Chapter 3
Uncertain Preferences and Imperfect Information in Decision Making

3.1 Vague Preferences

One of the main aspects defining solution of a decision problem is a preferences framework. In its turn one of the approaches to formally describe preferences is the use of utility function. Utility function is a quantitative representation of a DM's preferences and any scientifically ground utility model has its underlying preference assumptions.

The first approach to modeling human preferences was suggested by von Neumann and Morgenstern [70] in their expected utility (EU) model. This approach is based on axioms of weak order, independence and continuity of human preferences over actions set A. As it was shown by many experiments and discussions conducted by economists and psychologists, the assumption of independence appeared non-realistic [4,9,20]. There were suggested a lot of preferences frameworks which departs from that of EU by modeling a series of key aspects of human behavior.

Reconsideration of preferences framework underlying EU resulted in development of various advanced preferences frameworks as generalizations of the former. These generalizations can be divided into two types: rank-dependence generalization and sign-dependence generalization [11]. Advanced preference frameworks include various reconsiderations and weakening of the independence axiom [24,27,31,63] , human attitudes to risk and uncertainty (rank-dependent generalization), gains and losses [42,68] (sign-dependent generalization) etc.

Various notions of uncertainty aversion (ambiguity aversion) in various formulations were included into many preference models starting from Schmeidler's Choquet expected utility (CEU) [63] preferences framework and Gilboa and Schmeidler's maximin expected utility preferences framework [31], to more advanced uncertainty aversion formulations of Ghirardato, Maccheroni, Marinacci, Epstein, Klibanoff [10,12,21,23,25,27,28,44,45].

In [25] they provide axiomatization for $\alpha-$MMEU – a convex combination of minimal and maximal expected utilities, where minimal expected utility is multiplied by a degree of ambiguity aversion (ambiguity attitude). This approach allows differentiating ambiguity and ambiguity attitude. However, the approach in

R.A. Aliev: *Fuzzy Logic-Based Generalized Theory of Decisions*, STUDFUZZ 293, pp. 89–125.
DOI: 10.1007/ 978-3-642-34895-2_3

[25] allows to evaluate overall utility only comparing ambiguity attitudes of DMs with the same risk attitudes. The analogous features of comparative ambiguity aversion are also presented in [28]. Ambiguity aversion as an extra risk aversion is considered in [12,22].

In [43] they suggested a smooth ambiguity model as a more general way to formalize decision making under ambiguity than MMEU. In this model probability-relevant information is described by assessment of DM's subjective probabilistic beliefs to various relevant probability distributions. In contrast to other approaches to decision making under ambiguity, the model provides a strong separation between ambiguity and ambiguity attitude. To describe whether a DM is an ambiguity averse, loving or neutral it is suggested to use well known technique of modeling risk attitudes. More concretely, to reflect a considered DM's reaction to ambiguity it is suggested to use a concave nonlinear function with a special parameter α as the degree of ambiguity aversion – the larger α correspond to a more ambiguity averse DM. In its turn ambiguity loving is modeled by a convex nonlinearity. As opposed to the models in [21,25,28], the model in [43] allows for comparison of ambiguity attitudes of DMs whose risk attitudes are different.

The other important property included into some modern preference frameworks is a tradeoff-consistency which reflects strength of preferences with respect to coordinates of probabilistic outcomes.

The preference framework of the Cumulative Prospect Theory (CPT), suggested by Kahneman and Tversky [68], as opposed to the other existing frameworks includes both rank-dependence and sign-dependence features [11].

But are the modern preferences frameworks sufficiently adequate to model human attitudes to alternatives? Unfortunately, the modern preferences frameworks miss very important feature of human preferences: human preferences are vague [58]. Humans compare even simple alternatives linguistically using certain evaluation techniques such as "much better", "much worse", "a little better", "almost equivalent" [81] etc. So, a preference is a matter of an imprecise degree and this issue should be taken into account in formulation of preferences framework. Let us consider an example.

Suppose that Robert wants to decide among two possible jobs a_1, a_2 based on the following criteria: salary, excitement and travel time. The information Robert has is that the job a_1 offers notably higher salary, slightly less travel time and is significantly less interesting as compared to the job a_2. What job to choose?

Without doubt, evaluations like these are subjective and context-dependent but are often faced. If to suppose that for Robert salary is "notably" more important than the time issues and "slightly" more important than excitement then it may be difficult to him to compare these alternatives. The relevant information is too vague for Robert to clearly give preference to any of the alternatives. Robert may feel that superiority of the a_1 on the first criterion is approximately compensated by the superiority of the a_2 in whole on the second and the third criteria. But, at the same time Robert may not consider these jobs equally good. As a result, in contrast to have unambiguous preferences, Robert has some "distribution" of his preferences among alternatives. In other words, he may think that to some degree the job a_1 is as good as the job a_2 and, at the same time, that to some degree job a_2 is as good as the job a_1.

In this example we see that vagueness of subjective comparison of alternatives on the base of some criteria naturally passes to the preferences among alternatives. For example, the term "notably higher" is not sharply defined but some vague term because various point estimates to various extents correspond to this term – for a given point estimate its correspondence to a "notably higher" term may not be true or false but partially true. This makes use of interval description of such estimates inadequate as no point may partially belong to an interval – it belongs or not. It is impossible to sharply differentiate "notably higher" and not "notably higher" points. As a result, vague estimates (in our case vague preferences) cannot be handled and described by classical logic and precise techniques. Fuzzy logic [1,81] is namely the tool to handle vague estimates and there is a solid number of works devoted to fuzzy and linguistic preference relations [57,74]. This is due to the fact that vagueness is more adequately measured by fuzziness. As a result, fuzzy degree-based preference axiomatization is more adequate representation from behavioral aspects point of view as it is closer to human thinking. In view of this, linguistic preference relations as a natural generalization of classical preference relations are an appropriate framework to underlie human-like utility model.

Fuzzy preferences or fuzzy preference relations (FPRs) are used to reflect the fact that in real-world problems, due to complexity of alternatives, lack of knowledge and information and some other reasons, a DM can not give a full preference to one alternative from a pair. Preferences remain "distributed" reflecting that one alternative is to some extent better than another. In contrast to classical preference relations, FPR shows whether an alternative a is more preferred to b than alternative c is preferred to d.

Given a set of alternatives A, any fuzzy preference relation on A is a mapping $R: A\times A \to T$ where T is a totally ordered set. Very often fuzzy preference relation is considered as $R: A\times A \to [0,1]$ which assigns to any pair of alternatives $a,b\in A$ a degree of preference $R(a,b)\in[0,1]$ to which a is preferred to b. The higher $R(a,b)$ is, the more a is preferred to b. In other words, FPR is characterized by membership function $\mu_R(a,b)=R(a,b)$ which returns a degree of membership of a pair (a,b) to R. FPR is a valued extension of classical preference relations. For example, a weak order is a special case of FPR when $R: A\times A \to \{0,1\}$ with $R(a,b)=1$ if and only if $a\succeq b$ and $R(a,b)=0$ otherwise.

Consider a general case of a classical preference relation (CPR) implying that a is either strictly preferred, or equivalent or incomparable to b. This means that CPR is decomposed into a strict preference relation P, indifference preference relation I and incomparability preference relation J, that is, $\forall a,b \in A$ either $(a,b) \in P$, or $(a,b) \in I$ or $(a,b) \in J$. An important extension of this case to FPR can be defined as follows:

$$P(a,b)+P(b,a)+I(a,b)+J(a,b)=1$$

where $P,I,J: A \times A \to [0,1]$ are fuzzy strict, fuzzy indifference and fuzzy incomparability preference relations respectively.

Another important type of FPR is described by a function $R: A \times A \to [0,1]$ where $R(a,b)=1$ means full strict preference of a over b, which is the same as $R(b,a)=0$ (full negative preference) and indifference between a over b is modeled as $R(a,b)=R(b,a)=1/2$. In general, R is an additive reciprocal, i.e. $R(a,b)+R(b,a)=1$. This is a degree-valued generalization of completeness property of classical relation, and $R(a,b)>1/2$ is a degree of a strict preference. However, such an R excludes incomparability.

Consider yet another important type of FPR within which indifference is modeled by $R(a,b)=R(b,a)=1$, incomparability – by $R(a,b)=R(b,a)=0$ and completeness – by $\max(R(a,b),R(b,a))=1$.

FPR are a useful tool to handle vague preferences. Linguistic preferences, or linguistic preference relations (LPRs), sometimes called fuzzy linguistic preferences, are generalization of FPR used to account for a situations when a DM or an expert cannot assign precise degree of preference of one alternative to another, but express this degree in a form of linguistic terms like “much better”, “a little worse” etc. Indeed, under imperfect environment where relevant information is NL-based, there is no sufficient information to submit exact degrees, but is natural to express degrees in NL also corresponding to the kind of initial information.

To formalize LPR it is first necessary to define a set of linguistic terms as a set of verbal expressions of preference degrees which would be appropriate for a considered problem. As a rule, they consider a finite and totally ordered linguistic term set $T=\{t_i\},\, i \in \{0,...,m\}$ with an odd cardinal ranging between 5 and 13. Each term is semantically represented by a fuzzy number, typically triangular or trapezoidal, placed over some predefined scale, e.g. [0,1]. For example: “no preference” – (0,0,0), “slightly better” – (0,0.3,0.5), “more or less better” – (0.3,0.5,0.7), “sufficiently better” – (0.5,0.7,1), “full preference” – (0.7,1,1). The cardinality of the term set is usually an odd.

Consider a finite set of alternatives $A=\{f_i,\, i=1,2,...,n\ (n \ge 2)\}$. Then an LPR is formally defined as follows:

Definition 3.1 [36]**.** Let $A = \{f_i, i = 1,2,...,n \ (n \geq 2)\}$ be a finite set of alternatives, then a linguistic preference relation $\tilde{R}$ is a fuzzy set in A^2 characterized by a membership function

$$\mu_{\tilde{R}} : A^2 \to T$$

$$\mu(f_i, f_j) = \tilde{r}_{ij}, \forall f_i, f_j \in A$$

indicating the linguistic preference degree of alternative f_i over f_j, , i.e. $\tilde{r}_{ij} \in T$.

So, LPR is represented by a membership function whose values are not precise degrees in [0,1] but fuzzy numbers in [0,1]. This means that LPR is a kind of FPR if to recall that the latter is in general defined by MF whose range is an ordered structure.

Traditional Fuzzy Linguistic Approach (TFLA). TFLA preserves fuzzy information about degrees of preference by direct computations over fuzzy numbers and, as a result, is of a high computational complexity. There are various other approaches to modeling LPR, some of which allow reducing computational complexity of the TFLA or suggesting some reasonable trade-off between preserving information and computational complexity. One of them is referred to as ordinal fuzzy linguistic modeling (OFLM). This approach is based on an idea of the adopting symbolic computations [35] over indices of terms in a term set instead of direct computations over the terms themselves as fuzzy numbers. This makes the approach sufficiently simpler in terms of computational complexity than TFLA. In OFLM, they consider a finite linguistic term set with an odd cardinality and the terms described by fuzzy numbers over the unit interval [0,1]. Also, a mid term is used to express approximate equivalence of alternatives by a fuzzy number with a mode equal to 0.5 and labeled like "almost equivalent". The other terms are distributed around the mid term expressing successively increasing preference degrees to the right and their symmetrical counterparts to the left. For example: "sufficiently worse", "more or less worse", "slightly worse", "almost equivalent", "slightly better", "more or less better", "sufficiently better".

There exist also approaches to model uncertainty of preferences, other than FPR. These approaches accounts for comparison of ill-known alternatives under crisp (non-fuzzy) preference basis. In one of these approaches, which is used for modeling valued tournament relations, $R(x, y)$ measures the *likelihood of a crisp weak preference* $x \succeq y$ [13,17]. Formally, $R(x, y)$ is defined as follows:

$$R(x, y) = P(x \succ y) + \frac{1}{2} P(x \sim y),$$

where $x \sim y \Leftrightarrow x \succeq y$ and $y \succeq x$, which implies $R(x, y) + R(y, x) = 1$. Thus, uncertainty of preference is described by a probability distribution P over possible conventional preference relations, i.e., $T_i \subset A \times A$ with $x \succeq_i y \Leftrightarrow (x, y) \in T_i$ and $P(T_i) = p_i$, $i = 1,...,N$. Then

$$R(x, y) = \sum_{i: x \succ_i y} p_i + \sum_{i: x \sim_i y} \frac{1}{2} p_i$$

For more details one can refer to [14]. This approach to modeling uncertain preferences was the first interpretation of fuzzy preference relations in the existing literature and was considered in the framework of the voting theory. It is needed to mention that such relations may be considered fuzzy because a degree of preference is used whereas the approach itself is probabilistic. However, this degree is a measured uncertainty about preferences which are themselves crisp but are not known with certainty. So, in its kernel, this approach does not support an idea underlying FPR – preference itself is a matter of a degree.

Another application of this approach may be implemented when there exists a utility function which quantifies preferences between alternatives *a,b,c* on some numerical scale and the latter is supported by additional information in form of a probability distribution. Then, $R(x, y) = P(u(x) > u(y))$, where $u : S \to R$ is a utility function.

Other measures of uncertainty can also be used to describe uncertainty of preference. One of them is the possibility measure, by using thereof the uncertain preference is defined as

$$R(x, y) = P(x \succeq y),$$

where $P(x \succeq y)$ is the degree of possibility of preference. The use of the possibility theory defines $\max(R(x, y), R(y, x)) = 1$. At the same time, in terms of the possibility theory, $1 - R(x, y) = N(x \succeq y)$ is the degree of certainty of a strict preference.

A large direction in the realm of modeling uncertain preferences is devoted to modeling *incomplete preferences*. In line with transitivity, completeness of preferences is often considered as a reasonable assumption. However, transitivity is used as a consistency requirement whereas completeness is used as a requirement which exclude indecisiveness. The reasonability and intuitiveness of these basics are not the same: for completeness they may loss their strength as compared to transitivity because in real choice problems lack of information, complexity of alternatives, psychological biases etc may hamper someone's choice up to indecisiveness. From the other side, indecisiveness may take place in group decision making when members' preferences disagree. The issue that completeness may be questionable was first addressed by Aumann [6]:

"Of all the axioms of utility theory, the completeness axiom is perhaps the most questionable. Like others of the axioms, it is inaccurate as a description of real life; but unlike them, we find it hard to accept even from the normative viewpoint.

For example, certain decisions that [an] individual is asked to make might involve highly hypothetical situations, which he will never face in real life; he might feel that he cannot reach an "honest" decision in such cases. Other decision problems might be extremely complex, too complex for intuitive "insight," and our individual might prefer to make no decision at all in these problems. Is it "rational" to force decisions in such cases?"

If to assume that preferences are not complete, one has to reject the use numerical utility functions and has to deal with more complex representations. As it is argued in [56], the use of a numerical utility is naturally leads to loss of information and then should not be dogmatic if one intends to model bounded rationality and imperfect nature of choice. In [56] they suggest to handle incomplete preferences by means of a vector-valued utility function as its range is naturally incompletely ordered. The other main argument for such an approach is that the use of a vector-valued utility is simpler than dealing with preferences themselves and in this case well-developed multi-objective optimization techniques may be applied. The idea of incomplete preferences underlying the approach in [56] is realized by the following assumption: given the set of alternatives A there exist at least one pair of alternatives $x, y \in A$ for which neither $x \succeq y$ nor $y \succeq x$ is assumed.

There exist also other approaches dealing with incomplete preferences by means of imprecise beliefs and/or imprecise utilities. The following classification of these approaches is given in [54]:

1) Probabilities alone are considered imprecise. For this setting preferences are represented by a convex set of probability distributions and a unique, utility function u(). Such models are widely used in robust Bayesian statistics [41, 61,73];
2) Utilities alone are considered imprecise. In this setting preferences are represented by a set of utility functions {u(c)} and a unique probability distribution p(s). Such representations were axiomatized and applied to economic models by Aumann [6] and Dubra, Maccheroni and Ok [19];
3) Both probabilities and utilities are considered imprecise. This is represented by sets of probability distributions {p(s)} and utility functions {u()}. These sets are considered separately from each other allowing for all arbitrary combinations of their elements. This is the traditional separation of imprecise information about beliefs and outcomes. Independence of two sets is practically justified and simplifies the decision analysis. However, this approach does not have axiomatic foundations. From the other side, the set of pairs may be non-convex and unconnected [40,41].

In order to compare adequacy of FPR and incomplete preferences models, we can emphasize the following classification of preference frameworks in terms of increasing uncertainty: complete orders, FPR, incomplete preferences. The first and the third one are idealized frameworks: the first implies that preference is absolutely clear, the third deals with the case when some alternatives are absolutely not comparable. Incomplete preference deals with lack of any information which can elucidate preferences. This is a very rare case in the sense that in most of real-world situations some such information does exist, though it requires to be obtained. In its turn FPR implies that preference itself is not "single-valued" and should reflect competition of alternatives even if the related information is precise.

3.2 Imperfect Information

In real-life decision making problems DM is almost never provided with perfect, that is, ideal decision-relevant information to determine states of nature, outcomes, probabilities, utilities etc and has to construct decision background structure based on his/her perception and envision. In contrast, relevant information almost always comes imperfect. Imperfect information is information which in one or more respects is imprecise, uncertain, incomplete, unreliable, vague or partially true [79]. We will discuss these properties of imperfect information and relations among them.

Two main concepts of imperfect information are imprecision and uncertainty. Imprecision is one of the widest concepts including variety of cases. For purposes of differentiation between imprecision and uncertainty, Prof. L.A. Zadeh suggested the following example: "*For purposes of differentiation it is convenient to use an example which involves ethnicity. Assume that Robert's father is German and his mother's parents were German and French. Thus, Robert is 3/4 German and 1/4 French. Suppose that someone asks me: What is Robert's ethnicity. If my answer is: Robert is German, my answer is imprecise or, equivalently, partially true. More specifically, the truth value of my answer is 3/4. No uncertainty is involved. Next, assume that Robert is either German or French, and that I am uncertain about his ethnicity. Based on whatever information I have, my perception of the likelihood that Robert is German is 3/4. In this case, 3/4 is my subjective probability that Robert is German. No partiality of truth is involved.*" In the first case imprecision is only represented by partial truth and no uncertainty is involved. As Prof. L.A. Zadeh defines, such imprecision is referred to as strict imprecision or s-imprecision for short. In the second case, imprecision is only represented by uncertainty and no partial truth is involved.

Information is partially true if it is neither absolutely true nor absolutely false but in an intermediate closeness to reality. For example, suppose you needed to write down ten pages of a text and have already written 8 pages. Certainly 'the work is done' is not absolutely true and is not absolutely false, and, if to assume that all pages are written equivalently difficult, 'the work is done' is true with degree 0.8. Form the other side, 'the work is not done' is not true and is not false from viewpoint of intuition because it is not informative and requires to be substituted by a more concrete evaluation.

Another example on imprecision and uncertainty is provided by P. Smets:

"To illustrate the difference between imprecision and uncertainty, consider the following two situations:

1. John has at least two children and I am sure about it.
2. John has three children but I am not sure about it.

In case 1, the number of children is imprecise but certain. In case 2, the number of children is precise but uncertain. Both aspects can coexist but are distinct. Often the more imprecise you are, the most certain you are, and the more precise, the less certain. There seems to be some Information Maximality Principle that requires that the 'product' of precision and certainty cannot be beyond a certain critical level. Any increase in one is balanced by a decrease in the other."

Imprecision is a property of the content under consideration: either more than one or no realization is compatible with the available information [65].

One realization of imprecise information is ambiguous information. Ambiguous information is information which may have at least to different meanings. For example, a statement 'you are aggressive' is ambiguous because aggressive may mean 'belligerent' or 'energetic'. For example, homonyms are typical carriers of ambiguity.

Ambiguous information may be approximate, e.g. 'the temperature of water in the glass is between 40 and 50°C is approximate if the temperature is 47°C. Ambiguous information like 'the temperature is close to 100C' is vague. Such vague information is fuzzy, because in this case the temperature is not sharply bounded. Both 99C and 103 corresponds to this, but the first corresponds stronger. Correspondence of a temperature value to 'the 'temperature is close to 100C' smoothly decreases as this value moves away from 100C. In general, vague information is information which is not well-defined; it is carried by a 'loose concept'. The worst case of vague information is unclear information. Ambiguous information may also be incomplete: "the vacation will be in a summer month" because a summer month may be either June, July or August.

Uncertain information is commonly defined as information which is not certain. P. Smets defines uncertainty as a property that results from a lack of information about the world for deciding if the statement is true or false. The question on whether uncertainty is objective or subjective property is still rhetoric.

Objective uncertainty may be probabilistic or non-probabilistic. Probabilistic uncertainty is uncertainty related to randomness – probability of an event is related to its tendency to occur. Main kinds of non-probabilistic uncertainty are possibilistic uncertainty and complete ignorance. Possibilistic uncertainty reflects an event's 'ability' to occur. To be probable, an event has to be possible. At the same time, very possible events may be a little probable. The dual concept of possibility is necessity. Necessity of an event is impossibility of the contrary event to occur. Complete ignorance is related to situations when no information on a variable of interest (e.g. probability) is available. For case of probability complete ignorance may be described by a set of all probability distributions.

Objective uncertainty relates to evidence on a likelihood of phenomena. Subjective uncertainty relates to DM's opinion on a likelihood of phenomena. More specifically, subjective uncertainty is a DM's belief on occurrence of an event. Classification of subjective uncertainty is very wide and its primitive forms are, analogously to that of objective uncertainty, subjective probability, subjective possibility and subjective necessity. The structures of these forms of subjective uncertainty are the same as those of objective probability, possibility and necessity. However, the sources of them differ: subjective uncertainty is a DM's opinion, whereas objective uncertainty is pure evidence. For example, mathematical structure of subjective probability is a probability measure but the values of this measure are assigned on the base of a DM's opinion under lack of evidence. Analogously, subjective possibility and necessity are a DM's opinions on possibility and necessity of an event.

Unreliable information is information to which an individual does not trust or trusts weakly due to the source of this information. As a result, an individual does not rely on this information. For example, you may not trust to the meteorological forecast if it is done by using old technology and equipment.

Imperfect information is impossible to be completely caught in terms of understanding what this concept means (e.g. uncertainty concept), and thus, cannot be perfectly classified. Any classification may have contradictions, flows and changes of concepts.

In real-world, imperfect information is commonly present in all the components of the decision making problem. States of nature reflects possible future conditions which are commonly ill-known whereas the existing theories are based on perfect construction – on partition of the future objective conditions into mutually exclusive and exhaustive states. Possible realizations of future are not completely known. The future may result in a situation which was not thought and unforeseen contingencies commonly take place [26]. From the other side, those states of nature which are supposed as possible, are themselves vaguely defined and it is not always realistic to strictly differentiate among them. The outcomes and probabilities are also not well known, especially taking into account that they are related to ill-known states of nature. However, the existing theories do not pay significant attention to these issues. The most of the theories, including the famous and advanced theories, take into account only imperfect information related to probabilities. Moreover, this is handled by coarse description of ambiguity – either by exact constraints on probabilities (a set of priors) or by using subtle techniques like probabilistic constraints or specific non-linear functions. These are, however, approaches rather for frameworks of the designed experiments but not for real-world decision problems when information is not sufficiently good to apply such techniques. In the Table 3.1. below we tried to classify decision situations on the

Table 3.1 Classification of decision-relevant information

		Probabilities			
Outcomes	Utilities	Precise	Complete Ignorance	Ambiguous	Imperfect
Precise	Precise	Situation 1	Situation 2	Situation 3	Situation 4
	Fuzzy	Situation 5	Situation 6	Situation 7	Situation 8
Complete Ignorance	Precise	Situation 9	Situation 10	Situation 11	Situation 12
	Fuzzy	Situation 13	Situation 14	Situation 15	Situation 16
Ambiguous	Precise	Situation 17	Situation 18	Situation 19	Situation 20
	Fuzzy	Situation 21	Situation 22	Situation 23	Situation 24
Imperfect	Precise	Situation 25	Situation 26	Situation 27	Situation 28
	Fuzzy	Situation 29	Situation 30	Situation 31	Situation 32

base of different types of decision relevant information that one can be faced with and the utility models that can be applied. In this table, we identify three important coordinates (dimensions). The first one concerns information available for probabilities, the second captures information about outcomes, while the third looks at the nature of utilities and their description. The first two dimensions include precise information (risk), complete ignorance (absence of information), ambiguous information, and imperfect information. Two main types of utilities are considered, namely precise and fuzzy. Decision-relevant information setups are represented at the crossing of these coordinates; those are cells containing Situations from 1 to 32. They capture combinations of various types of probabilities, outcomes, and utilities.

The most developed scenarios are those positioned in entries numbered from 1 to 4 (precise utility models). A limited attention has paid to situations 5-8 with fuzzy utilities, which are considered in [5,7,29,51]. For the situations 9-12 with complete ignorance with respect to outcomes and with precise utilities a few works related to interactive obtaining of information were suggested. For situations 13-16, to our knowledge, no works were suggested. Few studies are devoted to the situations with ambiguous outcomes (situations 17-20) [37,38,39] and precise utilities and no works to ambiguous outcomes with fuzzy utilities are available (situations 21-24). For situations 25-32 a very few studies were reported including the existing fuzzy utility models [5,7,29,51]. The case with imperfect probabilities, imperfect outcomes, and fuzzy utilities (situation 32) generalize all the other situations. An adequate utility model for this situation is suggested in [3] and is expressed in Chapter 4 of the present book.

The probability theory has a large spectrum of successful applications. However, the use of a single probability measure for quantification of uncertainty has severe limitations main of which are the following [3]: 1) precise probability is unable to describe complete ignorance (total lack of information); 2) one can determine probabilities of some subsets of a set of possible outcomes but cannot always determine probabilities for all the subsets; 3) one can determine probabilities of all the subsets of a set of possible outcomes but it will require laborious computations.

Indeed, classical probability imposes too strong assumptions that significantly limit its use even in simple real-world or laboratory problems. Famous Ellsberg experiments and Schmeidler's coin example are good illustrative cases when available information appears insufficient to determine actual probabilities. Good discussion of real-world tasks which are incapable to be handled within probabilistic framework is given in [30]. In real problems, quality of decision-relevant information does not require the use of a single probability measure. As a result, probabilities cannot be precisely determined and are imprecise. For such cases, they use constraints on a probability of an event A in form of lower and upper probabilities denoted $\underline{P}(A)$ and $\overline{P}(A)$ respectively. That is, a probability $P(A)$ of an event A is not known precisely but supposed to be somewhere between $\underline{P}(A)$ and $\overline{P}(A)$: $P(A) \in [\underline{P}(A), \overline{P}(A)]$ where $0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1$; in more general formulation, constraints in form of lower and upper expectations for a random variable are used. In special case when $\underline{P}(A) = \overline{P}(A)$ a framework of lower and upper probabilities degenerates to a single probability $P(A)$. Complete lack of knowledge about likelihood of A is modeled by $\underline{P}(A) = 0$ and $\overline{P}(A) = 1$. This means that when likelihood of an event is absolutely unknown, they suppose that probability of this event may take any value from [0,1] (from impossibility to occur up to certain occurrence).

Constraints on probabilities imply existence of a set of probability distributions, that is, multiple priors, which are an alternative approach to handle incomplete information on probabilities. Under the certain consistency requirements the use of multiple priors is equivalent to the use of lower and upper probabilities. Approaches in which imprecise probabilities are handled in form of intervals $[p_1, p_2]$. Such representation is termed as *interval probabilities.*

An alternative way to handle incomplete information on probabilities is the use of non-additive probabilities, typical cases of which are lower probabilities and upper probabilities and their convex combinations. However, multiple priors are more general and intuitive approach to handle incomplete probability information than non-additive probabilities.

The most fundamental axiomatization of imprecise probabilities was suggested by Peter Walley who suggested the term *imprecise probabilities.* The behavioral interpretation of Walley's axiomatization is based on buying and selling prices for gambles. Walley's axiomatization is more general than Kolmogorov's axiomatization of the standard probability theory. The central concept in Walley's theory is the lower prevision concept which generalizes standard (additive) probability, lower and upper probabilities and non-additive measures. However, in terms of generality, the concept of lower prevision is inferior to multiple priors. Another disadvantage of lower prevision theory is its high complexity that limits its practical use.

Alternative axiomatizations of imprecise probabilities were suggested by Kuznetsov [47] and Weichselbergern [75] for the framework of interval probabilities. Weichselberger generalizes Kolmogorov's axioms to the case of interval probabilities but, as compared to Walley, does not suggest a behavioral interpretation. However, his theory of interval probability is more tractable in practical sense.

What is the main common disadvantage of the existing imprecise probability theories? This disadvantage is missing the intrinsic feature of probability-related information which was pointed out by L. Savage even before emergence of the existing imprecise probability theories:

Savage wrote [62]: "…there seem to be some probability relations about which we feel relatively 'sure' as compared with others…. The notion of 'sure' and 'unsure' introduced here is vague, and my complaint is precisely that neither the theory of personal probability as it is developed in this book, nor any other device known to me renders the notion less vague". Indeed, in real-world situations we don't have sufficient information to be definitely sure or unsure in whether that or another value of probability is true. Very often, our sureness stays at some level and does not become complete being hampered by a lack of knowledge and information. That is, sureness is a matter of a strength, or in other words, of a degree. Therefore, 'sure' is a loose concept, a vague concept. In our opinion, the issue is that in most real-world decision-making problems, relevant information perceived by DMs involves possibilistic uncertainty. Fuzzy probabilities are the tools for resolving this issue to a large extent because they represent a degree of truth of a considered numeric probability.

Fuzzy probabilities are superior from the point of view of human reasoning and available information in real-world problems than interval probabilities which are rather the first departs from precise probabilities frameworks. Indeed, interval probabilities only show that probabilities are imprecise and no more. In real-world, the bounds of an interval for probability are subjectively 'estimated' but not calculated or actually known as they are in Ellsberg experiment. Subjective assignments of probability bounds will likely inconsistent with human choices in real-world problems as well as subjective probabilities do in Ellsberg experiment. Reflecting imperfect nature of real-world information, probabilities are naturally soft-constrained.

As opposed to second-order probabilities which are also used to differentiate probability values in terms of their relevance to available information, fuzzy probabilities are more relaxed constructs. Second-order probabilities are too exigent to available information and more suitable for designed experiments.

Fuzzy probability is formally a fuzzy number defined over [0,1] scale to represent an imprecise linguistic evaluation of a probability value. Representing likelihoods of mutually exclusive and exhaustive events, fuzzy probabilities are tied together by their basic elements summing up to one. Fuzzy probabilities define a fuzzy set $\tilde{P}^{\rho}$ of probability distributions ρ which is an adequate representation of imprecise probabilistic information related to objective conditions especially when the latter are vague. As compared to the use of second-order probabilities, the use of possibility distribution over probability distributions [2,3] is appropriate and easier for describing DM's (or experts') confidence. This approach does not require from DM to assign beliefs over priors directly. Possibility distribution can be constructed computationally from fuzzy probabilities assigned to states of nature [7,80]. This means that a DM or experts only need to assign linguistic evaluations of probabilities to states of nature as they usually do. For each linguistic evaluation a fuzzy probability can then be defined by construction of a membership function. After this possibility distribution can be obtained computationally [7,80] without involving a DM.

We can conclude that fuzzy probabilities [8,50,51,66,76] are a successful interpretation of imprecise probabilities which come from human expertise and perceptions being linguistically described. For example, in comparison to multiple priors consideration, for majority of cases, a DM has some linguistic additional information coming from his experience or even naturally present which reflects unequal levels of belief or possibility for one probability distribution or another. This means, that it is more adequate to consider sets of probability distributions as fuzzy sets which allow taking into account various degrees of belief or possibility for various probability distributions. Really, for many cases, some probability distributions are more relevant, some probability distributions are less relevant to the considered situation and also it is difficult to sharply differentiate probabilities that are relevant from those that are irrelevant. This type of consideration involves second-order uncertainty, namely, probability-possibility modeling of decision-relevant information.

The existing utility theories are based on Savage's formulation of states of nature as "*a space of* ***mutually exclusive*** *and* ***exhaustive*** *states*" [62]. This is a perfect consideration of environment structure. However, in real-world problems it is naïve to suppose that we can so perfectly partition future into mutually exclusive objective conditions and predict all possible objective conditions. Future is hidden from our eyes and only some indistinct, approximate trends can be seen. From the other side, unforeseen contingencies are commonly met which makes impossible to determine exhaustive states and also rules out sharp differentiation to exclusive objective conditions. This requires tolerance in describing each objective condition to allow for mistakes, misperceptions, flaws, that are due to imperfect nature of information about future. From the other side, tolerance may also allow for dynamic aspects due to which a state of nature may deviate from its initial condition.

In order to see difficulties with determination of states of nature let us consider a problem of differentiating future economic conditions into states of economy. Commonly, states of economy can be considered as "strong growth", "moderate growth", "stable economy", "recession". These are not 'single-valued' and cannot be considered as 'mutually exclusive' (as it is defined in Savage's formulation of state space): for example, moderate growth and stable economy don't have sharp boundaries and as a result, may not be "exclusive" – they may overlap. The same concerns ‘moderate growth’ and ‘strong growth’ states. For instance, when analyzing the values of the certain indicators that determine a state of economy it is not always possible to definitely label it as moderate growth or strong growth. Observing some actual situation an expert may conclude that it is “somewhere between” ‘strong growth’ and ‘moderate growth’, but “closer” to the latter. This means that to a larger extent the actual situation concerns the moderate growth and to a smaller extent to the strong growth. It is not adequate to sharply differentiate the values related to ‘moderate growth’ from those related to the “strong growth”. In other words, various conditions labeled as “strong growth” with various extents concerns it, not equally. How to take into account the inherent vagueness of states of nature and the fact that they are intrinsically not exclusive but overlapping? Savage’s definition is an idealized view on formalization of objective conditions for such cases. Without doubt, in real-life decision making it is often impossible to construct such an ideal formalization, due to uncertainty of relevant information. In general, a DM cannot exhaustively determine each objective condition that may be faced and cannot precisely differentiate them. Each state of nature is, essentially, some area under consideration which collects in some sense similar objective conditions one can face, that is some set of “elementary” states or quantities [26]. Unfortunately, in the existing decision making theories a small attention is paid to the essence and structure of states of nature, consideration of them is very abstract (formal) and is unclear from human perception point of view.

Formally speaking, a state of nature should be considered as a granule - not some single point or some object with abstract content. This will result to some kind of information granulation of objective conditions. Construction of states of nature on the base of similarity, proximity etc of objective conditions may be adequately modeled by using fuzzy sets and fuzzy information granulation concepts [78]. This will help to model vague and overlapping states of nature. For example, in the considered problem economic conditions may be partitioned into overlapping fuzzy sets defined over some relative scale representing levels of economic welfare. Such formalization will be more realistic for vagueness, ambiguity, partial truth, impreciseness and other imperfectness of future-related information.

In real-life decision making a DM almost always cannot precisely determine future possible outcomes and have to use imprecise quantities like, for example, *high profit*, *medium cost* etc. Such quantities can be adequately represented by ranges of numerical values with possibility distribution among them. From the other side, very often outcomes and utilities are considered in monetary sense, whereas a significantly smaller attention is paid to other types of outcomes and utilities. Indeed, utilities are usually subjectively assigned and, as a result, are heuristic evaluations. In extensive experiments conducted by Kahneman and Tversky, which uncovered very important aspects of human behavior only monetary outcomes are used. Without doubt, monetary consideration is very important, but it is worth to investigate also other types of outcomes which are naturally present in real-life decision activity. In this situation it is not suitable to use precise quantities because subjective evaluations are conditioned also by non-monetary issues such as health, time, reputation, quality etc. The latter are usually described by linguistic evaluations.

In order to illustrate impreciseness of outcomes in real-world problems, let us consider a case of an evaluation of a return from investment into bonds of an enterprise which will produce new products the next year. Outcomes (returns) of investment will depend on future possible economic conditions. Let us suppose these conditions of economy to be partitioned into states of nature labeled as “strong growth”, “moderate growth”, “stable economy”, and “recession” which we considered above. It is impossible to precisely know values of outcomes of the investment under these states of nature. For example, the outcome of the investment obtained under “strong growth” may be evaluated “high” (off course with underlying range of numerical values). The vagueness of outcomes evaluations are resulted from uncertainty about future: impreciseness of a demand for the products produced by the enterprise in the next year, future unforeseen contingencies, vagueness of future economic conditions, political processes etc. Indeed, the return is tightly connected to the demand the next year which cannot be precisely known in advance. The investor does not really know what will take place the next year, but still approximately evaluate possible gains and losses by means of linguistic terms. In other words, the investor is not completely sure in some precise value of the outcome – the future is too uncertain for precise estimation to be reasonably used. The investor sureness is ‘distributed’ among various possible values of the perceived outcome. One way to model is the use of a probabilistic outcome, i.e. to use probability distribution (if discrete set of numerical outcomes is considered) or probability density function (for continuous set) over possible basic outcomes [30] to encode the related objective probabilities or subjective probabilities. However, this approach has serious disadvantages. Using objective probabilities requires good representative data which don’t exist as a demand for a new product is considered. Even for the case of a common product, a good statistics does not exist because demands for various years take place in various environmental conditions. The use of subjective probabilities is also not suitable as they commonly fail to describe human behavior and perception under ambiguity of information.

The use of probabilistic outcomes does not also match human perceptions which are expressed in form of linguistic evaluations of outcomes. Humans don’t think within the probabilistic framework as this is too strong for computational abilities of a human brain; thus, a more flexible formalization is needed to use. Fuzzy set theory provides more adequate representation of linguistic evaluations. By using this theory a linguistic evaluation of an outcome may be formalized by a membership function (a fuzzy set) representing a soft constraint on possible basic outcomes. In contrast to probabilistic constraint, a membership function is not based on strong assumptions and does not require good data. A membership function is directly assigned by a DM to reflect his/her experience, perception, envision etc. which cannot be described by classical mathematics but may act well under imperfect information. Fuzzy sets theory helps to describe future results as imprecise and overlapping, especially under imprecise essence of states of nature. Also, a membership function may reflect various basic outcomes' possibilities, which are much easier to determine than probabilities.

From the other side, the use of fuzzy sets allows to adequately describe non-monetary outcomes like health, reputation, quality which are often difficult to be defined in terms of precise quantities.

3.3 Measures of Uncertainty

Uncertainty is intrinsic to decision making environment. No matter whether we deal with numerical or non-numerical events, we are not completely sure in their occurrence. Numerical events are commonly regarded as values of a random variable. Non-numerical events can be encoded by, for example, natural numbers and then treated as values of a random variable. To formally take into account uncertainty in decision analysis, we need to use some mathematical constructs which will measure quantitatively an extent to what that or another event is likely to occur. Such constructs are called measures of uncertainty. The most famous measure is the probability measure. Probability measure assigns its values to events to reflect degrees to which events are likely to occur. These values are called probabilities. Probability is a real number from [0,1], and the more likely an event to occur the higher is its probability. Probability equal to 0 implies that it is impossible for an event to occur or we are completely sure that it cannot occur, and probability equal to 1 means that an event will necessary occur or we are completely sure in its occurrence. The axiomatization of the standard probability measure was suggested by Kolmogorov [46]. Prior to proceeding to the Kolmogorov's axiomatization, let us introduce the necessary concepts. The first concept is the space of elementary events. Elementary event, also called an atomic event, is the minimal event that may occur, that is, an event that cannot be divided into smaller events. Denote S the space of elementary events and denote $s \in S$ an elementary event. A subset $H \subseteq S$ of the space of elementary events $s \in S$ is called an event. An event H occurs if any $s \in H$ occurs. The next concept is a σ-algebra of subsets denoted $\mathcal{F}$.

Definition 3.2. σ-algebra [46]**.** A set $\mathcal{F}$, elements of which are subsets of S (not necessarily all) is called σ-algebra if the following hold:

(1) $S \in \mathcal{F}$

(2) if $H \in \mathcal{F}$ then $H^c \in \mathcal{F}$

(3) if $H_1, H_2, ... \in \mathcal{F}$ then $H_1 \cup H_2 \cup ... \in \mathcal{F}$

Now let us proceed to the Kolmogorov's axiomatization of a probability measure.

Definition 3.3. Probability Measure [46]**.** Let S be a space of elementary events and $\mathcal{F}$ is a σ-algebra of its subsets. The probability measure is a function $P:\mathcal{F}\to[0,1]$ satisfying:

(1) $P(H)\geq 0$ for any $H\in\mathcal{F}$.

(2) For any set $H_1,H_2,...\in\mathcal{F}$ with $H_i\cap H_j,...\varnothing$: $P(\bigcup_{i=1}^{\infty}H_i)=\sum_{i=1}^{\infty}P(H_i)$

(3) $P(S)=1$

The first condition is referred to as non-negativity. The second condition is referred to as additivity condition. The third condition implies that the event S will necessary occur. Conditions (1)-(3) are called probability axioms. From (1)-(3) it follows $P(\varnothing)=0$ which means that it is impossible that no elementary event $s\in S$ will occur. Let us mention that probability of a union $H\cup G$ of two arbitrary events is $P(H\cup G)=P(H)+P(G)-P(H\cap G)$. When $H\cap G=\varnothing$ one has $P(H\cup G)=P(H)+P(G)$.

Definition 3.3 provides mathematical structure of a probability measure. Consider now natural interpretations of a probability measure. There exist two main types of probabilities: objective probabilities and subjective probabilities. Objective probabilities, also called empirical probabilities, are quantities which are calculated on the base of real evidence: experimentations, observations, logical conclusions. They also can be obtained by using statistical formulas. Objective probabilities are of two types: experimental probabilities and logical probabilities. Experimental probability of an event is a frequency of its occurrence. For example, a probability that a color of a car taken at random in a city is white is equal to the number of white cars divided by the whole number of the cars in the city. Logical probability is based on a reasonable conclusion on a likelihood of an event. For instance, if a box contains 70 white and 30 black balls, a probability that a ball drawn at random is white is 70/100=0.7.

The use of objective probabilities requires very restrictive assumptions. For experimental probabilities the main assumptions are as follows:

(1) Experimentation (or observations) must take place under the same conditions and it must be assumed that the future events will also take place under these conditions. Alternatively, there need to be present clear dynamics of conditions in future;
(2) Observations of the past behavior must include representative data (e.g., observations must be sufficiently large).

As to logical probabilities, their use must be based on quite reasonable conclusions. For example, if to consider the box with balls mentioned above, an assumption must be made that the balls are well mixed inside the box (not a layer of white balls is placed under the layer of black balls) to calculate probability of a white ball drawn as 70/100=0.7.

From the other side, as Kahneman, Tversky and others showed [42] , that even if objective probabilities are known, beliefs of a DM don't coincide with them. As being perceived by humans, objective probabilities are affected by some kind of distortion – they are transformed into so-called decision weights and mostly small probabilities are overweighted, whereas high probabilities are underweighted. The overweighting and underweighting of probabilities also are different for positive and negative outcomes [68].

Subjective probability is a degree of belief of an individual in the likelihood of an event. Formally, subjective probabilities are values of a probability measure. From interpretation point of view, subjective probability reflects an individual's experience, perceptions and is not based on countable and, sometimes, detailed facts like objective probability. Subjective probabilities are more appropriate and 'smart' approach for measuring likelihood of events in real-life problems because in such problems the imperfect relevant information conflicts with the very strong assumptions underlying the use of objective probabilities. Real-life relevant information is better handled by experience and knowledge that motivates the use of a subjective basis.

Subjective probability has a series of disadvantages. One of the main disadvantages is that different people would assign different subjective probabilities. It is difficult to reason uniquely accurate subjective probabilities among those assigned by different people. Indeed, given a lack of information, people have to guess subjective probabilities as they suppose. As it is mentioned in [52], using subjective probabilities is a "symptom of the problem, not a solution". Subjective probability is based not only on experience but also on feelings, emotions, psychological and other factors which can distort its accuracy. The other main disadvantage is that subjective probability, due to its preciseness and additivity, fails to describe human behavior under ambiguity.

The use of the additive probability measure is unsuitable to model human behavior conditioned by uncertainty of the real-world, psychological, mental and other factors. In presence of uncertainty, when true probabilities are not exactly known, people often tend to consider each alternative in terms of the worst case within the uncertainty and don't rely on good cases. In other words, most of people prefer those decisions for which more information is available. This behavior is referred to as ambiguity aversion – people don't like ambiguity and wish certainty. Even when true probabilities are known, most people exhibit non-linear attitude to probabilities – change of likelihood of an event from impossibility to some chance or from a very good chance to certainty are treated much more strongly than the same change somewhere in the range of medium probabilities. Therefore, attitude to values of probabilities is qualitative.

Due to its additivity property, the classical (standard) probability measure cannot reflect the above mentioned evidence. Axiomatizations of such evidence required to highly weaken assumptions on a DM's belief which was considered as the probability measure. The resulted axiomatizations are based either on non-uniqueness of probability measure or on non-additivity of a measure of uncertainty reflecting humans' beliefs. The first axiomatization of choices based on a non-additive measure was suggested by Schmeidler [63]. This is a significant generalization of additive measures-based decisions because the uncovered non-additive measure inherited only normalization and monotonicity properties from the standard probability measure.

Nowadays non-additive measures compose a rather wide class of measures of uncertainty. Below we list non-additive measures used in decision making under ambiguity. For these measures a unifying term *non-additive probability* is used.

We will express the non-additive probabilities in the framework of decision making under ambiguity. Let S be a non-empty set of states of nature and $\mathcal{F}$ be a family of subsets of S. We will consider w.l.o.g. $\mathcal{F}= 2^S$.

The definition of a non-additive probability is as follows [63].

Definition 3.4 [63]**. Non-additive Probability.** A set function $v:\mathcal{F}\to[0,1]$ is called a non-additive probability if it satisfies the following:

1. $v(\varnothing)=0$
2. $\forall H,G\in\mathcal{F}$, $H\subset G$ implies $v(H)\le v(G)$
3. $v(S)=1$

The non-additive probability is also referred to as *Choquet capacity*. Condition (2) is called monotonicity with respect to set inclusion and conditions (1) and (3) are called normalization conditions. Thus, a non-additive probability does not have to satisfy $v(H\cup G)=v(H)+v(G)$. A non-additive probability is called *super-additive* if $v(H\cup G)\ge v(H)+v(G)$ and *sub-additive* if $v(H\cup G)\le v(H)+v(G)$, provided $H\cap G=\varnothing$.

There exist various kinds of non-additive probability many of which are constructed on the base of a set C of probability measures P over S. The one of the well known non-additive probabilities is the *lower envelope* $v_*:\mathcal{F}\to[0,1]$ which is defined as follows:

$$v_*(H)=\min_{P\in C}P(G) \tag{3.1}$$

The dual concept of the lower envelope is the *upper envelope* $v^*:\mathcal{F}\to[0,1]$ which is defined by replacing $\min$ operator in (3.1) by $\max$ operator. Lower and upper envelopes are respectively minimal and maximal probabilities of an event $H\subset S$. Therefore, $v_*(H)\le P(H)\le v^*(H),\forall H\subset S,P\in C$. Lower envelope is super-additive, whereas upper envelope is sub-additive. A non-additive probability can also be defined as a convex combination of $v_*(H)$ and

$v^*(H)$: $v(H)=\alpha v_*(H)+(1-\alpha)v^*(H)$, $\alpha\in[0,1]$. The parameter α is referred to as a degree of ambiguity aversion. Indeed, α is an extent to which belief $v(H)$ is based on the smallest possible probability of an event H ; $1-\alpha$ is referred to as a degree of ambiguity seeking.

The generalizations of lower and upper envelopes are *lower* and *upper probabilities* which are respectively super-additive and sub-additive probabilities. Lower and upper probabilities, denoted respectively $\underline{v}$ and $\overline{v}$, satisfy $\underline{v}(H)=1-\overline{v}(H^c)\ \forall H\in S$, where $H^c=S\setminus H$.

The special case of lower envelopes and, therefore, of lower probabilities are *2-monotone Choquet capacities*, also referred to as *convex capacities*. A non-additive probability is called 2-monotone Choquet capacity if it satisfies

$$v(H\cup G)\geq v(H)+v(G)-v(H\cap G),\ \forall H,G\subset S$$

A generalization of *2-monotone* capacity is an *n-monotone* capacity. A capacity is an *n-monotone*, if for any sequence $H_1,...,H_n$ of subsets of S the following holds:

$$v(H_i\cup...\cup H_n)\geq\sum_{\substack{I\subset\{1,...,n\}\\ I\neq\varnothing}}(-1)^{|I|-1}v\left(\bigcap_{i\in I}H_i\right)$$

A capacity which is *n-monotone* for all n is called *infinite monotone capacity* or a *belief function*.

The belief function theory, also known as Dempster-Shafer theory, or mathematical theory of evidence, or theory of random sets, was suggested by Dempster in [15], and developed by Shafer in [64]. Belief functions are aimed to be used for describing subjective degrees of belief to an event, phenomena, or object of interest. We will not explain this theory but just mention that it was not directly related to decision making. As it was shown in [33,34], axiomatization of this theory is a generalization of the Kolmogorov's axioms of the standard probability theory. Due to this fact, a value of a belief function denoted $Bel()$ for an event H can be considered as a lower probability, that is, as a lower bound on a probability of an event H . An upper probability in the belief function theory is termed as a *plausibility function* and is denoted Pl . So, in the belief functions theory probability $P(H)$ of an event H is evaluated as $Bel(H)\leq P(H)\leq Pl(H)$.

The motivation of using non-additive probabilities in decision making problems is the fact that information on probabilities is imperfect, which can be incomplete, imprecise, distorted by psychological factors etc. Non-additive measure can be determined from imprecise objective or subjective probabilities of states of nature. Impreciseness of objective probabilities can be conditioned by the lack of information ruling out determination of actual exact probabilities (as in Ellsberg experiments). Impreciseness of subjective probabilities can be conditioned by natural impreciseness of human beliefs. Let us consider the case when imprecise information is represented in form of interval probabilities. Given a set of states of nature $S=\{s_1,s_2,...,s_n\}$, interval probabilities are defined as follows [32].

Definition 3.5 [32]**. Interval Probability.** The intervals $\overline{P}(s_i)=[p_{*i},p_i^*]$ are called the interval probabilities of S if for $p_i\in[p_{*i},p_i^*]$ there exist $p_1\in[p_{*1},p_1^*],\ldots,p_{i-1}\in[p_{*i-1},p_{i-1}^*]$, $p_{i+1}\in[p_{*i+1},p_{i+1}^*]$,…, $p_n\in[p_{*n},p_n^*]$ such that

$$\sum_{i=1}^{n} p_i = 1$$

From this definition it follows, in particular, that interval probabilities cannot be directly assigned as numerical probabilities. The issue is that in the case of interval probabilities, the requirement to numerical probabilities to sum up to one must be satisfied throughout all the probability ranges. Sometimes, interval probabilities $\overline{P}(s_i)=\overline{P}_i$ can be directly assigned consistently to $n-1$ states of nature $s_1,s_2,...,s_{j-1},s_{j+1},...,s_n$, and on the base of these probabilities, an interval probability $\overline{P}(s_j)=\overline{P}_j$ for the rest one state of nature s_j will be calculated. For example, consider a set of states of nature with three states $S=\{s_1,s_2,s_3\}$. Let interval probabilities for s_2 and s_3 be assigned as follows:

$$\overline{P}_2=[0.2,0.3],\ \overline{P}_3=[0.5,0.6]$$

Then, according to the conditions in Definition 3.5, $\overline{P}_1$ will be determined as follows:

$$\overline{P}_1=[1-0.3-0.6,\,1-0.2-0.5]=[0.1,0.3].$$

Given interval probabilities $\overline{P}_i=[p_{*i},p_i^*]$ of states of nature $s_i, i=1,...,n$ a value $v_*(A)$ of a lower probability for an event A can be determined as follows:

$$\begin{aligned}
&v_*(A)=\min\sum_{s_i\in A} p_i\\
&s.t.\\
&\sum_{i=1}^{n} p_i = 1 \qquad\qquad (3.2)\\
&p_i \le p_i^*\\
&p_i \ge p_{*i}
\end{aligned}$$

A value $v^*(A)$ of an upper probability for an event A can be determined by replacing *min* operator by *max* operator in the above mentioned problem. Consider an example. Given interval probabilities $\overline{P}_1=[0.1,0.3]$, $\overline{P}_2=[0.2,0.3]$, $\overline{P}_3=[0.5,0.6]$, the values of the lower and upper probabilities v_* and v^* for $A=\{s_1,s_3\}$, obtained as solutions of the problem (3.2) are $v_*(A)=0.7$ and $v^*(A)=0.8$.

The above mentioned measures of uncertainty can be listed in terms of the increasing generality between the probability measure and the Choquet capacity as follows:

probability measure $\Rightarrow$ belief function $\Rightarrow$ convex capacity $\Rightarrow$ lower envelope $\Rightarrow$ lower probability $\Rightarrow$ Choquet capacity

In [48], it is suggested a decision model based on a new kind of measure called bi-capacity. Bi-capacity is used to model interaction between 'good' and 'bad' performances with respect to criteria. As compared to capacity, bi-capacity is a two-place set function. The values the bi-capacity takes are from [-1,1]. More formally, the bi-capacity is defined as a set function

$$\eta : W \rightarrow [-1,1], \text{ where } W=\{(H,G): H,G\subset I, H\cap G=\varnothing\}$$

satisfying

$$H\subset H' \Rightarrow \eta(H,G)\le\eta(H',G),\ G\subset G' \Rightarrow \eta(H,G')\le\eta(H,G)$$

and

$$\eta(I,\varnothing)=1,\ \eta(\varnothing,I)=-1,\ \eta(\varnothing,\varnothing)=0.$$

I is the set of indexes of criteria. The attributes in H are satisfied attributes whereas the attributes in G are dissatisfied ones. The integral with respect to bi-capacity as a representation of an overall utility of an alternative $f: I\rightarrow R$ is defined as follows:

$$U(f)=\sum_{l=1}^{n}(u(f_{(l)})-u(f_{(l+1)}))\eta(\{(1),...,(l)\}\cap I^+,\{(1),...,(l)\}\cap I^-), \qquad (3.3)$$

provided $u(f_{(l)})\ge u(f_{(l+1)})$; $I^+=\{i\in I: u(f_i)\ge 0\}$, $I^-=I\setminus I^+$ where $u(f_{(l)})$ is a utility of a value of *(l)*-th criterion for f, $\eta(\cdot,\cdot)$ is a bi-capacity.

In special case, when η is equal to the difference of two capacities η_1 and η_2 as $\eta(H,G)=\eta_1(H)-\eta_2(G)$, (3.3) reduces to the CPT model. In general case, as compared to CPT, (3.3) is not an additive representation of separately aggregated satisfied and dissatisfied criteria that provides more smart way for decision making.

The disadvantage of a bi-capacity relates to difficulties of its determination, in particular, to computational complexity. In details the issues are discussed in [48].

The bi-capacity-based aggregation which was axiomatized for multicriteria decision making [48] can also be applied for decisions under uncertainty due to formal parallelism between these two problems [53]. Indeed, states of nature are criteria on base of which alternatives are evaluated.

The non-additive measures provide a considerable success in modeling of decision making. However, the non-additive measures only reflect the fact that human choices are non-additive and monotone, which may be due to attitudes to uncertainty, distortion of probabilities etc, but nothing more. However, in real-world it is impossible to accurately determine precisely the 'shape' of a non-additive measure due to imperfect relevant information. Indeed, real-world probabilities of subsets and subsets themselves, outcomes, interaction of criteria, etc are imprecisely and vaguely defined. From the other side, attitudes to uncertainty, extent of probabilities distortion and other behavioral issues violating additivity are also imperfectly known. These aspects rule out exact determination of a uniquely accurate non-additive measure.

Above we considered non-additive measures which are used in the existing decision theories to model non-additivity of DM's behavior. Main shortcoming of using non-additive measures is the difficulty of the underlying interpretation. One approach to overcome this difficulty is to use a lower envelope of a set of priors as a non-additive probability and then to use it in CEU model. However, in real-world problems determination of the set of priors itself meets difficulty of imposing precise constraint determining what prior should be included and what should not be included into this set. In other words, due to lack of information, it is impossible to sharply constraint a range for a probability of a state of nature, that is, to assign accurate interval probability. From the other side, if the set of priors is defined, why to construct lower envelope and use it in the CEU? It is computationally simpler to use the equivalent model – MMEU. Let us consider very important a class of non-additive measures called *fuzzy measures*. Fuzzy measures have their own interpretations that do not require using a set of priors to define them and makes construction of these measures computationally simple. Finally, we will consider an effective extension of non-additive measures called *fuzzy-valued* fuzzy measures which have a good suitability for measuring vague real-world information.

The first fuzzy measure we consider is a *possibility measure*. Possibility means an ability of an event to occur. It was recently mentioned that probability of an event can hardly be determined due to a series of reason, whereas possibility of occurrence of an event is easier to be evaluated. Possibility measure has also its interpretation in terms of multiple priors.

Possibility measure [77] is a non-additive set function $\Pi : \mathcal{F}(S) \to [0,1]$ defined over a σ-algebra $\mathcal{F}(S)$ of subsets of S and satisfying the following conditions:

(1) $\Pi(\varnothing) = 0$

(2) $\Pi(S) = 1$

(3) For any collection of subsets $H_i \in \mathcal{F}(S)$ and any set of indexes I the following holds:

$$\Pi(\bigcup_{i \in I} H_i) = \sup_{i \in I} \Pi(H_i)$$

Possibility measure Π can be represented by *possibility distribution function,* or possibility distribution, for short. Possibility distribution is a function $\pi : S \to [0,1]$ and by means of π possibility measure Π is determined as follows:

$$\Pi(H) = \sup_{s \in H} \pi(s)$$

Condition (2) predetermines normalization condition $\sup_{s \in S} \pi(s) = 1$. Given S as a set of states of nature, possibility measure provides information on possibility of occurrence of an event $H \subset S$. A possibility distribution π_1 is more informative than π_2 if $\pi_1(s) \le \pi_2(s), \forall s \in S$.

The dual concept of the possibility is the concept of necessity. Necessity measure is a set function $N : P(S) \to [0,1]$ that is defined as $N(H) = 1 - \Pi(H^c)$, $H^c = S \setminus H$. This means, for example, that if an event H is necessary (will necessary happen), then the opposite event H^c is impossible.

From the definitions of possibility and necessity measures one can find that the following hold:

1) $N(H) \le \Pi(H)$
2) if $\Pi(H) < 1$ then $N(H) = 0$
3) if $N(H) > 0$ then $\Pi(H) = 1$
4) $\max(\Pi(H), \Pi(H^c)) = 1$
5) $\min(N(H), N(H^c)) = 0$

The possibility differs from probability in various aspects. First, possibility of two sets H and G provided $H \cup G = \varnothing$ is equal to the maximum possibility among those of H and $G : \Pi(H \cup G) = \max(\Pi(H), \Pi(G))$. In its turn probability $H \cup G$ is equal to the sum of those of H and G: $P(H \cup G) = P(H) + P(G)$.

Another difference between possibility measure and probability measure is that the first is compositional that make it more convenient from computational point of view. For example, given $P(H)$ and $P(G)$, we cannot determine precisely $P(H \cup G)$, but can only determine its lower bound which is equal to $\max(P(H), P(G))$ and an upper bound which is equal to $\min(P(H) + P(G), 1)$. At the same time possibility of $H \cup G$ is exactly determined based on $\Pi(H)$ and $\Pi(G): \Pi(H \cup G) = \max(\Pi(H), \Pi(G))$. However, the possibility of an intersection is not exactly defined: it is only known that $\Pi(A \cap B) \leq \min(\Pi(A), \Pi(B))$. As to necessity measure, it is exactly defined only for an intersection of sets: $N(H \cap G) = \min(N(H), N(G))$.

Yet another difference is that as compared to probability, possibility is able to model complete ignorance, that is, absence of any information. Absence of any information about H is modeled in the possibility theory as $\Pi(H) = \Pi(H^c) = 1$ and $N(H) = N(H^c) = 0$. From this it follows $\max(\Pi(H), \Pi(H^c)) = 1$ and $\min(N(H), N(H^c)) = 0$.

The essence of the possibility is that it models rather qualitative information about events than quantitative one. Possibility measure only provides ranking of events in terms of their comparative possibilities. For example, $\pi(s_1) \leq \pi(s_2)$ implies that s_1 is more possible than s_2. $\pi(s) = 0$ implies that occurrence of s is impossible whereas $\pi(s) = 1$ implies that s is one of the most possible realizations. The fact that possibility measure may be used only for analysis at qualitative, comparative level [69], was proven by Pytyev in [60], and referred to as the principle of relativity in the possibility theory. This principle implies that possibility measure cannot be used to measure actual possibility of an event but can only be used to determine whether the possibility of one event is higher, equal to, or lower than the possibility of another event. Due to this feature, possibility theory is less self-descriptive than probability theory but requires much less information for analysis of events than the latter.

One of the interpretations of possibility measure is an upper bound of a set of probability measures [18,62,72]. Let us consider the following set of probability measures coherent with possibility measure Π :

$$\mathcal{P}(\Pi) = \{P : P(H) \leq \Pi(H), \forall H \subseteq S\}$$

Then the upper bound of probability for an event H is

$$\overline{P}(H) = \sup_{P \in \mathcal{P}(\Pi)} P(H)$$

and is equal to possibility $\Pi(H)$. The possibility distribution is then defined as

$$\pi(s) = \overline{P}(\{s\}), \forall s \in S$$

Due to normalization condition $\sup_{s\in S}\pi(s)=1$, the set $\mathcal{P}(\Pi)$ is always not empty. In [18,72] they show when one can determine a set of probability measures given possibility measure.

Analogous interpretation of possibility is its representation on the base of lower and upper bounds of a set of distribution functions. Let information about unknown distribution function F for a random variable X is described by means of a lower $\underline{F}$ and an upper $\overline{F}$ distribution functions: $\underline{F}(x)\le F(x)\le \overline{F}(x), \forall x\in X$. The possibility distribution π then may be defined as

$$\pi(x)=\min(\overline{F}(x),1-\underline{F}(x)).$$

Baudrit and Dubois showed that a set of probabilities generated by possibility distribution π is more informative than a set of probabilities generated by equivalent distribution functions.

In order to better explain what possibility and necessity measures are, consider an example with a tossed coin. If to suppose that heads and tails are equiprobable, then the probabilities of heads and tails will be equal to 0.5. As to possibilities, we can accept that both heads and tails are very possible. Then, we can assign the same high value of possibility to both events, say 0.8. At the same time, as the result of tossing the coin is not intentionally designed, we can state that the necessity of both events is very small. It also follows from $\mathrm{N}(\{heads\})=1-\Pi(\{tails\}), \mathrm{N}(\{tails\})=1-\Pi(\{heads\})$. As this example suggests, we can state that possibility measure may model ambiguity seeking (hope for a good realization of uncertainty), where as necessity measure may model ambiguity aversion.

One of the most practically efficient and convenient fuzzy measures are *Sugeno* measures. Sugeno measure is a fuzzy measure $g: P(S)\to[0,1]$ that satisfies

(1) $g(\varnothing)=0$,

(2) $g(S)=1$;

(3) $H\subset G\Rightarrow g(H)\le g(G)$;

(4) $H_i\uparrow H$ or $H_i\downarrow H\Rightarrow \lim_{i\to+\infty} g(H_i)=g(H)$

From these conditions it follows $g(H\cup G)\ge \max(g(H),g(G))$ and $g(H\cap G)\le \min(g(H),g(G))$. In special case, when $g(H\cup G)= =\max(g(H),g(G))$, Sugeno measure g is the possibility measure. When $g(H\cap G)=\min(g(H),g(G))$, Sugeno measure g is the necessity measure.

The class of Sugeno measures that became very widespread due to its practical usefulness are g_λ measures. g_λ measure is defined by the following condition referred to as the λ-rule:

$$g_\lambda(H \cup G) = g_\lambda(H) + g_\lambda(G) + \lambda g_\lambda(H) g_\lambda(G),\ \lambda \in [-1, +\infty)$$

For the case of $H = S$, this condition is called normalization rule. λ is called normalization parameter of g_λ measure. For $\lambda > 0$ g_λ measure satisfy $g_\lambda(H \cup G) > g_\lambda(H) + g_\lambda(G)$ that generates a class of superadditive measures. For $\lambda > 0$ one gets a class of subadditive measures: $g_\lambda(H \cup G) <$ $< g_\lambda(H) + g_\lambda(G)$. The class of additive measures is obtained for $\lambda = 0$.

One type of fuzzy measures is defined as a linear combination of possibility measure and probability measure. This type is referred to as g_v measure. g_v measure is a fuzzy measure that satisfies the following:

(1) $g_v(\varnothing) = 0$

(2) $g_v(S) = 1$

(3) $\forall i \in N,\ H_i \in \mathcal{F}(S), \forall i \neq j$

(4) $H_i \cap H_j = \varnothing \Rightarrow g_v\left(\bigcup_{i \in N} H_i\right) = (1 - v) \bigvee_{i \in N} g_v(H_i) + v \sum_{i \in N} g_v(H_i),\ v \geq 0$

(5) $\forall H, G \in \mathcal{F}(S) : H \subseteq G \Rightarrow g_v(H) \leq g_v(G)$

g_v is an extension of a measure suggested by Tsukamoto which is a special case obtained when $v \in [0,1]$ [67]. For $v \in [0,1]$ one has a convex combination of possibility and probability measures. For purposes of decision making this can be used to model behavior which is inspired by a mix of probabilistic judgement and an extreme non-additive reasoning, for instance, ambiguity aversion. Such modeling may be good as reflecting that a person is not only an uncertainty averse but also thinks about some ‘average’, i.e. approximate precise probabilities of events. This may be justified by understanding that, from one side, in real-world situations we don’t know exactly the boundaries for a probability of an event. From the other, we don’t always exhibit pure ambiguity aversion by try to guess some reasonable probabilities in situations of ambiguity.

When $v = 0$, g_v is the possibility measure and when $v = 1$, g_v is the probability measure. For $v = 1$, g_v describes uncertainty that differs from both probability and possibility [59].

Fuzzy measures are advantageous type of non-additive measures as compared to non-additive probabilities because they mainly have clear interpretation and some of them are “self-contained”. The latter means that some fuzzy measures, like possibility measure, don’t require a set of priors for their construction. Moreover, a fuzzy measure can be more informative than a set of priors or a set of priors can be obtained from it. Despite of these advantages, fuzzy measures are also not sufficiently adequate for solving real-world decision problems. The issue is that fuzzy measures suffer from the disadvantage of all the widespread additive and non-additive measures: fuzzy measures are numerical representation of uncertainty. In contrast, real-world uncertainty cannot be precisely described – it is not to be caught by a numerical function. This aspect is, in our opinion, one of the most essential properties of real-world uncertainty.

The precise non-additive measures match well the backgrounds of decision problems of the existing theories which are characterized by perfect relevant information: mutually exclusive and exhaustive states of nature, sharply constrained probabilities. However, as we discussed above chapter, real-world decision background is much more 'ill-defined'. Essence of information about states of nature makes them rather blurred and overlapping but not perfectly separated. For example, evaluations like 'moderate growth' and 'strong growth' of economy cannot be precisely bounded and may overlap to that or another extent. This requires to use fuzzy sets as more adequate descriptions of real objective conditions. Probabilities of states of nature are also fuzzy as they cannot be sharply constrained. This is conditioned by lack of specific information, by the fact that human sureness in occurrence of events stays in form of linguistic estimations like "very likely", "probability is medium", "probability is small" etc which are fuzzy. From the other side, this is conditioned by fuzziness of states of nature themselves. When the "strong growth" and "moderate growth" and their likelihoods are vague and, therefore, relations between them are vague – how to obtain precise measure? Natural impreciseness, fuzziness related to states of nature must be kept as the useful data medium in passing from probabilities to a measure – the use of precise measure cannot be sufficiently reasonable and leads to loss of information. From the other side, shape of non-additivity of a DM's behavior cannot be precisely determined, whereas some linguistic, approximate, but still ground relevant information can be obtained. Fuzziness of the measure in this case serves as a good interpretation.

Thus, a measure which models human behavior under real-world imperfect information should be considered not only as non-additive, but also as fuzzy imprecise quantity that will reflect human evaluation technique based on, in general, linguistic assessments. In this sense a more adequate construction that better matches imperfect real-world information is *a fuzzy number-valued fuzzy measure.* Prior to formally express what is a fuzzy number-valued measure, let us introduce some formal concepts. The first concept is a set of fuzzy states of nature $\mathcal{S}=\left\{\tilde{S}_1,...,\tilde{S}_n\right\}$, where $\tilde{S}_i, i=1,...,n$ is a fuzzy set defined over a universal set U in terms of membership function $\mu_{\tilde{S}_i}:U\to[0,1]$. The second concept relates to comparison of fuzzy numbers:

Definition 3.6. [82]. For $\tilde{A},\tilde{B}\in\mathcal{E}^1$, we say that $\tilde{A}\le\tilde{B}$, if for every $\alpha\in\left(0,1\right]$, $A_1^\alpha\le B_1^\alpha$ and $A_2^\alpha\le B_2^\alpha$.

We consider that $\tilde{A}<\tilde{B}$, if $\tilde{A}\le\tilde{B}$, and there exists an $\alpha_0\in\left(0,1\right]$ such that $A_1^{\alpha_0}<B_1^{\alpha_0}$, or $A_2^{\alpha_0}<B_2^{\alpha_0}$.

We consider that $\tilde{A}=\tilde{B}$ if $\tilde{A}\le\tilde{B}$, and $\tilde{B}\le\tilde{A}$

The third concept is a fuzzy infinity:

Definition 3.7. Fuzzy Infinity [82]**.** Let $\tilde{A}$ be a fuzzy number. For every positive real number M, there exists a $\alpha_0\in(0,1]$ such that $M<A_2^{\alpha_0}$ or $A_1^{\alpha_0}<-M$. Then $\tilde{A}$ is called fuzzy infinity, denoted by $\tilde{\infty}$.

Now denote $\mathcal{E}_+^1=\left\{\tilde{A}\in\mathcal{E}\,\middle|\,\tilde{A}\ge 0\right\}$. Thus, $\mathcal{E}_+^1$ is a space of fuzzy numbers defined over $\mathcal{R}_+$. Let Ω be a nonempty finite set and $\mathcal{F}$ be σ-algebra of subsets of Ω. A definition of a fuzzy number-valued fuzzy measure as a monotone fuzzy number-valued set function suggested by Zhang [82] and referred to as a (z)-fuzzy measure, is as follows:

Definition 3.8. Fuzzy Number-Valued Fuzzy Measure ((z)-Fuzzy Measure) [82]**.** A (z) fuzzy measure on $\mathcal{F}$ is a fuzzy number-valued fuzzy set function $\tilde{\eta}:\mathcal{F}\to\mathcal{E}_+^1$ with the properties:

(1) $\tilde{\eta}(\varnothing)=0$;

(2) if $\mathcal{H}\subset\mathcal{G}$ then $\tilde{\eta}(\mathcal{H})\le\tilde{\eta}(\mathcal{G})$;

(3) if $\mathcal{H}_1\subset\mathcal{H}_2\subset...,\mathcal{H}_n\subset...\in\mathcal{F}$, then $\tilde{\eta}(\bigcup_{n=1}^{\infty}\mathcal{H}_n)=\lim_{n\to\infty}\tilde{\eta}(\mathcal{H}_n)$;

(4) if $\mathcal{H}_1\supset\mathcal{H}_2\supset...,\mathcal{H}_n\in\mathcal{F}$, and there exists n_0 such that $\tilde{\eta}(\mathcal{H}_{n_0})\ne\tilde{\infty}$, then $\tilde{\eta}(\bigcap_{n=1}^{\infty}\mathcal{H}_n)=\lim_{n\to\infty}\tilde{\eta}(\mathcal{H}_n)$.

Here limits are taken in terms of supremum metric d [16,49]. A pair $(\Omega,\tilde{\mathcal{F}}(\Omega))$ is called a fuzzy measurable space and a triple $(\Omega,\tilde{\mathcal{F}}(\Omega),\tilde{\eta})$ is called a (z) fuzzy measure space.

So, a fuzzy number-valued fuzzy measure $\tilde{\eta}:\mathcal{F}\to\mathcal{E}_+^1$ assigns to every subset of Ω a fuzzy number defined over [0,1]. Condition (2) of Definition 3.8 is called monotonicity condition. $\tilde{\eta}:\mathcal{F}\to\mathcal{E}_+^1$ is monotone and is free of additivity requirement. Consider an example. Let $\Omega=\{\omega_1,\omega_2,\omega_3\}$. The values of the fuzzy number-valued set function $\tilde{\eta}$ for the subsets of Ω can be as the triangular fuzzy numbers given in Table 3.2:

Table 3.2 The values of the fuzzy number-valued set function $\tilde{\eta}$

$\mathcal{H}\subset\Omega$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_1,\omega_2\}$	$\{\omega_1,\omega_3\}$	$\{\omega_2,\omega_3\}$
$\tilde{\eta}(\mathcal{H})$	(0.3,0.4,0.4)	(0,0.1,0.1)	(0.3,0.5,0.5)	(0.3,0.5,0.5)	(0.6,0.9,0.9)	(0.3,0.6,0.6)

Fuzzy number-valued set function $\tilde{\eta}$ is a fuzzy number-valued fuzzy measure. For instance, one can verify that condition 2 of Definition 3.8 for $\tilde{\eta}$ is satisfied.

Let us consider $\tilde{\eta}$ as a fuzzy number-valued lower probability constructed from linguistic probability distribution $\tilde{P}^l$:

$$\tilde{P}^l = \tilde{P}_1 / \tilde{S}_1 + \tilde{P}_2 / \tilde{S}_2 + \ldots + \tilde{P}_n / \tilde{S}_n$$

Linguistic probability distribution $\tilde{P}^l$ implies that a state $\tilde{S}_i \in \mathcal{S}$ is assigned a linguistic probability $\tilde{P}_i$ that can be described by a fuzzy number defined over [0,1]. Let us shortly mention that the requirement for numeric probabilities to sum up to one is extended for linguistic probability distribution $\tilde{P}^l$ to a wider requirement which includes degrees of consistency, completeness and redundancy that will be described in details in Chapter 4. Given $\tilde{P}^l$, we can obtain from it a fuzzy set $\tilde{P}^\rho$ of possible probability distributions $\rho(s)$. We can construct a fuzzy-valued fuzzy measure from $\tilde{P}^\rho$ as its lower probability function [55] by taking into account a degree of correspondence of $\rho(s)$ to $\tilde{P}^l$. A degree of membership of an arbitrary probability distribution $\rho(s)$ to $\tilde{P}^\rho$ (a degree of correspondence of $\rho(s)$ to $\tilde{P}^l$) can be obtained by the formula

$$\pi_{\tilde{P}}(\rho(s)) = \min_{\underline{i=1,n}}(\pi_{\tilde{P}_i}(p_i)) \ ,$$

where $p_i = \int_S \rho(s)\mu_{\tilde{S}_i}(s)ds$ is numeric probability of fuzzy state $\tilde{S}_i$ defined by $\rho(s)$. $\pi_{\tilde{P}_i}(p_i) = \mu_{\tilde{P}_i}\left(\int_S \rho(s)\mu_{\tilde{S}_i}(s)ds\right)$ is the membership degree of p_i to $\tilde{P}_i$.

To derive a fuzzy-number-valued fuzzy measure $\tilde{\eta}_{\tilde{P}^l}$ we suggest to use the following formulas [3]:

$$\eta(\mathcal{H}) = \bigcup_{\alpha \in (0,1]} \alpha\left[\eta_1^\alpha(\mathcal{H}), \eta_2^\alpha(\mathcal{H})\right] \tag{3.4}$$

where

$$\eta_1^\alpha(\mathcal{H}) = \inf\left\{\int_{\mathcal{S}} \rho(s)\max_{\mathcal{S}\in\mathcal{H}} \mu_{\mathcal{S}}(s)ds \,\middle|\, \rho(s) \in P^{\rho^\alpha}\right\},$$

$$\eta_2^\alpha(\mathcal{H}) = \inf\left\{\int_{\mathcal{S}} \rho(s)\max_{\mathcal{S}\in\mathcal{H}} \mu_{\mathcal{S}}(s)ds \,\middle|\, \rho(s) \in core(\tilde{P}^\rho)\right\}, \tag{3.5}$$

$$P^{\rho^\alpha} = \left\{\rho(s) \,\middle|\, \min_{\underline{i=1,n}}(\pi_{\tilde{P}_i}(p_i)) \geq \alpha\right\}, \ core(\tilde{P}^\rho) = P^{\rho^{\alpha=1}} \ , \mathcal{H} \subset \mathcal{S}$$

The support of $\tilde{\eta}$ is defined as $\text{supp}\,\tilde{\eta} = cl\left(\bigcup_{\alpha\in(0,1]} \eta^{\alpha}\right)$. So, $\tilde{\eta}_{\tilde{P}^l}$ is constructed by using $\mu_{\tilde{S}}(s)$ which implies that in construction of the non-additive measure $\tilde{\eta}_{\tilde{P}^l}$ we take into account impreciseness of the information on states of nature themselves. Detailed examples on construction of a fuzzy number-valued measure are considered in the upcoming chapters.

In this section we will discuss features of various existing precise additive and non-additive measures and fuzzy-valued fuzzy measures. The discussion will be conducted in terms of a series of criteria suggested in [72]: interpretation, calculus, consistency, imprecision, assessment, computation. The emphasis will be given to situations in which all the relevant information is described in NL.

Interpretation, Calculus and Consistency. Linguistic probabilities-based fuzzy-valued lower and upper probabilities and their convex combinations have clear behavioral interpretation: they represent ambiguity aversion, ambiguity seeking and their various mixes when decision-relevant information is described in NL. Updating these measures is to be conducted as updating the underlying fuzzy probabilities according to fuzzy Bayes' rule and new construction of these measures from the updated fuzzy probabilities.

Formal validity of the considered fuzzy-valued measures is defined from verification of degrees of consistency, completeness and redundancy of the underlying fuzzy probabilities as initial judgments.

Among the traditional measures, Bayesian probability and coherent lower previsions suggested by Walley [72] (these measures are crisp, non-fuzzy) are only measures which satisfy the considered criteria. Bayesian probability has primitive behavioral interpretation, on base of which the well-defined rules of combining and updating are constructed. Coherent lower previsions have clear and more realistic behavioral interpretation. The rules for updating, combining and verification of consistency of lower previsions are based on the natural extension principle [71,72] which is a general method. However, it is very complex both from analytical and computational points of view.

Possibility theory and the Dempster-Shafer theory, as it is mentioned in [71,72], suffer from lack of the methods to verify consistency of initial judgments and conclusions.

Imprecision. Fuzzy-valued lower and upper previsions and their convex combinations are able to transfer additional information in form of possibilistic uncertainty from states of nature and associated probabilities to the end up measuring of events. As a result, these measures are able to represent vague predicates in NL and partial and complete ignorance as degenerated cases of linguistic ambiguity.

Dempster-Shafer theory is a powerful tool for modeling imprecision and allows to model complete ignorance. However, this theory suffers from series of significant disadvantages [69]. Determination of basic probabilities in this theory may lead to contradictory results. From the other side, under lack of information on some elements of universe of discourse, values of belief and plausibility functions for these elements become equal to zero which means that occurrence of them will not take place. However, this is not justified if the number of observations is small.

Possibility theory is able to model complete ignorance and requires much less information for modeling than probability theory. Possibility measure, as opposed to probability measure, is compositional, which makes it computationally more convenient. However, possibility measure has a serious disadvantage as compared with the probability measure. This theory allows only for qualitative comparative analysis of events – it allows determining whether one event is more or less possible than another, but does not allow determining actual possibilities of events.

Dempster-Shafer theory, lower prevision theory and possibility theory can be considered as special cases of multiple priors representations [69]. In this sense, belief and plausibility functions can be considered as an upper and lower bounds of probability respectively. Possibility theory also can be used for representation of bounds of multiple priors and is used in worst cases of statistical information.

Possibility theory, Dempster-Shafer theory and coherent lower previsions as opposed to Bayesian probabilities are able to model ignorance, impreciseness and NL-based vague evaluations. However, as these theories are based on precise modeling of uncertainty, use of them lead to significant roughening of NL-based information. For example, linguistic description of information on states of nature and their probabilities creates a too high vagueness for these precise measures to be believable or reliable in real-life problems.

Assessment. Fuzzy-valued lower probability is obtained from the linguistic probability assessments which are practical and human-like estimations for real-world problems. Coherent lower prevision can also be obtained from the same sources, but, as a precise quantity, it will be not reliable as very much deviated from vague and imprecise information on states of nature and probabilities. From the other side, fuzzy-valued lower or upper probabilities are computed from fuzzy probabilities.

The other main advantage of fuzzy-valued lower probabilities and fuzzy probabilities constructed for NL-based information is that they, as opposed to all the other measures, don't require independence or non-interaction assumptions on the measured events, which are not accurate when we deal with overlapping and similar objective conditions.

Computation. The construction of unknown fuzzy probability, the use of fuzzy Bayes' formula and construction of a fuzzy-valued lower prevision are quite complicated variational or nonlinear programming problems. However, the complexity here is the price we should pay if we want to adequately formalize and compute from linguistic descriptions. However, as opposed to the natural extension-based complex computations of coherent lower previsions which involves linear programming, the computation of fuzzy-valued previsions is more intuitive, although arising the well known problems of nonlinear optimization.

Computations of coherent lower previsions (non-fuzzy) can be reduced to simpler computations of possibility measures and belief functions as their special cases, but it will lead to the loss of information.

Adequacy of the use of a fuzzy-valued lower (upper) probability consists in its ability to represent linguistic information as the only adequate relevant information on dependence between states of nature in real-life problems. The existing non-additive measures, being numerical-valued, cannot adequately represent such information. To some extent it can be done by lower previsions, but in this case one deals with averaging of linguistic information to precise values which leads to loss of information.

References

1. Aliev, R.A., Aliev, R.R.: Soft Computing and its Application. World Scientific, New Jersey (2001)
2. Aliev, R.A., Alizadeh, A.V., Guirimov, B.G., Huseynov, O.H.: Precisiated information-based approach to decision making with imperfect information. In: Proc. of the Ninth International Conference on Application of Fuzzy Systems and Soft Computing (ICAFS 2010), Prague, Czech Republic, pp. 91–103 (2010)
3. Aliev, R.A., Pedrycz, W., Fazlollahi, B., Huseynov, O.H., Alizadeh, A.V., Guirimov, B.G.: Fuzzy logic-based generalized decision theory with imperfect information. Information Sciences 189, 18–42 (2012)
4. Allais, M., Hagen, O. (eds.): Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of the Decisions Under Uncertainty with Allais' Rejoinder. D. Reidel Publishing Co., Dordrecht (1979)
5. Alo, R., de Korvin, A., Modave, F.: Using Fuzzy functions to select an optimal action in decision theory. In: Proc. of the North American Fuzzy Information Processing Society (NAFIPS), pp. 348–353 (2002)
6. Aumann, R.: Utility Theory without the Completeness Axiom. Econometrica 30, 445–462 (1962)
7. Borisov, A.N., Alekseyev, A.V., Merkuryeva, G.V., Slyadz, N.N., Gluschkov, V.I.: Fuzzy information processing in decision making systems. Radio i Svyaz, Moscow (1989) (in Russian)
8. Buckley, J.J.: Fuzzy Probability and Statistics. Studies in Fuzziness and Soft Computing, vol. 270. Springer, Heidelberg (2003)
9. Camerer, C., Weber, M.: Recent Developments in Modeling Preferences. Journal of Risk and Uncertainty 5, 325–370 (1992)
10. Casadesus-Masanell, R., Klibanoff, P., Ozdenoren, E.: Maxmin expected utility through statewise combinations. Economics Letters 66, 49–54 (2000)
11. Chateauneuf, A., Wakker, P.: An Axiomatization of Cumulative Prospect Theory for Decision Under Risk. Journal of Risk and Uncertainty 18(2), 137–145 (1999)
12. Chen, Z., Epstein, L.G.: Ambiguity, Risk and Asset Returns in Continuous Time. Econometrica 70, 1403–1443 (2002)

13. de Baets, B., de Meyer, H.: Transitivity frameworks for reciprocal relations: cycle transitivity versus FG-transitivity. Fuzzy Sets Syst. 152, 249–270 (2005)
14. de Baets, B., de Meyer, H., De Schuymer, B., Jenei, S.: Cyclic evaluation of transitivity of reciprocal relations. Soc. Choice Welfare 26, 217–238 (2006)
15. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics 38, 325–339 (1967)
16. Diamond, P., Kloeden, P.: Metric spaces of fuzzy sets. Theory and applications. World Scientific, Singapoure (1994)
17. Dubois, D.: The role of fuzzy sets in decision sciences: Old techniques and new directions. Fuzzy Sets and Systems 184, 3–28 (2011)
18. Dubois, D., Prade, H.: When upper probabilities are possibility measures. Fuzzy Sets and Systems 49, 65–74 (1992)
19. Dubra, J., Maccheroni, F., Ok, E.: Expected utility theory without the completeness axiom. J. Econom. Theory 115, 118–133 (2004)
20. Ellsberg, D.: Risk, Ambiquity and the Savage Axioms. Quarterly Journal of Economics 75, 643–669 (1961)
21. Epstein, L.G.: A Definition of Uncertainty Aversion. Review of Economic Studies 66, 579–608 (1999)
22. Epstein, L.G., Wang, T.: Intertemporal Asset Pricing under Knightian Uncertainty. Econometrica 62, 283–322 (1994)
23. Epstein, L.G., Zhang, J.: Subjective Probabilities on Subjectively Unambiguous Events. Econometrica 69, 265–306 (2001)
24. Epstein, L.G., Schneider, M.: Ambiguity, information quality and asset pricing. Journal of Finance 63(1), 197–228 (2008)
25. Ghirardato, P., Maccheroni, F., Marinacci, M.: Differentiating Ambiguity and Ambiguity Attitude. Journal of Economic Theory 118, 133–173 (2004)
26. Ghirardato, P.: Coping with ignorance: unforeseen contingencies and non-additive uncertainty. Economic Theory 17, 247–276 (2001)
27. Ghirardato, P., Klibanoff, P., Marinacci, M.: Additivity with multiple priors. Journal of Mathematical Economics 30, 405–420 (1998)
28. Ghirardato, P., Marinacci, M.: Ambiguity Made Precise: A Comparative Foundation. Journal of Economic Theory 102, 251–289 (2002)
29. Gil, M.A., Jain, P.: Comparison of Experiments in Statistical Decision Problems with Fuzzy Utilities. IEEE Transactions on Systems, Man and Cyberneticts 22(4), 662–670 (1992)
30. Gilboa, I.: Theory of Decision under Uncertainty. Cambridge University Press, Cambridge (2009)
31. Gilboa, I., Schmeidler, D.: Maximin Expected utility with a non-unique prior. Journal of Mathematical Economics 18, 141–153 (1989)
32. Guo, P., Tanaka, H.: Decision making with interval probabilities. European Journal of Operational Research 203, 444–454 (2010)
33. Halpern, J.Y.: Reasoning about uncertainty, p. 483. The MIT Press, Massachusetts (2003)
34. Halpern, J.Y., Fagin, R.: Two views of belief: Belief as generalized probability and belief as evidence. Artificial Intelligence 54, 275–317 (1992)
35. Herrera, F., Herrera-Viedma, E.: Linguistic decision analysis: steps for solving decision problems under linguistic information. Fuzzy Sets and Systems 115, 67–82 (2000)
36. Herrera, F., Herrera-Viedma, E.: Choice functions and mechanisms for linguistic preference relations. European Journal of Operational Research 120, 144–161 (2000)

37. Ho, J.L.Y., Keller, P.L., Keltyka, P.: Effects of Outcome and Probabilistic Ambiguity on Managerial Choices. Journal of Risk and Uncertainty 24(1), 47–74 (2002)
38. Hu, Q., Yu, D., Guo, M.: Fuzzy preference based rough sets. Information Sciences, 180: Special Issue on Intelligent Distributed Information Systems 15, 2003–2022 (2010)
39. Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., Platt, M.L.: Neural signatures of economic preferences for risk and ambiguity. Neuron 49, 765–775 (2006)
40. Insua, D.R.: Sensitivity Analysis in Multiobjective Decision Making. Springer, New York (1990)
41. Insua, D.R.: On the foundations of decision making under partial information. Theory and Decision 33, 83–100 (1992)
42. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under uncertainty. Econometrica 47, 263–291 (1979)
43. Klibanoff, P., Marinacci, M., Mukerji, S.: A smooth model of decision making under ambiguity. Econometrica 73(6), 1849–1892 (2005)
44. Klibanoff, P.: Maxmin Expected Utility over Savage Acts with a Set of Priors. Journal of Economic Theory 92, 35–65 (2000)
45. Klibanoff, P.: Characterizing uncertainty aversion through preference for mixtures. Social Choice and Welfare 18, 289–301 (2001)
46. Kolmogorov, A.N.: Foundations of the theory of probability. Chelsea Publishing Company, New York (1936)
47. Kuznetsov, V.P.: Interval Statistical Models. Radio i Svyaz Publ., Moscow (1991) (in Russian)
48. Labreuche, C., Grabisch, M.: Generalized Choquet-like aggregation functions for handling bipolar scales. European Journal of Operational Research 172, 931–955 (2006)
49. Lakshmikantham, V., Mohapatra, R.: Theory of fuzzy differential equations and inclusions. Taylor & Francis, London (2003)
50. Lü, E.-L., Zhong, Y.M.: Random variable with fuzzy probability. Applied Mathematics and Mechanics 24(4), 491–498 (2003), doi:0.1007/BF02439629
51. Mathieu-Nicot, B.: Fuzzy Expected Utility. Fuzzy Sets and Systems 20(2), 163–173 (1986)
52. Meredith, J., Shafer, S., Turban, E.: Quantitative business modeling. South-Western, Thomson Learning, USA (2002)
53. Modave, F., Grabisch, M., Dubois, D., Prade, H.: A Choquet Integral Representation in Multicriteria Decision Making. Technical Report of the Fall Symposium of Association for the Advancement of Artificial Intelligence (AAAI), pp. 22–29. AAAI Press, Boston (1997)
54. Nau, R.: The shape of incomplete preferences. Ann. Statist. 34, 2430–2448 (2006)
55. Nguyen, H.T., Walker, E.A.: A first Course in Fuzzy logic. CRC Press, Boca Raton (1996)
56. Ok, E.A.: Utility representation of an incomplete preference relation. J. Econ. Theory 104, 429–449 (2002)
57. Ovchinnikov, S.: On fuzzy preference relations. International Journal of Intelligent Systems 6, 225–234 (1991)
58. Pena, J.P.P., Piggins, A.: Strategy-proof fuzzy aggregation rules. Journal of Mathematical Economics 43, 564–580 (2007)
59. Pospelov, D.A.: Fuzzy Sets in Models of Control and Artificial Intelligence. Nauka, Moscow (1986) (in Russian)
60. Pytyev, Y.P.: Possibility. Elements of theory and practice. Editorial URSS, Moscow (2000) (in Russian)

61. Rigotti, R., Shannon, C.: Uncertainty and risk in financial markets. Econometrica 73, 203–243 (2005)
62. Savage, L.J.: The Foundations of Statistics. Wiley, New York (1954)
63. Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrita 57(3), 571–587 (1989)
64. Shafer, G.A.: Mathematical theory of evidence. Princeton University Press (1976)
65. Smets P Imperfect information: Imprecision – Uncertainty, `http://sites.poli.usp.br/d/pmr5406/Download/papers/Imperfect_Data.pdf`
66. Talasova, J., Pavlacka, O.: Fuzzy Probability Spaces and Their Applications in Decision Making. Austrian Journal of Statistics 35(2, 3), 347–356 (2006)
67. Tsukamoto, Y.: Identification of preference measure by means of fuzzy integrals. In: Ann. Conf. of JORS, pp. 131–135 (1972)
68. Tversky, A., Kahneman, D.: Advances in Prospect theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty 5(4), 297–323 (1992)
69. Utkin, L.V.: Risk Analysis and decision making under incomplete information. Nauka, St. Petersburg (2007) (in Russian)
70. von Neumann, J., Morgenstern, O.: Theory of games and economic behaviour. Princeton University Press (1944)
71. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)
72. Walley, P.: Measures of uncertainty in expert systems. Artificial Intelligence 83(1), 1–58 (1996)
73. Walley, P.: Statistical inferences based on a second-order possibility distribution. International Journal of General Systems 9, 337–383 (1997)
74. Wang, T.C., Chen, Y.H.: Applying fuzzy linguistic preference relations to the improvement of consistency of fuzzy. AHP Information Sciences 78(19), 3755–3765 (2008)
75. Weichselberger, K.: The theory of interval probability as a unifying concept for uncertainty. International Journal of Approximate Reasoning 24, 149–170 (2000)
76. Yager, R.R.: Decision Making with Fuzzy Probability Assessments. IEEE Transactions on Fuzzy Systems 7(4), 462–467 (1999)
77. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28 (1978)
78. Zadeh, L.A.: Toward a theory of fuzzy Information Granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and Systems 90(2), 111–127 (1997)
79. Zadeh, L.A.: Computing with words and perceptions—a paradigm shift. In: Proceedings of the IEEE International Conference on Information Reuse and Integration, Las Vegas, pp. 450–452. IEEE Press, Nevada (2009)
80. Zadeh, L.A., Aliev, R.A., Fazlollahi, B., Alizadeh, A.V., Guirimov, B.G., Huseynov, O.H.: Decision Theory with Imprecise Probabilities, Contract on Application of Fuzzy Logic and Soft Computing to communications, planning and management of uncertainty, Berkeley, Baku, vol. 95 (2009)
81. Zadeh, L.A., Klir, G., Yuan, B.: Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers By Lotfi Asker Zadeh. World Scientific Publishing Co. (1996)
82. Zhang, G.Q.: Fuzzy number-valued fuzzy measure and fuzzy number-valued fuzzy integral on the fuzzy set. Fuzzy Sets and Systems 49, 357–376 (1992)