# Soft Decision Analysis

Thomas Whalen
Robinson College of Business
Georgia State University
Atlanta, Georgia 30303-3083 USA

## 1. Soft Computing

In contrast to the traditional computational methods which regard precision as a *sine qua non*, soft computing tolerates, indeed welcomes, imprecision, uncertainty and partial truth. Imprecision and uncertainty are pervasive, and precision and certainty carry a cost which often outweighs their benefits even if achieving them (as opposed to merely assuming them) is possible at all. Soft computing exploits the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost.

Computing with words essentially means using a high-level language which lets the humans who develop, use, and maintain a computer program view the objects of computation in terms of quantitative linguistic variables expressed by words and propositions drawn from a natural language, e.g., small, large, far, thin, not very likely, low and declining, near San Francisco, etc. Humans can perform a wide variety of physical and mental tasks without any measurements and any conscious computations: driving and parking a car, playing golf, riding a bicycle, understanding speech and summarizing a story.

In Zadeh's view, this capability is essentially a matter of the brain's crucial ability to manipulate perceptions -- perceptions of distance, size, weight, color, speed, time, direction, force, number, truth, likelihood and other characteristics of physical and mental objects. He proposes computing with words as a foundation for a computational theory of perceptions. Developing such a theory can lead to great advances in our ability to understand how humans make perception-based rational decisions in an environment of imprecision, uncertainty and partial truth. Such insight can help us to train a new generation of managers to do these things even better for their organizations.

In business decision making, the moment of decision hinges on the perception that a course of action is "good enough." Simon [1949] refers to this phenomenon as "satisficing." Alternatives are examined in sequence until one is found that is rated as acceptable. At this point, the acceptable alternative is selected and the decision maker moves on to the next problem rather than pursuing the theoretically optimal alternative. The criterion of acceptability is flexible and dynamic; in other words, it is a fuzzy perception of acceptability, not a precise measurement.

Rather than focusing primarily on the automation of decision processes, the goal of the new management science should be the humanization of decision processes. Stakeholders in an organization are best served when the managers are free to make important decisions that confidently reflect the real goals of the organization. Three broad "troubles" can impair this freedom and confidence: the press of too many trivial decisions; information pathologies such as ignorance, misperception, or overload; and the need to compromise real perceptions of the situation and goals in order to fit overly restrictive tools. Soft computing offers powerful ways to automate the decisions that ought to be delegated to a sufficiently capable machine, to manage information in support of truly perceptive decision making, and to fit the tool to the problem rather than vice versa.

## 2. "Soft" Elements of Traditional Decision Analysis

From a standpoint of "computing with words," standard methods of decision analysis begin by converting the natural perceptions of how usual or unusual the possible states are into Bayesian estimates of subjective probability measured on a ratio scale, and converting the natural perceptions of how acceptable or unacceptable the possible outcomes of state-action pairs are into utility numbers on an interval scale using Von Neuman-Morgenstern [1947] utility theory or some similar method. Thus, these two elements of traditional decision analysis can be considered to belong at least to the borderland of soft techniques.

## 2.1 Uncertain Probabilities

**Bayes' Theorem** was developed a very long time ago (before 1763) by the good Reverend Thomas Bayes, who was a talented dabbler in mathematics and statistics as a hobby. It was published shortly after his death in the Philosophical Transactions of the Royal Society, with a foreword by his friend Richard Price, who remarked that Bayes' theorem was important because it applies to a problem posed by De Moivre: "To show what reason we have for believing that there are in the constitutions of things fixt laws according to which events happen, and that, therefore, the future of the world must be the effect of the wisdom and power of an intelligent cause; and thus to confirm the argument taken from final causes for the existence of the Deity." Unfortunately, Bayes was so far ahead of his times that his article sat in the prestigious Transactions of the Royal Society virtually unread until the mid 1930's.

The reason we need Bayes Theorem is because the two quantities P(A|B) and P(B|A) are not equal.

For example, consider an oil prospector who faces the decision whether to drill for oil on a site based on the information he already has, to let the site go, or to pay for a seismic test and drill if the seismologists' prediction is positive. He expects a profit of $4 Million if he drills and finds a major oil strike, a profit of $1 Million if he drills and finds a Minor dtrike, and a loss of $1 Million if he drills and finds a Dry hole.

To be specific, suppose our prospector's current knowledge leads him to extimate P(Major)=.40, P(Minor ) =0.20, and P(Dry ) = 0.40. Note that the prospector doesn't, and cannot, know the "true" probability of a major strike; after all, if the oil is down there it has been there for millions of years (probability = 1.0) and if it's not there it never has been and never will be (probability = 0). The meaning of subjective probability is that the prospector would be indifferent between a side bet that pays off if there is a major oil strike at his site, and a side bet with an known probability of success of exactly 40%.

Based on these probabilities, his "expected" profit for drilling without doing a sesmic test is .4*4M + .2*1M - .4*1M = 1.4 Million dollars.

Geologists have been doing oil exploration work for many years; methods have continued to improve and standardize, and the petroleum exploration industry is very familiar with their track record. There are many instances on record of how drilling attempts have turned out, and on what the geological predictions submitted before drilling had said. That means that it is possible to have a very good idea how accurate the predictions typically turn out to be. The information we are much more likely to be able to get is *P(Positive|Major)*, which amounts to the proportion of predictions preceding major oil finds that were positive. This form, while it is more indirect, is more valuable because it focuses on the underlying ability of the geologists while abstracting away the particular characteristics of the unique mix of sites they've been presented with in the past. The probability that the seismic test will be positive given that there is a major strike of oil waiting to be drilled is not the same as the probability of a major strike given that the seismic test was positive.

The thing that makes Bayes theorem possible is the fact that the following four quantities are all equal:

P(A|B)P(B) = P(A&B) = P(B&A) = P(B|A)P(A)

For example,

P(Major|Positive)xP(Positive)= P(Positive|Major)xP(Major).

We can use algebra to solve this for the probability of a major oil strike given that the seismic test was positive:

$$P(Major|Positive) = \frac{P(Positive|Major) \times P(Major)}{P(Positive)}$$

Our plan is now to get P(Positive|Major) from the historical record of the geological survey companies, calculate the overall probability of a positive test P(Positive), and use these to calculate the desired posterior probability P(Major|Positive) as well as the rest of the probabilities needed to fill out the decision tree.

We should also note that this implies that Major is *dependent* on Positive. If Major and Positive were *statistically independent* then we would find that P(Major|Positive) = P(Major), in which case Positive would *not* be a conditioning event. It doesn't matter whether Major causes Positive, Positive causes Major, they are both caused by a common factor (e.g., oil), or whether the situation is some complex mix of these, but there has to be <u>some</u> kind of connection in order to support inference.

Let's reconsider the probabilities for the oil prospector's situation. He really would have begun with 2 kinds of probability information:

***Prior Probabilities*** represent his initial beliefs or knowledge about the presence or absence of oil. *P(Major)* is one such probability. These prior probabilities may be objective or subjective probabilities, but are often subjective.

***Likelihoods*** are conditional probabilities that summarize the known performance characteristics of the tests. They are more typically objective, often based on relative frequencies. Whatever kind of testing may be involved in a problem, the conditional probabilities usually summarize how reliably such testing has performed in the past. For example, *P(Positive|Major)* should represent, for those past cases where the geologists did a study before a major oil find, the proportion of the time that study was positive.

That is the information available to us at the beginning. We want to use those probabilities to give us an indication of how testing will perform in *this particular case*. We want to use the above information to give us a revised view of the likelihood of the various outcomes of drilling after we know what the geologists prediction says. Just as *prior* means *before*, *posterior* means *after*. *P(Major|Negative)* is a ***Posterior Probability***, sometimes grandly referred to as a ***Bayesian Revised Probability***. In other words, once we know the outcome of the testing, it makes sense for us to revise what we believe to be the probability of a major oil find.

One useful way to look at the information involved in applying Bayes' Theorem is to describe the probabilities involved as:

P(Event) Our ***prior probability*** of an outcome that is our main interest. These probabilities, objective or subjective, reflect the circumstances of ***this particular situation***.

P(Prediction|Event) These ***likelihoods*** describe the expected performance of our source of additional information. They may arise from the "track record" of a market researcher, the past performance of a testing procedure, or from physical reality (as in counting black and white balls in the urns that statisticians so dearly love). In any case, these probabilities do not reflect the current situation at all. They describe the ***predictive power of the information source***.

P(Prediction & Event) These ***joint probabilities*** combine the above two kinds of probability information to reveal the probability that ***in this situation***, our information source will make a particular prediction ***and*** a particular outcome will happen.

P(Prediction) Sum the joint probabilities involving this prediction and you get this ***marginal probability*** that ***in this situation***, our information source will make this prediction.

P(Event|Prediction) The ***posterior probability*** that the event will occur, given a particular prediction. Given a prediction, this replaces our prior *P(Event)* in the decision process. These probabilities are what we seek in applying Bayes' Theorem.

Note that the prior probabilities and the marginal probabilities are unconditional. They do not depend on any particular event having occurred. The likelihoods and the posterior probabilities ***are*** conditional probabilities, but the conditioning goes in opposite directions!

Our oil prospector's ***Prior Probabilities*** were:

P(Major) =0.40 P(Minor ) =0.20 P(Dry ) =0.40

He got his ***Conditional Probabilities*** from the performance record of the firm he expected to hire, *Valdez et Cie*. *Valdez et Cie*.'s track record was

| ***Valdez et Cie*.** | Cases where Site later turned out to be a Major strike | Cases where Site later turned out to be a Minor strike | Cases where Site later turned out to be a Dry hole | Total Cases |
|---|---|---|---|---|
| Positive Prediction | 36 | 27 | 12 | 75 |
| Negative Prediction | 4 | 18 | 28 | 50 |
| Total Cases | 40 | 45 | 40 | 125 |

The next step is to convert these frequencies to likelihoods by appropriate divisions:

| ***Valdez et Cie*.** | Cases where Site later turned out to be a Major strike | Cases where Site later turned out to be a Minor strike | Cases where Site later turned out to be a Dry hole | Total Cases |
|---|---|---|---|---|
| Positive Prediction | $^{36}/_{40}$= .90 | $^{27}/_{45}$=.60 | $^{12}/_{40}$= .30 | 75 |
| Negative Prediction | $^{4}/_{40}$= .10 | $^{18}/_{45}$= .40 | $^{28}/_{40}$= .70 | 50 |
| Total Cases | 40 | 45 | 40 | 125 |

Now we need to construct the ***Joint Probabilities*** associated with the kind of prediction our oil prospector will get and the results he will receive if he drills.

P(Pos&Major)=
P(Pos|Major)P(Major) =
.9*.4= .36

P(Pos&Minor)=
P(Pos|Minor)P(Minor) =
.6*.2= .12

P(Pos&Dry)=
P(Pos|Dry)P(Dry) =
.3*.4= .12

P(Neg&Major)=
P(Neg|Major)P(Major) =
.1*.4= .04

P(Neg&Minor)=
P(Neg|Minor)P(Minor) =
.4*.2= .08

P(Neg&Dry)=
P(Neg|Dry)P(Dry) =
.7*.4= .28

These results will let us construct the ***Marginal Probabilities*** for the results of the geological studies.

P(Pos) = P(Pos&Major) + P(Pos&Minor) + P(Pos&Dry) = .36 +.12 +.12 =.60
P(Neg) = P(Neg&Major) +P(Neg&Minor) +P(Neg&Dry) = .04 +.08 +.28 =.40

Now we have everything we need to apply Bayes' Theorem to find the ***Posterior Probabilities*** that we need to evaluate the worth of the seismic test.

Let's do just one ***Posterior Probability*** the long way, then we'll see how to shortcut the rest.

$$P(Major|Pos) = \frac{P(Pos|Major)\times P(Major)}{P(Pos|Major)\times P(Major) \;+\; P(Pos|Minor)\times P(Minor) \;+\; P(Pos|Dry)\times P(Dry)}$$

$$P(Major|Pos) = \frac{.90\times .40}{.90\times .40 \;+\; .60\times .20 \;+\; .30\times .40}$$

$$P(Major|Pos) = \frac{.36}{.36 \;+\; .12 \;+ .12} = \frac{.36}{.60} = .60$$

Looking at that last calculation, you should be able to recognize that there is really an easy shortcut available to us, since we have already calculated all the ***Marginal*** and ***Joint Probabilities***. We could have done it like this:

$$P(Major|Pos) = \frac{P(Major\&Pos)}{P(Pos)} = \frac{.36}{.60} = .60$$

Now let's do all the rest of the ***Posterior Probabilities*** using the shortcut:

| | | |
|---|---|---|
| $P(Major|Pos) = \frac{P(Major\&Pos)}{P(Pos)} = \frac{.36}{.60} = .60$ | $P(Minor|Pos) = \frac{P(Minor\&Pos)}{P(Pos)} = \frac{.12}{.60} = .20$ | $P(Dry|Pos) = \frac{P(Dry\&Pos)}{P(Pos)} = \frac{.12}{.60} = .20$ |
| $P(Major|Neg) = \frac{P(Major\&Neg)}{P(Neg)} = \frac{.04}{.40} = .10$ | $P(Minor|Neg) = \frac{P(Minor\&Neg)}{P(Neg)} = \frac{.08}{.40} = .20$ | $P(Dry|Neg) = \frac{P(Dry\&Neg)}{P(Neg)} = \frac{.28}{.40} = .70$ |

As you can hopefully now see, Bayes' Theorem helps us in Decision Analysis in a fairly special set of circumstances:

- You have the opportunity, usually at a price, to get additional information before you must commit to an Alternative Action.
- You have likelihood information that describes how well that source of information should be expected to perform (often based on how it has performed in the past).
- You wish to revise your prior probabilities (the probabilities you would be forced to use if you couldn't get the additional information).

Don’t be misled by the atmosphere of mathematical precision attending these calculations; remember that their foundation is on the subjective prior probabilities that ultimately express the decision maker’s propensity to bet. Thus, for example, P(Major|Neg) = .10 means that, if (and only if) the prospector would be indifferent between a side bet on a major strike and a side bet with a known probability of 40% when he didn’t have seismic information, then he should rationally be indifferent between a side bet on a major strike and a side bet with a known probability of 10% when he possesses the results of a negative seismic test.

Finally, with these probabilities in hand we ccan calculate the prospector’s expected return given a Positive or a Negative Test. If the test is Positive, his expected return is .6*4M + .2*1M - .2*1M = $2.4 Million. If the test is negative, his expected return for drilling is .1*4M + .2*1M - .7*1M = -$0.1 Million. What the latter number really means is that, if he gets a Negative seismic test, he will not drill at all, for a return of zero. (Note that this ignores any sunk costs which are irrelevant to decision making because they’re the same for all outcomes currently under consideration.)

**Value of Imperfect Information**

We looked earlier (Churchill & Whalen, 2001) at the EVPI (Expected Value of Perfect Information), a convenient fiction. Here we have something more realistic, sometimes called Imperfect Information, and sometimes called Sample Information. The Expected Value of Sample Information (EVSI) represents the difference between our best EMV for the entire tree if sample information had not been available and the EMV we actually got. If the cost of the information is known, then we usually measure the value in excess of the cost; this is the Net EVSI. If the cost of the information is unknown (or negotiable!), then we just measure the value as though we could get the information for free; the Gross EVSI. That gives us a standard against which we can later compare the price at which someone might offer us the information.

The value of information almost always varies as a function of time. The correct time (that is, the correct point in the decision tree) at which to determine the real value of information such as the seismic test is at the time that we must decide whether to purchase the information or not.

The oil prospector’s expected return if he could get the seismic test for free would be P(Pos)*[ expected return for best bet given Positive] | P(Neg)*[ expected return for best bet given Negative] = .6*$2.4M + .4*$0 = $1,440.000. We have already seen that, without any seismic information, he could just drill now for an expected return of $1,400,000. Thus, the value of the information contained in the seismic test is $40,000. If he can get the test performed for a cost of $30,000 he should do it, but if the test costs $50,000 he should say “no thank you” and just take his chances drilling.

2.2 Utility Theory or “Knowing What You Want”

**Risk Aversion & Declining Marginal Utility**

--Driving a $10,000 car is more like driving a $60,000 car than it is like walking.

--Eating canned food at home every day is more like eating in a five-star restaurant every day than it is like starving.

--Living in a cheap apartment is more like living in a mansion than it is like living on the street.

--Using a $1000 personal computer is more like using a $5 million supercomputer than it is like using a pencil and paper.

In each case, if we take the dollar value of the three possibilities, the expected monetary value of a 50-50 chance between the best and the worst is substantially higher than the monetary value of the middle alternative, but nearly everyone would prefer the middle alternative to the gamble.

This phenomenon, known as the "declining marginal utility of money," shows that money is not an adequate "medium of exchange" for decision making under uncertainty.

The declining marginal utility of money (and most other goods) is a purely economic reason why decision makers are risk averse -- that is, they prefer less risky ventures with smaller expected payoffs to riskier ventures with higher expected payoffs. In addition, emotional, personality, and situational factors can act to amplify or decrease this element, giving rise to a grater or lesser degree of net risk aversion or even, in the case of gambling for example, in a positive preference for risk.

Von Neumann and Morgenstern [1948] originated an approach to quantifying a decision maker's attitude towards risk in a way that is well suited to supporting the process of decision making under risk. To understand their approach, first consider the following two decisions:

**Decision 1: Which is better?**
**--Toss a fair coin and get $190 on heads or $10 on tails**
**--Get $99 for certain.**

**Decision 2:**
**Suppose there is a lottery with 1,000 tickets. One of these tickets will be drawn at random, and the person who holds the winning number will receive $1,000.**
**Which is better?**
**--Toss a coin and get 190 tickets on heads or 10 tickets on tails**
**--Get 99 tickets for certain.**

The answer to decision 1 is a matter of individual preference, depending on the degree of risk aversion of the person or organization involved. However, in decision 2, tossing the coin is objectively better since it gives a 10% chance at $1000 while the alternative gives only a 9.9% chance at $1000.

The fact that the answer to decision 2 is independent of individual risk attitudes is a the foundation of the VonNeumann-Morgenstern approach, which in effect uses the artificial, controlled risk of a reference lottery to insulate the decision process from the difficulties raised by the decision maker's attitude toward actual risks inherent in the venture under consideration.

**Basic Reference Lottery Tickets (BRLTs)**

In the context of a particular monetary decision or set of decisions, let

W = a dollar amount greater than or equal to the best potential payoff.
L = a dollar amount less than or equal to the worst potential payoff.

Suppose your preferences among the following three options are as shown, from most to least preferred:

-- receive $X + 1¢ for sure
-- take a 40% chance at W and a 60% chance at L
-- receive $X - 1¢ for sure

Then your utility for $X in this context is defined as 40%.

Now suppose there are 100 numbered lottery tickets. If you hold the winning number (selected at random, independently of the actual venture's outcome) you will receive Q, otherwise your will receive L (which may be negative!). Your chance of winning the lottery is identical if you have

1 ticket for sure,
10% chance at 10 tickets,
1% chance at 100 tickets, etc

Von Neumann and Morgenstern showed that you can take risk aversion into account by using the following procedure to analyze a decision under risk:

1. convert all the payoffs to utilities,

2. find the action or strategy with the best expected utility just as we formerly found the decision or strategy with the best EMV

3. convert the best expected utility to a dollar amount known as the Certain Monetary Equivalent: the amount of cash whose utility equals the . calculated expected utility of the selected action or strategy.

This "CME" expresses how much cash the venture is worth to a risk-averse decision maker in the same way the EMV expresses how much cash the venture is worth to a risk-neutral decision maker.

**Macroeconomics of Risk Aversion**

From the microeconomic standpoint of an individual person or enterprise, risk aversion can be quite rational; it recognizes declining marginal utility of the goods involved, it can be an appropriate measure to minimize emotional distress, and it can also minimize the probability of ruinous loses. The latter reason is important because, if a person or enterprise is prevented from taking advantage of future opportunities by a disastrous outcome to an earlier venture, the "long run" which is essential to portfolio theory ceases to exist.

However, from a macroeconomic standpoint, the overall wealth of the community is greatest if everyone in it behaves in a risk-neutral fashion, making all decisions so as to maximize the expected monetary value of monetary ventures and expected utility of nonmonetary ones. [1] It is for this reason (among others) that governments encourage various forms of risk sharing, since participants in schemes such as partnerships, corporations, and insurance arrangements share some or all of the gains and losses of many ventures. This brings the personal gains and losses of the participants closer to the average gains to the community at large.

[1] An exception would be ventures with potential negative consequences so great as to endanger the long run well-being of the entire community; such ventures are usually regulated by government action.

## 3. Ordinal Methods

A direct generalization of traditinal decision analysis is to use fuzzy numbers for probabilities and utilities, computing a fuzzy expected utility for each action by the extension principle of fuzzy mathematics. Since natural perceptions of usuality and acceptability are more likely to be in the form of words than numbers, the fuzzy method has the advantage of a more natural representation. [Watson et al, 1979]

However, sometimes it is not reasonable to assume that the perceived usuality of the states of nature can be converted to a ratio scale of probability, even a fuzzy one, without excessive distortion. Similarly, sometimes it is not possible to assign utility scores on an interval scale, crisp or fuzzy, to adequately represent the perceived acceptability or unacceptability of outcomes. In such a case, it is necessary to rely on the ordinal properties of the perceived usuality and acceptability.

One method, widely used in practice even if frowned upon in theory, is to ignore the uncertainty about the states of nature, and base the decision entirely on the one state that is the most usual or likely one. Obviously, the appropriateness of this approach is directly proportional to the degree to which usuality is in fact concentrated in one state of nature.

Another ordinal method represents the attractiveness of each alternative action by the least desirable outcome for that action. The alternative selected is the one for which this worst-case value is most acceptable or least unacceptable.

The above two ordinal methods consider either the maximal usuality alone or the maximal disutility alone. The remainder of this paper is concerned with methods that try to capture the extremes of both utility and possbility within the confines of ordinal calculation'

### 3.1. Possibilistic Decisionmaking

To illustrate the ordinal approaches discussed above, we will look at a simple abstract decision problem in which utility is granulated to good, fair, and poor, and usuality is granulated to usual, plausible, and rare. To illustrate the simplest form of possibilistic decisionmaking, it is necessary to place these quantities in a complete ranking. For simplicity, assume that the membership of “good” in the fuzzy set of good outcomes is equal to the membership of “usual” in the fuzzy set of possible states, and similarly for fair and plausible and for poor and rare. Then the ordinal calculations to assess the overall attractiveness of the five actions are as follows:

| Utilities | | | |
|---|---|---|---|
| actions: | Usual State | Plausible State | Rare State |
| A1 | good | fair | poor |
| A2 | good | poor | fair |
| A3 | fair | poor | good |
| A4 | poor | good | fair |
| A5 | poor | fair | good |

| Equivalent Disutilities: | | | |
|---|---|---|---|
| actions: | Usual State | Plausible State | Rare State |
| A1 | good | fair | poor |
| A2 | good | poor | fair |
| A3 | fair | poor | good |
| A4 | poor | good | fair |
| A5 | poor | fair | good |

(Assuming mD("Low") =mU("Good") = p("Usual") etcetera.

To carry out the actual analysis, convert good, medium, and poor utility into low, medium, and high disutility respectively and assume that the membership of “high” in the fuzzy set of bad outcomes is equal to the membership of “usual” in the fuzzy set of possible states, and similarly for medium and plausible and for good and rare. The ordinal calculations to assess the overall attractiveness of the six actions is shown in the next table.

| | Usual State | Plausible State | Rare State | Overall Risk |
|---|---|---|---|---|
| A1 | min(low, usual) = low | min(medium, plausible) = medium | min(high, rare) = rare | max(low, medium, rare) = medium |
| A2 | min(low, usual) = low | min(high, plausible) = plausible | min(medium, rare) = rare | max(low, plausible, rare) = plausible = medium |
| A3 | min(medium, usual) = medium | min(high, plausible) = plausible | min(low, rare) = rare | max(medium, plausible, rare) = medium |
| A4 | min(high, usual) = high | min(low, plausible) = low | min(medium, rare) = rare | max(high, low, rare) = high |
| A5 | min(high, usual) = high | min(medium, plausible) = medium | min(low, rare) = rare | max(high, medium, rare) = high |

Ranking actions from lowest to highest disutility, using the second place as a tie-breaker, we have A1 and A2 tied for best, followed by A3, A4, A5.

## 3.2. Ordinal Probability Approaches

### Fuzzy Relational Ordinal Risk Minimization

Whalen & Schott [2000] have suggested an approach to commensurate ordinal decision making based on recent insights in knowledge granulation, computing with words, fuzzy relations as fuzzy x-y graphs, and second order fuzzy sets. In this new approach, the disutility of a state-action outcome is represented by a linguistic variable. The base values of this linguistic variable may be well-defined utility numbers, or they may be a purely abstract ordered set. The usuality of a state is similarly represented by a linguistic variable defined on a base variable of degrees of possibility.

The **threat of a state-action pair** is represented by a fuzzy relation formed by the Cartesian minimum of the disutility and the possibility. This, the possibility that the disutility of a state-action pair equals x is itself a fuzzy set of possibilities; the membership of possibility p in this fuzzy membership is the lesser of the membership of p in the usuality of the state and the membership of x in the disutility of the state-action outcome.

The **risk profile of an action**, then, is formed by the union of all the fuzzy relations corresponding to the threat of all the state-action pairs for that action. The possibility that the disutility of a given action is x is a fuzzy set of possibilities formed as the union of the fuzzy possibility that each possible outcome has disutility x. The risk profile of the action is the second order fuzzy set that associates each disutility x with its fuzzy possibility. This is equivalent to a fuzzy x-y graph for each action; the knowledge granules making up this graph have fuzzy coordinates defined by the disutility (x) and usuality (y) of the possible outcomes.

The second order fuzzy set defining the risk profile of a given alternative action is reduced to a first-order fuzzy set using a Sugeno integral. The first order possibility that the disutility of an action is x is the maximum over possibility grades (0 to 1) of the minimum of each possibility grade with its membership in the second order possibility that the disutility of the action is x. If $\tilde{r}$ is the second-order fuzzy set representing the risk profile of a particular action A, and x is a particular base value on the disutility scale, then $\mu_{\tilde{r}}(x)$ is itself a first-order fuzzy set of possibilities representing the second-order possibility that the actual disutility of action A is x. $\mu_{\mu_{\tilde{r}}(x)}(p)$ is the membership of a particular crisp possibility grade, p, in this second-order possibility. $\mu(x)$, the first-order possibility that the actual disutility of action A is x, is calculated by

$$\mu_r(x) = \max_p \left\{ \min\left(p, \mu_{\mu_{\tilde{r}}(x)}(p)\right) \right\}.$$

The final stage in the process is to convert the first order fuzzy risk profile of each alternative action back into words to allow the user to exercise judgment as to what action should be chosen and what rhetorical argument to use to justify that choice. Wenstop's [1980] classic approach to linguistic approximation does not work well for this purpose because the possibility distributions tend to be multimodal, so a method based on linear integer programming is used in this paper for demonstration purposes. Future research will evaluate the effectiveness of genetic algorithms and other methods for linguistic approximation in the context of risk minimization using computing with words.

### Quadratic Ordinal Psychophysical Optimization

Whalen and Wang [2000] apply quadratic programming to an ordinal interpretation of linguistic probability and utility terms. This interpretation incorporates a well-established finding of psychophysics: the degree to which stimuli must differ physically to be discerned perceptually is proportional to the magnitude for the stimuli according to a logarithmic law of human (and animal) perception.

If the utility of each state-action pair is known only roughly, for instance as "good," "fair," or "poor," we will represent this ordinally; each fair utility is greater than any poor utility and less than any good utility for a given attribute.

Over a century of research in the field of psychophysics indicates a very strong tendency for human perception to operate on a logarithmic scale, in which the "just noticeable difference" between stimuli is a constant proportion of the magnitude of the stimulus rather than a fixed incremental amount. Based on this, we require that each “fair” utility is greater than or equal to a fixed constant called the distinguishability ratio times any “poor” utility, and each “good” utility is greater than or equal to the same fixed constant times any “fair” utility. The distinguishability ratio is a generalization of the psychophysical concept of a decibel.

To represent the fact that two quantities with the same rough description need not be identical as long as their difference is not psychologically significant, the quantitative representation of each “fair” utility must be less than or equal to the distinguishability constant times that of every other “fair” utility, and similarly among “good” and “poor.” This automatically entails that each utility in a class is also greater than or equal to any other utility divided by the distinguishability ratio.

If the probability of each state is specified numerically, then the overall utility of an action is the sum of the unknown numeric representations of the utility of each of the action's state-action pairs, weighted by the numerical probability weights. This is a linear function with linear inequality constraints; thus, it is possible to find the maximum and minimum overall utility for each action by linear programming. More important, we can find the minimum and maximum difference between the overall utility of two actions. Alternative A dominates Alternative B if max{utility(A)-utiility(B)}<0; it is not sufficient that max{utility(A)}>max{utiility(B)} and min{utility(A)}>min{utiility(B)}, but Alternative A can dominate alternative B even if max{utility(A)}>max{utiility(B)}> min{utility(A)}>min{utiility(B).

A dominated action can be removed from consideration, leaving a short list of nondominated actions. These can be re-analyzed in several ways. One can use a finer grid of linguistic terms like “very low” or “upper medium” and repeat the analysis. One can also introduce additional inequalities into the linear programming formulation. For example, there may be two actions that both generate “low” utility in a particular state, but further introspection and/or economic analysis may indicate one is discernibly lower than the other. Finally, one can move to methods that are more decisive than the ordinal ones, but which require stronger assumptions.

If the utility of each state-action pair is specified numerically but the probability of the different attributes is only specified roughly, the linear programming problem is very similar to the one described above.

If utility is known only as good, fair, poor and probability is known only as low, medium, high, finding the maximum and minimum of the difference between two alternative actions, and thus identifying dominated alternatives, becomes a problem in quadratic programming. The difference in overall utility of between action i and action j is a weighted sum, $D_{ij} = \Sigma_k(U_{ik}*P_k - U_{jk}*P_k)$, in which each term is the product of two variables, the utility of action i in state k and the unknown probability, where these variables $U_{\bullet\bullet}$ and $P_\bullet$ are subject to a set of linear inequality constraints. If the maximum value of the quadratic function Dij, subject to the linear constraints, is negative then action j dominates action i.

## 4. Possibilistic Approaches

### 4.1. Commensurate L-Fuzzy Risk Minimization

Commensurate L-Fuzzy Risk Minimization ises L-fuzzy (variously rendered “lattice fuzzy” or “linguistic fuzzy”) sets [Goguen, 1967] for bad outcomes and possible states, with set memberships defined on an incompletely ordered abstract lattice. If all memberships are measured on a common lattice, it is possible in principle to take the minimum of the membership of a state-action outcome in the L-fuzzy set of bad outcomes and the membership of the corresponding state in the L-fuzzy set of possible states. But since the ordering is incomplete, it is not always possible to find an explicit minimum or maximum. Furthermore, it may not be meaningful to compare the membership of a state in the set of usual states with the membership of an outcome in the set of bad outcomes, since the two are qualitatively so different.

The Commensurate L-Fuzzy Risk Minimization technique deals with this by placing usuality set memberships on one incompletely ordered lattice and disutility set memberships on a separate incompletely ordered lattice. The version presented in [Whalen & Bronn, 1988] uses a (disutility, possibility) ordered pair to specify the membership of a state-action outcome in the set of outcomes that are both possible and bad. The risk of an alternative action is found by symbolically maximizing these pairs similarly to the way the previous method maximizes unresolved minima; the selected action is the one for which this symbolic structure is least risky. Recognizing that this process is not highly decisive, the method successively falls back to the L-fuzzy approach with a common incomplete lattice, and the revised possibilistic approach with a single complete weak order of membership grades.

### 4.2. Intuitionistic Decision Analysis

Intuitionistic decision analysis draws upon a parallel between the pair (belief, plausibility) in the Dempster-Shafer-Smets theory of evidence and transferable belief measures on the one hand, and the pair (membership, nonmembership) in the theory of intuitionistic fuzzy sets developed by Atanassof and his colleagues. Suppose we consider the "goodness" of the outcome of a (state, action) pair as its membership in the fuzzy set of good outcomes, and we consider the "badness" of the same outcome as its membership in the complement of the fuzzy set of good outcomes. According to the intuitionistic fuzzy set model, membership in the set of good outcomes is independent of its membership in the complement. Similarly, while the plausibility of a state must be greater than or equal to its belief, it is not completely determined by the belief in that state.

If we combine the goodness of a (state, action) pair with the belief of the corresponding action, we have a conservative measure of the positive effect of this potentiality on our well-being; its "expectation." Using the plausibility instead of the belief in connection with the goodness of the outcome gives a more optimistic measure, here called "hope." Pairing the badness of an outcome with the belief in the state gives its "cost," while pairing the badness with the plausibility yields a more pessimistic measure here termed "fear."

| | Plausibility of S | Belief in S |
|---|---|---|
| Goodness of (A,S) | Hope of (A,S) | Expectation of (A,S) |
| Badness of (A,s) | Fear of (A,S) | Cost of (A,S) |

Hope, expectation, cost, and fear all pertain to a (state, action pair) which may not occur (or may occur only partially). To evaluate a decision alternative, it is necessary to aggregate all the state-action outcomes for that alternative. the following table gives a suggested method for doing this; the supremum over states is taken for hope, fear, and cost, while the conservative measure of expectation is found by taking the infimum. The aggregates are given names based on Porter's [19xx] classic set of categories for evaluating a business enterprise or venture: Strength, Weakness, Opportunity, and Threat, often abbreviated as the SWOT model.[2]

| | |
|---|---|
| Opportunity of A = $\sup_S${Hopeof (A,S)} | Strength of A = $\inf_S${Expectation of (A,S)} |
| Threat of A = $\sup_S${Fear of (A,S)} | Weakness of A = $\sup_S${Cost of (A,S)} |

[2] Rowe, Mason, Dickel, Mann, Mockler; "Strategic Management: a methodological approach". 4th Edition, 1994. Addison-Wesley. Reading Mass.

## 5. Decision Making and Rhetoric

"Rhetoric is defined as "the art or study of using language effectively and persuasively." In addition to natural language discourse, the concept of “language” can include both the language of mathematics and modern methods of computing with words. “Using language effectively and persuasively” means leading someone towards making a desired decision;

In the context of decision analysis, the ‘desired” decision means a good decision rather than a foregone conclusion, and the person to be persuaded is oneself or one’s organization. We cannot base our search for a good decision on the outcome of the decision since we cannot know the future. Instead, we look for a decision that is backed up by a chain of evidence that is more compelling than that backing up any competing alternative. And such a chain of evidence, if it is not to be merely an inarticulate "gut feeling," must inevitably consist of an argument expressed in words, numbers, or some hybrid of the two. The art of constructing such arguments, viewed in its broadest conception, is the art of rhetoric.

The highest goal of decision analysis is to provide effective and persuasive elements that will inform the process of making good human decisions, not to automate human decisions away.

### References:


1. Simon. Herbert: "Theories of Decision Making in Economics and Behavioral Science," American Economic Review, 49:3, June 1959, pp. 253-80)
2. Zadeh, Lotfi: “From Computing with Numbers to Computing with Words: From Manipulation of Measurements to manipulation of perceptions,” Proceedings, Proceedings, Third International Conference on Applications of Fuzzy Logic and Soft Computing, Wiesbaden, 1998, p.1-2.
3. Zadeh, Lotfi: “A new Direction in Fuzzy logic: Toward Automated Reasoning with Perceptions,” Proceedings, Eighth International Fuzzy Systems Association World Congress, Taipei, 1999, Volume I, p. 1-5.