential.

In a nutshell, the credit assignment problem is to improve a complex system by modifying a part of it, given only an evaluation of its overall performance. In general, solving the credit assignment problem is impossible: the task is tantamount to generating many bits of information (a change to the internals of the system) from just a few bits of information (the system's input/output performance). However, the first four shared features of fuzzy controllers can solve this problem for the following reasons.

First, since it consists of only a few rules, the knowledge base of a fuzzy controller is a small system to modify. Second, the short paths between the fuzzy controller's inputs and outputs localize the effect of a change, making it easier to discover a change with a desired effect without producing undesired consequences. Third, because of the iterative way in which fuzzy controllers are refined, many observations of input/output performance are available for system improvement. Fourth, the continuous nature of the controller's parameters allows small quantities of performance information to be used to make small system changes.

Thus, what makes fuzzy controllers useful in practice is the combination of a rule-based formalism with numerical factors qualifying rules and the premises entering into rules. The principal advantage of rule-based formalisms is that knowledge can be acquired from experts or from experience incrementally. Individual rules and premises can be refined independently, or at least more independently than items of knowledge in other formalisms. Numerical factors have two main advantages. They allow a heuristic control system to interface smoothly with the continuous outside world, and they allow it to be tuned gradually — small changes in numerical factor values cause small changes in behavior.

None of the features contributing to the success of systems based on fuzzy logic is unique to fuzzy logic. It seems that most current fuzzy logic applications could use other numerical rule-based formalisms instead — if a human or a learning algorithm tuned numerical values for those formalisms, as is customary when using fuzzy logic. A quote from the originator of fuzzy heuristic control is relevant here:

> ... it should be remarked that the work on process control using fuzzy logic was inspired as much by Waterman and his approach to rule-based decision making as by Zadeh ... and his novel theory of fuzzy subsets.[23]

Several knowledge representation formalisms that are rule-based and numerical have been proposed besides fuzzy logic.[24,25] To the extent that numerical factors can be tuned in these formalisms, they should be equally useful for constructing heuristic controllers. Indeed, at least one has already been so used.[26]

Recapitulating mainstream AI

Several research groups are attempting to scale up systems based on fuzzy logic and lift the architectural limitations of current fuzzy controllers. For example, a methodology for designing block-structured controllers with guaranteed stability properties has been studied,[27] as have methodological problems in constructing models of complex systems based on deep knowledge.[28] Controllers with intermediate variables, thus with chaining of inferences, have also been investigated.[29]

However, the designers of larger systems based on fuzzy logic are encountering all the problems of scale already identified in traditional knowledge-based systems. It appears that the research history of fuzzy logic is recapitulating that of other areas in AI as well, particularly those dealing with knowledge engineering and state information.

The rules in the knowledge bases of current fuzzy controllers are obtained directly by interviewing experts. Indeed, the original motivation for using fuzzy logic in building heuristic controllers was that fuzzy logic is designed to capture human statements involving vague quantifiers such as "considerable." More recently, consensus has developed around the idea that research must focus on obtaining "procedures for fuzzy controller design based on fuzzy models of the process."[30] Mainstream work on knowledge engineering, however, has already transcended the dichotomy between rule-based and model-based reasoning.

Expert systems with knowledge consisting of *if-then* rules have at least two disadvantages. First, maintenance of a rule base becomes complex and time-consuming as the system size increases. Second, rule-based systems tend to be brittle: If an item of knowledge is missing from a rule, the system may fail to find a solution, or worse, may draw an incorrect conclusion. The main disadvantage of model-based approaches, on the other hand, is that it is very difficult to construct sufficiently detailed and accurate models of complex systems. Moreover, the models constructed tend to be highly application-specific and not generalizable.[31]

Many recent expert systems, therefore, are neither rule-based nor model-based in the standard way.[12] For these systems, the aim of the knowledge engineering process is not simply to acquire knowledge from human experts, but rather to develop a theory of the experts' situated performance (this is true regardless of whether the desired knowledge is correlational, as in present fuzzy controllers, or deep, as in model-based expert systems). Concretely, under this view of knowledge engineering. knowledge bases are constructed to model the beliefs and practices of experts and not "objective" truths about underlying physical processes. An important benefit of this approach is that the organization of an expert's beliefs provides an implicit organization of knowledge about the external process with which the knowledge-based system is intended to interact.

The more sophisticated view of knowledge engineering just outlined is clearly relevant to research on constructing more intricate fuzzy controllers. For a second example of relevant AI work, consider controllers that can carry state information from one moment to the next (mentioned as a topic for future research by von Altrock and colleagues[29]). Symbolic AI formalisms for representing systems whose behavior depends on their history have been available since the 1960s. Neural networks with similar properties (called recurrent networks) have been available for several years, and have already been used in control applications.[35] It remains to be seen whether research from a fuzzy logic perspective will provide new solutions to the fundamental issues of AI.

Applications of fuzzy logic in heuristic control have been highly successful, despite the collapse of fuzzy logic to two-valued logic under an apparently reasonable condition, and despite the inadequacy of fuzzy logic for general inference with uncertain knowledge. These difficulties have not been harmful in practice because current fuzzy controllers are far simpler than other knowledge-based systems. Theorem 1 is not an issue for fuzzy controllers because they do not perform chains of inference, and they are developed informally, with no formal reasoning about their rules that applies equivalences such as the one used in the statement of Theorem 1. Second, the knowledge recorded in a fuzzy controller is not a consistent causal model of the process being controlled, but rather an assemblage of visible correlations between sensor observations and actuator settings. Since this knowledge is not itself general-purpose, the inadequacy of fuzzy logic for general reasoning about uncertainty is not an issue. Moreover, the ability to refine the parameters of a fuzzy controller iteratively can compensate for the arbitrariness of the fuzzy logic operators as applied inside a limited domain.

The common assumption that heuristic controllers based on fuzzy logic are successful because they use fuzzy logic appears to be an instance of the *post hoc, ergo propter hoc* fallacy. The fact that using fuzzy logic is correlated with success does not entail that using fuzzy logic causes success. In the future, as fuzzy controllers are scaled up, the technical difficulties identified in this article can be expected to become important in practice.

Theorem 1 is a crisp demonstration of one of several deep difficulties of scale in AI: the problem of maintaining consistency in long sequences of reasoning. Other difficulties of scale can also be expected to become critical — in particular, the issue of designing learning mechanisms that can solve the credit assignment problem when the simplifying features of present controllers are absent.

Acknowledgments

The author is grateful to many colleagues for useful comments on earlier versions of this article.

References

1. D. Dubois and H. Prade, "New Results about Properties and Semantics of Fuzzy Set-Theoretic Operators," *Fuzzy Sets: Theory and Applications to Policy Analysis and Information Systems*, Plenum Press, New York, 1980, pp. 59–75.
2. D. Dubois and H. Prade, "An Introduction to Possibilistic and Fuzzy Logics," *Non-Standard Logics for Automated Reasoning*, Academic Press, San Diego, 1988.
3. M. Sugeno, "Fuzzy Measures and Fuzzy Integrals — A Survey," *Fuzzy Automata and Decision Processes*, Elsevier/North-Holland, New York, 1977, pp. 89–102.
4. R.R. Yager, "On a General Class of Fuzzy Connectives," *Fuzzy Sets and Systems*, Vol. 4, No. 3, Nov. 1980, pp. 235–242.
5. D. Dubois and H. Prade, "Gradual Inference Rules in Approximate Reasoning," *Information Sciences*, Vol. 61, No. 1–2, 1992, pp. 103–122.
6. D. van Dalen, *Logic and Structure*, second ed., Springer-Verlag, New York, 1983.
7. P. Smets, "Varieties of Ignorance and the Need for Well-Founded Theories," *Information Sciences*, No. 57–58, Sept.-Dec. 1991, pp. 135–144.
8. A. Cunningham and R. Smart, "Computer-Aided Parts Estimation," *Proc. 5th Innovative Applications of AI Conf.*, AAAI Press, Menlo Park, Calif., 1993, pp. 14–25.
9. A.J. Levy et al:, "Dodger: A Diagnostic Expert System for the Evaluation of Nondestructive Test Data," *Proc. 5th Innovative Applications of AI Conf.*, AAAI Press, Menlo Park, Calif., 1993, pp. 107–117.
10. D.D. Pierson and G.J. Gallant, "Diagnostic Yield Characterization Expert (DYCE): A Diagnostic Knowledge-Based System Shell for Automated Data Analysis," *Proc. 5th Innovative Applications of AI Conf.*, AAAI Press, Menlo Park, Calif., 1993, pp. 152–160.
11. G. Gallant and J. Thygesen, "Digitized Expert Pictures (Depict): An Intelligent Information Repository," *Proc. 5th Innovative Applications of AI Conf.*, AAAI Press, Menlo Park, Calif., 1993, pp. 50–60.
12. A. Hekmatpour and C. Elkan, "Categorization-Based Diagnostic Problem Solving in the VLSI Design Domain," *Proc. IEEE Int'l Conf. on AI for Applications*, IEEE Computer Society Press, Los Alamitos, Calif., 1993, pp. 121–127.
13. J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Francisco, Calif., 1988.
14. D.G. Burkhardt and P.P. Bonissone, "Automated Fuzzy Knowledge Base Generation and Tuning," *Proc. IEEE Int'l Conf. on Fuzzy Systems*, IEEE Computer Society Press, Los Alamitos, Calif., 1992, pp. 179–188.
15. L.A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Deci-

sion Processes," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 3, No. 1, Jan. 1973, pp. 28–44.

16. E.H. Mamdani, "Application of Fuzzy Algorithms for Control of Simple Dynamic Plant," *Proc. Institution of Electrical Engineers*, Vol. 121, No. 12, 1974, pp. 1585–8.
17. T. Yamakawa and K. Hirota, eds., special issue on applications of fuzzy logic control to industry, *Fuzzy Sets and Systems*, Vol. 32, No. 2, 1989.
18. C.C. Lee, "Fuzzy Logic in Control Systems — Parts 1 and 2," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 20, No. 2, Mar. 1990, pp. 404–435.
19. T. Yamakawa, "Stabilization of an Inverted Pendulum by a High-Speed Fuzzy Logic Controller Hardware System," *Fuzzy Sets and Systems*, Vol. 32, No. 2, 1989, pp. 161–180.
20. K. Uomori et al., "Automatic Image Stabilizing System by Full-Digital Signal Processing," *IEEE Trans. Consumer Electronics*, Vol. 36, No. 3, Aug. 1990, pp. 510–519.
21. M. Sugeno et al., "Fuzzy Algorithmic Control of a Model Car by Oral Instructions," *Fuzzy Sets and Systems*, Vol. 32, No. 2, 1989, pp. 135–156.
22. J.M. Keller and H. Tahani, "Backpropagation Neural Networks for Fuzzy Logic," *Information Sciences*, Vol. 62, No. 3, 1992, pp. 205–221.
23. E.H. Mamdani and B.S. Sembi, "Process Control Using Fuzzy Logic," *Fuzzy Sets: Theory and Applications to Policy Analysis and Information Systems*, Plenum Press, New York, 1980, pp. 249–265.
24. E. Sandewall, "Combining Logic and Differential Equations for Describing Real-World Systems," *Proc. First Int'l Conf. on Principles of Knowledge Representation and Reasoning*, R.J. Brachman, H.J. Levesque, and R. Reiter, eds., Morgan Kaufmann, San Francisco, Calif., 1989, pp. 412–420.
25. A. Collins and R. Michalski, "The Logic of Plausible Reasoning: A Core Theory," *Cognitive Science*, Vol. 13, No. 1, 1989, pp. 1–49.
26. C. Sammut and D. Michie, "Controlling a 'Black Box' Simulation of a Space Craft," *AI Magazine*, Vol. 12, No. 1, 1991, pp. 56–63.
27. K. Tanaka and M. Sugeno, "Stability Analysis and Design of Fuzzy Control Systems," *Fuzzy Sets and Systems*, Vol. 45, No. 2, 1992, pp. 135–156.
28. W. Pedrycz, "Fuzzy Modeling: Fundamentals, Construction, and Evaluation," *Fuzzy Sets and Systems*, Vol. 41, No. 1, 1991, pp. 1–15.
29. C. von Altrock, B. Krause, and H.J. Zimmermann, "Advanced Fuzzy Logic Control of a Model Car in Extreme Situations," *Fuzzy Sets and Systems*, Vol. 48, No. 1, 1992, pp. 41–52.
30. D. Driankov and P. Eklund, "Workshop Goals," distributed by the authors at IJCAI '91 Workshop on Fuzzy Control.
31. J.R. Bourne et al., "Organizing and Understanding Beliefs in Advice-Giving Diagnostic Systems. *IEEE Trans. Knowledge and Data Engineering*, Vol. 3, No. 3, Sept. 1991, pp. 269–280.
32. J. McCarthy and P.J. Hayes, "Some Philosophical Problems from the Standpoint of Artificial Intelligence," *Machine Intelligence*, Vol. 4, Edinburgh Univ. Press, Edinburgh, Scotland, 1969, pp. 463–502.
33. R.L. Watrous and L. Shastri, "Learning Phonetic Features Using Connectionist Networks," *Proc. 10th Int'l Joint Conf. on AI*, Morgan Kaufmann, San Francisco, Calif., 1987, pp. 851–854.
34. J.L. Elman, "Finding Structure in Time," *Cognitive Science*, Vol. 14, No. 2, 1990, pp. 179–211.
35. C.C. Ku, K.Y. Lee, and R.M. Edwards, "Improved Nuclear Reactor Temperature Control Using Diagonal Recurrent Neural Networks," *IEEE Trans. Nuclear Science*, Vol. 39, No. 6, 1992, pp. 2298–2308.

Charles Elkan is an assistant professor in the Department of Computer Science and Engineering at the University of California, San Diego. His main research interests are in artificial intelligence. With students and colleagues, he has worked recently on learning algorithms for DNA and protein sequence analysis, algorithms for reasoning about database queries and updates, methods of formalizing commonsense knowledge, and other topics. In the field of knowledge-based systems, his paper with A. Hekmatpour, "Categorization-Based Diagnostic Problem Solving in the VLSI Design Domain," won a best paper award at the 1993 IEEE Conference on Artificial Intelligence for Applications. Before joining UCSD in 1990, Dr. Elkan was a postdoctoral fellow at the University of Toronto. He earned his PhD and MS at Cornell University in computer science, and his BA in mathematics at Cambridge University. Charles Elkan can be contacted at the Dept. of Computer Science and Engineering, UCSD, La Jolla, CA 92093-0114; Internet: elkan@cs.ucsd.edu

RESPONSE

The Unique Strength of Fuzzy Logic Control

Hamid R. Berenji, Intelligent Inference Systems/NASA Ames Research Center

I am pleased to see that Elkan has revised his paper based on comments from fuzzy logic experts. His reference to Dubois and Prade indicates that he has realized, finally, that his alleged "new discovery" has long been known by specialists in fuzzy and multivalued logics.

Unfortunately, the new version still contains many misunderstandings and errors. I will briefly respond to some of them, avoiding a discussion of the supposedly startling proof about the purported inconsistency of fuzzy logic, which is covered in responses by Enrique Ruspini and others. I will confine my comments primarily to a fundamental misunderstanding that is the source of many of Elkan's mistaken assertions about the use of fuzzy logic in heuristic control and expert systems.

Elkan lists a number of powerful features of fuzzy-logic control, but then erroneously concludes that none is unique to fuzzy logic. He fails to realize that the unique strength of fuzzy-logic control is its dependence on fuzzy-set theory and its representational capabilities. The small number of rules typical in these systems is not the result of mere luck, but the direct consequence of the fuzzy predicates that appear in the rules. Each of these predicates covers a wide range of state variable values while facilitating interpolation of rule consequents. Fuzzy sets provide for a general, yet compact characterization of system state that requires fewer rules.

Elkan's assertion about the shallowness of fuzzy controller knowledge is simply wrong. Recent fuzzy-logic controllers, developed for more challenging tasks, use hierarchical fuzzy control methods.[1] Examples include the helicopter control developed by Sugeno and his collaborators at the Tokyo Institute of Technology (a system that can appear trivial only to those unfamiliar with control theory), and the controller for a three-linked inverted pendulum developed at Aptronix. In applications such as these, the result of the first level of control is used in deriving control rules for the second set, and so on. These examples prove that fuzzy-logic control systems can be developed to reason with considerable depth of complexity. Similarly, the control mechanisms for the local-motion control of SRI's autonomous robot[2] rely on several deliberation levels to determine the relevance level of each control rule (by evaluating the operational environment characteristics); to identify current goals and their state of achievement; to activate control rules according to the current context; and to blend their control recommendations.

At any rate, the "depth" of a reasoning process as Elkan seems to understand it is not even a well-defined measure of inferential system complexity. This is seen in the fact that the two-level forward chain $A \rightarrow (B \rightarrow C)$ is often "compiled" in real-time applications (such as control systems) into the single-level rule $A \wedge B \rightarrow C$ to simplify and speed computation. This simplification mechanism, which turns what Elkan would consider "complex" into an equivalent "simple" version, is used to introduce contextual and goal-dependence considerations into the reasoning chain both in the SRI's mobile robot controller and in our own two-goal inverted pendulum.

Using fuzzy sets to describe a general linguistic variable also significantly reduces the complexity of the search process in fuzzy systems that learn from experience. Elkan correctly points out that using fewer rules simplifies the credit assignment problem, but he fails to realize that this is a consequence of using fuzzy logic rather than an indicator of its current or future applicability. This feature is desirable in any control system, as is seen in the fuzzy-logic controller developed at NASA Ames for the Space Shuttle's rendezvous and docking operations.[3] This controller learns to improve itself from experience using reinforcement learning techniques,[4,5] a complex task that would have been very difficult, if not impossible, if other symbolic control techniques had been used.

In summary, I see two major misunderstandings in Elkan's paper. First, it relies on a theorem that is irrelevant to fuzzy logic to argue that the methodology is paradoxical. Second, it fails to note that the advantages provided by fuzzy-set constructs give fuzzy control a unique methodological strength — a fact Elkan mistakenly interprets as technological immaturity.

References

1. H.R. Berenji et al., "A Hierarchical Approach to Designing Approximate Reasoning-Based Controllers for Dynamic Physical Systems," in *Uncertainty in Artificial Intelligence*, P.P. Bonissone et al, eds. North-Holland, Amsterdam, 1991, pp. 331-343.
2. A. Saffiotti, E. Ruspini, and K. Konolige, "Blending Reactivity and Goal-Directedness in a Fuzzy Controller, Proc. Fuzzy Logic, in *Proc. Second IEEE Int'l Conf. Fuzzy Systems*, IEEE Computer Society Press, Los Alamitos, Calif., 1993, pp. 134-139.
3. H.R. Berenji et al., "Space Shuttle Attitude Control by Fuzzy Logic and Reinforcement Learning," in *Proc. Second IEEE Int'l Conf. Fuzzy Systems*, IEEE Press, Pistcataway, N.J., 1993, pp. 1396-1401.
4. H.R. Berenji and P. Khedkar, "Learning and Tuning Fuzzy Logic Controllers through Reinforcements," *IEEE Trans. on Neural Networks*, Vol. 3, No. 5, 1992, pp. 724-740.
5. H.R. Berenji, "An Architecture for Designing Fuzzy Controllers using Neural Networks," in *Int'l J. Approximate Reasoning*, Vol. 6., No. 2, Feb. 1992, pp. 267-292.

Hamid R. Berenji is a senior research scientist and principal investigator on intelligent control in the AI branch of the NASA Ames Research Center. He was a program chair for the IEEE International Conference on Neural Networks, and was a program cochair of the 1994 IEEE Conference on Fuzzy Systems. He serves on the editorial board of several technical publications, and is an associate editor of *IEEE Transactions on Fuzzy Systems* and *IEEE Transactions on Neural Networks*. He is a member of IEEE, and chairs the Neural Networks Council's Technical Committee on Fuzzy Systems. Hamid Berenji can be reached at berenji@ptolemy.arc.nasa.gov

RESPONSE

Broader Issues At Stake

A Response to Elkan

B. Chandrasekaran, Ohio State University

The fuzzy set approach has clearly captured the interest of many researchers around the world and has been used to build applications of various sorts, of which fuzzy control applications are currently the most prominent. The approach, however, remains controversial. While this controversy has many sources, there are relatively few places where the arguments are set out in a fashion that allows debate. It is thus useful to have both Charles Elkan's analysis of the fuzzy set approach to representing uncertainty, and his examination of which features of fuzzy set theory are responsible for the success of fuzzy control systems. In particular, I commend Elkan for making his arguments about these techniques in a nonpolemical way, letting technical arguments and results do most of the talking.

In Elkan's first argument, he claims that the axioms of fuzzy set theory, in conjunction with what appear to be a number of reasonable versions of logical equivalence between sentences, lead to a collapse of truth functions into just two values — a fate that fuzzy set theory was expressly meant to avoid.

As Elkan points out, a result similar to his collapse theorem was already known to researchers within the fuzzy set community (Dubois and Prade). My understanding is that they weren't too worried by this result, since they think that the traditional notion of logical equivalence or any of its variants should be abandoned for fuzzy sets. This response seems to me to be formally reasonable, but I think in practice it would be hard to work with a system in which logical equivalence itself is a fuzzy relation. Ultimately, we will have to see how much really interesting work is possible with this notion of fuzzy equivalence.

In the second argument, Elkan asserts that when fuzzy control systems that work well are analyzed, the real source of their success seems to be not the inferential capabilities of fuzzy set theory (derived from the theory's composition axioms) but rather a combination of things exclusive of fuzzy set axioms. Among these are the ability to represent certain things as continuous quantities rather than all-or-nothing quantities; certain heuristic techniques — that are themselves outside fuzzy set theory — to get the right parameters for the problems; and the fact that there is little complex rule-chaining going on. A number of alternatives and rivals to fuzzy set theory would work as well in those applications.

Part of Elkan's point — that the success of fuzzy control systems thus far is not really a full test or proof of the axioms and claims of fuzzy set theory — is actually an instance of a larger phenomenon in AI. I think that Elkan's point can be made against the claims of not only fuzzy control proposals, but also against a number of other proposals in AI, including the rivals of fuzzy sets, such as belief nets.

The general problem is a kind of credit allocation problem and can be stated as follows. Given some mechanism *M*, and some specific task *T*, suppose I write a program *P*, using *M* as the basis for the program. And, let us say that *P* does well in the task *T*. What conclusions can we draw about mechanism *M* from the success of *P* in tackling *T*? How much credit should *M* get for the success of *P*?

A historical perspective. In the late 1970's, rule-based expert systems were capturing the imagination of many people. Mycin and R1 were great successes. In the above terminology, rule-based languages would be *M*, Mycin and R1 would be the *P*'s, and simple diagnosis and configuration would be the corresponding tasks, *T*. The success of the two programs led to claims about the power of the rule-based mechanism. Similar examples involving other mechanisms, such as belief nets and truth maintenance systems, can be constructed.

In a series of articles (such as one from 1986,[1] for example), I made the following points regarding rule-based systems as a mechanism. The specifics of the mechanism were incidental in accounting for many aspects of why the programs worked. The mechanism was computation-universal, and of course could be used to implement any other mechanism or strategy. A higher order strategy — classification in the case of Mycin, or linear sequencing of subtasks in the case of R1 — was the problem-level strategy that was responsible for the programs' performance. Not only was the rule-based mechanism not the direct cause of the good performance, but they actually hid the reasons for success: The higher level strategies were programmed in the language of the lower level mechanism. The strategies had to be brought out by analysis, rather than seen by a direct inspection of the mechanism. The limitations and success of Mycin and R1 could be more insightfully analyzed by examining the adequacy of classification for diagnosis and linear subtasking for configuration design. Clancey also analyzed Mycin as a heuristic classifier[2] and pointed out the power such high-level analysis brought to building diagnostic systems. In the last decade or so, there has been a decisive shift in emphasis in the field of knowledge-based systems from mechanisms at the rule level to phenomena at the task level.

Thus, given an M-T-P triad, it is not always easy to decide exactly what the role of M was in the success of P in achieving that version of T. This is not to say that M's properties are irrelevant. There are several ways a given mechanism might play less of a role than is readily apparent, among them:

- M might simply be one among many perfectly reasonable lower level mechanisms to implement the causally more relevant higher level mechanism.
- M might have features which actually impede good performance for the class of problems in T. This might not be evident from the specific instance of T for which P was written. In this instance, the troublesome features of M might not have been used or their effect might be minimal. Fuzzy set theory has been successfully applied to simple versions of the control problem. As Elkan argues, however, the problematic features of the theory might start showing up as more complex versions of the control problem are encountered.
- In some cases, M has many more features than needed for capturing the essence of T. Hence, using M to build P for solving T calls for making commitments to details that are either irrelevant or that detract from building good P's. However, when such a program is built, it takes quite a bit of analysis to tell which features of M are necessary. There is often a tendency, especially among those who are enthusiasts of M for other reasons, to ascribe the success of P to those features of M that were actually incidental to P's success. Even more seriously, success with M might lead to its use for more complex versions of T, where these additional features actually make building successful P's more difficult. Elkan makes a good case for this possibility as fuzzy control approaches are applied to more complex control problems.

The history of Mycin is another source of wisdom about the role of uncertainty-handling mechanisms. When Mycin came out, much was made of the uncertainty-factor formalism. Debates raged about this formalism versus Bayesian formalism versus fuzzy set formalisms as an appropriate calculus. Cooper and Clancey got the idea of doing an experiment in which they coarsened the uncertainty factors in Mycin's knowledge base rules and examined how well the modified Mycin did in the same cases.[3] The modified Mycin solved the problems as well as the original Mycin.

How could this be? Clearly the calculus as such didn't play as fundamental a role in the ability of Mycin to solve the problems. The fine structure of uncertainty didn't really matter. The knowledge base had enough knowledge to establish or reject the conclusions in a near-definitive way. None of the conclusions were based on even moderate distinctions in uncertainty between the candidates. There were multiple ways to get to or reject conclusions, and even moderate changes in the uncertainties didn't matter. The correct conclusions were very strongly established, and the incorrect conclusions were very strongly rejected. Mycin did well, not because of the fine points of its uncertainty calculus — it would have done just as well with any of a number of alternative calculi — but because of the robustness of its knowledge base. This is another instance of the allocation of credit problem.

The nature of fuzzy theory

I have followed fuzzy set theory almost from its inception. The theory's claim that all senses of uncertainty in human knowledge cannot be reduced to some version of probability has always struck me as right. One of the most useful consequences of the fuzzy set movement has been the identification of different types of uncertainty. In particular, the theory suggests that many predicates such as "bald," "most," and "large" are neither binary predicates, nor are they simply probabilistic. This also seems to me to be true. However, the specific solutions offered and claims made by fuzzy set theory, and the way they have often been applied to problems like control, are problematic for me.

A psychological theory? At the heart of fuzzy set theory is an ambiguity about the nature of the theory, and how one goes about validating it. If it is a psychological theory — that is, a theory of how humans deal with certain types of uncertainty — we would need certain kinds of evidence about human behavior in uncertainty handling. I am unconvinced that fuzzy set theory is a psychological theory. I have not done an extensive literature survey, but the work of Kempton[4,5] raises doubts that human behavior in uncertainty handling follows the axioms of fuzzy set theory.

Even if it turns out that the theory does correspond to human behavior in this area, we must then decide what kinds of scaling and rationality properties the relevant human behavior has before it is used to make machines that make decisions.

Two relevant analogies are found in commonsense physical reasoning and reasoning about probabilistic uncertainty. We all have approximate rules about how the physical world behaves: "If we push this a little, this will move a moderate distance, while the other object would hardly move." We use such rules when we have to predict behavior in the physical world, but these rules are typically chained over a few steps. When a problem calls for many steps, these rules start accumulating large errors (to be expected), but curiously, they also start accumulating ambiguities of another sort. So many alternative possibilities are generated that we adopt all kinds of goal- and context-specific strategies to select a "future history" over other alternatives. Or, if we are physicists, we resort to a pencil and paper for more exact calculations even if what we really want are approximate answers. Clearly such approximate reasoning by humans does not scale up very well.

In the case of probability assessment behavior, human behavior is not always what an outside observer might regard as rational.[6] Thus, in addition to the scalability problem, there is the problem of rationality of human behavior as well.

The point that I want to make with these two examples is that, in many domains, automated decision systems should not be designed to emulate human behavior. Thus, even if fuzzy set theory turns out to be a model of how humans handle a certain type of uncertainty, we need additional arguments to make the theory the basis of automatic control.

A mathematical theory? On the other hand, fuzzy set could be a theory of an abstract mathematical system whose properties model some domain of human interest. Examples of such systems are arithmetic and deductive logic. The formalization of arithmetic starts with our intuitive notions about numbers, but it is not a psychological theory. It posits a world of numbers and operations on them, and the formalization is an attempt to capture the properties of this world. We can in fact construct the abstract world, recognize its objects as the familiar numbers and perform operations on them, and then verify those operations against the predictions of the axiomatization. For example, we can multiply 2 and 3, and check if the axiom system in fact generates the number 6 for the answer.

If fuzzy set theory is a theory of an abstract world whose constituents are uncertainties of certain types, and whose operations are the sort of things we do when we combine uncertainties, then the theory has to give two kinds of evidence. First, there must be evidence that such an abstract world indeed exists. Many abstract worlds that can be postulated fail to exist because their axioms lack a certain internal coherence. Second, it must give evidence that the fuzzy set axioms capture the operations of this world. Establishing that such an abstract world exists is actually quite hard. In fact, I think it is quite possible that there is no abstract world of uncertainty combination of the type that fuzzy set theory attempts to capture. In any case, fuzzy set theory has to worry about validation of its assumptions and about the existence of an abstract calculus for combining this kind of uncertainty.

What do I mean by "such an abstract world may not exist"? Again, the analogy of qualitative physics is relevant. We know there is a real physics, whose laws relate values of some state variables to the values of other state variables. If we have an exact value for the independent variables, we can calculate, using these laws, the exact values of the dependent variables.

The equations of physics are not a psychological theory. However, consider the ordinary commonsense reasoning about the physical world that I discussed earlier. People do make qualitative predictions about the physical world in response to qualitative changes in some state of the world. As I said, the qualitative rules people have cannot be chained into long inferences: The ambiguities multiply, resulting in too many possible future histories. Which one of the histories will be realized often depends on a more exact value for some variables than we can get from qualitative rules alone. I have described elsewhere a number of strategies people use to handle such an explosion of possibilities, but almost all of the strategies depend on the problem-solving goal and context.[7] The conclusion is not the result of applying an abstract, context-independent calculus. In short, there is no qualitative physics that is a homomorphism of the quantitative physics such that the qualitative physics gives answers that are just qualitative versions of the answers given by the quantitative physics.

With respect to uncertainty handling, many researchers seem to be looking for a similar abstract system that may not exist. They are looking for a calculus of uncertainty handling which has the following features:

- The semantics of its uncertainty terms capture the intuitive meaning of uncertainty terms that people use in their commonsense behavior.
- The operations of combination in the calculus capture human behavior when their uncertainties are combined.

This assumes that there is in fact a calculus that underlies the combining of uncertainties through human common sense. What if human behavior, in combining everyday uncertainties, is really governed by a combination of goal- and context-dependent strategies that make use of a rich body of domain-specific knowledge? What if this cannot be captured by a calculus of the type that fuzzy set or other theorists are looking for? If human conclusions are robust with respect to moderate changes in the uncertainty values of the constituents — as in the Mycin experiment by Cooper and Clancey — then the real explanation of human behavior is not given by a calculus, fuzzy or otherwise, but by the complex collection of situation- and goal-specific knowledge that people bring to bear on instances of the problem.

Like the case in qualitative reasoning mentioned earlier, people might in fact avoid anything like a chain of uncertainty combination. If the conclusion seems robust with respect to moderate changes in the uncertainty values of its constituents, people feel comfortable with the conclusion. If not, they might get additional data so that a robust conclusion can be reached, postpone making a decision, or make decisions that may not in general be considered the best, but that are fine for the specific goal at hand. In other words, the same values of uncertainties for two constituent beliefs would lead to a conclusion with an uncertainty value A in one situation, an uncertainty value B in another, additional information gathering in a third, explicit use of probability models in a fourth, and simply a shrugging of shoulders and no decision at all in a fifth. If this is the case, then the search for a calculus of the type fuzzy set theorists (and many others in the research community concerned with modeling uncertainty in reasoning) are looking for is likely to be futile. The issue is illustrated well in Elkan's example of his expert system, for which neither the probability scheme nor the fuzzy set approach was appropriate.

The problem with fuzzy set theory, in

my view, is not in the mathematics of the formal system. It is clearly a mathematical system of some interest. However, a theory of this type has to be judged either as a psychological theory or as a theory that has captured an abstract calculus that underlies some type of human reasoning. As I have just argued, an abstract calculus of this type may not exist.

The problem of context. In the 1980's, my colleagues and I were faced with a similar problem with uncertainty in medical diagnosis. Physicians have to come up with an assessment of the "likelihood" of some disease for which a number of data were potentially relevant. The relation between the data and the strength of belief in the disease was of course a classic example of uncertainty. For various reasons — not the least of which was that we didn't have the data needed to use the frequency version of the probabilities for this relationship — we needed a technique to model human expertise in this area. Bayesian approaches, fuzzy set theory, Dempster-Shafer theory, and uncertainty factor calculus were all available to us. All these calculi shared one important property or assumption about human expertise — that there was a situation- and goal-independent way of combining uncertainties.

For example, if two symptoms, $s1$ and $s2$, were relevant to making a decision about disease d, such calculi would provide ways in which evidence for $s1$ and $s2$ would be combined to give evidence about d, and additionally, that the rule of combination itself is independent of the specific labels for $s1$, $s2$, and d. If the evidence for $s1$ is large, and $s2$ is medium, the rule would specify what the evidence for d would be. But the rule cannot be one thing where $s1$ is "bilirubin," $s2$ is "alkaline phosphatase," and d is "liver disease," while another rule is used where $s1$ is "cholesterol level," $s2$ is "alkaline phosphatase," and d is "heart disease."

We found, however, that expert behavior in uncertainty combination in fact differed from context to context, and problem-solving goal to problem-solving goal. We had to resist the mathematical attractions of an abstract calculus. Instead, we developed a formalism in which we could incorporate the uncertainty-combining behavior of experts,[8] who were compiling a complex of background knowledge in such context-specific rules. It was also important to note that the chaining length was relatively small: Two or three steps were all that were used. If the problem called for much longer chaining, we took it as a sign that we were modeling the expert knowledge inaccurately, and sought additional pieces of knowledge that would shorten the chain.

Fuzzy set theory has done quite well as a formal mathematical system. Whether its theorems are interesting is a subjective opinion among mathematicians, but a large body of mathematical work exists. Where more work needs to be done is in establishing that fuzzy set theory actually captures something real and can make a pragmatic difference, for the right reasons.

I think Elkan has performed a service by initiating a debate about the properties of fuzzy set theory. I have argued that the points Elkan makes about fuzzy sets are really an instance of problems that apply to a number of other AI mechanisms and ideas, and specifically to many other proposals for subjective calculi for handling uncertainty. The issues raised are large in scope, and not only the fuzzy set community, but the AI community as a whole could benefit from giving them thought.

References

1. B. Chandrasekaran, "Generic Tasks in Knowledge-Based Reasoning: High-Level Building Blocks for Expert System Design," *IEEE Expert*, Vol. 1, No. 3, Fall 1986, pp. 23-30.
2. W.J. Clancey, "Heuristic Classification," *Artificial Intelligence* Vol. 27, No. 3, 1985, pp. 289-350.
3. B.G. Buchanan and E.H. Shortliffe, *Rule-Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project,* Addison-Wesley, Reading, Mass., 1984.
4. W. Kempton, "Category Grading and Taxonomic Relations: A Mug is a Sort of a Cup," *American Ethnologist* Vol. 5, No. 1, 1978, pp. 44-65; revised version reprinted in *Language, Culture, and Cognition: Anthropological Perspectives*, R.W. Casson, ed., Macmillan, New York, 1981.
5. W. Kempton, *The Folk Classification of Ceramics: A Study of Cognitive Prototypes*, Academic Press, San Diego, 1981.
6. A. Tversky and D. Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," *Science*, Vol. 185, 1974, pp. 300-306.
7. B. Chandrasekaran, "QP is More than SPQR and Dynamical System Theory," *Computational Intelligence*, Vol. 8, No. 2, 1992, pp. 216-222.
8. B. Chandrasekaran and S. Mittal, "Conceptual Representation of Medical Knowledge by Computer: MDX and Related Systems," *Advances in Computers*, Vol. 22, Academic Press, 1983, pp. 217-293.

B. Chandrasekaran is director of the Laboratory for AI Research and a professor of computer and information science at Ohio State University. His research interests include knowledge-based systems, using images in problem solving, and the foundations of cognitive science and AI. Chandrasekaran received his PhD from the University of Pennsylvania in 1967. He is editor-in-chief of *IEEE Expert*, a fellow of the IEEE and AAAI, and a member of the IEEE Computer Society. B. Chandrasekaran can be reached at the Department of Computer and Information Science, Ohio State Univ., 591 Dreese Labs, 2015 Neil Ave., Columbus, OH 43210-1277.

A Better Path to Duplicating Human Reasoning

Christopher J.S. deSilva and Yianni Attikiouzel, University of Western Australia

The paradox that arises from Elkan's Theorem 1 is mild in comparison to some of the logical problems that lurk behind the apparently innocent equations in Definition 1. In fact, although fuzzy logic has been promoted as a way of writing programs that carry out inference in the same way a person might, the equations of Definition 1 can lead inescapably to conclusions that no human being would accept.

Consider a simple example: You know that the airplane on which John Doe was traveling has crashed in some remote location, but you have no information about whether anyone on board has survived. In this situation, you might make the following assignment: t("John Doe is alive") = 0.5. The equations of Definition 1 would lead you immediately to t("John Doe is dead") = 0.5. While this is a reasonable assignment, it would in turn lead you to t("John Doe is both dead and alive") = 0.5. Thus, there is an element of truth in the statement "John Doe is both dead and alive." However, any rational person will argue that it is impossible for John Doe to be both dead and alive, so that the statement "John Doe is both dead and alive" must always be false, and have a truth value of zero.

We can imagine putting a fuzzy logic system to the Turing test on the matter of John Doe's well-being:

Interrogator: Is John Doe alive ?
Respondent 1: It is half-true that John Doe is alive.
Respondent 2: I don't know.

Interrogator: Is John Doe dead ?
Respondent 1: It is half-true that John Doe is dead.
Respondent 2: I don't know.

Interrogator: Is John Doe both dead and alive?
Respondent 1: It is half-true that John Doe is both dead and alive.
Respondent 2: It is impossible for John Doe to be both dead and alive.

While there is an element of caricature in this dialogue, it serves to highlight the problem. It is clear that if A is any proposition with a non-zero truth value, the equations of Definition 1 will lead to the conclusion that the truth value of the compound statement (A and (not A)) is also non-zero. This is a very simple example of how fuzzy logic diverges from human logic. It is to be expected that this divergence will increase with the complexity of the inference process.

Of course, people have been assigning truth values between zero and one to make inferences since the time of Laplace, on the basis of probability theory. As Cox has shown,[1] using the axioms of probability theory is essentially the only way to carry out this form of inference and remain consistent with human reasoning — any other way will lead to contradictions and inconsistencies. However, proponents of fuzzy logic appear to be unaware of Cox's work and that of Jaynes[2] and Tribus,[3] where the question of how to write programs that make inference based on incomplete knowledge is discussed.

As Cheeseman[4] pointed out for AI in general, the bottom line is that if you want to write a program or build a machine that will perform inference in the same way as people, then you must build the basic equations of probability theory into it, or face the inevitable outcome that it will not perform as required.

Perhaps the real paradox of fuzzy logic's success is that proponents hail it as a successful technology despite the fact that it is incapable of performing as they claim it can and does.

References

1. R.T. Cox, *The Algebra of Probable Inference*, Johns Hopkins Press, Baltimore, 1961.
2. E.T. Jaynes, "How Does the Brain do Plausible Reasoning?" Tech. Report 421, Microwave Laboratory, Stanford Univ., 1957.
3. M. Tribus, *Rational Descriptions, Decisions, and Designs*, Pergamon Press, New York, 1969.
4. P. Cheeseman, "An Inquiry into Computer Understanding," *Computer Intelligence*, Vol. 4, No. 2, Feb. 1988, pp. 58-66.

Yianni Attikiouzel is a professor of electrical and electronic engineering at the University of Western Australia where he is director of the Centre for Intelligent Information Processing Systems. His work has been published in and presented at more than 120 international journals and conferences, and he is the author of two books. He is a member of the Industry Research and Development Board of the Commonwealth Department of Science and Technology, and sits on its Services and Consumer Products Committee. Yianni Attikiouzel can be contacted at the Centre for Intelligent Information Processing Systems. Department of Electrical and Electronic Engineering. University of Western Australia, Nedlands, WA 6009 Australia; phone: 61 9 380 3134; fax: 61 9 380 1101; Internet: yianni@ee.uwa.edu.au

Christopher deSilva is a research fellow at the Centre for Intelligent Information Processing Systems at the University of Western Australia. He is currently working on the theory and application of artificial neural networks. His other research interests include syntactic pattern recognition and Bayesian inference. He can be reached at the Centre for Intelligent Information Processing Systems, Department of Electrical and Electronic Engineering, University of Western Australia, Nedlands, WA 6009 Australia; phone: 61 9 380 1765; fax: 61 9 380 1101; Internet: chris@ee.uwa.edu.au

Partial Truth is not Uncertainty

Fuzzy Logic versus Possibilistic Logic

Didier Dubois and Henri Prade, Université Paul Sabatier de Toulouse
Philippe Smets, Université Libre de Bruxelles

Charles Elkan has questioned fuzzy logic and cast serious doubts on the reasons for its success, arguing that "fuzzy logic collapses mathematically to two-valued logic." We completely disagree, and we especially object to two points:

(1) Elkan's proof uses too strong a notion of logical equivalence. The particular equivalence he considers, while valid in Boolean algebra, has nothing to do with fuzzy logic.
(2) Elkan claims that De Morgan's algebra "allows very little reasoning about collections of fuzzy assertions," although he correctly states that when logical equivalence is restricted to De Morgan algebra equalities' "collapse to two truth values is avoided."

Furthermore, Elkan fails to understand the important distinction between two totally different problems that fuzzy-set-based methods address.[1] These are the handling of *gradual* (thus non-Boolean) properties whose satisfaction is a matter of degree (even when information is complete) on the one hand, and the handling of uncertainty pervading Boolean propositions, the uncertainty being induced by incomplete states of knowledge that are represented by means of fuzzy sets, on the other hand.[2] The first problem requires the plain use of fuzzy sets, while the second is the realm of possibility theory[3,4] and possibilistic logic[5]. We now discuss in greater detail the points above and the distinction between truth functional fuzzy (multivalued) logic and non-fully compositional possibilistic logic.

Fuzzy logic equivalence is *not* classical. Elkan claims that in fuzzy logic, four requirements hold for any assertions A and B, t being a truth assignment function such that $\forall A$, $t(A) \in [0,1]$:

$$t(A \wedge B) = \min(t(A), t(B)) \quad (1)$$
$$t(A \vee B) = \max(t(A), t(B)) \quad (2)$$
$$t(\neg A) = 1 - t(A) \quad (3)$$
$t(A) = t(B)$ if A and B are logically equivalent. (4)

While Equations 1–3 are indeed the basic relations governing degrees of truth in fuzzy logic (as well as fuzzy set membership degrees) as proposed by Zadeh,[6] Equation 4 (where "logically equivalent" is understood in a stronger sense than the equivalences induced by 1–3) has never been seriously considered by any author in the fuzzy-set literature. (There are, as can be expected, a few erroneous attempts at the subject in a corpus of more than 10,000 published papers). Obviously, some classical logic equivalences still hold with fuzzy assertions obeying Equations 1–3, namely, those allowed by the De Morgan structure induced by 1–3, such as

$$A \wedge A \equiv A \ ; \ A \vee A \equiv A \ \text{(idempotency)}$$

$$A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C) \ ;$$
$$A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$$
(distributivity)

But other Boolean equivalences *do not* hold, for instance:

$$A \wedge \neg A \not\equiv \bot$$

since Equations 1 and 3 entail *only*

$$t(A \wedge \neg A) = \min(t(A), 1 - t(A)) \leq 1/2;$$

and

$$A \vee \neg A \not\equiv \top$$

since Equations 2 and 3 entail *only*

$$t(A \vee \neg A) = \max(t(A), 1 - t(A)) \geq 1/2$$

where $t(\bot) = 0$ and $t(\top) = 1$. Indeed, as many authors have emphasized, the failure of contradiction and excluded-middle laws is typical of fuzzy logic. This is natural with gradual properties like "tall." For example, in a given context, somebody who is 1.75 meters high might be considered neither as completely tall (tall with degree 1) nor as completely not tall (tall with degree 0). In this case, we might have, for example, $\mu_{\text{tall}}(1.75) = 0.5 = \mu_{\neg\text{tall}}(1.75)$.

To establish the collapse of fuzzy logic to binary logic, Elkan uses the logical equivalence

$$\neg(A \wedge \neg B) \equiv B \vee (\neg A \wedge \neg B) \quad (5)$$

postulated as being "plausible intuitively." If Equations 1–3 hold, the left-hand part of Equation 5 can be equivalently written in fuzzy logic as

$$\neg(A \wedge \neg B) \equiv \neg A \vee B$$

while the right-hand part can be equivalently written as

$$B \vee (\neg A \wedge \neg B) \equiv (\neg A \vee B) \wedge (B \vee \neg B),$$

which clearly relates to the excluded-middle law. Thus, it is expected that Equation 5 fails to hold in fuzzy logic — and indeed it can be checked, using Equations 1–3, that a counterexample to Equation 5 is provided by $t(A) = 0$, $t(B) = 0.5$, for instance. Thus, Elkan's claim of "a paradox in fuzzy logic" relies only on faulty assumptions, or at best on a logical equivalence, the rationale of which is far from natural in the scope of fuzzy logic.

Gradual and interpolative reasoning. Fuzzy logic is concerned with the handling of assertions like "John is tall" — assertions whose truth is a matter of degree due to gradual predicates within them. The degree of truth of compound expressions can be easily computed using Equations 1–3. (Although we restrict ourselves here to the operators minimum, maximum, and complement to one, there is a panoply of others[7,8] that enable us to model different kinds of AND and OR operations between properties in a multicriteria aggregation perspective.)

More than 20 years ago, R.C.T. Lee[9] provided the basic machinery for reasoning in fuzzy logic by extending the resolution rule in accordance with Equations 1–3. He established that if all the truth values of the parent clauses are greater than 0.5, then a resolvent clause derived by the resolution principle always has a truth-value between the maximum and the minimum of those of the parent clauses.

We can also use an implication operator to model "gradual rules,"[10] which express knowledge of the form "the more X is A, the more Y is B," such as, "the taller you are, the heavier you are." This is captured by the implication defined by

$$t(A \to B) \begin{aligned} &= 1 \text{ if } t(A) \le t(B) \\ &= 0 \text{ if } t(A) > t(B) \end{aligned} \qquad (6)$$

This implication is the natural counterpart of Zadeh's fuzzy set inclusion defined by the pointwise inequality of the membership functions.[6] It is also directly associated with Equations 1–3, since $A \to B \equiv \mathrm{T}$ if and only if $A \wedge B \equiv A$. Such an implication expresses a purely gradual relationship and has nothing to do with uncertainty. Besides, Takagi and Sugeno[11] have proposed an interpolation mechanism between n rules with fuzzy condition parts and nonfuzzy conclusions of the form "if X is A_i and Y is B_i then $Z = c_i$", by computing the following output when $X = x_0$ and $Y = y_0$ is observed

$$Z = \frac{\Sigma_i \gamma_i \cdot c_i}{\Sigma_i \gamma_i} \qquad (7)$$

where $\gamma_i = \min(\mu_{A_i}(x_0), \mu_{B_i}(y_0))$, $i = 1,n$. Again, this kind of "inference" (which is widely used in fuzzy control) has nothing to do with uncertainty handling, since only an interpolation between typical conclusions is performed, based on degrees of similarity between the input (x_0, y_0) and the prototypical values in the core of the fuzzy set $A_i \times B_i$. This similarity is measured by the coefficients γ_i which cannot be considered as degrees of uncertainty in any case. In spite of its apparently ad hoc nature, Equation 7 can be justified with one-premised rules using Equation 6 and viewing the rules as expressing "the more X is A_i and Y is B_i, the *closer* Z is to c_i" and using appropriately shaped membership functions.[12]

As this shows, contrary to Elkan's claim, some kinds of reasoning, as exemplified by Takagi and Sugeno's, and Lee's methods, can be handled in a De Morgan algebra framework.

Possibility theory and uncertainty. In addition to modeling the gradual nature of properties, fuzzy sets can be used to represent incomplete states of knowledge. In this second use, the fuzzy set plays the role of a possibility distribution that provides a complete ordering of mutually exclusive states of the world according to their respective levels of possibility or plausibility. For instance, if we know only that "John is tall" (but not his precise height), where the meaning of "tall" is described, in context, by the membership function of a fuzzy set (that is, μ_{tall}), then the greater $\mu_{\mathrm{tall}}(x)$ is, the greater the possibility that height(John) = x; the smaller $\mu_{\mathrm{tall}}(x)$ is, the smaller this possibility.

Given a [0,1]-valued possibility distribution π describing an incomplete state of knowledge, Zadeh[4] defines a so-called possibility measure Π such that

$$\Pi(A) = \sup\{\pi(x), x \text{ makes } A \text{ true}\} \qquad (8)$$

where A is a *Boolean* proposition (a proposition that can only be true or false). It can be easily checked that for Boolean propositions A and B, we have

$$\Pi(A \vee B) = \max(\Pi(A), \Pi(B)) \qquad (9)$$

but that we only have the inequality

$$\Pi(A \wedge B) \le \min(\Pi(A), \Pi(B)) \qquad (10)$$

in the general case (equality holds when A and B are *logically independent*). Indeed if $B \equiv \neg A$, $\Pi(A \wedge B) = \Pi(\perp) = 0$, while $\min(\Pi(A), \Pi(\neg A)) = 0$ only if the information is sufficiently complete for having either $\Pi(\neg A) = 0$ (A is true) or $\Pi(A) = 0$ (A is false). If nothing is known about A, we have $\Pi(A) = \Pi(\neg A) = 1$. By duality, a necessity measure N is associated to Π according to the relation (which can be viewed as a graded version of the relation between what is necessary and what is possible in modal logic)

$$N(A) = 1 - \Pi(\neg A) \qquad (11)$$

which states that A is all the more necessarily true as $\neg A$ has a low possibility to be true. It entails

$$N(A \wedge B) = \min(N(A), N(B)) \qquad (12)$$

and

$$N(A \vee B) \ge \max(N(A), N(B)). \qquad (13)$$

Equations 9, 11, and 12 should not be confused with Equations 2, 3, and 1, respectively. In 9, 11, and 12 we deal with Boolean propositions pervaded with uncertainty due to incomplete information, while 1–3 pertain to non-Boolean propositions whose truth is a matter of degree (the information being assumed to be complete). Very often, discussions about fuzzy expert systems or uncertain knowledge base systems get confused because of a lack of distinction between degrees of truth and degree of uncertainty. Fuzzy logic, as understood by Elkan, is a logic where the truth status of propositions is multiple-valued; that is, there are intermediary truth values between true and false (like "very true," "rather true," and so on). On the contrary, degrees of uncertainty apply to all-or-nothing propositions, and do not model truth values but express the fact that the truth value (true or false) is unknown. The uncertainty degrees then try to assess which one of "true" or "false" is the most plausible truth value. This distinction was made by one of the founders of subjective probability theory — De Finetti[13] — but with a few exceptions (including ourselves) it has been quite forgotten by the AI community in general and by Elkan in particular. Still, we consider this distinction a crucial prerequisite in any discussion about fuzzy sets and possibility theory and their use in automated reasoning.

Observe also that neither Π nor N are fully compositional with respect to $\wedge$, $\vee$, and $\neg$. This is not surprising, since the only way to map a Boolean structure on [0,1] by a fully compositional mapping f is to have $f(A)$ equal to 0 or to 1 for any A.[1] Truth-functionality in Equations 1–3 is preserved only by having A and B elements of a weaker structure, namely, a De Morgan algebra. Thus, logics of uncertainty cannot be fully compositional with respect to uncertainty degrees. This point is also recognized by Elkan in the case of probability measures, and dates back at least to De Finetti in the 1930s! Partial compositionality is possible, however; probabilities are compositional with respect to negation, possibilities with respect to disjunction, necessities with respect to conjunction. Based on his article, however, it seems that Elkan has not heard about possibility theory, which is another side of fuzzy sets.

Let us consider Elkan's watermelon example:

$$watermelon(x) \equiv redinside(x) \wedge greenoutside(x)$$

It is supposed that "for some melon m, evidence that m is red internally has strength 0.5, and m is green externally with strength of evidence 0.8." It is not clear what Elkan means by "strength of evidence" in the light of the above comments. We shall assume they are indeed degrees of uncertainty, rather than degrees of red and degrees of green. But then the only way to anchor this discussion in the fuzzy logic debate is to interpret these degrees in possibility theory. Elkan's watermelon sentence can be understood as $N(redinside(m)) \geq 0.5$ and $N(greenoutside(m)) \geq 0.8$, expressing that the available information makes us certain to the degree of 0.5 that m is red inside, and to the degree 0.8 that it is green outside. A direct application of Equation 12 leads to $N(watermelon(m)) \geq \min(0.5, 0.8) = 0.5$, a result also obtained under an equality form by Elkan by applying Equation 1 in an inappropriate way. However, he would like to conclude that "m is a watermelon with strength of evidence over 0.5." This seems a strange requirement, and one that a probabilistic model would not satisfy either (since $\mathrm{Prob}(A \wedge B) \leq \min(\mathrm{Prob}(A),\mathrm{Prob}(B))$. Indeed, we are not in a data fusion situation where two independent sources provide the same conclusion with various strengths,[14] but in a situation where the logical conjunction of two conditions is required to conclude that m is a watermelon (namely the inside redness of m and its outside greenness). Note that in case we have both $N(A) \geq \alpha$ and $N(A) \geq \alpha'$ as obtained from distinct arguments, we shall conclude that $N(A) \geq \max(\alpha,\alpha')$.

Reasoning with possibility theory. In possibilistic logic, first-order logic formulas are weighted by lower bounds of necessity or possibility measures, which reflect the uncertainty of the available information. Possibilistic logic[2,5] has been developed both at the syntactic level, where there is an inference machinery based on extended resolution and refutation (the lower bound of the resolvent clause necessity is the minimum of the lower bounds of parent clauses necessity measures), and at the semantic level, where a semantics in terms of a possibility distribution over a set of classical interpretations has been proved to be sound and complete with respect to the syntax. Due to the fact that a possibility distribution encodes a preferential ordering over a set of possible interpretations, possibilistic logic has been shown to capture an important class of nonmonotonic reasoning consequence relations[15] and has capabilities for handling partial inconsistency in knowledge bases.[5] Moreover, possibilistic assumption-based truth maintenance systems[16] based on possibilistic logic have been defined for dealing with uncertain justifications and ranking environments in a label; they have been successfully applied to a data-fusion application.[17]

However, possibility theory offers more general applications to reasoning with uncertain, imprecise, or fuzzy pieces of information by manipulating possibility distributions explicitly. An example of these reasoning capabilities is provided by the so-called generalized modus ponens,[18] which from a fuzzy fact "X is A'" (represented by a possibility distribution $\pi_X = \mu_{A'}$) and a fuzzy rule "if X is A then Y is B" (also represented by a possibility distribution $\pi_{Y|X}$), enables us to infer the possibility distribution restricting the possible values of Y by combining π_X and $\pi_{Y|X}$ and projecting the result on the domain of the variable Y. According to the multiple-valued logic implication $\rightarrow$ used to compute $\pi_{Y|X}$ from μ_A and μ_B, different kinds of fuzzy rules can be modeled. In particular, we can distinguish, for example, between the purely gradual rules already mentioned (of the form "the more X is A, the more Y is B") and certainty rules of the form "the more X is A the more *certain* Y is B." Thus, graduality can also be encountered in the expression of incomplete knowledge states pertaining to little-known relationships between variables (like the ones expressed by fuzzy rules).[2]

Expert systems with fuzzy rules have been designed that are not as simple as fuzzy controllers (where no chaining of rules is required, but only an interpolation between the conclusions of a parallel rules set). These expert systems, as expected by Elkan, do "knowledge-intensive tasks such as diagnosis, scheduling, or design," and include Cadiag-2,[19] Taiger,[20] RUM,[21] Milord,[22] OPAL.[23] All these systems were or are used in applications in one of the above-mentioned fields. These systems use some form of fuzzy set or possibility-theory-based inference mechanisms that is much more sophisticated than the three formulas proposed by Zadeh in 1965 (Equations 1–3) — and to which fuzzy set and possibility theory methods cannot be reduced. There are many other important works on fuzzy set and possibility theory-based inference systems in temporal, qualitative, and abductive reasoning, that, for the sake of brevity, we do not mention here.

Fuzzy logic is not as simple as Elkan seems to believe. In this respect, the absence of any mention in Elkan's discussion of Zadeh's possibility theory and approximate reasoning approach[4,18] is quite revealing.

In the literature, the expression "fuzzy logic" usually refers either to multiple-valued logic (as in the first part of Elkan's paper) or to fuzzy controllers. However, the two domains have very little in common, due to the fact that control engineers usually do not know about logic, and logicians do not know about control. In that sense, the first part of Elkan's article has very little relevance to his discussion on fuzzy control.

If the success of fuzzy logic is paradoxical, it is certainly not because of Elkan's collapsing property. More importantly, Zadeh's view of fuzzy logic seems to go far beyond multiple-valued logic, and is as much a framework for handling incomplete information as a methodology for capturing graduality in propositions. The concept of fuzzy truth values refers as much to the idea of a partially unknown truth value as to intermediate truth values. This is why we have emphasized the crucial distinction between the truth-functional handling of gradual properties and the possibilistic treatment of uncertainty (which is not fully compositional).

It is certainly true that the huge quantity of fuzzy set literature — whose quality is unavoidably inconsistent — does not contribute much toward helping newcomers have a synthetic, well-informed, and balanced view of the domain. Fuzzy controllers have encountered great success by providing an efficient way of implementing an interpolative mechanism, not only in small, but also in very large and complex problems. However, this should not obscure other existing applications, and the great potential of fuzzy set and possibility theory for AI applications in general.

References

1. D. Dubois and H. Prade, "An Introduction to Possibilistic and Fuzzy Logics," *Non-standard Logics for Automated Reasoning,* Academic Press, New York, 1988, pp. 287-326.
2. D. Dubois, H. Prade, and J. Lang, "Fuzzy Sets in Approximate Reasoning," *Fuzzy Sets and Systems,* Vol. 40, No. 1, March 1991. pp. 143-244.
3. D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York, 1988.
4. L.A. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility," *Fuzzy Sets and Systems,* Vol. 1, No. 1, Jan. 1978, pp. 3-28.
5. D. Dubois and H. Prade, "Epistemic Entrenchment and Possibilistic Logic," *Artificial Intelligence,* Vol. 50, No. 2, July 1991, pp. 223-239.
6. L.A. Zadeh, "Fuzzy Sets," *Information and Control,* Vol. 8, No. 4, June 1965, pp. 338-353.
7. R.R. Yager, "Connectives and Quantifiers in Fuzzy Sets," Fuzzy Sets and Systems, Vol. 40, No. 1, Mar. 1991, pp. 143-244.
8. D. Dubois, H. Prade, and R.R. Yager, eds., *Readings in Fuzzy Sets for Intelligent Systems,* Morgan Kaufmann, San Francisco, Calif., 1993.
9. R.C.T. Lee, "Fuzzy Logic and the Resolution Principle," *J. ACM,* Vol. 19, No. 1, Jan. 1972, pp. 109-119.
10. D. Dubois and H. Prade, "Gradual Inference Rules in Approximate Reasoning," *Information Sciences,* Vol. 61, No. 1-2, Apr., 1992, pp. 103-122.
11. T. Takagi and M. Sugeno, "Fuzzy Identification of Systems and its Applications to Modeling and Control," *IEEE Trans. Systems, Man and Cybernetics,* Vol. 15, No. 2, 1985, pp. 116-132.
12. D. Dubois and H. Prade, "Possibility Theory as a Basis for Preference Propagation in Automated Reasoning," *Proc. First IEEE Int'l Conf. Fuzzy Systems* (FUZZ-IEEE'92), IEEE Press, Piscataway, N.J., 1992, pp. 821-832.
13. B. De Finetti, "La Logique de la Probabilité," *Actes du Congrès Int'l. de Philosophie Scientifique,* Paris, 1935, Hermann et Cie Editions, 1936, pp. IV1-IV9.
14. D. Dubois and H. Prade, "Combination of Fuzzy Information in the Framework of Possibility Theory," in *Data Fusion in Robotics and Machine Intelligence,* Academic Press, New York, 1992, pp. 481-505.
15. D. Dubois and H. Prade, "Possibilistic Logic, Preferential Models, Nonmonotonic and Related Issues," *Proc. Int'l Joint Conf. Artificial Intelligence* (IJCAI '91), Morgan Kaufmann, San Francisco, Calif., 1991, pp. 419-424.
16. D. Dubois, J. Lang, and H. Prade, "A Possibilistic Assumption-Based Truth Maintenance System with Uncertain Justifications, and its Application to Belief Revision," *Lecture Notes in Artificial Intelligence,* Vol. 515, Springer-Verlag, Berlin, 1990, pp. 87-106.
17. F.F. Monai and T. Chehire, "Possibilistic Assumption-Based Truth Maintenance System: Validation in a Data Fusion Application," *Proc. Eighth Conf. Uncertainty in AI,* 1992, Morgan Kaufmann, San Francisco, Calif., 1992, pp. 83-91.
18. L.A. Zadeh, "A Theory of Approximate Reasoning," *Machine Intelligence,* Vol. 9, John Wiley & Sons, New York, 1979, pp. 149-194.
19. K.P. Adlassnig and G. Kolarz, "CADIAG-2: Computer-Assisted Medical Diagnosis Using Fuzzy Subsets," *Approximate Reasoning in Decision Analysis,* North-Holland, Amsterdam, 1982, pp. 219-247.
20. H. Farreny, H. Prade, and E. Wyss, "Approximate Reasoning in a Rule-Based Expert System Using Possibility Theory: A Case Study," *Proc. Information Processing '86,* North-Holland, Amsterdam, 1986, pp. 407-413.
21. P.P. Bonissone, S.S. Gans, and K.S. Decker, "RUM: A Layered Architecture for Reasoning with Uncertainty," *Proc. Int'l Joint Conf. Artificial Intelligence* (IJCAI), Morgan Kaufmann, San Francisco, Calif., 1987, pp. 891-898.
22. L. Godo et al., "Milord: The Architecture and the Management of Linguistically Expressed Uncertainty," *Int'l J. Intelligent Systems,* Vol. 4, No. 4, Winter 1989, pp. 471-501.
23. E. Bensana, E. Bel, and D. Dubois, "OPAL: A Multi-Knowledge-Based System for Industrial Job-Shop Scheduling," *Int'l J. Product Research,* Vol. 26, No. 5, 1988, pp. 795-819.

Didier Dubois is a research scientist at IRIT, University of Toulouse, and is director of research at the French National Center for Scientific Research. His research interests focus on the modeling of uncertainty in various branches of information technology, especially AI, decision analysis, and operations research. He is the coauthor of numerous technical papers, and coauthor with Henri Prade of a book on fuzzy sets (Academic Press, 1980) and a book on possibility theory (Plenum Press, 1988). He is a member of the Léa Sombé research group, which wrote *Reasoning with Incomplete Information*, and *Belief Revision and Updating* (John Wiley & Sons, 1990 and 1994). Dubois is on the editorial board of *Fuzzy Sets and Systems*, the *Journal of Intelligent Manufacturing*, the *International Journal of Approximate Reasoning*, and the *International Journal of General Systems*, among others. He is also an associate editor of *IEEE Transactions on Fuzzy Systems*. He received a doctorate degree in engineering in 1977 from ENSAE, Toulouse, and the Doctorat d'Etat in 1983 from Grenoble University. Didier Dubois can be reached at IRIT, Univ. Paul Sabatier, 31062 Toulouse Cedex, France; Internet: dubois@irit.fr

Henri Prade is a researcher at IRIT, University of Toulouse, and director of research at the French National Center for Scientific Research. His research interests include fuzzy sets and possibility theory, approximate reasoning nonclassical logics, AI and database systems, and operations research. He is the coauthor (with Didier Dubois) of a book on fuzzy sets (Academic Press, 1980) and a book on possibility theory (Plenum Press, 1988), and he has coedited three other books and published more than 200 technical papers. He is also a member of the Léa Sombé research group, authors of *Reasoning with Incomplete Information*, and *Belief Revision and Updating* (Wiley, 1990 and 1994) Prade is an associate editor of *IEEE Transactions on Fuzzy Systems* and is on the editorial board of several journals, including *Fuzzy Sets and Systems*, the *International Journal of Approximate Reasoning*, the *Information Sciences Series on Intelligent Systems*, and the *International Journal of Intelligent Systems*. He received a doctorate degree in engineering in 1977 from ENSAE, Toulouse, and the Doctorat d'Etat in 1982 from Paul Sabatier University, Toulouse. Henri Prade can be reached at IRIT, Univ. Paul Sabatier, 31062 Toulouse Cedex, France; Internet: prade@irit.fr

Philippe Smets is coordinator of the Institut de Recherches Interdisciplinaires et de Developpements en Intelligence Artificielle at the Université Libre de Bruxelles, where he is also a professor of medical statistics and director of the Laboratory for Medical Statistics. His research interests are in approximate reasoning, Bayesian theory, fuzzy set theory, and belief functions. He received his MD from the Université Libre de Bruxelles in 1963, his master's degree in experimental statistics from North Carolina State University, and his Ph.D. in medical statistics from the Université Libre de Bruxelles in 1978. Philippe Smets is on the editorial board of the *International Journal for Approximate Reasoning* and *Fuzzy Sets and Systems*. Philippe Smets can be reached at IRIDIA, Univ. Libre de Bruxelles, CP 194/6, 1050 Bruxelles, Belgium; Internet: psmets@ulb.ac.be

RESPONSE

Fuzzy Logic
An Interface Between Logic and Human Reasoning

Christian Freksa, University of Hamburg, Germany

Charles Elkan addresses two distinct areas of fuzzy logic: formal expressiveness and practical usefulness. He describes as a paradox that although the theory of fuzzy logic is not generally accepted, it is successfully used in many real-world applications. He also calls paradoxical the fact that these applications are predominantly found in the control domain.

I will not discuss here the alleged equivalence between fuzzy and two-valued logic; by choosing criteria established for the more restricted two-valued formalism, Elkan does not have a suitable framework for a meaningful comparison. To point out prerequisites for the practical usefulness of knowledge representation formalisms, I will focus on the role of fuzzy logic in linking two formally incommensurable worlds: the natural world of human perception and experience that leads to subjective cognitive concepts, and the formal world of classical logic that yields universal truth conditions.

Given the premise that there is no one-to-one mapping between human conceptual structures and the framework of classical logic, it is not important for the analysis of a formal representation structure if two logically equivalent expressions are evaluated identically; what we have to ensure is that derivations accepted in human reasoning can also be derived in our formalism.

Classical logic and human knowledge. In AI, propositions and various kinds of logic formalisms serve to represent and derive knowledge about formal or real domains. Traditionally, most effort has been put into the development of logically correct and consistent operations *within the formal representation;* however, little attention has been paid to the correspondence problem between the structure of these propositions and operations, on one hand, and the knowledge structure they are supposed to represent, on the other. When we represent formal domains (for example, card games or mathematical theorems), establishing this correspondence may not cause major problems. However, when we represent knowledge about a real domain, the correspondence between our formalism and the represented structure becomes a major issue.

A representation system consists of:

- a represented world, and the relations and operations in it;
- a representing world, and the relations and operations in it; and
- the correspondence between the two worlds.[1]

When representing knowledge about the *real* world, it is inherently impossible to prove something about the represented real-world knowledge; this part of the representation system is outside the formalism. We only can prove something within the representing formalism. Thus, the represented real world and its representation are *formally* incommensurable.

In expert systems, the knowledge engineer establishes the correspondence between the real and formal worlds, but he cannot prove its correctness; he depends on his perception and intuition to determine the equivalence between the two. Usually, a knowledge engineer relies upon assumptions to determine the validity of operations on a representation. These assumptions stem from his knowledge about formal logic, rather than from knowledge about specific properties of human reasoning. Nevertheless — as Elkan's article shows — this approach appears to be widely accepted for the treatment of human knowledge.

One of Lotfi Zadeh's main motivations for introducing the notions of fuzzy sets and fuzzy logic was his observation that real-world knowledge generally has a different structure and requires different formalization than existing formal systems. Contrary to established practice, a one-to-one correspondence between natural-language propositions and predicate calculus propositions can be shown to be inadequate. In particular, the instantaneous switch from truth to falsity can easily distinguish propositions in classical logic from those in natural language. In addition, numerous assumptions of the formally correct treatment of the propositions cannot be established in the corresponding source knowledge.

The fuzzy logic interface. Zadeh recognized the power of a formal approach to knowledge processing as well as the advantages of using soft knowledge in human reasoning. He thus took a first step in incrementally relaxing constraints imposed on existing formalisms to accommodate important properties of natural inference. This step was to generalize the classical notion of a set to the notion of a fuzzy set that allowed gradual membership. The choice of numerical degrees of membership was largely made for formal reasons: it provided a transparent way of formally treating the new notion. Using the familiar language of mathematics, the theory can easily be implemented in computer systems, while at the same time offering a better approximation to the associated human concepts.

Because human notions and concepts form the basis for reasoning in expert systems, the success of these systems depends upon the correspondence relation between human concepts and their formalization.

Studying the formal properties of the representation is insufficient.

Zadeh realized that it was much more important to have a good model of the semantics of human concepts and perform reasonable operations than to have a bad model and perform verifiably correct operations. He never insisted that his initial proposal for a fuzzy logic should be viewed as the final solution for representing human knowledge about the world; rather, he offered a model based on established notions that could easily be grasped by engineers and researchers alike as a step toward formalizing human reasoning. Because of this, Zadeh's basic notion of a fuzzy set stimulated enormous research activity in soft knowledge processing.

Zadeh's work also helped establish a radically different view of the status of expert knowledge. No longer is it viewed as a collection of absolute truths piped into an inference engine to derive all sorts of unexpected results; rather, it is now considered as a system of more or less soft constraints that are applied to specific situations to make reasonable decisions.

Soft knowledge is processed differently than logic clauses — the reasoning power is typically due to processing breadth rather than depth. The ability to use shallow processing to merge knowledge from different sources produced useful decisions. (Elkan uses the terms "deep" and "shallow" in two different senses: to distinguish general knowledge from specific knowledge, and to distinguish extensive and restricted knowledge propagation. I use the terms here in the second sense, which is the usual sense.) Elkan appears to attribute the fact that fuzzy systems employ only a few rules to the domain's simplicity. However, this fact can also be attributed to the important capability of summarizing complex knowledge into a dense and transparent description.

Success and limitations. The fuzzy set paradigm introduced a new concept of soft knowledge that helped characterize an important aspect of knowledge about complex environments. It also provides a language to bridge the gap between soft and shallow knowledge, on the one hand, and systematic and formal methods for dealing with it, on the other. This contribution might have a much more significant impact on human thought and the role of classical logic in systems analysis than the fuzzy set notion will have on the success of expert systems.

As the transition from crisp sets to fuzzy sets is a rather moderate step toward accounting for the nature of human concepts, we should not expect it to solve all our problems. In particular, fuzzy sets and fuzzy logic do not answer the fact that human concepts develop and are modified in an open world, while formal concepts are fixed in closed worlds, for the most part. Therefore, it is not surprising that successful applications of fuzzy logic are so far found mainly in well-defined closed domains like control problems which, to a large extent, share the properties of synthetic, formal problems. The way gradual membership is represented in fuzzy sets quite naturally suits such application domains.

The further we move from representing human knowledge about clearly delineated problems to representing concepts relating to open domains, the more we will have to overcome certain rigidities of the classical formal approaches.

Classical logic has proved extremely useful for solving formal problems specified in two-valued terms. Fuzzy logic is proving particularly useful for quasi-formal problems involving gradual transitions between various system states. For adequately formalizing less rigid domains, like the open world of human fuzzy concepts, we must relax the constraints on the formalisms even more. Specifically, numerical graduation of membership used in classical fuzzy logic is hardly justified for the representation of cognitive concepts; instead, less constraining ordering relations like partial orderings may be appropriate.

Considering the fact that it took 25 years to put fuzzy logic into wide use in the well-understood engineering domain of control, we should not be surprised if some barriers must be removed before fuzzy logic will be widely applied to more delicate areas of fuzzy reasoning.

For judging the quality of a representation formalism, I have proposed taking a representation-theoretical viewpoint: The correspondence between the represented domain and the formalism is at least as important as the representation's formal properties taken by themselves. This viewpoint permits a high-level characterization of the overall representation problem. I have also argued that real-world knowledge and formal knowledge are formally incommensurable. As long as the laws of human reasoning are not well understood, a good model of human reasoning should be expected to preserve some paradoxes; experimentation with the model may deepen the understanding and help resolve them.

Acknowledgments

I acknowledge stimulating discussions on this topic at the Tasso workshop 1993 in Bonn; at the panel discussion on Fuzzy Logic and AI at IJCAI '93 in Chambery, France; at the GI-Workshop "Fuzzy-Systeme '93" in Braunschweig, Germany; and valuable comments by Gerhard Dirlich.

References

1. S.E. Palmer, "Fundamental Aspects of Cognitive Representation," *Cognition and Categorization,* Lawrence Erlbaum, Hillsdale, N.J., 1978, pp. 259-303.

Christian Freksa is a professor in the Computer Science Department and the Cognitive Science Graduate Program at the University of Hamburg, Germany. His research interests include qualitative spatial and temporal reasoning, and representing incomplete, uncertain, and fuzzy knowledge about the physical world. He studied computer science and AI at the Technical University of Munich, the University of San Francisco, and the University of California, Berkeley, where he received his Ph.D. in 1981. His work has appeared in several journals and anthologies, including *Approximate Reasoning in Decision Analysis,* (North-Holland, 1982), *Artificial Intelligence Methodology, Systems, Applications,* (North-Holland, 1985), and in the *Artificial Intelligence Journal* (1992) Christian Freksa can be reached at the Fachbereich Informatik, Universität Hamburg, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany; fax 49-40-54715-385; Internet: freksa@informatik.uni-hamburg.de

Known Concerns About Fuzzy Logic

Oscar N. Garcia, George Washington University

I thank Charles Elkan for bringing into the open questions about fuzzy inferences that seem to bother him and others. I hope the result of this discussion will be a clearer understanding of many-valued logics in general, and fuzzy logic in particular. My comments address three topics: questions about Theorem 1, the "watermelon" example, and the issue of fuzzy logic in control.

Much of the confusion surrounding Theorem 1 stems from its rather unclear statement. I interpret the theorem to say the following:

"Let $f1(A,B) = \neg\,(A \text{ AND } \neg B)$ and $f2(A,B) = B \text{ OR } (\neg A \text{ AND } \neg B)$. Using Definition 1, *if one were to require* the following four equivalences —

(1) $f1(A,B) \leftrightarrow f2(A,B)$ AND
(2) $f1(\neg A,B) \leftrightarrow f2(\neg A,B)$ AND
(3) $f1(A,\neg B) \leftrightarrow f2(A,\neg B)$ AND
(4) $f1(\neg A,\neg B) \leftrightarrow f2(\neg A,\neg B)$

— then such a logic system would also require that $t(A) = t(B)$ or that $t(A) = 1 - t(B)$"

I can prove this supposition or "theorem" following the valuation t of Definition 1 for values of A and B in the interval [0,1]. Such valuation yields validity for the first of the four equivalences above except for $0 < t(A) < t(B) < 1 - t(A)$ when $t(f1) - t(f2)$ has the value $t(B) - t(A)$ if $t(B) < 1/2$ and the value $1 - t(A) - t(B)$ if $t(B) > 1/2$. The area where the equivalence is not satisfied is an isosceles triangle in the square $[1,0] \times [1,0]$ not including the isosceles sides. Similarly, for the other three equivalences, the non-overlapping triangles where the equivalences are not satisfied would cover the whole unit square — except for the isosceles sides, which constitute the two diagonals of the square. Thus, either $t(A) = t(B)$ in one diagonal of the square, or $t(A) = 1 - t(B)$ in the other.

Where is the catch? First, the fourth line of Definition 1 in Elkan's paper indicates that each side of a "logically equivalent" formula has the same evaluation. This is not a fair imposition, and Elkan need not choose such a formula to make his point. Just consider requiring the valuation of two "logically equivalent" formulas:

(5) $t(A \text{ AND } \neg A) \leftrightarrow t(\neg(A \text{ OR } \neg A))$

which, of course, only occurs in the bivalent case following Definition 1. Equivalences are tautologies, and while the arguments of t on each side of equivalence 5 are "logically equivalent" in classical logic, they are not so in fuzzy logic where the law of the excluded middle does not hold. Thus, it is not surprising that the attempt to evaluate these formulas using the classical bivalent logic interpretation of "logical equivalence" would not yield sound results. It can be easily shown that the manipulation of $f1$ or $f2$ in classical logic leads to a disjunction of a variable and its complement. We should not take a tautology that supports a rule base in one logic, use it in another logic that does not support that tautology, and expect it to work — and then go on to claim that a "collapse" of one logic to another has been proved. The requirement of "logical equivalence" in Definition 1 is therefore suspect. Elkan raises the question of why it is that intuitionistic logic is capable of rejecting the law of the excluded middle while fuzzy logic is not. While this is not directly relevant to the claimed "collapse," it is clear that intuitionistic logic is not used to the extent that fuzzy logic is used in controller design.

Another issue that might be troubling Elkan — implicit in his choice of the functions called f1 and f2 in my interpretation of his Theorem 1 — is what deductive tautologies (those involving implication, and particularly those known as the inferential implication tautology[1]) should be used in fuzzy logic if limited by Definition 1. This question is worthy of investigation, and has led to multiple alternatives to Zadeh's original definition of implication; however, it is beyond my concern here. My acquaintance with expert systems applications indicates that, in practice, value sets are categorized as designated (truth-like), antidesignated (false-like), and neutral (those for which insufficient knowledge exists for the model to be useful). A typical example for the real interval [0,1] would be antidesignated $A = [0,0.4)$, neutral $N = [0.4, 0.6]$, and designated $D = (0.6,1]$. (The complement of a designated value is antidesignated and vice-versa, while neutral values are the complements of other neutral values.) The object of expert systems is to mimic, as closely as possible, the reasoning of expert humans in terms of the best causality relations known to them, and to incorporate them in a knowledge-based model, often represented as a rule base.

Because fuzzy logic is known as a "normal" logic (all of its truth-assuming tautologies are included in classical logic, though the converse is not true, as can be shown in the case of equivalence 5) the puzzlement shown by Elkan is not novel. Indeed, the tautology involving the law of the excluded middle from classical bivalent logic does not hold in fuzzy logic, nor in many other normal many-valued logics. For those systems in which operations involving the designated and anti-designated defined sets coincide with those of classical logic, Shaw has enumerated the possible homomorphisms of any ordered designated system into one of 12 groups defined by the conjunction table of their designated (D), antidesignated (A), and neutral (N)

subsets.[2] (A deeper and more thorough algebraic approach to the theory of many-valued logics, including fuzzy, intuitionistic, and probabilistic inferencing is given by Bolc.[3]) For example, Shaw shows that the table for element conjunctions from these fuzzy logic subsets is

δ	A	N	D
A	A	A	A
N	A	N	N
D	A	N	D

This table is also valid for Lukasiewicz's n-valued logic, which, as one would expect, shares many analogies with fuzzy logic. The Lukasiewicz's logics are first defined in terms of negation and implication, and other operations are defined in terms of these two. The table of conjunctions above is also valid for Kleene's (strong) three-valued logic based on negation, conjunction, and disjunction operators. Kleene's logic has no true-assuming tautologies (formulas that always assume the highest truth value if more than one designated value is available). If only the operators of Definition 1 are used, and different independent operators are defined as part of fuzzy logic, then Elkan's point that only DeMorgan-like tautologies are possible in fuzzy logic is well taken, but of no great consequence as long as viable deductive laws are available. (A discussion of what those deductive laws could be and their relation to implication is treated nicely by Trillas, who characterizes a generic "modus ponens generating function."[1]) As these references point out, there is not only practical but theoretical credibility to the inferences proposed for fuzzy logic well beyond the limitation to DeMorgan equivalences suggested by Elkan. Elkan's acceptance of the so-called collapse as an established fact in his conclusions could be considered disingenuous by the finality with which he considers the hypotheses of Theorem 1 to be "apparently reasonable conditions."

The watermelon problem. Elkan's version of this example is a revision of the 1993 AAAI conference publication, where the watermelon model was in error because it was incomplete. The problem I find here is not directly related to the logic, but rather to the use and interpretation of the model. Elkan has not expressed what he calls "implicit" background knowledge in terms of rules. While he uses the equivalence operator to define *watermelon*(x), a knowledge engineer trying to identify watermelons might have given two rules (with different logical meaning) from (*insidered*(x) $\rightarrow$ *watermelon*(x)) AND (*outsidegreen*(x) $\rightarrow$ *watermelon*(x)) to indicate that the two predicates contribute separately to the implication of *watermelon*(x).

Alternatively, the engineer might have given the one rule with the conjoined antecedents. In the former case, in many expert systems shells the connotation that the conditions *insidered*(x) and *outsidegreen*(x) give "separate" results to argue the consequent from two different viewpoints would yield a heuristic function of the two valuations — somewhere between the values of *insidered*(m) and of *outsidegreen*(m). If the knowledge engineer had selected the model (*insidered*(x) AND *outsidegreen*(x)) $\rightarrow$ *watermelon*(x) for this example, it would connote the necessity to satisfy "simultaneously" both related conditions, and it can be argued that the conservative answer would be the "weakest link" answer (the minimum of the two valuations).

Fuzzy logic in control. It seems reasonable that the longer a chain of implications with uncertain predicates is — whatever the definition of the approximate deductive law — the more uncertain the result at the end of the chain will be (as in computing the range of values in worst-case designs). So it seems that it would be a good thing, in general, to have short inference chains and a small number of rules whenever possible. Furthermore, the fuzzifying and defuzzifying that takes place at times reminds me of the reshaping done in the analog transmissions of digital pulses to avoid signal deterioration through consecutive repeaters to distort information.

The fact that so many applications have been possible with short inference chains raises more interesting questions yet: Under what circumstances are long chains indispensable? How could long chains of inferences be avoided? However — make no mistake — even a set of one-layer rules requires some form of inference, and rule sets will increase their sequential complexity when hysterisis is taken into account.

Elkan repeats conventional wisdom when stating, "The basic problem is that the ways in which items of uncertain knowledge are combined must be carefully controlled to avoid incorrect inferences. Fixed, domain independent operators . . . do not work" to which I add: regardless of the logic system. We should not expect to find an exact function f such that $t(A*B) = f(t(A), t(B))$ for a logical operator * unless we know either the functional relations of occurrence between A and B or, equivalently, know that they are independent (and if that were the case, an exact analytical model could be built!) It is then not surprising that knowledge engineering and incremental learning methods are used in conjunction with parameter determination to compensate for this lack of generic knowledge, not the weakness of a logic system. So, what is new? The dogma of generality versus efficiency strikes again, and knowledge engineering and machine learning are not exempted.

Elkan's ability to generate interest in both the topic of nonclassical logics for AI and the need for more general understanding of many-valued logics and basic research on how it is applied, are important contributions that should be acknowledged. It is a good thing that the relatively smooth imprecisions of natural-language semantics — when contrasted with crisp symbolic approaches — are available without excessive complexity when simpler, closed-form, and linear designs are not forthcoming. This occurs frequently around those transitional regions of system operation where decision changes interface, and points to the value of vagueness in processing natural language — usually considered in the negative — as a useful, approximate, real-world engineering design tool, a fact not popularly noticed by researchers in natural-language processing. We can use fuzzy reasoning, as we do in everyday discourse, when more exact approaches are too complex, time-consuming, costly, or are just not available.

Acknowledgment

Thanks to Massoud Moussavi for some interesting and clarifying discussions on this topic.

References

1. E. Trillas and L. Valverde, "On Mode and Implication in Approximate Reasoning," *Approximate Reasoning in Expert Systems*, Elsevier, New York, 1985.
2. K.A. Shaw, "A Classification of Designated Logic Systems," master's thesis, George Washington Univ., Washington, D.C., 1988.
3. L. Bolc and P. Borowick, *Many-Valued Logics: Theoretical Foundations*, Springer-Verlag, New York, 1992.

Oscar N. Garcia is director of the Interactive Systems Program at the National Science Foundation, and a professor in the Department of Electrical Engineering and Computer Science at the George Washington University. His research interests are in artificial intelligence, with emphasis on speech recognition, knowledge acquisition and representation, multivalued logics to represent uncertainty, and human-computer interaction. He has also worked in computer architectures and parallel processing, testing of digital circuits, and arithmetic codes. He coauthored *Knowledge-Based Systems: Fundamentals and Tools*, published by IEEE Computer Society Press. He is a fellow of the IEEE and the AAAS, and past president of the IEEE Computer Society. He also served several terms as director of the American Federation of Information Processing Societies. He was awarded the Special Group Award for his work in the Computer Science and Engineering Model Curricula, the Richard E. Merwin Award for Distinguished Service to the IEEE Computer Society, the 1991 Professional Leadership Award from the IEEE US Activities Board, and the 1994 IEEE Richard M. Emberson award. He earned his BS and MS degrees in electrical engineering from North Carolina State University, and his PhD from the University of Maryland. Oscar Garcia can be reached on the Internet at: ogarcia@nsf.gov

RESPONSE

Elkan Goes Wrong — Again

George J. Klir and Bo Yuan, State University of New York, Binghamton

Elkan's article has three basic parts: a mathematical part consisting of one definition and one theorem; a discussion of the roles played by fuzzy logic in expert systems and control systems (based upon the mathematical part); and his appraisal of the roles of fuzzy logic and its likely significance in the future. Here we discuss the major fallacies we found in the first two; due to space limitations, we will not address the third, though we disagree with almost all of the author's opinions.

As is well known, Elkan's article is a revised version of his original paper, published last year. These two versions are not fully compatible, especially in the mathematical part. Here, we point out discrepancies between the two versions and address both alternatives.

In Definition 1, Elkan introduces a particular system of fuzzy logic by choosing the standard fuzzy operators for conjunction, disjunction, and negation, and by requiring that "$t(A)=t(B)$ if A and B are logically equivalent," where $t(A)$ and $t(B)$ are, respectively, the degrees of truth of arbitrary propositions A and B. Clearly, $t(A)$ and $t(B)$ are values in [0,1]. In the original paper, the term "logically equivalent" is defined as "equivalent according to the rules of classical two-valued propositional calculus." This is, of course, nonsense, since one logic system (in our case, a particular system of fuzzy logic) cannot be defined in terms of logical equivalence of another system (the more restrictive classical two-valued logic).

In the revised version, the meaning of the term in Definition 1 is not explicated. It is only remarked that "depending on how the phrase 'logically equivalent' is understood, Definition 1 yields different formal systems." Since the role of Definition 1 is to characterize a system of fuzzy logic, logical equivalence in this definition must be expressed in terms of all possible truth values of fuzzy propositions, that is, in terms of all real numbers in [0,1]. Specifically, two expressions in fuzzy logic based on the operators of Definition 1 are logically equivalent if and only if their truth values are equal for all possible assignments of truth values in [0,1] to logic variables involved.

The principal result (and the only mathematical result) of Elkan's papers, which purport to demonstrate "technical limitations of fuzzy logic," is Theorem 1. What is this result? The answer depends on which version of the paper you use. In the original version:

> For any two assertions A and B, either $t(B)=t(A)$ or $t(B)=1-t(A)$.

The theorem is supposed to apply to the system of fuzzy logic introduced by Definition 1. However, as explained above, the definition is based on the logical equivalence of two-valued logic and hence it is not a definition of a fuzzy logic system. The proof of the theorem is based on the fact that the sentences $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ (and seven other pairs of sentences obtained by exchanging and complementing A and B) are logically equivalent in classical two-valued logic. However, these sentences are not equivalent in a fuzzy logic that employs the logic operators of Definition 1. Hence the theorem has no relevance to this fuzzy logic. Let us turn now to the revised version of the theorem:

> Given the formal system of Definition 1, if $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are logically equivalent, then for any two assertions A and B, either $t(B)=t(A)$ or $t(B)=1-t(A)$.

The fundamental difference between the original and revised version of the theorem reflects the difference in the two versions of Definition. 1. In the revised version, the logical equivalence of $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ is employed as a condition in stating the theorem rather than a fact in proving it. If the notion of logical equivalence in the revised Definition 1 is understood as applying to all truth values in [0,1], in spite of its confusing characterization by the author (as discussed above), then the revised version of Theorem 1 is relevant to the fuzzy logic involved. However, the theorem is still stated incorrectly or, alternatively, its proof is incorrect. The proof depends on eight logical equivalencies, only one of which is included in the statement. The last paragraph of the proof is thus mathematically incorrect. It would be correct if logical equivalencies representing the seven implications listed in the paragraph were included as conditions in the statement of the theorem. Without these seven logical equivalencies as conditions, the theorem must be reformulated as follows:

> Given the formal system of Definition 1, for any two assertions A and B, if $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are logically equivalent, then the truth values $t(A)$ and $t(B)$ are constrained by the inequalities $t(A)+t(B)\geq 1$ or $t(B)\leq t(A)$.

In this case, the last paragraph of Elkan's proof is incorrect and must be excluded.

Assume that the statement of Theorem 1 and its proof are made compatible in one of the two ways we suggest. What then is the meaning of the resulting theorems — one with the single logical equivalence as a condition, and one with the eight logical equivalencies as conditions? These theorems basically show that the truth values of propositions within the system of fuzzy logic introduced by Definition 1 become appropriately constrained when additional extraneous conditions are imposed. With the eight conditions, the constraint is obviously more severe than with only one of them. If, for example, we required our system to satisfy $A \vee \neg A = 1$ then the truth values would become constrained to the set $\{0,1\}$, and the system would collapse to the classical two-valued logic.

All this is well known, and Elkan's theorem (when properly fixed) does not offer anything new. It is absurd, however, to constrain a system by extraneous requirements and then claim that the original system has "technical limitations." This is what Elkan attempts to do in his papers. The fact that every system of fuzzy logic must violate, under the assumption of truth functionality, some properties of Boolean algebra (and, hence, the classical two-valued logic) is a simple consequence of the decision to formulate logics that can deal with propositions that are not required to be either true or false, but may be true or false to various degrees.[1]

Elkan's remarks about the connection between fuzzy logic and intuitionistic logic also contain some errors. For example, it is not sufficient to characterize fuzzy logic by the rejection of the law of excluded middle. The system of fuzzy logic determines which properties of Boolean algebra are rejected. The system introduced by Definition 1, for example, rejects not only the law of excluded middle, but the law of contradiction as well.[2] This differs from intuitionistic logic, which rejects the law of excluded middle and the implication $\neg\neg A \rightarrow A$, but does not reject the law of contradiction and the opposite implication $A \rightarrow \neg\neg A$.[3,4] Other systems of fuzzy logic do not reject any of the mentioned laws; instead, they reject distributivity and idempotence.[2] Furthermore, de Morgan's laws are valid only in some systems of fuzzy logic. Another error is to consider the logical equivalence in Definition 1 as intuitionistic equivalence. A distinctive feature of intuitionistic logic is the operator of negation upon which it is based. For any proposition A, where $t(A) \in [0,1]$, the intuitionistic negation, $\neg A$, is defined by

$$t(\neg A) = \begin{cases} 1 \text{ when } t(A) = 0 \\ 0 \text{ otherwise} \end{cases}$$

This negation is not involutive, nor is it continuous — it acts as a defuzzifier. Clearly, there is no compatibility between intuitionistic logic and the fuzzy logic introduced in Definition 1.

Fuzzy logic applications. Elkan's discussion of fuzzy logic in expert systems reveals his confusion between degrees of truth in fuzzy logic and degrees of evidence expressed in terms of some fuzzy measures (probabilities, belief measures, and so on).[5] While the former are a matter of compatibilities of given objects with relevant fuzzy predicates, the latter result from information deficiency regarding the classification of a given (incompletely characterized) object in relevant crisp sets. While these two areas have distinct application domains, they can be combined, resulting in statistics with imprecise probabilities[6] or in fuzzified evidence theory,[7] for example.

In his discussion of fuzzy controllers, Elkan's lack of understanding is again revealed. He fails to understand that most of the simple fuzzy controllers on the market (we may call them the first generation of fuzzy controllers) are not explicitly based on fuzzy logic, but rather on the approximation of relevant control functions by fuzzy numbers that represent chosen linguistic states of the variables involved. This is similar to classical control, which is also not explicitly based on classical two-valued logic. It is well established that fuzzy controllers of this kind are universal approximators.[8,9]

While most existing fuzzy controllers are rule based, research on combining rule- and model-based approaches in designing fuzzy controllers is ongoing. Models employed in these controllers are expressed, in general, in terms of relations among relevant fuzzy variables. Hence, the use of fuzzy set theory (not necessarily fuzzy logic in the narrow sense) involves both parts of the controller — the rule-based part as well as the model-based part.

Elkan's papers do not contribute to knowledge. The mathematical part is fallacious; and, while some critical errors in the original version are corrected in the revised version, new errors are introduced and some statements become less specific. Even if we fix all the mathematical errors to help Elkan obtain his intended result, we find only that the result is trivial and well known: If one takes an axiomatic system and adds to it additional requirements, the system becomes more constrained. Given a free choice of requirements, one can constrain the system as he or she wishes. This is precisely what Elkan attempts, in an amateurish way. He tries to find requirements that would constrain a given system of fuzzy logic so severely that only two truth values are allowed. He then argues that this shows technical limitations of fuzzy logic. This sort of argumentation is absurd.

References

1. D. Dubois and H. Prade, "An Introduction to Possibilistic and Fuzzy Logics," *Non-Standard Logics for Automated Reasoning,* Academic Press, New York, 1988, pp. 287-315.
2. G.J. Klir and T.A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, Englewood Cliffs, N.J., 1988.
3. N. Rescher, *Many-Valued Logic,* McGraw-Hill, New York, 1969.
4. A.A. Zinov'ev, *Philosophical Problems of Many-Valued Logic,* D. Reidel, Dordrecht, Holland, 1963.
5. Z. Wang and G.J. Klir, *Fuzzy Measure Theory,* Plenum Press, New York, 1992.
6. P. Walley, *Statistical Reasoning with Imprecise Probabilities,* Chapman and Hall, New York, 1991.
7. J. Yen, "Generalizing the Dempster-Shafer Theory to Fuzzy Sets," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 20, No. 3, May-June 1990, pp. 559-570.
8. J.J. Buckley and Y. Hayashi, "Fuzzy Input-Output Controllers are Universal Approximators," *Fuzzy Sets and Systems*, Vol. 58, No. 3, Sept. 1993, pp. 273-278.
9. B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Approach to Machine Intelligence,* Prentice Hall, Englewood, N.J., 1991.

George J. Klir is distinguished professor of systems science in the Watson School of Engineering and Applied Science at the State University of New York, Binghamton. He is author or coauthor of 13 books, including *Fuzzy Sets, Uncertainty, and Information* (Prentice Hall, 1988) and *Fuzzy Measure Theory* (Plenum Press, 1992), and is the editor of seven books. He has also written more than 300 research papers, and has been the editor of the *Int'l Journal of General Systems* since 1974, and editor of the *Int'l Book Series on Systems Science and Engineering* since 1985. He is past president of the International Federation for Systems Research, and the North American Fuzzy Information Processing Society. He is currently president of the International Fuzzy Systems Association. He was recently awarded an honorary doctoral degree from the University of Economics in Prague, and the Gold Medal of Bernard Bolzano in mathematical sciences from the Czech Academy of Sciences. Klir can be reached at the Department of Systems Science and Industrial Engineering, SUNY, Binghamton, N.Y. 13902-6000.

Bo Yuan is a PhD candidate in systems science at the Watson School of Engineering and Applied Science at the State University of New York, Binghamton. His research interests are in fuzzy set theory, fuzzy logic, approximate reasoning, fuzzy control, neural networks, genetic algorithms, and their applications in industrial engineering. He received his BS and MS degrees from Shanghai Teachers University in 1985, and 1988, respectively. Yuan can be reached on the Internet at ba05074@bingvmb

Fuzzy Control

A Misconception of Theory and Application

E.H. Mamdani, Queen Mary & Westfield College, London

The argument in Charles Elkan's article has three steps. First, he provides a theorem that "proves" that fuzzy logic is deficient because it collapses to a two-valued logic. He then shows what makes the current applications of fuzzy logic successful, although this success may seem paradoxical. His final step shows how such a success cannot be guaranteed as applications scale up in the future — thus resolving the paradox. I expect other commentaries will deal with the misconceptions regarding the theorem. I focus attention here on the remainder of Elkan's argument.

The source of Elkan's paradox is the link between the first two steps formed by his statement that "One way to defend a calculus is to show that it succeeds in interesting applications." But first a couple of very different but relevant pointers.

Fuzzy logic control is successful because it replaces the classical PID controller. When tuned, the parameters of a PID controller affect the shape of the entire control surface. Because fuzzy logic control is a rule-based controller, the shape of the control surface can be individually manipulated for the different regions of the state space, thus limiting possible effects to neighboring regions only. Furthermore, the use of fuzzy mathematics provides interpolation between the adjoining regions, resulting in an overall smooth control surface — an important requirement in the control of continuous systems. This also suggests that fuzzy sets are an efficient way of representing continuous variables in rule-based systems.

Secondly, I have always felt that fuzzy logic has similarities with Boole's logic. That logic, originating over 150 years ago, was the first system of reasoning in the form of a calculus. However, after reading Boole's "Laws of Thought" it is difficult to discern whether Boole is concerned with a descriptive explanation of how people actually think, or with a prescriptive model of how they ought to think. AI research workers have seldom addressed this key distinction properly.

Within AI there are three distinct areas of research: the descriptive, the prescriptive, and what I call the applicative concerns. In the first area, researchers deal with descriptive theories about cognitive processes. These theories are very hard to prove experimentally (or more specifically to disprove experimentally — if one applies the Popperian view) because the level of control in experimental studies on human cognition is far below that in the natural sciences. The second group of researchers are concerned with prescriptive models: different reasoning systems and a variety of logics. Here, the issue is one of correctness of these models, variously defined. Again, it is not possible to use natural-science methods to devise controlled experiments that demonstrate the correctness of these models; correctness can only be dealt with by means of philosophical arguments (more on this later).

I belong to the third group of AI researchers, whose main concern is to build industrially successful artifacts. Such artifacts are successful in their own right, and do not owe their success to the underlying theory or a mathematical model. It is sad how many AI workers have lost the ability to distinguish between applications and well-designed controlled experiments set up to disprove a particular theory. Applications address the scientific needs of a specific domain, and cannot replace experiments conducted to test a theory. Many features of the domain knowledge introduced in an application also contribute to its success. There is a common misconception that models are created and then applied, and that success then legitimizes a model. This view is superficial, because an application's requirements seldom match the underlying axioms of the model exactly. The fixes that are added (defuzzification in fuzzy logic control) are instrumental in the industrial success — but often sit uncomfortably in the original theory. This is true of all applications inspired by prescriptive models.

The links between these three groups (descriptive, prescriptive and applicative) must be properly understood if one is to avoid the methodological trap Elkan has fallen into. In AI, the work of each group inspires the direction of the others — but that is all they do. The results of one group can never be used to legitimize the approach of another. Weak though these links are, they still play a significant role in scientific advances. My point is not to belittle the interplay between the three areas, but to point out that a misunderstanding of their relationship is clearly the source of Elkan's perception of the paradox.

What then is the relationship between fuzzy logic control applications and fuzzy logic itself? Precisely the same as that between Boole's laws of thought (a descriptive theory?), Boolean logic (a prescriptive model?), and logic circuits (an application) — namely, an effective tool presented itself that met many, though surely not all, of the application needs. However, the widespread success of logic circuits cannot be used to legitimize Boole's logic any more than the industrial success of fuzzy logic control legitimizes the philosophical correctness of fuzzy logic. Therefore, the question of a paradox — a central idea in Elkan's paper

— does not arise. Similarly, his argument on the philosophical deficiencies of fuzzy logic focuses on a theorem without fully discussing the assumptions and axioms it is based upon; this does nothing to argue against the adoption of fuzzy logic control. The terms "logic" in logic circuits and "fuzzy logic" in fuzzy logic control are purely incidental, and a matter of historical evolution.

The AI approach puts a much higher value on prescriptive mathematical models than they actually deserve. These models cannot be legitimized by controlled experiments or by application, nor can they be justified by some underlying descriptive theory (in spite of Boole). Prescriptive models can only be argued over at a philosophical level — an ability few AI researchers possess. Philosophical disputations about prescriptive models within informed groups such as Uncertainty in AI, have, nevertheless, helped to enlighten many difficult points. In the end, however, such disputations can never completely settle the matter.

Because AI researchers are mostly trained in mathematical skills, another frequently applied but false way of legitimizing prescriptive models is on the grounds of mathematical symmetries or some intrinsic sophistication of potential function. On rare occasions when models are abstracted from applications, the concern is no longer what led to the success of the application. Rather, the academic game of looking for the symmetries and the sophistication of the form or the soundness of the calculus begins.

Having rightly or wrongly detected a paradox, one then has to resolve it; in doing so, Elkan commits further errors. He has a lot to say about the small number of rules, the shallowness of fuzzy rule bases, and so on — implying that some beauty of the form often plays a significant role in assessing the worth of a model (and the intellectual enterprise of a researcher) rather than the content or industrial usefulness. To argue that fuzzy logic control is not worthy of industrial consideration because of its lack of complex form and structural sophistication, as Elkan effectively does in the final part of his paper, is to subscribe to an anti-inventions culture. Accentuating form without attention to the content is like praising beauty and ignoring the brain. To use the colloquial term, the scientific mythology within AI has created a "bimbo science."

The scenario worth keeping in mind is that since its inception, fuzzy logic has had its detractors and antagonists not least because the tag "fuzzy" is seen as debasing to the somber image of science. So incensed are some that they will clutch at any straw to rid us of fuzzy sets research, even through a paper based on mistaken interpretations and modish posturing. This scenario leaves me saddened, for reasons explained above.

It is the word "paradox" I find most baffling in Elkan's article. Science at its best is often counterintuitive; but paradoxical? Our accepted understanding of the scientific method is based on natural science and descriptive theories. But applying descriptive theories to computer science — which is dominated by prescriptive theories — cannot, in my opinion, work. New prescriptive theories often alienate many researchers, but they also inspire others to build novel applications. It may be that some of these applications are a runaway success. Rather than talking of "paradoxes," what is required at this point is a rigorous attempt to discover the secret of that success. Because this investigation is descriptive in nature, the traditional scientific method is likely to yield dividends. In the case of fuzzy control, this process is now underway.

E.H. "Abe" Mamdani is a professor in the electronic engineering department of Queen Mary & Westfield College, University of London. His research in fuzzy logic control began in the early 1970's, and he has acted as a consultant to several European industrial companies on applications of knowledge-based systems. His current research interests include reasoning under uncertainty, agent-based systems, and soft computing. He is a fellow of the IEEE and the Royal Academy of Engineers. Abe Mamdani can be reached through the Internet at e.h.mamdani@qmw.ac.uk

Fuzzy Logic

A Misplaced Appeal

Francis Jeffry Pelletier, University of Alberta

I have long found puzzling the acceptance and apparent success of fuzzy logic. We philosophically oriented logicians have pretty much sneered at fuzzy logic ever since it was introduced with that name.* Yet what can I say when I own an excellent fuzzy logic camera? I am grateful to Elkan for his explanation of this point of tension.

Basically, Elkan explains that the notion of "fuzzy logic" as it is used in control systems has nothing to do with the term as it is used in logic. That is, it has nothing to do with fuzzy logic as a formal system with rules of formation, evaluation, and inference. Fuzzy controllers are so-called because of a certain *analogy* with fuzzy logic, but in fact they do not embody, implement, or instantiate fuzzy logic.

For Elkan, the relationship between fuzzy controllers and fuzzy logic is rather like that between on-off light switches and predicate logic: Yes, there is a certain analogy between on-off and true-false, but it's *only* an analogy, a way of looking at light switches. There is nothing in the light switch corresponding to the connectives of sentence logic nor to predicates, names, and quantifiers of predicate logic. To identify the two, or to say that light switches implement or instantiate predicate logic, would be to ignore most of predicate logic and mistakenly fixate on just one insignificant aspect. According to Elkan, we should not be surprised that critiques of fuzzy logic have no impact on fuzzy control theory; the areas of fuzzy logic that get criticized are simply not employed in the control arena (whether practical or theoretical).

Elkan's theorem shows one of the difficulties surrounding fuzzy logic as a formal system. Supporters of fuzzy logic are without doubt tempted to respond to this by focusing on the assumption that logical equivalence in classical (or intuitionistic) logic is a warrant for formulas having the same truth value in fuzzy logic. I do not wish to enter this debate; instead, I will take this opportunity to point to some other features of a logical nature that have been used to criticize fuzzy logic and its claims of usefulness in various tasks.

Presentations of fuzzy logics have generally been semantic in nature, while the syntax — axioms and rules of inference — has generally been ignored. The basic semantic notion is that propositions can take any real value in [0...1] intuitively corresponding to "degrees of truth" of the proposition. Many advocates of fuzzy logic, especially those who want to replace classical logic as the medium of representation for ordinary reasoning and the description of natural-language phenomena, would like to "use" the semantics of fuzzy logic. That is, they are not interested merely in asserting theorems, nor in the uninterpreted formulas of fuzzy logic, but rather would like to be able to claim that a proposition is true to a certain degree, that it can be compared to another proposition which is true to some different degree, and that certain conclusions can be drawn from this comparison.

For example, it might be that "Sally is wealthy" is true to degree 0.7 while "Mike is wealthy" is true to degree 0.4. Now, we might wish to draw certain conclusions from this information, such as that Sally is wealthier than Mike, or wealthier to a certain degree than Mike. To do this, we need some way to "use the semantics." Technically speaking, we wish to have a kind of "autodescriptivity" in the logic: a way of mirroring the semantics within the syntax. This autodescriptivity is regarded by some authors as necessary for the adequacy of *any* many-valued logic,[2] for without it, the apparent many-valuedness is only illusory because we cannot *say* anything in a many-valued way. There are a number of ways of accomplishing this, depending on what sorts of operators are available within the language. The direct way is to have so-called parametric operators in the language: For each k, where $0 \le k \le 1$, there is a unary sentence operator J_k. The truth of such sentences is evaluated thus:

$$t(J_k[\Phi]) \begin{cases} = 1, \text{ if } t(\Phi) = k \\ = 0, \text{ otherwise.} \end{cases}$$

That is, a J_k operator says that the formula it operates on takes exactly the value k. Although there are other approaches, I will adopt this direct approach — that the language being used to "express the semantics" contains the parametric operators directly.

(There are many other ways to get their effect. Some writers allow constants — symbols that denote the truth values — others have "threshold operators," and so on. With suitable such other operators, we can indirectly define the parametric operators. Since there are innumerable truth values in the real range [0...1], the methods of autodescriptivity mentioned here cannot really be applied. Instead, we must consider the fuzzy logic generated by the *rationals* in the [0...1] interval. Attempts by fuzzy logicians to incorporate ever more inclusive — some would say obscure — operators indicates to me a lack of appreciation of what a logic is. For, if it can be shown that there is no algorithmic, deterministic procedure to determine the truth value of an arbitrary expression, then it is very unclear that there can be any use of the formalism.)

* It was studied by J. Lukasiewicz and A. Tarski in 1930 under the name "infinitely many-valued logics,"[1] and received intensive study by several mathematical logicians in the 1950s and early 1960s.[2]

Observation 1: Fuzzy propositional logic is not argument-complete. The first shortcoming of fuzzy logics concerns propositional logic (and hence any fuzzy logic, because they all contain propositional logic as a part). There is no theory of argumentation for fuzzy propositional logic such that whenever all premises of the argument are designated, then so is the conclusion. (Intuitively, some of the truth values are considered "good" or designated, while the others are undesignated. Exactly which ones are designated might vary from application to application. The point is that an argumentation theory is designed to take us from "good" premises to "good" conclusions, and never mislead us by deriving a "bad" conclusion from "good" premises. The present observation says that this cannot be done, ever; and this holds for any decision on what is designated, so long as at least one number is designated and at least one is undesignated.)

This result does not depend on there being (or lacking) any particular syntactic machinery around (other than the parametric operators); rather, there simply can be no such theory of argumentation. The proof of this is via the fact that fuzzy logics are not semantically compact. That is, it is *not* true for fuzzy logic that a set of formulas is satisfiable just in case every finite subset of it is satisfiable. For example the infinite set

$$\Gamma = \{ \neg J_k[p] \mid 0 \le k \le 1 \}$$

is not satisfiable, since the sentence letter p must take on one of the values $0 \le k \le 1$, whereas the membership condition in Γ says it doesn't. Yet any finite subset of Γ is satisfiable. Similar sets can be described using quantified sentences, such as

$$\Gamma' = \{ J_1[Fa_1], J_{1/2}[Fa_2], J_{1/4}[Fa_3], \ldots, \neg J_0[\forall x Fx] \}$$

Having noted this fact, it is an easy step to the conclusion that fuzzy logic, even fuzzy propositional logic, is not argument-sound, since all proofs are finite. Thus, there can be no adequate scheme for making inferences in general within fuzzy logic.

Observation 2: There is no normal form for fuzzy monadic predicate logic in which quantifiers have widest scope. The second shortcoming of fuzzy logic is found in the attempt to add quantifiers, even a simple monadic predicate logic. Fuzzy logic dictates that a universally quantified formula, such as $\forall x Fx$, takes the least value of all the substitution instances for x in the formula Fx, or the greatest lower bound if there is no such least value. An existentially quantified formula takes the greatest value of all the substitution instances, or the least upper bound if there is no such greatest value.

A sentence like $J_k[\forall x Fx]$ says that the greatest lower bound of the Fx's values is exactly k. There is no formula that has any quantifier outside the scope of J_k that has the same truth value. For example, $\forall x J_k[Fx]$ says that every individual instance of Fx has a greatest lower bound of exactly k, which is clearly wrong. $\exists x J_k[Fx]$ says that there is some particular individual that is F to exactly degree k, which is also wrong because there might not exist an object that has the greatest lower bound value. But the lack of a normal form makes it unlikely that there can be any method to detect theoremhood in fuzzy monadic predicate logic. Certainly resolution will not work.

So, not only can we not tell when a conclusion is validly derived in fuzzy logic (Observation 1), but we cannot even tell when a formula (even of monadic predicate logic) is a theorem. Surely together these two observations should give fuzzy control theorists pause; they show that fuzzy logic as an abstract theory reduces to stating intuitive principles without any way to generalize or use them. And, since fuzzy control theory is surely committed to using *something* — it follows that what it is committed to using is *not* fuzzy logic, just as Elkan said.

Observation 3: Full fuzzy predicate logic is not recursively axiomatizable. The real underlying reason that fuzzy logic fails to be of any logical interest does not have to do with the elementary fragments of propositional fuzzy logic and monadic fuzzy predicate logic, even though it is cute to note that even these elementary parts of fuzzy logic are not usable in the desired form. Instead, it is that full predicate logic is not really a logic.

This result was proved by Scarpellini[3] for infinite-valued Lukasiewicz logics, and the proof carries over to all the well-known modifications (such as adding parametric operators or various arithmetic operators) of this logic, which includes any of the fuzzy predicate logics ever described in the literature. The thrust of the proof is that the set of unprovable formulas of ordinary two-valued predicate logic can be mapped one-to-one into the set of valid (designated) formulas of fuzzy logic, for any closed or open range of values $(k\ldots1)$ that we designate. But the set of unprovable formulas of ordinary predicate logic is not recursively axiomatizable, and therefore neither is the set of valid formulas of fuzzy logic. Hence, they cannot even be adequately characterized or talked about coherently, except by example. Furthermore, fuzzy control theorists do not merely wish to appeal to examples of valid formulas of fuzzy logic, but to be able to characterize them in some way or other.

Lest my message be thought entirely critical of fuzzy control theory, let me point out that I believe that everything its proponents wish to do can be adequately carried out. (My camera works!) However, their appeal to fuzzy logic is misplaced. Every fuzzy logic application has an analogue in finitely many-valued logic, and each one of these is logically well-behaved. There are correct theories of argumentation for them, there are resolution-like theories of theorem-detection for them, and they are axiomatizable.

The only apparent advantage to fuzzy logic is that it seems to be a grand generalization of all those finitely many-valued logics — after all, we never know in advance which particular finite value might be needed for a specific application. However, it is an illusion to think that fuzzy logic is the correct generalization. It cannot be used and it has no reasonable logical foundation. But a "variable precision" finitely many-valued logic can do the sort of things desired. In such logics we have a superstructure of (say) three values. Having determined that some sentence takes one of these top-level values, we can then expand it to determine which of those three values it has at a lower level. For example, we grade essays as "good," "so-so," and "bad," but then given that an essay has been categorized as "so-so" we can look more closely at whether it is a good, so-so, or bad example of being so-so. And this process can continue for some finite number of times. Such a logic does not have any of the shortcomings that fuzzy logic does, and would seem to be the sort of thing that could form a logically adequate background theory for fuzzy control systems.

References

1. J. Lukasiewicz and A. Tarski, "Untersuchen über den Aussagenkalkül," 1930; translated as "Investigations into the Sentential Calculus" in Tarski's *Logic, Semantics, Metamathematics*, Oxford Univ. Press, 1956, pp. 38-59.

2. N. Rescher, *Many-Valued Logics*, McGraw Hill, New York, 1969.

3. B. Scarpellini, "Die Nicht-Axiomatieierbarkeit des Unendlichwertigen Prädikatenkalkuls von Lukasiewicz," *J. Symbolic Logic* Vol. 27, No. 2, June 1962, pp. 138-153.

Jeff Pelletier is a professor of philosophy and computing science at the University of Alberta. He has written more than fifty articles on philosophy of language and logic, computational linguistics and automated theorem proving, and ancient Greek philosophy. His is author of *Mass Terms: Some Philosophical Issues* (Reidel, 1979), *Parmenides, Plato, and the Semantics of Not-Being* (U. Chicago Press, 1990), and *The Generic Book* (U. Chicago Press, 1994), and is currently writing a book about automated theorem proving in non-classical logics. He is also interested in cognitive science verification of AI constructs. Pelletier was the Luce Professor of Cognitive Science at the University of Rochester and is a past president of the Society for Exact Philosophy. He received his BS in mathematics and education, and his MA in philosophy from the University of Nebraska in 1965 and 1966, and his PhD in philosophy from UCLA in 1971. He also received MS degrees in linguistics (1978) and computing science (1982) from the University of Alberta. Jeff Pelletier can be contacted at the Department. of Philosophy, University of Alberta, Edmonton, Alberta, Canada T6G 2E5; Internet: jeffp@cs.ualberta.ca

RESPONSE

On the Purportedly Paradoxical Nature of Fuzzy Logic

Enrique H. Ruspini, SRI International

Elkan's original paper purportedly showed that fuzzy logic was paradoxical in nature due to its reliance on formal bases that preclude truth values other than 0 and 1. Elkan has now modified some of his claims and arguments slightly, although he still depends on that result as his major source of insight into the technology.

We are now told, for example, that fuzzy logic is paradoxical because it is successful in many applications while its foundations remain under attack. Perplexing as this situation might be from a sociological viewpoint, it is hardly a logical self-contradiction, and describing it as a paradox is totally inappropriate.

Nor is a paradox implied by the claim that most theoretical fuzzy logic papers seem to deal with representation and reasoning methods, while most fuzzy logic applications have resulted in embedded controllers. The embedded controllers have been developed, of course, upon foundations provided by the representation and inferential methods of fuzzy logic. Elkan is not only unaware of this fact, but his overall analysis of the technology is colored by the strange notion that the depth and quality of deductive procedures in a controller are inferior to those in "sophisticated" reasoning systems.

For reasons of space, I will not discuss here Elkan's statements about application of fuzzy logic to control and other intelligent reasoning systems, but will confine my comments to the formal result (Theorem 1) that remains the major basis of his claims about the purported paradoxical nature of fuzzy logic. Other assertions about the methodology — arising in some cases from superficial analyses of relevant literature and issues, but mostly out of ignorance or plain confusion — are appropriately addressed by other respondents.

Starting from an axiomatic characterization of fuzzy logic proposed by Gaines, and assuming that logical equivalence in fuzzy logic means equivalence in the sense of classical logic (thus implying that all classical logic theorems are also fuzzy logic theorems), Elkan shows that Gaines' axioms imply that the only possible truth values are 0 and 1: fuzzy logic collapses into conventional logic.

Anybody acquainted with fuzzy logic, however, would not have much difficulty questioning Elkan's notion of logical equivalence; it is well known that many theorems of propositional logic are not valid in fuzzy logic. Assuming otherwise immediately leads to the result that Elkan finds so paradoxical. Applying, for example, the axioms of fuzzy logic to the law of the excluded middle $\alpha \vee \neg\alpha$, which is not a theorem of fuzzy logic, leads to the equation

$$\max(t(\alpha), 1 - t(\alpha)) = 1,$$

which only has the solutions $t(\alpha) = 0$ and $t(\alpha) = 1$.

Many theorems of classical propositional logic may also be used to derive this result. Elkan's unnecessarily lengthy proof — based on the conventional propositional logic equivalence of the formulas $\neg(a \wedge \neg b)$ and $b \vee (\neg a \wedge \neg b)$ — actually assumes the validity of the law of the excluded middle (to see this, simply expand the latter and note the conjunct $b \vee \neg b$).

Elkan's "shocking" discovery has been long known, and is discussed in elementary textbooks on fuzzy and multivalued logics.[1] For example, if (C,U,I) are negation, disjunction, and conjunction operators, respectively, that is —

$$\begin{aligned} t(\neg p) &= C(t(p)), \\ t(p \vee q) &= U(t(p), t(q)), \\ t(p \wedge q) &= I(t(p), t(q)) \end{aligned}$$

that satisfy the laws of excluded middle and contradiction, then the corresponding logics can be neither idempotent nor distributive. If Elkan had probed further, he could have proved that all continuous truth-functional multivalued logics "collapse" as well.

The definition of equivalence that Elkan describes as "apparently reasonable" is, therefore, patently unreasonable. The supposedly shocking result is just a well-known fact of little relevance to the practice of fuzzy logic. Simply stated, Elkan has found that fuzzy logic and the classical propositional calculus are different logical systems.

Not much is gained either by looking into seemingly more congenial quarters for alternative definitions. Elkan turns, for example, to intuitionistic propositional calculus (IPC) as another place to borrow notions of equivalence, feeling that his result is strengthened by the fact that the law of the excluded middle — a previous source of trouble — fails for both IPC and fuzzy logic. IPC is, however, based on a negation operator with different semantics than that of fuzzy logic (one is involutive while the other is not). Once again, one does not need a proof as extensive as Elkan's. The Gödel translation $\neg(\neg\alpha \wedge \neg\neg\alpha)$ of the law of the excluded middle is a theorem in IPC but not in fuzzy logic. Assuming that it is leads once again to the same incorrect conclusion: Fuzzy logic collapses. Elkan's theorem is, therefore, just as true for IPC equivalence as it was for classical equivalence, but it is also just as meaningless as before: All that has been proven is that fuzzy logic is neither classical nor intuitionistic logic.

This explanation, however, still does not answer a basic question: What is the meaning of the word *equivalent* in Gaines' Axiom 4:

$t(a) = t(b)$ if a and b are logically equivalent?

In classical logic, logical equivalence between two formulas α and β may be defined either as the validity of the formula $\alpha \leftrightarrow \beta$, or as the equality of the truth values of the formulas α and β for all possible assignments of truth values to their constituent propositional symbols. A quick inspection of the truth table of the $\leftrightarrow$ connective shows that these definitions are equivalent.

While this is very reasonable, something seems to be amiss here. How can we consider Axiom 4 before we even define logical equivalence? If equivalence means that the truth value of α is always equal to that of β, why do we need an axiom to state that this should be the case?

In multivalued logics, equivalence in the sense of the validity of $\alpha \leftrightarrow \beta$ is not the same as equivalence in the sense of equality of the truth values of α and β. For example, these notions yield the same relation in the Lukasiewicz L_3 logic, but not in the 3-valued logic of Bochvar (where, if α and β have the third value 1/2, then $\alpha \leftrightarrow \beta$ also has the third value 1/2). In these logics it is possible to consider several characterizations of the notion of logical equivalence, each having different formal properties.[2]

In multivalued logics in general, and fuzzy logic in particular, equivalence is usually defined in terms of the semantics of the $\rightarrow$ connective. Several such definitions have been proposed, notably by Zadeh, and by Trillas and Valverde.[3] Seeking a wide characterization of fuzzy logics, Gaines chose not to specify a particular semantics for the implication operator, instead requiring only the use of a reasonable notion of equivalence compatible with equality of truth values.

Those who have read Elkan's original paper wondered at the time why he had to seek definitions in other logics rather than proceeding along the lines I have sketched here. In the present article, Elkan at last considers a definition based on the semantics of the negation, disjunction, and conjunction operators, but not on that of the implication connective (see his last paragraph in the section on paradox). He concludes, however, that this leads to an extremely weak system where the only equivalences are the De Morgan axioms.

This statement, unlike previous claims, is not only irrelevant but false and misleading. Simple application of fuzzy logic operators for disjunction, conjunction, and negation immediately shows that the following laws of propositional logic also hold in fuzzy logic: commutativity of disjunction and conjunction; associativity of disjunction and conjunction; distributivity of disjunction (conjunction) with respect to conjunction (disjunction); idempotence of disjunction and conjunction; identity with respect to $\top$ and $\perp$; absorption with respect to disjunction and conjunction; absorption by $\top$ and $\perp$; involution; and, surely enough, the De Morgan laws.

All these properties give fuzzy logic considerable strength as a reasoning formalism, but their consideration alone — in the absence of definitions for the implication connective $\rightarrow$ and for the deductive rules of fuzzy logic (such as the generalized modus ponens) — cannot be the bases of any substantive argument, either pro or con, regarding the adequacy and correctness of fuzzy logic as a deductive methodology. Curiously, Elkan does not seem to feel that there is any need to discuss these matters, interpreting the independence of his theorem from any notion of implication as a sign of its universality and strength rather than as yet another indicator of its lack of relevance.

Elkan's arguments, arising from a meaningless result and a superficial and confused evaluation of the state of the art in fuzzy logic, do not provide any substantial insights into the methodology, its advantages, or its shortcomings. Given the weakness of his arguments, one can only be astonished at his conclusion that proponents of fuzzy logic are guilty of fallacious *non-sequitur* thinking (*post hoc, ergo propter hoc*). Those who propound the technology found their claims on solid theoretical results and on thousands of examples of its successful application. All that Elkan produces, on the other hand, is an irrelevant theorem and a rather shallow and mistaken discussion of a minor segment of the literature.

References

1. G. Klir and N. Folger, *Fuzzy Sets, Uncertainty, Information*, Prentice Hall, Englewood Cliffs, N.J., 1988, pp. 52–59.

2. N. Rescher, *Many-Valued Logic*, McGraw-Hill, New York, 1969, pp. 138ff.

3. R. Lopez de Mantaras, Approximate Reasoning Models, Ellis Horwood, Chichester, England, 1990.

Enrique H. Ruspini is a senior computer scientist with the Artificial Intelligence Center of SRI International. He received his doctoral degree from the University of California, Los Angeles, and has extensive research experience in the areas of approximate and commonsense reasoning, the calculus of evidence, knowledge-based systems, knowledge acquisition, inductive reasoning, and the representation and manipulation of uncertainty. He was an early contributor to the development of fuzzy-set theory and its applications, and pioneered its introduction to automatic classification and pattern recognition in 1969. His recent research has focused on the application of fuzzy-logic techniques to the development of intelligent control and signal processing systems. He is a Fulbright Fellow and a SRI Institute Fellow, and is one of the founding members of the North American Fuzzy Information Processing Society and a recipient of that society's King Sun Fu award. Dr. Ruspini was the General Chairman of the Second IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '93) and of the 1993 IEEE International Conference on Neural Networks. He can be reached at the Artificial Intelligence Center, SRI International, 333 Ravenswood Ave., Menlo Park, Calif. 94025; Internet: ruspini@ai.sri.com

Semantic Uncertainty of the Fuzzified Laws of Logic

Burhan Türksen, University of Toronto

The confusion surrounding Charles Elkan's article is generated by a lack of clear understanding of the four levels of knowledge representation: linguistic, metalinguistic, propositional, and computational. When we attempt to convert knowledge expressed in natural language into computable knowledge, at least three significant transformations occur between these four levels.

Linguistic expressions. *Linguistic expressions* are natural language expressions, such as

"inventory is low and demand is high,"
"inventory is low or demand is high," or
"inventory is not low,"

where "inventory" and "demand" are nouns, and "low" and "high" are adjectives. In the terminology of fuzzy set theory,[1] the nouns are *linguistic variables*, the adjectives are *linguistic values*, and "and," "or," and "not" are *linguistic connectives* that generate interval-valued fuzzy sets.[1,2]

A *metalinguistic* expression is a mapping from natural language to a symbolic language. For example, the metalinguistic forms of the linguistic expressions above are: "X_1 is A AND X_2 is B," "X_1 is A OR X_2 is B," and "X_1 is NOT A," where X_1 and X_2 are the metalinguistic representations of the linguistic variables, A and B are the metalinguistic representations of the linguistic values, and AND, OR, and NOT are the metalinguistic representations of the linguistic connectives. In short form, these metalinguistic expressions are represented as "A AND B," "A OR B," and "NOT A."

Propositional expressions. In the classical two-valued logic, there are at least two approaches that generate *propositional expressions* (normal forms) for every metalinguistic expression. The first is to assign the symbols $\cap$, $\cup$, and ^{c} to the basic metalinguistic connectives AND, OR, and NOT, respectively. Next, we form the canonical expressions of the basic metalinguistic expressions as $A\cap B$, $A\cup B$, and A^c. Then we derive all other propositional expressions with an application of $A\cap B$, $A\cup B$, or A^c, subject to the particular interpretations.

In the second approach, we first give an interpretation to a metalinguistic expression and define its meaning with a truth table. We then determine its normal forms from the truth table by the application of the "canonical form" generation algorithm. In this approach, two distinct but equivalent canonical forms are generated: the disjunctive normal form (DNF) and the conjunctive normal form (CNF). For example, DNF and CNF for "A AND B" are

$$\text{DNF}(A \text{ AND } B) = A\cap B =$$

$$\text{CNF}(A \text{ AND } B) = (A\cup B) \cap (A\cup B^c) \cap (A^c\cup B)$$

Fuzzy normal forms. It has been shown that fuzzy normal forms can be generated from the fuzzy truth table directly.[2] Depending on the set of axioms we impose, we get at least three different classes of fuzzy logics with their corresponding normal forms:

(1) If we assume $(\cap,\cup,^c)$ is a De Morgan logic such that only boundary and monotonicity conditions together with the involutive complementation are imposed, then we have the following FDNF and FCNF expressions for the first-level fuzzy logics:

$$\text{FDNF}^{(1)}(A \text{ AND } B) = (B\cap A) \cup (A\cap B)$$

$$\text{FCNF}^{(1)}(A \text{ AND } B) = (B\cup A) \cap (B^c\cup A) \cap (B\cup A^c) \cap (A\cup B) \cap (A^c\cup B) \cap (A\cup B^c)$$

(2) If we assume $(\cap,\cup,^c)$ is a De Morgan logic such that boundary, monotonicity, associativity, and commutativity conditions together with the involutive complementation are imposed, then we have FDNF and FCNF for the second-level fuzzy logics:

$$\text{FDNF}^{(2)}(\text{A AND B}) = (A\cap B) \cup (A\cap B)$$

$$\text{FCNF}^{(2)}(\text{A AND B}) = (A\cup B) \cap (A\cup B^c) \cap (A^c\cup B) \cap (A\cup B) \cap (A\cup B^c) \cap (A^c\cup B)$$

(3) If we assume $(\cap,\cup,^c)$ is a De Morgan logic such that boundary, monotonicity, associativity, commutativity, and idempotency conditions together with the involutive complementation are imposed, then we have FDNF and FCNF which are equivalent to the fuzzified extensions of the classical normal forms:[1]

$$\text{FDNF}^{(3)}(A \text{ AND } B) = A\cap B = \text{DNF}(A \text{ AND } B)$$

$$\text{FCNF}^{(3)}(A \text{ AND } B) = (A\cup B) \cap (A\cup B^c) \cap (A^c\cup B) = \text{CNF}(A \text{ AND } B)$$

In particular, it has been shown[1] that

$$\text{FDNF}^{(3)}(A \text{ AND } B) \subseteq \text{FCNF}^{(3)}(A \text{ AND } B)$$

In a similar manner, we can obtain FDNF and FCNF for the three classes of fuzzy logics and for all other metalinguistic expressions.

Computational expressions. At this level, symbolic elements of sets are assigned numeric values, and conjunction, disjunction, and complement operators are chosen. In Aristotle's logic, the assignments are

$$\mu_A : X_1 \to \{0,1\}, \text{ and } \mu_B : X_2 \to \{0,1\}$$

In Zadeh's fuzzy set theory and its logic, the assignments are

$$\mu_A : X_1 \to [0,1], \text{ and } \mu_B : X_2 \to [0,1]$$

Furthermore, for Zadeh's logic, the computational expression of the metalinguistic expression "A AND B" at the third level are

$$\mu_{\mathrm{FDNF}(A \text{ AND } B)}(a,b) = a \wedge b$$

$$\mu_{\mathrm{FCNF}(A \text{ AND } B)}(a,b) = (a \vee b) \wedge (a \vee N(b)) \wedge (N(a) \vee b)$$

where $a \in A$, $b \in B$ are elements of fuzzy sets A,B — that is, $a = \mu_A(x_1)$ and $b = \mu_B(x_2)$ — and $\wedge$ = min, $\vee$ = max, and $N(.)$ is the standard complement.

Interpretations

We can now reinvestigate and reinterpret the law of excluded middle for both the idempotent and nonidempotent operators as examples of our classification discussed above. For the idempotent class, an example is the min-max-standard complement triple. For the excluded middle expression, $\mathrm{FDNF}^{(3)}$ and $\mathrm{FCNF}^{(3)}$ are computed to be

$$(a \wedge N(a)) \vee (a \wedge a) \vee (N(a) \wedge N(a)) \leq a \vee N(a)$$

which results in

$$0.5 \leq a \vee N(a) \leq 1.0 \text{ for } a \in [0,1]$$

This is a type-II semantic uncertainty, that is, $\mu(\mu_A(x)) = a \vee N(a)$. However, it reduces to singletons as opposed to intervals. It is clear that in Zadeh's fuzzy logic we ought not to state that the excluded middle expression holds or does not hold. We ought to instead state that it is satisfied to the degree specified by $a \vee N(a)$ for $a \in [0,1]$.

For the nonidempotent class, consider the bold intersection-union-standard complement De Morgan logic where $T_B(a,b) = \max(0, a+b-1)$, $S_B(a,b) = \min(1, a+b)$, and $N(a) = 1-a$. For this case we obtain

$$0.0 \leq \mu_{\mathrm{FDNF}^{(2)}(A \text{ OR NOT } A)} \leq \mu_{\mathrm{FCNF}^{(2)}(A \text{ OR NOT } A)} = 1.0$$

and have an interval of graded values where the excluded middle expression is satisfied to a continuum of degrees in a subinterval of $[0,1]$ that is bound by its lower bound $\mu_{\mathrm{FDNF}^{(2)}(A \text{ OR NOT } A)} \in [0,1]$, and its upper bound $\mu_{\mathrm{FCNF}^{(2)}(A \text{ OR NOT } A)} = 1$.

Conclusions

I have demonstrated that there are three basic transformations between four levels of knowledge representation. Each metalinguistic expression is transformed to at least two propositional expressions known as the fuzzy disjunctive and conjunctive norms forms: FDNF and FCNF, respectively.

A consequence of this FDNF(.), FCNF(.) bounds is that classical expressions such as "excluded middle," "contradiction," and "equivalence," and any combination of two or more vague evidences, must be reinterpreted. The type-I fuzzy representation of linguistic expressions provides only a myopic interpretation of these expressions. These interpretations need to be restated: The fuzzified versions of the laws of classical logic hold to the degree specified by a type-II, second-order, semantic uncertainty computed by the membership of the membership grades, that is, $\mu(\mu_A(x)) = \mu^2{}_A(x)$. Thus, we cannot state, for example, that the law of excluded middle is satisfied or not. We can, however, state that the excluded middle expression is satisfied to a degree contained in the interval specified by:

$$[\mu_{\mathrm{FDNF}(A \text{ OR NOT } A)}(a, N(a)), \mu_{\mathrm{FCNF}(A \text{ OR NOT } A)}(a, N(a))]$$

Reinterpretations for contradiction, equivalence, and so on can be stated in a similar manner. In fact, this is the source of controversy surrounding Elkan's paper.

The essence of fuzzy set theory is that all vague statements should at least be interpreted first with type-I semantic uncertainty at the primary, elemental level. But when two or more vague concepts are combined with a linguistic connective, then we are confronted with a type-II, second-order, semantic uncertainty. This generates an interval where the location of a specific degree of membership is nonspecific in that interval.

Acknowledgments

This work was supported in part by the Manufacturing and Research Corporation of Ontario (MRCO), and the Natural Science and Engineering Council of Canada.

References

1. I.B. Türksen, "Interval-Valued Fuzzy Sets Based on Normal Forms," *Fuzzy Sets and Systems*, Vol. 20, No. 2, Sept. 1986, pp. 191–210.
2. I.B. Türksen, "Fuzzy Normal Forms," to appear in *Fuzzy Sets and Systems,* 1994.

Burhan Türksen is a principal investigator in the Expert Systems Laboratory of the Manufacturing and Research Corporation of Ontario at the University of Toronto. His research interests include fuzzy sets and logics, approximate reasoning, knowledge representation and inference, and manufacturing and process industries, emphasizing management decision support systems and intelligent systems control. Türksen edited the book series *Fuzzy Logic for Decision and Control*, and is on the editorial board of *Fuzzy Sets and Systems, Approximate Reasoning, Decision Support Systems, Information Sciences, Expert Systems and Applications, Turkish Operations Research Transactions, Fuzzy Logic Reports and Letters*, and the *Encyclopedia of Computer Science and Technology*.

Türsken received his PhD in systems management and operations research in 1969, and his MS and BS in industrial engineering in 1962 and 1960, respectively, all from the University of Pittsburgh. He is a senior member of the IEEE and a member of Alpha Pi Mu, IIE, CSIE, CORS, IFSA, NAFIPS, APEO, APET, TORS, and ACM.

Readers can contact him at the MRCO-Intelligent Fuzzy Systems Laboratory, Department of Industrial Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada; Internet: turksen@fuzzy.ie.utoronto.ca

RESPONSE

The Promising Future of Fuzzy Logic

Nader Vadiee and Mohammad Jamshidi, University of New Mexico

As Charles Elkan points out in his article, the foundation of fuzzy logic is the notion of partial truth and degrees of truth in any proposition stating facts about real-world objects, whether these objects are entities, events, relations, algorithms, systems, or machines. Facts and propositions are uncertain, ambiguous, and incomplete — and more importantly, they are goal-oriented, intentional, and subjective to the observer's perceptual capabilities, mental constructs, and meaning systems. In this philosophical view, the universe is seen as holistic, dynamic, and chaotic.

Fuzzy logic is basically a theory of human perception and cognition. It is concerned with the marvelous paradigm and methodology discovered by evolution and realized in our brains to cope with complexity, holism, dynamism, and chaos in the world around us.

The goal of a fuzzy expert system is to take in subjective, partially true facts that are randomly distributed over a sample space, and build a knowledge-based expert system that will apply certain reasoning and aggregation strategies to make useful decisions. These decisions are again approximate, and have partial degrees of truth and likelihood; the decisions and derived facts are reliable to the best of our available knowledge.

The important fact about these systems is that decisions made by them can be iteratively and adaptively improved, and as more such random/fuzzy facts accumulate, the results will converge to real precise facts. In this view of reality, no proposition is always 100% true for 100% of the observers and experts. Absolute certainty, absolute truth, and absolute objectivity are impossible because they require infinite pieces of information, infinite number of samples, and infinitely many observers. Fuzzy logic-based models are actually efforts in building our perceptual models and maps of reality, and not the reality itself.

Fuzzy logical equivalence. Like any other notion in fuzzy logic, the notion of logical equivalence is based on degrees of truth. The fourth postulate in Elkan's Definition 1 is not necessary, and can be replaced by classical implication relations. For example,

$$t(A = B) = t\{(A \rightarrow B) \text{ AND } (B \rightarrow A)\} = \min\{t(A \rightarrow B), t(B \rightarrow A)\}.$$

This expression can be used as the definition for degree of equivalence in fuzzy logic. For the special case where A and B have truth values of 1 or 0, the degree of logical equivalence is equal to 1 for the case of $t(A) = t(B)$. Consider the following special case. Using the classical implication relation that was generalized by Zadeh for fuzzy logic,[1,2] we have

$$t(A \rightarrow B) = \max\{\min[\, t(A), t(B)],\ 1-t(A)\}$$
$$t(B \rightarrow A) = \max\{\min[\, t(B), t(A)],\ 1-t(B)\}$$

for $t(A) = 0$, $t(A = B) = 1 - t(B)$. For $t(B)$ being a number between 0 and 1, the degree of equivalence will be in a range from 0 to 1.

As seen in the above equations, for the case where $t(A) = t(B)$, $t(A = B)$ is always greater than 0.5 in fuzzy logic, which means a strong logical equivalence. Two propositions could be logically equivalent in a fuzzy sense without $t(A) = t(B)$.

We agree with Elkan's point that the last postulate of Definition 1 is the most controversial piece. He has in fact provided his own answer for preserving the continuum of degrees of truth.

Fuzzy expert systems. The types of uncertainty captured by fuzzy logic are vagueness, incompleteness, and ignorance. An example of this is the fuzzy expert systems developed for Japan's Stock Exchange Market in Tokyo. The Nikkei average has been reportedly gone consistently higher using fuzzy logic.[3] However, real applications of fuzzy expert systems have, for the most part, been kept out of the public eye because much of the work is proprietary.

As far as the domain independence of fuzzy operators is concerned, it is well known that max-min operations are default operations, and there are many different definitions suggested by the fuzzy logic research community for "and," "or," and implication operation.[1,2] Of course, aggregation operators are important and context dependent, but they can be a part of the knowledge to be learned and gathered from the expert.

Consider Elkan's watermelon example about the context dependency of the "and" aggregators: If being red inside and green outside are believed to be mutually reinforcing pieces of evidence toward being a watermelon, then the logical proposition could read:

> If X is *red* inside and X is *green* outside, then X is a watermelon is *very true*.

In this example, Elkan is using the fourth postulate to reach an intuitively incorrect conclusion. Based on the definition of the fuzzy logical "and" operation, t(red inside and green outside) is simply the degree to which an object is "red inside" and "green outside" and does not have anything to do with being a watermelon. The degree of being a watermelon depends on the other circumstantial information as well as the degree of being red inside and green outside. This "other piece of information" is the degree of logical equivalence that must be provided by the expert.

Fuzzy expert systems have been used in many applications. For example, Parkinson & Duerre have used both expert systems and fuzzy expert systems to choose the most suitable new "technology" for oil recovery.[4] In the case of classical expert systems, the sharpness of the boundaries of crisp variables involved in this application led to wrong conclusions based on the *if-then* rules. The fuzzy expert system took care of all the limiting (worst-case) problems and made natural conclusions. Although these worst-case problems are not the most common for this application (in which the recovery technologies are somewhat outdated), their occurrence will become the rule rather than the exception in future years. As it is now, major oil reserves in the US cannot be recovered by the old technologies.[4]

Fuzzy control. Most of the current applications of fuzzy logic are fuzzy expert control systems. Fuzzy controllers are expert control systems that smoothly interpolate between otherwise crisp (or predicate logic-based) rules. Rules fire to continuous degrees and the multiple resultant actions are combined into an interpolated result. The basis of fuzzy control is provided by processing uncertain information and saving energy through the use of commonsense rules and natural-language statements.

The use of sensor data in practical control systems involves several tasks that are usually done by a person, such as an astronaut adjusting the position of a satellite or putting it in the proper orbit, or a driver adjusting a car's air conditioning unit, and so on. All such tasks must be performed based on an evaluation of the data according to a set of rules that the person has learned from experience or has been trained in. Often, if not most of the time, these rules are not crisp (based on binary logic), that is, they involve common sense and human judgment in the decision making process. Such problems can be addressed by a set of fuzzy variables and rules that, if calculated and executed properly, can make expert decisions.

Fuzzy logic has given a new definition to the causality in dynamic systems. Fuzzy relational equations[1,2] are indications of the notion of degree of causality between input and output variables in a dynamical system. Like any other notion, causality is not a matter of black or white, or yes or no; instead, the cause-and-effect relation itself is a matter of degree. As Elkan correctly observes, the advent of the fuzzy chip, which came on the market in 1987, is a major force behind the spread of industrial applications of fuzzy logic control.

In reference to the use of words such as "image stabilization" for fuzzy logic camcorder image stabilizing systems, or "grade logic " for fuzzy logic, Elkan brings out the common difficulties that English-speaking Western communities have with this new technology, and with the innocent word "fuzzy." It is not surprising in light of this bias that manufacturers chose alternative words in their advertisements in the US, and to a lesser extent, in other English-speaking countries.

As far as the standard architectures of fuzzy control are concerned, a small number of rules are an advantage for fuzzy control systems. This is evidently a result of interpolative reasoning and the ability to aggregate the overlapping pieces of fuzzy information. Elkan brings up the point indicated by Sugeno and his colleagues — that the knowledge recorded in a fuzzy controller typically reflects immediate relations between the inputs and outputs of the system to be controlled, as opposed to a deep causal model of the system.[5] Although this point of view is accurate, it is also true that this is the exact manner in which human experts summarized their expertise — by capturing the causal links between the inputs and outputs of the systems and putting them in the form of a set of linguistic rules. The expert might have deep knowledge of the system's causal relationships, but it is hard to access that type of knowledge in the form of linguistic protocols. For example, the knowledge of an operator with 20 years of experience at an electric power substation cannot be tapped in a few simple linguistic rules to offer a deep knowledge about the transience and stability of a power system.

Short development times have been a big advantage of fuzzy logic in control systems. To achieve quick design periods, simple rules have been used thus far to put the designer in the ball park, and although approximate and crude, through tuning and adaptation the rules are fine tuned for better performance of the overall system. It is true that most current applications of fuzzy logic could use other rule-based formalisms, but these come with costs in terms of memory, efficiency, development times, and longer compilation of vague linguistic types of knowledge. For example, consider the following type of proposition:

> Most experts believe if X is A, then Y is B is *very true* and *fairly likely*.

There are techniques that can handle this type of vague logical proposition that have elements of both probability and possibility.[1,2] Elkan brings up the brittleness of rule-based systems caused by a missing piece of information. This is not the case for fuzzy rule-based expert systems. As mentioned earlier, this is due to the interpolative capabilities of fuzzy logic's continuous aggregation of the rules and elastic semantics assigned to the symbols, as defined by the membership functions.

Fuzzy control, as we mentioned earlier, constitutes a major application area of fuzzy logic. With most control systems, based on some real data from certain sensors, some decision must be made through a decision process. Fuzzy controllers are nonlinear controllers that provide rather reasonable robustness and adaptiveness with the changing environment — be it unmodelled dynamics in the system, external disturbance, or simply a lack of precise knowledge about the plant that is being controlled.

The subjectivity in fuzzy modeling is a blessing rather than a curse. The subjectivity in the definition of the terms is compensated for by the subjectivity of the conditional rules used by an expert. Because the set of variables and their meanings, as represented by corresponding membership functions, are compatible and consistent with the set of conditional rules used, the overall outcome turns out to be objective, meaningful, and reliable. Fuzzy mathematical tools and the calculus of fuzzy *if-then* rules opened the way for the automation and use of a huge body of human expertise that has gone untapped for years in industry. Fuzzy logic has provided a mechanism to share, communicate, and transfer a wealth of human technical expertise into computers. This has reversed the trend of machine tyranny: We are now forcing computers to think like people, rather than the other way around. This is the beginning of a new era in the applications of AI, neural networks, fuzzy logic, genetic algorithms, and probabilistic reasoning within a bigger picture called *soft computing*.[6] The fundamental issues of AI can only be solved with an orchestrated application of fuzzy logic, neural networks, genetic algorithms, and probabilistic reasoning.

Elkan is distinguished from most critics of fuzzy logic because he seems to have sincerely studied the subject from both a theoretical and applied point of view. It seems to us, however, that Elkan's primary contact with fuzzy logic has been through open literature rather than industrial applications and the tremendous activity across the industrial world.

Some of the shortcomings that Elkan attributes to applied fuzzy logic are due to the gap that exists between theory and application, despite the revolution in the industrial use of fuzzy logic. We believe, however, that it is too soon to scrutinize this gap. For example, at the University of New Mexico's CAD Laboratory for Intelligent and Robotic Systems, fuzzy logic technology is being put on a chip to be embedded in a new generation of controllers with large industrial and technology transfer implications.[7] We are trying to introduce the next generation of fuzzy expert systems capable of handling truth qualifiers, quantifiers, rule interaction, chaining, and hierarchical rule structures.

By starting to think in terms of a holistic, relativistic, probabilistic, and possibilistic knowledge structures, we believe scientific thinking is entering its new major stage of maturity. Crisp, binary, deterministic, first-principle-based approaches in modeling the real world belong to the childhood years of science. The scientific thought that began with Aristotelian logic and was followed by Laplacian determinism has reached its limitations — particularly when it comes to understanding human systems. In the last hundred years, we have witnessed the development of quantum mechanics, and with it, probabilistic notions of microcosm, relativistic mechanics for macrocosm, and more recently, fuzzy logic and chaos theories. The emergence of these theories have a philosophical implication that points toward a probabilistic and possibilistic picture of reality.

Fuzzy logic — with the help of probability theory — will provide yet another powerful tool in an engineer's or scientist's toolbox for coping with complexity and nonlinearity in real-world systems. It will also furnish answers that are never 100% accurate and certain, but are acceptable within the constant constraints of real time, energy, memory, and resources.

References

1. N. Vadiee, "Fuzzy Rule-Based Expert Systems I," in *Fuzzy Logic and Control: Hardware and Software Applications*, M. Jamshidi, N. Vadiee, and T. Ross, eds., Vol. 2, Prentice Hall, Englewood Cliffs, N.J., 1993.
2. N. Vadiee and M. Jamshidi, "A Tutorial on Fuzzy Rule-Base Expert System Models I: Mathematical Foundations," *J. Intelligent and Fuzzy Systems*, Vol. 1, No. 2, 1993, pp. 171-188.
3. M. Jamshidi, T.J. Ross, and N. Vadiee, *Fuzzy Logic With Industrial Applications,* National Short Course Notes, Univ. of New Mexico, Albuquerque, N.M., 1991-94.
4. M. Jamshidi, N. Vadiee, and T. Ross, eds., *Fuzzy Logic and Control: Hardware and Software Applications*, Vol. 2, Series on Environmental and Intelligent Manufacturing Systems, Prentice Hall, Englewood Cliffs, N.J., 1993.
5. M. Sugeno et al., "Fuzzy Algorithmic Control of a Model Car by Oral Instructions," *Fuzzy Sets and Systems,* Vol. 32, No. 2, 1989, pp. 135-156.
6. F. Aminzadeh and M. Jamshidi, eds., *Soft Computing: Fuzzy Logic, Neural Networks, and Distributed Artificial Intelligence*, Prentice Hall, Englewood Cliffs, N.J., 1993.
7. G.W. Parkinson and K. Duerre, "A Comparison of Crisp and Fuzzy Logic Methods for Screening Enhanced Oil Discovery Technologies," *Fuzzy Logic and Control: Hardware and Software Applications*, M. Jamshidi, N. Vadiee, and T. Ross, eds., Prentice Hall, Englewood Cliffs, N.J., 1993.

Nader Vadiee is a teaching and research assistant at the CAD Laboratory of Intelligent and Robotics Systems, Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque. He received his BS and MS degrees from Shiraz University (formerly Pahlavi University), Shiraz, Iran, in 1975 and 1978, respectively. He will receive his PhD from the University of New Mexico in September, 1994. Vadiee has more than 20 years of teaching and research experience at engineering schools, where his work has focused on robot control, neural networks, and fuzzy logic. His doctoral dissertation is on cognitive systems in integrated neuro-fuzzy architecture for nonlinear control and system identification. He is co-editor of *Fuzzy Logic and Control: Software and Hardware Applications*, and is a member of the IEEE Control and Computer Societies. Vadiee can be reached at the CAD Laboratory for Intelligent and Robotic Systems, Dept. of Electrical and Computer Engineering, Univ. of New Mexico, Albuquerque, N.M., 87131; Internet: Manimar@carina.unm.edu

Mohammad "Mo" Jamshidi is a professor of electrical and computer engineering at the University of New Mexico, Albuquerque where he is director of the Computer-Aided Design Laboratory for Intelligent and Robotic Systems and holds the AT&T Professorship of Manufacturing Engineering. He is also a consultant with the Los Alamos and Oak Ridge National Laboratories, and is the founder of TSI Enterprises. He has more than 300 publications, including 32 books. His latest book is *Soft Computing: Fuzzy Logic, Neural Networks, and Distributed AI* (Prentice Hall, 1994). He is editor of several publications, including the *International Journal of Computers and Electrical Engineering,* and was founding editor of *IEEE Control Systems Magazine*, and cofounding editor of *International Journal of Environmentally Conscious Manufacturing.* He earned his BS cum laude from Oregon State University, Corvallis, in 1967, and his MS and PhD from the University of Illinois at Urbana-Champaign in 1969 and 1971, respectively. He is a member of several honor societies, a fellow of the IEEE, and recipient of the IEEE Centennial Medal and IEEE Control Systems Society Distinguished Member Award. He is the honorary chaired professor at Nanjing Aeronautical Institutes, Nanjing, China, and Xia'n Institute of Technology, Xia'n, China. Jamshidi can be reached at the CAD Laboratory for Intelligent and Robotic Systems, Dept. of Electrical and Computer Engineering, Univ. of New Mexico, Albuquerque, N.M., 87131.

RESPONSE

Toward a Framework for Fuzzy Dynamic Systems

Pei-Zhuang Wang, Sie-Keng Tan, and Shaohua Tan, National University of Singapore

Fuzzy logic, according to Lotfi Zadeh, can be broadly considered as the union of fuzzified crisp logics. Its primary aim is to provide a formal, computationally oriented system of concepts and techniques for dealing with modes of reasoning that are approximate rather than exact. Charles Elkan's claims are derived mainly from entangled interpretations of fuzzy logic stemming from his mathematical approach to the formal system and intuitionistic approach to the practical system. Here we examine the mathematical structures of classical and fuzzy logic, and then point out that Elkan's view of the standard version of fuzzy logic is not valid. We then attempt to envisage fuzzy logic, in its practical aspect, as a dynamic system that will enhance control and expert systems.

Mathematical aspect of fuzzy logic. It is well known that the fundamental rules of classical logic are governed by the structure of a Boolean algebra, defined as follows:

Definition 1. A *Boolean algebra* $(B, \vee, \wedge, ^c, o, i)$ is a system consisting of a nonempty set B together with two binary operations $\vee$ and $\wedge$, a unary operation ^{c}, and two nullary operations o and i on B, that satisfies the following axioms for any elements $a, b, c \in B$:

(1) Commutative laws:
$a \vee b = b \vee a$
$a \wedge b = b \wedge a$

(2) Associative laws:
$(a \vee b) \vee c = a \vee (b \vee c)$
$(a \wedge b) \wedge c = a \wedge (b \wedge c)$

(3) Absorption laws:
$(a \vee b) \wedge b = b$
$(a \wedge b) \vee b = b$

(4) Idempotent laws:
$a \vee a = a$
$a \wedge a = a$

(5) Distributive laws:
$(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c)$
$(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)$

(6) Involution law:
$(a^c)^c = a$

(7) De Morgan's laws:
$(a \vee b)^c = a^c \wedge b^c$
$(a \wedge b)^c = a^c \vee b^c$

(8) Laws of Excluded Middle:
$a \vee a^c = i$
$a \wedge a^c = o$

Here, the nullary operations o and i are commonly known as the least element and the greatest element of the Boolean algebra.

Due to the pointwise definition of the operations used in the theory of two-valued logic, we can consider classical logic as a theory that is based upon the particular Boolean algebra $(\{0,1\}, \vee, \wedge, \neg, 0, 1)$ where 1 and 0 represent respectively the *true* and *false* of a certain statement or an assertion, and the operations $\vee$, $\wedge$, and $\neg$ are defined according to the usual rules of the logical connectives *or, and,* and *not,* respectively. Zadeh's innovation of fuzzy logic, on the other hand, is an attempt to generalize the classical two-valued logic. Instead of the two values *true* and *false* represented by the set $\{0, 1\}$, he considered the interval $[0,1]$ to be the range of the truth value of any assertion, and replaced the binary operations $\vee$ and $\wedge$ on $\{0,1\}$ by the binary operations max and min on $[0,1]$. The unary operation $\neg$ is also replaced by c where $c(\alpha)$ is defined to be $1-\alpha$ for any $\alpha \in [0,1]$. Under these operations, the system $([0,1], \max, \min, c, 0, 1)$ satisfies all the axioms of a Boolean algebra except the laws of excluded middle. Such a system is known as a soft algebra defined as follows:

Definition 2. A *soft algebra* $(S, \vee, \wedge, ^c, o, i)$ is a system consisting of a nonempty set S together with two binary operations $\vee$ and $\wedge$, a unary operation ^{c}, and the nullary operations o and i on S, that satisfies the axioms 1 to 7 stated in Definition 1.

Thus, fuzzy logic theory can be seen as a theory based on the structure of a soft algebra. It is clear that every Boolean algebra is a soft algebra, but not vice versa. Hence, soft algebra is a more general system than Boolean algebra. Consequently, propositions that are valid in classical logic may not be valid in fuzzy logic. For example, if we view each assertion A as a set in a universe U and identify the truth value $t(A)$ of the assertion A by its characteristic function $\mu_A : U \to \{0,1\}$, then the two compound statements $(A \wedge B^c)^c$ and $B \vee (A^c \wedge B^c)$ are logically equivalent according to the rules of classical two-valued propositional calculus; however, in the context of fuzzy logic, these two statements with truth value in $[0,1]$ are not equivalent. (For example, take $t(A) = 0.3$ and $t(B) = 0.6$, then $t((A \wedge B^c)^c) = 0.7$, whereas $t(B \vee (A^c \wedge B^c)) = 0.6$.

In his article, Elkan views a standard version of fuzzy logic as a system that satisfies the four postulates given in the following definition:

Definition 3. Let A and B be arbitrary assertions. Then

$$t(A \wedge B) = \min \{t(A), t(B)\} \quad (1)$$
$$t(A \vee B) = \max t(A), t(B)\} \quad (2)$$
$$t(\neg A) = 1 - t(A) \quad (3)$$
$$t(A) = t(B) \text{ if } A \text{ and } B \text{ are logically equivalent,} \quad (4)$$

where "logically equivalent" means equivalent according to the rules of classical two-valued propositional calculus.

Certainly, under these postulates, one can prove that for any assertions A and B, either $t(A) = t(B)$ or $t(A) = 1 - t(B)$. However, the main issue here is that postulate 4 is generally not valid in the realm of fuzzy logic, as shown in the above example. Thus, as a rigorous mathematical system, postulate 4 should not be included as a postulate in the formal system of fuzzy logic. Although Elkan has stated that in fuzzy logic applications it is unclear whether or not postulate 4 is assumed, and that in theoretical work it is often used explicitly, he still imposes this postulate in the formal system of fuzzy logic and then claims that the standard version of fuzzy logic collapses mathematically to two-valued logic. This type of condemnation, in our view, is an impediment to the growth of our knowledge.

Practical aspect of fuzzy logic. At the present stage, fuzzy mathematics is viewed in two ways. First, it is a theory that conforms with the precise and rigorous principles of mathematics to deal with fuzzy objects. In other words, it is a strictly mathematical theory to study the objects in a fuzzy environment. Second, fuzzy mathematics is a metamathematical theory that involves fuzzy proof techniques and fuzzy theorems with their applications. This latter status has yet to be fully developed. We have already seen many successful applications of fuzzy logic that the use of conventional mathematics could not achieve. We envisage that as fuzzy mathematics develops further, applications will be even more convincing and prominent.

Elkan and other researching technologists perhaps view fuzzy logic as fuzzy mathematics in its second status, and this might explain why Elkan has been unable to find a real-world expert system that uses fuzzy logic as its primary formalism for reasoning under uncertainty.

Fuzzy logic, in the practical aspect, deals mainly with fuzzy quantization, its meaning and means. In this respect, Elkan regards those operators in fuzzy logic as fixed and domain-independent. In fact, fuzzy quantization is introduced precisely for the purpose of generating domain-specific quantities. The numerous forms of fuzzy operations suggested in the literature were created to cater to the domain-specific needs.[1,2] Other semantically dependent formulations of fuzzy operations and inference relations have also been proposed.[3,4]

Elkan's paper does bring up some valid points in the discussion of the status quo of fuzzy control. Indeed, the present fuzzy controllers are mostly structurally shallow, and in most cases, the controllers simply deal with no more than a simple static fuzzy mapping of the sensory and actuation signals. However, this is not the whole picture of fuzzy control. In fact, the success of Aptronix's simulation of the two-stage inverted pendulum using a fuzzy controller is a fuzzy logic application that is not structurally shallow. When fuzzy logic is used as a way of quantization, it can serve as our quantity basis for modeling dynamic systems in the real world. This leads to the notion of fuzzy dynamic systems. Obviously, fuzzy dynamic systems are more complex, as they describe dynamic evolution of certain fuzzy quantities, not simple points or numbers. Undoubtedly, in the light of such a theory, many important issues such as stability, controllability, and observability can properly be addressed, and it may also serve to bring the seemingly diverging model-based or rule-based methodologies into a unifying framework.

An appropriate theory for fuzzy systems has not yet been developed in fuzzy control. The main task is to establish a framework in which fuzzy controllers of deeper structures can be described properly and handled with ease. Elkan has predicted a tough time ahead for fuzzy logic in general, and for fuzzy control in particular. We, too, predict a tough time ahead in working out a meaningful and acceptable framework for fuzzy-based dynamic system theory. However, we remain optimistic. We believe that such a framework will emerge.

References

1. I.R.Goodman and H.T. Nguyen, *Uncertainty Models for Knowledge-Based Systems*, Elsevier, New York, 1985.
2. E.H. Ruspini, "Approximate Reasoning: Past, Present, Future," Tech. Note 492, SRI Int'l, 1990, pp. 1–22.
3. S.K. Tan, P.Z. Wang, and E.S. Lee, "Fuzzy Set Operations Based on the Theory of Falling Shadows," *J. Mathematical Analysis and Applications,* Vol. 174, No. 1, 1993, pp. 242–255.
4. S.K. Tan, P.Z. Wang, and X.Z. Zhang, "Fuzzy Inference Relations Based on the Theory of Falling Shadows," *Fuzzy Sets and Systems,* Vol. 53, 1993, pp. 179–188.

Pei-Zhuang Wang is professor and head of the Fuzzy Information Processing and Fuzzy Computing National Laboratory, Beijing Normal University, Beijing, China. He is also a researcher at the Institute of Systems Science, National University of Singapore. He has served in a number of national and international societies including as the vice president of the International Fuzzy Systems Association (1991-93), and chair of the Chinese Chapter of the International Fuzzy Systems Association. He is honorary chair of Aptronix, San Jose, Calif. He is on the editorial board of several international and local technical journals on fuzzy logic, and has been given awards by various organizations including China's National Personal Ministry and Ministry of Education. He has published five books and more than 100 technical papers in international journals and conferences. His research interests are in fuzzy mathematics theory, probability, and AI systems.

Sie-Keng Tan is a senior lecturer in the Department of Mathematics, National University of Singapore. He earned his BS and PhD from Nanyang University, Singapore, and his MPhil from Queen Mary College, University of London, England. His main research interest is in the mathematical structure of the theory of truth-value flow inference of fuzzy logic. S.K. Tan is a member of the International Fuzzy Systems Association (IFSA).

Shaohua Tan is a senior lecturer in the Department of Electrical Engineering, National University of Singapore. He has been working in the areas of systems and control, digital signal processing, speech processing, and artificial neural nets. His current research interests are in intelligent control and pattern recognition using neural net and fuzzy inference techniques. He received his BE degree in mechanical engineering from the Beijing Institute of Technology, China, in 1982, and ME and PhD degrees in electrical engineering from the Katholieke Universiteit Leuven, Belgium, in 1984, and 1987, respectively. Shaohua Tan can be reached at the Department of Electrical Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511; Internet: eletansh@leonis.nus.sg

RESPONSE

Misrepresentations and Challenges

A Response to Elkan

Ronald R. Yager, Iona College

The comments made by Charles Elkan can be classified into three categories. The first are those that are technically incorrect and should not have been allowed to pass an unbiased review process. The second are those that are truly challenging, and in point of fact help show the great representative power of fuzzy subsets. The third category are those pertaining to the practical application of fuzzy control. Some of these latter comments are quite reasonable — though not as damning as Elkan tries to make them. I shall address these issues in turn.

The first category — those that are completely incorrect — is dominated by Theorem 1. Rather than wasting considerable space on addressing this "theorem" when I am certain that other respondents will effectively show its complete absurdity, I shall make only a few comments. The key issue here is of course the last premise,

$t(A) = t(B)$ if A and B are logically equivalent.

In most texts on logic,[1] the definition of logical equivalence is specified the other way. Usually, one says that A and B are logically equivalent if propositions A and B attain the same truth value for all models of the constituent atoms. However, in this case, rather than defining the *undefined* concept of logical equivalence in terms of a well-defined idea of attaining the same truth value, Elkan tries to define the idea of attaining the same truth value from the undefined concept of logical equivalence. Once having made this error, the author then compounds it by imposing a requirement that is completely antithetical to the idea of fuzzy logic:

$$t(\neg(A \wedge \neg B)) = t(B \vee (\neg A \wedge \neg B)).$$

First we note that

$$t(\neg(A \wedge \neg B)) = t(\neg A \vee B),$$

and hence

$$\begin{aligned} t(A \vee (\neg\neg A \vee \neg B)) &= t((\neg A \vee B) \wedge (B \vee \neg B)) \\ &= \min[t(\neg A \vee B), t(B \vee \neg B)]. \end{aligned}$$

Thus we have the requirement

$$t(\neg A \vee B) = \min[t(\neg A \vee B), t(B \vee \neg B)].$$

For this condition to hold for every B requires that

$$t(B \vee \neg B) = 1$$

for all B. However, this condition is the law of the excluded middle and is exactly what fuzzy logic was constructed not to support. In fact, I have suggested a measure of fuzziness based upon the lack of satisfaction of this condition.[2] Furthermore, the condition $t(B \vee \neg B) = 1$ implies that

$$\max[t(B), 1 - t(B)] = 1.$$

Hence, either

$$t(B) = 1 \text{ or } t(B) = 0\ (1 - t(B) = 1).$$

Thus, Elkan has essentially assumed that the logic is a binary logic.

As to the second category, Elkan raises the issue of defining the concept of watermelon in terms of the constituent concepts of redness on the inside and greenness on the outside. His basic contention is that the definition of watermelon in terms of these constituents should exhibit a characteristic of reinforcement. Essentially, he correctly requires that multiple confirmations to the constituents' criteria should reinforce each other, while disconfirmations of the constituents' criteria should also reinforce each other in the other way. The issue raised here is an interesting and challenging question. However, rather than showing the limitations of fuzzy logic, this problem illustrates the power of fuzzy logic to model sophisticated aggregation requirements.

It is fundamental to a comprehensive understanding of the agenda of fuzzy logic to appreciate the pervasive nature of its ability to model continuity and graduality in all concepts. In using fuzzy logic, we are not confined to only using the idea of fuzziness (graduality) in the definition of the predicates (redness/greenness), but we can also apply the concept of fuzziness to the operators used to connect the predicates. In addressing this important issue, we must call upon fuzzy logic's ability to provide connectives lying between the logical *and* and logical *or*. Consider the definition of watermelon suggested by Elkan,

$$watermelon(x) = redinside(x) \perp greenoutside(x).$$

Elkan correctly shows that if we interpret $\perp$ as a pure logical *and*, defined as the $\min(\wedge)$ we end up with a result that doesn't provide the appropriate property of reinforcement. Similarly, using a pure logical *or*, defined as the $\max(\vee)$ also leads to unsatisfactory results. The key point is that in fuzzy logic we are not restricted to these two extremes as we are in binary logic.

Recently my colleagues and I introduced a new class of fuzzy connectives, called uninorms,[3] that provide the exact type of aggregation postulated as being required by Elkan. Consider the situation that for some object m we have $t(redinside(m)) = a$, and $t(greenoutside(m)) = b$. Our problem is to provide an aggregation operator $\perp$ to implement the connection between these values. Formally, letting $d = t(watermelon(m))$ we require some aggregation R such that

$$d = R(a, b).$$

The question is, what form should R take to capture the type of reinforcement desired by Elkan? As I will show, uninorms provide the appropriate aggregation. These uninorms, which generalize the idea of t-norms (*and* operators) and t-conorms (*or* operators) and lie between these extremes, do exactly what Elkan requires.

A *uninorm* is a mapping[3]

R: $[0, 1] \times [0, 1] \rightarrow [0, 1]$

having the properties:

(1) Commutativity: $R(a, b) = R(b, a)$
(2) Monotonicity: $R(a, b) \geq R(c, d)$ if $a \geq c$ and $b \geq d$
(3) Associativity: $R(R(a, b), c) = R(a, R(b, c))$
(4) There exists an identity element $\delta \in [0, 1]$ such that for all a, $R(a, \delta) = a$.

If $\delta = 1$, this reduces to the t-norms that are essentially pure *and* aggregations that include the min operator, while if $\delta = 0$, it reduces to the t-conorms which are essentially pure *or* aggregations and include the max operator. Thus, the logical *and* and *or* are extremes of this class.

Of particular interest is a property of these uninorms called the *upward reinforcement characteristic*. For the uninorm, the upward reinforcement characteristic is captured in the following:

$$R(a, b) \geq a \quad \text{if } b > \delta$$
$$R(a, b) \leq a \quad \text{if } b < \delta$$

We select a value δ to be our neutral point; values above δ are considered as confirming and those below δ as disconfirming. (Actually, δ can be a range; however, for the present purpose we'll consider δ as a point.) Now, assume that both a and b are above δ. In this case we have

$$a \leq R(a, b)$$
$$b \leq R(b, a) = R(a, b)$$

and thus

$$R(a, b) \geq \max[a, b].$$

Hence there is a reinforcement in the positive direction when both criteria are "confirmed."

Now assume that both a and b are below δ. In this case we have that

$$a \geq R(a, b)$$
$$b \geq R(b, a) = R(a, b)$$

and thus

$$R(a, b) \leq \min(a, b).$$

Hence there is negative reinforcement if both are below the neutral value.

Finally, consider the case where one is below and one above, $a \leq \delta$ and $b \geq \delta$. In this situation we see

$$a \leq R(a, b)$$
$$b \geq R(b, a) = R(a, b)$$

and hence we get

$$a \leq R(a, b) \leq b$$

and thus there is no reinforcement.

With the aid of these uninorm fuzzy aggregation operators, we can capture the type of aggregation Elkan desires.

Finally, Elkan's comments on the use of fuzzy logic in heuristic control — while in some points are quite valid — manifest a type of "fuzzy bashing" that is all too common in the AI community. For example, Honda's choice of the term "grade logic" has much less to do with their concern for any scientific resistance to fuzzy logic methodology than to the simple marketing expedient that "fuzzy" is not the type of word that sells cars.

In a recent book on fuzzy modeling and control,[4] we look carefully at the process of building fuzzy logic controllers. The reasons we found for the success of these controllers are not in complete agreement with those Elkan suggests.

First of all, the fact that most fuzzy controllers are built with a small number of rules should be seen as one of the powers of this technology. An essential feature of the fuzzy approach is the ability to generalize — in a way, to reduce the necessity for detail.

Elkan fails to mention a feature I think is essential to the success of the fuzzy modeling approach: the partitioning of the input variable space into regions that allow a simplification of the modeling process. Closely related to this is the idea of partial matching, which lets us smoothly combine solutions from different regions as we get near the boundary.

Elkan correctly observes that most fuzzy controllers are shallow (requiring no chaining between the rules) and usually directly connect the input to the output. I think it is here that these systems might have trouble in the future. However, the reason for these potential problems is not found in the paradigm of fuzzy modeling, but in the choice of the implication operative.

To me, Elkan's reference to the 1980 comment by Mamdani and Sembi[5] is most disturbing. Rather than seeing these remarks as I believe Mamdani meant them— as a statement of the power of the symbiotic relationship between the paradigms of AI (in this case rule-based systems) and the knowledge-representation capability of fuzzy logic — Elkan has chosen to interpret this as a sign of the weakness of fuzzy logic. However, if we discard the obvious misrepresentations, Elkan's paper can serve as a challenge to fuzzy researchers to continue improving the valuable tool of fuzzy logic.

References

1. E. Mendelson, *Introduction to Mathematical Logic*, Van Nostrand Reinhold, New York, 1964.
2. R.R. Yager, "On the Measure of Fuzziness and Negation Part I: Membership in the Unit Interval," *Int. J. General Systems*, Vol. 5, 1979, pp. 221-229.
3. R.R. Yager and A. Rybalov, "Uninorm Aggregation Operators," Tech. Report #MII-1407, Machine Intelligence Institute, Iona College, New Rochelle, N.Y., 1994.
4. R.R. Yager and D.P. Filev, *Essentials of Fuzzy Modeling and Control*, John Wiley & Sons, New York, 1994.
5. E.H. Mamdani and B.S. Sembi, "Process Control Using Fuzzy Logic," *Fuzzy Sets*, P.P. Wang and S.K. Chang, eds., Plenum Press, New York, 1980, pp. 249-266.

Ronald R. Yager is director of the Machine Intelligence Institute and professor of information systems at Iona College, and is a research fellow of the Knowledge Engineering Institute, Guangzhou University, China. He is on the scientific committee of the Fuzzy Logic Systems Institute, Iizuka, Japan, and is copresident of the International Conference on Information Processing and Management of Uncertainty, Paris. He is editor-in-chief of the *International Journal of Intelligent Systems*. He also serves on the editorial board of a number of other journals, including the *IEEE Transactions on Fuzzy Systems*, the *Journal of Approximate Reasoning*, and *Fuzzy Sets and Systems*. He has published more than 350 articles and has edited 13 books. He is co-author of a new book, *Essentials of Fuzzy Modeling and Control* (John Wiley & Sons, 1994). He received his undergraduate degree from the City College of New York, and his PhD from the Polytechnic Institute of Brooklyn. Ronald Yager can be reached at the Machine Intelligence Institute, Iona College, New Rochelle, NY 10801; Internet: rry1@iona.bitnet

RESPONSE

Why the Success of Fuzzy Logic is not Paradoxical

Lotfi A. Zadeh, University of California, Berkeley

Elkan's paper consists of two almost unrelated parts. In the first section, Elkan arrives at the conclusion that an apparently reasonable version of fuzzy logic collapses mathematically to two-valued logic. In the second section, he questions the value of fuzzy logic in control applications and concludes that fuzzy logic does not provide an effective tool for dealing with the problem of uncertainty in knowledge-based systems. As I see it, the first conclusion is based on faulty reasoning, while the second reflects a misconception of what fuzzy logic is and a misunderstanding of the role it plays in control and knowledge-based systems applications.

It is easy to show why Elkan's mathematical analysis is faulty. What he really shows is that fuzzy logic is not consistent with the law of the excluded middle. This, of course, applies in general to multivalued logical systems.

The law of the excluded middle asserts that the truth value of any logical expression of the form $B \vee \neg B$ is T (true). The law of contradiction asserts that the truth value of any logical expression of the form $B \wedge \neg B$ is F (false).

Immediate consequences of these laws in two-valued propositional calculus are as follows:

If p is logically equivalent to q then p is also logically equivalent to $q \wedge (B \vee \neg B)$.

If p is logically equivalent to q then p is also logically equivalent to $q \vee (B \wedge \neg B)$.

If p is logically equivalent to q then $p \wedge (B \vee \neg B)$ is logically equivalent to $q \vee (B \wedge \neg B)$.

Now let us consider Elkan's Theorem 1. Starting with the valid equivalence

$$\neg(A \wedge \neg B) = \neg A \vee B$$

which is an expression of De Morgan's law, we can replace the right-hand member with $(\neg A \vee B) \wedge (B \vee \neg B)$, which is in turn equivalent to

$$B \vee (\neg A \wedge \neg B).$$

Consequently, we can assert the logical equivalence

$$\neg(A \wedge \neg B) \equiv B \vee (\neg A \wedge \neg B), \qquad (1)$$

which is the example used in Elkan's proof.

What we see, then, is that Elkan's example uses a disguised form of the law of the excluded middle. As should be expected, Equation 1 is not a logical equivalence in multivalued logic because the law of the excluded middle does not hold, in general, in multivalued logic. In sum, what Elkan shows in a roundabout way is that the law of the excluded middle does not hold in multivalued logic. There is no justification whatsoever for jumping from this obvious fact to the conclusion that fuzzy logic collapses to two-valued logic.

Turning to his analysis of fuzzy logic applications, Elkan's conclusion reflects a misunderstanding of what fuzzy logic is, and a faulty analysis of the reasons for its success. First, it must be clarified that the term "fuzzy logic" is used in two different senses. In its narrow sense, fuzzy logic is a logical system that is an extension of multivalued logic and serves as a foundation for approximate reasoning. What is important to note is that even in its narrow sense, the agenda of fuzzy logic is quite different from that of traditional multivalued systems.

In its wider sense — the sense in which it is predominantly used today — fuzzy logic is a much broader theory that is fuzzily synonymous with "fuzzy set theory," that is, the theory of classes with unsharp boundaries. In this perspective, fuzzy logic in the narrow sense is one of the many branches of fuzzy logic, among which are fuzzy arithmetic, fuzzy probability theory, possibility theory, fuzzy relations, and so on. It should be noted that fuzzy logic in the narrow sense plays a very minor role in fuzzy control, just as classical logic plays a very minor role in classical control theory.

In his article, Elkan fails to differentiate between fuzzy logic in the narrow sense and fuzzy logic. In the first part, he interprets fuzzy logic in its narrow sense. But in the second part, he interprets fuzzy logic in its wide sense, since most applications of fuzzy logic — especially in the realm of control — do not involve fuzzy logic in the narrow sense. However, narrow fuzzy logic plays an essential role in the management of uncertainty in expert systems.[1] In what follows, fuzzy logic will be used in its wide sense.

What are the reasons for the rapid growth in the number, variety, and visibility of fuzzy logic applications? The reasons are not those given in Elkan's article. What fuzzy logic offers, above all, is a methodology for representing and analyzing dependencies that are approximate rather than exact. In this methodology, the key concepts are:

- a linguistic variable, whose values are words rather than numbers;
- a canonical form, which expresses the meaning of a proposition as an elastic constraint on a variable;
- a fuzzy if-then rule and rule qualification, in particular probability qualification and possibility qualification;
- interpolative reasoning; and
- a fuzzy graph.

Through the use of techniques based on these concepts, fuzzy logic makes it possible to exploit the tolerance for imprecision and uncertainty. In so doing, fuzzy logic has proved to be successful where traditional approaches have failed or yielded inferior results.

Most fuzzy logic applications involve the use of what might be called the *calculus of fuzzy rules*.[2-4] The use of fuzzy rules in conjunction with interpolative reasoning greatly reduces the number of rules that are needed to describe imprecise dependencies, and makes it much easier for humans to articulate them. Consider, for example, the rules that people use (consciously or subconsciously) in parking a car, filling a tub with hot water, crossing a traffic intersection, or riding a bicycle. How would Elkan describe the rules that govern human behavior in these and similar instances?

What is actually used in most control applications is a subset of the calculus of fuzzy rules, which can be called the *calculus of fuzzy graphs.*[3-5] In this calculus, a function $f: U \rightarrow V$ is approximated to by a fuzzy graph f^*, which is a disjunction of Cartesian products of the form

$$f^* = \Sigma_i A_i \times B_i$$

where A_i and B_i, $i=1,\ldots, n$, are values of linguistic variables, and Σ_i represents the disjunction (union) of Cartesian products $A_i \times B_i$. For example, a fuzzy graph of a function may be expressed in a coarse way as

$$f^* = \text{small} \times \text{small} + \text{medium} \times \text{large} + \text{large} \times \text{small}$$

which is equivalent to the set of rules

If X is small then Y is small.
If X is medium then Y is large.
If X is large then Y is small.

The use of fuzzy graphs results in data compression, which is one of the key — though perhaps not widely recognized — advantages of using fuzzy rules. Elkan's analysis makes no reference to this point, and fails to identify the use of the fuzzy graph concept as one of the principal tools in the application of fuzzy logic to control.

Today, fuzzy logic applications in control and consumer products are far more visible than fuzzy logic applications in knowledge-based systems. Does this mean, as Elkan surmises, that fuzzy logic is limited in its applicability to simple systems? Not at all. What it means is that fuzzy logic can be applied easily and effectively to the conception and design of "high machine IQ" control systems and consumer products — applications that in most cases involve replacing a trained operator or an experienced user with a fuzzy rule-based system. In the case of knowledge-based systems, what has to be replaced is an expert rather than an operator. This is an inherently more complex problem, no matter what approach is used.

Basically, what differentiates control applications from knowledge-based systems applications is that in control the main problem that has to be addressed is that of imprecision. By contrast, in the case of knowledge-based systems, one has to come to grips with both imprecision and uncertainty.

In applying fuzzy logic to control systems, it is generally sufficient to employ categorical rules — rules that involve no quantifiers, probabilities, or possibilities. In the realm of control, the calculi of fuzzy rules and fuzzy graphs provide the necessary tools for exploiting the tolerance for imprecision and lead to systems that are simpler, more robust and have higher machine IQ than systems designed by conventional methods. Recently published books[6-8] provide easily understandable accounts of the methodology of fuzzy logic control and explain why the applications of fuzzy control are growing rapidly in visibility, variety, and number. It is very likely that it will not be long before familiarity with fuzzy control will be an essential qualification for control engineers and system designers.

In the case of knowledge-based systems, two sources of difficulty are that the rules are frequently probability-qualified, and that the qualifying probabilities are not compositional. More specifically, assume that we have two rules of the form

If p_1 then q_1 $\quad (P_1)$
If p_2 then q_2 $\quad (P_2)$

where p_1, q_1, p_2, and q_2 are propositions, and P_1 and P_2 are qualifying probabilities. Assume that we wish to compute the qualifying probability, P, in the combined rule

If (p_1 and p_2) then (q_1 and q_2). $\quad (P)$

The problem is that P cannot be computed as a function of P_1 and P_2 without making some assumptions about conditional independence or, equivalently, invoking the maximum entropy principle. Such assumptions tend to be ad hoc and hard to justify. What this implies is that the problem of inference from probability-qualified propositions may not have a satisfactory solution within the framework of classical probability theory.

In this connection, it should be noted that Elkan gives the impression that there are many expert systems that do not employ fuzzy logic and that provide effective ways of dealing with uncertainty and imprecision. This is not the case. As a test, which of the systems that he as in mind could provide an answer to the following question:

If X is small then it is very likely that Z is large.

If X is large then it is not likely that Z is large.

What is the probability that Z is large if X is medium?

What this example points to is that the conventional approaches to the management of uncertainty in expert systems fail in four important respects:

(1) They do not provide the means for dealing with the fuzziness of antecedents and consequents.
(2) They assume that probabilities can be estimated as crisp numbers.
(3) They do not offer a mechanism for inference from rules in which the qualifying probabilities are fuzzy.
(4) The rules for composition of probabilities depend on unsupported assumptions about conditional independence.

Fuzzy logic addresses some — but not all — of these problems.[1] More specifically, fuzzy logic allows the antecedents and/or consequents and/or qualifying probabilities to be fuzzy. Furthermore, fuzzy logic makes it possible to estimate probabilities as fuzzy rather than crisp numbers. There remain, however, two problems. First, the composition of qualifying probabilities can lead to fuzzy probabilities that are insufficiently specific or, equivalently, insufficiently informative. Second, inference in fuzzy logic reduces, in general, to the solution of a nonlinear program. Standard techniques for the solution of such programs may be computationally expensive. What we do not have as yet are approximate, inexpensive techniques for inference from fuzzy-probability-qualified fuzzy if-then rules. However, we do have an effective method of inference from possibility-qualified rules within a branch of fuzzy logic known as possibilistic logic.[9]

There are many statements in Elkan's articles that relate to ill-posed questions or reflect a misunderstanding of what fuzzy logic is, as well as an inadequate familiarity with its literature. I will comment here on just a few of these statements.

In the section on fuzzy logic in expert systems, Elkan states, "there is still debate as to what types of uncertainty are captured by fuzzy logic." Obviously if the boundaries of what constitutes fuzzy logic are not defined, this is not a well-posed issue. In this context, what is important to realize is that any theory X can be fuzzified by generalizing the concept of a crisp set in X to a fuzzy set, leading to a theory that can be called fuzzy X. For example, classical probability theory can be generalized to fuzzy-probability theory; topology to fuzzy topology; neural network theory to fuzzy neural network theory; control to fuzzy control; arithmetic to fuzzy arithmetic; modal logic to fuzzy modal logic; resolution to fuzzy resolution; temporal logic to temporal fuzzy logic; Mycin to fuzzy Mycin; chaos to fuzzy chaos, and so on. Many such generalizations have already been described in the literature and many more will be made in the future. What is gained from fuzzification is greater generality and better approximation to reality.

Given that any theory can be fuzzified, the question of what types of uncertainty are captured by fuzzy logic loses much of its meaning. For example, when probability theory is fuzzified, it becomes a part of fuzzy logic. In this broad perspective, then, fuzzy probabilistic uncertainties fall within the scope of fuzzy logic. The same applies to any type of uncertainty that I can think of.

In the same section, Elkan reports that his search of the literature revealed no published reports of an expert system that uses fuzzy logic as its primary formalism. This is somewhat surprising, since there are, in fact, many such examples. Among them is Cadiag-2, the well-known large-scale medical diagnostic system.[10-13] Another well-known and commercially available system is FRIL,[14] which is Prolog-based and has a highly sophisticated system for the management of uncertainty. Still another example is the Yamaichi Securities Fund, and there are many more (see Table 1 on page 46).[15-17]

Elkan also seems to suggest that expert systems that combine grades of membership using operators other than max and min are not valid examples of the use of fuzzy logic. This position is hard to understand since the use of *t*-norms, *t*-conorms, and other connectives is now a standard part of fuzzy logic.[18]

The issue of the management of uncertainty in expert systems presents many complex and difficult problems. There is no system at present that is free of serious shortcomings, and it would be unrealistic to expect that such systems will be developed in the foreseeable future. But Elkan's statement that "experience shows that fuzzy logic is rarely suitable in practice for reasoning about uncertainty" reflects inexperience in the use of fuzzy logic. I advise Elkan to study with care the extensive literature on the management of uncertainty in expert systems based on the use of fuzzy logic. A good starting point would be the treatises by Dubois-Prade on possibility theory and approximate reasoning, and the books on fuzzy expert systems.[16,19] There is little doubt that, in coming years, the growth in familiarity with fuzzy logic will lead to its wide acceptance as a key component of information systems and knowledge engineering methodologies.

I compliment Elkan on writing a provocative article that is likely to contribute to further discussion of the strengths and limitations of fuzzy logic. Fuzzy logic has been and still is somewhat controversial. With the passage of time, however, the controversies will abate and fuzzy logic is likely to become a standard tool for the conception and design of intelligent systems. Indeed, it would not be surprising if, in retrospect, the skeptics will find it hard to understand why they failed to realize that fuzzy logic is a phase in a natural evolution of science — an evolution brought about by the need to find an accommodation with the pervasive imprecision of the real world.

References

1. L.A. Zadeh, "The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems," *Fuzzy Sets and Systems,* Vol. 11, 1983, pp. 199-227.
2. L.A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 3, 1973, pp. 28-44.
3. L.A. Zadeh, "On the Analysis of Large Scale Systems," *Systems Approaches and Environment Problems*, H. Gottinger (ed.), Vandenhoeck and Ruprecht, 1974, pp. 23-27.
4. L.A. Zadeh, "The Calculus of Fuzzy If-Then Rules," *AI Expert*, Vol. 7, No. 3, 1992, pp. 22-27.
5. L.A. Zadeh, "Fuzzy Logic, Neural Networks and Soft Computing," *Comm. ACM*, No. 37, March 1994.
6. D. Driankov, H. Hellendoorn, and M. Reinfrank, *An Introduction to Fuzzy Control*, Springer-Verlag, Berlin, 1993.
7. M. Jamshidi, N. Vadiee, and T. Ross, eds., *Fuzzy Logic and Control,* Prentice Hall, Englewood Cliffs, NJ, 1993.
8. A. Kandel and G. Langholz, eds., *Fuzzy Control Systems*, CRC Press, Boca Raton, 1994.
9. D. Dubois and H. Prade, "Epistemic Entrenchment and Possibilistic Logic," *Artificial Intelligence*, Vol. 50, 1991, pp. 223-239.
10. K.P. Adlassnig, "A Fuzzy Logical Model of Computer-Assisted Medical Diagnosis," *Methods of Information in Medicine*, Vol. 19, 1980, pp.141-148.
11. K.P. Adlassnig, "Fuzzy-Set Theory in Medical Diagnosis," *IEEE Trans. Systems, Man, and Cybernetics*, Vol. 16, 1986, pp. 260-265.
12. K.P. Adlassnig and G. Kolarz, "Representation and Semiautomatic Acquisition of Medical Knowledge in Cadiag-1 and Cadiag-2," *Computers and Biomedical Research*, Vol. 19, 1986, pp. 63-79.
13. K.P. Adlassnig, "Uniform Representation of Vagueness and Imprecision in Patient's Medical Findings Using Fuzzy Sets," in *Proc. Cybernetics and Systems '88*, Kluwer Academic, Dordecht, 1988, pp. 685-692.
14. J.F. Baldwin, "Fril: A Fuzzy Relational Inference Language," *Fuzzy Sets Systems*,

Table 1. Fuzzy logic tools and products.
(Source: Sammy Wong and Nelson Wong, Computer Science Dept., Chinese University of Hong Kong.)

COMPANY	PRODUCT	DESCRIPTION
American Neuralogix	NLX 230 fuzzy microcontroller	Has 8 digital inputs, 8 digital outputs, 16 fuzzifiers; holds 64 rules. Evaluates 30M rules/sec.
	ADS230 fuzzy microcontroller development system	PC-compatible system uses NLX 230 with analog and digital I/O.
	NLX 110 fuzzy pattern correlator	Correlates eight 1-Mbit patterns; expandable to 256 n-bit patterns
	NLX 112 fuzzy data correlator	Performs pattern matching on serial data streams.
Aptronix	Fide (Fuzzy Interference Development Environment)	Runs under MS Windows on 386/486 PCs. Supports development, fuzzy simulation, debug tracing, and 3D display of control surfaces. Real-time code generation for microcontrollers. Software implementation of fuzzy logic in C. Complete tutorial information and phone support.
Byte Craft	Fuzzy-C	Preprocessor translates fuzzy source code into C source code.
Fuzzy Systems Engineering	Manifold editor	Runs under MS Windows 3.1 on 386/486 PCs. Edits rules in a matrix display. Lets users view fuzzy sets graphically.
	Manifold graphics editor	Runs under MS Windows 3.1 on 386/486 PCs. Color graphics display of rules and fuzzy sets. Lets users view designs in 3-D map and slice formats.
Hitachi American	H8/300 and H8/500	Microcontrollers
Hyperlogic	Cubicalc	Software for developing fuzzy-logic applications. Runs under MS-Windows with 286 or higher processor. Simulates fuzzy and nonfuzzy systems.
	Cubicalc-RTC	A superset of Cubicalc. Provides runtime compiler support and libraries for linking. Compatible with Microsoft C and Borland C.
	Cubicalc runtime source code	Generates C source code for use in compiling to a specific processor.
	Cubicard	Includes Cubicalc-RTC and PC-based hardware for analog and digital I/O.
Inform Software	Fuzzytech Explorer Edition	Introductory fuzzy-logic system. Software runs under MS Windows. Accepts two inputs, one output, five fuzzy membership sets per variable, and 125 rules. Includes tutorial.
	Fuzzytech MCS-96 Edition	Full fuzzy development system for MCS-96 microcontrollers. Generates optimized assembly code.
	Fuzzytech Online Edition	Lets users debug and modify fuzzy-logic systems while they are running. Generates C source code.
Integrated Systems	RT/Fuzzy Module	Simulation and code generation of fuzzy logic for real-time systems.
Metus Systems Group	Metus	Fuzzy-logic development and simulation system. Runs under MS DOS. Provides high-level modeling and low-level development for embedded applications.
Modico	Fuzzle 1.8	PC-based fuzzy-logic shell. Generates source code for C and Fortran.
Motorola	Fuzzy-logic kernel for microcontrollers	Fuzzy processing kernels for 68HC05 and 68HC11 microcontrollers. Includes fuzzy knowledge-base generator to create code for kernel.
	Fuzzy-logic educational kit	Interactive training tool provides good introduction for understanding and using fuzzy logic. Runs under MS Windows. Includes demonstration version of Fide (from Aptronix).
Togai Infralogic	TILShell + fuzzy C development system	Complete fuzzy development system generates C code and includes debug, fuzzy-simulation, and graphical-analysis tools. Tutorial included.
	Microcontroller evaluation packages	Fuzzy development systems for Hitachi H8/300, H8/500, and HMCS400; Intel 8051; and Mitsubishi 37450.
	Microcontroller production licenses	Unlimited production license.
	FC 110	Digital fuzzy-logic processor (IC)
	FC 110 development system	Hardware and software development system for FC 110. Versions support IBM PC/AT bus, Sbus, and VMEbus.

Vol. 14, 1984, pp. 155-174.

15. K.S. Leung, W.S. Wong, and W. Lam, "Applications of A Novel Fuzzy Expert System Shell," *Expert Systems: The Int'l J. Knowledge Engineering*, Vol. 6, 1989, pp. 2-10.
16. N. Vadiee and M. Jamshidi, "A Tutorial on Fuzzy Rule-Based Expert System Models I: Mathematical Foundations," *J. Intelligent and Fuzzy Systems*, Vol. 1, No. 2, 1993, pp. 171-188.
17. S.D. Wee, R.E. Larew, and F.C. Hadipriono, "Angular Fuzzy Logic for An Expert System for Pavement Maintenance and Rehabilitation Strategy in Ohio," *Proc. Fifth Int'l. Fuzzy Systems Association World Congress '93*, Vol. 2, 1993, pp. 211-214.
18. H.J. Zimmerman, *Fuzzy-Set Theory and its Applications*, 2nd ed., Kluwer Academic Publishers, Dordrecht, 1990.
19. C. Negoita, *Expert Systems and Fuzzy Systems*, Benjamin Cummings, 1985.

Lotfi A. Zadeh is professor emeritus of electrical engineering and computer sciences at the University of California, Berkeley, where he is director of the Berkeley Initiative in Soft Computing. Until 1965, his work centered on system theory and decision analysis. Since then, his research interests have shifted to the theory of fuzzy sets and its applications to artificial intelligence, linguistics, logic, decision analysis, expert systems, and neural networks. His current research focuses on fuzzy logic and soft computing. He is an alumnus of the University of Teheran, MIT, and Columbia University, and was awarded honorary doctorates by the Paul-Sabatier University, France, and the State University of New York, Binghamton, in recognition of his development of the theory of fuzzy sets. He was also awarded honorary doctorates by the University of Dortmund, Germany (1993), and the University of Granada, Spain (1994). Zadeh has held visiting appointments at the Institute for Advanced Study in Princeton, N.J., MIT, IBM Research Laboratory, SRI International, and Stanford University. He is a fellow of the IEEE, AAAS, ACM, and AAAI, and is a member of the National Academy of Engineering and the Russian Academy of Natural Sciences. He has been the recipient of numerous honors, including the IEEE Education Medal (1973), the IEEE centennial medal (1984), the Honda Foundation's Honda Prize (1989), the Certificate of Commendation for AI Special Contributions Award from the International Foundation for Artificial Intelligence (1992), the IEEE Hamming Medal (1992), the Grigore Moisid Prize (1993), and the Rufus Oldenburger Medal from the American Society of Mechanical Engineers (1993). Lotfi Zadeh can be reached at the Computer Science Division, Dept. of EECS, University of California, Berkeley, Berkeley, CA 94720; Internet: zadeh@cs.berkeley.edu

Elkan's Reply

The Paradoxical Controversy over Fuzzy Logic

The responses to my article provide an exceptionally wide range of perspectives on the current state of research on fuzzy logic and its applications. Overall, I find that with most commentators I agree more than I disagree. I shall try here to steer a middle course between simply repeating points of agreement and narrowly countering points of disagreement.

The foundations of fuzzy logic. Some commentators take a more extreme position than I do concerning the coherence of fuzzy logic. I do not agree with Attikiouzel that "if one wishes to write a program or build a machine that will perform inference in the same way as human beings, then one must build the basic equations of probability theory into it, or face the inevitable outcome that it will not perform as required" Neither humans nor machines always require formal rigor to act successfully in the world, nor is success always guaranteed by rigor. Successful controllers and expert systems can use heuristic, shallow knowledge and therefore they can use arbitrary reasoning formalisms such as certainty factors or fuzzy logic. I also do not agree that "Proponents of fuzzy logic appear to be unaware of Cox's work and that of Jaynes and Tribus"; for evidence see the debate in a recent issue of *IEEE Transactions on Fuzzy Systems.*[1]

However, I am uncomfortable with the dogmatism evinced by many of the advocates of fuzzy logic or some of its many variants. For example, Dubois, Prade, and Smets say that I fail "to understand the important distinction between ... properties whose satisfaction is a matter of degree" and uncertainty "induced by incomplete states of knowledge." Later they write that the AI community has forgotten this distinction. It appears to me that the AI community has not forgotten this very binary distinction, but rather has implicitly rejected the claim that it is a uniquely important distinction. A particular concern that I have is whether the distinction is really well defined. On the one hand, there may be multiple types of imprecision and vagueness. Is the domain-independent imprecision involved in "around 1.80m" the same as the human-specific imprecision involved in "tall"? On the other hand, it may be possible to model some types of imprecision probabilistically. For example, the degree of truth of the assertion "1.80m is tall" might be modeled as the probability that an individual with height 1.80m would be labeled as tall given incomplete knowledge, that is, given no other information on the individual.

Overall, I am wary of the enterprise of even making an attempt to classify the types of uncertainty. A complete and consistent analysis of all the many varieties of uncertainty involved in human thinking and revealed in human language is a philosophical goal that we should not expect to achieve soon. Moreover, this aspiration is a variant of the quest for formal rigor criticized above as neither necessary nor sufficient for engineering success. As Freksa points out, it is always the case that "the represented real world and its representation are *formally* incommensurable." Therefore, however ideal the logics that one has at hand, knowledge engineering is always a tentative activity that can never succeed completely.

More varieties of uncertainty may well exist in the case of shallow knowledge than in the case of deep knowledge, because shallow knowledge is intrinsically domain-specific and of restricted generality. As Garcia points out, the reasoning in my watermelon example relies on important background knowledge that is not expressed in terms of rules. But it is not a fair reply to the example to call for this implicit background knowledge to be made explicit. The deep knowledge that underlies a given fragment of shallow knowledge may often be impossible or too expensive to make explicit. It is precisely then that the deep knowledge becomes implicit background knowledge that must be used tacitly in tuning the allowed interactions between the items of explicit shallow knowledge. To quote Garcia, "The dogma of generality versus efficiency strikes again, and knowledge engineering and machine learning are not exempted."

Fuzzy logic in expert systems. Only three of the responses give references in an attempt to dispute the claim that there are very few deployed expert systems that actually use fuzzy logic as their principal formalism for reasoning about uncertainty. Moreover, most of the references given actually support this claim.

Before I discuss these references one by one, it is worth emphasizing that I use the term "expert system" to designate a reasoning system that applies a large base of explicit knowledge to perform a task requiring complex inference, such as diagnosis, scheduling, or design. A fuzzy controller is a knowledge-based system of a different nature. If a fuzzy controller is called an expert system, this blurs some important distinctions. As Zadeh writes, "what differentiates applications to control from applications to [general] knowledge-based systems is that in control the main problem which has to be addressed is that of imprecision. By contrast, in the case of knowledge-based systems, one has to come to grips with both imprecision and uncertainty."

As I discussed in my paper, another important difference is that most controllers do not have to remember and reason about the history of the portion of the outside world that they deal with. Most fuzzy controllers have no internal state, while expert systems retain considerable state information.

Dubois, Prade, and Smets give five references, the latest of which is five years old. The Cadiag work of Adlassnig and his colleagues is indeed impressive.[2] However, it is especially difficult to deploy medical expert systems in the real world, in comparison, say, to applications in manufacturing. Both the cited paper and more recent papers on Cadiag-2[3,4] state only that Cadiag-2 systems are undergoing clinical trials.

Similarly, the paper on Taiger[5] does not claim that the system has been deployed, and I could not find any further papers on this system. The cited paper on RUM[6] states it is a "development environment," and the only published application built using it is described as a "prototype."[7] Finally, OPAL[8] is described in the cited paper as "under development," and Milord[9] is said to be a "shell." More recent versions of Milord use finite multiple-valued logics rather than fuzzy logic.[10]

Vadiee and Jamshidi say that "The Nikkei average has reportedly gone consistently higher using fuzzy logic." This statement is difficult to understand, let alone to believe; the only citation is to the authors' own unpublished course notes. The other application they mention is a system for choosing oil recovery methods. According to the journal paper on this system it uses the Clips shell, which is not founded on fuzzy logic.[11]

Zadeh gives three examples of expert systems using fuzzy logic as their primary formalism for reasoning about uncertainty: Cadiag-2 again, FRIL,[12] and a system for securities trading with no citation. Recent papers indicate that FRIL is a "programming language"[13] and that the trading system has only been "tested."[14] Zadeh also cites papers on systems for acupuncture diagnosis and pavement maintenance from the 1993 International Fuzzy Systems Association World Congress, but I do not have access to these papers.

The theorem. Except for Klir and Yuan, no commentators dispute the mathematical validity of the theorem given in my paper, but several commentators disagree with the assumptions made in its statement. Dubois, Prade, and Smets say it relies "at best on a logical equivalence the rationale of which is far from natural in the scope of fuzzy logic." In my opinion, the opposite is true. The equivalence between $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ is a natural one to use (perhaps inadvertently) in compiling a knowledge base of fuzzy logic sentences, and compilation into single-level rules "to simplify and speed computation" is mentioned by several commentators, Berenji in particular.

As Garcia and other commentators point out, the theorem can also be proved by considering much simpler equivalences such as $A \wedge \neg A = \neg(A \vee \neg A)$ or $A \wedge \neg A = B \wedge \neg B$. The reason the proof given uses a more complicated equivalence is that, as just mentioned, it is more natural in some intuitive sense. Intuitively speaking, in $A \wedge \neg A = B \wedge \neg B$ the two sides are irrelevant to each other, and $A \wedge \neg A = \neg(A \vee \neg A)$ is obviously similar to the law of excluded middle.

The phrase "obviously similar" in the statement above is vague. One interpretation of the theorem is that if we reject the law of excluded middle, then we must also reject many other equivalences that are not obviously similar to this law, but that are nevertheless interchangeable with the law using only the first three postulates of Definition 1. When Yager gives a derivation of the law of excluded middle from $t(\neg(A \wedge \neg B)) = t(B \vee (\neg A \wedge \neg B))$, this is an alternative statement of the theorem, not a demonstration that the theorem is absurd.

Overall, I am saddened by the hostility visible in the comments by Yager and by Klir and Yuan. I will refrain from responding line by line to their remarks on the different versions of my theorem and its proof. It is quite usual in the history of mathematics for a theorem that attracts interest to be restated and reinterpreted over time, and for similarities with previous results to be noticed later. For a similar but friendly exegesis of the development of the statement and proof of a far deeper and more important theorem the reader can consult *Proofs and Refutations* by Imre Lakatos.[15]

The theorem is technically correct as stated and proved both here and in my AAAI '93 paper. Klir and Yuan say that either the statement or the proof of the theorem is incorrect, because the "proof depends on eight logical equivalencies, only one of which is included in the statement." This claim is based on a misreading of the statement of the theorem, where the condition "if $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ are logically equivalent" must be understood as asking for a schema of logical equivalences, in which A and B may be replaced by any assertions, including assertions of the form $\neg C$.

The success of fuzzy control. Perhaps the most important contention of my paper is that the success of fuzzy controllers has little to do with the theory of fuzzy logic or fuzzy sets. Several commentators confirm this. For example, Klir and Yuan say that "most of the simple fuzzy controllers on the market ... are not explicitly based on fuzzy logic." Dubois, Prade, and Smets write that "Takagi and Sugeno have proposed an interpolation mechanism ... this kind of 'inference' (which is widely used in fuzzy control) has nothing to do with uncertainty handling," and Pelletier writes that "those areas of fuzzy logic that get criticized are simply not employed in the control arena."

It is a general property of systems that use only shallow knowledge that numerical uncertainty values can be tuned, if necessary, to overcome arbitrariness in the operators used for combining uncertainty values. Alternatively, within reason, the operators can be adjusted to match given numerical values. As Chandrasekaran reminds us concerning Mycin, a system based on shallow medical knowledge: "The fine structure of uncertainty didn't really matter." Several commentators support my specific contention that this property is one reason for the success of heuristic controllers using fuzzy logic. For example, Wang, Tan, and Tan write that "...numerous forms of fuzzy operations ... were created to cater to the domain-specific needs."

I do not agree with Ruspini that the term "paradox" should only be used to mean "logical self-contradiction," so I believe that it is fair to call the lack of connection in fuzzy systems between theory and practice an apparent paradox. All paradoxes have the property that once resolved, they no longer appear paradoxical. To paraphrase a statement by Türksen, there are no paradoxes, only limited or partial understanding. The paradox that fuzzy controllers have had real industrial success, while fuzzy logic itself is still under attack mathematically, is resolved by understanding the distinction between a scientific experiment designed to confirm or disconfirm a theory and an engineering application of the theory. Fuzzy controllers are applications, not experiments that could validate theoretical claims about fuzzy logic. On this point I agree with Mamdani: "There is a common misconception that models are created and then applied and the success then legitimizes a model."

Overall, the response by Mamdani is particularly trenchant and thought-provoking. Where we disagree, I think the cause is a misunderstanding. I do not argue that fuzzy control "is not worthy of industrial consideration because of its lack of complex form and structural sophistication." Rather, I argue that this simplicity is vital to the industrial success of the current generation of fuzzy controllers, but that fuzzy controllers for more complex applications will run into the same problems of complexity that other knowledge-based systems do today. It is the case that the "philosophical deficiencies of fuzzy logic" do *something* "to argue against the adoption of fuzzy logic control": These deficiencies are what makes scaling-up difficult.

Many research teams are actively working on scaling-up fuzzy controllers. A common feature of the research prototypes developed by these teams is the use of ideas for organizing large intelligent systems first proposed by mainstream AI researchers. For example, the SRI autonomous robot mentioned by Berenji uses "several deliberation levels to determine the relevance level of each control rule ...; to identify current goals and their state of achievement; to activate control rules according to the current context; and to blend their control recommendations." The main novelty here compared to classical robot architectures is the idea of interpolating smoothly between different suggested actions — but this idea is also found in other AI work, such as that of Brooks.[16]

The ability to interpolate between the conclusions of several rules is an important advantage of fuzzy control methodologies. As Yager writes, "the fact that most fuzzy controllers are built with a small number of rules should be seen as one of the powers of this technology," and as Berenji writes, "Fuzzy sets provide for a general yet compact characterization of system state that requires fewer rules." However, interpolation is a purely local operation, where the conclusions of a few rules describing responses to nearby input parameter configurations are blended. It is therefore difficult to see how interpolation could reduce the amount of knowledge needed to capture a complex, multidimensional input/output mapping by more than one order of magnitude compared to other approaches.

Klir and Yuan write that "... fuzzy controllers of this kind [that do interpolation] are universal approximators." This fact is true, but less significant than it may appear at first sight. Given suitable smoothness constraints, many mathematical formalisms can be used as universal approximators of multidimensional input/output mappings. For example, any continuous function can be approximated to any desired degree of accuracy by a polynomial of sufficiently high order. Neural networks with hidden layers are also universal approximators.[17] The important question is how complex an approximation must be allowed to be to achieve a given level of precision. As recognized by Kosko and Isaka,[18] the number of rules required by a fuzzy controller — which is the number of patches used to approximate its control surface — grows exponentially with the dimensionality of the controller and the level of precision demanded. From a formal point of view, fuzzy controllers thus do not enjoy a clear advantage over other formalisms for approximating smooth functions. Of course they are still pragmatically very useful.

References

1. J.C. Bezdek, ed., *IEEE Trans. Fuzzy Systems*, special issue on fuzziness versus probability, Vol. 2 No. 1, Feb. 1994.
2. K.P. Adlassnig and G. Kolarz, "Cadiag-2: Computer-Assisted Medical Diagnosis Using Fuzzy Subsets," in *Approximate Reasoning in Decision Analysis*, North-Holland, Amsterdam, 1982, pp. 219–247.
3. K.P. Adlassnig et al., "Approach to a Hospital-Based Application of a Medical Expert System," *Medical Informatics*, Vol. 11, No. 3, July-Sept. 1986, pp. 205–223.
4. K.P. Adlassnig, H. Leitich, and G. Kolarz, "On the Applicability of Diagnostic Criteria for the Diagnosis of Rheumatoid Arthritis in an Expert System," *Expert Systems with Applications*, Vol. 6, No. 4, Oct.-Dec. 1993, pp. 441–448.
5. H. Farreny, H. Prade, and E. Wyss, "Approximate Reasoning in a Rule-Based Expert System Using Possibility Theory: A Case Study," *Proc. 10th World Computer Congress* (IFIP), North-Holland, Amsterdam, 1986, pp. 407–413.
6. P.P. Bonissone, S.S. Gans, and K.S. Decker, "RUM: A Layered Architecture for Reasoning with Uncertainty," *Proc. 10th Int'l Joint Conf. Artificial Intelligence* (IJCAI), Morgan Kaufmann, San Francisco, Calif., 1987, pp. 891–898.
7. P.P. Bonissone and S. Dutta, "MARS: A Mergers and Acquisitions Reasoning System," *Computer Science in Economics and Management*, Vol. 3, No. 3, 1990, pp. 239–268.
8. E. Bensana, G. Bel, and D. Dubois, "OPAL: A Multi-Knowledge-Based System for Industrial Job-Shop Scheduling," *Int'l J. Production Research*, Vol. 26, No. 5, May 1988, pp. 795–819.
9. L. Godo et al., "Milord: The Architecture and the Management of Linguistically Expressed Uncertainty," *Int'l J. of Intelligent Systems*, Vol. 4, No. 4, 1989, pp. 471–501.
10. C. Sierra and L. Godo, "Modularity, Uncertainty, and Reflection in Milord II," *Proc. 1992 IEEE Int'l Conf. Systems, Man, and Cybernetics*, IEEE Computer Society Press, Los Alamitos, Calif., Vol. 1, 1992, pp. 255–260.
11. W.J. Parkinson et al., "Using an Expert System to Explore Enhanced Oil Recovery Methods," *Computers and Electrical Engineering*, Vol. 20, No. 2, March 1994, pp. 181–197.
12. J.F. Baldwin and S.Q. Zhou, "A Fuzzy Relational Inference Language," *Fuzzy Sets and Systems*, Vol. 14, No. 2, Nov. 1984, pp. 155–174.
13. J.F. Baldwin and T.P. Martin, "Fast Operations on Fuzzy Sets in the Abstract FRIL Machine," *Proc. IEEE Int'l Conf. Fuzzy Systems*, IEEE Press, Piscataway, N.J., 1992, pp. 803–810.
14. H. Tanaka, "Fuzzy Modeling in Operations Research," *Japanese J. Fuzzy Theory and Systems*, Vol. 4, No. 1, 1992, pp. 75–83.
15. Imre Lakatos, *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge Univ. Press, New York, 1976.
16. R.A. Brooks, "Intelligence Without Representation," *Artificial Intelligence*, Vol. 47, No. 1–3, Jan. 1991, pp. 139–159.
17. H. White, "Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings," *Neural Networks*, Vol. 3, No. 5, 1990, pp. 535–549.
18. B. Kosko and S. Isaka, "Fuzzy Logic," *Scientific American*, Vol. 269, No. 1, July 1993, pp. 76–83.

On Elkan's Theorems: Clarifying Their Meaning via Simple Proofs

Radim Bělohlávek,[1,†] George J. Klir[2,*]
[1]*Department of Computer Science, Palacký University, Tomkova 40, CZ-779 00, Olomouc, Czech Republic*
[2]*Department of Systems Science and Industrial Engineering, Binghamton University—SUNY, Binghamton, NY 13902-6000, USA*

This article deals with the claims that "a standard version of fuzzy logic collapses mathematically to two-valued logic" made by Charles Elkan in two papers [Proc 11th National Conf on AI, Menlo Park, CA: AAAI Press, 1993, pp 698–703; IEEE Expert 1994;9:3–8]. Although Elkan's effort to trivialize fuzzy logic has been questioned by numerous authors, our aim is to examine in detail his formal arguments and make some new observations. We present alternative, considerably simpler proofs of Elkan's theorems and use these proofs to argue that Elkan's claims are unwarranted. © 2007 Wiley Periodicals, Inc.

1. INTRODUCTION

In 1993, Charles Elkan presented a paper at the 11th National Conference on Artificial Intelligence, which was also published in the conference proceedings.[1] In the paper, Elkan attempted to show that "as a formal system, a standard version of fuzzy logic collapses mathematically to two-valued logic." For this purpose, he employed one theorem, to which we refer in this article as *Elkan's first theorem*. One year later, a debate regarding a revised version of Elkan's paper[2] was organized in *IEEE Expert* (August 1994), in which 15 invited responses to the paper (written by 22 authors) were published. In these responses, various misconceptions and other shortcomings of the paper were identified. Later, some of these shortcomings were examined more thoroughly in additional papers.[3–8] The theorem in Elkan's revised paper, to which we refer as *Elkan's second theorem*, is different from the theorem in his original paper. In both cases, proofs of the theorems are quite long. Because it is common to take the length of a proof as a measure of profundity of the proven theorem, Elkan's theorems may look on the surface to be quite profound. The purpose of this article is to demonstrate the contrary. This is

*Author to whom all correspondence should be addressed: e-mail: gklir@binghamton.edu.
†e-mail: radim.belohlavek@upol.cz.

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 22, 203–207 (2007)
 Published online in Wiley InterScience (www.interscience.wiley.com). • DOI 10.1002/int.20194

accomplished by (1) presenting short proofs of both theorems, and (2) by using these proofs to show that assumptions upon which the theorems are based define formal systems that are not capable of dealing with fuzziness.

The article is organized as follows. Elkan's theorems are presented in Section 2. This is followed in Section 3 by presenting simple proofs of both theorems. Based on the insight from these proofs, we then argue that the crucial assumptions employed in these theorems assert in each case that the formal system of concern is, in fact, a system of "nonfuzzy logic."

2. ELKAN'S THEOREMS

In his first theorem, Elkan considers a logical system with connectives of conjunction ($\wedge$), disjunction ($\vee$), and negation ($\neg$). A truth degree from $[0,1]$ is assigned to each formula A (Elkan uses "assertion" instead of "formula"); this truth degree is denoted by $t(A)$. The properties of the system that Elkan requires are summarized in the following definition[1]:

Definition 1. *Let A and B be arbitrary assertions. Then*

- $t(A \wedge B) = \min\{t(A), t(B)\}$
- $t(A \vee B) = \max\{t(A), t(B)\}$
- $t(\neg A) = 1 - t(A)$
- $t(A) = t(B)$ *if A and B are logically equivalent*

In addition to this, Elkan says,[1] "In the last case of this definition, let 'logically equivalent' mean equivalent according to the rules of classical two-valued propositional calculus." Then, Elkan asserts and proves for this system the following theorem.

Theorem 1 (Elkan[1]). *For any two formulas A and B, either* $t(B) = t(A)$ *or* $t(B) = 1 - t(A)$.

In his second theorem, Elkan considers again a system with properties given in Definition 1, but he explains the meaning of the term "logically equivalent" in Definition 1 very differently. He says,[2] "Depending on how the phrase 'logically equivalent' is understood, Definition 1 yields different formal systems. A fuzzy logic system is intended to allow an indefinite variety of numerical truth values. However, for many notions of logical equivalence, only two different truth values are possible given the postulates of Definition 1." Then, Elkan asserts and proves the following theorem.

Theorem 2 (Elkan[2]). *Given the formal system of Definition 1, if* $\neg(A \wedge \neg B)$ *and* $B \vee (\neg A \wedge \neg B)$ *are logically equivalent, then for any two formulas A and B, either* $t(B) = t(A)$ *or* $t(B) = 1 - t(A)$.

3. SIMPLE PROOFS OF ELKAN'S THEOREMS AND ASSOCIATED REMARKS

In this section, we present simple proofs of both Elkan's theorems and make remarks regarding the meaning of assumptions upon which the theorems are based.

Simple proof of Theorem 1. Observe that $A \wedge \neg A$ is logically equivalent to $B \wedge \neg B$ (in classical two-valued calculus). By direct application of the assumptions, $\min(t(A), 1 - t(A)) = t(A \wedge \neg A) = t(B \wedge \neg B) = \min(t(B), 1 - t(B))$ for A and B. Because $\min(t(\theta), 1 - t(\theta))$ equals $t(\theta)$ or $1 - t(\theta)$, for any θ, we conclude $t(B) = t(A)$ or $t(B) = 1 - t(A)$. ■

In the following remark we argue that the last assumption of Definition 1 is unnatural for fuzzy logic and, in fact, makes Elkan's logical system a "nonfuzzy logic."

Remark 1 (on assumptions of Theorem 1). Note first that assumption $t(A \vee B) = \max\{t(A), t(B)\}$ of Definition 1 was not used in our proof, and, therefore, it can be omitted from the first version of Elkan's theorem. As to the rest of the assumptions, both $t(A \wedge B) = \min\{t(A), t(B)\}$ and $t(\neg A) = 1 - t(A)$ are quite reasonable and, in fact, are often used in applications of fuzzy logic. Let us now concentrate on the last assumption, that is, on

$$t(A) = t(B) \text{ if } A \text{ and } B \text{ are logically equivalent}$$

In our proof, we used a particular instance of this assumption. Namely, we took $A \wedge \neg A$ and $B \wedge \neg B$ as a pair of logically equivalent formulas in our proof. It is interesting to use these formulas to show that the last assumption makes Elkan's system defined by Definition 1, in a sense, a system of "nonfuzzy logic."

We argue as follows: $t(\theta \wedge \neg\theta)$ gives us in fuzzy logic nontrivial and useful information about formula θ. To see this, assume that θ denotes the assertion "x is a red apple." Observe that $t(\theta \wedge \neg\theta)$ ranges between 0 and 0.5. If $t(\theta \wedge \neg\theta) = 0$, then we have $t(\theta) = 0$ or $t(\neg\theta) = 0$, that is, $t(\theta) = 1$. That is, if $t(\theta \wedge \neg\theta) = 0$, then either θ is completely false or θ is completely true. On the other hand, if $t(\theta \wedge \neg\theta) = 0.5$, then we have $t(\theta) = 0.5$, that is, θ is completely a borderline proposition. In general, the closer $t(\theta \wedge \neg\theta)$ is to 0.5, the more borderline a case of being a red apple θ describes; the closer $t(\theta \wedge \neg\theta)$ is to 0, the more clear-cut a case θ describes. That is, $t(\theta \wedge \neg\theta)$ can be taken as a measure to which θ describes a borderline case. Now, the assumption that $t(A \wedge \neg A) = t(B \wedge \neg B)$ for every pair of propositions A and B says that the degree to which A describes a borderline case is the same as the degree to which B describes a borderline case. Needless to say, such an assumption is absurd *because* the aim of fuzzy logic is just the opposite. Namely, the principal aim of fuzzy logic is to enable us to describe a whole spectrum of cases, ranging from completely borderline cases to completely clear-cut ones. Using an analogy with probability theory, the last assumption is analogous to assuming that a probability assignment is restricted to 0 and 1 only. This would be saying right in the beginning that a probability calculus is, in fact, not capable of dealing with randomness. In the same way, to accept the last assumption of Definition 1 is to say right in the beginning that our fuzzy logic is, in fact, not capable of dealing with fuzziness.

Simple proof of Theorem 2. Denote $a = t(A)$ and $b = t(B)$. Then we have

$$\max(b, 1 - a) \le \max(b, 1 - b) \tag{1}$$

Indeed, we have $\max(b,1-a) = 1 - \min(a, 1-b) = t(\neg(A \wedge \neg B)) = t(B \vee (\neg A \wedge \neg B)) = \max(b, \min(1-a, 1-b)) = \min(\max(b, 1-a), \max(b, 1-b))$, which is clearly equivalent to Inequality (1) because $x \le y$ iff $x = \min(x, y)$. Now, Inequality (1) is equivalent to $1 - a \le \max(b, 1-b)$. Without loss of generality, we can assume that $a \le 1 - a$ (otherwise take $\neg A$ instead of A). Then we have $\max(a, 1-a) = 1 - a \le \max(b, 1-b)$ and, by symmetry, also $\max(b, 1-b) \le \max(a, 1-a)$. Then $b = a$ or $b = 1 - a$ immediately follows. ■

As in the case of Theorem 1, we argue in the following remark that the crucial assumption of Theorem 2 is unnatural from the point of view of fuzzy logic.

Remark 2 (on assumptions of Theorem 2). Unlike Theorem 1, Theorem 2 uses a particular scheme of equivalent formulas as one of its assumptions. Nevertheless, it can be shown in a manner similar to that in Remark 1 that this assumption is unnatural and makes Elkan's system a "nonfuzzy logic." Namely, following our proof of Theorem 2, we arrived at $\max(a, 1-a) = \max(b, 1-b)$ for $a = t(A)$ and $b = t(B)$. Because $\max(a, 1-a) = \max(b, 1-b)$ is equivalent to $\min(a, 1-a) = \min(b, 1-b)$, the assumption of equivalency of $\neg(A \wedge \neg B)$ and $B \vee (\neg A \wedge \neg B)$ yields $t(A \wedge \neg A) = t(B \wedge \neg B)$. Now, the argument in Remark 1 showing that the logical system in Elkan's first theorem is, in fact, a system of "nonfuzzy logic" applies to the second Elkan's theorem as well.

4. CONCLUSIONS

The motivation for introducing and developing fuzzy logic has been its capability to capture borderline cases, in which propositions are not required to be true or false, but are allowed to have intermediate truth degrees. Although both Elkan's theorems purportedly deal with a system of fuzzy logic (Definition 1), they are, in fact, based on assumptions that exclude all borderline cases. This fact was already recognized by some authors who responded to Elkan's papers. The intent of the simple proofs of Elkan's theorems and their analysis presented here is to make this fact more transparent.

Acknowledgments

R. Bělohlávek acknowledges support by grant MSM 6198959214.

References

1. Elkan C. The paradoxical success of fuzzy logic. In: Proc 11th National Conf on Artificial Intelligence. Menlo Park, CA: AAAI Press; 1993. pp 698–703.
2. Elkan C. The paradoxical success of fuzzy logic. IEEE Expert 1994;9:3–8.
3. Dubois D, Prade H. Can we enforce full compositionality in uncertainty calculi? In: Proc 12th National Conf on Artificial Intelligence, Seattle, Washington; 1994. pp 149–154.
4. Nguyen HT, Kosheleva OM, Kreinovich V. Is the success of fuzzy logic really paradoxical? Toward the actual logic behind expert systems. Int J Intell Syst 1996;11:295–326.
5. Pacheco R, Martins A, Kandel A. On the power of fuzzy logic. Int J Intell Syst 1996;11: 779–789.
6. Novák V. Paradigm, formal properties and limits of fuzzy logic. Int J Gen Syst 1996;24:1–37.
7. Trillas E, Alsina C. Elkan's theoretical argument, reconsidered. Int J Approx Reason 2001;26:145–152.
8. Trillas E, Alsina C. Standard theories of fuzzy sets with the law $(\mu \wedge \sigma')' = \sigma \vee (\mu' \wedge \sigma')$. Int J Approx Reason 2004;37:87–92.