SOCIETY OF ACTUARIES Forecasting and Futurism Section

ISSUE 9 | JULY 2014

Forecasting & Futurism

NEWSLETTER

If More Precision Is Always The Answer, Have We Forgotten The Question?

By Dave Snell

Actuaries
Risk is Opportunity.®

ISSUE 9 | JULY 2014

Forecasting & Futurism

NEWSLETTER

Published by the Forecasting and Futurism Section Council of the Society of Actuaries

This newsletter is free to section members. Current issues are available on the SOA website *(www.soa.org)*.

To join the section, SOA members and non-members can locate a membership form on the Forecasting and Futurism Section Web page at *www.soa.org/forecasting-futurism*.

EDITORIAL

If More Precision Is Always The Answer, Have We Forgotten The Question?

By Dave Snell

"*It is the mark of an instructed mind to rest satisfied with that degree of precision which the nature of the subject limits, and not to seek exactness where only an approximation of the truth is possible.*" —Aristotle, 384-322 BC

The theme of this issue is precision. We are not discussing it in the usual sense of how do we get even more precise; but in a contrary one of whether we are already too focused on more details and more decimal places. Actuaries can be too oriented towards precision—sometimes when it does not make sense. The advent of inexpensive computers has made it easier to do thousands of stochastic runs, to carry intermediate results to many decimal places, and to exhaustively analyze myriad criteria.

Yet the models seem, if anything, even more fragile than previous, simpler ones. When people don't do what is logical, or expected, the self-correction mechanisms in some models can cause precipitous market falls; and natural disasters, such as the Japanese tsunami, result in unnaturally dire consequences when we focus on the minutiae of failsafe mechanisms and ignore common (less common?) sense.

Alberto Abalo, our chairperson, starts us off with a quote from Shakespeare ("That which we call a Rose, by any other Name …") and a common sense question about our section name *"What's in a Name?"* Is it still an accurate reflection of who we are? Some members would like us to change the F&F section name to better reflect the sophisticated, advanced analytics we do. In fact, at a recent meeting with several SOA Board members, one proposal was to form a new SOA section that would embrace predictive modeling, 'Big Data', and other topics that we have been using and writing about (right here in our newsletter) for almost five years now—a few of them had no idea there was already a section, Forecasting & Futurism, that was doing this! Have we all become so enamored with the trees (especially those involving the Bayesian branching and hidden Markov models) that we have become blind to the forests (other than random forests and similar machine learning techniques)? Alberto raises some important points. If we ignore the opportunity to rebrand and explicitly put some sort of advanced analytics into our name, we may lose membership of those who want to be a part of this initiative; but if we abandon our Futurism appellation, we risk turning off (and away) those who came to us to learn about Delphi studies, behavioral economics, and other "softer" sciences that help us to step off the analytics treadmill, smell the roses, and see the bigger picture.

Geof Hileman helps us see the bigger picture with his poignant article *"Roughly Right."* Geof suggests five key practices that we all should keep in mind. I'll mention

CONTINUED ON **PAGE 4**

only one of them here: "be aware that precision implies confidence." I once worked on a valuation project for a client who insisted upon calculating and saving the seriatim valuation of millions of policies on a quarter-by-quarter basis for the next 80 years; and all the calculations were done to four decimal places. When I suggested projecting for less years, keeping less decimals, or perhaps using annual calculations to save both computing time and storage space, the suggestion was soundly rejected because the present value calculated would differ (in total) by thousands of dollars. None of the input assumptions held up after even one quarter, let alone 320 quarters, but that didn't matter because the focus was on precision, not value. Sometimes we start looking at trees and then become obsessed with the tiny twigs on the ends of the branches. We lose the ability to see, and therefore to explain the overall picture, and consequently we find ourselves displaced by the "communicators"—those quants who can speak in terms the client understands. I loved Geof's first sentence supporting his insistence on simple corroborating models: "As fascinating as you may find neural networks, genetic algorithms, or negative binomial regressions, you were hired because your client (using this term loosely) would rather not know about these things."

Next, we have another contribution from Charles Brass, our Futurist from Down Under (Australia). Charles wrote the article ***"The Past Is No More Certain Than The Future—Decision Making In The Face Of Unavoidable Uncertainty."*** He reminds us how two independent juries (one for the criminal case; one for the civil case) each came to unanimous but opposite decisions about the murders allegedly committed by O.J. Simpson. He also points out the responsibility we have as futurists: "Futurists acknowledge the power that past performance might bring to the future, but they also explicitly recognize the possibility of 'wild cards' which might change the picture completely." Remember that portion of our section name?

Which brings us to the question ***"How Do YOU Forecast?"*** wherein Doug Norris describes our F&F fourth annual contest. This time, instead of an iPad, we are offering a $500 credit in the Apple store. Informally, we are calling it an "iPrize." We know you want it. Here it is. Doug explains the rules, the scoring criteria, and oh yeah, the purpose of the contest. It's basically, to advance the actuarial profession. Wouldn't you want to be known as the winner of an SOA contest to advance the profession? Learn how to enter in Doug's contest announcement.

Learning is a major focus for us; and Jeff Heaton has contributed an article about how machines can learn. ***"An Introduction To Deep Learning"*** delves into how Google and other leaders in the machine learning area teach a neural network much faster than the former, multiple-hidden-layer approach. IBM's Watson uses Deep Learning (among an ensemble of other learning techniques), and Jeff explains the simultaneous supervised and unsupervised nature of this training that makes this methodology "deep." He also explains the neat "bag of words" algorithm that helps us deal with unstructured data. It is a simple concept that works well with unknown text from a book, or from a large text-oriented database like Wikipedia.

Unstructured data is often synonymous with Big Data; and the term is used and misused a lot. Richard Xu and his colleague, Dihui Lai, dispel some of the confusion about Big Data in their article ***"Big Data In Life Insurance—Does It Exist? If So, How Should We Handle It?"*** They address some of the ways to deal with the mounting challenges of capacity and speed as data scales up rapidly in size. Hadoop was once just the name of a toy elephant; but there is nothing toy-like about how it has been employed to handle very large datasets. Five exabytes supposedly represents all the words ever spoken by human beings; but according to IBM, the new SKA telescope initiative will generate over an exabyte of data every day.[1] How will we cope with big data? Read Richard and Dihui's article for some hints.

Hints, hunches, opinions, and collaborative ideas are the mainstay of progressive think tanks; and the SOA Delphi Study on Long-Term Care Financing Solutions had a lot of them. Ben Wolzenski and Ron Hagelman carry on a simulated dialog in their article ***"A Conversation About The Delphi Study On Long-Term Care Financing Solutions"*** to explain the six Principles generated by the Delphi study participants as well as the legislative background applying to this growing concern for our aging population.

Compared to some of the more conventional actuarial forecasting techniques, Delphi studies seem a little vague at times. The answers are often not numbers. Often they are free form text, which has to be analyzed to understand the nuances of meaning.

Like the real world, the answers to a Delphi study questionnaire are sometimes a bit "fuzzy." I'm actually an advocate of fuzzy set theory and fuzzy logic (some colleagues suggest that I might be fuzzy more than I intend to be). Fuzzy set theory appears to be a superset of the set theory we learned in school. In fact, the former set theory is now referred to as "crisp" set theory. A cool aspect of fuzzy logic is that it tends to work better (than crisp logic) with problems involving incomplete or imprecise data. Since precision is our theme this issue, we have two articles on fuzzy logic.

In my article, ***"Warm And Fuzzy … And Real!"*** I take a nonconventional approach to explaining fuzzy logic. I use only one greek character, μ (mu), which represents membership in a given fuzzy set; and I avoid almost all the fancy mathematical distribution descriptions and set theory symbols. My purpose is to try to convey the basic ideas unobscured by these artifacts of too many graduate courses in statistics. Fuzzy logic is something that we learn as children. It is not that difficult! OK, I can take basic terms only so far, and the planned Part 2 article next issue will have to bring back the Greeks; but here is a chance to warm up to the concept without as much angst.

Jeff Heaton extends the angst-free fuzzy zone by teaching how you can use fuzzy logic without even having to do any calculations yourself. In his article ***"Fuzzy Logic In R,"*** Jeff shows that everything you need for your fuzzy-logic-in-a-can experiments is included in the programming language `R`, supplemented by a "sets" package. You can get started just by giving the `R` command:

```
install.packages("sets")
```

Jeff shows how building a fuzzy logic application can be almost as simple as combining some LEGO blocks. He told me that he is planning a Part 2 as well. It will be a similar programming tutorial using the fuzzy add-on features of Python instead of R. Between the two languages, you can program the vast majority of data science applications.

That's a summary of the articles in this issue. However, let's return to the issue of the F&F section name; and the rose metaphor from Albert.

The study of "meaning-making," signs, signification and communication is called Semiotics. One of the most famous fictional books employing semiotics is Umberto Eco's ***Il nome della rosa (The Name Of The Rose)***. All of this symbolism is appealing to me; and we'd like to sponsor another contest, which I am calling The Name Of The Rose contest. The winner will be the person who submits the best suggested name for our section (to be submitted to any council member by Aug. 15, 2014, and judged by our section council), and gives a compelling argument why this is a better fit than our current F&F. In case of a tie, the winner will be randomly chosen from the top entries. If no names are deemed by the council to be better than the current one, the council reserves the right to reject all entries.

The prize will be a dozen roses for your significant other. It's not an iPrize; but it is likely to make you look really good in his or her eyes.

IF MORE PRECISION … | FROM PAGE 5

Finally, since our stated theme this issue is precision, I'd like to leave you with one more quote:

"so far as the laws of mathematics refer to reality, they are not certain. And so far as they are certain, they do not refer to reality … mathematical precision does not correspond to reality." (Albert Einstein, 1921).[2]

Enjoy the issue! ▼

ENDNOTES

1 http://highscalability.com/blog/2012/9/11/how-big-is-a-petabyte-exabyte-zettabyte-or-a-yottabyte.html

2 Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing, By Nazmul Siddique, Hojjat Adeli, p.20, John Wiley & Sons, Ltd., 2013

Dave Snell

Dave Snell, ASA, MAAA, is technology evangelist at RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *dave@actuariesandtechnology.com.*

CHAIRPERSON'S CORNER

What's In A Name?

By Alberto Abalo

"A rose by any other name would smell as sweet."

— William Shakespeare, Romeo and Juliet

Forecasting and Futurism.

Loyal readers of this newsletter will readily recognize the dual aspects of our mission and interests in that name. *Forecasting* acknowledges our commitment to introducing actuaries to new quantitative tools and methods successfully used outside of our profession to forecast or predict future events. Equally important, the term *Futurism* recognizes that numbers alone won't help us reach our goal. These pages continually affirm that, despite past assumptions regarding the infallibility of ultra-sophisticated forecasting tools, the world does not seem to feel obligated to follow our models.

Outside of our close-knit group, the perception of our section's activities is a bit murkier. To some, *futurism* conjures up visions of tin-foil hats and 1950s' science fiction movies. How many of those people associate our section with the topics that have come to define it over the past five years? Are we reaching those actuaries who would benefit from learning more about predictive modeling, artificial societies, neural networks, Delphi studies, complexity science, and genetic algorithms?

With this concern in mind, a section member recently posed a seemingly innocuous question to the council: would a name change bring greater attention and perceived relevance to the topics we discuss? Here are some responses on our LinkedIn page:

- "I think the term *futurism* embraces some of the aspects of our section that transcend numerics and that we would be limiting our scope and our responsibility to society by dropping it. …"
- "I have to be honest, before I became a Forecasting/Futurism member, I didn't understand the difference between our section and the [Actuary of the Future] section. …"
- "Our meeting presentations and webcasts introduce and endorse emerging techniques, and they are quite well received. Does our section name put folks off? Or is it still a rose by any other name? ..."

As a member of the council, I have always believed the topics this section investigates are vital to our profession. As long as we stay true to the spirit of our mission and nurture our intellectual curiosity, our section, by any name, will continue to contribute significantly. On the other hand, ignorance can only lead to irrelevance. So why not consider a rebranding? As of this writing, names we are considering include Predictive Methods, Advanced Analytics and Behavioral Methods, Predictive Analytics and Futurism. Forecasting and Futurism still has its champions too. What are your thoughts?

A final note: I invite you to be a more active member of the section. By the time you read this, I will be ending my term as Council chair. Serving on the Council has been an incredibly rewarding experience, both professionally and personally. The content and quality of this newsletter were what first inspired me to join. I was introduced to the science (or sciences) of complexity, discovered the mind-blowing nature of complex adaptive systems, and learned what an artificial society was. Through my membership in the section, I had the pleasure of speaking to experts about these topics (and the privilege of coercing them to speak at industry meetings). There is nothing more professionally fulfilling than making an impact through your own contributions. I may not have a crystal ball, but I predict this will remain the case, wherever the future leads us.

Enjoy the newsletter!

Alberto Abalo

Alberto Abalo

Alberto Abalo, FSA, CERA, MAAA, is a principal at Oliver Wyman in Atlanta, Ga. He can be reached at alberto.abalo@oliverwyman.com.

Roughly Right

By Geof Hileman

A proverb, often incorrectly attributed to Keynes, states that it is better to be roughly right than to be exactly wrong. Whether or not we've heard the concept stated so explicitly, we've all put its wisdom into practice. When my daughter asks me about the weather, I'll summarize the hourly forecast that I read online ("74 degrees with 20 percent humidity and a 10 percent chance of rain starting at 4 p.m.") as, "It's going to be perfect." She didn't need to project the path of a tropical storm—she just wanted to know whether she needed to wear a jacket to school.

Much like with weather, we are working in an era marked by tremendous amounts of data and by sufficient computing power to analyze those data. There is a huge temptation to build models that take advantage of these factors without great regard to what is really necessary to answer the question at hand. Most actuaries enjoy working with detailed data—that's what drew many of us in to the profession. It's critically important to remember, to torture another idiom, while we may love the trees, our stakeholders generally only want to see the forest. The depth of available data increases the risk that we will focus on the details of a problem rather than on the broader principles.

I'm not suggesting that there isn't a place for complex, data-rich models in the actuarial world. In fact, I believe that judicious use of emerging modeling approaches can set actuaries apart from our analytical peers from other disciplines, potentially even from within the same organizations. To that end, I am suggesting five key practices that will allow actuaries to continue using highly complex models to answer business questions without losing our audience along the way.

First, complex models should be supported and explained through the use of corroborating simple models. As fascinating as you may find neural networks, genetic algorithms, or negative binomial regressions, you were hired because your client (using this term loosely) would rather not know about these things. Corroborating models can be as simple as a graph that places a projected value in its historical context. This graph would be accompanied by an explanation of why the projected value is different from what might be suggested by extrapolating recent history. We should steer clear of explanations such as "this is what the data indicate" or "this is what our model says." These platitudes are generally indicative of shortcuts around finding the real-world cause behind changes in the data and models. Supporting models can also be simpler approaches or methods more familiar to the stakeholder that point in the general direction of a result from a more sophisticated approach.

Second, modeling approaches should be back tested and the results publicized. Comments referring to actuarial "black boxes" are rarely complimentary. However, the blackness of the box is certainly in the eye of the beholder. What is science to some will appear as voodoo to the uninitiated. In order to prevent this from becoming a barrier to our work being perceived as trustworthy, we must either train or reassure. In many cases, training our stakeholders in the ways of our models is not practical or desirable to either party. However, we can reassure others by demonstrating the historical accuracy of the same models that are producing our future forecasts. This must be done in a concise and understandable manner—lest we introduce additional black boxes—but is a critical step in gaining trust in our methodologies.

Third, be aware that precision implies confidence. This truth is often used to the advantage of marketers or attorneys who wish for their audience to believe something. For example, requested damages in lawsuits are often developed to much greater accuracy than necessary just to lead the jury to believe more fully in the arguments supporting the judgment. We must be very careful to not fall into this trap as well. While point estimates are often required (you have to book a specific dollar amount in reserve and file a specific premium), there are many cases where ranges of estimates are more appropriate. While statistical techniques can sometimes be used to generate precise confidence intervals, sometimes statistical rigor is not possible or even necessary. By discussing a range of estimates, actuaries can provide more value to their stakeholders by painting a more complete picture of the potential impacts of a decision.

Fourth, the a priori assumptions of both the actuary and the

stakeholder should be considered when building models and communicating results. On the front end, there is a temptation to dive full steam into the model building without first considering our expectations. If we are, as the Ruskin quote goes, in the business of substituting "facts for appearance and demonstrations for impressions," we must first consider the appearances and impressions. When the facts and demonstrations become evident, the degree to which they deviate from the initial assumptions will guide the degree of rigor necessary to test and explain the models. On the back end, the explanations of our results should be compiled with the a priori assumptions of the stakeholders in mind. If we are simply validating what they already thought they knew, then there is far less need for a detailed validation of our methods. However, if our models suggest a dramatic change in direction, then more care should be taken to manage the inevitable and reasonable scrutiny that will come our way.

Finally, we must ensure that the information being provided from our work points stakeholders to the more fundamental questions at hand. Sure, there's a premium to establish. But the individuals running the company don't really care what the actual premium is—they need to know the likely impacts of that premium on the business. From a financial perspective, running with this example, don't just say you've priced for a certain margin—that exact margin is, in the end, going to be exactly wrong! Explain the range of possible outcomes and the impacts of each.

This is an exciting and dangerous time for the modeling actuary. The proliferation of data and analytical techniques has opened up doors to solve problems we have been previously unable to tackle, but the same advances have brought analytical professionals from other disciplines into spaces traditionally led by actuaries. As we develop models and prepare results in this environment, I will close with three key questions that I believe ought to be asked whenever complex actuarial modeling results are shared with others:

- Am I conveying an appropriate level of confidence in my results and not leading stakeholders to trust them more than I do myself?
- Am I trying to help others develop a deeper understanding of the business or to make myself and my work sound impressive?
- Are my results helpful in expanding beyond narrow analytical questions toward addressing more fundamental issues?

Geof Hileman

Geof Hileman, FSA, MAAA, is director of actuarial studies with Kennell and Associates, Inc., in Raleigh, N.C. He can be reached at *ghileman@kennellinc.com.*

The Past Is No More Certain Than The Future

Decision-Making In The Face Of Unavoidable Uncertainty

By Charles Brass

Around midnight on June 13, 1994, Nicole Simpson and Ron Goldman were found stabbed to death outside Simpson's house while Simpson's two small children slept inside. Four days later a warrant was issued for the arrest of Orenthal James (O.J.) Simpson, Nicole's ex-husband and father of the two children. After a police chase, much of which was broadcast live on prime-time TV, Simpson was charged with two counts of murder. Simpson had not long retired from a stellar sports career, and was very much a public figure. His 134-day trial was filmed, and broadcast, live. Nine months later he was acquitted and the murders are still considered an open case by the Los Angeles Police Department.

During the 19th century, thanks largely to the work of Isaac Newton, most scientists were convinced that the only thing stopping them from fully understanding the universe was gaining access to a sufficiently powerful computing machine. Newtonian mechanics described the world as following simple mechanical rules, and most scientific research supported this view.

By the early 20th century, scientists like Albert Einstein and Walter Heisenberg had shown that what appeared to be a predictable and mechanical universe was, at its core, uncertain, chaotic and unpredictable.

Not surprisingly, these insights caused many people a lot of angst. Even Einstein seemed hopeful that further research might remove this uncertainty when he said that he didn't believe God would play dice with the universe.

Charles Brass

Charles Brass is chair of the futures foundation—an Australian based foresight practice. He can be reached at *cab@fowf.com.au.*

During the following 100 years, the uncertainty built into the universe has been completely confirmed. In fact, in some ways things have gotten worse. As one of Einstein's successors put it: "Not only does God play dice with the universe, but sometimes he hides the dice where we can't find them!"

Human beings crave certainty and regularity, and most people have a hard time coming to terms with the ultimate uncertainty that underpins our world.

Some people attempt to reduce the impact of uncertainty on their lives by suggesting it only applies to the future and not to the past. As one old adage puts it: "We may not know where we are going, but at least we all know where we have been."

A moment's reflection, however, suggests that this is not the case.

There is as much uncertainty about the past as there is about the future, but this has at least as much to do with human frailty as with the underlying rules of the cosmos.

First of all, none of us could have been everywhere at all times in the past, and even if we could the fallibility of human memory and recollection inevitably introduces uncertainty.

I am reminded of a terse interchange I once heard between a radio broadcaster and an eminent Australian historian. The historian interrupted a tirade from the broadcaster against a particular interpretation of history by saying: "Please note what it says on the front of the book—Manning Clark's history of Australia. If you don't like it, write your own."

Arguments among historians are interesting, but there are many times when we are required to make vital decisions *now* about past events, and here uncertainty and ambiguity can be devastating.

An obvious example is our legal system. Most legal cases focus on events from the past, and seek to both make decisions about the rightness or wrongness of these events and lay out the consequences for the perpetrators.

In the example with which this paper began, two people were brutally murdered. Our society views such actions as abhorrent and uses our legal system to bring the perpetrators to account.

In this case, the ex-husband of one of the victims was accused of both murders and brought to a criminal trial that began only seven months after the murders were committed. One hundred fifty witnesses were called over a nine-month period, and the jury took four hours to unanimously find O.J. Simpson not guilty. Had he been found guilty he would have faced a lifetime prison sentence (the district attorney declined to seek the death penalty). No one else has ever been charged with these crimes.

Some two years later, the victims' families commenced civil proceedings against O.J. Simpson for damages in the wrongful death of the victims. Four months later a different jury unanimously found "there was a preponderance of evidence to hold Simpson liable for damages" and awarded over $45 million to the plaintiffs.

My point is not whether Simpson is guilty or innocent, but to notice that the legal system has no difficulty imposing serious penalties (death, life in jail or huge financial payments) *today* based on an investigation of events that happened in the past.

In this case, both trials took place relatively quickly after the murders and all relevant witnesses were able to testify in both trials—and two different juries came to unanimous but opposing decisions.

Our legal system recognizes that there is rarely certainty about past events, and has evolved elaborate protocols and techniques for establishing the truth "beyond reasonable doubt."

Lawyers and judges (not to mention the general public) are probably not particularly happy when two conflicting conclusions are reached about basically the same set of events, but the system still manages to validate both decisions (Simpson is not in jail for this[1], but has relinquished most of his assets).

FUTURISTS … TEND TO CONSIDER SIMPLE EXTRAPOLATIONS OF THE PAST INTO THE FUTURE AS RATHER SHALLOW AND UNSATISFACTORY FUTURES PRACTICE.

The inevitable uncertainty about what actually took place does not prevent binding and far-reaching decisions from being made.

Even now, 15 years later, this case stirs passionate debate in America. It is still studied today because of the way forensic DNA evidence was collected and presented, and, given that O.J. Simpson has black skin, there are racial overtones as well. As recently as three years ago books and films were being produced, making various claims about guilt and innocence and the conduct of the two trials.

Again, my point is not about the individual circumstances of this case, but to note that there are occasions when unambiguous decisions need to be made today about ambiguous situations that occurred in the past—and that we have developed systems designed to make such decisions as intelligently as possible.

It is worth noting some of the ways in which our judicial system deals with the inevitable uncertainty it faces every day.

First, all parties have the right to be vigorously represented by professional counsel.

Second, everyone is sworn to "tell the truth, the whole truth and nothing but the truth" (as they see it) and there are severe penalties for deliberately lying.

Then, the entire process is conducted before a learned professional (judge or mediator) who sometimes is also the decision-maker and sometimes guides others to decide.

There are also elaborate rules, protocols and procedures about how the entire process is conducted, and how people's views will be heard.

Those charged with making decisions are given guidelines about the basis on which they will decide. In cases with the most serious consequences, the criterion is "beyond reasonable doubt"; otherwise it is "on the balance of probabilities," and if a firm decision cannot be made against the relevant criterion, the most conservative option is always followed ("innocent until proven guilty").

The system is also deliberately multi-layered, with many options to review each decision if required. The original decision is binding, unless appealed, but once appealed no irreversible action is taken until the appeal is decided.

And finally, the system is founded on precedent. What has worked well in the past is validated and repeated, and what hasn't is discarded.

Futurists face much the same dilemma as those in our legal system. Futurists work with people who need to make decisions today about the future—a time and place that is inevitably uncertain.

The legal system has had centuries to develop its protocols. Futurists as professionals have been around for fewer than 50 years. Nonetheless, futures studies are beginning to develop tools, techniques and systems designed to improve the quality of decisions made today about events that are yet to occur.

It is probably not surprising that some of the elements incorporated into the legal system are also becoming part of the futurist's toolkit. Most prominent is the emergence of professional, practicing futurists. Just as the legal system has spawned specialist lawyers, barristers and judges to assist those who seek clarity about the impact of past events on their lives, so futurists have emerged to assist those seeking to explore the future of their country, company, community or even their individual lives.

Just as the legal system encourages differing perspectives to be robustly put forward before any decisions are made, so many futurists encourage the creation of alternative future scenarios that are then explored for their likelihood and desirability.

Good futures practice also encourages the accumulation, and analysis, of as much data as possible. Futurists call this "environmental scanning" where lawyers call it "collecting evidence."

Futurists, like judges, never forget that human beings, with all their fallibilities and frailties, are an integral part of the process. So, they both couch their analysis in terms of probabilities and likelihoods, and they always remain open to the possibility that new information might revise their conclusions.

Futurists are also acutely aware of the importance of precedent—though they think about it a little differently than do lawyers. Futurists acknowledge the power that past performance might bring to the future, but they also explicitly recognize the possibility of "wild cards" that might change the picture completely. They also tend to consider simple extrapolations of the past into the future as rather shallow and unsatisfactory futures practice.

Futurists, like judges, might be somewhat discomfited when contradictory conclusions arise from the same analysis (such as the two different decisions in the two O.J. Simpson trials—or the opposing views about the impact of climate change; Is it global warming or the precursor to an ice age?), but the best practitioners learn to embrace this uncertainty and look for a frame of reference within which both can be accommodated (in the case of climate change either outcome—heating or cooling—is sufficiently dire to warrant urgent action today).

Uncertainty is a fact of life. It cannot be willed away. In the words of American philosopher Ken Wilber, this means "all truth is partial" and "any single perspective is always only a part of the picture." This does not provide an excuse for inaction, or a lack of decisiveness—but it does mean key decisions should be made intelligently and with sufficient humility to acknowledge their potential frailties.

Society has learned to accept decisions made by our legal system as binding, while acknowledging that sometimes they get things wrong.

Similarly, we are slowly beginning to recognize that good futures practice can lead to much better decisions about the future, even when we don't actually know what the future holds. ▼

ENDNOTES

[1] In September, 2007, he was arrested for subsequent felonies (including armed robbery) and found guilty. He is serving his sentence at the Lovelock Correctional Center in Nevada.

FORECASTING & FUTURISM FOURTH ANNUAL CONTEST

How do YOU Forecast?

By Doug Norris

Since our inception, the Forecasting & Futurism section has been dedicated to promoting and educating the actuarial profession with respect to innovative and leading edge predictive techniques. In SOA meeting sessions, live webinars, and publications (including this newsletter), we have explored a variety of topics including predictive modeling, genetic algorithms, agent-based models, neural networks and artificial intelligence, Delphi techniques, and Bayesian networks.

Where do we go next? That's where you come in.

THE CONTEST

In the past, our annual Forecasting & Futurism contest has concentrated on a specific subset of predictions. For instance, last year's contest was dedicated to the study of genetic algorithms. This year, we are breaking the contest wide open—we are looking for you to explore and develop a technique (literally *any* technique) that advances the science of actuarial predictions and forecasting. You may choose to:

- Investigate an interesting application of a well-established actuarial technique, such as predictive modeling, risk adjustment, or reserve modeling,
- Explore and develop an approach that has been promoted by the section in the past, such as genetic algorithms, hidden Markov models, or Delphi techniques, or
- Advance something entirely new to the profession altogether.

As you can see, you could choose to take this contest in just about any direction you wish.

CONTEST SCORING

With that said, the direction taken has to be productive. Contest entries will be scored by multiple judges, using a 100-point scale with the following criteria:

Doug Norris

Doug Norris, FSA, MAAA, PhD., is a consulting actuary at Milliman Inc. in Denver, Colo. You can reach him at *doug.norris@milliman.com*.

- (25 points max) How useful is the technique to the actuarial profession? Is this something that actuaries can and will actually use? This is the most important single criterion.
- (20 points max) How understandable is the approach to an actuarial audience? A technique might be very useful, but if no one can follow it, then it's not going to be used.
- (20 points max) How easy would it be for another actuary to reproduce your work? Have you developed the technique enough that a qualified actuary could pick it up and run with it?
- (15 points max) How sophisticated is the technique (or extension) developed? What methods are built? How much territory did you cover?
- (10 points max) How flexible is your technique? Could it easily apply or extend to other applications? Will it appeal to a broad section of actuaries, or only a very few?
- (10 points max) How creative is your approach?

This contest is equal parts art and science, so feel free to let your artistic side run wild. What's the best way to ensure that another actuary can follow your work? The answer probably depends upon the technique being developed—a written report, an Excel workbook, or a programming code, any combination of the above, or something else altogether could work best. Consider your audience.

THE PRIZE

For the first three annual Forecasting & Futurism contests, we have awarded an Apple iPad to the contest winner. We have viewed this as a "must have" gadget for many modern actuaries.

This prize has two limitations—first, being a "must have" gadget means that (by definition) many actuaries already have it. If you walk around one of the SOA's meetings, you'll notice a lot of actuaries walking down the halls with their noses buried in their iPad (how do you tell an extroverted actuary? They look at your iPad when they're talking to you). Second, between now and October, something new and even cooler could surface as the next "must have" gadget.

Therefore, this year's contest winner will receive $500 in credit to spend at the Apple store on any gadgetry of their choice (including an iPad, of course). The winner will also receive high-quality bragging rights and curriculum vitae material; just ask Jeff Heaton, the winner of last year's contest.

CONTEST SUBMISSION

To participate in the contest, you must be a member of the Forecasting & Futurism Section at the time of submission. Joining the section can be done on the SOA website. All entries must be submitted to Christy Cook at ccook@soa.org no later than Sept. 30, 2014. Christy will confirm receipt of your entry via e-mail, so please resubmit your entry if you do not receive an e-mail confirmation.

The winner will be announced in Orlando at the 2014 SOA Annual Meeting in October, at the section breakfast. Entrants do not need to be present in order to win, although being present would allow the winner to throw their arms up into the air vigorously.

THE RULES

All entrants must be current members of the SOA's Forecasting & Futurism Section at the time of entry submission. In the event that all entries are sufficiently wide of the target, the Section Council reserves the right to award no prize.

The Section reserves the right to substitute the cash equivalent of the contest prize if necessary, and the contest winner is responsible for any taxation issues appropriate to their region.

The Section and the Society of Actuaries may use submission information in publications or other SOA venues without further involvement of the entrant. For instance, we published Jeff Heaton's winning entry in the December 2013 edition of this newsletter.

Questions, concerns, or compliments about this year's contest may be sent to my e-mail address (*doug.norris@milliman.com*). If you have a potential submission idea that you'd like to talk through with me, I'd be happy to do that. I also like talking about healthcare reform, hockey, or high-altitude hiking. ▼

An Introduction to Deep Learning

By Jeff Heaton

Deep learning is a topic that has seen considerable media attention over the last few years. Many large technology companies have invested heavily in deep learning. In January of 2014, Google purchased DeepMind (a deep learning startup) for $400 million. Deep learning is being applied to the fields of robotics, computer vision, and natural language processing. Deep learning is successful because it learns by a hierarchical system of features that bears similarity to the human mind. Deep learning also works well with modern technologies such as grid computing and General Purpose Graphics Processing Units (GPGPU).

Deep learning does hold great potential for data science. However, deep learning works somewhat differently than many of the more familiar statistical models. In this article I will introduce deep learning and show how it relates to other techniques in the field of data science. I will also show how deep learning has application to the type of unstructured data seen by the insurance industry.

TOWARD COMPOSITE MODELS

Initially, you may want to compare deep learning to statistical models and machine learning models such as neural networks, support vector machines, linear regression, generalized linear models (GLM) and others. It is very important to remember that deep learning is not a specific model. Rather, deep learning is a means of combining several models together to form a composite model. The individual components will retain autonomy and can be trained independently.

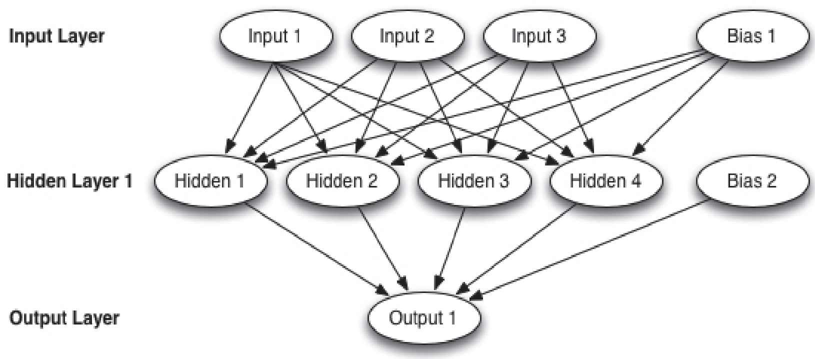
Over the last five years, boosting and ensemble learning have become two very popular techniques for producing composite machine learning models. Neither boosting, nor ensemble learning, specifies exactly what models make up the resulting composite model. The primary high level difference between boosting and ensemble learning is that boosting uses a homogeneous set of models, whereas an ensemble is heterogeneous. An ensemble is much like an orchestra producing one song with many different instruments.

Like boosting and ensemble learning, deep learning also produces a composite model. However, the deep learning model offers some unique features that are not seen in other machine learning models. Deep learning allows individual parts of the model to be trained independently of the others. Deep learning is typically applied to neural networks. However, this is by no means a necessity. Yichuan Tang, of the University of Toronto, introduced the use of deep learning for support vector machines.[1]

DEEP LEARNING ARCHITECTURE

Consider the typical multi-layer perceptron (MLP), or neural network. Such a network has an input layer, zero or more hidden layers, and an output layer. Most neural networks contain one single hidden layer. Figure 1 shows just such a network.

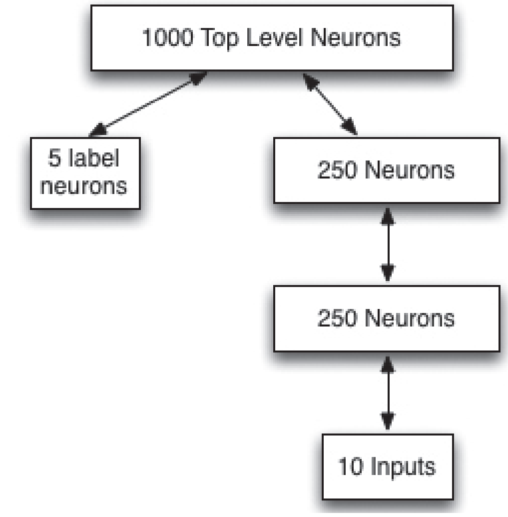
Figure 1: Shallow Multi-Layer Perceptron (MLP, Neural Network)



The above diagram shows the inputs, hidden layers, outputs and bias neurons. Weights connect these neurons together. Weights control the sigmoidal curve of the neuron's output. Bias neurons allow the neuron's sigmoidal output curve to be shifted left or right in the x direction. Most neural networks are shallow, and have a single hidden layer. However, it is possible to create neural networks with two or more hidden layers, as seen in Figure 2.

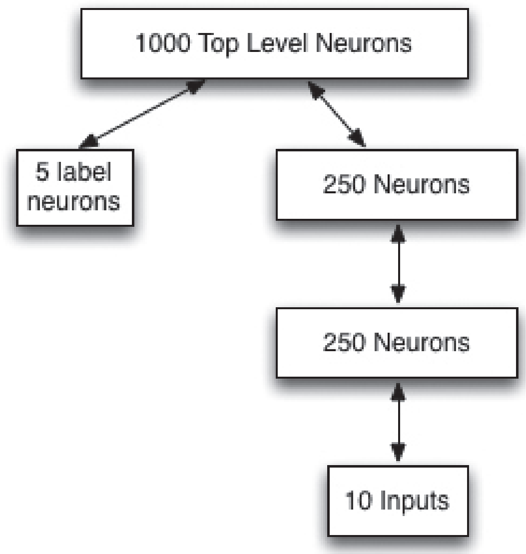
CONTINUED ON **PAGE 18**

Figure 2: A Deeper MLP (3 hidden layers)



The above neural network contains a total of three hidden layers. Most research indicates that more than a single hidden layer is counterproductive.[2] Furthermore, additional hidden layers greatly lengthen the training time for the neural network. Is a deep belief network simply a neural network that has a large number of hidden layers? Yes and no.

Figure 3: shows an overview of deep learning architecture.



Deep learning recognizes that you may not always have labels for all of the data you have collected. Deep learning allows the network to be trained using both supervised, and unsupervised techniques. You might not know the desired outcome for every item in your data set. This is OK with deep learning.

This approach very much models the way the human brain functions. A human child sees many different types of vehicles before they ever learn the difference between a car and motorcycle. However, years of learning taught that human to identify what features a vehicle has. Features describe how many tires a vehicle has, its shape, color and size. All of these features are rolled up into the person's final classification of what sort of vehicle this is.

Simultaneous supervised and unsupervised training is what lets a DBNN get away with being "deep." It is not practical to train a traditional neural network in both a supervised and unsupervised manner at the same time. The vanishing gradient problem causes the backpropagation derivatives to shrink as each new layer is added. Additionally backpropagating through many layers is computationally expensive.

Training a DBNN is usually accomplished by the following steps (shown in Figure 3).

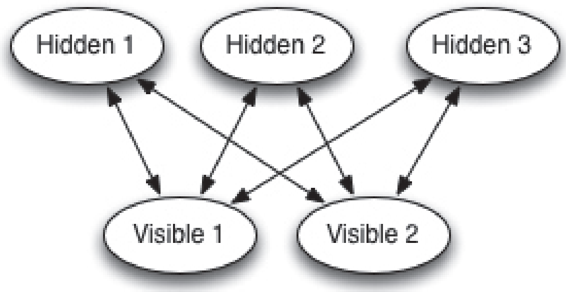
1. Train the first layer (250 neurons) with the 10-input data provided in an unsupervised way.

2. The first layer has now learned a representation of the data that is used to train the second layer in an unsupervised way.

3. This process continues until we have trained the top 1,000 neuron layer.

4. Finally, we use our labels to train a logistic regression (or similar model) based on the features extracted from the top 1,000 neuron layer. Labels identify what we are ultimately trying to predict with our model and data. For example, in an underwriting system, labels might be the final underwriting decision.

Figure 4: Restrictive Boltzmann Machine (RBM)



The key to this process is that we are hierarchically learning features first from the training data, and then features built upon features from the lower levels. We are not training the entire model at once. Each layer is trained independent of the others. Finally, the labels that we have are used to perform a more traditional gradient-based fine tuning of the final output of the model. This training method is what truly separates a DBNN from a regular neural network with a large number of hidden layers.

RESTRICTIVE BOLTZMANN MACHINES

Deep learning does not imply what makes up each level of the model. However, DBNN's are usually made up of Restrictive Boltzmann Machines (RBM). An RBM is essentially a simple neural network made up of visible and hidden elements. A sample RBM is shown in Figure 4.

The RMB is said to be restricted, because connections only occur between visible and hidden nodes. Some variants of RBM do allow lateral connections among visible nodes. However, no RBM model allows connections among the hidden nodes.

A full discussion of RBM's is beyond the scope of this article. However, one of the most challenging aspects of an RBM is that all input and output is binary. You cannot directly use continuous numbers with an RBM. One of the biggest challenges, for using an RBM, is to construct your input data as a binary vector. For computer vision problems, the input is often a pixel map. For non-graphical data, you need to get a little more creative.

DEEP LEARNING AND UNSTRUCTURED DATA

Unstructured data is a very active, and challenging, area of data science research. There are many different ways to handle unstructured data. A common task in unstructured data is to classify documents. You might want to cluster similar documents, or you might want to find similar documents given a starting example document. Most statistical models, DBNN's included, require the input data to be represented as a numeric vector.

There are many different ways to represent a document as a numeric vector. One of the most common is the "Bag of Words" algorithm. For example, to create a 2,000 element vector, the "Bag of Words" algorithm proceeds as follows.

1. Remove all "stop words" (i.e., "the," "and," "or," etc.) from the document.

2. Remove all punctuation from the document.

3. Change all words to a common stem (e.g., "people" becomes "person").

4. Perform a frequency count of all remaining words.

5. Arrange the counts of the top 2,000 alphabetically (or any consistent ordering). This is your input vector.

Because input vectors must be consistent you must always choose the same 2,000 words over all documents that you will classify. For example, if you were classifying Wikipedia articles you would build your 2,000 word vector of the most common "non-stop words" in Wikipedia. Unfortunately, this word frequency vector is not binary, as required by a DBNN. To convert the frequency vector to binary you typically establish a threshold count. Any word that has a frequency above this count is represented by 1, otherwise 0.

Attending Physician Statements (APS) are a common form of unstructured data seen in the insurance industry. Using machine learning models to classify and compare APS statements could be very useful to the life insurance industry.

CONTINUED ON **PAGE 20**

AN INTRODUCTION TO DEEP LEARNING | FROM PAGE 19

I am currently researching the applicability of deep learning, as well as other machine learning algorithms, to APS analysis.

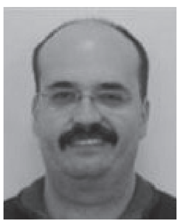
GETTING STARTED WITH DEEP LEARNING

One of the best sources of information for deep learning is the site *http://www.deeplearning.net.* This site is maintained by some of the most active researchers in the field of deep learning. This site includes a very helpful tutorial at the following URL.

http://www.deeplearning.net/tutorial/

The Python programming language is a very popular choice for deep learning research. All of the examples contained at the above URL are written in Python. They also make use of the Theano Mathematical package for Python. Theano is described as a CPU (central processing unit) and GPU (graphics processing unit) math expression compiler. Theano handles the mathematical processing behind deep learning.[3] Theano is capable of using a higher-end GPU to speed up computations by up to 140 times. GPU's in the $500 USD range can typically achieve this level of performance.

The above tutorials start with familiar statistical models, such as logistic regression. New techniques and models are then added as the tutorial progresses eventually to deep learning.



Jeff Heaton

Jeff Heaton is EHR data scientist at RGA Reinsurance Company and author of several books on artificial intelligence. He can be reached at *jheaton@rgare.com.*

ENDNOTES

1. Deep Learning using Linear Support Vector Machines, http://deeplearning.net/wp-content/uploads/2013/03/dlsvm.pdf
2. How many hidden layers should I use? http://www.faqs.org/faqs/ai-faq/neural-nets/part3/
3. Theano, http://deeplearning.net/software/theano/

Big Data in Life Insurance
Does it exist? If so, how should we handle it?

By Richard Xu and Dihui Lai

Predictive modeling is a growing capability in the life insurance industry. There are more and more discussions about how applications of predictive modeling can be used to increase production or to efficiently manage risks. At the same time, the term "big data" is commonly used in public media and within the actuarial community.

DO WE REALLY HAVE BIG DATA?

The expression "big data" is not consistently applied and can have different meanings in different situations. According to Wikipedia, big data is "a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications." In principle, data should be considered "big" when it is close to a magnitude of billion gigabytes[1] (exabyte). Data sets of this size are typically found in areas such as genomics, climate science, astronomy and nervous system connectomics. Strictly applying this definition, big data in the life insurance domain is probably more of a marketing term than a reality. Current applications of predictive modeling for life insurance are predominantly based on structured data, which is readily available from existing internal systems (e.g., policy data, claim data). Even if we include many of the external data sources that are becoming available to life companies, we are still a long way from reaching the exabyte limit.

Even if data assets of life insurance companies are not as large as some other industries, this does not mean that existing data management tools or modeling technique are sufficient to handle the ever-expanding data sets. There are many business analytic projects involving huge amounts of data that result in a traditional approach being ineffective or even impossible. One example is a traditional life experience study where study output can easily reach tens of millions of records. Building a predictive model on data of that size can already prove to be challenging. So, practically speaking, the life insurance industry is indeed facing "big data" in that the data is big enough that it can no longer be effectively processed or analyzed using traditional methods.

HOW CAN WE HANDLE BIG DATA?

Vague as the term "big data" is, the solutions to the challenges it creates can vary. Two major issues arise as a company's data volume increases: capacity and speed. Upgrading the hardware (memory and processors) can be a simple and inexpensive solution. If necessary to go beyond the limits of a desktop PC, a terminal server provides great memory capacity and has the advantage of incorporating multiple processors. This can be one possible solution for a reasonably large data set.

Cluster computing techniques are also relevant to the topic of big data analysis. This approach, including Massively Parallel Processing (MPP) and Hadoop system, partitions and processes data across a number of distinct but interconnected computing nodes. The final result is assembled once the individual bits and pieces are completed. MPP has a longer history than Hadoop and has the advantage of using SQL as its interface.[2] Hadoop, on the other hand, processes data in parallel using a MapReduce framework.[3] Although powerful, a cluster solution can be expensive to construct and maintain.

Other than attacking big data with these atomic tools, it is sometimes more efficient to solve memory or speed issues using better memory allocation techniques or algorithms. The R package *ff* provides a disk-stored data structure that can be accessed as if it were in RAM. Additionally, the R package *bigmemory* is especially good for dealing with large matrices of data. However, these types of packaged solutions are best for solving specific problems and might not be ideal for problems that go beyond the intended scope. Commercial software such as SAS or Revolution R provide a better ability to deal with large data sets and are generally better integrated with large data packages.

After all, problems dealing with big data are usually case-specific and solutions will depend greatly on the nature of the data set. In the following section we will demonstrate a real-world example of how big data can be approached.

CASE STUDY: BUILDING A GLM USING BIG DATA

The generalized linear model (GLM) is widely accepted as an efficient tool in insurance analytics. The standard *glm* function provided in R meets most of the everyday demands and applications of GLM. However, the *glm* function becomes less efficient when faced with big data. The fitting procedure can be very slow on a data set of several million records. Additionally, the calculation process might not even complete on a regular desktop PC due to memory overflow.

For this case study, we used a 5.64 million record data set that was initially used for industry post-level term lapse study and demonstrate the efficiency of the available GLM routines from three different R packages (*primary R, biglm* and *RevoScaleR*[4]). The lapse model we tested consists of 16 independent parameters and assumes that lapses follow a Poisson distribution.

All functions tested for this experiment (summarized in Table 1) finished the job in a reasonable amount of time. The *glm* function required about 2 GB memory and four minutes to finish, while directly calling the *glm.fit* function shortened the procedure significantly. In comparison to the built-in *glm* function, *bigglm* is much more economical in terms of memory allocation, while the routine requires a comparable amount of time to finish. The *rxGlm* function is excellent in speed (finishing the procedure in less than a minute), yet this function requires only slightly more memory than *bigglm*. In summary, the built-in *glm* function is flexible and easy to use, but not ideal for big data. The *bigglm* function is excellent in memory efficiency, but *rxGlm* is superior in computing speed.

CONTINUED ON PAGE 24

BIG DATA … | FROM PAGE 23

Function	Elapsed Time (s)	Memory* (Mb)
glm	185	2408
glm.fit	78.3	1056
bigglm	209	2.3
rxGlm	28.8	43.5

Table 1: Comparison of GLM Function Using Different R Packages. The CPU time is evaluated using the built-in function proc.time in R and the memory usage is evaluated using a wrapped-up gc function. The model is run on a PC desktop (Intel core i7-3770 CPU 3.4GHz and 12 GB).

Building a successful model requires construction of multiple models and then selecting the best among the candidates. The procedure can be computationally intense and time-consuming. Optimizing the model selection procedure would be beneficial to modelers. By default and without any add-on packages, R only uses one core for processing. Parallel computing packages such as *multicore*, *snowfall* can take advantage of multi-core features and speed up the model selection tasks. For the case study, we tested the *snowfall* package to demonstrate the power of parallel computing. The same data set was used for this test. Six variables (in other words, six models) are tested for significance.

The results showed that the built-in *glm* function failed to complete due to memory error. The *bigglm* function finished the routine in 16 minutes, with <10 MB of memory usage while parallelizing the procedure reduces the time by half. The usage of *snowfall* only reduces the procedure by about 10 seconds when *rxGlm* is used as the core function. Overall, parallelism can speed up the model selection procedure but can put some stress on the memory demands.

Approach	Elapsed Time (s)	Memory (Mb)
lapply*+glm	Memory overflow	
snowfall+glm	Memory overflow	
lapply+bigglm	1004	<10
snowfall+bigglm	497	<10
lapply+rxGlm	54.4	44.4
snowfall+rxGlm	43.2	204

Table 2: Comparing the Serial Approach and Parallel Approach. The evaluation of CPU time and memory is the same as what is described in Table 1. *lapply is a built-in R function that enables the process of a list of models serially.

CONCLUSION

Big data is no doubt a big topic in the world of insurance and will become even bigger in the future. Tools are available to help us, but we must be careful in making our decision. Depending on the nature of projects and data attributes, the optimal solution can vary. In the case study presented, we see that Revolution R provides the best solution if speed is the priority, while *biglm* should be considered if memory is of greater concern. Big data is on its way and will no doubt present challenges. To be successful, companies need to prepare.

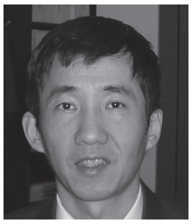
REFERENCE

[1] Francis, Matthew. 2012. Future Telescope Array Drives Development of Exabyte Processing. *Ars Technica*, April 2.

[2] *Massively Parallel Processing* (DW)—*A Technical Reference Guide for Designing Mission-Crtical DW Solutions*, http://technet.microsoft.com/en-us/library/hh393582.aspx

[3] Borthakur, Dhruba. 2007. The Hadoop Distributed File System: Architecture and Design.

[4] Rickert, Joseph B. 2011. The RevoScaleR Data Step White Paper.



Richard Xu

Richard Xu, FSA, Ph.D., is senior data scientist at RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *rxu@rgare.com.*

Dihui Lai

Dihui Lai, Ph.D., is data scientist analyst at RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *dlai@rgare.com.*

A Conversation About the Delphi Study on Long-Term Care Financing Solutions

By Ben Wolzenski and Ron Hagelman

Editor's note: *Ron Hagelman is co-chair of the LTC Think Tank, and both Ron and Ben were members of the project team for "Land This Plane"—How America Should Deal With The Pending Long Term Care Crisis: A Research Study Using The Delphi Method.*

Ben: Ron, thanks for your time today. I'd like to explore your thoughts on the use of the Delphi Method to study LTC financing solutions. But first, please remind our readers of the state of LTC financing in the United States.

Ron: The CLASS Act [Community Living Assistance Services and Supports] was demonstrated to be actuarially unsound and was surgically removed from The Affordable Care Act. A new Commission on Long Term Care was created when the American Taxpayer Relief Act of 2012 was signed into law on Jan. 2, 2013.The Commission was composed of 15 appointees from the President and both parties in the Senate and House of Representatives. The Final Report was published on Sept. 18, 2013 on a nine to six bi-partisan vote. There is recognition of the clear and present danger of inaction. The report delivered to the President ends with "We request your highest attention to this Report and urge you to take action to maintain momentum toward creating a Long Term Services and Supports system that will meet the needs of all Americans with functional or cognitive needs now and in coming generations." While the Commission supported "a comprehensive and balanced approach to public and private responsibility" its members ultimately failed to agree on a specific financing option.

Ben: I guess that is not surprising, considering that the Commission was composed of political appointees. What expertise did the LTC Think Tank bring to the issue?

Ron: The LTC Think Tank has more than 70 member "experts": regulators, actuaries, market/sales leaders and insurance company executives. We have worked constantly to try to find consensus and discern ways to move private insurance solutions forward. A smaller group was needed for this type of subjective opinion consensus building. We therefore created a Long Term Care cohort of 40 experts selected proportionally by discipline category.

Ben: It is time to remind our readers of how the Delphi method works. The Delphi method provides a means of gathering the opinions of a panel of experts and finding the extent to which there is a consensus among them. It consists of anonymous rounds of opinion gathering, which are fed back to the panel of experts (again, anonymously) for their further responses. In group meetings or discussions there can be issues with deferral to hierarchy or aggressive behavior. Having the opinions anonymously compiled and fed back by a moderator helps to remove these effects. In addition, the use of multiple anonymous rounds allows panelists to revise their views as the study progresses, without any pressure to maintain a publically expressed opinion. So, Ron, how did the Delphi method play out in this study?

Ron: The final study results generated six principles that received the support of large majorities of the panelists. For ease for the readers, we've agreed those principles will be shown in bullet point format. They are:

- Principle 1: A Robust and Efficient LTC System—Need for a robust and efficient LTC system: 88 percent of panelists agreed.
  Private insurance should be a part of solution: 100 percent of panelists agreed;
- Principle 2: Social Insurance—Social Insurance is a necessary part of the solution: 88 percent of panelists agreed;
- Principle 3: Changes in Medicaid—Medicaid Reform, tighten eligibility: 79 percent of panelists agreed;
- Principle 4: Changes to Regulations and Legislation—Outlined a need to modify the NAIC Model Act;
- Principle 5: An Active Government Role—Need government sponsored public awareness: 92 percent of panelists agreed; and

- Principle 6: Improved Marketing and Sales—Improve LTCI Training: 83 percent of panelists agreed.

Ben: This three-round Delphi study produced surprising consensus from our diverse group of panelists. Ron, how would you wrap this up?

Ron: What is clear to all those committed to the eventual success of this market is that business as usual will not get us to the level of market penetration required to achieve critical mass in solving what many believe may be America's largest unprotected risk. The Delphi method has allowed us to identify all the moving parts of a new future, better focused on potentially achievable reforms. We have, I believe, developed a true road map for a new and reinvigorated campaign to increase company participation, look for leadership from public sources to increase awareness and tax incentives, achieve needed reforms from existing government programs, obtain more flexible product options from the NAIC, produce greater professional agent participation in LTC risk abatement, explore reinsurance and take another hard look at new and innovative product design.

[See the complete Delphi studay at http://www.soa.org/Research/Research-Projects/Ltc/research-2014-ltp-ltc.aspx.]

Ben Wolzenski

Ben Wolzenski, FSA, MAAA, is managing member at Actuarial Innovations, LLC in St. Louis, Mo. He can be reached at *bwolzenski@rgare.com.*

Ronald R. Hagelman

Ronald R. Hagelman, Jr., CLTC, CSA, LTCP, is president of America's Long Term Care Insurance Experts, a nationally recognized network of BGA's (brokerage general agents) specializing in long-term care insurance. His office is in New Braunfels, Texas and he can be reached at *ron@rmgltci.com.*

Warm and Fuzzy … And Real!

By Dave Snell

(This is part one of a two-part article. Part two will be in our next issue; and both parts will be incorporated into a presentation at the 2014 SOA Annual Meeting in Orlando (October 25) and a Forecasting & Futurism Webcast in December, 2014)

Fuzzy logic is not new. It has been around for a long time.

The previous two sentences contain a few examples of fuzzy logic in our real life environment.

- New—when does new begin or end?
- Around—nearby? How near? How prevalent?
- Long—how long is long? How many years, months, hours, minutes, seconds?

We learn fuzzy logic as children, well before we enter formal schooling:

"Don't touch that! It's hot."

"We had lots of fun and good food at Grandma's kitchen."

So, why should it come as a surprise to us to learn that fuzzy logic is often a better methodology than "crisp" logic for many actuarial modeling situations?

Crisp set theory and crisp logic are more recent terms for what we used to consider set theory and Boolean logic. In crisp set theory an item is either a member of the set or not a member of the set. We can easily say that 0.96 is a member of the set {-5.7, 0.96, 7} and that -2.5 is not a member of that set. Fuzzy set theory deals with sets where membership does not have to be strictly in or out. Take Tall for example. A person 69 inches might be considered Very Tall for a 10 year old, or Tall for an adult woman; but Not Tall for a basketball player.

I'm six feet tall; and I used to consider myself a little taller than the average male. When my wife and I first started dating, I met her 6' 4" brother and her 6' 6" cousin. Her dad liked me even though he thought I was a bit "short."

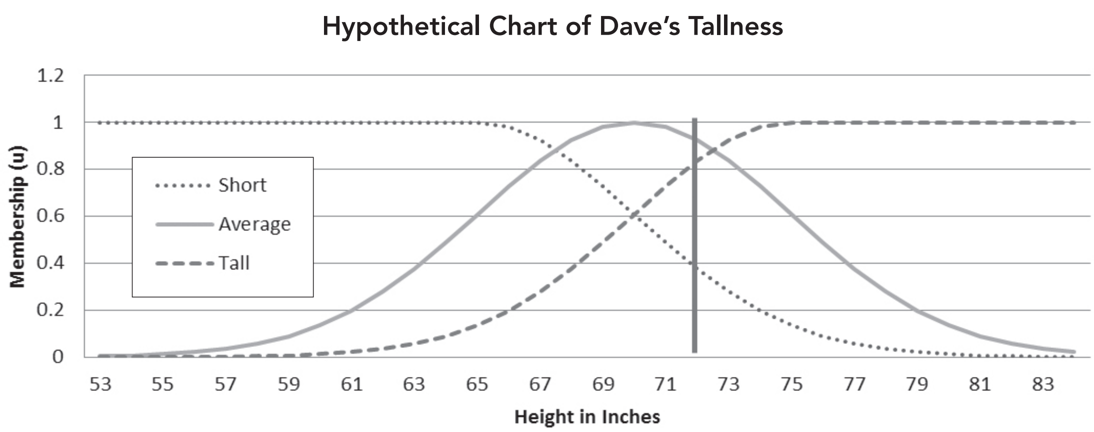


Figure 1 - A hypothetical chart of 'Tallness'. Although it is often tempting to consider membership in a fuzzy set as the probability of being in that set, that is misleading. Note that one data point can be a member of several sets, and the membership values do not have to sum to 1.

Soon after we moved from Connecticut, on the East coast of the United States, to St. Louis, in the Midwest, we were eating in a restaurant and the waitress asked if we wanted any dessert. She offered sherbet among the selections, but pronounced it sherbert [sic]. I picked up on this right away. It was the same way we mispronounced sherbet in my section of Connecticut; and I asked if she was from back East. She enthusiastically said "yes!" ... she was from East St. Louis! Our definitions of East differed by about 1,000 miles. Some of my Asian friends would consider my definition as laughable, as they think of Japan as back East.

Tall, East, Near, Hot and *Many* other adjectives are what Fuzzy Logic considers Linguistic Variables. Like the more conventional variables we use in our actuarial models, they can take on specific values (72 inches, 86 degrees longitude, 3.8 miles, 40 degrees Celsius, 7,583,278); but they usually imply a range and that range is relative to some other ranges. It is not necessary to tell your child that the food is 160.53 degrees Fahrenheit. The more important information, that you can say quickly and your child can understand immediately, is that it is hot, and might burn his tongue.

Likewise, a life insurance underwriter has neither the time nor the data to determine that an applicant for this $5,000,000 policy will live for another 17.45 years with a standard deviation of 5.6 years. She is under time (and data) constraints; and must quickly decide if this person is a preferred, standard, substandard, or uninsurable risk.

Her decision may be based on a glycohemoglobin blood test (aka A1c - longer term sugar level) result in the normal range, a body mass index (BMI) of overweight, but not obese, and a family history (one or more close family members) of diabetes but a blood pressure only slightly elevated over normal for the applicant's age and gender. She may have results from several blood tests for this applicant, and she compares them to the 'reference range' values such as shown in Figure 2.

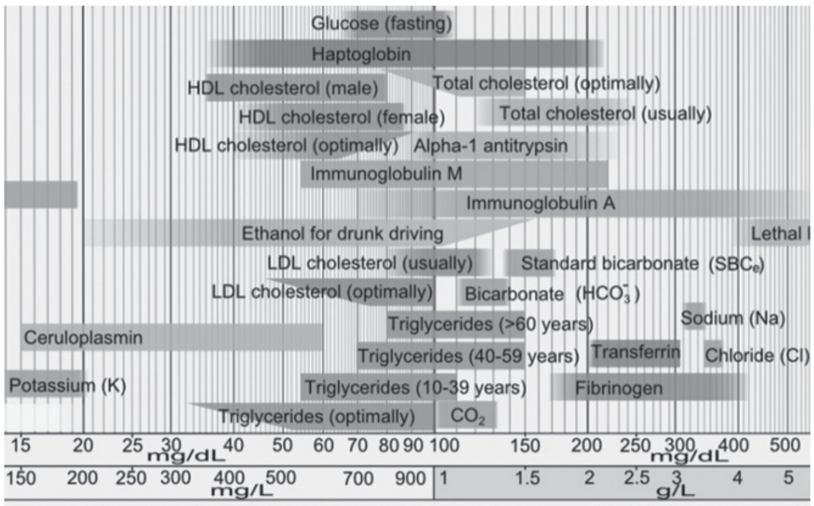


Figure 2 - Subset of common blood reference ranges—source: author's subset of excellent (award winning) Wikipedia image //upload.wikimedia.org/wikipedia/commons/thumb/c/cb/Blood_values_sorted_by_mass_and_molar_concentration.png contributed by Mikael Häggström, MD and released under the Attribution-Share Alike 3.0 Unported license

Fuzzy logic provides a way to work with these linguistic variables and reach a quantitative (if desired) answer.

According to the Mayo Clinic, the normal fasting blood sugar range for an individual without diabetes is 70-100 mg/dL (3.9-5.6 mmol/L).[1] Does that mean that every "normal" person without diabetes will have a fasting blood sugar level in that range? If you have 69.9 mg/dL or 100.1 mg/dL does that automatically make you less healthy than an individual with 70 or 100 milligrams/deciliter? Is the range truly that crisp as in Figure 3?

CONTINUED ON **PAGE 30**

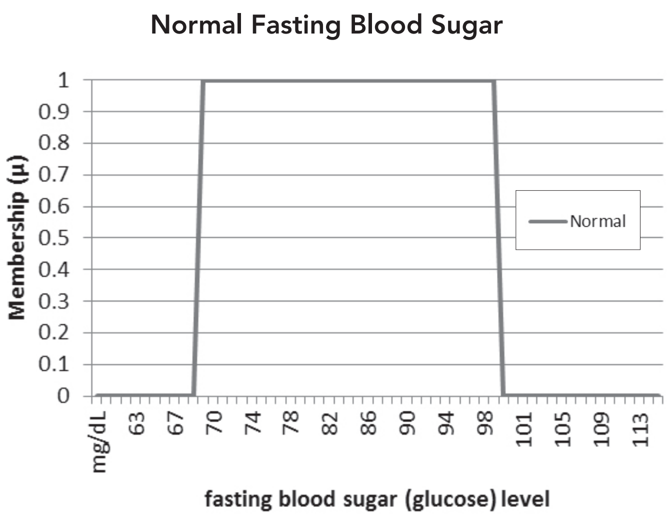


Figure 3 - Crisp set theory representation of "Normal" fasting blood sugar

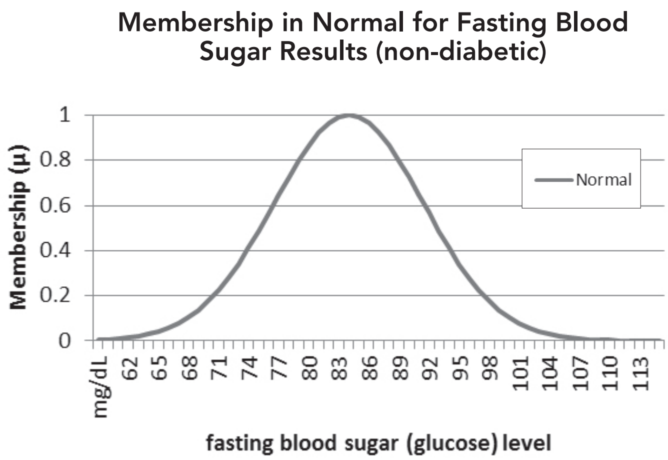


Figure 4 - Fasting Blood Sugar (Glucose) Results (assuming a normal distribution with mean 85 mg/dL and standard deviation of 7.5 mg/dL)

Actually, no! One drawback of this binary classification approach is the conflict between sensitivity and specificity. Sensitivity measures the proportion of actual positives which are correctly identified as such. Also called a True Positive, this measures the percentage of people tested for dread disease X who actually have dread disease X. Specificity measures the True Negative rate—those people who do not have dread disease X and are correctly diagnosed as not having it. In general, laboratory testing attempts to maximize specificity, even if it means missing a few positives.[2]

A reference range is usually a set of values 95 percent of the normal population falls within. A better view of this might be that of Figure 4:

ONE DRAWBACK OF THIS BINARY CLASSIFICATION APPROACH IS THE CONFLICT BETWEEN SENSITIVITY AND SPECIFICITY.

Now that we know the potential advantages of fuzzy logic, how do we apply it?

It's as simple as one, two, three:

1. Fuzzification – convert your input and output to linguistic values, utilizing ranges and membership functions.

2. Apply rules (from your experience or knowledge base) using fuzzy logic.

3. Defuzzification – convert your results to the form you want (often a numeric result).

OK, that's probably not apparent, so let's look at a very simple example in order to better understand this.

Let's assume that Applicant James, age 25, has applied for a $20,000 life insurance policy. James lives in a state considered "medium" for cocaine usage; but he works five miles away in a state classified as "high" for such usage. Assume also, that we are back in 1996, when the following chart may have applied to the situation.

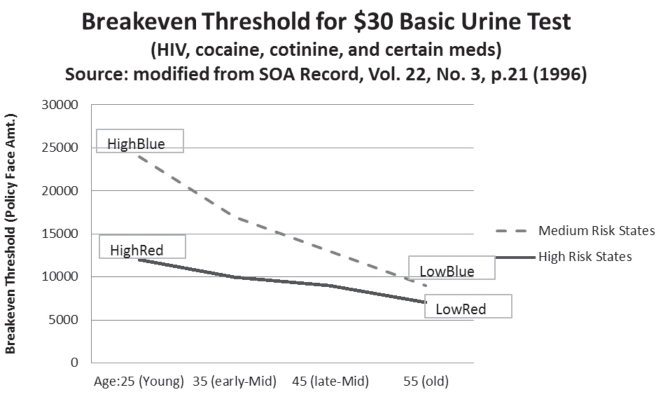


Figure 5 – Breakeven threshold (point at which test becomes cost effective) based upon applicant age and state of residence. Some states have much higher incidence of cocaine usage. Note that prices and state characteristics may have changed significantly since 1996. Chart built from Table in SOA Record, Vol. 22, p.21 and modified for this example.

Let's assume that an underwriting rules table has been developed for this condition. Here is a portion of the rules that were developed:

1. If Age is Young and HighStateActivity is Significant then Threshold is HighRed

2. If Age is Old and HighStateActivity is Significant then Threshold is LowRed

3. If age is Young and HighStateActivity is Not Significant then Threshold is HighBlue

4. If age is Old and HighStateActivity is Not Significant then Threshold is LowBlue

Obviously, a real situation would have more rules, since I have not even covered the two mid-age groups and we would normally have more information and criteria.

One advantage of this type of rule set is that it uses a more natural language. Underwriters are used to using natural language terms such as *Overweight, Obese, Hypertension, Diabetic,* etc. versus a series of numbers. Plus, using these as parameters, the definition of terms like *Obese* and *Hypertension* can be refined (and they have been as standards have been changing to reflect new medical study results) and the same rules can apply.

Here we shall define HighStateActivity as the membership in the High risk state; and we'll say that it is *Significant* if that membership is greater than 0.50, and not *Significant* it is 0.50 or less.

Our definitions of *HighBlue, LowBlue, HighRed, and LowRed* here are going to be very simple. We shall make them the endpoint values for the Blue line (the dashed line if you are not seeing this in color) and the Red line (the solid line). According to the underlying table from the SOA Record, this would mean that they would be as follows: *HighBlue*=$24,000; *HighRed*=$12,000; *LowBlue*=$9,000; *LowRed*=$7,000.

Since there are no mandates for how you choose your membership functions,[3] we can go ahead here and say that James is Young with a membership of 1.00 and Old with a membership of 0.0 and then we'll just have to deal with one linguistic variable, the HighStateActivity, for our example. We will assume a distribution of membership (μ) in the two states according to the carefully prepared proximity study (Not!) below in Figure 6 and the similarly prepared time chart of Figure 7 as follows:

CONTINUED ON **PAGE 31**

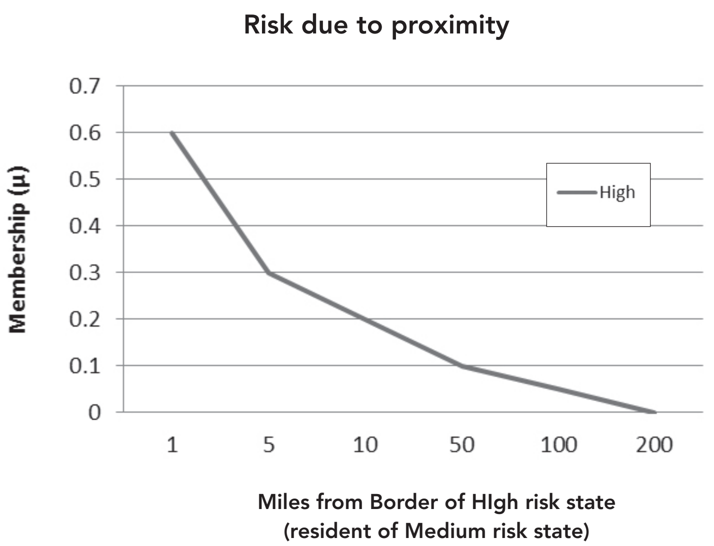


Figure 6 - Assumed memberships for resident of Medium risk state living near High risk state (hypothetical values)

Distance to Border (miles)	μ^{High}_{miles}
1	0.60
5	0.30
10	0.20
50	0.10
100	0.05
200	0

Membership due to proximity to the High state

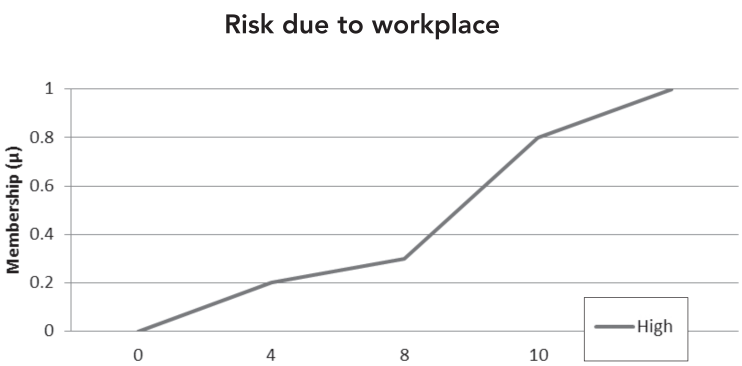


Figure 7 - Assumed memberships for resident of Medium risk state who works in nearby High risk state (hypothetical values)

Time at work (hours)	μ^{High}_{miles}
0	0
4	0.2
8	0.3
10	0.7
12+	1.0

Membership due to work in the High state

We know that James resides in the Medium risk state so his membership in that state is

μ^{Medium} = 1.0 (see note[4] for discussion)

I am going to assume here that James works 9 hours per day at his job in the High risk state, just so that we can use some fuzzy set theory here (and even do an interpolation … but no fancy stuff). What do we have going here? Essentially, we have a union of two sets. The two sets have some overlap (James must spend some time in the High risk state in order to go to work there) but he may also spend time there after work. Since he lives only five miles away, he might go there on weekends or evenings; but this is not required.

In fuzzy logic, there are several ways to handle membership in the union of two fuzzy sets. A very popular (and simple) one, that Lotfi Zadeh (founder of fuzzy logic) proposed, is to take the maximum of the two set memberships.

Thus, $\mu^{High} = \max(\mu^{High}_{miles}, \mu^{High}_{work}) = \max(0.3, (.3+.8)/2) = .55$ which is *Significant*, so we will use our *HighBlue* and *HighRed* result values in the defuzzification process.

and $\mu^{Medium} = 1.0$ since James lives in the Medium risk state

James applied for $20,000 of coverage.

The threshold for his age (25) would be

$$\frac{\mu^{High}}{\left(\mu^{High}+\mu^{Medium}\right)} \times HighRed + \frac{\mu^{Medium}}{\left(\mu^{High}+\mu^{Medium}\right)} \times HighBlue$$
$$= \frac{.55}{(1+.55)} \times HighRed + \frac{1.0}{(1+.55)} \times HighBlue$$
$$= \frac{1.0}{(1+.55)} \times \$12{,}000 + \frac{1.0}{(1+.55)} \times \$24{,}000 = \$19{,}742$$

so a urine test would be cost effective (but just barely).[5] Fuzzy logic has given us an alternative way of addressing a problem with incomplete data.

Recap:

1. Fuzzification—We converted our input and output to linguistic values (*Young, Old, HighStateActivity, Significant* for input; *HighRed, LowRed, HighBlue, LowBlue* for output), utilizing ranges and membership functions.

2. We applied rules (based on *age, stateactivity, significance*) using four fuzzy logic rules we defined to determine our outcome (threshold).

3. Defuzzification—We converted our results to the form wanted (in this case, we just took the endpoint values and computed an average, weighted by membership).

This was a contrived example where I tried to avoid nearly all mathematics and programming. In practice, *HighBlue* and the other linguistic values and variables would be shapes where you would use centroids, matrices, and various types of distribution functions for memberships.

Yes, you could have done this example with crisp logic. Most destinations can be reached by several paths. Fuzzy logic offers a more natural language, a way to deal with imprecise or incomplete data, and a way to group items together so that complexity is reduced, rule sets can be smaller, and speed of solution can be increased. Consider it as one more arrow in your quiver of actuarial tools.

FUZZY LOGIC OFFERS A MORE NATURAL LANGUAGE, A WAY TO DEAL WITH IMPRECISE OR INCOMPLETE DATA, AND A WAY TO GROUP ITEMS TOGETHER SO THAT COMPLEXITY IS REDUCED.

More sophisticated examples would also be likely to employ hedging. If *Tall* has a membership value in the range from 0 to 1.0, then *Very Tall* could be defined as the square of this value; and *Nearly Tall* might be the square root of the *Tall* membership value. In this way, we keep a consistent relationship, where the *Very Tall* is more selective than *Tall*, which in turn is more selective than *Nearly Tall*.

The logical question (fuzzy or crisp) you may be asking is "why isn't fuzzy logic in wider use in the actuarial profession?" In the actuarial area, fuzzy logic is still a relatively new paradigm. It is a shift from old ways of thinking; and that results in initial resistance from those more comfortable with their older toolset.

George Klir and Bo Yuan stated this eloquently in their book, *Fuzzy Sets and Fuzzy Logic – Theory and Applications*:[6]

"Each paradigm, when proposed, is initially rejected in various forms (it is ignored, ridiculed, attacked, etc.) by most scientists in the given field. Those who support the new paradigm are either very young or very new to the field and, consequently, not very influential. Since the paradigm is initially not well-developed, the position of its proponents is weak. The paradigm eventually gains its status on pragmatic grounds by demonstrating that it is more successful than the existing paradigm in dealing with problems that are generally recognized as acute. As a rule, the greater the scope of a

CONTINUED ON PAGE 34

WARM AND FUZZY … | FROM PAGE 33

paradigm shift, the longer it takes for the new paradigm to be generally accepted."

Surprisingly, although fuzzy logic was first proposed in the United States,[7] it was most enthusiastically accepted in Asia. Today, your Canon or Minolta camera probably has a fuzzy logic control circuit to stabilize the pictures you take. Your Honda or Nissan auto transmission selects the optimal gear ratio for your driving style and the engine load conditions; and my Toyota Prius even knows when to switch to the electrical motor or gasoline engine, or both, for the best mix of power and fuel economy. Your Sharp refrigerator decides when to turn on defrost or cooling cycles based on your needs. The newer washing machines from Korea and Japan adjust their strategy based upon the level of dirt, the water level, the fabric type and the size of the load. In Japan, "Fuzzy" has become a sort of quality seal proudly displayed on consumer products. One theory about the difference between Asian embracement of fuzzy logic is that it more closely fits with the concept of yin-yang, where contrary forces interact to various degrees in the natural world. Whatever the reason, it appears that the West was slower to adopt fuzzy logic.

Figure 8 - Yin-Yang nature of the natural world is better exemplified by fuzzy, rather than crisp set theory.

An encouraging recent exception in the actuarial area is an excellent research paper jointly sponsored by the CAS/CIA/SOA Joint Risk Management Section.[8]

As actuaries, we have a natural inclination towards precision. Yet, as Matisse so aptly reminded us, "Precision is not truth."[9] Reality is a bit more fuzzy, and fuzzy logic is better suited for the cases where we have imprecise data and incomplete subject matter expertise.

Next issue, we'll go into more depth and examples of the mechanics involved with fuzzy logic. You can get very sophisticated with matrix algebra, exotic distribution functions for the fuzzification and a host of defuzzification techniques. In the meantime, please read Jeff Heaton's article ""Fuzzy Logic in R" in this issue. We tried to coordinate in this issue so that this article could focus on the "Why" and some theory, and his on the "How" for a jump start. Jeff shows how to use the host of packages available for plug-and-play processing of fuzzy logic in the programming language R.

As Lotfi Zadeh, the founder of fuzzy logic said in 1973,

"We must exploit our tolerance for imprecision.[10]

Enjoy being less crisp, and more real!

Dave Snell

Dave Snell, ASA, MAAA, is technology evangelist at RGA Reinsurance Company in Chesterfield, Mo. He can be reached at *Dave@ActuariesAndTechnology.com.*

END NOTES

1. http://www.mayoclinic.org/diseases-conditions/diabetes/expert-blog/blood-glucose-target-range/bgp-20056575
2. ALU101 Textbook – 5th Edition, p. 115, Association of Home Office Underwriters.
3. Develop your membership function to fit your problem. Sometimes it is determined heuristically and sometimes it is a subjective decision based on your experience or intuition. The fuzzy logic literature shows a lot of triangular, trapezoidal, Gaussian and bell-shaped functions. We'll investigate some of them next time; but the focus of this article was to keep the mathematics very simple.
4. You might argue that the membership in the Medium risk state should decrease as some threshold is passed of membership in the High risk state; and you may be correct! There is much subjectivity in the choice of membership function distributions. One answer or standard does not fit all situations.
5. Obviously, we are applying group methods to a single individual; and we might be wrong. In general though, we expect James to be a representative sample of the group.
6. Klir, George and Yuan, Bo [1995], Fuzzy Sets and Fuzzy Logic – Theory and Applications, Prentice Hall P T R, Upper Saddle River, New Jersey,1995, p.30
7. Lotfi Zadeh, a Professor at University of California, Berkley, is considered the founder of fuzzy mathematics, fuzzy set theory, and fuzzy logic. He published his seminal work, "Fuzzy sets", in 1965
5. Shang, Kailan and Hossen, Zakir [2013] "Applying Fuzzy Logic to Risk Assessment and Decision-Making", CAS/CIA/SOA Joint Risk Management Section. Note: Arnold Shapiro and others have also written research papers on the utilization of fuzzy logic. Search for fuzzy logic on the SOA website for a current list of actuarial papers.
6. Henri E. B. Matisse, 1869-1954, as quoted in Ross, Timothy [2010] "Fuzzy Logic with Engineering Applications, Third Edition, John Wiley and Sons, Ltd., UK.
7. L.A. Zadeh,"Outline of a new approach to the analysis of complex systems and decision processes",IEEE Trans. Syst., Man, Cybernetics, SMC-3 (1973), pp. 28–44

Fuzzy Logic in R

By Jeff Heaton

R is a programming language designed for statistical computing. R is widely used in scientific, actuarial and data science computing. The greatest strength of the R programming language is the many third party packages contributed by R's user community. In this article, I will introduce you to fuzzy logic programming in R. This article assumes that the reader already has knowledge of fuzzy logic. If you need a review of fuzzy logic please read "***Warm and Fuzzy … And Real!***" by Dave Snell. It is also in this issue.

There are several different R packages available for fuzzy logic programming. This article assumes that you are using the "sets" package. If "sets" is not already installed, it can be installed with the following command.

```
install.packages("sets")
```

You only need to install the "sets" package once. After installation, R programs can make use of "sets" by invoking the following command.

```
library(sets)
```

If you get an error from the above command, then "sets" is not properly installed on your system.

I will now show you how to set up a basic fuzzy system. This system will implement a very simplistic underwriting rating system. This will allow you to define linguistic variables about a potential insured. Fuzzy rules will be defined based on those linguistic variables. You will also be able to perform fuzzy inference and ultimately defuzzify to an underwriting rating.

You must first define the range and granularity of your universe. The universe used for this example will be between 0 and 40, with a granularity of 0.1. The inputs for all of your variables must fit within this range. Additionally, the granularity will specify the accuracy of the fuzzy inferences. At a more superficial level, the range also defines the x-axis of your plots. I used the following range and granularity for this example.

```
sets_options("universe", seq(from = 0,
to = 40, by = 0.1))
```

If your individual variables use vastly different ranges, it may be useful to normalize the variables to more consistent ranges.

Linguistic Variables

Linguistic variables allow the use of descriptive words such as "underweight" or "obese" to describe normally numeric variables. Underwriters use many different variables to assign a rating to a potential insured. For this example we will only consider the hemoglobin A1c (HbA1c) blood test, a hypertension class and body mass index (BMI). We will place this set of linguistic variables into a set named "variables."

```
variables <-

 set(
```

Starting with BMI—we define several linguistic values, such as "under," "fit," "over," and "obese." We define the mean for each of these in BMI. For simplicity, I assign a standard deviation of 3.0 to each. There are a variety of fuzzy membership functions available to define your variables. For BMI, I am using a normal distribution.

```
   bmi =

    fuzzy_partition(varnames =

    c(under = 9.25, fit = 21.75,

     over = 27.5, obese = 35),

     sd = 3.0),
```

For the linguistic variable "a1c" I use a conic fuzzy membership function, with a radius of five. I define linguistic values of "l" (for low), "n" (for normal) and "h" (for high).

CONTINUED ON **PAGE 36**

These are assigned to actual a1c test values.

```
a1c =
 fuzzy_partition(varnames =
 c(l = 4, n = 5.25, h = 7),
FUN = fuzzy_cone, radius = 5),
```

The linguistic variable "rating" also defines its membership with a cone. This set defines the underwriter rating for the proposed insured. Underwriter ratings can range from 10 (decline) to 1 (preferred). I define three linguistic values in this range. The linguistic variable "DC" is decline, "ST" is normal and "PF" is preferred.

```
rating =
 fuzzy_partition(varnames =
  c(DC = 10, ST = 5, PF = 1),
  FUN = fuzzy_cone, radius = 5),
```

Finally, I define linguistic variable "bp" to represent blood pressure. Here I normalize the systolic and diastolic readings to a single value. The value 0 represents normal and 30 represents severe hypertension. It does not matter, for this example, exactly how you normalize the actual systolic and diastolic values. I suggest using a table, similar to the following URL.

http://www.mayoclinic.org/diseases-conditions/high-blood-pressure/in-depth/blood-pressure/art-20050982

I use a normal distribution fuzzy membership function, with a standard deviation of 2.5. I define linguistic values of "norm" (normal), "pre" (prehypertension), "hyp" (hypertension) and "shyp" (severe hypertension).

```
 bp =
  fuzzy_partition(varnames =
  c(norm = 0, pre = 10, hyp = 20,
   shyp = 30), sd = 2.5)
)
```

Now that the linguistic variables have been defined, rules can be created.

FUZZY RULES

Fuzzy rules are used to link the linguistic variables of "bmi," "a1c," and "bp" to the linguistic variable "rating." I use three different rules for this example. You can see this rule set here.

```
rules <-
 set(
  fuzzy_rule(bmi %is% under || bmi %is%
obese || a1c %is% l,
        rating %is% DC),
  fuzzy_rule(bmi %is% over || a1c %is% n
|| bp %is% pre,
        rating %is% ST),
  fuzzy_rule(bmi %is% fit && a1c %is% n
&& bp %is% norm,
        rating %is% PF)
 )
```

The first rule states that the rating will be DC (decline) if the BMI is "under," BMI is "obese" or the a1c is "low." The double-pipe (||) represents "or," and "%is%" represents a fuzzy "is" operator. The rules are relatively readable as English sentences.

The second rule specifies that the rating will be ST (standard) if BMI is “over,” a1c is “norm,” or bp is “pre.”

Finally, the third rule states that the rating will be PF (preferred) if BMI is “fit,” a1c is “norm,” and bp is “norm.” Notice here that we use the “and” operator, represented by the double ampersand (&&). Of course, this is just a simple example. The above rules are not meant to define an actual underwriting system.

Now that the rules and linguistic variables have been defined, we can build a system. This is done with the following R code. You can also “print” and “plot” this system.

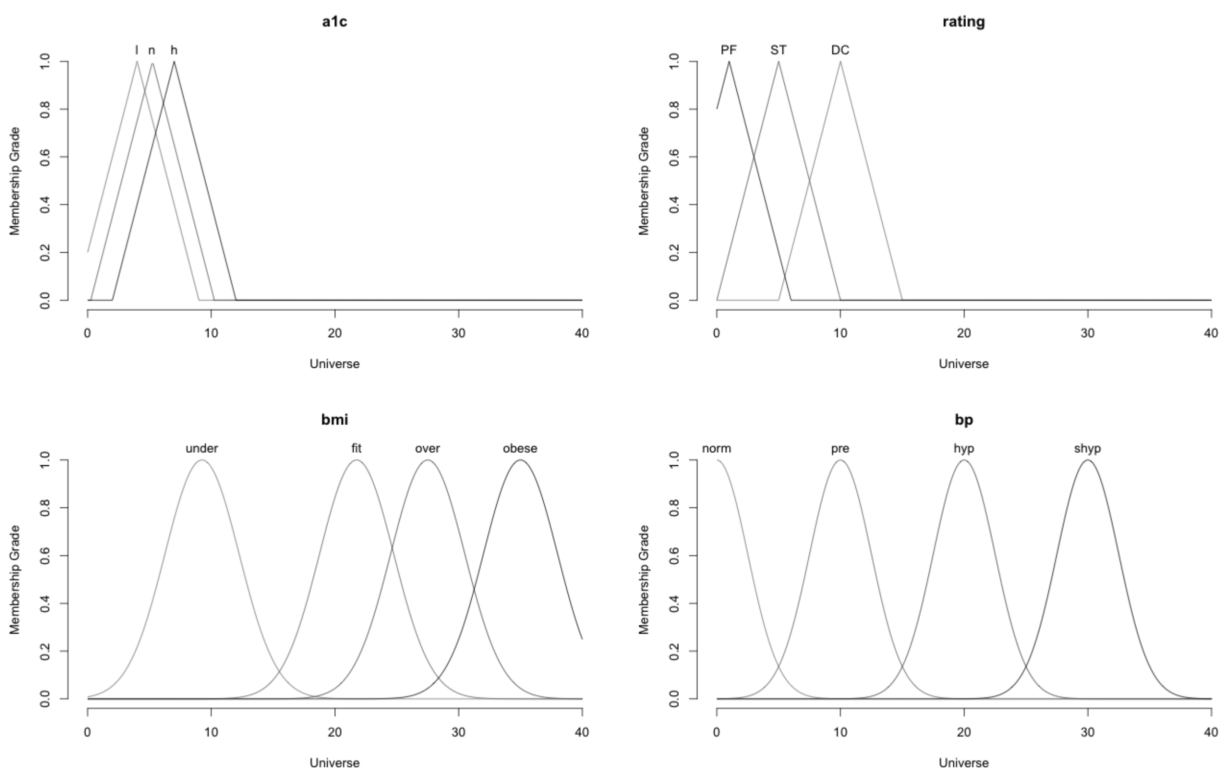
```
system <- fuzzy_system(variables, rules)

print(system)

plot(system)
```

The plot of this system can be seen in Figure 1.

Figure 1: Linguistic Variable System



CONTINUED ON **PAGE 38**

FUZZY INFERENCE AND DEFUZZIFICATION

We can now infer underwriter ratings from the above system. The process of fuzzy inference allows us to specify values for a1c, rating and bmi. This will give us a percent membership in the "rating" linguistic variable. Consider a proposed insured with a BMI of 29, a1c of five, and bp rating of 20. The following command would infer the rating into the variable "fi."

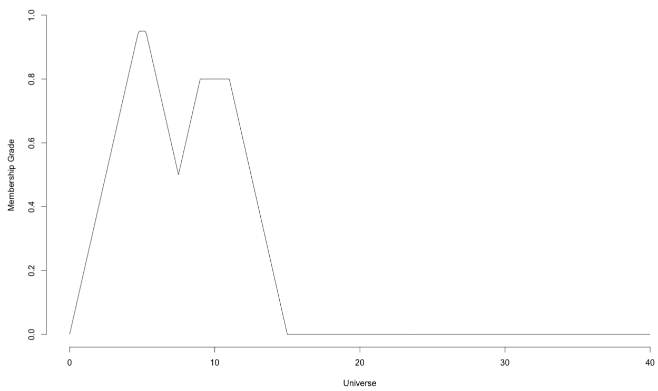
```
fi <- fuzzy_inference(system, list(bmi = 29, a1c=5, bp=20))
```

There is not a single value for "fi"; rather, it is a percent membership in each underwriter rating. The following command plots this to a chart.

```
plot(fi)
```

This chart can be seen in Figure 2. You will notice that the chart has two different membership peaks. One is near 7 and the other near 10.

Figure 2: Inferred Rating Membership



Defuzzification is the process where this membership is taken back to an actual number. There are several algorithms for defuzzification. The following command performs a defuzzification. The defuzzified rating is 7.445238.

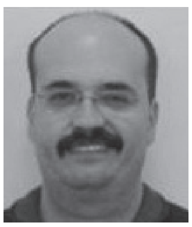
```
gset_defuzzify(fi, "centroid")
```

Once you have completed your inferences, it is considered good practice to clear the fuzzy sets. This is done with the following command.

```
sets_options("universe", NULL)
```

CONCLUSIONS

Fuzzy logic offers many advantages over the more traditional "crisp logic" that most computer programs are composed of. Because rules are inferred, it is not necessary to create the vast number of rules that most traditional rule engines grow into. The R source code for my example can be found at this link: *http://www.soa.org/news-and-publications/newsletters/forecasting-futurism/default.aspx*. If you would like to read more about the R "sets" package, you can visit its home page at the following link: *http://cran.r-project.org/web/packages/sets/sets.pdf*



Jeff Heaton

Jeff Heaton, is EHR data scientist at RGA Reinsurance Company and author of several books on artificial intelligence. He can be reached at *jheaton@rgare.com.*

SOCIETY OF ACTUARIES Forecasting and Futurism Section

NONPROFIT
ORGANIZATION
U.S. POSTAGE
PAID
SAINT JOSEPH, MI
PERMIT NO. 263

Forecasting & Futurism

N E W S L E T T E R

475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
p: 847.706.3500 f: 847.706.3599
w: www.soa.org