Risk Modeling for Policy Making

Ronald R. Yager

Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, USA
yager@panix.com

Summary. Supporting policy makers requires tools to aid in decision making in risky situations. Fundamental to this kind of decision making is a need to model the uncertainty associated with a course of action, an alternative's uncertainty profile. In addition to this we need to be able to model the responsible agents decision function, their attitude with respect to different uncertain risky situations. In the real world both these kinds of information are to complex, ill defined and imprecise to be able to be realistically modeled by conventional techniques. Here we look at new techniques arising from the modern technologies of computational intelligence and soft computing. The use of fuzzy rule based formulations to model decision functions is investigated. We discuss the role of perception based granular probability distributions as a means of modeling the uncertainty profiles of the alternatives. Tools for evaluating rule based decision functions in the face of perception based uncertainty profiles are presented. We suggest a more intuitive and human friendly way of describing uncertainty profiles is in terms of a perception based granular cumulative probability distribution function. We show how these perception based granular cumulative probability distributions can be expressed in terms of a fuzzy rule based model.

1 Introduction

Policy decisions run the gamut from taxation to health care to education to allocation of resources in combating terrorism. Almost all domains of human experience are effected by local, national or trans-national policy decisions. The support of decisions involving policy in most cases require tools to address issues related to a desire to satisfy multiple, often conflicting, goals and a need to negotiate between numerous, often adversarial, constituencies. In addition choices must be made in the face of uncertainty and associated risks. Further compounding any formal attempt to support policy decisions is the imprecision in much of the information provided by the participating agents. In this work we introduce some tools to address issues related to uncertainty and risk management. We are particularly concerned with problems inherent

R.R. Yager: *Risk Modeling for Policy Making*, Studies in Computational Intelligence (SCI) **117**, 1–18 (2008)
www.springerlink.com

in the imprecision of our knowledge of uncertainty and the imprecision in the characterization of the policy makers risk tolerance.

The need for risk management arises when we have to make a choice involving a risky alternative. One component of a risky alternative is the uncertainty of the payoff (outcome) resulting from its selection, there are more than one possible outcome. Making decisions in the face of uncertain outcomes requires some of representation of our knowledge of uncertainties associated with the possible outcomes, for example probabilities. Often this information is impossible to obtain precisely and may require an imprecise and fuzzy characterization. Here we shall take advantage of Zadeh's [1–4] work on perception based probability information.

A fundamental difficulty that arises when making decisions involving alternatives with uncertain outcomes is the comparison of the alternatives. This is do to the fact that the multiplicity and complexity of these types of the alternatives makes their direct comparison almost impossible. Here we use rule based valuation functions to circumvent this difficulty.

An additional feature that distinguishes a risky alternative from one that is simply uncertain is that at least one of its possible outcomes is bad, 'undesirable' or 'disturbing.' The concept of undesirable is fuzzy and often involves aspects of human perception. Let us try to provide some intuition. Consider a financial decision in which we can make a profit of either $50, $100 or $200. In this case while we have uncertainty with respect to the outcome and a preference for 200 over 100 over 50, we don't have a risky alternative because none of the payoffs are undesirable. On the other hand, consider an alternative with payoffs $\{-\$10,000,\ \$50,\ \$200\}$. This can be considered as a risky alternative because in addition to there being an uncertainty with respect to the outcome, it has at least one undesirable outcome. As another example we can consider is a person who has a non-life threatening medical disorder and is offered a treatment that can either cure his disorder or kill him. This can be clearly seen as a risky alternative. The determination of whether a particular outcome is undesirable is often subjective and context dependent. It is very much dependent on the current state of the decision maker, what in some situations would be considered as disturbing may in other situations not be considered disturbing.

A fundamental point that we want to make here is that the construction of decision functions involving these "risky" alternatives often involves some kind of categorization of outcomes with respect to their being undesirable or bad. From a formal point of view decision making with risky alternatives requires that the possible outcomes be expressed on a scale that is richer then an ordinal scale. The scale used must be of a bi-valent nature [5], having positive and negative members, and thereby enabling the capturing of concepts good and bad. An additional feature is that the concepts used to specify "bad" and "good" outcomes are generally fuzzy and imprecise.

We should note that in addition to comparing risky alternatives risk management involves another important aspect, the creation of new alternatives to better satisfy the needs of the participants. Since this process of alternative creation is generally domain dependent we shall not focus on this important issue. However the tools developed here can play an role in the part of risk management focusing on alternative creation.

2 Modeling the Valuation Function

One approach to addressing the problem of comparing alternatives having uncertain outcomes is to use a valuation function. These functions map the possible payoffs associated with an uncertain alternative into a single scalar value called its valuation. The association of a scalar value with an alternative allows us to easily compare alternatives. Conceptually these valuation functions can be viewed as a mechanism to enable the responsible decision maker to reflect their preferences among different uncertain situations. Statistics such as expected value, median and variance have historically been used to help provide valuation functions. With the consideration of risky alternatives the nature of the decision makers' preferences between different uncertain situations becomes more complex then can be captured by these simple statistics. In order to capture the decision makers preference in these situations we need more sophisticated structures for modeling the valuation functions.

One approach to modeling a decision makers preference structure, i.e. valuation function, is to use a rule based [6]. A rule base consists of a collection of statements, rules, each of which expresses the decision makers valuation (attitude) about a particular uncertain situation. The totality of these individual components constitutes the decision makers preference function. The use of a rule base allows a decision maker to express their preferences in a modular fashion. The facility of using a modular expression of their valuation greatly eases the task of formulating the function.

In Fig. 1 we see how this rule base (knowledge base) is used. An alternative is presented to the rule base which then provides a value for the alternative. The value V is some score associated with the alternative.

Fuzzy system modeling [6, 7] provides a well established framework for constructing these types of models used to capture the decision makers' valuation function in the form of a rule base. An individual component rule in the preference rule base is of the form

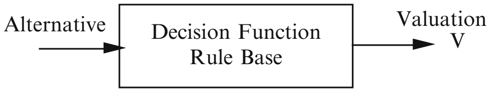

Fig. 1. Rule representation of decision function

$$\text{If } \textbf{antecedent} \text{ then V } \textit{is } \text{S}_i$$

where the term **antecedent** describes some characterization of a risky alternative. An example could be "if an alternative has a very bad outcome with a substantial probability of occurrence then give it a very low value."

In this approach we use predicates to construct the antecedent. Here we use Pred_i to indicate a predicate corresponding to some property or feature of an alternative. For any alternative A we can calculate $\text{Pred}_i(A)$, the degree to which A satisfies the predicate. The antecedent of a rule may consist of a single predicate or a collection of predicates connected by some logical or other aggregation procedure. Typically the antecedent can be expressed in terms of properties associated with surrogate features of the uncertainty profile of an alternative. Things like variance, probability of particular situations, expected values are examples of these features. The consequent of the rule, V *is* S_i indicates a valuation of an alternative that satisfies this rule.

Given a collection of rules[1]

$$R_i\text{: If } \text{Pred}_i \text{ then V } \textit{is } S_i$$

the general procedure for working with these rules is as follows. For the alternative A we calculate $\text{Pred}_i(A)$, the degree R_i is valid for this alternative. This gives us a collection of pairs $(\text{Pred}_i(A),\ S_i)$. We then aggregate these pairs to get an overall valuation for the alternative being valuated, $V(A) = \text{Agg}_i(\text{Pred}_i(A),\ S_i)$. The methodology used to aggregate these pairs depends upon the structure underlying the partitioning of the uncertainty profile space by the rules. We note in fuzzy systems modeling the most common aggregation is a weighted average

$$V(A) = \frac{\sum\limits_i \text{Pred}_i(A)\ S_i}{\sum\limits_i \text{Pred}_i(A)}$$

Our focus here shall be on the formulation and evaluation of some types of predicates needed to describe antecedents in these rule based models of valuation functions.

3 Valuation Functions and Uncertainty Profiles

Formally a risky alternative is characterized by an **uncertainty profile**. In part an uncertainty profile consists of a collection of possible outcomes (payoffs) that can occur as a result of selecting this alternative. We shall denote this collection of possible payoffs as X. In addition a uncertainty profile usually contains information about the realizability of each of the payoffs. A general

[1] Here for simplicity we assume the antecent is composed of just one predicate. As we noted more generally the antecent can involve multiple predicates.

framework for expressing this information can be had in terms of a monotonic set function $\mu : 2^X \to [0, 1]$ having the properties **1**. $\mu(\emptyset) = 0$, **2**. $\mu(X) = 1$ and **3**. $\mu(A) \geq \mu(B)$ if $B \subseteq A$ [8]. Here μ provides a measure of the belief of finding the actual payoff in the subset A. If as is often the case in many applications we assume μ is additive, $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A \cap B = \emptyset$ then μ is a probability measure.

In the following we assume that the measure associated with the uncertainty profile of an alternative is best captured by a probability model. Thus we are assuming that the payoff of a risky alternative is a random variable **R**. One of our concerns here is with the characterization of the features of this random variable that can be used as predicates in the antecedent of the rules used in the rule base definition of the valuation function. We must emphasize that the representation of the features used must be such that we can evaluate the degree of satisfaction of the associated predicate for an alternative given our knowledge of the uncertainty profile of the alternative. Well established features associated with a random variable are expected value, variance, model and median. A typical example of the use of these features in a rule based is the form

"If the expected payoff *is* ***high*** then V *is* ***good***"

Here the expected value is the feature being used. The predicate here is "the expected payoff *is* ***high***." Thus for a given alternative we must determine the degree to which this is true. Specifically if we have the uncertainty profile of the alternative expressed in terms of a random variable with known probability distribution we can calculate the expected value. With ***high*** expressed as a fuzzy set we can calculate the degree to which the predicate is satisfied. Another example would be a rule of the form

If the expected payoff is **high** and the variance is **small** then
V is **very good**.

Here our antecedent consists of two predicates connected by an "and." The second predicate, the "variance is **small**" uses as its feature the variance. Here then for a given alternative we would calculate its expected value and its variance from its uncertainty profile. We then calculate the satisfaction of each of the two predicates and then take the "anding" of these two values. Using results from multivalued logic [9] we could use the minimum of these values as the "and." It important to emphasize that with the use of predicates and these rules we have circumvented the issue of combining expected values and variances.

In policy making decisions in which we have risky alternatives the responsible decision maker's mental preference structure is generally more complex then that which can expressed simply using the basic features such as expected value and variance. Making decisions in risky environments require us to use more sophisticated features of an alternatives uncertainty profile.

One feature of an uncertainty profile that can play an important role in the formulating decision rules in the face of risky alternatives is the probability of some subset of payoffs. An example of a rule using this type of feature is

"If the probability of having a severe loss *is* **low** then the value of the alternative *is* **high**."

In this case the feature used in the rule is "the alternative's probability of having a severe loss." The predicate here is the degree to which this feature attains a value that is considered as **low**. The process of evaluating this antecedent predicate involves the following. We represent the concept "low probability" as a fuzzy subset, **LOW**, of the unit interval. If Prob(S) is the probability of having a severe loss under the alternative then the degree to which the predicate is satisfied is **LOW**(Prob(S)), the membership grade of value Prob(S) in the fuzzy subset **LOW**.

The issue now becomes that of obtaining Prob(S), the probability of having a severe loss under the alternative. The determination of this depends upon our definition of severe loss and our knowledge about the uncertainty profile associated with the alternative. Initially we shall assume complete information about the probability associated with the random variable, the uncertainty profile of the alternative. If **R** is a continuous random variable, we assume the availability of the probability density function f. If the random variable is discrete we assume the availability of the probability mass function. In addition to our knowledge of the uncertainty profile we need a definition of the concept of "severe loss." Here we can use fuzzy sets to help in the definition. More generally as we shall see the combined use of fuzzy sets with probabilistic information provides a very powerful way to express features that can play a role in constructing intelligent decision making functions. Let us look at this closer.

Consider the payoff random variable whose uncertainty is captured by its probability density function f(x). Let us calculate the "probability of a severe loss." In order to obtain this we first need a definition of the term "severe loss." We define the concept of a severe loss as a fuzzy subset S on X such that S(x) is the degree to which an outcome x satisfies the concept of being a severe loss. Using this definition and the probability density function f(x) we obtain the probability of a severe loss as [10]

$$\mathrm{Prob(S)} = \int_{\mathrm{R}} \mathrm{f(x)\ S(x)\ dx}$$

We note if S is a crisp subset then this becomes $\mathrm{Prob(S)} = \int_{\mathrm{x \in S}} \mathrm{f(x)\ dx}$. For example if S is defined crisply as "any payoff less or equal a" then $\mathrm{Prob(S)} = \int_{-\infty}^{\mathrm{a}} \mathrm{f(x)\ dx}$.

In similar manner we can define the concept of a large payoff as the fuzzy subset L obtain $\mathrm{Prob(Large\ Payoff)} = \int_{\mathrm{R}} \mathrm{f(x)\ L(x)\ dx}$. More generally if E is any linguistically expressed description of the payoff space which can be represented as a fuzzy subset E then we can obtain $\mathrm{Prob(E)} = \int_{\mathrm{R}} \mathrm{f(x)\ E(x)\ dx}$. We emphasize the subjective nature of the concept E and the related fuzzy subset E. This situation comes with positives and negatives. While this allows a user to introduce the concepts needed to describing their preferences it requires a definition be supplied either by the user or via some default supplementary mechanism.

Note: In the case in which the random variable describing the payoffs is discrete and captured by a probability mass P then $\mathrm{Prob}(E) = \sum P(x)\ E(x)$.

4 Perception Based Granular Probability Distributions

In the complex environment of policy making the information needed to fully detail the probability measure associated with an alternative's uncertainty profile may only be partially or imprecisely available.

Techniques such as the Dempster–Shafer theory of evidence [11] provide useful structures for representation of an alternative's uncertainty profile in the cases of lack of precise knowledge about the exact probability measure. Another approach recently developed by Zadeh [4] is rooted in the observation that much of the information appearing in an alternative's uncertainly profile is based upon the perceptions of the decision maker. In the light of this understanding Zadeh [4] has introduced the idea of **P**erception **B**ased **G**ranular (PBG) probability distributions to address situations in which we have less than perfect information about the uncertainty profile. We now consider the situation where this is the case.

Zadeh [4] observed that the type of probability information associated with an uncertainty profile is generally a reflection of perceptions as well as measurements by the decision making entity. He suggested that an appropriate way of representing this type of information is with a **P**erception **B**ased **G**ranular (PBG) probability distribution. With the aid of a PBG probability distribution the human can very naturally express their perceptions of an uncertainty profile. As we shall see a PBG probability distributions generalize the idea of ordinary probability distribution.

Let **R** be a random variable whose domain X is a subset of the real line. A PBG probability distribution consists of a collection of tuples $(A_i,\ Q_i)$. Within each tuple A_i is an imprecise element from the domain X of **R** represented as a fuzzy subset of X. Q_i is an amount of probability allocated to that range, generally having a imprecise linguistic nature and expressed as a fuzzy subset of the unit interval. For example if **R** takes its values in the interval $X = [-10 \text{ to } 10]$ then an example of a such a PBG probability distribution is

$$(\text{low, about } 0.5), (\text{near zero, about } 0.3), (\text{near } 10, \text{ about } 0.2)$$

In order to further discuss PBG probability distributions we must first distinguish between two types of situations regarding the underlying domains. The first is when X is a continuous subset of the real line, $X = [a,\ b]$, and the second is when X is discrete $X = \{x_1, \ldots, x_n\}$.

We first consider the case in which X is discrete. Here the underlying measure is a probability distribution P, whose actual values are unknown. The PBG probability distribution is providing partial information about the underlying probability distribution. Let us look at this situation. First we recall with $X = \{x_1, \ldots, x_n\}$ then a valid probability distribution P on X is a collection $[p_1, \ldots, p_n]$ such that $\text{Prob}(x_i) = p_i$ and $p_i \in [0, 1]$ and $\sum_{i=1}^{n} p_i = 1$. We shall let P_X be the set of all valid probability distributions on X.

Formally a PBG probability distribution induces a possibility distribution over all the valid probability distribution over X. Let $K = \{(A_i, Q_i)|i = 1, \ldots, m\}$ be a PBG probability distribution on X. If $\prod_K$ is the induced possibility distribution then for each valid probability distribution, $P \in P_X$, $\prod_K(P)$ indicates the possibility that P is the actual probability distribution on X.

With $P = [p_1, \ldots, p_n]$ in the following we describe one approach to determine $\prod_K(P)$ given $K = ((A_i, Q_i)|i = 1, \ldots, m\}$.

(1) For each A_i calculate $\text{Prob}(A_i)$ using P: $\text{Prob}(A_i|P) = \sum_{j=1}^{n} A_i(x_j)\ p_j$
(2) For each i calculate, $\tau_i = Q_i(\text{Prob}(A_i|p))$. This is the compatibility of P with Q_i
(3) $\prod_K(P) = \text{Min}_i[\tau_i]$

In the case in which $X = [a, b]$, it is continuous, the random variable is characterized by a probability measure. Here the PBG probability distribution is only providing partial information about underlying probability measure. We note that a valid probability measure f associated with X is such that $f(x) \geq 0$ for all $x \in [a, b]$ and $\int_a^b f(x)dx = 1$. We let F_X be the collection of all valid probability measures on X. In this case a PBG probability induces a possibility distribution over the set F_X. Again we shall assume $K = ((A_i, Q_i), I = 1, \ldots, m)$ is the PBG probability distribution corresponding to the uncertainty profile. We let $\prod_K$ be the induced possibility distribution over F_X. Here $\prod_K(f)$ indicates the possibility that f can be the actual probability measure given K. We determine $\prod_K(f)$ as follows:

(1) For each A_i we calculate $\text{Prob}(A_i|f) = \int_a^b f(x)\ A_i(x)\ dx$
(2) For each i calculate, $\tau_i = Q_i(\text{Prob}(A_i|p))$. This is the compatibility of f with Q
(3) $\prod_K(f) = \text{Min}_i[t_i]$

Let us look at this nature of the PBG probability distribution in more detail. As we shall subsequently see a PBG probability distribution is essentially a generalization of the idea of an ordinary probability distribution. Consider the PBG probability distribution $((A_i, Q_i), i = 1, \ldots, m)$. First we note that each Q_i is a fuzzy number drawn from the unit interval I, it is normal and unimodal. In particular there exists an $r \in [0, 1]$ such that $Q_i(r) = 1$. In addition since it is unimodal, there exist two values a_i and $b_i \in I$ such that

1. $Q_i(r)$ is non-decreasing for $r \in [0,\ a_i]$
2. $Q_i(r) = 1$ for $r \in [a_i,\ b_i]$
3. $Q_i(r)$ is non-increasing for $r \in [b_i,\ 1]$

One implication of the unimodality of the granular probabilities is the interval nature of the associated level sets [12]. Thus if Q_i^α is the α-level set of Q_i, $Q_i^\alpha = \{r/Q_i(r) \geq \alpha\}$, then $Q_i^\alpha = [l_i(\alpha),\ u_i(\alpha)]$. It is also the case that the unimodality of Q_i implies that if $\alpha > \beta$ then $Q_i^\alpha \subseteq Q_i^\beta$, the level sets are nested.

We should note two special cases of these granular probabilities. The first is the case when Q_i is a precise value q_i in I, $Q_i = \{q_i\}$. The second is when Q_i is an interval, $Q_i = [a_i,\ b_i]$. Here $Q_i(r) = 1$ for $r \in [a_i,\ b_i]$ and $Q_i(r) = 0$ for $r \notin [a_i,\ b_i]$.

Generally the A_i are human comprehensible concepts associated with the space X. As discussed by Gardenfors [13] concepts on a domain are expressed as convex subsets. Thus formally the A_i are normal and unimodal, they are fuzzy numbers from the domain X. Two special cases of A_i are singletons and crisp intervals.

5 Evaluating Decision Functions with PBG Uncertainty Profiles

Previously we indicated that the rules based approach for specifying the decision making entities valuation function can involve rules in which we have antecedent terms of the form:

$$\text{If Prob(Fuzzy Event) } is \text{ Large then} \ldots \tag{I}$$

Here we shall investigate a method for evaluating the satisfaction of this type of antecedent by risky alternatives for this case in which an alternative's uncertainty profile is expressed in terms of a PBG probability distribution.

We first formalize the above antecedent. Let **R** indicate the payoff associated with the alternative being evaluated. Formally it is a random variable on real line. In order to formalize the antecedent in I we let F be a fuzzy subset of the domain of **R**, this corresponds to a general fuzzy event. In addition we let Q be a fuzzy probability corresponding to what we generically denoted as Large in (I). Using these notations our rule becomes

$$\text{If Prob}(\mathbf{R}\ is\ \text{F})\ is\ \text{Q then} \ldots\ldots.$$

Let us use W to indicate the variable corresponding to the "probability of the event **R** *is* F." Using this notation we can express our rule as

$$\text{"If W } is \text{ Q then} \ldots\text{."}$$

The firing of this rule is determined by the compatibility of the value of W with the fuzzy subset Q.

We now consider a risky alternative whose uncertainty profile is expressed using the PBG probability distribution $K = ((A_i, Q_i), i = 1, \ldots, m)$. Here A_i is a fuzzy subset of X and Q_i is a fuzzy subset corresponding to amount of probability, a fuzzy number in the unit interval.

The task of evaluating the degree to which the risky alternative under consideration satisfies the rule can be formulated as follows. We need to determine the compatibility of the value of W, the probability of the event **R** *is* F with Q, given that all we know about **R** *is* K, $((A_i, Q_i), i = 1, \ldots, m)$.

Consider the firing of the rule "If W *is* Q then ......" If we know that the probability of the event **R** *is* F is precisely equal to the value b, W = b, then the degree of firing τ is simply Q(b). More generally, if the value for W is a fuzzy probability B, then using the established procedure in fuzzy systems modeling we obtain as the firing level $\tau = \text{Max}_y[Q(y) \wedge B(y)]$, we take the maximum of the intersection of Q and B.

The situation we are faced with is slightly different than either of these. Instead of knowing the value of W, the probability of **R** *is* F, all we have is the PBG probability distribution K on **R**. In this case our task becomes to calculate the value of W from our information about **R**.

If instead of having a PBG probability distribution we had an ordinary probability distribution $P = [(x_i, p_i)]$, p_i being the probability that $\mathbf{R} = x_i$ then to calculate W, probability that **R** *is* F, we use

$$W = \sum_{i=1}^{n} F(x_i)p_i$$

We must now extend this approach to our situation where we have the PBG probability distribution $K = [(Ai, Q_i), i = 1, \ldots, m]$. With K we have that both A_i and Q_i are fuzzy subsets. The fact that A_i is not crisp conceptually provides more difficulty than the fuzziness of Q_i.

If we temporarily consider the situation in which Q_i is precise, $Q = q_i$ and A_i is an interval we can get some insight into how to proceed. We shall also for simplicity assume that F is a crisp subset. In calculating W we are essentially obtaining the sum of the probabilities of the possible values of **R** that lies in F. When A_i is an interval it is difficult to decide whether the probability is associated with element in F or not. To get around this problem we must obtain upper and lower bounds on W. The actual probability lies between these values.

Using this idea for the more general situation where all the objects are fuzzy we obtain

$$\text{Upper}_F = \sum_{i=1}^{n} \text{Poss}[F/A_i]\ Q_i$$

$$\text{Lower}_F = \sum_{i=1}^{n} (1 - \text{Poss}[\bar{F}/A_i])Q_i$$

where $\text{Poss}[F/A_i] = \text{Max}_x[F(x) \wedge A_i(x)]$ and $\text{Poss}[\bar{F}/A_i] = \text{Max}_x[(1 - F(x)) \wedge A_i(x)]$. Essentially we see that $\text{Poss}[F/A_i]$ is the degree of intersection of A_i and F while $1 - \text{Poss}[\bar{F}/A_i]$ is the degree to which A_i is included in $\bar{F}$. There values are closely related to the measures of plausibility and belief in Dempster–Shafer theory [11].

At this point we must draw upon some of results from fuzzy arithmetic [14]. We recall if A and B are two fuzzy numbers then their sum $D = A \oplus B$ is also a fuzzy number such that

$$D(z) = \underset{\substack{x,\ y\ \text{s.t.} \\ x+y=z}}{\text{Max}} [A(x) \wedge B(y)].$$

We also note that if α is a scalar then α A is a fuzzy number D such that

$$D(z) = \underset{\substack{x\ \text{s.t.} \\ \alpha x=z}}{\text{Max}} [A(x)]$$

More generally if $D_1, \ldots, D_n$ are fuzzy numbers and $\alpha_1, \ldots, \alpha_n$ are nonnegative scalars then

$$D = \alpha_1 D_1 \oplus \alpha_2 D_2 \oplus \cdots \oplus \alpha_n D_n$$

is a fuzzy number such that

$$D(z) = \underset{\substack{x_i\ \text{s.t.} \\ \Sigma_i \alpha_i x_i=z}}{\text{Max}} [\text{Ai}(x_i)]$$

The point we can conclude from this digression is that we have available to us the facility to calculate the values Upper_F or Lower_F. More specifically if we denote $\lambda_i = \text{Poss}[F/A_i] \in [0,1]$ then Upper_F is a fuzzy number H defined on the unit interval such that for all $z \in [0, 1]$

$$H(z) = \underset{\substack{z_i\ \text{s.t.} \\ \Sigma_i \lambda_i Z_i=z}}{\text{Max}} [\text{Min}_i[Q_i(z_i)]$$

If we denote $\gamma_i = 1 - \text{Poss}[\bar{F}/A_i] \in [0, 1]$ then Lower_F is a fuzzy number L defined on the unit interval such that for all $z \in [0,1]$

$$L(z) = \underset{\substack{z_i\ \text{s.t.} \\ \Sigma_i \gamma_i Z_i=z}}{\text{Max}} [\text{Min}_i[Q_i(z_i)]$$

We must now consider the relationship between the fuzzy subsets H and L. In anticipation of uncovering this we look at the relationship between $\lambda_i = \text{Poss}[F/A_i]$ and $\gamma = 1 - \text{Poss}[F/A_i]$. Here we use the fact that F and A_i are normal, they have at least one element with membership grade 1. Assume

$\gamma = \alpha$, then $\mathrm{Max}_x[(1 - F(x)) \wedge A_i(x)] = 1 - \alpha$. Since A_i is normal there exists some x^* where $A_i(x^*) = 1$ and therefore $(1 - F(x^*)) \wedge 1 = (1 - F(x^*)) \leq 1 - \alpha$ hence $F(x^*) \geq \alpha$. Since $\lambda_i = \mathrm{Max}_x[F(x) \wedge A_i(x)] \geq F(x^*) \wedge A_i(x^*) \geq \alpha$. Hence we get $\lambda_i \geq \gamma_i$ for all i. Thus we see that $L = \sum_{j=1}^{n} \gamma_j Q_j$ and $H = \sum_{j=1}^{n} \lambda_j Q_j$ where $\lambda_j \geq \gamma_j$ for all j.

Before preceding we want to introduce a type of relationship between fuzzy numbers

Definition 1. *Let* G_1 *and* G_2 *be two fuzzy numbers such that*

$$\begin{array}{ll} G_j(x) \textit{ is non-decreasing} & \text{for } x \leq a_j \\ G_j(x) = 1 & \text{for } x \in [a_j, b_j] \\ G_j(x) \textit{ is non-increasing} & \text{for } x \geq b_j \end{array}$$

where $a_1 \leq a_2$ *and* $b_2 \geq b_1$. *If in addition we have*

$$G_1(x) \geq G_2(x) \quad \text{for all } x \leq a_1.$$

$$G_2(x) \geq G_1(x). \quad \text{for all } x \geq a_2$$

we shall say G_2 *is to* **the right of** G_1 *and denote this as* $G_2 \geq_R G_1$

This relationship $G_2 \geq_R G_1$ can be equivalently expressed in terms of level sets. If $G_i(\alpha) = [a_i(\alpha),\ b_i(\alpha)]$ is the α level set of G_i, then the relationship $G_2 \geq_R G_1$ is equivalent to the condition that for each $\alpha \in [0, 1]$ we have $a_1(\alpha) \leq a_2(\alpha)$ and $b_1(\alpha) \leq b_2(\alpha)$.

It can be shown that if $G_2 = \sum_{i=1}^{n} \lambda_i\ Q_i$ and $G_1 = \sum_{i=1}^{n} \gamma_i\ Q_i$ where $0 \leq \gamma_i \leq \lambda_i \leq 1$ for all i and the Q_i are non-negative fuzzy number then $G_2 \geq_R G_1$. From this it follows that $H \geq_R L$, the upper bound is always to the right of the lower bound.

Earlier we indicated that the value of W, the probability that **R** is F, lies between the H and L. In particular, we have the following constraints on the value of W:

W is greater that or equal L
and
W is less than or equal H.

If we let L^* indicate the fuzzy subset *greater than or equal L* and let H^* indicate the fuzzy subset *less than or equal H* then W *is* E where $E = L^* \cap H^*$. It is the intersection of the fuzzy subsets L^* and H^*.

Let us now calculate L^* and H^* from L and H. L^* is obtained as

$$L^*(x) = \mathrm{Max}_y[\mathrm{GTE}(x, y) \wedge L(y)]$$

where GTE is the relationship "greater then or equal" defined on $[0,\ 1] \times [0,\ 1]$ by

$\mathrm{GTE}(x,\ y) = 1$ if $x \geq y$
$\mathrm{GTE}(x,\ y) = 0$ if $x < y$

Here $L(x)$ is non-decreasing for $x \leq a_1$ and $L(x) = 1$ for $x \in [a_1, b_1]$ it is non-increasing for $x \geq b_1$. It is easy to show that in this case that L^* is such that $L^*(x) = L(x)$ for $x \leq a_1$ and $L^*(x) = 1$ for $x \geq a_1$.

Similarly for H^* we have $H^*(x) = Max_y[LTE(x, y) \wedge H(y)]$ LTE is the relationship "less then or equal" defined on $[0, 1] \times [0, 1]$ by

$LTE(x, y) = 1$ if $x \leq y$
$LTE(x, y) = 0$ if $x > y$

If $H(x)$ is a fuzzy number with value one in the interval $[a_2, b_2]$ then H^* is a fuzzy number such $H^*(x) = 1$ for $x \leq b_2$ and $H^*(x) = H(x)$ for $x > b_2$.

Combining L^* and H^* to get E, the possible values for W, we have $E = H^* \cap L^*$ hence $E(x) = H^*(x) \wedge L^*(x)$. From this we get

$E(x) = L(x)$ for $x \in [0, a_1]$
$E(x) = 1$ for $x \in [a_1, b_2]$
$E(x) = H(x)$ for $x \in [b_2, 1]$

Returning to our concern with determining the firing level of the rule

If W *is* Q then

when our input is W = K we now use this E to calculate the firing level of the rule as

$$\tau = Max_x[Q(x) \wedge E(x)]$$

6 Cumulative Distribution Functions

Here consider the situation where the information about the uncertainty profile of an alternative is available in terms of a cumulative distribution function and more generally a **P**erception **B**ased **G**ranular **C**umulative **D**istribution function., PBG-CD function.

If **R** is a random variable that takes its value on the real line we recall that a cumulative distribution is a function such that F(x) is the probability that $\mathbf{R} \leq x$. Formally F is a function $F : [-\propto, \propto] \rightarrow [0, 1]$ which is monotonic, $F(x) \geq F(y)$ if $x > y$. We note F is available whether **R** is discrete or continuous. If **R** is discrete then $F(x) = \sum_{\substack{i \text{ s. t.} \\ x_i \leq x}} p_i$. If **R** is continuous with probability density f then $F(x) = \int_{-\infty}^{\infty} f(x)\, dx$. In many real applications we can assume that the domain of F is bounded, there exists some value x_* s.t. such that $F(x) = 0$ for $x \leq x_*$ and some x^* such that. $F(x) = 1$ for all $x \geq x^*$.

With the availability of the CDF we can easily provide the information needed to determine the firing level of a rule of the form

If Prob(A) *is* then....

If A is a crisp subset, $A = \{x/a_1 \leq x \leq a_2]$ then $Prob(A) = F(a_2) - F(a_1)$ and the firing level is $Q(F(a_2) - F(a_1))$. If A is a fuzzy subset we must look a little more carefully at the situation. Here we shall assume A is a fuzzy number, the fuzzy subset A is of the form

$$\begin{array}{ll} A(x) = 0 & \text{for } x \le b_1 \\ A(x) \ge A(y) & \text{for } b_1 \le y < x \le a_1 \\ A(x) = 1 & \text{for } a_1 < x \le a_2 \\ A(x) \le A(y) & \text{for } a_2 \le y \le x \le b_2 \\ A(x) = 0 & \text{for } x \ge b_2 \end{array}$$

We now define a fuzzy subset $\tilde{a}_1$ such that

$$\begin{array}{ll} \tilde{a}_1(x) = A(x) & \text{for } b_1 \le x \le a_1 \\ \tilde{a}_1(x) = 0 & \text{elsewhere} \end{array}$$

We also define the fuzzy subset $\tilde{a}_2$ such that

$$\begin{array}{ll} \tilde{a}_2(x) = A(x) & \text{for } a_2 \le x \le b_2 \\ \tilde{a}_2(x) = 0 & \text{elsewhere} \end{array}$$

$\tilde{a}_1$ and $\tilde{a}_2$ are fuzzy numbers which allow us to express Prob(A) $= F(\tilde{a}_2) - F(\tilde{a}_1)$. In order to obtain Prob(A) we need to obtain $F(\tilde{a}_2)$ and $F(\tilde{a}_1)$. Since the processes needed to obtain these values are similar we shall only concentrate on $F(\tilde{a}_2)$. Using Zadeh's extension principle [15,16], since $\tilde{a}_2$ is a fuzzy number of real line, then $F(\tilde{a}_2)$ is a fuzzy subset of the unit interval such that $F(\tilde{a}_2) = \bigcup_x \{\frac{\tilde{a}_2(x)}{F(x)}\}$ and since $\tilde{a}_2(x) = A(x)$ for $x \in [a_2,\ b_2]$ and $\tilde{a}_2(x) = 0$ elsewhere then $F(\tilde{a}_2) = \bigcup_{x \in [a_2,\ b_2]} \{\frac{A(x)}{F(x)}\}$. Here $F(\tilde{a}_2)$ is a fuzzy number. In this case the possibility that $F(\tilde{a}_2)$ takes the value z is $\underset{\substack{x \in [a_2,\ b_2] \\ F(x)=z}}{\text{Max}} [A(x)]$. The monotonic nature of the cumulative distribution function F and the special form of $\tilde{a}_2$ results in a form of $F(\tilde{a}_2)$ as shown in Fig. 2. We emphasize that $F(\tilde{a}_2)$ is a fuzzy number of the unit interval such that its membership grade is one at the value $F(a_2)$, and monotonically decreases to zero at the value $F(b_2)$. In the range from zero to $F(a_2)$ and $F(b_2)$ to 1 its membership value is also zero.

Some special situations are worth pointing out. If F is such that it is constant, $F(x) = k$, in the range $x \in [a_2,\ b_2]$ then it can be shown that $F(\tilde{a}_2)$

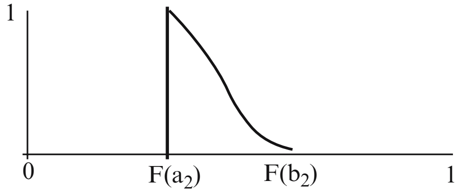

Fig. 2. Fuzzy subset $F(\tilde{a}_2)$

is a singleton set, $F(\tilde{a}_2) = \{\frac{1}{k}\} = \{\frac{1}{F(a_2)}\}$. Another special case occurs if F is a discrete function. Specifically if F is such that within the interval $[a_2,\ b_2]$ it jumps at the points $a_2 + \delta_1,\ a_2 + \delta_2,\ a_3 + \delta_3$ where the amounts of these jumps are $\Delta_1,\ \Delta_2,\ \Delta_3$. In this special case we get

$$F(\tilde{a}_2) = \left\{ \frac{1}{F(a_2)},\ \frac{A(a_2+\delta_1)}{F(a_2)+\Delta_1},\ \frac{A(a_2+\delta_2)}{F(a_2)+\Delta_1+\Delta_2},\ \frac{A(a_3+\delta_3)}{F(a_2)+\Delta_1+\Delta_2+\Delta_3} \right\}$$

The significant point here is that here $F(\tilde{a}_2)$ is a discrete function reflecting the discrete nature of F.

In a similar way we can show generally $F(\tilde{a}_1)$ is a fuzzy number of the unit interval such that $F(\tilde{a}_1) = \bigcup_{x \in [b_1,\ a_1]} \{\frac{A(x)}{F(x)}\}$, see Fig. 3.

Using these fuzzy values for $F(\tilde{a}_2)$ and $F(\tilde{a}_1)$ we obtain $\mathrm{Prob}(A) = F(\tilde{a}_2) - F(\tilde{a}_1)$ as a fuzzy number of unit interval having nonzero membership grade in the interval $(F(a_2) - F(a_1))$ to $(F(b_2) - F(b_1))$. Here if we let PA be the fuzzy subset denoting the value Prob(A) then

$PA(z) = 0$	$z < F(a_2) - F(a_1)$
$PA(z) = 1$	$z = F(a_2) - F(a_1)$
$PA(z)$ is decreasing	$F(a_2) - F(a_1) < z < F(b_2) - F(b_1)$
$PA(z) = 0$	$z > F(b_2) - F(b_1)$

In some practical situations it may be much more efficient to defuzzify $F(\tilde{a}_1)$ and $F(\tilde{a}_2)$ and use these scalar values to obtain a scalar value for Prob(A).

Let us consider the defuzzification of $F(\tilde{a}_2)$ which we recall was $F(\tilde{a}_2) = \bigcup_{x \in [a_2,\ b_2]} \{\frac{A(x)}{F(x)}\}$. Letting d_2 denote the defuzzified value of $F(\tilde{a}_2)$ we get

$$d_2 = \frac{\int_{a_2}^{b_2} F(x)\ A(x)\ dx}{\int_{a_2}^{b_2} A(x) dx}.$$

We observed that if F(x) is constant, $F(x) = k$ in the range $[a_2,\ b_2]$, then $d_2 = k$. Actually as we have already pointed out if $F(x) = k$ in the range a_2 to b_2 then $F(\tilde{a}_2)$ is itself a constant value k, no fuzziness exists.

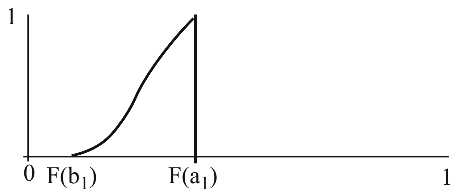

Fig. 3. Fuzzy subset $F(\tilde{a}_1)$

In many real situations it may be difficult for a decision maker to obtain a precise manifestation of the cumulative distribution of the payoff of a risky alternative. In these cases a decision maker may be only able to obtain a imprecise characterization of the underlying cumulative distribution in the form of what we shall call a **P**erception **B**ased **G**ranular **C**umulative **D**istribution function, PBG-CD function. A PBG-CD is a granular description of the cumulative distribution function in a form that is widely used in fuzzy modeling [6]. When using a PBG-CD we partition the range R into fuzzy intervals $B_1, \dots, B_n$. We then express the value of F in each one of these fuzzy ranges using a fuzzy subset of the unit interval F_i. With PBG-CD function we have a rule based representation of the cumulative distribution function F

If U is B_1 then F is F_1.
....
If U is B_i then F is F_i.
....
If U is B_n then F is F_n.

In working with the fuzzy rule based description of the underlying function we can draw upon the well established literature of fuzzy systems modeling.

In order to find the value of F at some value for U, a, we proceed as follows. We first obtain the firing level of each rule $\tau_i = B_i(a)$. We then calculate $\omega_i = \frac{\tau_i}{\sum_{i=1}^{n} \tau_i}$. Using this we calculate F(a) as the fuzzy subset $\mathbf{F_a} = \sum_{i=1}^{n} \omega_i F_i$. Here we get for F(a) a fuzzy subset of the unit interval such that $\mathbf{F_a}(y)$ is the possibility that F(a) assumes the value y. We can apply a defuzzification operation on $\mathbf{F_a}$ to obtain a scalar value.

In the following example we illustrate the generation of a perception based granular CD function

Example. We consider an investment alternative in which the investor has the following perceptions of the outcome of his investment.

He is certain that he won't lose more then \$500 dollars
He believes his chances of losing more then \$100 is about 10%
He believes his chances of losing any money is 20%
He feels that there is about a 90% chance that he will win at most \$500
He is certain that he won't win more then a \$1,000

We can use this to construct a rule based description of the cumulative distribution function. In particular if $F(U) = \text{Prob}(\mathbf{R} \leq U)$ with **R** being the random payoff then the rule base is

If U *is* **less then \$500** then F *is* **zero**
If U *is* "**near −\$100**" then F *is* **about 10%**
If U *is* **zero** then F *is* **about 20%**
If U *is* **about \$500** then F *is* **about 90%**
If U *is* **greater then 1,000** then F *is* **100%**

7 Conclusion

We focused on the issue of decision making in risky situations. We discussed the need for using decision functions to aid in capturing the decision maker's preference among these types of uncertain alternatives. The use of fuzzy rule based formulations to model these functions was investigated. We discussed the role of Zadeh's perception based granular probability distributions as a means of modeling the uncertainty profiles of the alternatives. We look at various properties of this method of describing uncertainty and showed how they induced possibility distributions of the space of probability distributions Tools for evaluating rule based decision functions in the face of perception based uncertainty profiles were presented. We considered the situation in which uncertainty profiles are expressed in terms of a cumulative distribution function. We introduced the idea of a perception based granular cumulative distribution and describe its representation in terms of a fuzzy rule based model.

References

1. Zadeh, L. A., From computing with numbers to computing with words-From manipulation of measurements to manipulations of perceptions. *IEEE Transactions on Circuits and Systems* 45, 105–119, 1999
2. Zadeh, L. A., A new direction in AI – toward a computational theory of perceptions. *AI Magazine* 22(1), 73–84, 2001
3. Zadeh, L. A., Toward a logic of perceptions based on fuzzy logic. In: Novak, W., Perfilieva, I. (Eds.) *Discovering the World with Fuzzy Logic*. Physica-Verlag: Heidelberg, pp. 4–28, 2001
4. Zadeh, L. A., Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. *Journal of Statistical Planning and Inference* 105, 233–264, 2002
5. Yager, R. R., Using a notion of acceptable in uncertain ordinal decision making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 241–256, 2002
6. Yager, R. R., Filev, D. P., *Essentials of Fuzzy Modeling and Control*. Wiley: New York, 1994
7. Takagi, T., Sugeno, M., Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics* 15, 116–132, 1985
8. Klir, G. J., *Uncertainty and Information*. Wiley: New York, 2006
9. Klir, G. J., Yuan, B., *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall: Upper Saddle River, NJ, 1995
10. Zadeh, L. A., Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications* 10, 421–427, 1968
11. Yager, R. R., Liu, L., (A. P. Dempster and G. Shafer, Advisory Editors) *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Springer: Berlin Heidelberg New York, 2008
12. Dubois, D., Prade, H., *Fuzzy Sets and Systems: Theory and Applications*, Academic Press: New York, 1980

13. Gardenfors, P., *Conceptual Spaces: The Geometry of Thought*. MIT Press: Cambridge, MA, 2000
14. Dubois, D., Prade, H., Fuzzy numbers: an overview. In: Bezdek, J. C. (ed.) *Analysis of Fuzzy Information Vol. 1: Mathematics and Logic*. CRC Press: Boca Raton, FL, pp. 3–39, 1987
15. Zadeh, L. A., Fuzzy sets. *Information and Control* 8, 338–353, 1965
16. Yager, R. R., A characterization of the extension principle. *Fuzzy Sets and Systems* 18, 205–217, 1986