RECORD, Volume 26, No. 1*

Las Vegas Spring Meeting
May 22–24, 2000

Session 8PD
Fuzzy Logic

Track: Computer Science

Moderator: TIMOTHY P. SWANKEY
Panelists: RICHARD A. DERRIG†
FRED A. WATKINS‡
Recorder: TIMOTHY P. SWANKEY

Summary: The panelists cover:
- *The basic concepts underlying fuzzy logic*
- *Current applications to risk classification and pricing*
- *Future possible uses of fuzzy logic*

Dr. Fred A. Watkins: I hope that before I have finished, I will have managed to convey to you something about fuzzy logic. What it is, and then maybe, what it is not, as well. At the end, perhaps you will be in a position to decide if fuzzy logic is something that you might apply in your daily business.

What is fuzzy logic? Well, let me jump back before I answer that and ask, "What is logic?" I will answer my own question, by saying logic is the business of deciding what follows from what. That sounds straightforward, but it is important to point out that with respect to logic, the "what" that I mentioned is understood implicitly to be well defined.

There is an old example that everybody sees in logic class, if you take one of these in school. It goes like this. It's called modus ponens, which says that all men are mortals, Socrates is a man, therefore, Socrates is mortal. You have an implication, *a* implies *b*, you have a thing *a*, and you know that from *a* implies *b* and *a*, then you get *b*. Okay, but the deal is they are implying that Socrates is a man, implicit in there is the knowledge, the certainty that you know what a man is. Therefore, you can answer yes or no, unequivocally, absolutely, now and for all time, always was, always will be, no ifs, no ands, no buts, no how, no way, Socrates is a man, and that is true. That is the nature of logic.

Note: The charts referred to in the text can be found at the end of the document.

†Dr. Derrig, not a member of the sponsoring organizations, is Senior Vice President with Automobile Insurers Bureau of Massachusetts in Boston, MA.

‡Dr. Watkins, not a member of the sponsoring organizations, is President of Hyperlogic Corporation in Escondido, CA.

It is important to keep that in mind with respect to logic, truth is absolute. No ifs, no ands, no buts, no how, no way. Well, all right, that's how it is in logic as passed down to us from Aristotle on. Some people call it Boolean logic for historical reasons. It is all the same thing really.

Fuzzy logic is the new kid on the block and it is systematic, what follows from what, with respect to concepts that don't admit a simple yes or no decision rule. Here is one, "light rain". All right, I think everybody knows what light rain is, at least by experience, but I defy you to define it. I absolutely defy you so to do. You can come up with all sorts of measures, and no matter how you do it, if I'm in the audience, I will target you, I will say, "Well, your measurements are approximate. How does that fit your scheme? Who knows how much water there will be?"

All kinds of questions can arise, but nonetheless, we know that when water is falling out of the sky onto the ground, we have rain. We just don't know what light rain is, exactly. Then what, if I say moreover, that when we have rain, then there is breeze. If I can couple that with the previous, where there is light rain outside, then I think it is safe to conclude there is some breeze out there. I don't want to quantify it exactly, because I don't have the data, I don't have the measurements. I don't have a lot of things that would be required to give that a proper, absolute answer. Otherwise, I could probably say, it is a bit breezy outside and in that regard, it would be true to some degree at least. Fuzzy logic is logic that deals with those kinds of questions.

Let me amplify a little bit about truth, because that is the big word that shows up in this. When we have something that is fuzzy, it is fuzzy because it is not well defined. It is not possible to say "x" is this quality, or that quality, absolutely. Let me see if there is an example from physics. "The electron is a particle." You could say that. I could ask, "Well, is that true?" You can decide for yourself whether it is, but I think a physicist will say, "Well, it is true, at least, sometimes. It depends on how you treat the electron, today."

Or you can say, "Alright, it is a wave." Same game, it depends on how many splits are in the room, or whatever. In some sense, the electron is a particle on the one hand, wave on the other, and a mixture of the two. When you are not looking at it, it's whatever it is, and then when you do look at it, it decides whether it's going to act like a particle or a wave today. You never can really say again, "The electron is a particle." You can't say that that is true, because it is not. So, then you can't say that it is false either, because it is not.

Fuzziness is an attempt to quantify that kind of truth. Shades of gray come up a lot when you hear people talking about fuzzy logic. Gray is a one-dimensional thing. We have zero, dead black. We have one, pure white. We have shades of gray, numbers in between zero and one. So, first off, the truth in a fuzzy sense is a number between zero and one. Absolute zero is absolute falsehood; absolute one is absolute truth; and somewhere in between is fuzzy.

Here are some grades of truth that I could give you just for the sake of discussion. One, absolutely true, in the Aristotelian, Boolean sense, is no ifs, no ands, and all that. Similarly, zero, absolutely false, no ifs, no ands, no buts, that's all.

From here we have the things that show up in the world, nearly truths, almost certainly truths, basking in a little bit of haze, things that are generally true. I might be guilty a little bit of misrepresentation here, because the descriptions that I'm showing sort of act like probability kind of things, with an infrequent exception.

If it is true in all these cases, and it is false in all those cases, that is one kind of fuzziness. But the kind I am trying to get across to you is the kind that is not clear.

For example, suppose you take some comment that I'm telling you today, and you write a letter to the editor of a newspaper. The editor publishes this letter. The next day you see it in the editorial page, and it is your letter, but it's been edited. The words that you see in print are not the words that you wrote. That is the question, is this your letter? Did they print your letter? If you are an Aristotelian logician, the answer has to be "no", because it is not your letter.

Did they publish your letter? And, if the words in the letter they publish are not yours, absolutely, 100%, down to punctuation, capitalization, paragraphing, and indenting, and so on, the answer has to be "no". But at the same time, I would be uncomfortable with that flat "no", because if you hadn't done something on your own, this letter would never have shown up in the paper. So, to some extent they did publish your letter.

That's the kind of thing I mean, when we talk about fuzziness. It is not clear whether they published your letter or not. You can't say "yes" or "no." Your impulse is to say "yes and no". Whenever that happens, we are talking fuzzy. Having said something about what it is, we can be a little bit more rigorous, because in the end what we are going to want to do is, do this stuff with computers.

This is how you do it in a technical sense. The domain of truth is the real numbers between zero and one inclusive, what we call a unit interval. Every truth-value sits in there. If I have two things, A and B, for which I know the truth-values, T_A and T_B, then using this scheme, I can say the truth of the conjunction $A \wedge B$, is just the smaller of the truth-values of A and of B. It is a definition. Then the dual definition, or the T of the disjunction of A and B, we call that $A \vee B$, is just the larger of the truth of A and the truth of B. Then there is "not". To define the truth of ~A, simply subtract the truth-value of A from one.

Just for fun, I will include, "implies". What is the truth of the implication, $A \Rightarrow B$? Well, there are lots of possible expressions for the implication operator. The $\min(1,1-T_{A+}T_B)$ is just one of them that I happen to like, but there are many, many, many others.

It is not the case that all the problems in the world can be handled with one or two notions. Usually, we have lots of notions, hundreds of them, thousands, maybe millions. For that, we need a way to break down large complicated expressions, usually involving parentheses, into smaller units, until finally, we can bring the little rules to bear.

Here are DeMorgan's Rules.

$A \wedge (\bigvee_k B_k) = \bigvee_k (A \wedge B_k)$ AND distributes over OR

$A \vee (\bigwedge_k B_k) = \bigwedge_k (A \vee B_k)$ OR distributes over AND

$\neg(\bigvee_k A_k) = \bigwedge_k (\neg A_k)$ NOT distributes over OR

$\neg(\bigwedge_k A_k) = \bigvee_k (\neg A_k)$ NOT distributes over AND

You have an A, a thing that you know its truth, and a bunch of B's, and you know all of their truth values, and you are trying to find out what is the truth of A and the "or" of all those B's, finitely many, infinitely many, I don't know and don't care. Then, that result is just the "or" of the A "and" ed against each of the B_k's. You know how to do that. We just did that before. So this large, infinite class of statements reduces to an infinite class of simple statements, each of which you can compute.

Now, for the mathematically trained, I really can't say "max" in the case that the number of k's are infinite, but, they'll replace that with supreme number, some other magic word, but it means basically the same thing.

Using these rules, and the definitions I gave on the previous page, you can reduce any expression involving logic to atomic terms, compute their values, and finally come up with a value for the truth of the expression.

Chart 1 is an illustration of DeMorgan's Rule. I've got the A, and here we have 3 B's, B_1, B_2, B_3. We are looking at the conjunction of A, that is to say what is common with A, and the union, the "or" of all the B's.

By inspection, you can see that the A intersected with the union of all those B's is the same as the union of all the pieces $A \wedge B_k$, for each of the k's, one, two and three.

So you can study Chart 1 at your leisure, it is just an illustration of how the De Morgan's Rules work in the simple case of this finitely many, in fact only three, sets.

The heart of the matter is "What is fuzziness?" That is what you have to recognize, when you are trying to solve a problem in the real world, and you are asking yourself the question, "Is it fuzzy?" This is the one you've got to be able to deal with, because there are a lot of problems out there that are not clear for one reason or another that aren't fuzzy. Something is properly fuzzy if, and only if, it is neither true absolutely, nor false absolutely.

There is an axiom in logic, and by logic, I mean Aristotelian logic. It is called the Law of the Excluded Middle. It says, for any proposition A, A is either true, or it is false, and that is all. There is no middle ground, hence the name. That thing has a dual. It is never the case that A is true and false simultaneously. That is called the Law of Noncontradiction. So you can't have these things where you want to say "yes and no". That is specifically disallowed by axiom. You can't have that.

A fuzzy thing is and is not simultaneously to some degree. It may not be the case that it is and is not to the same degree. But it is the case to some degree that you can say it is, and in the same breath, to some degree you can say it is not.

Let me finish off with this business of a relationship of fuzzy logic to regular logic. Fuzzy logic enjoys a reduction property, which says basically that if you apply fuzzy logic in the case where things are not fuzzy, then the conclusions you draw from fuzzy logic will coincide with the conclusions you draw from regular Aristotelian logic, and that is nice.

In some sense, then, fuzzy logic is a generalization of Boolean logic. You've got to be careful here, because fuzzy logic is defined in terms of math, and math in turn is defined in terms of logic. If you are not careful, we will have circularity. So, you have to be careful when you say fuzzy logic generalizes logic. But in this sense, it does.

If you don't worry about origins, and use it to generate results, the results you get will coincide with those of Aristotelian logic in every case where the objects you are dealing with are in fact not fuzzy. Table 1 is an example, this is the truth table, if you will, for an A and a B and the result of $A \wedge B$. So when either one of them is false, the "and" is false, and only when they are both true, is the result of the "and" true. If you identify with the zero, the false, and identify with the one, the true, then you see, you get a correspondence, and the same thing for "or", and the same thing for "not".

TABLE 1
REDUCTION PROPERTY
Given Boolean values, Fuzzy Logic reproduces
the Boolean Truth Tables for AND OR:

AND			Fuzzy AND	Fuzzy OR	OR		
F	F	F	min(0,0)=0	max(0,0)=0	F	F	F
F	T	F	min(0,1)=0	max(0,1)=1	F	T	T
T	F	F	min(1,0)=0	max(1,0)=1	T	F	T
T	T	T	min(1,1)=1	max(1,1)=1	T	T	T

In the interest of completeness, I must tell you that "max" and "min" are not the only things we can use. In fact, mathematically, you can use any t-norm for the conjunction operator, any t-norm at all. What is a t-norm? Suffice it to say, there are t-norms in the world, there is a whole bunch of them, infinitely many, and among all of them, "min" is the largest. If you apply the t-norm that you have picked out to any pair of numbers, you will get a number that's no more than the "min" of the two numbers you start with.

Similarly, if you replace the t-norm "min" with any odd t-norm, than you must replace the "max" with the corresponding t-conorm. "Max" is distinguished among all the t-conorms in that it is the smallest. These are the extreme t-norms, that is why I like them, they are special in that way, but you don't have to use them.

It turns out, if you really want to get into the mess, you can even play with the negation operator, "one minus". I mention that only for completeness. It is only of interest, I think, to the technicians, but some people get a lot of mileage out of it.

We've been talking a lot about fuzzy. It is of no use, if you can't come up with truth-values. You may know all the formal steps of the mathematics, but without knowing what the truth of a statement in the real world is, it doesn't do you much good.

So, let's attack that problem, briefly. For any A, a statement T_A is the degree to which A is true. Truth can coincide with probability.

It is important to distinguish between fuzzy and probability, because, although, actuaries are supposed to know what probability is, a lot of times you don't, because probability is a very rarified thing when you get right down to it. But nonetheless, it has certain properties and one of these properties is that a probability is called a measure. A measure is a function defined on sets. It is well defined. The sets are well defined. Everything is clear. In the model of probability, you are saying that something either happened or it didn't happen, but whatever that thing was, it is well understood, it is well defined.

You can say, "yes, it happened", or "no, it didn't happen" after the fact, and, there's no ifs, no ands, no buts, about that.

Here is Little Billy, who every afternoon, comes home from school, and looks in the refrigerator. In that refrigerator, about half the time, he finds a nice, ripe, red apple that his mother puts there, and he proceeds forthwith to devour it. That is my hypothesis. Now, today, Billy shows up with his friend, Jack. Jack beats Billy to the fridge, looks inside, slams the door shut, and says, "Billy, is there an apple in the refrigerator, yes or no?" Well, that's a probabilistic question, at least the way I've set it up for you. Jack knows whether the apple is there or not, so it either is or it isn't. The only problem is Billy doesn't know. I hinted that half the time Billy's mother puts the apple in there, so we are placed with a fifty-fifty decision, more or less. That is probability. The event in question is clear. The apple is in the fridge or it is not.

Now, the next day, Billy comes home, goes into the fridge and he pulls out this nasty, rotten, ugly, stinky thing that maybe one day was an apple. Now, I ask the question again, "Is there, or was there an apple in the refrigerator, on that day?" Some of you will say, "Yes, an apple is an apple." Then, if I confront you with this, you may want to take it back. If I say, "Well, if it is an apple, then you should eat it." Point being, it is not clear whether it was an apple. It is to some degree, but it is not to some degree, as well.

When does the sand pile become a mountain? When does the ocean become a lake? All these questions are vague, fuzzy.

One way that we deal with all this is to introduce the notion of fuzzy sets. Here's the definition: a fuzzy set is just a function. It maps the universe of discourse, "X", whatever it is you are talking about, into the unit interval. It doesn't have to be onto. Mostly, you find that it goes all the way up to one somewhere, and all the way to zero somewhere, but it doesn't have to.

The fuzzy function "one", the one that takes the value one for everything in it that is the same as X itself, can identify with the set X. It doesn't have to have much of a range. But, once we have made that definition, than I can define the operations on fuzzy sets. This leads to an algebra function. You have two functions, you combine them pointwise by the "min". The value of F_A "min" F_B, is the value at A and in X, the value of B and in X. Take the smaller of those two numbers and define that to be the value of the "min" function.

Similarly, the value can be found for union of fuzzy sets, or similarly for compliment of a fuzzy set, by using "one minus". All these things can be done pointwise. What you get back is another function that maps the universe of discourse, into the unit interval. This all is an instance of exploiting an isomorphism between logic on the one hand, and sets on another hand, and even lattice algebra on the other hand.

Maybe you did Venn diagrams in school. If you've ever done Venn diagrams, you know you draw the circles. Chart 1 is an example of one. The fact that you can do that, harks back to this isomorphism. Isomorphism means basically, "the same thing". If you look at it right, you see that everything that happens has an exact analog someplace else. A good example of that is all real numbers under addition, positive real numbers under multiplication, and the map that connects those two would be the exponential function, or going back the other way, the logarithm. Because of the way exponentials work, you find that a + b maps exactly to $(e^a)(e^b)$ and that's one-to-one. That's an isomorphism.

So, once we have these fuzzy sets, we can get truth-values from fuzzy sets. Fuzzy sets defined over a set of the universe, may be measurements of some kind. Once you take a measurement, then you apply the fuzzy set to it, and you get a truth-value. In the real world, all of the truth-values that you are going to find will come ultimately from measurement. So, the "X", the universe of discourse, in every case, will be a set of measurements that you could take.

Once you have these fuzzy sets, that is to say a way of getting truth-values, then we can say how to do inference. Inflation is either high or low or medium. If it is low, and if gold price is low, then buy some gold. So, in principle, you can imagine having a whole bunch of these functions lying around that give you truth-values. You go into the world, you take some measurements, you plug these measurements into the functions, you get some numbers, you crunch the numbers according to the rules that I gave you before, and out comes a response of some sort. In particular, it will be a truth of an implication, "Should I buy gold? Yes or no?"

Chart 2 is an illustration of what we are talking about when we partition the variable's value into descriptive sets. We partition all those that are basically the same, in the sense that I am using them today. We have small, which is largest on the left-hand side of all possible measurements. We have medium, which is largest in the middle. We have large, which is largest on the right. Canonically ordering the numbers, smallest to largest, left to right.

With this simple system, that given something, produces the "opposite", more or less. If the thing is small, than the output should be big. If the thing is middle sized, the output should be middle sized, if the thing is large, the output should be small.

There is a set of rules, and there is a response. The more rules you have, the better you can make that curve look like what you want it to look like.

So, this leads to the notion of computing with words. There are a whole bunch of adverbs that you can apply to these adjectives that we have been describing. "Very" translates to "square the function that is the fuzzy set". "Somewhat" says "take the square root of that function, after you do that, then apply it normally", and so on, and so on. I have an expression, "not very large". The "not" is the "one minus" part, and then the "very" is the square.

In general, this is how you do it. Evaluate the state of nature against your fuzzy set collections. For each rule that you have, evaluate the weight of that rule by the truth that you get from the corresponding input fuzzy set, add up all those functions and that gives you a function again.

Then, compute the centroid of the sum, and that is the response of your system. Do this once, or do this once a second, or once every millisecond or whatever, depending on your requirements.

If you do, you will end up with something like this: for any given situation, there will be the fuzzy sets, they add up to something. The centroid is the place where the something would balance, if it were made out of one piece of uniformly dense material, and that is your result, that location. The centroid is kind of nice, because, in one number it can summarize everything that is in a function.

A fuzzy decision engine, is what we call a commutative diagram (Chart 3). You start out at "X", that is where your measurements are, and you want to infer something in "Y", which is where your results are. You assume that there is some rule that gets you from X to Y. Maybe you know what it is, maybe you don't.

The claim of the fuzzy engine is that you can get from X to Y by going up to the fuzzy set in X, doing a fuzzy system inference from the fuzzy set in X to the fuzzy set in Y, and then go back from the fuzzy set in Y back down to Y. It should be the case, that whether you go across the bottom, or around the top, you get to the same result. That is why we say the diagram is commutative. It doesn't matter which way you go, you get the same result. Cognitive maps are a variation on this.

I hope this has given you a brief intro that you can use with respect to what fuzzy logic is all about. Perhaps, you could look at our company home page to see more details: www.hyperlogic.com Now, having said all that, we are going to get an opportunity to find out how to actually apply this in actuarial context from Dr. Derrig.

Dr. Richard A. Derrig: I am assuming that you've now been spiritually moved, that you are somewhere in the middle between truth and falsehood. What I would like to do is to tell you what this might have to do with insurance operations, and what this might have to do with actuarial work.

Let me start by giving you a mnemonic that I found helpful to keep in context, because fuzzy logic is not the tip of the iceberg, but it is part of a general suite of techniques, that go under the rubric of artificial intelligence. I use "GANN Rifles". It helps me to remember that the techniques that are around are genetic algorithms, neural networks, rule induction, and then fuzzy logic, and the final one is expert systems.

There are more, and I am sure Fred can tell you many more, but this is the suite that is out there in the practical world. It is very nicely explained in *Intelligent System for Finance and Business* by Suran Goonatilake and Philip Treleaven. The combinations are nicely outlined in *Intelligent Hybrid Systems* by Suran Goonatilake and Sukhdev Khebbal.

I almost never get to explain references for my panel information, but I think it might be useful. In the book category, I utilized the Goonatilake books. *Practical Applications of Fuzzy Technologies* by H.J. Zimmerman is the latest one. It's practical application of fuzzy technologies. It shows the entire range of venues for applying fuzzy concepts.

For general articles and papers, you would not want to miss *The Fuzzy Mathematics of Finance* by J. J. Buckley, and I'll explain a little bit of that when I go along.

Finally, in that category, if you are one of those people that likes to go back and read the original Einstein paper on relativity, *Fuzzy Sets, Information and Control,* by Zadeh is the beginning of all this.

Finally, in the insurance related material, you will of course want to go out and acquire all of my papers, but beyond that, there is a paper by Jean Lemaire, 1990, called *Fuzzy Insurance*, in ASTIN Bulletin. That is the best place to find an outline of all of this, rewritten in insurance terms, actuarial terms. Jean is an actuary and chairman of the insurance department at the Wharton School. Also, Khris Ostaszewski wrote a book in 1993 for the Society on applications of fuzzy set methods called *An Investigation into Possible Applications of Fuzzy Sets Methods in Actuarial Science*.

With that out of the way, let me bring you into insurance in a little more systematic way. I want to reiterate what Fred had said about comparing fuzzy logic with probability, because you will see why the difference arises in insurance.

I'll take two different kinds of applications that are common. One is clustering. Within clustering, there is pattern recognition, there are multiple cluster memberships, and there is, for us, risk classification. On the other side, among many applications, we'll look at prices, or premiums charged for insurance products. We will also be looking at income to an insurance company, and my favorite, taxes.

Fred had it right on point when he compared fuzzy logic with probability. Probability measures randomness, and it measures whether or not the event occurs. The key part of it is that randomness dissipates with further knowledge.

That's the fact in that these are all experiment based. It is either true or it is false, when it is finally resolved. What is the difference? The difference is in fuzziness, we are measuring vagueness, we are not measuring probabilities of outcome. We are measuring the extent to which the event occurs. That is the extent of truth, as Fred put it. The key part is that vagueness doesn't dissipate over time, or further knowledge. It doesn't go away.

Let's start with clustering. Pattern recognition is one area that has found a lot of applications in and from mathematics. One way you might have seen it, which I'll label "old view", is "given *n* objects, divide them into *c* subsets, clusters".

The idea is you want homogenous things within the cluster. Similarity can be based on multiple features or criteria, but each object is in one and only one cluster. This is the truth of whether or not it inhabits a cluster.

For example, in Massachusetts, we're actually in the process of trying to decide for rating purposes, which town is in which rating territory for automobile insurance. In most states, we would do this sort of in an actuarial arena. But, we are Massachusetts, so we do this in public, and try to speak in terms of empirical base formulas, and so forth. At any event, at the end of the day, every town will be rated in one and only one territory. It will be presumably associated with towns that are like that town.

That leads always to a border problem. That is if this town is over here, and that town is over there, and they are close, well, aren't they sort of like each other? The new view, or the fuzzy view is that objects can be members of one or several clusters, but with varying strengths of membership. So, one town can be 20% high, 30% middle, and the rest low.

When you look out at a population of potential insured's they don't all fall into nice rock solid risk categories. They are almost always gray. You still might have to decide whether or not to write the risk, or to accept the book of business, but in any event, the real problem is not certain.

Chart 4 happens to be a classification of automobile insurance claims in terms of degree of suspicion of fraud. So, if you go from the left to the right, the first clump was a valid claim, in the middle they were suspected of being fraudulent, but somewhere around a sufficient level of five on a zero to ten scale. The catagory on the extreme right was labeled Planned Fraud by experts reviewing these claims. That would have been the old-time or crisp classification. Every claim would have been somewhere, and you would have acted on what cluster it was in.

It turns out that there were really six features to each of these claims that had potential to identify fraudulent claims. That is what you see on the right. There is a six dimensional vector and what is at the bottom is all zeroes, and what is at the top are larger numbers, sevens and eights, and those were various suspicions of fraud based on different characteristics.

Applying fuzzy clustering to this gives you what you see as the graphic. That is, for the most part, claims will generally fall into a single category, but there are claims, and those are highlighted by having multiple dots, that share features across the different categories. Those are to be treated just the way you would think. They are treated as being not necessarily in this category or that category, if you do this mathematically or arithmetically, you would get strengths, too.

So, you can act on even fuzzy information, by something like this as a classification device. The key parts of this that make it fuzzy are multiple dimensions, some kind of assessment that would have been crisp, and then some mechanism for combining the mathematical features into fuzzy clusters.

Pricing—we all price somehow. Everybody sets premiums or deals with premiums or wants to know what the correct premium is. In my experience, where fuzzy comes into the insurance world, is in two places. It comes in enabling you to describe variables in terms of fuzzy numbers and fuzzy arithmetic computational combinations. So you can quantify things in terms of fuzzy numbers instead of ordinary numbers.

The second general area is in decision making. The key to the fuzzy part is that generally you have more information than you need to make a decision. Generally, parts of those pieces of information are conflicting, so that real people in real situations make decisions even with conflicting information. This is a way of trying to systematize that so that the decision is made. That is Fred's defuzzifier.

Eventually, you've got to get out of the fuzzy world and do something. You do it through various rules. For example: you could set linguistic rules, which are obviously fuzzy concepts, like these:

- If we are writing a *large* amount of business, but earning *low* profits, then *raise* the rates.
- If we are writing a *moderate* amount of business, and earning *moderate* profits, *do nothing*.
- If we are writing a *small* amount of business, but earning *large* profits, then *lower* the rates.

These are the kinds of things where the italicized words are obviously quantitatively oriented variables, but are not precise, and are not easily quantified. This turns into a pricing model, where you combine those variables and out comes an action. The action is whether you change prices or not. Where would you find this? Well, you might find it in a company modeling its books of business that is wide spread, perhaps even over countries, where the number of variables that affect things are really very large and have a high tendency to be conflicting.

For example: you can see that there might be a fuzzy actuary out there, who is actually an algorithm, in which case you guys have to find another job, who has an algorithm that says in one of its components, if Greenspan is going to raise the interest rates, we've got to do what?

Whatever the business questions are, they are presumably, decision models based on real business decisions that have to be made, like price changes, but there are many others. Sell the bonds, for example.

To determine a price for an insurance product, first of all, decide what's the cost. In the property-casualty world, the price is the expected cost plus an expected profit margin. I would assume that is true in your world, as well. One of the problems with that notion is that you have to forecast the cost, since these insurance products usually go into the future, and you have to figure out what a fair profit is, and include taxes in your calculations.

The solutions as we've found, at least conceptually, if not in actuality, are fuzzy trends into the future, premiums or taxes.

What is the question on trends, forecasting? We all try to figure out what the expected claim cost is for the period of coverage, or the expected payout for a book of reserves you've got. There are all sorts of things. We're always trying to estimate the future costs. The expected costs are always generated out of historical data, and then moved from this economic climate to whatever is in the future.

Why are they fuzzy? Well, they are fuzzy not because of those the reasons above, because everybody does those. They tend to be fuzzy for other reasons that you just sweep under the rug. For example: How much historical data to use? How many periods; sixty months, or seventy-six years? What kind of forecasting model should we use? If we are fitting curves, and running off curves, what kind of curves do we use? How do we test whether or not these are good ones? There are statistical measures of fit that we could use. How about a reasonability of results? There isn't a CEO I know that will listen to an actuary explain how really good the R-squares are. "You really ought to take this projection when the number makes no sense at all." I think you have seen that situation where they just say forget it. The marketing department wins, even though you've already decided that other answers are really more reasonable.

Finally, there is really an incompatibility possible between fit or mathematical purity, let's say, and reasonability. Dave Cummins and myself saw this when we were trying to establish good forecasting methods, just good, not best, and using a range or a stochastic basis, if you will, for types of models that you would use in forecasting.

We were measuring based on how accurate they were, whether they were biased, and finally, whether they were reasonable. What finally dawned on us is that forecast methods that are nearly as accurate and unbiased may not produce expected or forecasted values that are anywhere near each other.

So, if you do these kinds of optimization problems, where you end up choosing the best, and then you use that model, you may be throwing away something very valuable that is second best. Vilfredo Pareto knew that a long time ago, but the idea is that if you get into a fuzzy world, believe me, it is a different way of thinking of everything, you end up not throwing away anything. All the information that you have counts, and you are not in the business of getting rid of it.

Expected claim costs in a practical sense are really fuzzy if you set it up as a fuzzy problem, but they can be made crisp or defuzzified by using weighted averages. That is really the analog to what the solution is in the forecasting game.

On the premium side, a premium is fair in the economic sense if it equals the present value of all liabilities. That is presumably what you are charging to a customer, the present value of all costs including capital costs. Fair premiums in our world are calculated using a risk-adjusted discounted cash flow model.

What we use to decide, thus why this is a fuzzy problem, is to divide the uncertainty into two different types or levels. The first level uncertainty includes cash flow magnitudes, cash flow patterns, risk free interest rates, risk adjustments, and tax rates. If you were solving this problem, I'm sure you would take all these into consideration. You would consider them vital to deciding what the price is.

For example, you would always take into account the magnitude of the cash flows, the patterns of the cash flows. You'd worry about risk free interest rates, if you were looking for present values. You'd be looking for risk adjustments, if you wanted to make a profit in the world. Finally, my favorite, you always worry about taxes.

The second level of uncertainty covers historical data quality, development methods, trend or forecast methods, expense allocation, surplus allocations, capital markets, and insurance markets. Who worries about data quality? I mean, you use this data, but do you know, is it right? If it is the government, it is always right. If it is your company, well, you know how people enter data.

The forecast method, well, we already decided that's kind of fuzzy. There's a whole other world, called allocations. Usually, the accountants are involved in this one. If you are trying to price something, you have to allocate some expenses to this. How confident are you that the expense allocation is really good?

That's not to mention allocating capital across lines of insurance. The capital markets and the insurance markets themselves don't lend themselves to probabilistic modeling.

Dave Cummins and I looked at fuzzy premiums for the *North American Actuarial Journal* in 1997, calculating fuzzy fair premiums out of fuzzy parameters, where the parameters, risk-free rates, were fuzzy. You know, will Alan Greenspan change them or not? Is that probabilistic? It is nowhere near probabilistic.

The crisp parameters are the flow patterns in the tax rates.

In order to do all this, I said part of this was decision making and the other part is calculation and variable setting, in terms of fuzzy, so there are fuzzy numbers. There is fuzzy arithmetic that goes with it, and there are fuzzy present values, and there is fuzziness of the premium. Finally, there is fuzzy net present value.

If you remember your finance or capital budgeting kinds of things, sensible people only take positive net present value projects. Right? Well, this model is the only one that I have ever seen where you can at least set up a model where the ordinary net present value would be negative, and you would still accept the project because the centroid is really positive, but the mean is negative. That, to me, is marketing it actuarially; that is why companies actually do it.

Fuzzier prices are fuzzier than we thought. This is just sort of a general view of fuzzy numbers.

Finally, taxes. Let me just say you can set up balance sheets, and you can set up expected portfolios. So you can fuzzify all of that.

Mr. Timothy P. Swankey: How many came in here knowing something about this fuzzy stuff, networks, and so on? About half of you.

Did we help you out any, or are you still where you were when you walked in the door? We have someone interested in the game. Fred, do you want to take on the game?

Dr. Watkins: The illegal crop game was an exercise in this thing we call the fuzzy cognitive map. It actually began as an exercise in dynamic systems by Bart Kosko. It was an effort to analyze the dynamics of the situation that we have when we have basically three game players. There is the grower, we all know who they are, and then we have the informer, we all know who they are, and then we have the rip-offs, unless you are hip, you don't know who they are. They are the guys that go out in the woods, find the goods, and steal it, or hold up the farmer, or even maybe the informer sometimes. Anyway, these three players constitute the people in the game.

The idea of all this was to analyze the dynamics of these operators, these players, in day-to-day life, and make comments on the policy that we have currently.

The cognitive map ultimately is a matrix; it is a symmetric matrix, really. It has rows and columns that are the same. It all comes back to writing down what's going on. Here is a shot at what goes on out there in XYZ County, or whatever. We have the growers. When you have growers, then you have informers. The presence of growers stimulates the presence of informers. We put an arrow with a one on it, a plus number, to indicate that growers stimulate informers. If you have informers, of course, that exerts presumably backpressure on the growers.

It is supposed to be the case that if you had more informers, then the number of growers declines. Although, some places I hear, that's not quite true. The more growers there are, probably, the more rip-offs there would be. It makes sense. The more informers there are, probably, depresses the number of rip-offs.

In some sense, the informers replace the rip-offs, if you're going to be hard and cold about it. Finally, the more growers there are, the cash per sale has to drop, just by supply and demand. Of course, the more cash there is per sale, the more growers there will be, demand and supply. Also, the more cash there is, the more rip-offs there will be. The more rip-offs there are, the more cash per sale there will be.

That's because when the transactions become rarer and rarer, the dollar value of each one will probably rise. This is a stupid, but more or less, intelligible picture model of the dynamics of XYZ County agriculture. We can take this kind of data and put it in the form of a matrix. I can write down rows and columns, informers, rip-offs, growers, and cash per sale, and write down the sense of those arrows, plus one or minus one.

I have now, believe it or not, captured the dynamics of the situation in a matrix, as long as I'm willing to follow the arithmetic rules, which are very simple. I start with some initial condition, which in this case is lots of growers, or some growers at least, and a reasonable amount of cash per sale, whatever. No informers, and no rip-offs, and let it rip. I mean by that, I apply the matrix to that initial condition vector, and then I'll get a result, a vector. It will have various numbers in it. If the number is positive, I'll put in a one, if the number is negative, I'll put in a 0, and if nothing happened, I'll leave it alone. I will then take that result and push it back through the matrix, again, and do the same thing over, and over, and over, again. Until what I get out of it is a sequence of vectors that is predictable, what they call a limited cycle. It will either be one vector that always gives you the same number, or maybe two of them, or three of them, or whatever. That is, if you will, the equilibrium behavior of this model.

In the Gulf War, Dan Rather had some fellow on the TV and the fellow said the ecology of the Gulf was ruined, basically. Some period later, six months, nine months, a year, I don't know, the same two were back on the television. Dan says, "Red Adair and his people have cleaned up the place, we've got birds chirping, things are good. What have you got to say about this?" The guy says, "Well, I told you what the model said." In that same sense you must be a little sanguine about using this kind of thing. You have no authority for saying that the dynamics of the model and the dynamics of the real world are the same.

Intuitively, there is some kind of connection and there are examples where the prediction that comes from the fuzzy cognitive map and the actual opinions of experts have coincided, so there's something there. If nothing else, it is worth looking at as a way to fix ideas.

Here are the rules: We take the matrix, we apply the initial condition vectors, we reiterate through the matrix, over and over and over again. Each time we threshold the results so that the result is again a vector with zeroes and ones, only. A positive result gives you a one. A negative result gives you a zero. No change, leave it alone. In the end, you get a map of what is supposed to be the result.

For example, if it turned out that the limit cycle was a single vector with nothing but informers that would say that our policy leads to nothing but informers in the place we are talking about. All the growers are gone. All the rip-offs will be gone. It will be great, I guess.

Or it might say something else; maybe there will be nothing but rip-offs, who knows? Whatever you get is something that is intelligible to you in terms of the model. Then you can decide whether it fits reality.

Let me say one last thing about this kind of procedure. I drew my matrix of the game. I drew my little picture of the cells with the arrows. That is just my opinion. Each of you could do this too. We could all take our matrices, each of us, and combine those matrices in a natural way. Add up all the similar items, and then divide by the number of them, and we'd still end up with a matrix.

It would just have different numbers in there. It wouldn't be my numbers; it would be the consensus of everybody in the room. Then we'd do the same stuff, and then we would have the results of the Committee of Actuaries on the problem in XYZ County.

If we tried to do it in words, together, if we tried to have any kind of committee-oriented consensus, we would probably fail.

By using the matrix method, we have a mechanical procedure for providing a systematic, reasoned, and repeatable response to the varied and sundry opinions of a large group of so-called experts.

It is a useful technique. If you look on the web, you can probably find it by name, fuzzy cognitive map (FCM). It is a useful technique. I don't say that it is something that you should depend on, because it is dangerous. But, in the right hands it can be very useful.

Dr. Richard A. Derrig: What I would like to finish with, are four predictions. I think you are going to run into fuzzy things, maybe not run into, but they should be there. Somebody will run over you if you don't have them.

The first is the fuzzy controller for insurance. The idea is for the few companies yet to be interactive with the Web. The only thing you can put out there are fuzzy prices. You can't put out crisp prices.

The second is fuzzy data mining. That's not just coming, it's here already. Where you would be interested in something like that, besides the usual data mining questions, is data cleaning, making the data correct. Finding the bad data is a fuzzy operation.

The third out of four is something called fuzzy agents. Your agent may in fact be a fuzzy controller, or at least partially a fuzzy controller. This involves the interaction between the insurance company and the policyholder climate.

The last one that I would like to leave you with is that you could really make the exam structure fuzzy, because there is nothing fuzzier than the answers to those actuarial exams.

From the Floor: Among the "GANN Rifles", could you position fuzzy logic relative to neural networks, and so forth, and the AI setup tools?

Dr. Watkins: Yeah, we can. I'll take a shot at that. If you are an AI guy, then you believe in the word "symbol". To an AI guy, "symbol" is a technical term that means a list. The list has a sequence of attributes attached, and on this list one does logic in the formal sense.

If you do neural networks, you have, as opposed to AI and its highly structured organization of knowledge, a completely unstructured mess, but it is no longer done with logic, it is now done with numbers. If we partition knowledge processing algorithms according to structure or not, numeric or not, then AI is structured, not numeric.

Neural nets are unstructured numeric, and then there is fuzzy, which is structured. But, because it is done with these functions that I told you about, it is numeric.

It is possible to draw a chart with structured and unstructured along one axis, numeric and nonnumeric along the other. In structured, numeric, you have fuzzy. Structured, nonnumeric has AI. For unstructured, numeric, we have neural nets. There is nothing in the unstructured, nonnumeric category as far as I know.

Dr. Derrig: Fuzzy is really the only one on my list that's a conceptual framework, unless you want expert systems in the conceptual. It is the only one that really changes the whole world, as you look at it. The other ones are some of the tools, genetic algorithms for example, to get there.

Neural networks are tools that take advantage of function approximations. Rule induction is another tool that allows you to turn this stuff into English and decision steps.

Finally, expert systems are, you know, a bunch of people in a room that you organize. I would divide it that way. The big thing is the change of view, while the others are tools.

CHART 1

"DeMorgan" Rule Illustrated

$$A \wedge (\bigvee_k B_k) = \bigvee_k (A \wedge B_k)$$

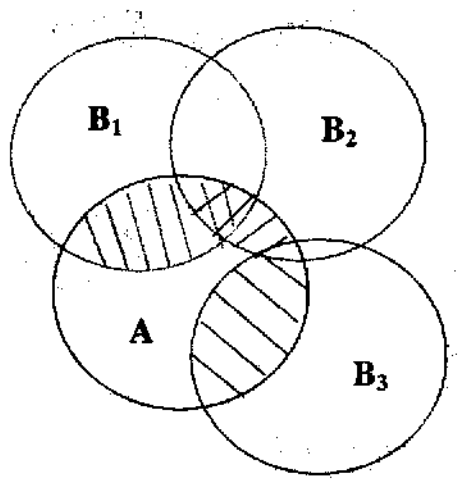
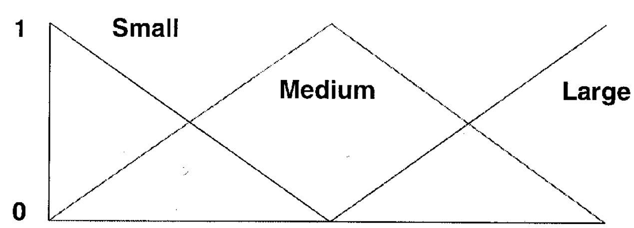


CHART 2

Resolution into Set Values

- Partition a variable's range into descriptive sets:



- A measurement is SMALL, MEDIUM, or LARGE to some degree.

CHART 3

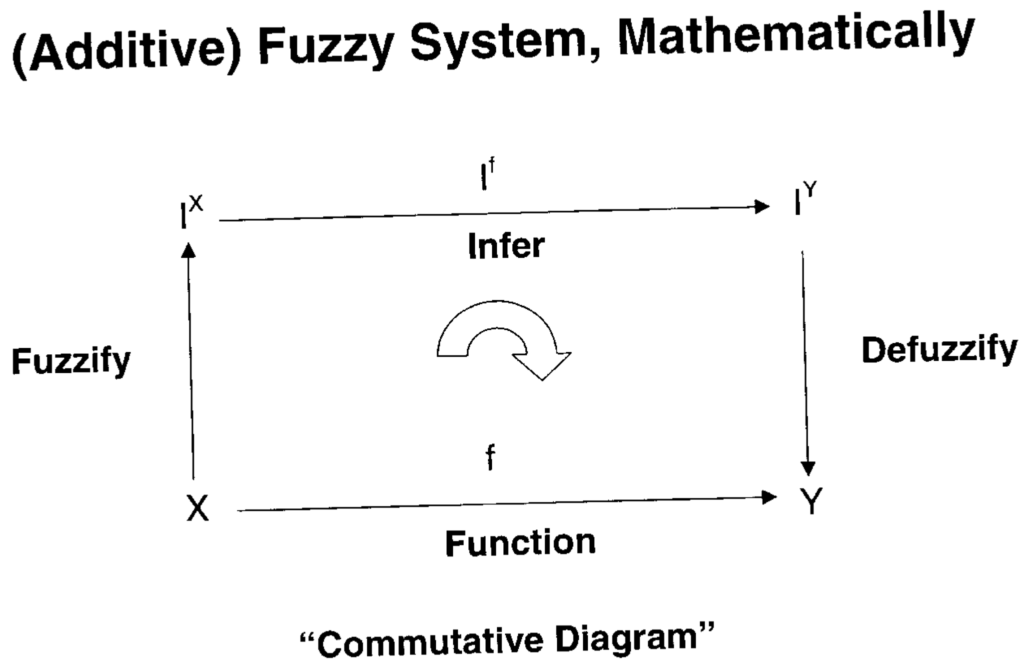
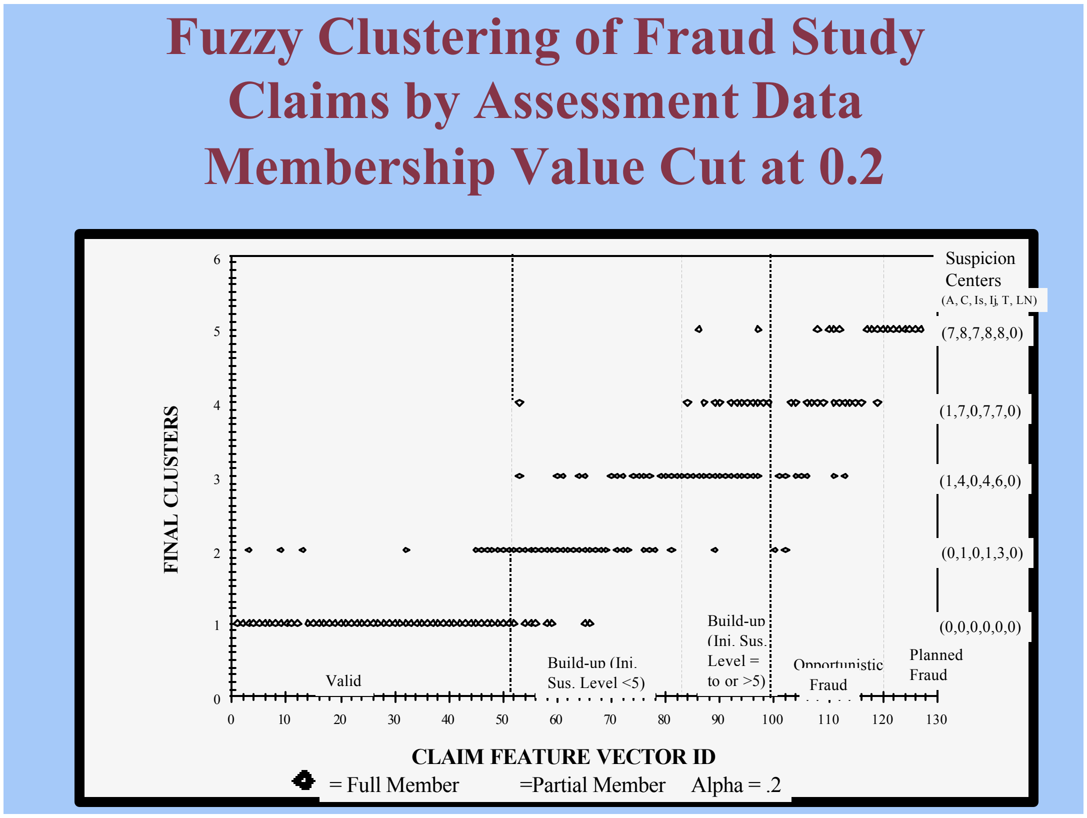

(Additive) Fuzzy System, Mathematically
I^f
I^X
I^Y
Infer
Fuzzify
Defuzzify
f
X
Y
Function
"Commutative Diagram"

CHART 4


Fuzzy Clustering of Fraud Study
Claims by Assessment Data
Membership Value Cut at 0.2
FINAL CLUSTERS
0
1
2
3
4
5
6
Suspicion
Centers
(A, C, Is, Ij, T, LN)
(7,8,7,8,8,0)
(1,7,0,7,7,0)
(1,4,0,4,6,0)
(0,1,0,1,3,0)
(0,0,0,0,0,0)
Valid
Build-up (Ini.
Sus. Level <5)
Build-up
(Ini. Sus.
Level =
to or >5)
Opportunistic
Fraud
Planned
Fraud
0
10
20
30
40
50
60
70
80
90
100
110
120
130
CLAIM FEATURE VECTOR ID
= Full Member
=Partial Member
Alpha = .2