Chapter 2
Various Approaches to Decision Making

Life is the art of drawing sufficient conclusions from insufficient premises
—Samuel Butler (Telling it like it is, Paul Bowden, p. 88)

Abstract In general, two paradigms deal with the categorization of decision theories: the descriptive and normative theories. Descriptive theories are based on empirical observations and on experimental studies of choice behaviors. But the normative theories specifically assume a rational decision-maker who follows well-defined preferences of behaviors as well as obeys certain axioms of rational processes. The axiomatic approach plays an important role in formulating these theories of decision making. In this process, theories of decision making are often formulated in terms of deterministic axioms. But these axioms do not necessarily account for the stochastic variation that attends empirical data. Moreover, a rigorous description of the decision process is provided only through real/time perception. Then it is possible to avail the real-time decisions by repetitive application of the fundamental cognitive process. In such a situation, the Bayesian framework provides readily applicable statistical procedures where typical inference questions are addressed. This framework offers readily applicable statistical procedures, and it is possible to address many typical inference questions. But, in many cases, the applicability of algebraic axioms comes into question concerning viability, especially, when the application connected to empirical data arises. Again, the probabilistic approach to decision making needs to be investigated properly in order to study the empirical data. In such cases, where typical inference questions are addressed, the Bayesian framework provides readily applicable statistical procedures. Attempt has been made to study the various aspects of the Bayesian approach to analyze the observational data found in the new empirical findings on decision making.

Keywords Normative model · Canonical approach · Axiomatic approach · Bayesian approach · Bayes' rule · Dempster–Shafer theory

S. Roy, *Decision Making and Modelling in Cognitive Science*,
DOI 10.1007/978-81-322-3622-1_2

In the traditional approach to decision making or judgment, a comparison is made between a decision and a judgment (that is, a decision or judgment about what to do), and a standard one or "benchmark". This leads to an evaluation of whether a particular decision or judgment is good or bad relative to the standard one. Decision-making behavior is present in almost every sphere of life and so is studied in many very different fields, from medicine and economics to psychology, with major contributions from mathematics, statistics, computer science, artificial intelligence and many other branches of the scientific–technical disciplines. However, conceptualization of decision making and the methods applicable for studying it with success vary greatly. This has resulted in fragmentation of this field. According to Wang, "Decision making is one of the 37 fundamental cognitive processes modelled in the Layered Reference Model of the Brain (LRMB)" , where each of the disciplines explores specifically the special aspect of decision making connected to a particular problem in that discipline. So, the basic problem for the conceptualization of decision making is to address what kind of methods are needed to be applied and followed successfully, to study it. These methods vary in substantial ways, from problem to problem, the results of which lead to the fragmentation of this field into many categories.

The criteria for the success of this theory cover the whole decision cycle. They consist of, e.g., the framing of a particular decision which has been reached based on beliefs, goals, etc. They should also include the background knowledge of the decision-maker regarding the formulation of the decision options. It establishes more preferences compared to others, together with making commitments which ultimately initiate the making of some new decisions. These decisions may lead, again, to some other newly developing indecision. The next step of the problem concerned necessitates in bringing steps where, in turn, it is needed to incorporate the reasons towards the consecutive cycle about previous decisions as well as the rationales for them. This process, ultimately may lead either to the revision or abandonment of commitments that already exist. In this way, the theory of decision making depends on the successful roles played by other high-level cognitive capabilities like problem solving, planning and collaborative decision making. Three domains are responsible for assessing the canonical approach: artificial intelligence, cognitive and neuropsychology and decision engineering. These standards can be achieved or provided in "normative models". 'Normative' means relating to a model or an ideal standard. In this approach, people are primarily concerned in making their decisions logically and rationally.

The normative process is typically contrasted with informative (referring to the standard descriptive, explanatory or positive content) data. This can be termed as supplemental information, for example, additional guidance, tutorials, history, supplemental recommendations and commentary, as well as the background development together with relationships with other elements. Informative data is not a basic requirement and doesn't compel compliance. They are valuable because of their sets of rules or axioms which are derived from utility theory in economics and probability theory. These can be used to test the predictions about human behaviors. Biases are studied if the behavior deviates from the predictions of normative models. There are other approaches to decision making than the normative approach, e.g., the rational theory of decision making.

Thus decision making is linked with a vast set of phenomena and processes (Rangel et al. 2008; Sanfey and Chang 2008). For example, all human beings when in normal, healthy, mental and physical states, try to make decisions resulting in their natural control of their actions: "I would rather prepare for the exam instead of joining that evening party", i.e., some kind of self-conscious control of decisions and actions comes into play. But voluntary control of actions can be realized only from the angle of pure reflexes. When hit by the doctor in the right spot, the patient's foot moves without having any intention. But regarding reflexes, it can happen due to earlier learned experiences which could be encountered in most common experiences. An example of such reflexive decision making, i.e., of taking action (controlling the speed or when to stop, etc.), is when you start to apply your car's break even before you become aware of your conscious decision, and it happens as fast as a reflex. The classic experiment in the research field of human volition was performed by Libet (1985a, b). However, after a modified and modern version of this experiment, the authors (Soon et al. 2008) sought to predict the final response outcome, based on the activity patterns observed via localized brain waves. They observed the activity patterns in the frontopolar cortex V, seconds ahead of the "urge to move", which enabled them to predict the final choice (Soon et al. 2008). While the latter result was above chance, however, one cannot arrive at a reliable prediction. Research advances of modern neuroscience lead us to classify the approaches to decision making as follow in the next section.

2.1 Decision Making (On the Basis of Human Decisions)

To start with, on the basis of human decisions, decision making can be categorized so:

I. Decision making, developed for money making, i.e.,

   - Mathematical formalization, made in game theory.
   - Phenomenological models of rationality.

II. Decision Making (with application from and of neuroscience):
These are models based on learning theories from artificial intelligence (AI) research, i.e.:

   - Based on the findings of neural correlates of realistic models of decision making in the brain.
   - Models arising from understanding the cellular basis of cognition.

2.1.1 Canonical Approach and Normative Models

In general, two paradigms can be stated as dealing with the categorization of decision theories into descriptive and normative theories. Descriptive theories are based on empirical observations and on experimental studies of choice behaviors. But the normative theories specifically assume a rational decision maker who follows well-defined preferences of behaviors, as well as obeys certain axioms of rational processes. Many researchers (e.g., Zachery et al. 1982) have proposed three components of decision making as the main contributors, identifying them as:

1. *decision situation*
2. *decision maker*
3. *decision process*

Even if different decision makers possess cognitive capacities varying greatly in degree, the human brain shares similar and recursive characteristics and mechanisms, with all the core cognitive processes being interconnected within the mechanism of the brain. The canonical approach to decision making can be assessed in three domains:

1. cognitive and neuropsychology
2. decision making in artificial intelligence
3. decision engineering

All these standards can be achieved or provided in "normative models". Normative, by definition, relates to an ideal standard of model, or is based on what is considered to be the normal or correct way of doing something. Typically, the informative data approach is in contrast to the normative approach. The informative data approach can be applicable and related to standardized positive contents. Here, the data could have the nature of a descriptive or explanatory nature, which can be, for example, supplemental information or recommendations, additional guidance, etc. This also includes data from the tutorials, or from background history, development, and relationships with other elements. It is interesting to note that informative data does not require and compel compliance. They are valuable because of their set of rules or axioms which are derived from utility theory in economics and probability theory. These can be used to test the predictions about human behaviors. Biases are studied if the behavior deviates from the predictions of normative models.

There are other approaches to decision making than normative models, which vary in the extent they correspond to observed choices. But the normative approach can be stated as a rational theory of decision making. However, there exist differences between these three functions—descriptive, normative and prescriptive when choosing models, where the evaluation criteria are interesting to study. For example, empirical validity, i.e., the extent to which they correspond with the observed data, determines the evaluation of descriptive models, whereas normative models are evaluated only when the theoretical adequacy, i.e., the degree to which they provide acceptable idealization or rational choices, is provided. Lastly, the ability in helping people to make better decisions, i.e., their pragmatic values, are deciding factors for the evaluation of the prescriptive models. However, difficulties are faced in defining as well as in evaluating all these three criteria, which could be encountered by the students of philosophy of science. It is also a fact, nevertheless, that the criteria are obviously different, thus indicating that the argument posed in the case of normative model might not be an argument against a descriptive model and vice versa.

Next, considering the property of stochastic dominance, this condition is considered as the cornerstone of rational choice, and any theory contrary to this can be regarded unsatisfactory from a normative standpoint. A descriptive theory, on the other hand, is expected to be responsible in accounting for the observed violations of stochastic dominance problems 2 and 8 as stated and explained in Tversky and Kahneman (1986). Only prescriptive analysis could be developed in order to eliminate and reduce such violations. The failure of dominance therefore could be the only answer to this kind of violation culminating in the proper counter-example to a normative model. It is due to this fact that this could be the observation to be explained by the descriptive model, and it appear to be a challenge for a prospective model.

2.1.2 The Axiomatic Approach

The axiomatic approach, belonging to the first category of approaches, plays an important role in formulating theories of decision making (Roberst 1979). In this process, theories of decision making are often formulated in terms of deterministic axioms that do not necessarily account for stochastic variations that attends empirical data. But, a rigorous description of the decision process is only possible when provided through real-time perception. Then it is possible to avail the real-time decisions by repetitive application of the fundamental cognitive process. In many cases, however, the applicability of algebraic axioms comes into question regarding viability, especially when applications connected to empirical data arise. The axiomatic approach is intended to characterize the fundamental principles of human decision making, which provides the necessary inputs to find out the essentially important and sufficient conditions needed for the existence of numerical representations. Usually, the deterministic axioms are the real base upon which the decision-making theories are formulated, and these axioms do not take into account the stochastic variation associated with empirical data. But the probabilistic approach to decision making, again, needs to be investigated properly to study the empirical data. In such cases, where typical inference questions are addressed, the Bayesian framework provides readily applicable statistical procedures. This situation arises when algebraic axioms are applied to empirical data, and the applicability of these axioms is to be tested. It is well known that employing a prior distribution is the key idea of the Bayesian framework where the parametric order constrains the implications of a given axiom. Here, the Bayesian framework is discussed, considering this as a suitable probabilistic approach to handle the empirical data employing a prior distribution that represents the parametric order constraints, implied by a given axiom. Usually, to estimate the posterior distribution, modern methods of Bayesian computation such as Markov chain Monte Carlo (McMC) is used. The advantage of following this method is that it provides the needed information enabling an axiom to be evaluated. Specifically, the descriptive adequacy of a given model is assessable only when we adopt the Bayesian *p*-value as the criterion. In turn, this can assess the descriptive adequacy of a given axiomatic model so that it becomes possible to select the deviance information criterion (DIC) among a set of candidate models.

Bayesian framework, thus, can be tested in this way and at the same time helps us to test already established, well-known axioms of decision making which also includes the axioms of monotonicity of joint receipt and stochastic transitivity. It is a well-known fact that monotonicity is a property of certain types of digital-to-analog-converter (DAC) circuits where the analog output always increases or remains constant whenever the digital input increases. This characteristic becomes an important factor in many communications applications in which DACs are used. It is interesting to note that such applications can function in the presence of nonlinearity, but not in the presence of non-monotonicity.

2.1.3 Bayesian Probabilistic Approach

Since this book studies modelling in the cognitive domain, specifically we will focus our attention on the Bayesian approach in the context of brain functions. Our thoughts, though abstract, are determined by the actions of specific neuronal circuits in our brains. The new field under consideration is known as "decision neuroscience" which is meant to discover these circuits, thereby mapping thinking on a cellular level. This way, cognitive neuroscience, a joint investigation with the aim of investigating the nature of human intelligence, is connected to decision making. Particular emphasis is given to improving these functions through cognitive neuroscience, based on the knowledge of the neural basis on which human beings make decisions. This includes two aspects, i.e., how we learn the value of good, as well as the actions to be taken.

Among all global brain functions, arguably, the ability to predict the outcome of future events is universally the most significant. The results of a given action depend not only on the ability to anticipate the amount of the outcome, but also on sensory stimuli. This information is fed by not only the outside world but also from previously learned experience or inherited instincts. As a next step, applying prior knowledge to decision making, as well as that to judgment, is then absolutely needed to develop a theory of inference. Typically, it has been found that the Bayesian models of inference are useful in solving such problems involve probabilistic frameworks. According to Atram (1985: 263–264):

> By nature, human minds everywhere are endowed with common sense. They possess universal cognitive dispositions that determine a core of spontaneously formulated representations about the world. The world is basically represented in the same way in every culture. Core concepts and beliefs about the world are easily acquired, yet they are restricted to certain cognitive domains and are rather fixed. Not all advocates of domain specific biological constraints on cognition are so concerned to stress that only one outcome is permitted.

The basic question, addressed by the Bayesian framework is to find out the way how one updates beliefs and achieves the result, makes inferences, and applies the observed data. However, appropriate application of Bayesian analysis in the human cognitive domain has remained largely unexplored. Only a few attempts have been made (Griffith et al. 1998; Yu and Smith 2007; Berniker and Körding 2008; etc.) to model or explain human-brain dynamics rather than the cognition mechanism (although the two are definitely related). In this regard, Bayes paradigm may offer a possible solution by offering a tractable version of the so-called hierarchical modelling of the Bayes paradigm, which may be termed as the "Bayesian Brain Hypothesis" (Friston 2003, 2005; Knill and Pouget 2004; Doya et al. 2007). According to Griffith et al. (2000: 29):

> Cognitive processes, including those which can be explained in evolutionary terms, are not "inherited" or produced in accordance with an inherited program. Instead, they are constructed in each generation through the interaction of a range of developmental resources. The attractors which emerge during development and explain robust and/or widespread outcomes are themselves constructed during the process. At no stage is there an explanatory stopping point where some resources control or program the rest of the developmental cascade. "Human nature" is a description of how things generally turn out, not an explanation of why they turn out that way. Finally, we suggest that what is distinctive about human development is its degree of reliance on external scaffolding.

Thus, this characteristic turns out to be a major virtue of the hierarchical, predictive coding account. Following Clark (2012), we can arrive at the conclusion that this method effectively implements a computationally tractable version of the so-called Bayesian Brain Hypothesis. "*But can Bayesian brains really be the same as a predictive brain? Or is the claim merely informal or imprecise, shorthand for something which is formally and factually false?*"

However, even besides being an intrinsically interesting and suggestive one, demonstration of behavior through this model needs to be established. Especially, the need resides in the observation of facts by which one could arrive at strong conclusions regarding the shape of the mechanisms that generates those behaviours. In the hierarchical predictive coding framework, the brain is assumed to represent the statistical structure of the world at different levels of abstraction by maintaining different causal models which are organized, accordingly, on different levels of a hierarchy. In such a case, each level of abstraction obtains input from its corresponding subordinate level. Now, predictions, for the level below, are made by employing a feed-backward chain. This means that the errors obtained between the models' predicted and the observed (for the lowest level) or inferred (for higher levels) input at that level, are used. These are carried out:

(a) In a feed-forward chain to estimate the causes at the level above and
(b) To reconfigure the causal models for future predictions.

Ultimately, this will produce the needed output for stabilizing the system when the overall prediction error is minimized. In his famous paper, Griffith (2000) beautifully expressed his views regarding the "developmental systems" perspective to replace the idea of a genetic program. This idea appears to be clearly and boldly convergent with the recent works in psychology involving situated/embodied cognition and the role played by, so to say, external 'scaffolding' in cognitive development. He stated as follows:

> Developmental Systems Theory (DST) is a general theoretical perspective on development, heredity and evolution. It is intended to facilitate the study of interactions between the many factors that influence development without reviving 'dichotomous' debates over nature or nurture, gene or environment, biology or culture. (Paul E. Griffith, Discussion: three ways to misunderstand developmental system theory, Biology and Philosophy (2005); 20; 417)

At this stage of development, we can now add, for example, the evolutionary accounts of art, religion and science (Mithen 1996), family dynamics, conscience (de Waal 1996), categorization (Atram 1990), cooperation (Sober and Wilson 1998) and cheating detection (Cosmides and Tooby 1992). On the other hand, many other researchers (Byrne and Whiten 1988; Whiten and Byrne 1997) proposed new theories of human cognitive evolution. Not only that, but a strident movement has been started, even in psychology and anthropology during the same period. The arguments made against earlier theorists in that discipline concerned neglecting the basic fact about the mind. Their proposal considered theories of human cognitive evolution as, basically, a product of evolution (Cosmides and Tooby 1992). This gives rise to the necessity of taking into consideration and paying attention to this context, and much research is yet to be accomplished. Thus, the very idea, which once we would have considered as a failure of rationality, now appears to be likely interpreted as good evolutionary design. Gigerenzer (2000) opined that, under the condition of our evolving statistical reasoning abilities, statistical information would accumulate in the form of natural frequencies. So, nothing appears wrong with our rationality if we like to use it in the domain for which it was designed, maybe, exclusively for this purpose. Currently, much attention has been focused on the suggestion that the evolved regularities in human cognition are very rich and domain specific. This fact has profound effects on the conceptualization of innate theories of the domain, such as grammar, physics, biology and psychology, to name a few of some of the most widely discussed spheres (Carey 1985a, b; Keil 1979; Welman 1990; Wellman and Gelman 1992; Gunner and Maratsos 1992; Pinker 1994). In fact, we do not construct these theories on the basis of the evidence available to us; instead, we inherit them from our ancestors. Such claims are commonly expressed in terms of biological "constraints" on the human mind and believed to be constructed in order to reason in particular ways.

Very recently, several theories of human memory have revolutionized our perspective on human cognition. They propose memory as a rational solution to computational problems posed, especially by the environment. The first rational model of memory offered the idea along with convincing demonstrations that human memory appears to be well adapted in a remarkable way to environmental statistics. But proper assumptions for these ideas are needed, at a bare minimum, as long as the form of environmental information is represented in memory. Several probabilistic methods have been developed recently for representing the latent semantic structure of language in order to delineate the connections to research in computer science, statistics and computational linguistics.

Now, before going into the details of the Bayesian model for cognitive science, let us elaborate the Bayesian framework itself. The Bayesian rule of decision theory is named after Thomas Bayes, who lived in the eighteenth century. After the death of Bayes, his friend Richard Price sent Bayes' papers to 'Philosophical Transactions of the Royal Society' for publication. In 1974, the paper "***Essay towards solving a problem in the doctrine of chances***" was published and became the mainstream of statistics. It is well known that among different branches of statistics, the two major paradigms in mathematical statistics are the frequentist and Bayesian statistics. Under the uncertainty of information contained in empirical data, methods based on the statistical inference and decision making are efficiently handled by Bayesian methods. Bayesian inference has tremendous impact on different kinds of popular components among formal models of human cognition (Chatter et al. 2006). This approach provides a general methodology which can be derived from an axiomatic system.

But Kwisthout (2011) raised some serious questions about a few difficulties while discussing the computational tractability of this method. They pointed out that it is necessary to focus on the cause estimation steps present in the feed-forward chain. They argued: instead one should question whether the predictive coding framework specifies satisfactorily the causes, given the necessary steps to be taken so that these steps can be both Bayesian and computationally tractable. In the Bayesian interpretation of predictive coding (Friston 2002), the estimation of the causes means finding the most probable causes ϑ_m, when the input μ are given for that level and the current model parameters θ, i.e.,

$$\Pr(\vartheta | u; \theta) = \arg\ \max\ \vartheta_m$$

Now, if ϑ_m input has a maximum a posteriori probability (MAP), the idea that Bayesian inference is implemented by the predictive coding appears to be primarily dependent on this step. Also, in addition to this, the idea that the predictive hierarchical coding, as a next step, makes Bayesian inference tractable is dependent on the presumed existence of a computational method that could be tractable. But, computing MAP—exactly or approximately—is already found to be computationally intractable because of the presence of causal structures (Shimony 1994; Abdelbar and Hedetniemi 1998). It is a well-known fact that the existence of tractable method is clearly dependent on the structural properties of the causal models. So, as opined by Blokpoel et al. (2012: 1):

> At present, the hierarchical predictive coding framework does not make stringent commitments to the nature of the causal models, the brain can represent. Hence, contrary to suggestions by Clark (2012), the framework does not have the virtue that it effectively implements tractable Bayesian inference….. three mutually exclusive options remain open: either predictive coding does not implement Bayesian inference, or it is not tractable, or the theory of hierarchical predictive coding is enriched by specific assumptions about the structure of the brain's causal models… assume that one is committed to the Bayesian Brain Hypothesis, then the only third survives instead of first two options.

2.1.4 Bayesian Statistics

The basic idea of Bayesian statistics is that the probabilities in this framework are interpreted as rational, conditional measures of uncertainty and loosely resemble the word probability in ordinary language. Here, in various propositions, probability represents the degree of belief whereas, the Bayes rule updates those beliefs. This is based on some kind of new information and the strength of this belief which can be represented by a real number lying between 0 and 1. Its other basis can be traced to the idea of expressing uncertainty about the unknown state of nature in terms of probability. Basic characteristics of this kind of statistics are that it is always possible to update the obtained probability distribution in the light of the new data, which solves many technical problems associated with standard statistics.

2.1.5 Bayes' Rule

Bayes' rule in probability theory and its various modes of applications can be stated as follows: Let us consider two random variables A and B. Now, applying the principle of probability theory, a joint probability of these two variables, P (a, b) can be written by taking particular values of a and b for A and B respectively, as the product of the conditional probability of P (a) and P (b), i.e., when A will take on value a given B, and B takes on value b for given A, we have:

$$P(a,b) = P(a|b)P(b).$$

Using a symmetrical argument, i.e., without having preference of choosing A over B in factorization of joint probability, P (a, b) can be written as:

$$P(a,b) = P(b|a)P(a).$$

Rearranging the above two equations we have:

$$P(b|a) = P(a|b)P(b)/P(a).$$

This is known as Bayes' rule by which the conditional probability of b, i.e., $P\ (b|a)$ given a, can be obtained from the conditional probability of a, given b. It is important to note that in terms of random variables, this is nothing but the elementary result of probability theory. Usually, Bayes' rule can be thought of as a tool, updating our belief about a hypothesis A in the light of new evidence B. More technically expressed, our posterior belief $P\ (A|B)$ is calculated multiplying our prior belief $P\ (A)$ by the likelihood function $P\ (B|A)$ that B will occur if A is true.

For example, the observed data D can be analyzed within a statistical framework considering a formal probability model $\{p\ (D|\omega), \omega \in \Omega\}$ for some unknown value of ω over the parameter space Ω. According to the Bayesian approach, prior to the data being observed, it is necessary to assess a prior probability distribution $p\ (\omega|K)$ over the parameter space Ω, as it describes the available knowledge K about ω. Then the next step is to obtain posterior distribution from Bayes' theorem with probability density $p\ (\omega|D, A, K)$:

$$p(\omega|D, A, K) = p(D|\omega)p(\omega|K) / \int_{\Omega} p(D|\omega)p(\omega|K)d\omega$$

with A the assumptions made about the probability model. Here, all the available information about ω, after the data D is observed, is contained in the posterior distribution.

2.2 Decision Making and Statistical Inference

Statistical inference is a decision-making process under uncertainty that deduces the properties of an underlying distribution by analysis of data and draws conclusions, for example, about populations or scientific truths from the results. There are many modes of Bayesian and non-Bayesian approaches for performing statistical inferences which have been discussed by various authors in different contexts. This includes statistical modelling, data-oriented strategies and explicit use of designs and randomization in analyzes. Furthermore, there are broad theories (frequentists, Bayesian, likelihood, design-based, etc.) with numerous complexities (missing data, observed and unobserved confounding biases) for performing inference. At present, we focus our discussions on the Bayesian approach to decision making, especially in the cognitive domain.

Generally, when there are two or more possible courses of action, one faces a decision problem. Consider the class of possible actions be designated by A and for each, $a \in A$, let us denote the set of relevant events by Γ_a. By denoting the consequence of a chosen action as $c\ (a,\ \gamma) \in C$ where the event is denoted by $\gamma \in \Gamma_a$, then the decision problem can be described by the class of pairs $\{\Gamma_a, C_a\}$. Here, the possible actions are considered to be mutually exclusive. Then again, minimum collection of logical rules is required for "rational decision making". So as to avail a minimum collection of these types of logical rules, it is possible to propose different set of principles. For this reason, the rational decision making can be defined as a logical, multistep model for choice. It is expected to make a choice only between alternatives that will gain maximum benefits for themselves, but at the minimum cost. This follows an orderly path starting from problem identification through the expected solution.

In fact, it can be expressed in a straightforward way, i.e., this model not only assumes that the decision maker has full or perfect knowledge about the information regarding the alternatives supplied as a prior information, but also expects to have the time, cognitive ability, response and resources needed for the evaluation of each choice against the others. Basically, Bayesian decision theory is based on two basic tenets:

- Decision maker's choice or preferences are affected by new information through its effect on his rather beliefs than his taste.
- Decision maker's posterior beliefs represented by posterior probabilities are obtained by updating the prior probabilities (representing his prior beliefs) based on Bayes' theorem.

Probability theory has become the focus of attention only recently in cognitive science for the following reasons:

- In cognitive science, the focus was mainly on the computational aspect but not on the nature of inferences, probabilistic or not.
- For uncertain situations, such as in psychology and artificial intelligence, formal approaches are mainly based on non-probabilistic methods, for example, non-monotonic logic and heuristic techniques.
- The applications of probabilistic methods, in some cases, are considered to be too restrictive for the cognitive domain.

However, after the remarkable technical progress in mathematics and computer science using probabilistic models (Yuille and Kersten 2006), it has been possible to substantially reduce these restrictions. Tversky and Kahneman (1983) and their collaborators, in their classic works, suggested that human cognition is non-optimal, non-rational and non-probabilistic. But the answer depends on the definition of the word "rational". The cognitive scientists observe that we live in a world of uncertainty and that rational behavior depends on the ability to process information effectively despite various types of ambiguities. We investigate decision making in the domain of neuroscience in the next chapter. Presently, we focus our attention on understanding Bayesian probability theory in the domain of human cognition. As opined by Suedfeld (1992: 435), in his famous article "Cognitive managers and their critics":

> The performance evaluation of decision-maker performance almost always results in the finding that leaders do not live up to the criteria of rationality and complex thinking espoused by the researcher. However, these criteria are not necessarily correct or even relevant. Decision-makers must cope with uncertain, ambiguous, changing, inadequate, and/or excessive information; high threat and reward; different time perspectives and pressures; and a multiplicity of values, goals, constraints and opportunities. As cognitive managers, they need to make good meta decisions (i.e., deciding what strategy to adopt and how much time and efforts to expend on particular decisions). No simple prescription, whether it advocates adherence to formal logic, understanding of the laws of probability, or maximal complexity of information search and processing, can adequately guide this effort.

Thus, in connection to the rationality of thinking and decision making, cognitive scientists observe that we live in a world of uncertainty. Rational behavior depends on the ability to process information effectively despite various types of ambiguity. But, then "are people rational?"

This is a complex question of "Action selection", quite a fundamental decision process for us on which the states, both of our body and the surrounding environment, depends. It is a well-known fact that the signals in our sensory and motor systems are corrupted by variability or noise. Our nervous system needs to estimate these states which depend on several factors including also the associated roles of these very factors. To select an optimal action, it is necessary to combine these state-estimates with knowledge of the potential costs or rewards of different action outcomes. Different mechanisms have been employed for studying the nervous system and the decision problems connected to this, specially, for the estimation followed by proper applications. It is interesting to note that the results obtained so far emphasize that human behavior is quite close to that predicted by Bayesian decision theory. This theory not only defines optimal behavior in a world characterized by uncertainty, but also describes coherently various aspects of sensorimotor processes.

Among many objectives meant for the nervous system, the central and primary one is to sense the state of the world around us and to affect this state in such a way so that it becomes more favorable and easy to us to deal with than the previous one. Sensory feedback is used for measuring this state. But, practically speaking, this information is available with questionable precision and subject to noise. Only limited levels of precision are available connected to our sensory modalities, which vary depending upon the situation. It is a quite well-known fact that vision becomes limited under dim light or extra-foveal conditions, hearing turns out to be unreliable for weak sounds, and proprioception drifts without calibration from vision (Hoyer et al. 2003; Ma 2008; Denève 2008; Beck et al. 2011; Zemel et al. 1998; Pouget et al. 2003).

Furthermore, our actions always cannot claim to have a deterministic outcome as these are subject to noise and uncertainty. Same negative comments could be made about our motor commands, which, in each stage, are associated with various kinds of noise, and even the properties of our muscles, themselves being subject to various kinds of fluctuations, vary day to day, not to mention the variations in strength, i.e., our bodily systems having varying degrees of fatigue linked with health, both physical and mental. Therefore, for successful achievements of its objectives, the nervous system must apply a method for the integration of sensory information into a cohesive whole and, at the same time, this must be done before choosing among actions leading to uncertain outcomes. Thus we arrive at the conclusive point when the nervous system faces a substantially crucial question, i.e., how the noisy information could be tackled so that only the positive contributions would be available and the negative avoided?

As we are already aware, a statistical framework is necessary for efficiently managing uncertain information. Bayesian integration is the mathematical framework by which uncertain information from multiple sources can be calculated and combined optimally. This framework can estimate the results in a coherent and maximally accurate way which it derives from a set of observations. Therefore, this method can be applied to integrate sensory information about the world and, then, to use this uncertain information to make choices about how to act accordingly in a fruitful manner. Above all, the Bayesian framework supplies us with a principled formalism with the help of which it is possible to track down the ways how an optimal nervous system senses its surrounding world and then acts upon it. To illustrate these ideas, let us consider the act of descending a staircase: as a first step to achieve this, the nervous system must sense the stairs, and then act accordingly for transporting us down them. Based on our familiarity with walking down stairs, we have preliminary but strong expectations for certain parameters like the distance between steps, their height and their general shape. Quite often these expectations are strong enough. As a result, when we descend stairs without looking at our feet or in the dark, we feel quite comfortable taking stairs without even observing them, though, normally, we first try to assess what is the proper distance we need to cover for a single step and other possible criteria to be taken care of, if necessary. In fact, vision does not provide perfect measurements, on the contrary, and provides us with an approximate estimate that might be very near to the actual measure of the step's height. Bayes' rule, on the other hand, defines how to combine our expectations of the step's height etc., without the visual sense which is expected to make an optimal estimate of the same. Related to another important point, an example could be put like this: the moment we start to take a step, we simultaneously receive sensory information about our on-going motion and also of height of the stairs. As a next step, we are able to combine this sensory information with the action we have just chosen that makes an optimal estimation of where we are and where are we headed. Finally, the Bayesian framework is utilized in choosing how we should step, given all our expectations and sensory estimates of the steps.

2.2.1 Bayesian Probability and Cognitive Domain

The Bayesian approach to inference has drawn much attention to the subject by the community working in the broad spectrum of cognitive science. During the last decade, many authors addressed issues like animal learning (Courville et al. 2006), motor control (Körding and Wolpert 2006a, b), visual perception (Yuille and

Kersten 2006), language processing (Xu and Tenenbaum 2008), semantic memory (Steyvers et al. 2006), language acquisition, etc., using Bayes' rule. Many such research programs have continued in order to study these aspects starting in the last decade or so. However, the great challenge is to learn about the mysteries of the human mind, which goes beyond the data experienced or, in other words, how the mind internalizes the external world out of the noisy data collected through the sensory organs. This is the most challenging aspect of cognition, not only from the point of view of computation, but also from the version of the age-old problem of induction in Western thought. The same problem also lies at the core of the debates that deal with building machines having human-like intelligence (robotics).

The various methods of probability theory not only characterize uncertainty in information but also manipulate the information itself. Not only that, this helps also for optimization. Bayesian rules, quite contrary to other variants of probability theory, help cognitive scientists to define the rules of rationality. This rule updates the belief of the agents from new data in the light of information as well as prior knowledge. As mentioned earlier, the uncertainties in the decision problem, considered to be unknown numerical quantities, can be represented, for example, by Γ (possibly a vector or matrix) possessed by the agents. Classical statistics are directed towards the use of sample information in making inferences about, say, Γ. These classical inferences are, for most of the cases, made without taking into consideration the use, i.e., for which purpose they are required. But in decision theory, on the other hand, an attempt is made to combine the sample information with other relevant aspects of the problem with the goal of arriving at the best decision. In addition to the sample information, two other types of information are typically needed, the first of which is the knowledge of possible consequences of the decision taken. Often this very prior knowledge can be quantified by determining the loss that could be incurred for each possible consequence of the decision. This loss, generally, can be quantified by determining the loss that would be incurred for each possible decision and for the various values of Γ. The Bayesian rule combines information in an optimal way, based on prior belief, with information from observational data. Within the framework of Bayesian decision theory, it helps during the choice of the action to maximize the performance related to particular tasks (Jacobs 1999, 2002). In Bayesian models, probability computations are applied to explain learning and reasoning, instead of hypothesis spaces of possible concepts, word meanings, or causal laws. The structure of the learners' hypothesis space reflects their domain-specific prior knowledge, while the nature of the probability computations depends on the domain-general statistical principles. Bayesian models of cognition thus combine both approaches which have historically been kept separate for their philosophical differences-providing a way to combine structured representations and domain-specific knowledge with domain-general statistical learning. Battaglia et al. (2003) made a comprehensive overview of Bayesian modelling and Bayesian networks. According to their findings, the use of sensory information is found to be satisfactorily efficient in making judgments and for the guidance of successful actions in the world. Added to this, it has been argued that the brain must represent and make use of the information gathered about uncertainty, both for perception and action in its computations. Applications of Bayesian methods, already a time-tested technique, have been successfully accomplished in building computational theories for perception, as well as for sensor motor control. Not only that, but sufficient evidence has been provided in the case of psychophysics which has been establishing examples and proof that 'Bayes' optimal' is the only answer to human perceptual computations.

With these specific characteristics, the 'Bayesian coding hypothesis' states that the brain represents sensory information probabilistically, i.e., in the form of probability distributions. Many of the proposals deal with several kinds of computational schemes. This type of approach, of course, puts special emphasis on the viability and degree of success connected to these schemes, with the aim of making their model successful in dealing with populations of neurons. Neurophysiological data on this hypothesis, however, is almost non-existent. Due to this insufficiency of data available to neuroscientists, this situation poses a major challenge to test these ideas experimentally, i.e., how to determine through which possible process neurons code information about sensory uncertainty will be successfully applied. They mainly focused on three types of information processing. Specifically, these are:

- Inference
- Parameter learning
- Structure learning.

These three types of operations are discussed in the context of Bayesian networks and human cognition. These types of Bayesian networks become more and more popular in the field of artificial intelligence and human cognition since the factorization of a joint distribution is expressed by graph theory where a network contains nodes, edges and probability distributions. The models, developed following a Bayesian framework, do not follow the usual algorithm or process level. On the other hand, this characterizes more the usual, traditional cognitive modelling, but in the spirit of "*Marr's computational theory*".

Marr is best known for his pioneering work on vision. But, before starting that work, in his three seminal publications, he proposed computational theories of the cerebellum (1969), neocortex (1970), and hippocampus (1971). In each of these mind-boggling papers, crucial: he presented new ideas, which continue to influence modern theoretical thinking to the present day. His 'cerebellum theory' was motivated by two unique features of cerebellar anatomy:

1. The cerebellum contains vast numbers of tiny granule cells, each receiving only a few inputs from "mossy fibers";
2. Among Purkinje cells, present in the cerebellar cortex, each receives tens of thousands of inputs from "parallel fibers", but, surprisingly, only one input from a single "climbing fiber" is extremely strong.

According to Marr's proposal, the granule cells encode combinations of mossy-fiber inputs and the climbing fibers carry "teaching" signals. This signal instructs their corresponding Purkinje cell targets to modify the strength of their synaptic connections with parallel fibers. Though neither of these ideas is universally accepted, yet both of them form the essential elements of viable modern theories.

To be very precise, Hubel and Wiesel (1974) found several types of "feature detectors" in the primary visual area of the cortex, and these primarily motivated Marr to introduce the theory of the neocortex. Generalizing those observations, he proposed that cells in the neocortex are a kind of flexible categorizer, i.e., they learn, primarily, the statistical structure of their input patterns and become sensitive to frequently repeated combinations. Again, the theory of the hippocampus (named "archicortex" by Marr) was developed from their discovery by Scoville and Milner (1957). The latter demonstrated that destruction of the hippocampus produce amnesia but only for the new memories or recent events, while this damage leaves the memories of earlier events that occurred years ago intact. It is interesting to note that Marr described his theory as a phenomena of "simple memory" only: The basic idea behind it was that the hippocampus, by strengthening connections between neurons, could rapidly form a simple type of memory traces. Marr (1982) emphasized that, for carrying out the information processing tasks, vision could very well be considered as responsible. Not only that, but any such task, he argued, could be described in three levels:

(i) computational theory
(ii) specific algorithms, and
(iii) physical implementation.

These three levels correspond roughly to:

(i) not only defining the problem but also setting out the way for solving the problem, at least minimally, so that, in principle, it can be solved
(ii) designing a detailed simulation of the process and be available at hand and
(iii) building a working system for carrying out the proposed problem.

The important point to be noted here is that, in this problem, the levels can be considered independently. As a result, it should be possible to mimic the algorithms underlying biological visions in robots: the only difficulty to be overcome is to find out the possible and viable ways for the physical implementation of these criteria. This concept of independent levels of explanation remains as a "mantra" in vision research even today. Marr made a serious attempt to set out a computational theory for vision in its totality, emphasizing that the visual process passes through a series of processes, where each corresponds to a different representation, say, from retinal image to '3D model' representation of objects.

Remarkably, Marr's paper preceded only by two years a paper by Bliss and Lømo (1973) that provided 'the first clear report' about the long-term potentiation in the hippocampus, a type of synaptic plasticity, very similar to what Marr hypothesized. This vital observation was reported in Marr's paper also as a footnote that mentioned a preliminary report of that discovery. Though Marr's theory regarding his understanding about basic concept of hippocampal anatomy has turned out to contain errors and hence has become less important, nobody can deny his work regarding the basic concept of the hippocampus as a temporary memory system, which remains a part of a number of modern theories (Willshaw and Buckingham 1990). At the end of his paper, Marr promised a follow-up paper on the relations between the hippocampus and neocortex, but no such paper ever appeared. However, some phenomena are described in a more satisfactory way at the algorithmic level or at the neuro-computational level. Here, it is important to mention that all the models of human cognition based on different levels of computation are not applicable to Bayesian analysis. In fact, problems related to inductive inference help fruitfully to solve the problems related to Bayesian rules, and in a more natural manner.

A great deal of theoretical and experimental work has been done in computer science, on inductive inference systems, i.e., the systems that try to infer general rules from examples. But, it still remains a far-reaching goal to suggest a successful and applicable theory for such a system. To start with, it is necessary to survey highlights and explain the main ideas that have already been developed in the study of inductive inference. At the same time, special emphasis should be given on the relationships between the general theory and the specific algorithms and implementations. This is needed for surveying the essential characteristics, both the positive and difficulties related to the techniques which have already been developed. However, the Bayesian approach is becoming more and more relevant and successful in many areas of cognitive science, whereas, previously these areas were conceived in terms of traditional statistical inference (Anderson 1990). Many perceptual phenomena can be explained parsimoniously using a Bayesian approach (Knill and Richards 1996). Bayesian inference fits well with all of Marr's levels of description. It is a useful tool in describing a problem at the level of computational theory, especially as adopted in current biology. A Bayesian model of a motion illusion occurs, i.e., when a narrow, low-contrast rhombus, is moved to the right—it appears to move down as well. This can be understood by: (i) considering the set of stimuli that could have produced the edge-motion signals the observer receives; and (ii) including a 'prior' assumption that objects tend to move slowly (adapted from Weiss et al. 2002). Weiss et al. (2002: 598) formulated a model of visual perception, applying standard estimation theory in which they stated:

> The pattern of local image velocities on the retina encodes important environmental information. Although humans are generally able to extract this information, they can easily be deceived into seeing incorrect velocities…. (i) There is noise in the initial measurements and (ii) slower motions are more likely to occur than faster ones. We found that specific instantiation of such a velocity estimator can account for a wide variety of psychophysical phenomena.

According to this approach, these velocities would, eventually, all fall along a line for a high-contrast edge. The 'aperture problem' is considered to be behind this fact which makes it impossible to measure exactly the velocity of an edge. The cause behind this is that no local motion of the signal movement in the direction of the edge is produced. The line appears blurred for the low-contrast stimulus.

Because, moving at other velocities, edges produce the same motion signals, having some noise in the very same system. Barlow (2001) reviewed this problem in detail where he discussed Marr's notion of independence between levels, together with the theories of neural architecture in the brain that might carry out this kind of inference. According to his views, this technique can be developed to deal with the generic quantities without reference to specific stimuli (reviewed by Barlow 2001). Gibson (1979) advocated that Bayesian approaches also demonstrated their effectiveness in the results when applied not only to the evolutionary but also to the ecological perspective. He advanced the theory of direct perception and proposed that perception and action are one and the same thing. Not only that, but an observer and the environment can be considered as an inseparable pair, which, as per his views, in principle, uses real-life stimuli under natural environmental conditions. The core concepts to be taken into consideration in dealing with such problems are, among others: invariance, affordances, and pi-numbers. According to Gibson (1979: 235), a few preliminary observations in ecological psychology can be stated as follows:

> Prolonged distortion leads to recovery. When one puts on distorting spectacles, e.g., so that straight lines get curved, the visual system adapts to this. After some time, a few hours or days, the lines are seen as straight again. Perception is not the sum of simple sensations. How an observer recognizes visual objects such as human faces or animals cannot be derived from research on the perception of simple points and lines. In laboratory conditions, stimuli are impoverished and poor in information and the percept represent marginal phenomena. Perception must be studied in real-life conditions, not in laboratory conditions. In real-life conditions perceptual information is rich in structure and, hence, in information.

As per his views, the basic need of a simple organism with a simple behavioral repertoire is only to divide information about the organism's state into a small number of categories, basically with respect to the world. Following this technique, the simple organism concerned can use its motor system while moving between these categories (this, however, remains the sole way by which it can get the knowledge about the success of its motor movement). But a greater number of states are required for reliable discrimination whenever dealing with the case of a more complex behavioral repertoire. In generating different motor outputs, this requirement becomes essential for the sensory systems so that it can evolve as per the requirement which helps an organism in discriminating between the contexts. At various stages of a task, the sensory parameters become most helpful in discriminating, as well as controlling movements, quite differently. Again, from Gibson and his collaborators (Clutton-Brock et al. 1979):

> This leads to a view which states light enters the visual system as an optic array with highly complex, but structured, and rich in information. Moreover, by moving around in the environment, the flow of information over the senses is considered to be the essential source of information for the organism. The organism scans the environment in the course of perceiving. Hence, the observer and the environment are an inseparable pair; and perception and action cannot be separated from each other (Clutton-Brock et al. 1979).

This leads to a view which states that the cortex is a pool from which evidence can be drawn. Thus, according to the demands of the task, in each consecutive moment, the neurons loaded with the most relevant information may be located in quite different parts of the cortex. Rizzolatti and Craighero (2004) proposed a theory based on their experimental observations in relation to the role of a category of stimuli. This appears of great importance for primates. They opined after their observations:

> Humans in particular, formed by actions done by other individuals … to survive, must understand the actions of others…without action understanding, social organization is impossible. In the case of humans, there is another faculty that depends on the observation of others' actions: imitation learning. Unlike most species, we are able to learn by imitation, and this faculty is at the basis of human culture … on a neurophysiological mechanism–the mirror-neuron mechanism … play a fundamental role in both action understanding and imitation… we describe first the functional properties of mirror neurons in monkeys… characteristics of the mirror-neuron system in humans… those properties specific to the human mirror-neuron system that might explain the human capacity to learn by imitation.

2.3 Dempster–Shafer Theory

The Dempster–Shafer theory (DST) is basically a theory of belief functions, also referred to as evidence theory. It is a general framework which deals with uncertainty in reasoning and has understandable connections to many other frameworks, e.g., possibility and probability, including imprecise probability theories. This framework can be generalized as a mathematical theory of evidence (Dempster 1967; Shafer 1976) for tackling problems connected to uncertain information. Dempster–Shafer (DS) structure (Shafer 1976; Klir and Yuan 1995) on the real line, though quite similar to a discrete distribution, is different for the characteristic which requires the locations at which the probability mass is assigned, rather than precise points. These are sets of real numbers, termed as 'focal elements'. Each focal element has a non-negative mass assigned to it. The basic probability assignment is zero but corresponds to the probability masses with these focal elements. The basic probability assignment can be related to the correspondence of probability masses associated with focal elements. However, in DS structure, the focal elements may overlap one another rather than be concentrated at distinct mass points, as found in a conventional discrete probability distribution. The idea here is to develop a belief function and a plausibility function based on either the sample observations and/or other evidence from prior knowledge. These beliefs and plausibility are expected to serve as lower and upper bounds respectively, with respect to the actual probability of occurrence meant for a particular event. Here, although DS structure exploits traditional probability theory, it gives more general structure to the uncertainty underlying the phenomenon. It is well known that a decision is affected by many such factors, i.e., information of both objective as well as subjective nature. Various models have been developed to handle objective and subjective information, such as interval numbers (Deng et al. 2011; Kang et al. 2011; Xu et al. 2006); fuzzy set theory (Deng et al. 2010; Ngai and Wat 2005) Dempster–Shafer theory of evidence, and so on.

In the decision-making process, at first, the objective information is collected and then the decision makers join their subjective preferences with the objective information to reach a decision. It is not always easy to combine the objective information with the subjective, e.g., in political decision making to select the optimal economic policy in a country. In decision making based on Dempster–Shafer theory (DST), the information is usually assumed to be exact numbers. In real situations, the information is usually imprecise or vague. DST is used based on the interval of numbers. Recently, a new approach to DST has been developed by Merigo, Ramon and many others Merigo and Casanova 2008; Merigo et al. 2009; Merigo and Engmenn 2010) to handle uncertain information using the method of uncertain-induced-aggregation operators. Both DST of belief functions and Bayesian probability theory (BPT) are two distinct frameworks dealing with uncertain domains. Although they have important differences due to their underlying structures in semantics, representations and the rules for combining and marginalizing representations, they have significant similarities, too. Cobb et al. (2006) discussed in detail about the differences and similarities between these two frameworks and finally came to a conclusion, stating that the two frameworks have "R*oughly the same expressive power*". The extra advantage of DST is that it needs weaker conditions than BPT. The belief function model or DST is shown to be a generalization of the Bayesian model (Shafer and Srivastava 1990). In addition, under an uncertain environment, DS theory has the distinct advantage of being a theory of reasoning. It is able to express the "uncertainty" by assigning the probability to the subsets of the set composed of **N** objects, instead of assigning it to each of the individual objects. Furthermore, this theory has the ability of combining pairs of bodies of evidence or belief functions to derive a new evidence or belief function.

Recently, Fox et al. (1993, 2003, 2005, 2010) proposed a unified approach to decision making called "a canonical theory of decision making" where they address the following questions:

- How can we understand the dynamic lifecycle of decision making from the situations and events that make a decision necessary, to influence their prior knowledge, beliefs, and goals which determine how a decision will be framed, preferences arrived at, and commitments to actions made (Fox and Das 2000)?
- What are the general functions that underpin and constrain the processes that implement such a lifecycle for any kind of cognitive agent, whether the agent is natural or artificial?
- How does decision making, conceived in this very general way, fit within cognitive science's strategic objective of having a unified theory of cognition (UTC) that can cut across psychology, computer science, artificial intelligence (AI) and neuroscience (e.g., Newell 1990; Anderson 2007; Shallice and Cooper 2011)?

- How can we apply this understanding to decision engineering, drawing on insights into how decisions are and/or ought to be made to inform the design of autonomous cognitive agents and decision support systems (e.g., Fox et al. 2003, 2010; Lepora et al. 2010)?

So, this appears to be a very ambitious theory. According to Fox et al. (2010), it is necessary to establish a framework where discussions between decision researchers in various communities are possible. This theory, however, advocates an interdisciplinary approach and needs more careful and analytic work to make it comprehensive.

2.3.1 Cognition and Emotion in Human Decision Making

Since in this book, among other things, we wish to draw the attention of scholars to this very intriguing but interesting line of research, i.e., to the new empirical evidence related to gambling and decision making, it is worthwhile to consider the *interaction between cognition and emotion in human decision making*. Much research of varied natures, i.e., investigations underlying biological, psychological or social factors, have been ongoing that are hypothesized to contribute to gambling behavior.

In recent decades, the gradual expansion of the availability of gambling facilities, specifically, in the most of affluent Western countries, are much more common than those available in most parts of the East. Due to this, there is a necessity for, and considerable interest in, conducting research, especially, in a field involving those people who develop problematic levels of gambling. In order to address this complex social and psychological problem, a large body of research has been conducted in order to understand the determinants of gambling behavior. Evidence now exists that biological, psychological and social factors are all interlinked for the development of problematic levels of gambling. However, the theoretical explanation for gambling has lagged behind the advances in empirical works until now. Clark (2010, 2012), discussed the two dominant approaches to gambling behavior: **cognitive** and **psychological**. A number of erroneous beliefs has been identified in cognitive approaches whereas case-control differences between groups of 'pathological gamblers' and 'healthy control' have been identified with the psychological approach. In short, impaired decision making is responsible and a key feature among many with neuropsychiatric disorders.

Franken et al. (2007), in their book "*Impulsivity is Associated with Behavioural Decision-making Deficits*" concluded, as a result of his experimental observations, that impulsivity and behavioral decision-making deficits are always associated with each other. Not only that, but impaired decision making is the origin and is responsible as key features of many neuropsychiatric disorders. In their experiment, they noted also the task performances in a healthy population whose scores indicated high and low impulsivity depending on different kinds of behavioral decision-making tasks. This reflected orbitofrontal functioning measures “included tasks” that assess decision making with and without a learning component and choice flexibility. The results in their experiments on decision-making performance pointed to the fact that subjects, specifically those exhibiting high impulsivity, display an overall deficit, showing weaknesses in learning reward and punishment associations. Not only that, but they are prone to be wrong in making appropriate decisions (e.g., in the reversal-learning and Iowa gambling tasks where it is, absolutely needed). Together with these observations, it has also been pointed out that the impaired adaptation of choice behavior is present, according to changes in stimulus-reward contingencies (reversal-learning task). On the other hand, the simple, non-learning components of decision making based on ‘reward and punishment’ (Rogers decision-making task) appears to be relatively unaffected. Their results indicate that, in response to fluctuations in reward contingency, the impulsivity observed is associated with a decreased ability in altering the choice of behavior. Furthermore, these findings also establish evidence that supports the notion that trait ‘impulsivity’ is associated with decision making, a function of the ‘orbitofrontal cortex’ only. The impulsivity is known to be associated with behavioral decision-making deficits followed by impaired decision making. This is considered to be a key feature of many neuropsychiatric disorders. So, the cultural, as well as biological dependence are to be taken into account, especially, in the case of gambling data where human decision making and judgment are critical.

Another typical gambling disorder (GD) has been noted in the case of those gamblers who face problems in stopping their gambling behavior once the process gets started. The hypothesis behind this behavioral patterns states that, on a neuropsychological level, it is the result of the cognitive inflexibility of pathological gamblers, which reflects in their impaired performance on neuro-cognitive inflexibility, and is task measured and reward-based. Their suggestion is that the cognitive inflexibility in GD is the result of an aberrant reward-based learning, and, ‘*not based*’ on a more general problem with cognitive flexibility. Boog et al. (2014), based on their detailed experiments on this problem, observed flexibility and the pattern of problems and suggested that GD is the result of an aberrant reward-based learning, not based on a more general problem connected to cognitive flexibility. These results indicate the cause behind GD to be due to the dysfunction of the orbitofrontal cortex, the ventrolateral prefrontal cortex, and also the ventral regions of the striatum in gamblers. It is quite well known how people make decisions, keeping in their mind varying levels of probability and risks associated with the steps of actions taken. But, as stated earlier, even in the present state of research, we are not that much aware of the missing information, not yet available. For this reason, little is known of the neural basis of decision making. Due to this prime ambiguity, as a consequence, the probabilities are also uncertain. In decision theory, ambiguity about probabilities is not at all welcome as it affects substantially the choices. Using functional brain imaging, Hsu et al. (2005) pointed out the role of different parts of the brain in decision making and have shown that the level of ambiguity in choice correlates positively with activation in the amygdala and orbitofrontal cortex but negatively with the striate system. They found positive correlation of striate activities with expected rewards. Not only that, but it has been observed that neurological subjects with orbitofrontal lesions were insensitive to the level of ambiguity and risks in behavioral choices. Finally, they suggested based on their observations that the response of a general neural circuit is dependent on the different degrees of uncertainty, contrary to decision theory.

Next, it is important to mention the vital role of noise, its imminent effects on nervous system and consequently the outcome regarding decision making (Körding and Wolpert 2006a, b). Let us explain the scenario in the following way: if a steady rate is considered an important criterion for determining the intensity of a constant sensory or motor signal, then any variation in the inter-spike intervals will cause fluctuations in the rate, which will appear as unwanted noise. But, if the timing of each spike carries extra information, then this very variability turns out to be a crucial part of the signal. We argue here that both temporal and rate coding are used, to varying degrees, in different parts of the nervous system, and this enables the nervous system to discriminate between complex objects. As a result, graceful movements are generated.

Nearly 60 years ago, it was Claude Shannon (see Gold and Shaden 2001), who developed a theory regarding the information that can be transmitted in a noisy communication channel. This theory has been found to be very effective for applications in many branches of science, especially computer science. In computers, this theory has become popularized in terms of the 'bits and bytes' of information and has even been applied to the information content of the universe (Noel-Smith 1958). Application of this theory to neural systems appears to be quite logical. In a noisy or uncertain environment, the survival of an organism is dependent on how rapidly it is able to collect the necessary crucial information. However, it is not at all clear whether information theory proposed by Shannon is applicable to biological system (Gatlin 1972). Gatlin addressed this problem beautifully in her book where she treated this problem from the angle of classical information theory but with different interpretations. She applied the concepts of information theory developed by Shannon and others to the living system in the process. But the information theory and the second law of thermodynamics are redefined and ingeniously extended in the context of the concept of entropy. Finally, let us conclude with Körding and Wolpert's remarks in this regard:

> Action selection is a fundamental decision process for us, and depends on the state of both our body and the environment. Because signals in our sensory and motor systems are corrupted by variability or noise, the nervous system needs to estimate these states. To select an optimal action these state estimates need to be combined with knowledge of the potential costs or rewards of different action outcomes. We review recent studies that have investigated the mechanisms used by the nervous system to solve such estimation and decision problems, which show that human behavior is close to that predicted by Bayesian Decision Theory. This theory defines optimal behavior in a world characterized by uncertainty, and provides a coherent way of describing sensory- motor processes (Körding and Wolpert 2006a, b: 319).

References

Abdelbar, A.M. & Hedetniemi S. (1998); Artificial Intelligence; 102(1), 21-38.

Anderson, John R. (1990); "*Cognitive Psychology and its implications (3rd edition)*": W.H. Freeman; NewYork (series of books on Psychology).

Anderson, John (2007); "*How Can the Human Mind Occur in the Physical Universe*"? Oxford University Press.

Atram, Scott (1985); American Anthropologist, New Series; **87**,298-310; ibid; Atram, Scott (1990); "*Cognitive Foundations of Natural History: Towards an Antropology of Science":* Cambridge University Press.

Barlow, Jessica A. (2001); Journ. Child. Lang.; **28**, 201-324.

Barry R Cobb, Prakash P Shenoy & Rafael Rumi (2006); Statistics and Computing; **16**; 293-308.

Battaglia P.W, Jacobs R.A, et al (2003); J. Opt. Soc. Am. A., **20**, 1391–1397.

Berniker, M. & Körding, K. (2008); Nature Neuroscience, **11**(12): 1454-61.

Beck, J.M., Latham, M., Latham, P.E., & Pouget, A. (2011): *J. Neurosci.* **31**:15310-15319 and references therein.

Bliss T., & Lømo T. (1973) *Journ. Physiol* **232**(2):331–56.—PMC 1350458. PMID4727084.

Blokpoel Mark, Johan Kwisthout and Iris van Rooij (2012) Frontiers in Psychology **3**, 1

Boog, M, Höppner, P. et al (2014) Front. In Human. Neuroscience; **8**, 1.

Byrne, R. & Whiten, A. (1988) *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*: Oxford University Press.

Carey, Susan (1985a, b);(a) *Are children different kind of thinkers than adults";* In S. Chapman, J. Segal & R. Glaser (Eds); *Thinking and Learning skills*; **2**, 485-518; Hilside, NJ, Erlbaum. (b)"*Conceptual change in childhood"*; Cambridge; MA, MIT press

Chatter, N, Tenenbaum, J.B. & Yuille, A. (2006); Trends in Cognitive Sciences; **10**(7), 287-291.

Clark, L. (2010); Philos. Trans. R. Soc. London, B. Biol. Sci.; **365**; (1538):319-330.

Clark, L., Studer, B. et al (2012); Frontier in Neuroscience (decision Neuroscience);**6**,46.

Clutton-Brock, S.D. Albon, R.M. Gibson. (1979). Animal behavior, **27**(1), 211–225.

Courville, A.C., Daw, N.D. & Touretzky, D.S. (2006) Trends in Cognitive Sciences**; 10**, 294-300.

Cosmides, L & Tooby, J. (1992); *Cognitive adaptations for social exchange;* In J. Barkow, L. Cosmides, & J. Tooby (Eds.); "*The adapted mind"*, New York: Oxford University Press.

Dempster, A. (1967) *Annals of Mathematics and Statistics*; **38**, 325-33.

Deng, X, Zhang, Y. et al. (2010) *Journ. Inform. & Comput. Science*; **7**, 2207-2215.

Denève, S., *Bayesian Spiking Neurons I: Inference*, Neural Computation, 20, 91-117 (2008);ibid; *Bayesian Spiking Neurons II: Learning*, Neural Computation, 20, 118-145 (2008).

Deng, Y., J., Wu, J.X., et al (2011) ICIC Express Letters, **5**, 1057-1062.

Doya, K. et al (2007) *Bayesian Brain: Probabilistic approach to Neuoral Coding*; The MIT Press; edition (Jan 1).

F.B.M. de Waal (1996); Good Natured; The Origins of Right and wrong in Humans and other Animals: Harvard University press.

Friston, K.J., Glaser, D.E. et al (2002); NeuroImage, **16**:484-512.

Fox E., Hix, D., Nowell, L.T. et al; (1993); Journal of the American Society for Information Science, **44**(8): 480–491.

Fox J. & Das S.K. (2000); Safe and Sound; Artificial intelligence in hazardous applications: American association for Artificial intelligence, The MIT press, Cambridge MA, USA.

Fox, J., Beveridge,M., et al (2003) AI Commun. **16**, 139–152.

Fox, E., Karana, M., et al (2005); Acm *Transactions On Information Systems (*Tois*)*, 23(2): 147–168.

Fox, C.W., Billington, P., Paulo, et al (2010): "*Origin-Destination Analysis on the London Orbital Motorway Automated Number Plate Recognition Network";* Euro. Transport Conference.

Franken, I H A. et al. et al (2007); Alcoholism-Clinical & Experimental Research, **31**, 919-927. doi: http://dx.doi.org/10.1111/j.1530-0277.2007.00424.x

Friston, K.J, Holmes, A.P. et al. (2003); Dynamic causal modeling; Neuroscience, **19** (04,1273-302.

Friston, K.J., (2005); Philos. Trans. R. Soc., London, B. Biol. Sci.; **360**(1456),815-36.

Gatlin Lila L. (1972); *Information Theory and the Living System*; Columbia University Press, NY, & London.

Gibson, J.J. (1979); The ecological approach to visual perception (Houghton Mifflin).

Gigerenzer, G. (2000); *Adaptive thinking, Rationality in the real world*: New York; Oxford University Press.

Gold, Joshua I and Shaden N. Michael (2001) Trends in Cognitive science; 5(1); **10**-16.

Griffith, T.D., Rees, G et al (1998); Nature.com; in Percept Psychology, **45**, 371-327.

Griffith Paul & Stotz K. (2000); *Synthese*; **(122)**, 29-51.

Gunner, M.R. & Maratsos, M. (1992); in *Language and Modularity and constraints cognition*: The Minnesota symposium on Child Psychology; 25, Hillside, NJ, Lawrence, Erlbaum associates.

Hoyer, P.O. & Hyvarinen, A. et al. (2003); Neuroimage; **19** (2), 253-260.

Hubel, D. H. & Wiesel, T. N. (1974); *Brain Research*; Elsevier.

Jacob, R.A. (1999); Vision. Res., **39**, 3621-3629;

Jacobs, R.A. (2002). Trends. In Cogn. Sci.; **6**: 345-50.

Kang, B., Zhang, Y.X. et al (2011) Journ. Inform. & Comp. Sci.; **8**, 842-849.

Keil, F (1979); *Semantics and conceptual development*; Cambridge, MA, Harvard University Press.

Klir, G.J. & Yuan, B. (1995); Proc. Ninth ACM-SIAM Symp. on Discrete Algorithm, 668-677.

Knill, D.C & Pouget, A. (2004) Trends *in Neurosciences;* **27** pp. 712-719.

Knill, D. & Richards, W. (1996) *Perception as Bayesian Inference;* Cambridge University Press.

Körding, K.P. & Wolpert, DM (2006); *Probabilistic mechanisms in sensor motor control*; *Novartis Found Symp.,* **270**:191-8; discussion: 198-202, 232-7.

Körding KP, Wolpert DM (2006) Trends Cogn. Sci**, 10**(7):319-26.

Kwisthout, J. (2011) Int. Journ. Approx. Reasoning, **52**, 1452–1469.

Lepora, N., Evans, M., Fox, C.W., et al (2010); *Naive Bayes texture classification applied to whisker data from a moving robot:* Proc. Int. Joint/Conf. Neural Networks (IJCNN).

Libet, B. (1985a) Behavioral and Brain Sciences **8**; pp. 529-66.

Libet, B. (1985b) Journal of Theoretical Biology **114**; pp. 563-70.

Marr David, (1969) *J. Physiol.*, **202**:437–470;

Marr David (1970 *Proceedings of the Royal Society of London B"*, **176**:161–234;

Marr David 1971) *Phil. Trans. Royal Soc. London"*, **262**:23–81;

Marr David (1982) *Proceedings of the Royal Society of London B"*, **214**:501–524. Markov, N.T, Ercsey - Ravasz, M.M. et al (2014); Cerebral Cortex, **24**:17 -36.

Merigo, J.M. & Casanova, M. (2008); "*Procd.IMPU'08";* (eds.; L. Magdalena, M. Ojedo-Aciego, J.L. Verdegay), 425-432.

Merigo, J.M. & Gil-Lafuente, et al (2009*); "Intelligence System Design & Applications"*; ISDA'09.

Merigo, J.M. & Engmann, K.J. (2010); *Procd. World Congress on Engineering*, **1**, WCE2010, June 30-July 2, London, UK.

Mithen Steven (1996) *The Prehistory of Mind: Cognitive Origins of Arts, religion and Science*; NewYork: Thomas and Hudson.

Ming, Hsu, Bhatt, Meghana, etal (2005); Science; **310** (5754),1680-1683.

Newell, Allan (1990) *United theory of cognition;* Cambridge, MA, Harvard University press.

Ngai, E.W.T. & Wat, F.K.T. (2005); Fuzzy decision support system for risk analysis in e-commerce development, Decision Support Systems; **40**; 235-255.

Noel-Smith, P.H. (1958); Choosing, deciding and doing; Analysis; **18,** 63–69.

Pinker, S. 1994. *The language instinct;* NY: Morrow.

Pouget A., Dayan, P. & Zemel, R.S. (2003); Annual Review of Neuroscience, **26**; 381-410.

Rangel, A., Camerer, C. & Montague, R. (2008); Nature Reviews Neuroscience; **9**; 545-556.

Rizzolatti G. & Laila, Craighero (2004); Ann. Rev. Neuroscience; **27**; 169-172.

Roberts, F. S. (1979); *Measurement theory with applications to decision making, utility and the social sciences*; London, UK: Addison-Wesley.

Sanfey A.G. & Chang L.J. (2008) *Annals of the New York Academy of Science*; **1128**(1)**;** 53-62.

Scoville, William Beecher & Milner, Brenda (1957) Journal of Neurology and Psychiatry; **20** (1):11–21.

Shallice, Tim & Cooper, Rick (2011) *The Organization of Mind*: OUP, Oxford (ISBN(s); 9780199579242).

Shafer, G. (1976) *A Mathematical Theory of Evidence;* Princeton University Press, Princeton.

Shafer, G. & Srivastava, R.P. (1990) A Journal of Practice and Theory, **9** (Suppl.), 110-48.

Shimony, S.E. (1994) Artificial Intelligence; **68**; 399-410.

Sober Elliot & Wilson David Sloan (1998); *Unto Others: The Evolution and Psychology of Unselfish Behaviour*; Cambridge, MA; Harvard University Press.

Soon, C.S., Brass, M. Heinze, H.J. et al (2008); Nature Neurosccience; **11**; 543-545.

Steyvers, M., Griffiths, T.L. & Dennis, S (2006): TRENDS in Cognitive Sciences, **10**(7), 327-334.

Tversky A. & Kahneman, D. (1983) Psychological Review, **90**(4); 293-315.

Tversky A. & Kahneman, D. (1986); The journal of Business; **59**(4); S251-S278

Weiss, Y., Simoncelli, E.P. & Adelson, E.H. (2002); Motion illusions as optimal percepts; Nat. Neurosci; **5;** 598–604.

Welman, H.M. (1990); *The Child's theory of Mind*; Cambridge; M A; MIT Press; A Bradford book.

Wellman Henry, M & Gelman, Susan (1992) Annual review of Psychology, 43, 337-375.

Willshaw DJ, Buckingham JT (August 1990); *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 329 (1253): 205–15.

Zachery, W., Wherry, R. et al (1982); Decision situation, decision process and functions: Towards a theory –based framework for decision-aid design; Procd. 1882 Conf. on Human factors in Comp. System.

Whiten, A. & Byrne, R.W. (1997); *Machiavellian Intelligence II : Extensions and Evaluations (Eds.)*; Cambridge University Press.

Xu, D. L. et al (2006); European Journal of Operational Research; **174;** 1914- 194.

Xu, F., & Tenenbaum, J. B. (2008); *Word learning as Bayesian inference; Psychological Review;* (The Cambridge Handbook of Computational Psychology, Ron Sun (ed.); Cambridge University Press.

Yu Chen & Lind Smith (2007); Cognition; **106**(3); 1558-1568.

Yuille, A.L. & Kersten, D. (2006); Trends in Cognitive Neuroscience; **10**(7); 301-308

Zemel, R.S., Dayan, P. & Pouget, A. (1998); Neural Computation; **10**; 403-430.

Springer

http://www.springer.com/978-81-322-3620-7

Decision Making and Modelling in Cognitive Science
Roy, S.
2016, XV, 165 p. 2 illus. in color., Hardcover
ISBN: 978-81-322-3620-7