

Book Review

Review of the book John Kay and Mervyn King, *Radical Uncertainty: Decision Making Beyond the Numbers*, W. W. Norton and Co., New York, 2020

Vladik Kreinovich*

Department of Computer Science, University of Texas at El Paso, El Paso, TX, USA

Warning about this review. This is not a technical book, it does not have a single formula, but it has many ideas. What this review does is translates some of these ideas into the language familiar to us – researchers in uncertainty and fuzziness.

Will the authors of the book agree with this interpretation? Probably they would not even understand it: the authors never mention fuzzy logic, and even when they talk about probabilities, they are sometimes confused: e.g., they are criticizing a financial CEO for claiming (actually, correctly) that the stock price experienced a change which was 25 times larger than the standard deviation – the authors erroneously claim that such deviations are not possible, while in reality, they are only impossible for normal distributions.

We could get snobbish and dismiss the whole book based on its few mathematical mistakes. But Zadeh taught us – and may of us, who have tried collaborating with scientists and engineers know from experience – that many people who are not very mathematically educated can (and do) have great ideas. Our goal is not to dismiss these ideas, but to help their authors describe them in correct terms. This is what this review is trying to do.

The book has many ideas, I am sure this review misses many of them – and hopefully, other readers will help describe these other ideas in precise terms.

What is radical uncertainty? First answer: it is imprecise probabilities. The book claims that it is about radical uncertainty. What is it?

The book provides several answers to this question. The first answer is related to the fact that in many application areas, people use traditional statistical techniques, techniques that assume that we know the exact probabilities of different alternatives. One such application area is economics.

In practice, we rarely have full information about the probabilities, there are many possible probability distributions which are all consistent with our observations. However, since the traditional techniques requires us to use some probability distribution, we select the most reasonable one and use it.

How can we select a distribution? A natural idea is to take into account that different probability distributions have different uncertainty – which is numerically described by them having different entropy values. A natural idea is to preserve the uncertainty of the situation. For example, if all we know that the corresponding quantity is distributed on the interval $[0, 1]$, then it is, in principle, possible that this quantity is equal to 0.5 with probability 1 – but this particular distribution does not reflect the original uncertainty. It is therefore reasonable to select, among all possible distribution, the one with the largest possible value of uncertainty – i.e., the one for which entropy is the largest possible.

In particular, if all we know is the mean and the standard deviation, then out of all possible probability distributions with these values of mean and standard

*Corresponding author. Vladik Kreinovich, Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA. E-mail: vladik@utep.edu.

deviation, the distribution with the largest entropy is the normal (Gaussian) one. This is one of the reasons why normal distributions are actively used in many applications areas.

Often, this works well, but in some cases, the use of normal distributions leads to a disaster. One such example is economics. One of the reasons for the 2008 crisis was that economists assumed that random fluctuations are Gaussian, and thus, deviations larger than 6 standard deviations are practically impossible. So, when the stock went down by 25 standard deviations, this was a phenomenon that previous models – based on Gaussian distribution – could not predict, and, what was worse, situations for which Gaussian-based models could not help.

The presence of such big deviation shows that the probability distribution is different from Gaussian. This and similar situations made the authors conclude that in situations when we do not have full knowledge of a probability distribution, we should not select one of the possible distributions, we should consider all possible distributions as possible scenarios. This idea – whose technical name is *imprecise probabilities* – is what the authors call *radical uncertainty*.

What is radical uncertainty? Second answer: it is computing with words. How can we describe this radical uncertainty? A natural idea is to describe it by numbers, be it bounds on probabilities, or – as in fuzzy logic – our degrees of certainty that some values are possible. But how accurate are these numbers?

Yes, we can, e.g., describe our uncertainty in degrees by making degrees themselves fuzzy numbers – what is called type-2 fuzzy logic – but again, we need to assign some numerical degrees. So we can use type-2 logic, we can use type-3 logic to take care of this uncertainty – but no matter what we do, we only get an approximate description of what we have in mind.

How can we get an accurate description? Well, if you ask an expert doctor how he treats patients, if you ask an expert driver how she drives a car, these experts will not give numerical answers, they will provide answers in terms of words from natural language. We in fuzzy community understand this very well – the fact that an important part of our knowledge is described in terms of natural language was one of the main motivations for Zadeh’s fuzzy logic.

But here the book authors deviate from Zadeh: while Zadeh was looking for ways to describe this knowledge by using numbers, the authors recommend dealing directly with words – the idea that is

very similar to Zadeh’s vision of computing with words. This is what the authors mean by “going beyond numbers”.

What the authors emphasize is that they do not mean computing with words as it is often described in fuzzy-related publications – we translate from natural language into type-1 or type-2 fuzzy, deal with numbers, and then translate back into natural language. What they encourage is a development of techniques that would process words directly, without using numbers at all. Ok, maybe we can use numbers to motivate, to understand – but at the end, decisions should be made by processing words and texts directly. How? Well, this is a vision, not a technical text.

They do, however, show a pathway to this vision. Since the book’s emphasis is on decision making, the authors believe that we need to analyze – and emulate – how people make successful decision by communicating in natural language, by exchanging imprecise narratives and ideas.

Going beyond numbers does not mean numbers are useless: they are needed to make AI explainable. According to the authors, we should not make decisions based only on numbers, but this does not mean numbers are useless.

Many situations are very complex, and related decisions are very complex. Often, such situations are based on expertise of a large number of people from different competence areas. No single person can describe all the motivations for this decision – just like no single person can explain all the details of design and production of a modern airplane: some can explain the engine design, others the seat arrangement, yet others anti-fire precautions, etc.

To get the resulting complex group-decision accepted, it is important to supplement this complex decision with a simple explanation – an explanation which is simplified but understandable to everybody. This explanation can use words, it can use numbers – and it must be hierarchical: for each of the somewhat fuzzy premises, we can ask the Why question and get a more detailed (and more accurate) explanation, etc. In such simplified explanations, numbers are often irreducible – we all have intuition about linear, quadratic, exponential etc. dependencies, intuition that cannot be easily replaced by natural language words. This is why formulas were invented in the first place: it is easy to understand the formulation of the Pythagoras theorem when it is presented in the usual form $c^2 = a^2 + b^2$, but when described by words –

how it was in earlier times – it becomes almost incomprehensible: “the square of the hypotenuse is equal to the sum of the squares of the two sides”; we can memorize this formulation like we memorize poems, but it does not make this formulation easy to use.

The idea that we need a simple explanation may sound innovative for economics, but this is the usual way how physics operates: no matter how complex the problem, the physicists analyze the problem, understand what are the most important factors and which factors can be, in the first approximation, ignored, and come up with back-of-the envelope estimates – that usually provide a very good first-approximation description of real-life phenomena. This can be (and is) later supplemented by detailed complex calculations, but the explanation remains.

There is a known story about Einstein’s widow: when she was invited to visit a modern observatory with complex telescopes and powerful computers for processing data, the hosts explained that this is all done to understand how the Universe functions. To this explanation, she replied that her husband used to answer this question by using the back of the envelope.

To us in fuzzy, the importance of computing with words is clear – and it is clear to many folks who applied fuzzy techniques in engineering, but, interestingly, in economics, these methods are still practically not used. This is an important area of potential research and applications for us.

Difference between risk and uncertainty: an important aspect of decision making. Since this book is focused on decision making under uncertainty, the authors devote some time to differentiating uncertainty from risk. A naive mathematical approach equates risk with uncertainty, but from the commonsense viewpoint, they are different – and it is not always clear how to properly describe this difference.

One important example that the book provides relates to decision making. Most decision making models assume that rational people usually want to minimize risks – so if two stocks have the same expected return but different standard deviations, people should always prefer the stock with the smaller standard deviation.

This indeed works when standard deviations are high, but when they get smaller, many people have

the opposite preference: they prefer an alternative that has a slightly larger standard deviation to the one that has practically no standard deviation at all.

Why? (Notice that we follow the book’s recommendation of always asking the Why question :-). At first glance, this seems counterintuitive – it looks like one of the examples of irrational human behavior to add to the ones listed by Kahneman in his 2011 book (*Thinking, Fast and Slow*) and in his papers. But the authors explain that this behavior is quite rational. Indeed, we want to invest in a financial instrument which is most probable to survive possible future shocks. Big shocks are rare, but small events that affect the stock’s value happen all the time. As a result of these small events, economies fluctuate a little bit: it is reasonable to immediately react to changes. If an instrument does not fluctuate, this means that its reaction to these changes is too delayed – and so, in the case of the big shock, it may react too late and lose all its value.

In general, it is important to be flexible: trees and other plants move in the wind, this shows that they are flexible, and this allows many of them survive very strong winds, winds that topple inflexible poles. Similarly, a building that is flexible and whose top moves a little bit under a strong wind or under a small earthquake is more probable to survive a strong earthquake. On the other hand, inflexible systems may not survive a big change – like the dinosaurs did not survive the asteroid.

When explained this way, the danger of low standard deviation (or, in economic slang, low volatility) becomes clear, but, interestingly, this danger is not taken into account in the current economic models.

Warning about the book. Hopefully, this review raised your interest in this book. This is great, but may I repeat the warning with which I started this review: this is not a technical book, this book has no formulas at all, it is a book full of interesting ideas – sometimes, ideas described in not always accurate and correct mathematical form.

But, as the famous line from the movie “Some like it hot” says, “Nobody’s perfect”. Would I prefer a book that is mathematically correct but has no new ideas? Absolutely not. We can correct mathematical mistakes and make ideas shine, but if there are no ideas, the book is hopeless.

So, enjoy the book – but beware!