The Dissertation Committee for Mahdi Habibi Certifies that this is the approved version of the following dissertation:

Development of a Theory for Objective Assignment of Prior Probabilities within the Context of a Decision

Committee:

Robert B. Gilbert, Supervisor

Larry W. Lake, Co-Supervisor

Randall J. Charbeneau

Amit Bhasin

James S. Dyer

Farrokh Nadim

Development of a Theory for Objective Assignment of Prior Probabilities within the Context of a Decision

by

Mahdi Habibi

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2020

Dedication

To my loved ones

Acknowledgements

This achievement would not have been possible without help and support from my supervisors, my committee members, and my family and friends.

I owe my deepest gratitude to Prof. Robert Gilbert for granting me the opportunity of working with him. If it was not for his patience, sympathy, and consistent support, this work would not be reached the level it has.

I am also very grateful to my co-supervisor Prof. Larry Lake for his challenging and profound questions as well as his valuable inputs to the study.

I am very much thankful to the committee members of my dissertation, Professors Dyer, Charbeneau, Nadim, and Bhasin, for their critical and constructive comments and discussions that extremely enhanced the quality of my work.

I am indebted to my best friend of more than twenty years, Hossein Roodi, for his lasting patience, support, and encouragement. I am confident that without Hossein's, my life would not be as enriched as it has been along the time.

I would also thank my colleagues in Prof. Gilbert's decision research group, Dr. Azadeh Mostofi and Kai Feng who provided a friendly and productive environment for research, and whose insightful comments had valuable contributions to my work. I also acknowledge other members of our geotechnical engineering group at the University of Texas who provided a scientific atmosphere to live and conduct research.

Finally, these acknowledgments would not be complete without thanking my family and friends for their unconditional love and support during this long journey. Without them this journey may have ended differently!

Development of a Theory for Objective Assignment of Prior Probabilities within the Context of a Decision

Mahdi Habibi, Ph.D.

The University of Texas at Austin, 2020

Supervisors: Robert B. Gilbert and Larry W. Lake

Two theories are being presented and advanced. While the theory of *Philosophical processualism* is general and aims at supporting the application of mathematics to guide action with precision, the computational theory of *Decision Entropy* is specific and aims at establishing a set of objective rules for assigning Bayesian priors, that are non-informative regarding the decision at hand. Besides, the concept of probability is being analyzed from a process-centered perspective, and three types of probability are being introduces: propensity, possibility, and credibility. The highlighted distinction is hoped to settle a number of long-standing disputes regarding the interpretation of probabilities. Also, three properties of objectivity, transparency, and defensibility are offered to characterize inductive systems. To evaluate defensibility, a set of seven theoretical principles are being offered including the principle of "unbiased evaluation," which can be interpreted as the informativeness of priors for the case of Bayesian inductions.

Philosophical processualism is an original perspective advanced as the foundation for an epistemic machinery called *pragmatic mathematics*; a system of linguistically manifested mental constructs aimed at guiding purposive actions with precision. Philosophical processualism relies upon a process-centered interpretation of causality, which sees an event as a constellation of changes made by a process in certain states of the world and/or the mind. By taking concept as the fundamental constituent of the purposive agents' evaluation processes, the view presents its ontological account of concepts and elaborate on how concepts could look like, where they could be present, and how they could come to realization. Philosophical processualism opposes Platonism by asserting that every concept is the outcomes of cognitive processes unfolding in time and space and is not an abstract entity in the so-called third realm, to which mind can gain access through unknown metaphysical processes.

Probability concept is being analyzed from a process-centered perspective, and it is being divided into three types of propensity, possibility, and credibility. Propensity values are relative repetitions of processual outcomes, as they have come to realization. Possibility values for the outputs of a more-to-less process are the relative contributions of inputs; i.e. the relative number of inputs associated to every output. Credibilities are imaginary relative weights assigned to (the parameters of) the processes, who are hypothesized to deliver the intended outcomes. Credibilities are only means to the end of assessing propensities and/or possibilities, whose exact values are unknown. Bayesian priors are prime examples of credibilities. While the imaginary nature of credibilities allows subjects to assign credibilies of their own preference, their assignment might not be justifiable to others. Alternatively, it is possible to establish a set of theoretical assignment rules on purely logical grounds, and to justify the designations based on their effects, or lack thereof, on the assessed propensities and/or possibilities. Due to their rule abidance, theoretically constructed credibilities may be described as objective, even though imaginary and only existing in subjects' minds. The class of credibilities called non-informative are the ones, whose assignment are aimed at not informing (certain aspects of) the assessed propensities and/or possibilities.

A set of properties including objectivity, transparency, and defensibility are defined to characterize an inductive assessment procedure. Objectivity is concluded to be the result of transparency and community acceptance. Although every communal rule is contractual by nature, its justification makes the procedures defensible, especially when a community is deciding whether to adopt it as the rule. A set of seven theoretical principles are proposed to evaluate the defensibility of an inductive system, namely (I) Evaluative Orientation, (II) Investigative Prioritization, (III) Explicative Sufficiency, (IV) Evaluative Inclusion, (V) Credibility Conception, (VI) Artifact-Reality Division, and (VII) Unbiased Evaluation.

Decision Entropy Theory (DET) offers an original method for assessing credible outcomes of uncertain processes by incorporating Bayesian probabilities. Since DET is aimed at guiding purposive action with precision and through the use of mathematical measures, it falls under the category of pragmatic mathematics. DET aims at representing uncertainty in an objective and defensible way. The motivation to develop the theory is to account for the possibility of events occurring that are beyond our range of experience.

The theory characterizes uncertainty in the context of making a decision; the case of maximum uncertainty corresponds to the maximum entropy for the possible outcomes of the decision. Therefore, the starting point for assessing probabilities, i.e., the non-informative prior probabilities of possibilities before information is included, depends on the decision at hand. Decision Entropy Theory is developed from the following principles that describe the case of no information or maximum uncertainty in making a decision between various alternatives.

1. If no information is available about the probabilities of decision outcomes, then a selected alternative is equally probable to be or not to be the preferred alternative.

2. If no information is available about the probabilities of decision outcomes, then the possible differences in preference between a selected alternative and the preferred alternative are equally probable.

3. If no information is available about the probabilities of decision outcomes, then the possibilities of learning with new information about the selected alternative compared to the preferred alternative are equally probable.

To illustrate the theory, two examples from engineering are being worked out:

(a) selecting the appropriate design wave height for offshore structures, and

(b) assessing the value of test wells before committing to developing a hydrocarbon resource play.

The examples highlight the following points:

- The prior sample space depends on the decision, meaning that the importance of extreme uncertainty depends on its consequences to the decision and the availability of feasible decision alternatives to compensate for these consequences.
- The prior sample space emphasizes the possibilities that distinguish two alternatives from one another.
- The prior sample space can affect the final decision, even when substantive data are available to inform (update) this sample space.
- It is unreasonable to assume that the probability distribution for the frequencies can be established based entirely on experience because that precludes the possibility of events beyond our experience. Such events can be particularly important where experience tends to be limited.
- The value of information is enhanced when leaving open the possibilities for making excessive gains and losses.
- It is possible to rationally balance between relying entirely on historical data versus not relying on them at all. Within this balance, direct information can be combined with information from analogous cases of the past to inform the decision.

Table of Contents

List of Tables

CHAPTER 2

List of Figures

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

APPENDIX B

Chapter 1 Introduction

1.1 Acquiring Knowledge by Experiencing and Reasoning

There are things that we do know and there are (many) other things that we do not know. Learning is the process through which, by observing specific aspects of the reality, data, and by using our human reasoning faculties, we manage to add more to what we already knew. Two central questions regarding learning are determining how we can use what we already know to inform what we currently do not know, as well as determining the (relative) weights we have to give to evidence and reasoning in order to conclude a new learning. To answer the aforementioned questions, we can adopt any of the following approaches.

In the first approach, we assert that until we make new observations, we take the set of our past observations, as of today, as the (sole) representation of the reality. The approach relies on a philosophy whereby (empirical) evidence is the only source of knowledge and results in (implicitly) ignoring unknowns, as they are yet to be observed and, at the moment, the only way to their existence is to hypothesize them by the aid of human faculties. Thus, the only way to learn new things is to wait until we observe them. This approach is mainly pursued in empirical science including experimental physics.

In the second approach, most of our knowledge, is taken to be the logical consequence of what we already know. This approach heavily relies on imposing/ postulating (logical) structures/ orders on the reality, where (states of) certain aspects of the world are (deemed) to be in a particular type of relationship. It is such a relationship between knowns and unknowns that makes the later a logical consequence of the former, i.e. by considering our current knowledge, we can deduct the existence and the character of what we are yet to observe. The approach relies on a philosophy whereby the real-world is deemed to follow the (logical) orders/ structures devised by human minds and takes observation as a (sometimes secondary) support to demonstrate the validity of postulated orders. This mode of learning is mainly pursued in theoretical science including theoretical physics and philosophy.

Following the second approach to learn unknowns, if one believes in the truth of these (imposed) orders, prediction/ characterization of unknowns look logical, reasonable, and natural. By the same token, an easy way to fool oneself into taking a (unobserved) character as real, is to postulate a certain structure on the reality and deduct that character accordingly.

1.2 BACKGROUND

1.2.1 A Random Act/ Process and the Propensity Values for its Outcomes

An Act is taken by an agent in order to achieve a desirable outcome. While a deterministic act/ process always results in the same outcome, a random act/ process sometimes result in one outcome and other times in different outcomes such that the agent is unable to foresee the (exact) outcome every time he takes the random act. The lack of knowledge and the possibility of not achieving the intended consequence make random acts risky. The significance of the risk depends on the undesirability of unintended outcomes as well as the natural tendency of the process to result in such outcomes, which can be quantified by the measure termed (physical) probability. The (physical) probability/ (natural) propensity/ relative frequency, P_P, for a given outcome, E_i, of a (repeatable) random act/ process, A_R, triggered under specified conditions/ circumstances, C, is "defined" as the proportion of the times that taking the act results in the outcome, n_{Ei}, relative to the total number of (possible) repetitions (for the act), n_T , Eq. (1-1).

$$P_P(E_i|A_R, C) = \frac{\mathrm{n}_{E_i}}{\mathrm{n}_T} \tag{1-1}$$

If the random process may result in any of the m possible outcomes/ events, contained in the set $E = \{E_1, \dots, E_m\}$, the set of propensity values for these outcomes defines the (propensity) distribution (function) of the random act, $D_{AR} = \{P_{P1}, \dots, P_{Pm}\}$. Propensity values for a desirable and an undesirable outcomes of a random act are respectively called the reliability and the risk of the random act.

1.2.2 Estimation/ Approximation of Propensity Values based on Statistical Observations

To make the decision whether to take the risky act, the agent needs to assess the significance of its risk and for that, he needs to know the propensity values for various outcomes of the process. Propensity calculations, Eq. (1-1), require the knowledge of the outcome of the process in all (possible) instances/ repetitions of the act. In many cases, however, available data is limited to the outcomes of a few instances; thus, (true) propensity values have to be estimated/ approximated with the aid of the propensities observed in the (small) set of instances, a sample.

Sampling is the act of observing/ learning about the outcome of a random process, as it is being produced or as it was. Thus, observational outcomes are also random and sampling itself is being considered a random process, a statistical experiment. A random sample is a sequence/ series/ collection, where each member/ element is an outcome of the random process. The set enclosing all possible observations of a statistical experiment, random samples, is called the "sample space" (for the act of sampling). If we denote a possible observational outcome, a sample, as I_j, the sample space of the statistical experiment composed of a sequence of n observations can be denoted as $SS_n = \{I_{1,} \dots , I_m{}^n\}$.

The (sequential) outcomes of sampling, as a random act, also do realize with a (natural) propensity if the act of sampling is being repeated, with the same number of observations and under the same conditions.

The (physical) probability/ (natural) propensity/ relative frequency, P_P, for a given observational sequence, I_j, of a (repeatable) sampling from the outcomes of a random act, S_{AR}, performed under specified observational conditions/ circumstances, C_S, is "defined" as the proportion of the times the act of sampling results in the observation of the sequence of interest, n_{Ij}, relative to the total number times a sample of such length could be observed, n_{Ts}, Eq. (1-2). Since an observational sequence is made up of single observations (of the outcomes of an underlying random act), the propensity of each sequential (sampling) outcome is related to the propensity values for its constituent outcomes. In other words, propensity of a sample data, $P_P\ (I_j)$, is a function of the propensity distribution for the (underlying) random act, D_{AR}.

$$P_P(I_j|S_{AR}, C_S) = \frac{n_{I_j}}{n_{T_S}} = P_P(I_j|D_{AR}) \tag{1-2}$$

The connection established in Eq. (1-2) between the propensity of a given (sequential sample) observation and the propensity distribution of the random act allows one to compare various sampling outcomes in terms of the chance for their realization. This ability is quite helpful when the propensity distribution of the random act is known and a comparison between sampling outcome is needed, e.g. calculating the (physical) probability for the realization of a winning combination in a (well-defined) game of chance. However, in many applications, the (underlying) propensity distribution is unknown; thus, the main challenge for the acting agent is to hypothesize all possible distributions for the random process, enclosed in the set of statistical hypotheses $H_D = \{D_{AR1}, \ldots , D_{ARt}\}$, and to estimate the propensity distribution of the random act based on the only concrete evidence in hand, a (sample) observation.

Since the chance of observing a given piece of (statistical) information, $P_P(I_j)$, depends on the (underlying) propensity distribution, D_{AR}, it is possible to define a measure that allows a comparison between possible distributions based on the level of support provided by the observed sample. The measure, named "likelihood" by R.A. Fisher (1922), is the output of an (algebraic) function that associates each possible propensity distribution, D_{ARk}, with the probability of the realization of the (observed sample) information, if that distribution is the true/ real/ natural propensity distribution of the (underlying) random process, $P_P(I_j|D_{ARk})$, Eq. (1-3).

$$L_j(D_{AR_k}) = L(D_{AR_k}|I_j) = P_P(I_j|D_{AR_k}) \quad (1\text{-}3)$$

1.2.3 Bayesian Approach to Estimation/ Approximation of Propensity Values

The existing methods of propensity estimation can be broadly categorized into Frequentist and Bayesian, which are different in terms of their methods, results, features, abilities, and essentially their foundations. While both approaches use (statistical) observations as the essential component of their methods and may even use the same measures to characterize observations, e.g. Likelihood, they differ in terms of the assumptions they have to make, beyond data, in order to reach conclusions. The need to assume (beyond data) arises from the insufficiency of data for propensity calculations. While frequentist (propensity estimation) methods have a data-centered approach and focus on finding a propensity distribution that best fits the data, e.g. the distribution whose likelihood (given the observed sample data) is maximum, the Bayesian approach to estimation (of propensities) tries to offset the (influence of) data based on the amount, and consequently the sufficiency, of the data at hand.

A main feature of the Bayesian approach is the incorporation of all possible propensity distributions in its estimation procedures, i.e. the set of statistical hypotheses $H_D = \{D_{AR1}, \ldots, D_{ARt}\}$. The incorporation materializes by assigning a degree of credibility/ plausibility, C_R, to each credible/ plausible propensity distributions, D_{ARk}, and to adjust (the propensities of) each distribution by its credibility. The final result of the procedure is a set of all-inclusive propensity estimates, “credible propensity” values, for various outcomes of the random act, and subsequently any sequential sample of the realizations. Eq. (1-4) presents the formal relationship for calculating the credible propensity for the outcome, E_i, of the random process, A_R, under condition C, where the credible propensity is being calculated as the average of the propensity values suggested by various plausible distribution, $P_P(E_i|D_{ARk})$, weighted by the credibility values of the distributions, $C_R(D_{ARk})$.

$$P_B(E_i|A_R, C) = {P_P}^T(E_i|A_R, C) = E[P_P(E_i|A_R, C)]$$
$$= \sum_k P_P\left(E_i \middle| D_{AR_k}\right) \times C_R\left(D_{AR_k}\right) \tag{1-4}$$

Since Eq. (1-4) is a variant form of the theorem/ law of total probability, “credible propensity” is as well called “total propensity,” ${P_P}^T$. Also, Eq. (1-4) has a form similar to the definition of the mathematical expectation of a random set. Because it is possible to show that the same algebra governing (physical) probabilities (of the outcomes of a random process), contained in the set $P_{AR} = \{P_{P1}, \ldots , P_{Pm}\}$, can be applied to the degrees of credibility, contained in the set $C_H = \{C_{R1}, \ldots , C_{Rt}\}$, Cox (1946) and (1961), credible propensity can also be called “expected propensity,” $E(P_P)$. Because credibility/ plausibility values are exclusive to Bayesian approach, credible propensity can also be called “Bayesian propensity,” P_B.

Bayesian estimation of propensities, as presented in Eq. (1-4), requires credibility values, $C_H = \{C_{R1}, \ldots , C_{Rt}\}$, for various credible propensity distributions, $H_D = \{D_{AR1}, \ldots , D_{ARt}\}$. This entails a (mathematical) definition of credibility, according to which a specific credibility value is assigned to each propensity distribution. Such a rule, Eq. (1-5), was first introduced by Thomas Bayes (1763), where an initial/ starting/ prior credibility value for a propensity distribution, $C_R(D_{ARk})$, is being revised/ adjusted/ updated by multiplying it with a factor of the likelihood of the distribution given a specific (sample) observation, $L_j\ (D_{ARk})$. The factor is the inverse of the Bayesian/ total/ expected/ credible propensity of the observed sample, $P_B(I_j)$, Eq. (1-6).

$$C_R\left(D_{AR_k}|I_j\right) = \left[\frac{1}{P_B\left(I_j\right)}\right] \times L_j\left(D_{AR_k}\right) \times C_R\left(D_{AR_k}\right) \tag{1-5}$$

$$P_B\left(I_j\right) = \sum_k P_P\left(I_j\middle|D_{AR_k}\right) \times C_R\left(D_{AR_k}\right) \tag{1-6}$$

Bayesian multiplication, as defined in Bayes' theorem, Eq. (1-5), provides a direct and transparent way to assign a revised (degree of) credibility to each credible propensity distribution such that it reflects the (relative) level of support given to the distribution by statistical data, $L_j\ (D_{ARk})$. The theorem, however, requires a starting/ initial/ prior (degree of) plausibility for each propensity distribution to begin with, $C_R(D_{ARk})$. The collection of starting credibility values for all plausible propensity distributions, $H_D = \{D_{AR1}, \dots, D_{ARt}\}$, is called the prior credibility distribution, $C_H = \{C_{R1}, \dots, C_{Rt}\}$, which follows the laws of probability theory and can be manipulated accordingly.

Since Bayes' theorem does not help in establishing the prior credibility distribution, a separate (mathematical) rule is needed in order to define/ assign a credibility distribution. Defining the prior credibility distribution, however, has turn out to be extremely challenging and the definitions/ assignments introduced, to the point, have remained controversial. So, the questions are why the assignments of the past were challenging and controversial. Since our goal is to help the cause of an agent, who is considering whether to take a risky act, we have to (1) interpret the answers we may find for the aforementioned questions and (2) to assign/ define prior probabilities within the context of making risky choices.

1.3 MOTIVATION

This research is motivated by (1) the importance and the influence of assigned credibility values on the analysis of risky decisions and (2) the shortcomings of the existing methods in establishing credibility values objectively, yet wisely. The differences between various methods and schools of thoughts matter most when data available is limited and inductions about propensities are more guided by the assumptions made rather than the few relevant observations. When data are abundant, on the contrary, the inductions made by various methods are heavily influenced by data and the assessments made by various methods are similar. The imposing challenge is that in many practical situations, where we must make a risky decision, relevant data are limited and assumptions make huge impacts on the analysis performed and the recommendations made.

1.3.1 Black Swans; Rare (Low Propensity) Events with Significant Impacts

Often times we have to make our choices in the face of extremely limited information. In reference to the realization of such unperceived possibilities, Taleb (2007) reminds us the observation of “Black Swans” upon westerners’ discovery of the Australian continent, which invalidated the (perceived) fact that all swans are white.

Examples of the realization of Black Swan events are numerous in engineering. For the first half a century of offshore hydrocarbon production in the Gulf of Mexico, the largest hurricane wave heights impacting production facilities were all less than 26 m. Then, out of the five major hurricanes occurred between 2004 and 2008, four experienced wave heights greater than 26 m and as a result, many platforms designed based on improbability of wave heights greater than 26 m were destroyed (e.g., Energo 2010). In another recent example, a 2014 landslide in Oso, Washington caused a debris run-out ten times further than the estimated maximum distance based on historical data (GEER, 2014). The result, was the destruction of an entire community, including 43 lives, who could have remained alive if we had not perceived such run-out improbable.

Another historical Black Swan in the field of petroleum exploration occurred in 1968, when operators began the exploration of Santa Barbara Channel offshore California. While there was a consensus within the oil industry that these potential reserves were certain and even greater than those of Los Angeles basin (OGJ, 1967), the outcome of the exploration turned out to be disastrous as only two of the thirteen tracts drilled had commercial oil (OGJ, 1968). The main reason for the failure was the extension of the existing knowledge from the geologic structures and trends of onshore basin to the federal lands in the offshore channel (Newendorp and Schuyler, 2000). If experts, had not taken for granted the relevance of the historical data from the (perceived) onshore analogue, they may not have incurred such great losses.

1.3.2 Deficiencies of Existing Method for Estimation of the Propensity Values of Rare Events

Existing approaches for estimating physical probabilities/ propensities with limited information are lacking, especially when propensity values of interest are small. Frequentists estimate propensities of physical events that occur repeatedly given specific observational/ sampling circumstances mainly based on the available data. When data is limited, which is essentially the case for small-propensity events, frequentist estimates are expected to be far from reality; i.e. they do have a large standard error, in frequentist terminology. On the other side, the methodology of Bayesians allows for offsetting the influence of the data on estimates by introduction of the concept of "credibility" for a given set of propensity values, a statistical hypothesis. Credibility makes it possible to also consider (many other statistical) hypotheses that are less supported by the data, but still credible, to be considered in propensity estimation.

The inferential procedures used by Bayesians for incorporating observations into credibility values rely on the use of Bayes' theorem, which takes an (assumed) set of prior/starting probabilities and gives a revised/updated/final set of probabilities. Since Bayes' theorem does not provide any guidance on the choice of the starting credibility values, Bayesians are divided on the choice of prior and its meaning into roughly two camps of subjectivists and objectivists.

Subjective Bayesians believe that individuals are implicitly able to assign prior probabilities given their background information and the implicit assumptions they do make. Nonetheless, (1) the assignment, (2) the application, and (3) the (theoretical and real) justification of the priors set in the subjective Bayesian approach faces a number of challenges. The challenge with subjective assignments of prior credibility values, is the vagueness and obscurity of its assignment process, where a given set of priors is being postulated without informing the process through which the analyst has reached it.

The challenge with the application of subjective assignments, in analyses aimed at guiding a non-personal acts, arises from the inability to examine the assignment and to verify its accordance with the (technical) conventions of a field. This leaves the "faith" in the analyst himself, and not the work itself, as the sole source for the legitimacy of the analysis and its recommendations (see section 4.6.1 for further details). Thus, in many applications, which demand the analysis process and its starting points to be examinable, subjective priors are set aside (Kass and Wasserman, 1996).

Finally, the deeper challenge with subjective priors are their justification, both in regards to the theoretical position they may occupy within the framework of the Bayesian inference and in regards to the reality they may represent. As I will argue (section 4.3), prior credibility is a theoretical term/ construct and is only meaningful within the context of a Bayesian theory aimed at inducting/ estimating propensities, i.e. they are not observable/ measurable in the real-world (independent of the theory). Thus, the only way to set (prior and posterior) credibility values is to use a theoretical assignment rule/ law/ definition and no amount of real-world observation/ information, in itself and without pertinent theoretical definitions, can either set credibility values or make assignments existent in the real world.

1.3.3 Deficiencies of Existing Objective Methods for Establishing Bayesian Credibilities

In contrast to subjectivists, objective Bayesians intend to provide the process of establishing prior probabilities with objectivity and transparency. The application of objective methods to establish non-informative priors for a decision, however, is not free of challenges. The most important difficulty arises from the fundamental difference between the context for which existing objective procedures have been developed and the context of making a risky decision.

As an example to illustrate the challenges with existing Objective Bayesian approaches, consider assessing/ estimating the propensities for different possible occurrence rates of waves exceeding 20 m at an offshore location. A common Objective Bayesian approach is to establish the prior credibilities by maximizing the entropy of information, which produces a uniform probability distribution over the range of possible values.

Figure 1.1 shows two possible interpretations of this approach: a uniform credibility distribution for the mean occurrence rate or a uniform credibility distribution for the logarithm of the mean occurrence rate (since it can vary over many orders of magnitude).

Figure 1.2 then shows the updated credibility distribution (the one used for making decisions about developing infrastructure at the site) for each prior probability distribution given the available information about historical hurricanes at this location; there is a significant difference in these two distributions and they very well could lead to different decisions. For example, there is a 50-percent probability that the 100-year wave height is 30 m for one distribution and more than a 95-percent probability for the other distribution (Figure 1.2).

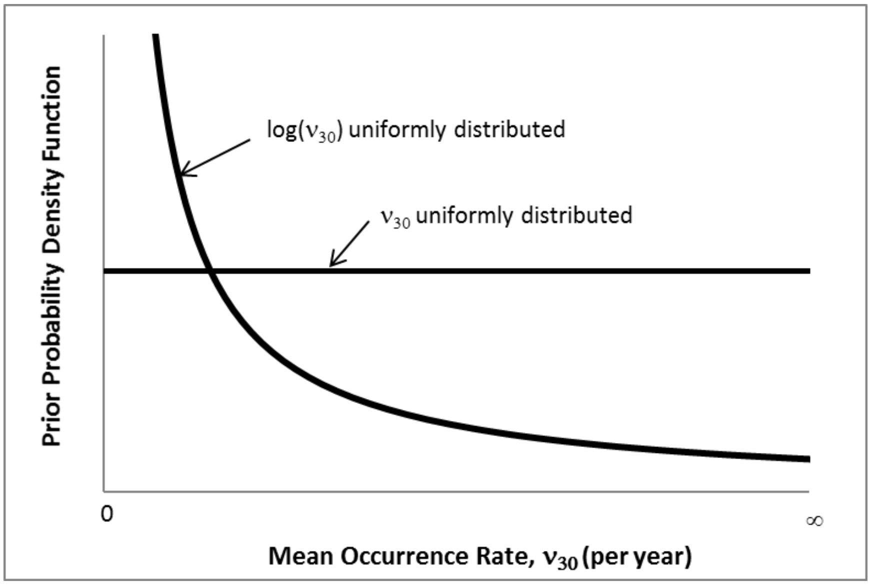

Figure 1.1. Alternative prior credibility density functions for mean occurrence rate of maximum wave height exceeding 30 m (from Gilbert, Habibi and Min, 2012)

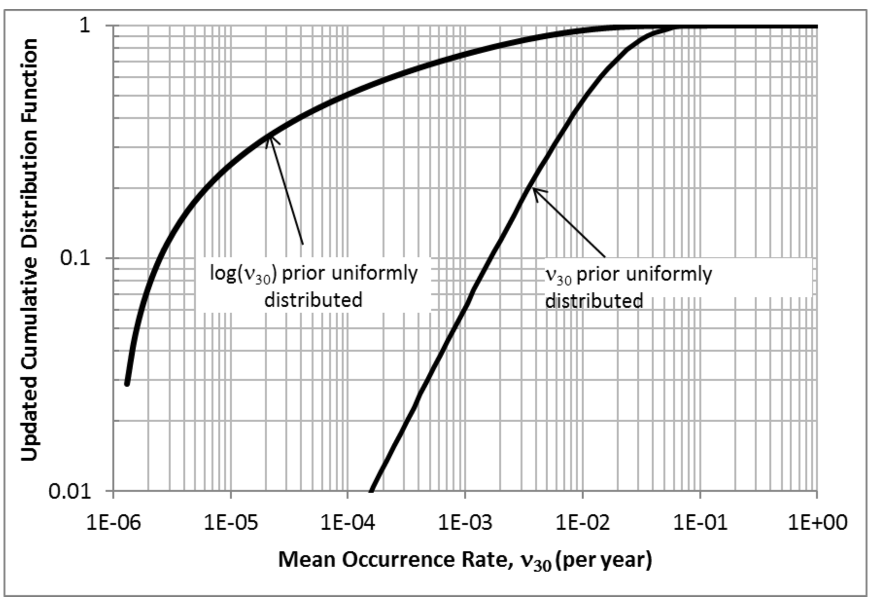

Figure 1.2. Alternative updated cumulative credibility distribution functions for mean occurrence rate of maximum wave height exceeding 30 m given no occurrences in 60 years (from Gilbert, Habibi and Min 2012)

Making decisions rationally requires considering possibilities with considerable uncertainties and assessing/ estimating their propensities based on limited information. An Objective Bayesian approach is being developed with the intent to establish prior probabilities by maximizing entropy of the information in the context of making a decision. This theory, named Decision Entropy Theory, is described in Appendix B. Advancing the development of Decision Entropy Theory is the goal of this research.

1.4 OBJECTIVES

The goals I have tried to achieve in my research are as follows

- Review, analyze, and synthesize the background on prior probability and utility
- Establish a philosophical and logical foundation to define an "objective, transparent and defensible" method for establishing non-informative prior propensities
- Describe and illustrate the principles of a theory that establishes non-informative decision-based prior probabilities (Decision Entropy Theory)
- Identify the practical consequences of applying a theory that establishes non-informative decision-based prior credibilities
- Evaluate whether a theory that establishes non-informative decision-based prior probabilities is "objective, transparent and defensible"

1.5 METHODOLOGY

To achieve each of four objectives set for this research, I have provided with a list of tasks I have performed and a short description of each task.

1.5.1 Review, analyze, and synthesize the background on prior probability and utility

- Task 1-1: to review and analyze the cause of controversies in statistics
- Task 1-2: to review and analyze the background on the concept of probability
- Task 1-3: to review and analyze the background on the Bayesian probabilities, in general, and prior probabilities in specific
- Task 1-4: to review and analyze the background on Bayesian inference
- Task 1-5: to synthesize the unite the reviews and analyses I performed in Tasks 1 to 4

1.5.2 Establish a philosophical and logical foundation to define an "objective, transparent and defensible" method for establishing non-informative prior propensities

- Task 2-1: to distinguish between the role of assumptions and information in statistical inductions
- Task 2-2: to define various types of probability including an objective Bayesian probability
- Task 2-3: to scrutinize the concept of information within the framework of statistical inference and to define the state of no-information/ignorance
- Task 2-4: to incorporate the concept of the "Bayesian robot" in the doctrine of objective probabilities

- Task 2-5: to emphasize on context-dependency of all types of probabilities
- Task 2-6: to incorporate the concept of "consensus" in the foundations of objective Bayesian approach
- Task 2-7: to establish the concepts specific to the theory of decision entropy

1.5.3 Describe and illustrate the principles of a theory that establishes non-informative decision-based prior probabilities (Decision Entropy Theory)

- Task 3-1: to present the principles and mathematical axioms
- Task 3-2: to investigate the uniqueness and consistency of the proposed approach
- Task 3-3 to develop a framework for accommodating multiple decision alternatives.
- Task 3-4 to create and present simple, illustrative examples
- Task 3-5 to evaluate the theory in the context of "objective, transparent and defensible"

1.5.4 Identify the practical consequences of applying a theory that establishes non-informative decision-based prior credibilities

- Task 4-1: to formulate a set of simplified decisions, including those involving the value of obtaining additional information to support making a decision
- Task 4-2: to develop general analytical and/or non-dimensional numerical solutions to these problems.
- Task 4-3: to explore numerical approaches to solve complicated problems

1.6 OVERVIEW

This research tackles the task of assessing the outcome of a process, when we are extremely uncertain about the processual tendencies due to the limitations of "relevant" observations. I have tried to offer answers to the questions of "when" we may face extreme uncertainty, "how" we can assess it, and "what" is the interpretation of the made assessment. Answering the three questions required different kinds of investigations, which is evident in the diversity of the contents presented in the dissertation. While answering to the question of "when to expect" was (partly) done through examining historical cases and inductively extending the lessons to the cases in the future, answering to the question of "how to assess" was done by developing mathematical formulae on the foundation of theoretically-postulated principles. Finally, answering to the question of "what interpretation" was done through analyzing existing concepts and developing new ones.

The question of "when to expect" is being addressed in this Chapter, Section 1.3.1, as well as early parts of Chapter 4, Sections 4.1 and 4.2. By refereeing to the concept of "Black Swan," I have proposed that we must "always" consider the possibility for the realization of unprecedented events and take their extreme consequences into account, if we are facing a non-trivial decision. The claim regarding the potential for the realization of Black Swans is justified by citing historical cases and by reasoning inductively, that if such a phenomenon has been repeatedly taking humans by surprise, it is likely to do so in the future, and we need to have the foresight of considering yet-to-come Black Swans in our simulations of future.

To review the historical cases of Black Swans, I have mainly referred to the literature, including the well-known book by Taleb (2007); however, I have highlighted a number of cases as I see them both humbling and enlightening. One case, described in Section 1.3.1, was the 1968 hydrocarbon exploration in Santa Barbara Channel offshore California, where spatial proximity of the prospect with a highly productive area mislead oil companies about the profitably of the venture and cause them to lose a huge fortune with little return. The other case, described in Section 4.1.1, was the 1998 colossal collapse of the hedge fund LTCM, in the aftermath of a series of economic turmoil abroad, which was assessed as improbable by Nobel prize-winning frequency models that LTCM was using to valuate financial options.

The question of "how to assess" is being investigated in Chapters 4, 5, 6, where Chapter 4 offers the principles of a method developed to address the question, Decision Entropy Theory (DET), and Chapters 5 and 6 illustrate the application of the theory to two practical problems. Also, an extensive description of DET including its mathematical formulation is being presented in Appendix B.

It is being suggested that a way to quantitatively address Black Swan outcomes is to first consider the set of every possible outcome of significant consequence, and then to assign Bayesian prior probabilities to possibilities, before incorporating any empirical observation. After adopting a measure of information to precisely quantify the degree of non-informativeness for any set of theoretically-assigned Bayesian priors, DET offers a hierarchy of "significant variables" within the context of rational decisions, whose theoretically-assigned Bayesian priors must be minimally informative. The significant variables, as postulated by the principles of DET, are (I) “decision outcome,” (II) “difference in preference,” and (III) “learning potential/information value.” Nevertheless, minimally informing significant variables does not guarantee the non-informativeness of other variables, and on the contrary, it may dictate an informed distribution for another insignificant variable, if the variable depends on significant variables. The level of such informativeness is determined by the specifics of the relation between insignificant and significant variables, including the linearity of their relation.

The question of “what interpretation” is being approached in Chapters 2, 3, and in Appendix A. In Chapter 3, I have suggested that three concepts of propensity, possibility/potentiality, and credibility, encompass all senses of probability; thus, all probabilistic assessments are of observational, and/or potential, and/or credential nature.

By defining credibilities as imaginary relative weights assigned to a set of possibilities, I have argued that credibility values cannot be obtained from empirical observations and must be assigned without resorting to empirical justifications. After declaring Bayesian probabilities as credibilities, I have defined non-informative priors as theoretically-established credibilities, whose assignment are aimed at not informing (certain aspects of) the assessment. As so, the assessments made by non-informative Bayesian probabilities are of potential-credential nature, and those based on partially-informed probabilities, are of observational-potential-credential nature.

The suggested interpretations of probability are coming from a more fundamental process-centered worldview, *Philosophical Processualism*, which I have presented in Chapter 2. This theoretical account puts “concepts” at the center of human understanding and suggests ways in which we can come to share, not only perceivable concepts, but also imaginary concepts, such as credibilities, with precision.

Also, to increase the precision and the clarity of the concepts used throughout the research, a glossary of terms is being offered in Appendix A. The compendium is better understood as the outcome of my efforts to analyze various concepts, rather than to collect a dictionary of technical terms. Since I have developed the definitions across a rather long timespan, there is considerable variation in the definitions' tone and format, as well as in their generality and complexity.

Chapter 2 Philosophical Processualism; Foundations of Pragmatic Mathematics

2.1 INTRODUCTION

The chapter introduces a philosophical perspective, I have labeled “Philosophical Processualism” due to the centrality of the concept of process in its tenets. While it seems to me that the perspective is of immense potential and may become the basis for the development of a rather extensive philosophy of science, currently, the view is at its infancy. The philosophical nature of perspective is due to its focus on the logical and critical examination of the rational grounds for the concepts we do use in mathematics, as well as natural and social sciences. While the intended application of Philosophical Processualism, makes its metaphysical concerns narrower than a general philosophical theory, establishing its position regarding a number of metaphysical issues are necessary and unavoidable.

2.2 THE CONCEPT OF PROCESS

The concept of process is central to the philosophy of probability I am proposing here. The concept is closely associated with the belief in causality, where each object has somehow come to existence. The “how” of realization is being captured by the concept of process, in which the outcome is conceived to be generated by a mechanism taking some inputs, and delivering output under specific circumstances.

In continue, I will introduce processes by giving a definition, elaborating on the significance of conceiving processes to our cognition and action, and by stating the objectives we may achieve by such conception. I will then elaborate on our two fundamental ways of understanding concepts including the concept of process. I will also present the (general) types of processes we deal with, and finally, enlist and describe the cognitive operations we employ to theoretically organize processes. The term operation is used as an equivalent to the term process. The use of interchangeable names is aimed at avoiding terminological monotony, which occur if we label every realization mechanism a process.

2.2.1 Introducing Processes

A process is a set of mechanisms that cause changes in a set of attributes of the objects involved. The definition of a process includes reference to a number of other concepts with labels such as object, attribute, change, and mechanism. Although I elaborate on these terms in the next section, I do recommend to take them as primitive concepts, meaning that defining them in terms of other concepts adds little clarity, and the audience must ultimately rely on their (default) cognitive apparatus for understanding.

The concept of process is significant for both cognition and action, and its significance is due to the embedment of "change" in the fabric of the existence. In terms of cognition, there are few worldly phenomena not involving changes of some sort, and in terms of action, its objective almost involves bringing a change to the worldly affairs or preventing such changes from realization (status quo). Existence, from a more broad and general perspective, has started and continued with unexpected changes, revolutions, as well as gradual changes, evolutions. The study of change enables man both to understand the worlds he encounters, and to act in accordance with his understanding. For such a study, I argue, the concept of process is essential.

The main achievement of forward-looking studies is predicting the changes due to be realized in the future. The forecasts may then be used to verify the accuracy of the predictive machinery, in pure cognitive studies, or as the basis for purposive actions, in applied studies. In either case, contemplating the concept of process enriches the developed predictive machinery by considering the causal grounds of the predicted changes. Besides, the contemplation emphasizes the fact that a predictive machinery, similar to a physical machinery, relies on a set of processes to deliver its output. If the machinery is intended to be copied and different copies are expected to yield consistent output, the implanted processes ought to be known and to be imitable; otherwise, the apparatus remains isolated and the goal of duplication becomes unfulfilled.

2.2.2 Cognizing Processes

Our ability to cognize processes stems from our wide-ranging ability to comprehend things, and it is this ability that makes our experiences meaningful and sensible. Since cognition and the resulting sense of meaningfulness occur inside the mind and are hidden from direct sensory observations, theorizing about cognitive mechanisms partly relies on external manifestation of our cognition. The appearances are in the form of behavior including linguistic utterances, as well as our reflections on the conscious part of mental processes.

2.2.3 Perception and Conception

Humans can understand things through a variety of methods including the processing of their sensory experiences to create percepts of things outside their minds. Perception, as a process resulting in meaningful mental experiences, must be differentiated with sensation, which only results in the reception of stimuli from one's external environment by internal sensors. Figure 2.1 presents a schematic diagram depicting the complexity of perception process, where external stimuli is being received and processed by a sensor and the outcome is being reprocessed by a modeler to deliver a percept. The meaningfulness of the sensory experience, in contrast to the customary assumption, does not come from the work of the sensors, but mostly from the work done by the modelers. Meaning, generally speaking, has more to do with (the output of) models than data.

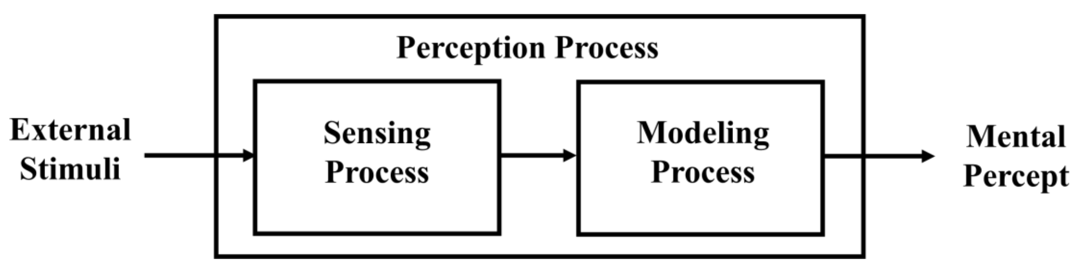

Figure 2.1. A sketch of perception process, characterized by processing of sensation by a modeler to yield a meaningful sensory experience

Humans can also expand their understanding by processing of the things inside their minds, including externally-triggered percepts, to form concepts. Figure 2.2 presents a schematic diagram depicting the sub-processes of perception including simplification and modeling processes. There are a number of similarities and differences between the two processes of perception, portrayed in Figure 2.1, and conception, portrayed in Figure 2.2. One major difference between the two is the input to the processes, where perception takes stimuli generated in the outside environment, whereas conception takes the (already processed) entities existing inside the mind. Apart from different input sources, the output for both processes have the shared feature of being formed inside the mind. The similarity in features, can also be reflected by similarity in names, where we relabel percepts as "perceptual concepts." The new labeling unites all meaningful things for the mind under a group, whose name include the term "concept."

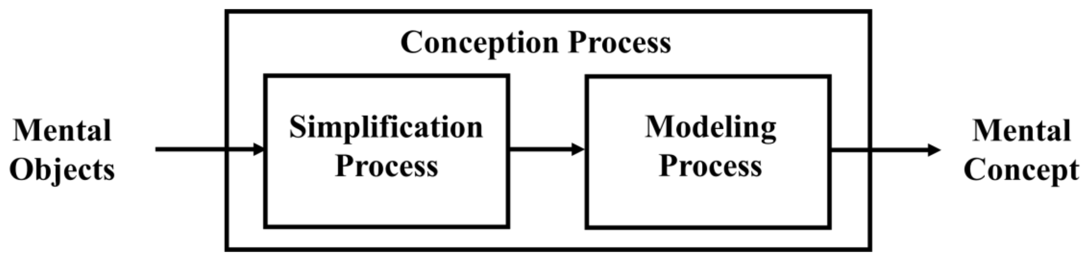

Figure 2.2. A sketch of conception process, characterized by simplifying a set of mental objects and processing them by a modeler to new meaningful mental objects

Another similarity between the processes of perception and conception is the involvement of the sub-process of modeling, that going through which seems to bring meaning to mental entities. The significant semantic role of the model can be explained by its function. A model's job is not just to combine or rearrange components, but to add things not present in input. In a geometric analogy, a model connects some dots, not just by associating them, but by creating and adding many more points of specific spatial position to the few points presented to the model. Thus, the sub-process of modeling along with the sub-process of simplification, depicted in Figure 2.2, may not be considered equivalents of the two processes of analysis and synthesis.

A modeling process, when it results in the development of concepts expressible in human language, may be also be called an interpretation process, and its outcome an interpretation. An example interpretation, is the process-oriented philosophy such as the one I am presenting here for the case of cognition, which could not be produced simply by analyzing and synthesizing the current knowledge. My interpretation, as a new coherent conception, has come to realization by creating new entities and by using them to connect existing data, in the forms of empirical observations and theoretical conceptions made by other researchers. In a broader perspective, if I had followed a heavily empirical approach and overly relied on observations, I would have possibly obtained semantically-deficient conceptions. The meaning has come more from (internally postulated) models than (externally received) data.

2.3 ONTOLOGY OF CONCEPTS

The account of cognitive processes I put forward in the previous section, introduces concepts as things in the mind, without elaborating on how they do look like. The answer to the later question clarifies the ontology of concepts within the framework a process-centered philosophy I am advocating. Below, I will discuss how concepts could look like, where they could be present, and how they could come to realization. Besides, the offered ontology establishes the foundations of the processual philosophy and will be used in developing pertinent views toward a range of issues including mind, language, and an epistemic machinery, which I do label as "pragmatic mathematics;" a system of linguistically manifested mental constructs aimed at guiding action with precision.

2.3.1 Mental Representations

Mental representation is one of the common phrases used to label objects of the mind. The term representation not only distinguishes between what concept refers to, the referent, but also indicates the existence of a difference between the form of the concept and its referent. The difference in the appearance depends on the way representation may look like. Although it is tempting to hypothesize about THE appearance of the concepts, as many philosophers have done, I think it is more realistic to introduce an open set of possible appearances for the concepts. My pluralist approach roots in my belief in (1) multi-modality of human cognition, and (2) cross-populational modal variabilities. While cognitive multi-modality of humans enables them to form their mental representations in a variety of ways, modal variability across human populations results in differences in individual tendencies to use various modes of representation.

Visual representation seems to be the most common way of forming concepts, in which the referent is being represented in a seeable format. For percepts, the representation could be a mental image of the referent, similar to what mind experiences when it really sees the object. In a more sophisticated method, pictorial images can be substituted with visual symbols and enable the mind to conduct various kinds of symbolic manipulations. Examples include representation of (a) a referent by a word, or (b) the non-perceptual concept of number with a numeral, or (c) the concept of process with an arrowed box, such as one I have used throughout this chapter!

2.3.2 Abstract Objects

It is also possible not to specify the form of mental entities and simply regard them as abstract objects. The approach may be warranted as we can only learn about a very small part of our mental objects and processes, which is accessible to our consciousness. However, I suspect the main motivation of the advocates of the approach, from Plato in antiquity to Gottlob Frege in 19th century to prominent contemporary philosophers, is to endow humans with the ability to directly access absolute and timeless truth. Such a privilege, would be endangered if one has to address the potential differences between individuals' mental representations of the same referent. By evading the question of the appearance of concepts, it is possible to overlook individual differences and assert the accessibility of the truth. Nevertheless, to ignore potential variations in individual conceptions, the theorist needs to deny mind as the place for the concepts, and to introduce another place, both outside physical world and human minds; the third realm.

I see the Platonist view of cognition, where concepts are outside time and space, as inconsistent with the process-oriented account of cognition, in which every object is the outcome of a process unfolding in time and space. While we may not know the exact process underlying an outcome, our ignorance does not strip the process out of its spatiotemporal medium. However, by rejecting the Platonist theory of concepts, an alternative theory needs to explicate the phenomena the Platonist view was suited to explain, albeit with minimum recourse to metaphysical constructs such as the third realm. Among other things, we must explain the exemplification of universal properties; the seemingly eternal recurrence of the same properties in seemingly infinite number of instances. For example, it must be clarified whether there physically exist a property of redness, as we do perceive it, apart from the instances of red objects; and if so, how?

2.3.3 Universals as Outcomes of Typical Processes

I propose a two-step strategy for addressing the problem of universals. The first step involves the use of the familiar philosophical distinction of type vs token, and regrading a universal property as a type and its realized instances as tokens. The second, and more important, step involves the introduction of a generative process with the ability to produce tokens of a type. The introduction will not be merely nominal, but aspectual, where a number of aspects of the token-generation process are being explored. The reason for the significance of the second step is its reliance on a materially tangible process, as a means of explication. The lack material hypotheses of the kind, I believe, is one of the reasons for non-ending disputes in philosophy; whereas in science, the combination of the hypotheses with the empirical investigations regarding the feasibility and the specifics of the hypothesized processes, results in further discrimination among the contending explanations and fewer argumentation.

I offer the concept of "code" to refer to the means through which a process executes its mechanism in the way it does. In an analogy with cooking, a code is like the recipe for preparing a meal, containing the information regarding the participant objects in the process and their properties, as well as their levels of participation and the changes required for triggering and maintaining the process. The analogy is not perfect, as cooking is a process guided by a purposive agent, whereas many token-generation processes advance without the involvement of biological agents and the recipe is being activated and read in an automatic fashion.

The code of a process is embedded in the material participants of the process in the form of their physicochemical arrangements. The code needs not to be in a single objects and parts of which could be distributed among all participants. Once all objects of certain character take the appropriate spaciotemporal position relative to each other, the process commences according to the specifics of the objects' physicochemical arrangements. Even though the recent description of the code closely fits our understanding of chemical reactions, it can be used to comprehend not chemical processes, as well.

2.3.4 Example Processual Codes

For instance, one may ask about the code for the process of the gravitation of heavenly bodies and wonder about the way spatially remote objects may participate in the process. One answer is that the mere possession of the mass by two spaced objects triggers a gravitational process, in which objects move to close the gap. While many readers may be doubtful of the description and find it somehow unbelievable, it is noteworthy that these readers share the same doubts with the developer of an influential theory of gravitation, Isaac Newton, regarding his own discoveries. Based on some accounts, it was Newton's skepticism about the feasibility of such simple code that motivated him to introduce force, as an intermediate theoretical construct, which did provide a more believable code for the process. Either way, the code for the process of gravitation is stored in each and every particle of both celestial bodies in the form of the physicochemical arrangements we may simply come to label as mass.

In another example, we may inquire about the code, according to which some objects look red-colored to us. While one may be content with a general description of the code for the color property and attribute it to the specifics of the physicochemical arrangements of the particles of the colored object, another may remain skeptical of the description and enquires about the details of the arrangement. To find the details, we need to learn about the underlying process resulting in the overlying feature. The learning, among other things, requires the development of novel hard means enabling us to sense the participating objects and to conceive their attributes. Such process, may result in characterizing the arrangement in terms of the wavelength and the intensity of the lights emanating from the colored-object.

The skeptic may accept the presence of the delineated arrangement, but ask why the light coming from the object has such wavelength. Even though the skeptic's inquiry may seem to be about the "why" of the influencing features, further explanation may be provided by answering "how" the influencing features come to possess their current arrangements. In other words, we need to undercover one more layer and learn about a deeper process resulting in the target intensity and wavelength of the light. By conducting more research, we may find out that light's features are determined in a process, where emission spectra and light absorption of the surface of the colored object come to play.

The unimpressed skeptic, can pose another why question and send us to conduct another inquiry resulting in the discovery of subatomic particles termed quarks and in the introduction of quark's properties as determinants of the deeper arrangements within the colored object. In any case, the code of the process is partly captured by the physicochemical arrangements of the object we may simply come to label as color properties.

2.3.5 Tower of Physical Processes

The chain of investigations we conducted to find the code for colored appearing process can be used as a prototype for framing other enquiries aimed at discovering the code for other attributes of the objects. One lesson we learned is the endless nature of such enquiries, where each discovery can entice us to ask another why question and motivates us to find another how answer; though, at a deeper level. To capture this endless nature, I sketched a visual schema (Figure 2.3), in which a series of processes stack up to eventually produce the target attribute perceivable to us. The processual tower, as an imagery of the concept, differs from a physical tower in that the first is being developed from top to bottom and the second is being constructed from the bottom up.

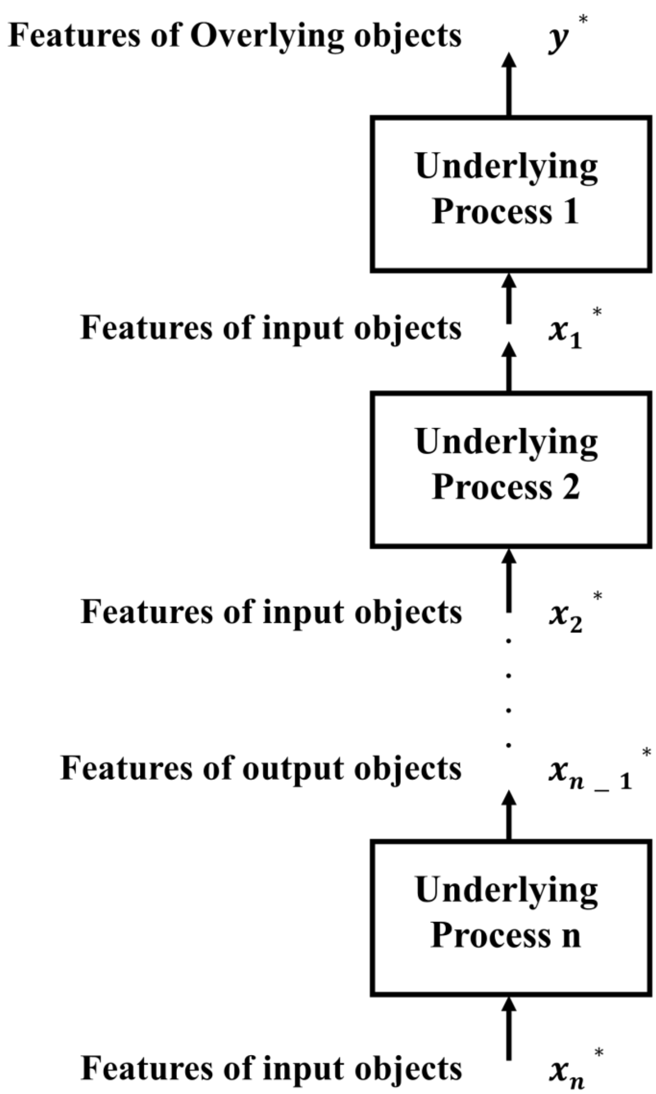

Figure 2.3. A sketch of non-ending tower of physical process, resulting in an apparent feature perceivable to humans

A pragmatist, after conceiving the notion depicted by non-ending tower of processes, may seek to find where, if any, she/he can abort the investigation and declare sufficiency for her/his practical purposes. Fortunately, the answer could be positive, depending on (1) the precision of our processual knowledge at levels above a specific height along the tower, and on (2) our sensory access to the state of the objects participating in processes higher than the specific height. The two conditions for investigative sufficiency are the resultants of the goals the pragmatists may possibly pursue. The goals could either be (a) the prediction of the appearance of overlying objects, as they might influence the outcome of another process, or (b) the alteration of the overlying appearance by manipulation of the features of the intermediate objects.

The degree to which the pragmatist could achieve her/his aims depends on the degree to which her/his processual and sensual knowledge is available and precise. If the current knowledge above a specific level enables determination of the overlying feature for any to-be-generated token, the pragmatist is good to go and can stop further inquiry; otherwise, she/he needs to go deeper. The level of determinacy of processual and sensual knowledge can be analogically described as the tightness of the grasp over the tower. As soon as one could hold a tight grip on the tower, she/he gains predictive and (potentially) manipulative power over the process. The tight grip, in statistical language, translates into a perfect correlation between the features of the intermediate object, at a given level of the tower, and the features of the appearing objects.

2.3.6 Processual Interpretation of Tokens and Types

My answer to the philosophical question regarding the ontology of universal categories; e.g. objects, properties, or relationships; can also be used to give an account of the distinction between a type and its tokens. If so, the claim that every universal category is out there in the world in the form of a processual code, can be used to elucidate the type-token distinction.

Tokens, can be defined in terms of the process generating them. At every occasion, when the code of a process is being activated, the same token does come to realization. The term “same” exclusively refers to common non-spaciotemporal aspects of instances; otherwise, each token is a distinct, at least because of the difference between the time (and potentially the location) of its generation, compared to other instances. Thus, in a more precise labeling approach toward the inter-relation among tokens, we may label tokens “similar” and correspond the similarity with the “same” non-spaciotemporal features, which they share.

While the concept of type can be similarly elucidated in terms of a generative process, we may first discuss an alternative definition of a type, based on its tokens. The discussion matters not only because of practical concerns, as the extensional account of types are commonly employed in science and engineering, but because of epistemological concerns, as many theories in mathematics including mathematical statistics are founded on such conception.

A type, in an extensional account, is defined as the set of its tokens. It is possible to use the shared feature as the basis of relating a number of tokens, and to form an extensionally defined set, i.e. by enlisting its members and specifying the extent of the set. However, such a finite set, my not be labeled as the type of the tokens and could only be considered a finite sample of tokens. A set that could be more intimately associated with the concept of the type, is the imaginary set containing each and every token generated in the past, along with the ones, yet to-be-generated in the future. The challenge with such an infinite set is its abstractness, as we can only access a finite subset of already-generated tokens, which have become accessible to us. Nevertheless, one wonders how accepting an abstract set as the definition of a type differs from considering type as an abstract entity inaccessible to us. In other words, gathering tokens in a set, either a finite sample or an infinite unavailable population, does little in explaining how the tokens are related to the type.

To identify a type, I suggest that we forgo the temptation of giving an extensive definition and instead, identify it with the "typical code," according to which tokens come to realization. After discovering the code, the type is completely defined and the record of the realized tokens could be considered as a mere tally of historical events, differing only in terms of their spaciotemporal positions.

2.3.7 Processual Codes; General vs Detailed Descriptions

A philosophical theory of type may only marginally help in offering the details of the code for a specific type. In fact, any general philosophical theory, including the one I have put forward, can only present the nature of types from a very broad and general perspective. The code for every specific type is initially hidden and can only be discovered through contextual inquiries, where we use all sorts of material and mental means, to form a body of knowledge termed natural philosophy, in earlier times, or science, in modern times. In fact, the constant expansion of scientific knowledge is the result of mankind's effort to decode more of specific types.

A processual code for a certain type (of object, property, or relation), at any point in time, may only be partially available, regardless of the amount of efforts made to develop the natural philosophy of that type. For many contextual types, our collective knowledge is partial and imprecise and can yield correct predictions of the state of a fraction of to-be-generated tokens. Besides, we have the tendency of trying to establish meaningful connection among an ever-larger number of contextual types, and to form more comprehensive and less contextual types. Due to the desire and will for perfecting and generalizing the natural philosophy, the descriptions of the typical codes we see as plausible changes along the time; mostly in an evolutionary fashion and sometimes in a revolutionary manner. Since my aim here is not to develop a philosophy of science, I leave it here by emphasizing that postulating a universal code for types, which follows a metaphysical belief in causality, does not imply that we do or will have the code for a certain type.

2.3.8 Philosophical Processualism; A Pluralist Perspective

My proposal that every universal property is out there in the world in the form of a processual code enabling the generation of its tokens, may seem inconsistent with my main thesis that concepts are inside the mind and not out there in the world. The apparent inconsistency, may also be perceived due to the contentions of the philosophical schools of thoughts historically advocating various elements of my position. While the claim that universals and types do exist makes me a realist, the assertion of universals' being out there in the world and not in an abstract realm, makes me an Aristotelian realist and in opposition to Platonist realists; however, the assertion that we understand the worlds available to us only by means of our concepts, seems to make me a conceptualist.

Nonetheless, I will demonstrate how it is possible to hold the mentalistic thesis without compromising the conception of the processual code. In fact, I will use the concept of the generative code, not only to support the mentalism of concepts, but also to elaborate on necessary conditions for a shared perception, when an external referent may result in the formation of similar perceptions across a population of minds.

A process-centered perspective can be employed to further support the idea of taking concepts as mental entities. The support may be more if we first clarify the case for perceptual concepts, which are meant to represent things in the physical world rather than the worlds of human imagination. The elaboration strategy will be similar to the 2-step analysis I gave for the age-old philosophical conundrum of universals. In the first step, I use the type vs token distinction and consider a percept generated in an individual's mind as a token of the corresponding perceptual type; and in the second step, I introduce the generative process with the ability to produce the target perceptual tokens.

My take on tokens can be used to explain the realized objects of cognition. Taking an instance perception as a token means that at every occasion, when the person is being exposed to the same external stimuli, the same mental representation is being formed in her/his mind. The term "same," as I clarified before, refers to the aspects of representation associated with non-spaciotemporal features shared among the external stimuli.

Likewise, my account of types can be employed to elucidate the concept of cognitive universals. The goal is to explain what makes humans to have the same perception of the same external stimuli emanated at a variety of spaciotemporal positions, assuming the phenomenon does occur. While a universal sensory experience can be ascribed to the presence of a typical code, the challenge is to introduce the generative process, which uses the typical code to generate the experiences. More specifically, we must clarify whether the generative process is out in the world, as Aristotelian realists claim, or is inside the mind of the experiencer, as conceptualists claim; and if the latter one, we must answer whose mind is the place for the process.

2.3.9 Universals as Outcomes of Multi-World Processes

I propose that the process resulting in the realization of percepts is a two-world serial process, the first one in the physical world outside human mind and the second inside human mind. It follows that the percept's processual code is a composition of two codes, one for the tower of physical processes and the other for perception process. Figure 2.4 illustrates the sequence of processes, which eventually yield a percept at one's mind.

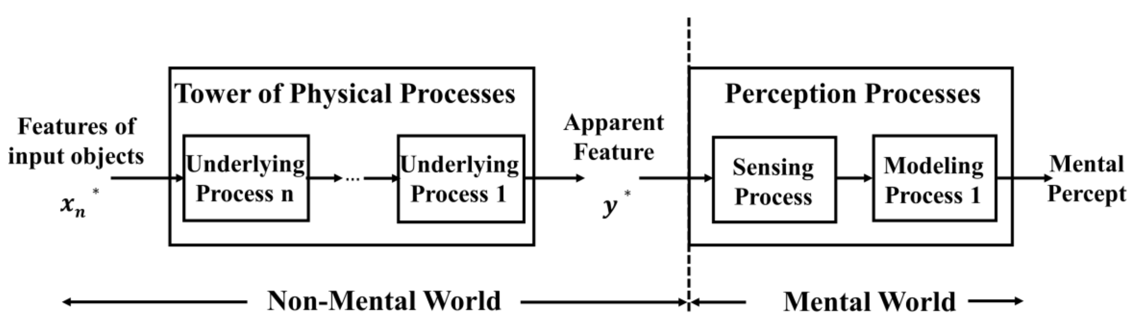

Figure 2.4. A sketch of the sequences of multi-world processes involved in formation of a perception

The multi-world account of universal percepts is in contrast with the positions of both Aristotelian realists, who believe that types are out there in the world independent of our minds, and conceptualists, who believed that universals do not exist outside of the perceiving mind. The position of the multi-world account is that even though we may hypothesize the existence of universals without perceiving them, both existence of universals and their qualitative specifics, as percepts, depend on our minds. Every universal does exist because of the uniformity of its pertinent typical codes, one in the non-mental world and the other in the mental world; without each, the universal would cease to exist.

The feasibility and the utility of developing perspectives that exclude human mind from their accounts of universals, is questionable. Such theories, may need to explain the feasibility of their imaginary conditions, and to elaborate who and how has come to be aware of the existence of universals. Is development of a form of awareness regarding universals possible, without the presence of a subject? If not, and the existence of a subject is necessary for obtaining the information concerning the presence of the property, how the subject can obtain such information? Until we answer the how-questions about the feasibility and the specifics, the imaginary scenarios and their potential answers may negligibly expand our understanding. Responding to the later questions requires theoretical development of the kind we are pursuing in our process-centered account, and may possibly result in constructs similar to ours.

Ignoring the role of the mind by philosophical theories may root in some implicit assumptions the theorists are making. It appears that the development of such perspectives; which are very common across a wide range of human inquiries; is only possible because (1) we, in the first place, have come to be aware of our environment the way we do, and (2) we implicitly assume that our way of perceiving the world is the default and is ubiquitous. To me, the label "made in human universe" is invisibly stamped all over our hypotheses, and I see it impossible to form any understanding, let alone a hypothesis, outside "human universe."

2.3.10 Dependence of Universals on Cognitive Processes

The dependence of the universals on our minds, both in terms of their existence and their qualitative specifics, can be illustrated by an example. Let's consider the conceptual category of properties (of objects) and focus on the much-discussed property of color. A textbook introduction to the philosophical problem of universals goes like this: we see distinct and different objects, who are all red-colored. Is there any such thing as redness aside from red-colored objects? To answer this question and other questions of the kind, philosophers have developed their opposing theories of universals, including the process-centered account I elaborated above. Since we have already provided a number of answers to the question, I suggest we put the assumptions implicit in the framing of the problem under scrutiny, and ask whether it is possible for a red-colored object not to be red, at all? And if so, how? My answer, is a surprising yes!

In general, every entity (either an object, a property, or a relation) does only exist FOR subjects, who are able to recognize the entity as such. For the case of the color property, an object could only be red-colored TO the subject with the abilities of (a) sensing the color property, in general, and (b) the red color, in specific; for subjects without the ability (a) the object is not colored at all, and for subjects with the ability (a), but without the ability (b) the object is colored, but not red. The proposal, can be further demonstrated with a review of color-perception abilities of a number of other animals. While a kind of fish named skate lacks ability (a) and only sees in black and white, dogs do have abilities (a) and (b), but with a limited extent compared to those of humans; for dogs, a red-colored object may look like what humans perceive as greyish brown to light black. In contrast, most birds not only have the ability (a), but have a stronger ability (b), and unlike humans, are able to differentiate between various shades of red, rather than bunching a wide red spectrum into a single type.

Characterizing alternative perception mechanisms and the resulting differences between their generated percepts, may turn out to be a difficult task, if not impossible. In the above example, we reviewed how dogs may see a red-colored object like what we see as greyish brown to light black. One may ask about the reason for our imprecise knowledge of the doggy perception and for ascribing a spectrum of colors to the outcome of a perceptive process, less discriminating than ours.

A simple explanation, could be given by reviewing the visual schemas in Figure 2.1 and Figure 2.4, illustrating the two sub-processes involved in perception, namely sensing and modeling. Since sensing abilities are closely related to the physical construct of sensors; hard devices used to receive external stimuli; it is possible to analyze sensors at the service of dogs, and to compare them with those of humans. The comparison reveals that in contrast to humans whose eyes are equipped with three types of cones to sense the three colors of blue, green, and red, dogs' eyes are only equipped with two types of cones to sense the colors of blue and green. The observed hard difference, however, is not sufficient in answering the posed question, as percepts are the product of a joint process between hard sensors and soft modelers. Eventually, it is the specifics of the soft modeler, i.e. its typical code, that determines how the outcoming percept would looks like to the perceiving mind. So, to find out how an object, perceived by humans as red-colored, looks like to a dog, you have to be a dog!

2.3.11 Processual Uniformity Hypothesis

Considering an entity (either an object, a property, or a relation) as a universal, relies on two processual uniformity assumptions, one non-mental and the other mental. While uniformity of a non-mental process captures the belief that numerous activations of a process under similar non-spaciotemporal conditions, consistently yields the same outcome, uniformity of a mental process captures the belief that when the subjects are being exposed to the same signals, they interpret them in a similar fashion.

The universality of percepts among humans, must not to be considered a natural law, but a physical hypothesis, whose validity ultimately depends on the (potential) existence and the specifics of the differences among percepts formed in the minds of humans. When perceptual universality is being used as a postulate in theories aimed at predicting perception-dependent changes in the behavior of subjects, the outcoming predictions must be examined and adjusted for deviations from the universality. However, the more realistic theories are those, who allow deviations from the universality and take effect of the deviations on predictions into account.

2.3.12 Worldviews; a Cognitive Explanation

Every philosophical account of universals, including ours, depends on the perspective, according to which it has been developed. The term perspective is used to indicate that the differences among the answers provided by various accounts, is not only due to methodical differences, but mainly due to their diverging worldviews. The simplest approach to worldviews is a descriptive one, either by giving an exclusive portrayal of one view; such as one I have presented for the processual perspective; or by comparing diverging views and highlighting the similarities and differences among their positions; such as one presented in philosophy textbooks, handbooks, and encyclopedic entries. Nevertheless, the main scholarly challenge may not be in giving descriptions, but in developing explanations about the reason for the divergence of worldviews among the brightest human minds, who have tried to make a broad sense of the world they are living in. Below, I try to take a small step and offer a hypothesis aimed at explaining some of the underlying reasonings for divergence among views toward universals.

I suggest that we take each worldview as a distinct cognitive perspective. If so, a process-oriented account of realizations may be able to explain the divergence of worldviews by hypothesizing the cognitive processes underlying various worldviews. In other words, different takes on universals is due to differences in the conceptions of the mental processes, whose outcomes are types.

2.3.13 Example Explanation: Nominalism

To illustrate the method, I take the nominalist position toward universals and I hypothesize a kind of cognitive process, which may result in development of the perspective. The reason for choosing nominalism is its striking position toward universals, which looks bizarre and preposterous to our common-sense minds. Nominalists, deny the existence of universals, neither in human minds, nor in the material world (outside human mind), or in the metaphysical realm of the abstract entities. A universal, nominalists assert, only exists as a name given by our minds to a set of empirically-observed tokens, and informs nothing typical beyond the similarity of tokens' appearances.

A nominalist position, I propose, may be the result of conceiving the types as the output of cognitive processes generated by, what I do label as, an "experiencing-organizing mind," which is able to process bodily experiences into percepts, and to organize the generated percepts into sets. When the mind is exposed to tokens with different appearances, its organizing abilities allows it to separate and group them into sets. Besides, an experiencing-organizing mind is capable of externally manifesting its internally created divisions by naming the sets and revealing them to other minds. The formation and the presentation of named groups is the basis of the nominalists' claim that universals do not have an existence beyond those names. Figure 2.5 schematically illustrates the chain of processes resulting in the formation and presentation of a named group to the world apparent to humans.

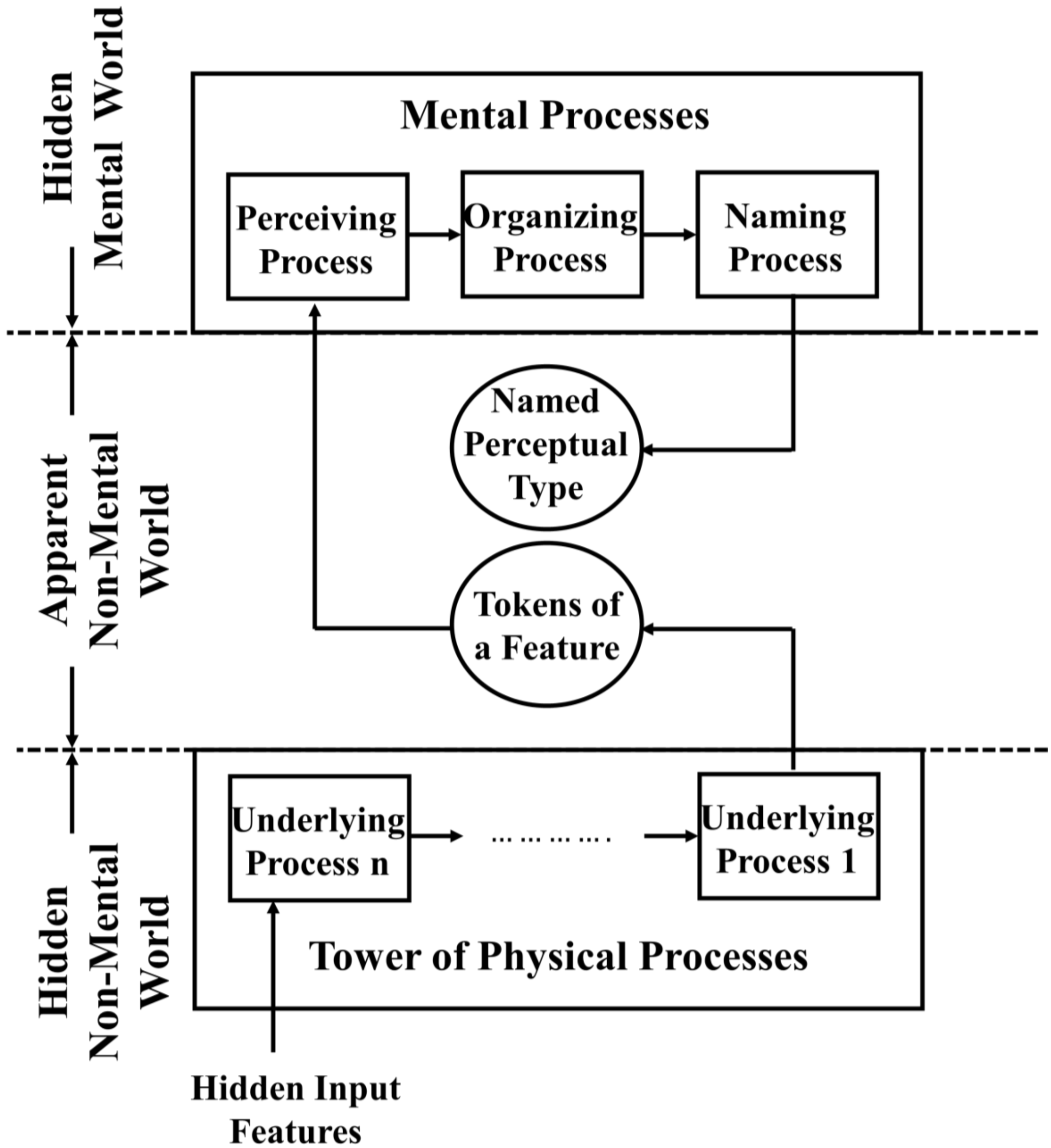

Figure 2.5. A sketch of a possible nominalist conception of the processes generating a nominal type

The nominalist denial of universals is possibly not due to their recognition of "experiencing-organizing mind," but due to their lack of recognition of, what I do label as, "questioning-hypothesizing" mind. While the focus of an experiencing-organizing mind is on forming organized description of observations, the focus of a questioning-hypothesizing mind is on inquiring about the reasons for the similarity of the grouped tokens. By inquiring, the mind introduces the possibility for the existence of something beyond or beneath the observations, that is making the appearances similar. To find an answer to the inquiry, the mind needs not only to imagine the things beyond or beneath the observations, but also to hypothesize how the imagined objects and/or their properties are related to the appeared tokens.

2.3.14 Explaining Specific Types

Nominalists may not supplant a general explicative-associative account of universals to their descriptive-collective account, due to the imaginary nature of general hypotheses. Even our processual account of universals and the concept of typical code may not sway nominalists in our direction, as it fails to provide what matters to nominalists. Since the nominalist position is being developed by focusing on the non-hidden world of appearances, what may change their position is observations of different kinds than the ones provided by similar tokens. Such observations may be provided by a different kind of philosophy, natural philosophy, which focuses his resources to explain specific types and tries to supplements the soft means of a questioning-hypothesizing minds with the hard means of new devices.

The new hard devices enable natural philosophers to look beyond or beneath the apparent tokens and to observe entities hidden to man’s default hard sensors. Observing previously hidden object and/or their properties, not only makes plausible the idea that there is something beyond or beneath the apparent tokens, it also helps the hypothesizing mind to explain the realization of the apparent tokens by associating them with the previously hidden entities. While new experiences make it difficult for the nominalists to shrug the proposed association off as a mere imagination, he may soon come back to repeat his nominalist assertion, this time for the newly observed types, and claim that there is nothing beyond what has been observed to the point. To react, the natural philosopher has to embark on another inquiry and uncover one more level of the hidden tower of physical processes, with the hope of satiating the nominalist. However, one may wonder about the possibility of a positional change under any amount of new observations. Maybe, the nominalist is averse to believing in anything beneath or beyond observables!

2.3.15 Non-observed Entities

My examination of the nominalist position on universals had motivations beyond offering a new analysis of the opposing positions on one the oldest and most abstract problems in philosophy. Such abstract debates, to me, are ubiquitous and are very similar to disputes on many unsettled foundational issues across a wide spectrum of fields in physical and social sciences, as if the “disagreement” itself is a (higher order) universal, relative to the (lower order) field-specific disputed universals. If such disagreements are taken to be instances of a universal, uncovering its typical code will illuminate the circumstances under which a token dispute will be introduced to a scholarly exchange.

I suggest that any hypothesis involving unobserved entities will generate a dispute, whose settlement would only be possible either by observing entities (or their associates), or by taking them as primitive entities. Thus, it might be tempting to remove all non-observed or non-postulated entities from our proposals and to make them undisputable. In fact, there was a movement among logical positivists in the 20^{th} century to cleanse theoretical science and philosophy of such terms, which eventually faded with limited success.

The inability to eliminate non-observed entities from hypothesized relationships could have a variety of reasons. For one thing, science and humanities contain so many proposals that include unobserved entities and reconstructing them needs enormous efforts beyond anyone's abilities. For the other, there might not be an easy and clear way for the reconstruction of many theories, and if it were, each field's experts would probably have done it before philosophers. Even when it is possible to drop non-observables from the relationships among entities, their preservation is an indication of the theoretician's belief regarding their existence and the role of the entities in making the relationship meaningful. While the hypothesis may turn out to false upon further discoveries, it is such works of imagination that may guide us to discover more about beneath or beyond! If we had limited ourselves to experiences, we would all be nominalists!

2.3.16 Cognitive Explanation of Perspectival Opposition

The cognitive approach we followed for analyzing the nominalist perspective can also be used to analyze other positions, including our processual perspective (Table 2.1). In the approach, the positions taken by the perspective are seen as the outcomes of a cognitive process, where the mind is conceived to plays a certain role due to the abilities ascribed to it. For the case of nominalism, the recognized mental faculties include experiencing, organizing, and labeling, and as a result, a nominalist-envisioned mind acts as a classifier and namer of the outside world.

Orientation, I propose, is another property that, in combination with the perceived mental abilities, causes perspectives to have different positions toward universals. Orientation can be evaluated relative to either of the two references of (a) human mind, and/or (b) the apparent world, where (a) can be described as outward, inward, or both, and (b) can be described as within, beneath, and beyond. While orientation (a) reveals the theoretician's awareness and recognition of the influence of cognitive processes on our knowledge, orientation (b) reveals theoretician's psychological predispositions toward a suite of fundamental concepts including order, law, uniformity and truth.

An outwardly orientation toward human mind indicates that the philosopher is either unaware of the relationship between one's cognitive processes on her/his knowledge, or discounts it. In contrast, an inwardly orientation reveals the philosopher's emphasize on the role of the mind in shaping human knowledge; in an analogy, she/he sees knowledge as a participant of an internal theater, directed by the mind. In between the two extreme positions, an inward-outward orientation recognizes (1) the significant role of human cognition on the nature and the character of knowledge, as well as the role of the differences among individual cognitive processes on diversity of knowledge perspectives across human population.

Most prominent perspectives including Aristotelian realism, Platonist realism, and even nominalism, embrace a mentally-outward orientation. Aristotelian realists see things as out there in the world and believes that humans can discover those things “as they are.” By not elaborating on the mechanisms enabling humans to understand things as they are, Aristotelian realists implicitly take the claimed epistemological ability as a metaphysical postulate.

Also, the other major group of realists, the Platonists, see things as instances of the forms existing outside human mind, but also outside of the material world, and in a 3rd non-spaciotemporal world. Platonists take even more metaphysical postulates than Aristotelian realists; not only they postulate the existence of a 3rd abstract realm, they also take connections between each of the two physical and mental worlds and the abstract world, as granted.

Although the nominalists’ position that universals are only names in the mind and they do not exist in the physical or in the mental worlds, may appear to be mentally-inward, their real position might be as mentally-outward as those of the realists’. While nominalist do not set the mind aside and ascribe minimal capabilities of organizing and naming it, their focus is all on observables in the material world. By denying anything beyond observables in the outside word, nominalists reveal their mentally-outward orientation, as well as their minimum reliance on implicit metaphysical postulates.

On the other side, conceptualists' position is mentally-inward, as they see nothing but the mind and its processes, and as such, it leaves little to be justified by metaphysical postulates. While the full commitment of conceptualism to the mind, makes it an ideal foundation for fields such as cognitive science, social sciences, and humanities, where the phenomena are mainly shaped by the workings of the minds of humans. However, its mentalistic perspective hinders its contribution to physical sciences and engineering, where phenomena are commonly perceived to be shaped independently from the workings of the minds of humans.

In between, is Philosophical Processualism, which does not focus on physical or mental worlds at the expense of the other, and sees them jointly from a more comprehensive perspective. Philosophical Processualism also resorts to few metaphysical postulates in establishing the relation between the two worlds, as it introduces a specific physical process through which, features of the outside world can be sensed and interpreted by the mind. Philosophical Processualism's explicative sufficiency, one may argue, has become feasible due to contemporary advances in physical and cognitive sciences, which were unavailable to philosophers of the past. Regardless, the processual perspective, as we demonstrated, has more explicative power than perspectives of the past and can even explicate the conception of each past perspective, using known cognitive processes. Table 2.1 summarizes the features of Philosophical Processualism, in comparison with existing perspectives.

Table 2.1. Characteristics of Processual Perspective in comparison with other perspectives, in terms of perspectives' prioritization of worlds, mind, and mental abilities

Perspective	**Position on Existence of Universals**	**Place of Universals**	**Significant Worlds**	**Role of the mind**	**Emphasized Mental Abilities**
Aristotelian Realism	Acceptance	External (material)	External (material)	Neutral Discoverer (of Outside)	Experiencing, Reasoning, Revisioning
Platonist Realism	Acceptance	Third Realm (non-material and non-mental)	Third Realm (non-material and non-mental)	Connector to Third Realm	Reasoning, Accessing the third realm metaphysically
Conceptualism	Acceptance	Mental	Mental	Director of a Mental Theater	Conceiving, Reasoning
Nominalism	Denial	Nowhere (only as mental Names)	External and Mental	Classifier & Namer of Outside	Experiencing, Organizing, Labeling
Processualism	Acceptance	Joint worlds (both external and mental)	External and Mental	Observer & Hypothesizer (of Outside & Inside)	Experiencing, Organizing, Questioning, Hypothesizing, Revisioning

Orientation toward the apparent world, is another kind of orientation causing mentally-outward perspectives to take different positions regarding universals. The orientation determines the perspective's answer to the question whether there are hidden things connected to the apparent things, and if so, where are the hidden things? Since the only way to consider non-observed entities and their relationships is through human imagination, the answer to both questions require taking a metaphysical stance regarding the legitimacy and the quality of relating objects of human imagination to observed entities.

I suggest, we interpret each stance in the light of its advocates psychological predispositions toward fundamental concepts including order, law, uniformity and truth.

Without getting into the details of the suggestions, its general scheme can be constructed based on the following opposing ideas.

Idea1: The only way for two ordered systems of associated entities, one inside human mind, and the other in the physical world, to exactly correspond each other is for both to correspond a model system, unaffected by each. In other words, both physical and mental systems must be copies of a source.

Idea 2: The how of the correspondence between two systems does not matter. What it does matter is whether we can (someday) observe the correspondence between imagined and observed entities.

Idea 3: We have to declare whatever, which has not been observed, as imaginary and unreal, at least for the time being.

2.4 CONCEPTS; MANIFESTATION

An external manifestation of the contents of one's mind provides a way for others to become aware of the contents and act upon their awareness. The efficacy of the process in making other agents cognizant depends on their ability to perceive the external manifestation and to interpret the percept by associating it to a concept in their minds.

Manifestation of the concepts can also be studies in terms of their effects on others' conception of an individual action, where the act is a process that the purposive agent triggers and/or maintains to create new or to modify existing entities inside her/his/its mind or outside in the environment. If other cognizant agents assume that an individual's action is dependent on the presence and the specifics of a set of concepts inside her/his/its mind, they can then associate the perceptual aspects of the action to individual’s mental contents. In other words, the (manifested aspects of the) action becomes a symbol of the (hidden) concepts. The mental concepts associated with the action, can also assist others in providing an alternative account of the individual’s purpose, in which the motivation is explained in terms of the mental contents rather than the apparent environmental changes caused by the action.

Manifestation of concepts in a community of agents can also become a means for collective actions aimed at pursuing common goals among the participants. Whether the goals be mental or environmental, the role of the manifestation is to facilitate the division of labor. When the common goal is to advance a cognitive process, the division is (entirely) cognitive, and when the goal is to advance a physical process, the division is partly physical and partly cognitive, as any action taken by purposive agents requires pertinent cognitive support.

2.4.1 Linguistic Manifestation

The ability of community members to associate actions with concepts can be used by them to take certain actions for the pure purpose of sharing mental contents. Since in communicative processes, the taken acts are symbolic and provide nothing beyond perceivable manifestation of the concepts, the community members can come to develop a system of communication by entirely relying on a set of symbols; a language.

A language can be defined as a conventional system of perceivable symbols used by community members as a means of communication. Since percepts can be created through external stimuli of various kinds, linguistic symbols can be in auditory, visual, or kinesthetic formats. For example, natural human languages, which are by default in auditory format, have also been presented in other formats, e.g. visual form of written languages and sign languages, or kinesthetic form of braille language.

Development of a language adds a third world of linguistic expressions to the two worlds of physical entities outside human mind and of cognitive entities inside human mind, whose relationships could be investigated. Since some aspects of the relationship between physical and mental entities were being studied before, I focus the attention on the connection between mental entities and their linguistic expressions, and examine how they can be translated into another. The term "translation" indicates the directionality of the relationship, where the processes involved in converting meaning into a linguistic form have to do the opposite of the processes converting the linguistic expressions into entities comprehensible for the mind. Figure 2.6 gives a visual depiction of the manifestation processes, where a mind reveals its meaningful contents by encoding it into linguistic expressions, and the interpretation process, in which a mind converts linguistic forms into comprehensible concepts.

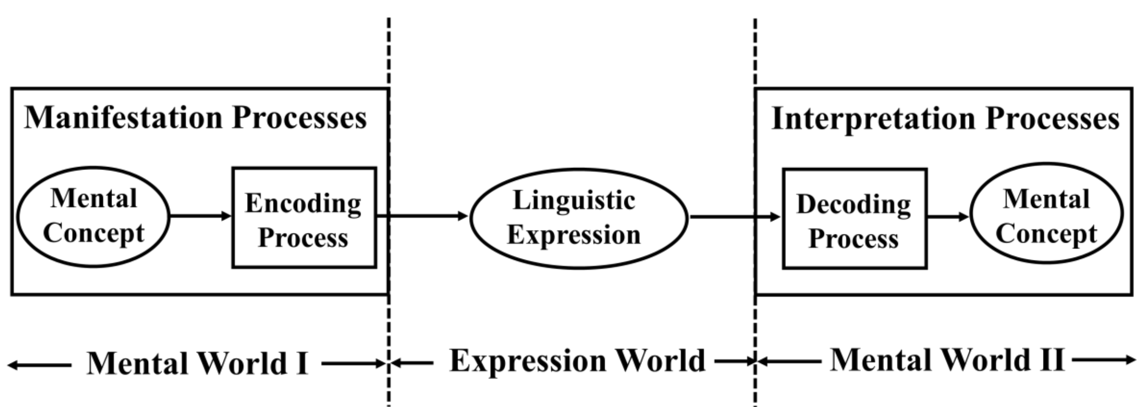

Figure 2.6. Processes involved in the translation of mental entities to/from linguistic expressions

The translation processes illustrated in Figure 2.6, could be studied in a systematic fashion to provide detailed answers to two different "how questions;" (1) how to form linguistic expressions, and (2) how to interpret meaning. Table 2.2 summarizes a comparison between the two, in which the contents of the mind is labeled as thought and eternalized expressions as language. The distinguishing features of the two worlds, according to Table 2.2, is the latency of mental contents and the symbolism of linguistic manifestations. While different labels may be used for the unit of mental contents, my preference for the word "concept" is to emphasize the cognitive dimension of knowledge in processualism, and to contrast it with other perspectives, where one may choose the label "referent" to ignore the role of the mind when denoting things in the outside world, for the case of Aristotelian realists, or when denoting imaginary things inside the mind, for the case of Platonists.

Table 2.2. Characteristics of mental contents in comparison with their linguistic manifestations

	Feature	**Unit**	**Other Unit Labels**	**Linguistic Study Discipline**	**Study Subject**
Thought	Hidden	Concept	Referent	Semantics, Pragmatics	Meaning
Language	Apparent	Symbol	Signifier, Denotation, Name	Phonology, Morphology, Syntax	(Apppearance of) Formation

I propose to take conventional acquaintance as a necessary condition for the translation of cognitive contents to and from linguistic expressions. The necessity can be explained based on the latency of mental contents and the symbolism of linguistic manifestations. The only way for the use of a purely symbolic form as a means for communicating a concept among individuals is for the participants to agree on the representativeness of the symbol. If so, for a new member of a linguistic community to generate expressions understandable to other members, she/he has to embark on a convention learning process and to learn how others linguistically refer to a concept.

Being exposed to a linguistic community may also assists the individual to use her/his mental faculties to acquire new concepts, which were otherwise unavailable to him/her. Although the mental faculties of an experiencing-organizing mind suffice for the purpose of learning the name of the concepts already present in her/his mind, developing new concepts by reliance on linguistic expressions requires additional faculties available to questioning-hypothesizing mind. Upon encountering a new symbolic form, a questioning mind wonders whether the form is content-less or is a lexical concept, i.e. a linguistic expression of a concept; and to find an answer, the mind enters an investigation process to form a hypothesis regarding the meaning of the term.

Concept acquisition may be facilitated and accelerated through the use of language. Generally, a new concept may be acquired through a set of processes, collectively labeled as thinking, where the mind defines the new concept in terms of other (already existing) concepts and the specific way they are connected to each other. In other words, the new concept becomes a member of a conceptual system composed of concepts associated in certain ways. The development process can be enhanced by the use of language in two ways: (1) by associating labels rather than the concepts themselves, and (2) by using linguistic association rules that are partially content-independent, e.g. apply to all concepts of a conceptual category. The gained efficiency can be analogically explained by the ease of moving a set of handled containers, in comparison to those without handles.

A conceptual system, as the outcome of a conceptual development process, can be visually imagined as a network of various connected entities, organized based on their types. Independent of the use of language in system's formation process, the language can be used to name the imagined entities and to facilitate the communication. In an example system, whose concepts are of the fundamental categories of objects, properties, and relationships, the system can be defined based on two sets of organized associations; (1) objects related to each other in a hierarchical fashion based on their types, and (2) properties of each object related to properties of a subset of other objects.

Figure 2.7 gives a visual illustration of such exemplary conceptual system, in which the sketch on the left depicts the hierarchical relation between objects, and the sketch on the right depicts the relation between the attributes of the objects. One apparent difference between the right and the left depictions is the larger number of connections between properties of the objects than those between the objects. The intuitive reason I may offer for the presence of direct relation among properties of indirectly connected objects is that direct connection between objects not only cause some of their properties to depend on each other, but also cause the properties of other objects connected through them to depend. While it could be the case that the relation between properties of directly connected objects might capture the relation between attributes of indirectly connected objects, the emphasis on depicting the relation may be due to the need for defining one attribute based on the other.

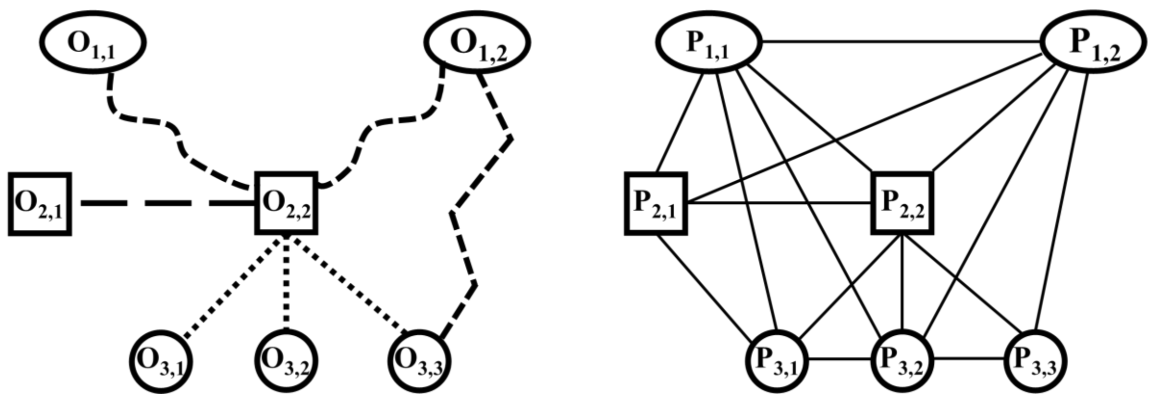

Figure 2.7. A visual representation of a linguistic system; Left: associations between token objects of three different types, Right: associations between objects' properties

A natural human language can be used to express meaningful systems of various degrees of perceivability, in which the degree indicates the relation between the number of perceivable entities to the imaginary ones. The meaningfulness of the system does not depend on its degree of perceivability and on the presence of the imaginary entities, but on the mind's ability to associate every entity with a mental content; e.g. when the mind can imagine an imaginary entity, it feels as if the linguistic expression is meaningful. In other words, meaningful entities could be perceivable and/or conceivable, depending on the degree to which the mind relies on sensory means for comprehending them.

Since meanings are in the minds of individuals, a fundamental question is whether the conceived meaning of an entity in a subject's mind is the same as the meaning conceived in another subject's mind. The question, reframed within the modern context of objectivity and subjectivity, is whether a subjectively conceived entity can be objective, or at least possess some objective aspects? And if so, how can we assess the objectiveness of the subjectively imagined concepts?

I propose a communal verification-based account of conceptual objectivity, in which a conceived entity by members of a community is considered to be objective, if the subjects consistently deliver similar manifestations of the concept in any number of agreed-upon verification processes. While the stated operational definition gives a way for assessing conceptual objectivity, it does not offer a method to community members for ensuring the objectivity of their conceptions. The assurance can be provided by processual homogeneity, where using the same processual codes for (a) the conception and (b) the manifestation results in the similarity of the generated concepts and their manifestations. However, due to the frequent observations of natural variabilities among biological agents, it is realistic to assume conceptual heterogeneity across a community and to focus on finding ways for facilitating processual assimilation.

Processual assimilation can be pursued by capitalization on the learning abilities of the subjects and by providing them with instructions on how to to create and/or with feedbacks on how to modify their concepts. The nature and the extent of instructions and/or feedbacks depend on the degree of perceivability of the target entity, where less perceivable entities are in need of more extensive training. To simplify exemplification, I offer two distinct cases: (a) an entity of high perceivability, and (b) an entity of low perceivability, and demonstrate the differences between the extents of the instruction and the verification process for each case.

Consider (a); the process of imagining a red-colored entity by subjects across a community, and the task of assimilating the imagined perceptual concepts by conforming them to an agreed-upon perception, in which the color has a certain hue. The instruction and verification programs could be as simple as training subjects by presenting them with different shades of red and specifying the target hue, and verifying the correctness of the subjects' imagination by asking them to identify the agreed-upon hue among the shades in a color wheel. Further instructions in the form of offering distinct names for varying shades of red may assist subjects to sharpen and expand their visual imagination and linguistic manifestation capabilities.

Consider (b); the process of conceiving an imaginary entity, e.g. the number three hundred twenty-seven billion, by subjects across a community, and the task of assimilating the imagined concepts by conforming them to an agreed-upon conception, who possess certain properties. An instructional program can provide subjects with recipes on how to construct such an imaginary object, and the objects of the same kind, as well as recipes on how to relate such imaginary objects to each other. As a result of participating in the program, the subjects become able to consider the intended imaginary entity in their minds and reveal the compliance of their conception by manifesting it through a certain linguistic description, e.g. "327,000,000,000."

The above example, may cause a nominalist to raise the objection that conceiving an imaginary entity is nothing but introducing a name without any contents, and such entity is essentially meaningless. In specific, the nominalist can claim that manifestation of the number "three hundred twenty-seven billion" as "327,000,000,000" is just a formal transformation, where a symbolic form, defined in a natural human language, is translated to another form, defined in a numeral system such as Hindu-Arabic. In fact, the nominalists' claim is likely to be true, when the imaginary object is property-less or isolated!

I assert that the significance of an imaginary entity is not due to its name, but due (a) to its properties, (b) to its relations with other entities and/or their properties, and (c) to its participation in certain processes. Similar to physical objects, who have multitude of properties and are related to each other, imaginary objects can also possess different properties and be associated with other non-physical objects. In either case, the conceptions of a system, either of physical or of imaginary entities, can be linguistically described, such the example visually illustrated in Figure 2.7.

While conceiving a property-less or isolated imaginary entity might be nothing but a name assignment, conceiving a property-possessing, and/or other-connecting, and/or process-participating is about exploration of properties, relations, and processes. Nevertheless, the name of an imaginary entities could be assigned based on properties it possesses, the relations it does have, or the processes it participates. I claim that imaginary entities called numbers are single-property beings, i.e. their values, and are named accordingly. However, each number is connected to infinite other numbers through their values; for instance, the natural number (with the value of) 327,000,000,000 is related to both its preceding natural number, 326,999,999,999 and the number 1, and their relationship can be expressed as 327,000,000,000 = 326, 999,999,999 + 1.

The systemic nature of an imaginary entity not only makes the entity multifaceted, but also does the evaluation of the objectivity for its conception by subjects, as the entity can be described and assessed in terms of its relation with many other entities in the system. As a result, the accordance of the linguistic manifestation of the conception by a subject with an agreed-upon form, may only provide a partial verification of the concept's normal compliance. For instance, a multifaceted verification program may reveal the objectivity of a subject's conception of the number 327,000,000,000 in some regards, e.g. the name or the value property, and its non-compliance in other aspects, e.g. value relations.

I suggest that an imaginary entity may have a meaning beyond what is bestowed by the systems it belongs to, and that imaginary systems are not semantically self-reliant. The idea is that in every imaginary development project, the properties its imaginary entities do possess, the relationships they are in, and the processes they participate in, can be traced back to a set of percepts and/or to cognitive processes, who have perceptive roots. In another words, imaginary properties, relationships, and processes are all "grounded," and are either abstracted from or analogous to perceivable realities; and so, a subject who cannot perceive, would not conceive!

The thesis of the perceptual roots of imaginary constructs can be illustrated through an example. Consider the imaginary number 327,000,000,000, and see if we can uncover some of its perceptual origins through investigating the number's properties and/or the relationships it does have with other numbers and their properties. Since numbers are single-property imaginary entities and are named according to their value property, an uncovering strategy could be tracking the relationship between property value of 327,000,000,000 and the property values of other numbers associated with it. A promising association to investigate is the recursive relationship between the value of a number, its preceding natural number, and the number one; here 327,000,000,000 = 326, 999,999,999 + 1. Since the preceding number is also in a recursive relation with its own preceding number, and so on, the value of the number 327,000,000,000 is in relation (2-1) with so many repetitions of the number valued as one.

$$327{,}000{,}000{,}000 = \sum_{\boldsymbol{i=1}}^{327{,}000{,}000{,}000} \mathbf{1} \tag{2-1}$$

The meaning of any other natural number, following an analogous approach, can be tied to the meaning of the three concepts of (a) the number with value property one, (b) the equality (of an apparent feature), and (c) repetition (of the objects with the same feature). The number one, I propose, is grounded in our ability to perceive objects as distinct wholes, who are separate from other objects. The concept of equality is grounded in our ability to perceive a certain property in a variety of objects and our inability to further differentiate between the specifics of our perceptions; e.g. perceiving two red-colored objects as having the same red shade. The concept of repetition, which is closely associated to the concept of equality, is rooted in both our ability to compare objects in terms of their certain properties, and our life experiences of finding objects with similar appearances, whom we cannot differentiate; e.g. identical twins.

Set is another fundamental imaginary concept, whose perceptual roots may be simply explained by our ability to distinguish entities from another. However, further grounding must be presented for the formation of two distinct types of sets: finite, and infinite. While the concept of finite set is grounded in the limits of our perceptual abilities and opportunities, including our attentional resources, the concept of infinite set is rooted in the limitless ability of typical codes in generating tokens.

Finite sets can be easily understood from an empirical perspective and be attributed to one's simultaneous awareness and ignorance of the things outside her/his mind. Both awareness and ignorance can have a natural character, due to the presence (or lack thereof) of our abilities to cognize certain types of things and our opportunities to conceive tokens of the types within our spaciotemporal limits. Besides, the ignorance could also be intentional and a result of the willful choice of allocating our attentional resources, where we deliberately ignore entities possessing or lacking certain properties.

In contrast to finite sets, who we can be understood intuitively and on the basis of material limitations, infinite sets are anathema to any empirical tenet and can only be understood from a processual perspective. I propose to take an infinite set as the collection of all entities that can be potentially generated by a certain processual code, where the realization of the potential does not depend on the code itself, but on the material requirements necessary for the implementation of the code. Consequently, when the code for an imaginary process is available to us, we can generate any of its outcomes (almost) at will, given that we have provided enough material power to keep our code-containing minds up and running.

For example, by having the recursive code for the generation of natural numbers, we can imagine 327,000,000,000, as the number whose property value is 326, 999,999,999 more than the number valued as one. Our ability to imagine the target entity has little to do with the fact that we may have never conceived it before, and is due to our access to the processual code that can generate the entity. To generalize from the example, we can interpret the claim that the set of natural numbers is infinite as a statement about the recursive code, which allows the generation of any imaginary number based on its value difference with the number one.

The generation of the (conception of) outcomes can be extended to physical processes, whose codes we have imagined and translated into one of our imaginary languages. In fact, a prediction of a future event in the outside world seems to be the result of us running a mental simulation of the corresponding processual code under some presumed conditions. For example, by having the codes for the orbital movements of the planets in the solar system, and assuming that the codes remains unchanged into the future, we may predict the relative positions of the planets in 327,000,000,000 years from now. Again, what seems to be limitless is not the reality itself, but is our ability to imagine the reality to be such!

2.5 CONCEPTUAL CATEGORIES

Concepts may be distinguished by their features and those with common features can be grouped into common concept types. The process of typification can be continued to form hierarchies of more general types, which include more concepts but are characterized by less common features. The set of most general distinct types of concepts can be labeled as conceptual/semantic categories, where a category specifies the broadest feature type for cognizable things under the category. Since antiquity, when Aristotle first introduced categories; though for the contents of the linguistic expressions rather than the thoughts (behind them); prominent philosophers including Kant and Husserl, have come to introduce their list of semantic and/or syntactic categories. Because my goal here is to provide just enough of background knowledge needed to support the fundamental categories I intend to present, especially the category of processes, I abstain from going through the details of the categories conceived by other philosophers. The interested readers can consult the wealth of reviews and analysis presented by modern scholars of philosophy. A number of such discussions are cited in the reference section of the script.

The set of categories I conceive as fundamental include the three categories of (1) objects, (2) properties, and (3) relationships. A number of concrete examples for each conceptual category includes a celestial body or a biological species as objects, color or mass as properties of objects, and mutual gravitation of masses or dependence of buyers' demand on the market price of a commodity as relationships (among properties of objects).

While defining either of the three categories in terms of the other two may seem feasible, I, as a principle, do suggest to take these categories as both primitive and distinct. Humans, I postulate, are able to cognize things as wholes, and in addition, are able to understand attributes objects may have, as "aspects" of the whole. The distinction between an object and its properties implicitly assume that we may not be able to present a precise definition of an object by enlisting those of its properties, which we have come to be aware of them. My assumption is in line with the fundamental postulate of Gestalt psychology, asserting that the totality of an object is different and more than the sum of its parts. The postulate is empirically supported for the case of perceptual concepts by observations indicating organisms' tendency to perceive entire patterns and configurations rather than pieces.

The totality postulate, nevertheless, seems somehow unreasonable. The counter intuitiveness of the postulate, I believe, arises from an implicit comparison of mental processes with physical ones, where belief in causality entails that things cannot come to existence out of nowhere and the components of a whole must come from somewhere. In other words, physical generative processes are of analytics and/or synthetic nature, and are governed by the laws of conservation.

To solve the puzzle, we must explain how it is possible for the mind to insert things not sensed by bodily sensors. An explanation, I claim, lies in the complexity of perception process, where external stimuli must first be transformed to a readable format for the mind, Figure 2.1. Once transformed, any other format-compatible internal mental process, such as one labeled in Figure 2.1 as modeling process, can manipulate the intermediate representation to yield a new one. This way, a new object, e.g. a curve, can be superposed upon a set of existing objects, e.g. some points, to form a coherent whole containing starting objects.

2.6 PHYSICAL PROCESSES

An outcome of a physical process, as well as other components of the process including its trigger and circumstances, can be described by a number of aspects/ characteristics, where each aspect may take any of the possible states/ conditions. The (mechanics) of the process can be represented by the relationship between the components of the process, summarized in a (set of postulated/ hypothesized) relationships between the characteristics describing these components. When the states, at which the initial and circumstantial aspects are, remain unchanged in subsequent runs of the process, it is expected that the states of the aspects characterizing the outcome also remain static. On the contrary, fluctuation in certain initial and circumstantial aspects of the process results in the variation in particular characteristics of the outcomes of the process. When the mechanics of the process and/or the pattern governing the fluctuations (in subsequent runs of the process) is unknown, the process is called a "random process."

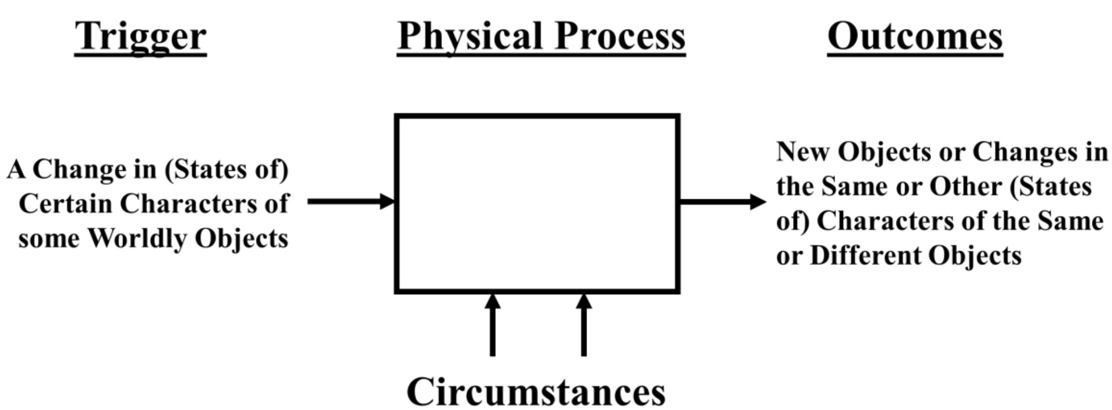

Figure 2.8. A schematic diagram depicting a physical process, characterized by the trigger, circumstances, the process, and the outcome

Material objects of the world can be characterized by the state of their various attributes and both, the objects and their attributes, are the outcomes of the physical process the objects have gone through. A physical process, in its general forms, can be described by the relationship governing the attributes characterizing the (the initial, the circumstantial, and the final) components of the process. A change in the process occurs when one (or more) aspect(s) of the (components of the) process shifts its current state to another possible state. While a change in some aspects of the initial and circumstantial components may influence some aspects of the final outcome, it may not influence other aspects.

Although knowing the (structural) relationships governing a process enables prediction of the changes caused in the attributes of the outcome given a change in initial and circumstantial attributes, it neither inform (1) the possible states each attribute can take, nor (2) the (natural) tendency to take some states more than the others in various runs of the process. However, if we know the answers to questions (1) and (2) for all attributes of the initial and circumstantial components of the process, in a quantitative format, we can answer the questions for the attributes of the final outcome. Quantitatively, the (structural) relationships governing the process are represented by a set of algebraic equations, and the (natural) tendency to discriminate among the states of the initial and circumstantial attributes is represented by the joint propensity distribution for the attributes. These two sets of equations, if combined, result in a (specific) joint propensity distribution for the attributes of the (final) outcome of the process.

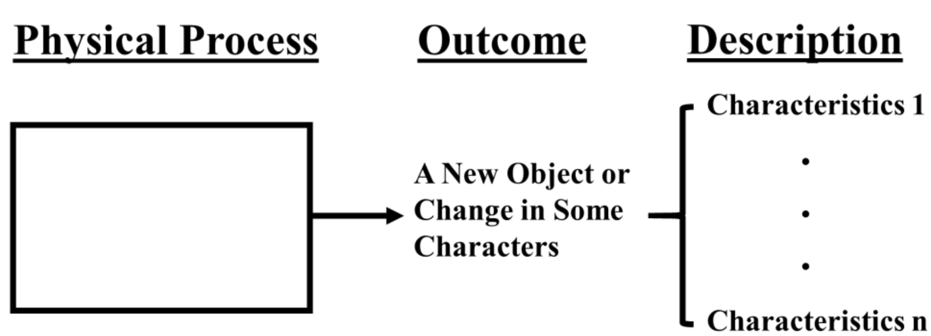

Figure 2.9. The multi-faceted nature of physical objects

An outcome of a physical process, as well as other components of the process including its trigger and circumstances, can be described by a number of aspects/ characteristics, where each aspect may take any of the possible states/ conditions. The (mechanics) of the process can be represented by the relationship between the components of the process, summarized in a (set of postulated/ hypothesized) relationships between the characteristics describing these components. When the states, at which the initial and circumstantial aspects are, remain unchanged in subsequent runs of the process, it is expected that the states of the aspects characterizing the outcome also remain static. On the contrary, fluctuation in certain initial and circumstantial aspects of the process results in the variation in particular characteristics of the outcomes of the process. When the mechanics of the process and/or the pattern governing the fluctuations (in subsequent runs of the process) is unknown, the process is called a “random process.”

The outcomes of a physical process can have different characteristics, where each (character) can take a number of states.

The real-world outcomes of a physical process can have different characteristics, where each (character) can take a number of states. If the realization of the various states of a character in the products of the subsequent runs of the process are unpredictable (in advance), given the existing level of knowledge, that aspect/ character of the process is random/ stochastic and the process can be called a random/ stochastic process. Describing (and labeling) the process as random only corresponds to the characters, whose states we are unable to predict.

2.6.1 Quantitative Models of Physical Processes

Material objects of the world can be characterized by the state of their various attributes and both, the objects and their attributes, are the outcomes of the physical process the objects have gone through. A physical process, in its general forms, can be described by the relationship governing the attributes characterizing the (the initial, the circumstantial, and the final) components of the process. A change in the process occurs when one (or more) aspect(s) of the (components of the) process shifts its current state to another possible state. While a change in some aspects of the initial and circumstantial components may influence some aspects of the final outcome, it may not influence other aspects.

2.7 SUMMARY AND CONCLUSIONS

2.7.1 Brief Summary

Philosophical Processualism is an original perspective advanced as the foundation for an epistemic machinery called *pragmatic mathematics*; a system of linguistically manifested mental constructs aimed at guiding purposive actions with precision. Philosophical Processualism relies upon a process-centered interpretation of causality, which sees an event as a constellation of changes made by a process in certain states of the world and/or the mind. By taking concept as the fundamental constituent of the purposive agents' evaluation processes, the view presents its ontological account of concepts and elaborate on how concepts could look like, where they could be present, and how they could come to realization. Philosophical Processualism opposes Platonism by asserting that every concept is the outcomes of cognitive processes unfolding in time and space and is not an abstract entity in the so-called third realm, to which mind can gain access through unknown metaphysical processes.

2.7.2 Conclusions

The following conclusions can be made based on the contents of this chapter:

- *Philosophical Processualism* is an original perspective advanced as the foundation for *pragmatic mathematics*; a system of linguistically manifested mental constructs aimed at guiding purposive actions with precision.
- A process is a set of mechanisms that cause changes in a set of attributes of the objects involved.
- If a machinery is intended to be copied and different copies are expected to yield consistent output, the implanted processes ought to be known and to be imitable.
- Theorizing about cognitive mechanisms partly relies its external manifestations, as cognition and the resulting sense of meaningfulness occur inside the mind and are hidden from direct sensory observations.
- The meaningfulness of a sensory experience does not come from the work of the sensors, but mostly from the work of the modelers, who do not only combine or rearrange sensory input, but also add things not present in input.

- The meaning has more to do with the output of (internally postulated) models than (externally received) data.
- An Interpretation is the outcomes of a model.
- Every concept is the outcomes of cognitive processes unfolding in time and space and is not an abstract entity in the so-called third realm, to which mind can gain access through unknown metaphysical processes.
- Investigation of universals from an observational view yields a nominalists account, where a universal is seen as merely a name for the set of instances, who all share a common feature.
- Each universal can be explained by a processual code containing the information on the specifics of objects and their properties, whose presence trigger and advance the process resulting in the exemplification of the universal phenomena.
- Every exemplification of a universal property is the result of a single implementation of the pertinent processual code.
- A processual tower is a visualization of a series of processes stack up to eventually produce a target attribute that is perceivable to us.
- The code of a process is embedded in the material participants of the process and in the form of their physicochemical arrangements. Once all objects of certain characters take the appropriate spaciotemporal position relative to each other, the process commences according to the specifics of the objects' physicochemical arrangements.

- Processual inquiries are potentially endless, as every observable event is the output of a (macro) process, whose input may come to be the output of another (micro) process, and so on.
- The complete knowledge of the code for the macro process along with the ability to sense its input allows for the precise prediction of the observable event and makes the consideration of micro processes unnecessary for prediction purposes.
- While the processual knowledge of the micro processes may not be of predictive value, it may facilitate processual manipulation such that a process delivers a desirable event.
- The Processual inquiry can be visually depicted as going down a tower of physical processes, in which a series of processes stack up to eventually produce the observable attributes. Also, complete processual knowledge above a processual level, can be depicted as a tight grip around the tower, from that level above.
- The extensional definition of a type, as the set of its tokens, is imprecise. While the finite set of observed tokens lacks sufficiency and can only be accepted as the definition for a sample (of tokens), the infinite set containing yet-to-be-generated tokens lacks accessibility due to the inclusion of entities without spaciotemporal specification.
- Processual ontology of universals can be extended to elucidate type-token distinction, where a universal property can be taken as a type and its

realized instances as tokens. In other words, the type is being defined as a typical code.

- A philosophical theory of types offers little about the specifics of certain processual codes, and the details can only be discovered through contextual scientific inquiries.
- The availability of many detailed codes, at any point in time, is at best partial.
- By applying the concept of processual code to the domain of the mind, cognitive universals and their cognitive exemplifications are being defined.
- A process-centered perspective can demonstrate the consistency of the mentalism of concepts with the realism of universal physical phenomena.
- Perception is a two-world serial process, the first one in the physical world outside human mind and the second inside human mind; thus, a percept's processual code is the composition of two codes, one for the tower of physical processes and the other for the perception process.
- The multi-world account of universal percepts is in contrast with the positions of both Aristotelian realists, who believe that types are out there in the world independent of our minds, and conceptualists, who believed that universals do not exist outside of the perceiving mind.

- The belief in the existence of every universal relies on the implicit assumption of the uniformity of its two non-mental and mental typical codes.
- The universality of percepts among humans, must not to be considered a natural law, but a physical hypothesis, whose validity ultimately depends on the (potential) existence and the specifics of the differences among percepts formed in the minds of humans.
- The divergence of positions among philosophical worldviews, may be explained by the differences in the cognitive processes they implicitly assume and rely upon.
- Nominalists' denial of universals is possibly not due to their recognition of the "experiencing-organizing mind," but due to their lack of recognition of, "questioning-hypothesizing" mind
- By remaining at the level of observations, nominalists ignore the possibility for the existence of entities beyond or beneath the perceived appearances.
- Any hypothesis involving unobserved entities, is likely to generate a dispute, whose settlement would only be possible either by observing entities (or their associates), or by taking them as primitive entities.
- The tempting idea of making scientific proposals undisputable by removing all non-observed or non-postulated entities is elusive; not only because of the limitations of resources for running such a huge project, but

due to losing the potential contribution of hypothetical entities to the advancement of human knowledge.

- The orientation toward (a) human mind, and/or (b) the apparent world is also causing philosophical perspectives to diverge. (a) can be described as outward, inward, or both, and (b) can be described as within, beneath, and beyond.
- Most prominent perspectives including Aristotelian realism, Platonist realism, and even nominalism, embrace a mentally-outward orientation, and only conceptualists' position is mentally-inward. Philosophical Processualism is in-between, as it does not focus on physical or mental worlds at the expense of the other, and sees them jointly from a more comprehensive perspective.
- A language adds a third world of linguistic expressions to the two worlds of physical entities outside human mind and of cognitive entities inside human mind, whose entities could be related by directional acts of translation.
- Conventional acquaintance is a necessary condition for the translation of cognitive contents to and from linguistic expressions.
- The proliferating role of language in development of conceptual systems are attributed to (1) the association of labels rather than the concepts themselves, and to (2) the application of linguistic association rules that are partially content-independent. The gained efficiency is being

analogically depicted by the ease of moving a set of handled containers, in comparison to those without handles.

- The meaningfulness of a mixed system composed of perceivable and imaginary entities is not being influenced by its degree of perceivability, but by mind's ability to associate every entity, either perceivable or imaginary, with a mental content.
- In a communal verification-based account of conceptual objectivity, a conceived entity by members of a community is considered to be objective, if the subjects consistently deliver similar manifestations of the concept in any number of agreed-upon verification processes.
- Processual assimilation can be pursued by capitalization on the learning abilities of the subjects and by providing them with instructions on how to create and/or with feedbacks on how to modify their concepts.
- The significance of an imaginary entity is not due to its name, but due (a) to its properties, (b) to its relations with other entities and/or their properties, and (c) to its participation in certain processes.
- The common nominal misinterpretation of numbers may be due to dual nature of numerals, as manifestations of both the name and the value property of numbers, whose relations give a systemic meaning to numbers.
- An imaginary entity may have a meaning beyond what is bestowed by the systems it belongs to, and that imaginary systems are not semantically self-reliant.

- Imaginary properties, relationships, and processes are all “grounded,” and are either abstracted from or analogous to perceivable realities.
- The meaning of a natural number, is tied to the meaning of the three concepts of (a) the number with value property one, (b) the equality (of an apparent feature), and (c) the repetition (of the objects with the same feature). The grounds for each of the three concepts can be traced back to a certain set of perceptual cognitive processes.
- While the concept of finite set is grounded in the limits of our perceptual abilities and opportunities, including our attentional resources, the concept of infinite set is rooted in the limitless ability of typical codes in generating tokens.
- An infinite set can be interpreted as the collection of all entities that can be potentially generated by a certain processual code.
- In processual interpretation of the concept of prediction, whenever the code for the physical process is being translated into one of our imaginary languages, running a mental simulation of the code yields a prediction.

Chapter 3 Principles for Objective Assignment of Credibilities

3.1 INTRODUCTION

In the following, I pose the philosophical challenge I intend to tackle, namely interpretational/perspectival reconciliation regarding the concept of probability. I will then elaborate on the merits of my inquiry and present a schema of the fruits of the investigation.

3.1.1 Problem Statement

The fundamental philosophical challenge about the concept of probability regards its ontology; i.e. answering the questions what probability "is," where does it exist, and how to define it. The answers to these questions are co-related such that a specific answer to what questions is expected to guide the answer to when and how questions. Each set of answers, can be labeled as an "interpretation" of probability.

Probability, as historical investigations reveal, e.g. (Fienberg, 2006a), has had a dual nature in the mind of pioneers, who were first to conceive the notion. From one side, probability was thought as having an epistemic nature, representing the uncertainty man does have regarding the reality of (the reference of) a linguistic construct, whether stated in human common language or in a more abstract language such as formal logic. Probability, as an epistemic measure, gives the degree by which man has *knowledge*, *episteme* in Greek, about truth/falsity of the statement, assuming man knows how truth can be eventually revealed. From the other side, probability was thought as having a physical nature, giving a ratio that can be calculated based on the measurements made in the physical/phenomenal world.

The categorization of probability interpretations into epistemological and physical, is intended to characterize all the varieties of probability interpretations conceived by many intellectuals, based on their most fundamental differentiator. In reality, within the past three hundred years, probability as interpreted by each scholar who has focused on the topic, has differences relative to the notion conceived by the others. For example, the British statistician Irving Good (1972) titled one of his papers “46,656 varieties of Bayesians,” to indicate the vast differences between the views of scholars whose work rely on a specific kind of probabilities, namely Bayesian probabilities. Nevertheless, the twofold classification of probability interpretation, seems to succeed in highlighting the fundamental feature that distinguishes conceptions, all other differences aside.

The duality of probability interpretation, seems to be the main reason for probability to be a contentious notion, giving rise to many different treatises, arguments, and debates on the nature of probabilities. These vast differences, as assessed by modern philosophers of science, e.g. (Galavotti, 2005), seems irreconcilable, and has resulted in compartmentalization of acceptance and the use of each interpretation given the field of study, e.g. (Gillies, 2000). The compartmentalization not only has resulted in the technical developments of each field wholly relying on a single interpretation, but also in the vehement rejection of the other interpretation by some scholars of each field, who are most ideologically invested in the interpretation founding their technical work. The battle is not unlike of those apologists embarking on an intellectual crusade to prove the righteousness of the religious believes they are advocating.

The most fundamental challenge regarding probability interpretations regards their reconciliation. While current assessments testify on irreconcilability of the two notions, one may question the assessment and ask whether the two interpretational perspectives could become compatible, and if so, how? Answering reconcilability questions, may become easier by answering the following simpler questions. Could the interpretational variations only be nominal and caused by varying labels? If not, and the differences are substantial rather than nominal, are there also substantial commonalities between the interpretations? If so, how the interpretations could be associated with each other? Could these associations indicate a reference to a bigger whole; the elephant in the room?

3.1.2 Significance

My inquiry into the philosophical foundations of probability is expected to make advancements along the two directions of cognition and application. In terms of cognition, by pursuing the goal of interpretational/perspectival reconciliation, it is expected to provide a more coherent interpretational system for the concept of probability. In terms of application, by pursuing the goal of expanding the application of epistemic probability, it is expected to close the application divide among practitioners of different fields of study. Currently, the application of each interpretation of probability is limited to certain fields of study. The physical interpretation of probability is more of the interest of physical sciences, where the goal is to inquire about community-verifiable characterization of the world. The epistemological interpretation of probability is more of the interest of social sciences, especially economics, where the goal is to inquire about the behavior and the evaluation of subjects.

3.1.3 Objectives

The philosophical development I am putting forward pursues two objectives:

(1) Providing a wholistic view comprising of different interpretations of probability. This view is founded on a process-centered notion of causality and insists on procedural transparency in evaluating probabilities.

(2) Introducing a fundamental character common among all interpretation of probability. The character, I claim, is the more-to-less nature of relationships, and I try to propose ways to account for various such relationships.

3.2 Observation, Theoretical Reasoning, and Learning Unknowns

3.2.1 Possible Approaches to Learn Unknowns

There are things that we do know and there are (many) other things that we do not know. Learning is the process through which, by observing specific aspects of the reality, condensed in data, and by using our (human) reasoning faculties, we manage to add more to what we already knew.

Two central questions regarding learning are determining how we can use what we already know to inform what we currently do not know, as well as determining the (relative) weights we have to give to evidence and reasoning in order to conclude a new learning. To answer the aforementioned questions, we can adopt any of the following approaches.

3.2.1.1 Observation

Observation is the act/ process of sensing a specific character/ aspect of the material world and being able to distinguish the state at which the character is. Thus, carrying an observation out requires two sets of means; (1) a set of means for sensing the (target) character and (2) a set of means to interpret the sensed character by mapping it to a state, among all states that character can take. For example, to observe the color (character) of an object, a human needs a set of means to receive the sensory signals pertinent to the color, as well as a set of means to interpret received signals and to map it to a (mental) representation for the state of the character. If the light emitted from an object has a wavelength in the range 620 to 740 nanometers, a specific representation (of the color) is being formed inside human mind. This mental representation, in the spoken language English, is labeled “red.” In the example, the color is a material character/ aspect of the world and specific colors, e.g. red, are the (possible) states for that character.

The observational approach, it seems, is the primary mode of learning among biological agent, whereby the agents uses its (sensory) observations to characterize the outside world. Humans, along the time, have succeeded in taking the approach beyond the levels other biological agents practice it, in terms of (1) the scope, (2) the precision, and (3) the communication of the observations.

In terms of the extent, humans have managed to build a variety of devices that can sense/observe aspects of the real-world, which are not sensible by human (built-in) sensory means; thus hugely expand the scope of experimental learning. In terms of the precision, humans have managed to standardize the characterization of specific aspects of the world such that every human can have the same understanding of the states at which a (standardized) characteristics is. Attributes such as length, mass, and time are prime examples of aspect, for which humans have managed to establish standards to precisely map a specific state of the character to an indicator, (mostly) a number. Standardization of observable characters are examples of the application of applied reasoning to mathematical modeling.

In terms of communication, humans have managed to develop new means, other than means available to biological agents, for recording and/or communicating their observations to other fellow humans. These means can generally be called languages, of which spoken and written languages are certain sub-classes. For example, a numerical indicator is a description comprehendible for the agents who know pertinent parts of elementary mathematical langue. The advent of precise languages, combined with the precise (standardized) characterization, enables humans to record and inform the state of certain (worldly) character in a system in an “observationally objective” fashion.

In the (pure) observational/experimental approach to learning, we assert that until we make new observations, we take the set of our past observations, as of today, as the (sole) representation of the reality. The approach relies on a philosophy whereby (empirical) evidence is the only source of knowledge and results in (implicitly) ignoring unknowns, as they are yet to be observed and, at the moment, the only way to (humanly) perceive their existence is to hypothesize them by the aid of human faculties, other than his observational faculties. Thus, the only way to learn new things is to wait until we observe them.

The observational approach to acquiring (new) knowledge is the method used in experimental science. An (epistemological) philosophy that is obsessed with the observational mode of learning is called empiricism, in which experience is the (only) source of new knowledge. In other words, whatever that is not being observed is not considered as knowledge.

3.2.1.2 Theoretical Reasoning

Reasoning is the mental process humans consciously use to expand their comprehension. Reasoning, it seems, is exclusive to humans and is one of the factors that has caused humans to accelerate differentiating themselves from other biological agents and to ascend to their superior position on the planet earth. A number of definitions in Appendix A.1, provide with further details of my views on various aspects of reasoning, as a special mode of thinking, and its symbiotic relation with human civilization.

In reasoning, an agent manipulates a number of mental constructs/ objects, as input, to obtain a number of different constructs, as output. An object of the mind, in its most basic form, can represent (some aspect of) a concrete (worldly) object as perceived by human senses. In contrast, an abstract (mental) construct represents a thing that does not have a physical referent, i.e. it does not refer to an object existing in any specific time and/or space. Purely abstract objects are created in thought processes, where worldly aspects of a simple abstract object is taken away to create a mental object independent of the world (outside human mind). Prime examples of such abstract constructs are mathematical objects, e.g. numbers.

I separate reasoning into the two class of “abstract” and “applied” reasoning. In an abstract reasoning, the mind operates on a set of (pure) abstract objects, arranged in a certain way, in order to find the outcomes of that arrangement, themselves other abstract objects. The process, can be conducted by any set of (self-consistent) rules including the general rules established in symbolic logic, as well as any specific set of rules, called axioms, established for an abstract theory. Mathematics is mainly the result of the application of (pure) abstract reasoning to (arrangements) of mathematical objects. A mathematical arrangement is a structure, in which mathematical objects of different groups/set are associated in a certain way; thus, resulting in characteristics of the arrangement to take certain states. The processes required for finding the resulting characteristics is often referred as “deductive reasoning.”

While mental constructs subject to (pure) abstract reasoning, as designed, have little relationship with the objects of the material world, the aim of applied reasoning is to establish analogies between (1) the concrete (aspects of interest) and certain (pure) abstract objects, and more importantly between (2) the arrangements of the concrete characteristics of interest and the arrangements of abstract objects associated with those concrete objects. The process to establish such relationships is generally called (abstract) “modeling” and in specific, when the mental representations are mathematical objects, it is called “mathematical modeling.” The biggest advantage of, and the motivation for, abstract modeling is that it allows the characteristics of a specific abstract structure to be extended to the concrete associates of pertinent abstract objects.

The first step to mathematical modeling is to associate the (characters of) concrete objects of interest to a set of mathematical objects. When mathematical objects are numbers, or algebraic variables as an extension of numbers, this step only requires the (states of the) concrete characters to be associable/mappable to numbers. While (standardized) observable characters, by their design, meet the requirement, e.g. time and distance, other characters, as long as they remain observable, can be associated with numbers according to any (arbitrary) scheme.

The second step to mathematical modeling is to associate the arrangements between concrete (characters of the) objects of interest to a set of (theoretical) arrangements connecting mathematical objects. In contrast to the first step to mathematical modeling, which could be resolved by adoption of any assignment scheme, as long as it remains unique, the arrangements sought for in the second step are not optional. Adoption of any (theoretical) structure among abstract objects as the model of reality can only be true if observations made on observable characters exhibit the adopted relationship. The processes required for taking this (second) step is often referred as "inductive reasoning." A number of terms defined in Appendix A.1 further elaborate the differences between mathematics and mathematical modeling

The learning through (theoretical) reasoning occurs when both steps of (mathematical) modeling, (1) mapping observable (states of) characteristics to (numerical) variables and (2) adopting/ inducting functional relationships among variables, observables as well as non-observables, have been taken. New knowledge then comes as a logical/ deductive consequence of what we already know, i.e. the adopted functional relationships.

The theoretical approach to learning relies on a philosophy whereby the real-world is deemed to follow the (logical) orders/ structures devised by human minds and takes observation as a (sometimes secondary) support to justify the validity of postulated orders. This mode of learning is mainly pursued in theoretical science including mathematical physics as well as philosophy.

3.2.1.3 Authoritative Inquiry

Following the theoretical approach to learn unknowns, if one believes in the truth of these (imposed) orders, prediction/ characterization of unknowns look logical, reasonable, and natural. By the same token, an easy way to fool oneself into taking a (unobserved) character as real, is to postulate a certain structure on the reality and deduct that character accordingly.

This approach heavily relies on imposing/ postulating (logical) structures/ orders on the reality, where (states of) certain aspects of the world are (deemed) to be in a particular type of relationship. It is such a relationship between knowns and unknowns that makes the later a logical consequence of the former, i.e. by considering our current knowledge, we can deduct the existence and the character of what we are yet to observe.

3.2.2 Challenges for Approaches to Learn Unknowns

Following the second approach to learn unknowns, if one believes in the truth of these (imposed) orders, prediction/ characterization of unknowns looks logical, reasonable, and natural. By the same token, an easy way to fool oneself into taking a (unobserved) character as real, is to postulate a certain structure on the reality and deduct that character accordingly.

3.3 Probability and its Interpretations

The controversies about the concept of probability, is rooted in different interpretations of the notion. While the two broad interpretations can be labeled physical and epistemological, various combinations of interpretation, function, and structure can become the basis for further nominal distinction. Each assigned name, can be perceived as the label for a probability kind/type. The distinguished concepts, in relation to the two broad interpretations, that may be considered as physical, epistemological, or both.

Nominal distinction, in addition to functional differentiation, is also aimed at conceptual precision; an essential means for preventing cognitive confusion. Besides, the distinction could facilitate conceptual and functional association, elaborating the conditions under which, various types may be related, as well as the properties of the relationships. Such associations constitute the building blocks of a wholistic interpretational perspective, which I aim at developing here.

The classification of probabilities into types has a long history and scholars have proposed diverse kinds and have labeled them according to the functionality or the property of the type, or they have used the same label as those of scholars before them. In fact, the act of classifying different types of probability is so central to a philosophy of probability that Good (1959) believes that probability typology addresses half of the problems in philosophy of probability.

I introduce three fundamental kinds of probabilities, namely propensity, possibility, and credibility. The types, given the long history of probability classifications, are expected to be found in the works of past scholars, potentially, each under a variety of names. The reason for emphasizing on these types is the belief that the three have the power to explicate and include all other types of probability, e.g. Frequentist and Bayesian probabilities. In other words, propensity, possibility, and credibility are necessary and sufficient types in the process-oriented philosophy of probability, which I am putting forward.

3.3.1 Probability as a Means of Representation

The introduction of the probability types based on a combination of interpretation, function, and structure, and distinguishing them by distinct labels, raises the question whether we do need the label probability for a purpose other than a reference to a collection of more specific concepts? And if we do, what could be the purpose? My suggestion is that we define a probability solely as a member of a numeric set, whose values are between zero and one and collectively, add up to one.

My definition portraits probability as an abstract mathematical measure, constrained by limitations imposes by a set of axioms such as those suggested by Kolmogorov (1933). The definition positions probability as a means of mathematical representation, which can be used to represent the specifics of the relationship between an object with other objects in any non-mathematical world, e.g. the world of physio-chemical processes. In other words, probability can be used to describe mathematical constructs build to model non-mathematical phenomena. The concepts that make a probability meaningful within the context it is being applied, have labels such as propensity and credibility.

3.3.2 Possibility

The possibility, $\boldsymbol{P_b}(\overrightarrow{\boldsymbol{y}^*})$, of a specific character such as $\overrightarrow{\boldsymbol{y}^*}$, as an outcome of the characterization process taking a character such as $\overrightarrow{\boldsymbol{x}^*}$, specified based on criteria $\boldsymbol{C_i}$, and assigning a (different) character based on criteria $\boldsymbol{C_o}$, is defined as the proportion of the number of states, characterized according to criteria $\boldsymbol{C_i}$, whose character based on criteria $\boldsymbol{C_o}$ is $\overrightarrow{\boldsymbol{y}^*}$, relative to the total number of input characters that the process is defined to take, Eq. (3-1).

$$\boldsymbol{P_b}(\overrightarrow{\boldsymbol{y}^*}) = \frac{\boldsymbol{n}(\vec{\boldsymbol{x}}|\overrightarrow{\boldsymbol{y}^*})}{\boldsymbol{n}(\vec{\boldsymbol{x}})} \tag{3-1}$$

Figure 3.1 presents a visual schema of such characterization process. To further illustrate the definition of possibility, I provide some examples. Consider a characterization process which takes the outcomes of a cartesian production process orderly joining the members of the numeric set $\{\mathbf{0}, \mathbf{1}\}$ with the numeric set $\{\mathbf{2}, \mathbf{3}\}$, and assigning the sum of the values of the two numeric components, as $\boldsymbol{y}$. The possibility of output character being 3 is $\frac{\mathbf{2}}{\mathbf{4}}$, as in two of the four input states $\boldsymbol{y} = \mathbf{3}$. In another example, consider a (re)characterization process which takes the temperature of objects between 50°F to 77°F , and delivers their temperature in Celsius. The possibility that the temperature be between 20℃ to 25℃ is $\frac{\mathbf{1}}{\mathbf{3}}$, as only the last third of the total domain maps to the specified range.

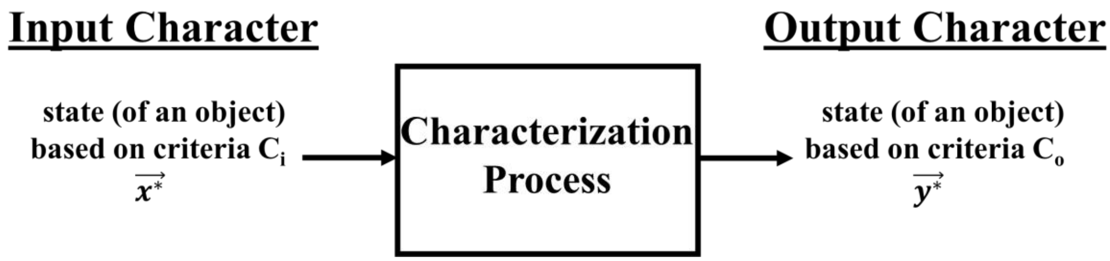

Figure 3.1. a schematic diagram illustrating the process of (re)characterization

While the definition of possibility presented in (3-1) has similarities with the classic definition of probability, there are also differences between the two measures. The most important difference is the directionality of the definition of possibility, which is only defined for the output of the characterization process.

Possibility values for the outputs of a more-to-less process are the relative contributions of inputs; i.e. the relative number of inputs associated to every output. Possibilities are named such as they only depend on the processual code, i.e. the specifics of input-output relation, and are independent of empirical observations. Possibilities may also be called logical or structural probabilities, contrasting propensities, which may be called physical probabilities or relative frequencies. Propensities for the outcomes of a process are due to the combined effects of (a) the propensities for corresponding conditions, and (b) the possibilities imposed by the specifics of condition-outcome relation.

3.3.3 Propensity

Propensity of an outcome of a physical process gives the preferential tendency of the process to generate the outcome relative to other outcomes of the process. Since the quantity characterizes a physical process, its measurement requires empirical observations.

The way to measure such tendency is to observe the outcomes of numerous runs of the process and to find the proportion of the runs resulting in the target outcome. The number of runs that is sufficient for propensity measurements is the number beyond which the quotient is less sensitive to variations in the observations. Eq. (3-2) presents the quotient, resulting from dividing two observed values, where $\boldsymbol{n}(\overrightarrow{\boldsymbol{y}^*})$ is the number of outcomes with the target character, $\overrightarrow{\boldsymbol{y}^*}$, and $\boldsymbol{n}(\overrightarrow{\boldsymbol{y}})$ is the total number of outcomes generated by the outcome through the observation.

$$\boldsymbol{P_p}(\overrightarrow{\boldsymbol{y}^*}) = \frac{\boldsymbol{n}(\overrightarrow{\boldsymbol{y}^*})}{\boldsymbol{n}(\overrightarrow{\boldsymbol{y}})} \tag{3-2}$$

Figure 3.1 presents a schematic diagram illustrating the concept of propensity of an outcome of the process. As it can be seen, the propensity of a process to generate a specific outcome is related to the propensity of the realization of the input characters, $\boldsymbol{P_p}(\overrightarrow{\boldsymbol{x}^*})$ and $\boldsymbol{P_p}(\overrightarrow{\boldsymbol{e}^*})$, as well as the possibility of the outcome to be generated by the mathematical representation of the process.

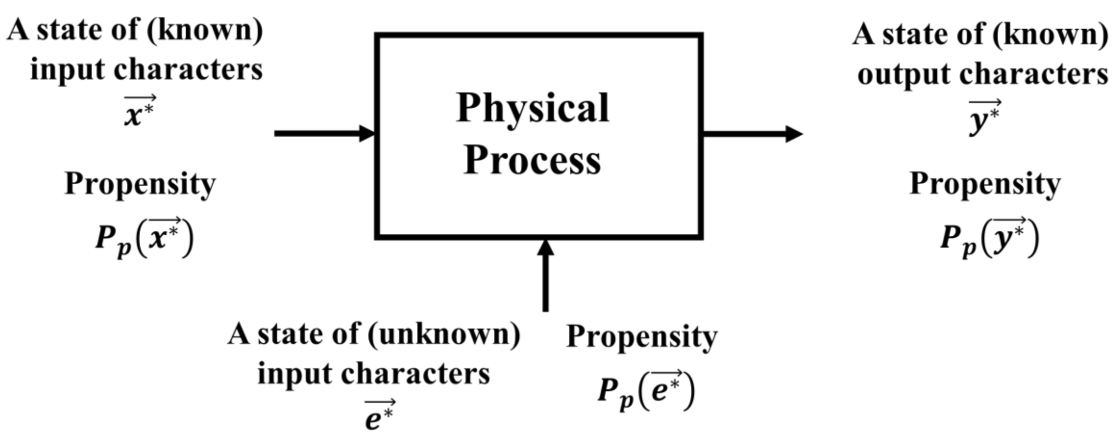

Figure 3.2. a schematic diagram illustrating the concept of propensity within the context of a physical process

3.3.4 Credibility

Credibility is the degree an evaluating agent assigns to its evaluation. Figure 3.3 presents a schematic diagram illustrating the two types of credibility evaluation processes, where the top diagram depicts a propensity evaluation process, and the bottom a possibility evaluation process.

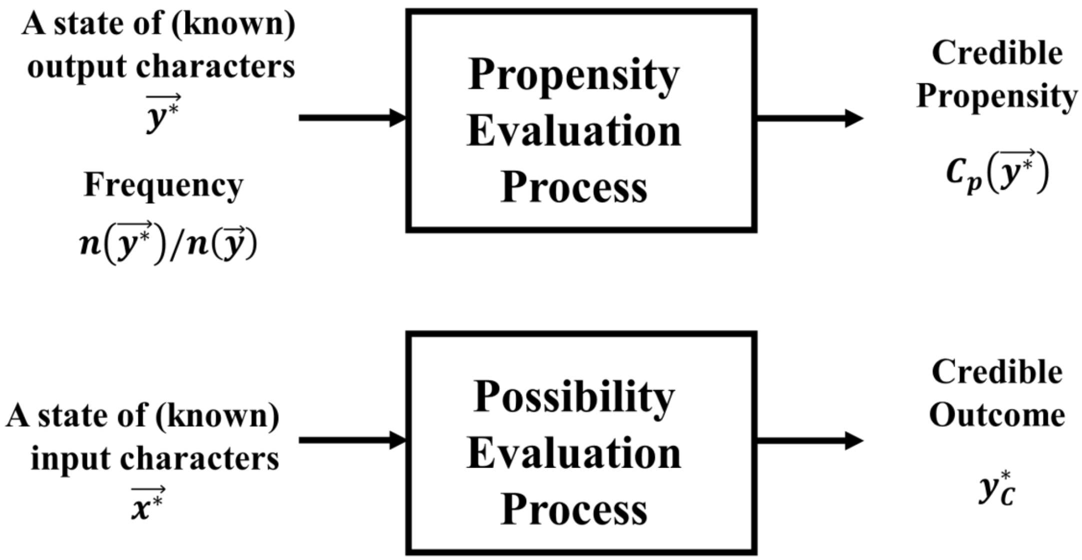

Figure 3.3. a schematic diagram illustrating the two types of credibility evaluation processes; top: propensity evaluation process, down: possibility evaluation process

Credibilities are imaginary relative weights assigned to (the parameters of) the processes, who are hypothesized to deliver the intended outcomes. The parameters of a process are the (structural) values that specify the propensity values for the input or the output and/or the input-output relation, which determine the possibilities. Credibilities are only means to the end of assessing propensities and/or possibilities, whose exact values are unknown. The motivation for the use of credibilities are the desire to account for a more comprehensive set of possible processes, who might be otherwise ignored in propensity/possibility assessments, due to the lower level of support provided by available knowledge (for them). Bayesian priors are prime examples of credibilities.

3.4 Evaluating Probabilities

Evaluating probabilities is centered around using credibilities to evaluate propensities and possibilities.

Since credibilities are imaginary, their values cannot be obtained from empirical observations and must be assigned without resorting to empirical justifications. While the imaginary nature of credibilities allows subjects to assign credibilies of their own preference, their assignment might not be justifiable to others. Alternatively, it is possible to establish a set of theoretical assignment rules on purely logical grounds, and to justify the designations based on their effects, or lack thereof, on the assessed propensities and/or possibilities. Due to their rule abidance, theoretically constructed credibilities may be described as objective, even though imaginary and only existing in subjects' minds. The class of credibilities called non-informative are the ones, whose assignment are aimed at not informing (certain aspects of) the assessed propensities and/or possibilities. Decision Entropy is one such theory, whose aim is not to inform the assessments made within the framework of analyzing decisions. The principles of the theory precisely specify the assessments, regarding which the credibility assignments must remain neutral.

3.4.1 Learning; Conceptual and Perceptual

My conjecture is that learning (of an agent) can be categorized into two classes of perceptual and conceptual. While perceptual learning solely relies on empirical observations, conceptual learning involves both empirical observation and conceptualization of structural relationships between neighboring states. Figure 3.4 presents a schematic diagram, which can assist in illustrating the two types of learning. The perceptual learning involves making observations on concurrent values of $\vec{x}$ and $\vec{y}$, whose results can be tabulated. However, the conceptual learning requires going beyond observations and imagining how a specific change in stimuli may result in a change in response.

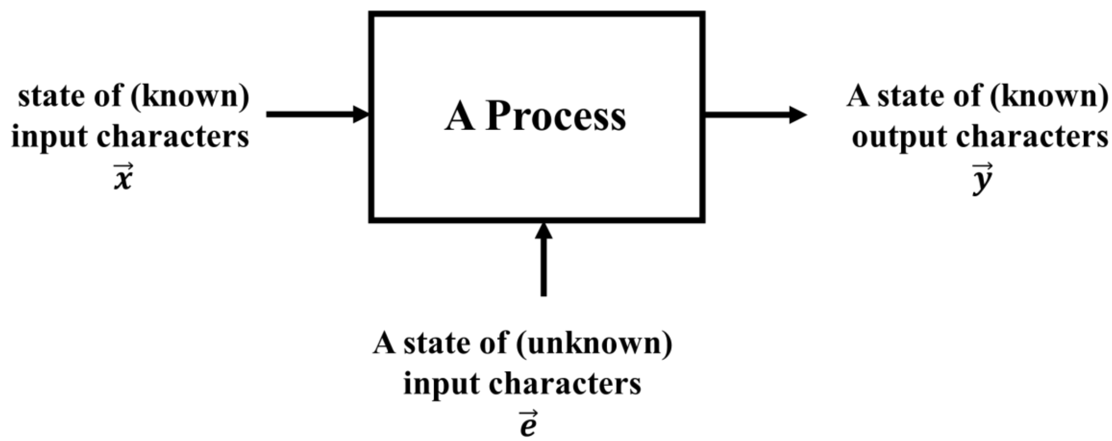

Figure 3.4. a schematic diagram illustrating the two types of learning

3.5 BAYESIAN ROBOT

3.5.1 Historical Background

The idea of an inferential robot/computer program, which could do Bayesian inference by itself, was first put forward by Jaynes (1957) and then refined by himself and his followers along the time, e.g. (Tribus, 1969) and (Jaynes, 2003). The robot receives the information in the possession of subjects as an input and gives its Bayesian estimate as an output, based on the algorithm embedded in its program. The algorithm is written based on a set of desiderata developed to ensure the rationality and the consistency of the inference. An example of such desiderata for plausible/probable inference was proposed by the Cox (1946 and 1961).

A consequence of the development of an objective approach to statistical induction is that the procedure developed has the potential to be automated and to be run with minimal involvement of subjects. Although such objective inference relies on the information elicited from subjects, the process of assigning probabilities is done with no on-the-fly guidance from the subject. This is sharp contrast with the way most humans form their probability judgements when questioned about uncertain quantities (O'Hagan et al, 2006). We believe the idea of robot is worthy to be included in the foundations of objective Bayesian approaches and we will do so. In addition to existing developments regarding the robot, which mainly focuses on the role of information (Tofolli, 2004), we will also emphasize on the role of the assumptions required for making an induction.

3.5.2 Evaluating Robots; Type A vs Type B

Robot/machine A is a complex robot with the ability to both perceive the external signals, as well as imagining variety of relationships between them. However, the mechanics of the evaluation of the robot is murky, and as a result, cannot be taught/transferred to other fellow machines of the same type. In the following we aim at elaborating on the potential mechanisms, through which this robot may learn.

Robot/machine B, is a simple robot with as good as of an ability to perceive the external signals, relative to robot A, but a simpler imagining apparatus, mainly consisting of mechanically constructed model available to its memory. Although simple, evaluation processes used by robot B has the opposite character of robot A, i.e. transparency. This allows the robot to transfer the mechanics of its evaluation to other fellow machines of the same type. In the following we aim at elaborating on the potential mechanisms, through which this robot may learn.

3.6 PRINCIPLES OF RATIONAL INDUCTION

In the following, I introduce a number of principles, which I deem essential for a rational, consistent, and transparent inductive reasoning. Together, the principles form a flexible framework that can act as the conceptual foundation of any axiomatic system to be developed upon the framework.

3.6.1 Evaluative Orientation Principle

The orientation principle can be stated as follows:

- Inductive analysis is a goal-oriented act and must be directed toward answering a set of target questions

A goal-oriented act involves triggering and sustaining a physical process to obtain an outcome of certain characteristics. However, the focus of a study aimed at assisting a goal-oriented act is on the (set of) characteristics of the outcomes (of the process) that are of interest to the purposive agent.

3.6.2 Investigative Prioritization Principle

The prioritization principle can be stated as follows:

- Target questions must be prioritized in order to direct the inductive analysis to approach questions according to their importance

The principle insists on theoretical and pragmatical necessity of ordering the variables of interest. The ordering becomes important as the unbiased inductive approach toward variables higher in the hierarchy inevitable results in informing the variables lower in the hierarchy, given the structural/deterministic relationships between the higher and lower variables.

3.6.3 Explicative Sufficiency Principle

The explanation principle can be stated as follows:

- A (limited) number of distinct explanations/ hypotheses must be formulated such that each provide definitive answers to target questions

The explanation principle is aimed at setting the determinacy of the mathematical system representing the problem, such that the system does not become indeterminate.

3.6.4 Evaluative Inclusion Principle

The inclusion principle can be stated as follows:

- When data is limited such that an explanation/hypothesis does not stand out, assessments based on (the validity of) different hypotheses must be combined in order to form an inclusive/ collective assessment.

The inclusion principles is a form of adaptation to demands from the inductive investigation. When the objectives of inquiry demands the singling out of a possible answer, a means must be devised to distinguish such answer from others.

3.6.5 Credibility Conception Principle

The credibility conception principle can be stated as follows:

- To form an inclusive assessment, each explanation/ hypothesis must be differentiated with a degree of credibility/ plausibility relative to other explanations.

Credibility is a theoretical construct devised to compensate for the lack of information. Since credibility does not exist in the real world, (1) there is no true value, for a credibility (to be estimated), and (2) no amounts of information, in itself, can give its value. Credibility value has to be established given an (arbitrary) definition.

3.6.6 Artifact-Reality Division Principle

The artifact-reality division principle can be stated as follows:

- Credibility must be defined to explicitly account for (1) what we know, observations, and (2) what we arbitrarily devise to reason beyond what we know, explanation/ hypothesis, such that the contribution of (1) and (2) in the collective assessment be tractable.

The principle aims at providing methodical transparency and justifiability, such that each step of the evaluation be transparent and justifiable.

3.6.7 Unbiased Evaluation Principle

The unbiased evaluation principle can be stated as follows:

- Credibility values must not be established such that the inclusive assessment favors any of the possible answers to the target questions before incorporation of observations.

The principle aims at highlighting the fact that only observation is allowed to tip the outcome of the inductive evaluation toward any of the potential answers. In other words, before incorporation of the observations, it is rational not to distribute credence to any answer, more than the other.

3.6.8 Applications of the Principles

The following is a list of the application of the aforementioned seven principles within the context of a decision, where the following variables are paramount.

- Right Decision
- Consequence of an Action (Prospect of an Act)
- Consequence of an Exploration (Prospect of Learning)

Inductive analysis will be conducted such that the evaluation does not favor possible values of the objective variable, before inclusion of the observations.

3.7 Summary and Conclusions

3.7.1 Brief Summary

The concept of probability from a process-centered perspective is being divided into three types of propensity, possibility, and credibility. Credibilities are imaginary relative weights assigned to (the parameters of) the processes, who are hypothesized to deliver the intended outcomes. Since credibilities are imaginary, their values cannot be obtained from empirical observations and must be assigned without resorting to empirical justifications. The class of theoretically-established credibilities called non-informative are the ones, whose assignment are aimed at not informing (certain aspects of) the assessed propensities and/or possibilities. Decision Entropy is one such theory, whose aim is not to inform the assessments made within the framework of analyzing decisions. The principles of the theory precisely specify the assessments, regarding which the credibility assignments must remain neutral.

3.7.2 Conclusions

The following conclusions can be made based on the contents of this chapter:

- Unsettled conceptual debates on the nature of probability are due to the agglomeration of various quantitative measures, who all share a number of mathematical properties, but do have different natures and functionalities.
- From a process-centered perspective, probability can be divided into three types of propensity, possibility, and credibility.
- Propensity values are relative repetitions of processual outcomes, as they have come to realization.
- Possibility values for the outputs of a more-to-less process are the relative contributions of inputs; i.e. the relative number of inputs associated to every output.

- Possibilities are named such as they only depend on the processual code, i.e. the specifics of input-output relation, and are independent of empirical observations.
- Possibilities may also be called logical or structural probabilities, contrasting propensities, which may be called physical probabilities or relative frequencies.
- Propensities for the outcomes of a process are due to the combined effects of (a) the propensities for corresponding conditions, and (b) the possibilities imposed by the specifics of condition-outcome relation.
- Credibilities are imaginary relative weights assigned to (the parameters of) the processes, who are hypothesized to deliver the intended outcomes. The parameters of a process are the (structural) values that specify the propensity values for the input or the output and/or the input-output relation, which determine the possibilities.
- Credibilities are only means to the end of assessing propensities and/or possibilities, whose exact values are unknown.
- The motivation for the use of credibilities are the desire to account for a more comprehensive set of possible processes, who might be otherwise ignored in propensity/possibility assessments, due to the lower level of support provided by available knowledge (for them).
- Bayesian priors are prime examples of credibilities.

- Since credibilities are imaginary, their values cannot be obtained from empirical observations and must be assigned without resorting to empirical justifications.
- While the imaginary nature of credibilities allows subjects to assign credibilies of their own preference, their assignment might not be justifiable to others.
- It is possible to establish a set of theoretical assignment rules on purely logical grounds, and to justify the designations based on their effects, or lack thereof, on the assessed propensities and/or possibilities.
- Theoretically constructed credibilities may be described as objective due to their rule abidance, even though they remain imaginary and only existing in subjects' minds.
- The class of credibilities called non-informative are the ones, whose assignment are aimed at not informing (certain aspects of) the assessed propensities and/or possibilities.
- Decision Entropy is a non-informative theory, whose aim is not to inform the assessments made within the framework of analyzing decisions. The principles of the theory precisely specify the assessments, regarding which the credibility assignments must remain neutral.

Chapter 4 Theoretical Decision-Based Assignment of Bayesian Probabilities

4.1 INTRODUCTION

In the following, I introduce a number of concepts and methods central to the development of a theory aimed at objective assessment of uncertainties through the assignment of non-informative Bayesian priors.

4.1.1 Black Swan Outcomes

Black Swan is a metaphorical character label for rare but consequential processual outcomes, whose occurrence takes humans by surprise as they were deemed impossible. The three attributes of a Black Swan Outcome (BSO), according to Taleb (2007), are (a) extremely low frequency, (b) extremely high impact, and (c) retrospective predictability by humans, where combined extremity of the first two attributes make the BSO unpredictable but significant. While lying beyond the range of common experiences make rare events unforeseeable, the ensuing unpreparedness makes their occurrence of a BSO much more impactful that what it would, if humans had not failed to predict the outcome. Taleb (2007) claims that while history is directed and shaped by extreme outliers much more than by frequent but expected outcomes, their historical significances are toned down due to the human tendency to explain BSOs after their realizations and incorporate their records in our predictive machinery, as if BSOs were both destined and expected to occur.

Among many BSOs throughout history of the world, a number of contemporary extreme outliers with significant impacts on society, economy, and politics in the United States include (a) deliberate plane crashes into the twin towers of World Trade Center of New York city in Sep 11 2001, (b) the early 2000's stock market devaluations, collectively known as (the burst of) dot-com bubble, resulting from excessive speculative trading of the stocks of companies who were offering internet-related products and services, and (c) the global financial crisis of 2007-8 due to bursting of housing bubble in the U.S.

A more enlightening contemporary BSO in the U.S., though less significant in terms of its socio-economic impact, was the 1998 collapse of the hedge fund named Long-Term Capital Management (LTCM), which the winners of 1997 Nobel prize in Economics, Robert Merton and Myron Scholes, had helped to create. The firm was trading financial products called derivatives based on the assessed prices calculated by the mathematical model named after the Nobel laureates Black-Scholes-Merton; a model that its developers claimed to had reduced, or even eliminated, risk from financial markets by calculating risks of losses with decimal place precision. After four years of astonishing returns up to 40%, in 1997 LTCM started to make losses due to the effects of Asian financial crisis. A month after 1998 default of Russian government on its bonds, LTCM lost about half of its $4 billion capital; an outcome whose risk was assessed to be (almost) zero by the same model that had brought the Nobel prize for the laureates less than a year before. LTCM, whose founders were not considering it a hedge fund but a financial technology company, was eventually bailed out by the US Federal Reserve to prevent the spread of its collapse across the whole financial market. The arrogant intellectuals who had claimed to tamed the risk, were financially and reputationally destroyed by a BSO, deemed improbable by their own model!

The prominence of the term Black Swan in the minds of western intellectuals is partly due to the dramatic change in its modern application, compared to its ancient uses. Since every observed swan in the old world was white, the assertion “all swans are white” was frequently used in illustrative examples of categorial syllogism, to the extent that even Aristotle took the color property of white as the necessary condition for swans; i.e. the color of a swan’s feathers must be white. Later, the term black swan came to become widely cited in ancient philosophical discussions of the concept of impossibility. Nevertheless, the observation of black-feathered swans in the newly-discovered continent of Australia in the 17^{th} century, caused the term black swan to refer to a logical fallacy resulted from the use of empirical observations to form a general rule. Since Scottish philosopher David Hume (1748) had questioned the rationality of inferring such rules based on uniform past experiences, later philosophers including John Stuart Mill (1843) used the term black swan within the context of philosophical justification of Induction. In recent times, the term was revived by Levantine scholar Nicholas Nasim Taleb (2007), and in reference to the realization of a specific subset of the outcomes that violate the uniformity of past observations; the seemingly improbable outcomes, whose monstrosity hugely impacts our lives.

4.2 MOTIVATION

In the following, I declare accounting of Black Swan Outcomes as the main practical motivation for the research and propose that theoretically-assigned non-informative Bayesian credibilities can be a means of such accounting. So, I declare the development of a defensible method for the assignment of Bayesian priors as a theoretical motivation for the research.

4.2.1 Accounting for Black Swan Outcomes

The main practical motivation for the theory we are advancing here is to be able to account for yet-to-be-realized Black Swan Outcomes (BSO) in the theoretical assessments of processual uncertainties. The seemingly non-ending surprise of human kind by a wide variety of BSOs throughout history may be explained by the innate tendency to heavily rely on our memorable experiences to form cognitions and actions. Our naïve empiricism seems to makes us either negligent and/or ignorant; negligent as we may not track and accumulate further historical records, and ignorant as we may not account for our knowledge deficiencies whether in terms of its extent or its relevance.

The approach chosen to account for BSO in the assessments of uncertainties must relieve us from our natural negligence and ignorance. While tracking historical records and accumulating past records ensures that no past observation is ignored, there will be a point where all available data is exhausted and the limits of experience is being reached. We must note that the only means to go beyond the experience is to use logical reasoning such that the resulting assessment be defensible. If so, the assessed probabilities must not be taken as empirically-grounded predictions of to-be-observed frequencies, but as rational assessments made with the aim of moderating empirically-based expectations by accounting for unknowns.

4.2.2 Directing Credible Proportional Induction

The theoretical motivation for the research is to develop a defensible method for directing the assignment of non-informative Bayesian credibilities, in the sense described in the previous chapter. Also, the offered method provides with a theoretical means needed to fulfil the practical desire of accounting for yet-to-be-realized Black Swan Outcomes. In other words, when it comes to extreme outcomes beyond the range of observations, the credible proportions assessed by theoretically-established non-informative Bayesian priors reflect the consideration of unobserved possibilities.

Philosophically, the theory we are advancing takes Hume's (1748) position regarding the legitimacy of inductive inference (from observed to unobserved) as a principle, and views the assumption that future will be the same as the past (processual uniformity along the time), as rationally unjustified. Besides, we propose that the so-called uniformity assumption is likely to result in neglecting or ignoring Black Swan Outcomes, which are often non-existent in historical records.

To relax the blinding uniformity assumption, two limiting cases of stationary and evolving processes are conceivable regarding the specifics of the processual knowledge we are lacking. While the parameters of a stationary process remain unchanged along the time, available observations only inform the parameters within the range of historical records and beyond, the (values of the) parameters must be assessed without reliance on observations. In contrast, we may have a process whose past parameters are completely known to us, but we do not know whether the process has started to evolve and to take different parameters after a certain time. The scenario is similar to the case, where we need assess a target process and we have a complete knowledge of an analogous process, whose similarities makes the existence of a relation between processes plausible.

4.3 DECISION ENTROPY THEORY

4.3.1 General Description of the Theory

Decision Entropy Theory (DET) is a computational machinery aimed at assessing unknown aspects of a process in the light of the available information, where sought-for aspects depend on the processual parameters; (structural) values specifying collective or relative individual features of processual input and/or output and/or the (deterministic) input-output relation. The theory is intended to be objective and defensible in the sense described in Chapters 2, and 3, meaning that (a) its assessment procedures are transparent and imitable and (b) the constraints specifying the process are set in compliance with a set of intuitively logical guiding principles.

The theory assigns non-informative Bayesian priors to sets of possibilities, representing the aspects of the process that are unknown to us. The priors are assigned to possibilities by considering a hierarchy of variables and such that prior credibilities minimally inform the variables, whose exact values are the subjects of the inductive inquiry. The adopted measure for the information contained in a distribution is Shannon's entropy, where the minimally informative distribution is the one with the maximum entropy. In the simplest case of the hierarchy composed of a single variable, whose only constraint is its boundaries, the uniform distribution is the one with the maximum entropy.

4.3.2 Principles of the Theory

The following three principles specify the hierarchy of variables, as the targets of the inquiry, and elaborate the distributions that are minimally informative:

1. If no information is available about the probabilities of decision outcomes, then a selected alternative is equally probable to be or not to be the preferred alternative.

2. If no information is available about the probabilities of decision outcomes, then the possible differences in preference between a selected alternative and the preferred alternative are equally probable.

3. If no information is available about the probabilities of decision outcomes, then the possibilities of learning with new information about the selected alternative compared to the preferred alternative are equally probable.

The mathematical formulations of the axioms corresponding the three principles are presented in Eqs. (4-1), (4-2), and (4-3), respectively. A complete description of the theory including its mathematical formulation is being presented in Appendix B.

$$
\begin{aligned}
&Maximize\ H_{rel}\left(Preference\ Outcome \middle|\ A_i\ Selected \cap S_{n_A}\right) \\
&\quad = -P\left(A_i\ Preferred \middle| A_i\ Selected \cap S_{n_A}\right) \\
&\quad \times ln\left[P\left(A_i\ Preferred \middle| A_i\ Selected \cap S_{n_A}\right)\right] \\
&\quad - P\left(A_i\ \overline{Preferred} \middle| A_i\ Selected \cap S_{n_A}\right) \\
&\quad \times ln\left[P\left(A_i\ \overline{Preferred} \middle| A_i\ Selected \cap S_{n_A}\right)\right] - \ln(2) \\
&\quad = -\Big\{P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle| A_i\ Selected \cap S_{n_A}\right] \\
&\quad \times ln\left\{2P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0 \middle| A_i\ Selected \cap S_{n_A}\right]\right\} \\
&\quad + \left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle| A_i\ Selected \cap S_{n_A}\right]\right\} \\
&\quad \times ln\Big\{2\left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle| A_i\ Selected \cap S_{n_A}\right]\right\}\Big\}\Big\}
\end{aligned}
\tag{4-1}
$$

$$Maximize\ H_{rel}\left(Information\ Potential|\ A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A}\right)$$

$$= \sum_{p=1}^{n_{\Delta u_{i<0}}} \begin{matrix} -P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_i)_p \middle| A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A}\right] \\ \times\ ln\left\{P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_i)_p \middle| A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A}\right]\right\} \end{matrix}$$

$$-\ ln\left(n_{\Delta u_{i<0}}\right)$$

$$= -\sum_{p=1}^{n_{\Delta u_{i<0}}} \begin{matrix} P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_i)_p \middle| A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A} \right. \\ \times\ ln\left\{n_{\Delta u_{i<0}} P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_i)_p \middle| A_i\ \overline{Preferred} \cap A_i\ Selected \right.\right. \end{matrix} \quad (4\text{-}2)$$

$$Maximize\ H_{rel}\left(Value\ of\ Information|\boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right) =$$
$$= -\left\{P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right.$$
$$\times \ln\left\{2P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\}$$
$$+\left\{1 - P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\}$$
$$\left.\times \ln\left\{2\left\{1 - P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\}\right\}\right\}$$
$$+\left\{-\sum_{q=1}^{n_{vi>0}} \begin{matrix} P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = (vi)_q\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right] \\ \times \ln\left\{n_{vi>0} P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = (vi)_q\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\} \end{matrix}\right\}$$
$$\times\left\{1 - P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle|\ \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\}$$

(4-3)

4.3.2.1 Evolutions of DET

While the general objectives and the fundamental measures underlying the theory of Decision Entropy have remained unchanged, its principles, which specify the hierarchy of variables, have evolved along the time. For example, the applications presented in Chapters 5 and 6 were conducted at the time when theoretical principles were defined in terms of the concept of utility difference and the decisions between multiple alternatives were made based on the exhaustive collection of pair-wise decisions. The current version of the theory relies on the conception of *information potential*, which was first introduced by Mostofi (2018), and is defined as a quantitative measure of the maximum potential effect of information for each decision outcome is the difference between the utility value of selected alternative and the maximum utility value. The mathematical formulation of information potential, labeled as $\Delta u(\vec{\theta}_k, A_j\ Selected)$ by (Mostofi, 2018), is presented in Eq. (4-4):

$$\Delta u(\vec{\theta}_k, A_j\ Selected) = u(\vec{\theta}_k, A_j\ Selected) - \max\,[\Delta u(\vec{\theta}_k, A_i\ Selected)\ for\ all\ i] \tag{4-4}$$

Where

$$\Delta u(\vec{\theta}_k, A_j\ Selected) = 0 \quad if\ \ A_j\ is\ the\ preferred\ alternative$$

$$\Delta u(\vec{\theta}_k, A_j\ Selected) < 0 \quad if\ \ A_j\ is\ not\ the\ preferred\ alternative$$

4.3.3 Implementations of DET

The theory of decision entropy has been implemented to assess the uncertainties in a variety of practical engineering problems within the past years. Min (2008) applied the theory to a number of petroleum engineering problems, where parameters characterizing production, transmissibility, and spatial variability of an oil reservoir were uncertain. Figure 4.1 compares the Value of Perfect Information for a production unit obtained from assessments made by the proposed theory with those from a generic assessment.

Decision Entropy Theory has also been applied to assess the risk of a number of natural hazards. While Mostofi (2018) provided a general assessment of uncertainties for landslides, as complicated multi-hazard processes, Mostofi et al. (2019) conducted a case study for the catastrophic landslide of March 2014 in Oso, Washington and concluded that if the cost of prohibiting the development in the region for the next century becomes less than about 1/6 the cost of another massive landslide, the preferred decision is to avoid the risk and to prohibit civil developments. For the case, when the alternative decision of accepting the risk is preferred, Figure 4.2 presents the updated assessments of the parameters of renewal model describing the occurrence of characteristic landslides (Mostofi et al., 2019)

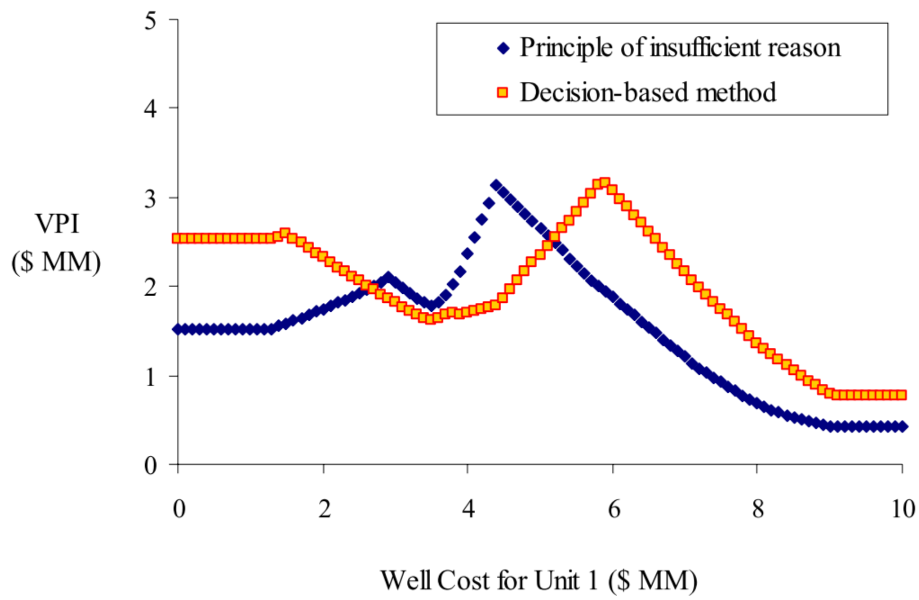

Figure 4.1. Variation of the value of perfect information (VPI) with well cost for an example production unit (Min, 2008)

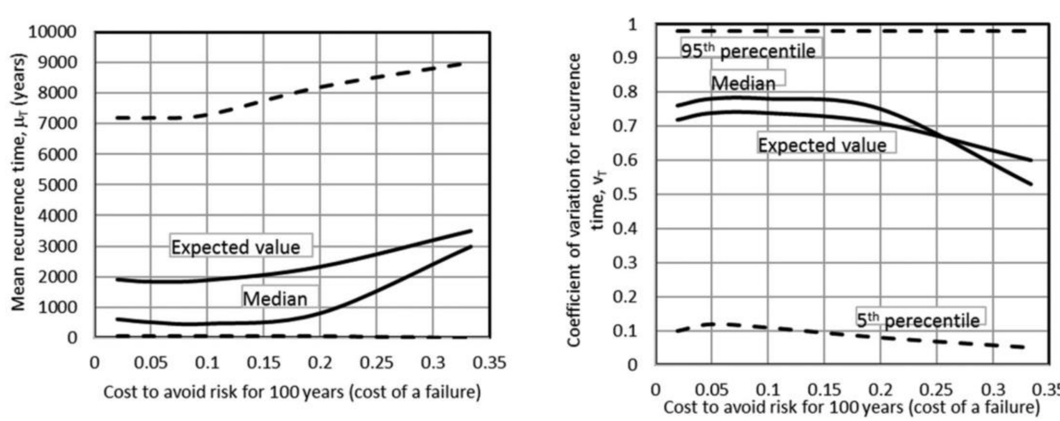

Figure 4.2. Informed assessments of the parameters describing the occurrence of characteristic landslides at Oso if one chooses to accept the risk (Mostofi et al., 2019)

Decision Entropy Theory has also been used to assess the effect of natural hazards on engineering structures. Feng et al. (2019) have evaluated the rehabilitation decision for a rockfill dam in Norway, where the overtopping of water in the dam reservoir is uncertain. It has been demonstrated that a decrease in the amount of available information and/or an increase in the cost of failure relative to the cost of rehabilitation results in an increase in the value of perfect information about the hazard curve increases (Feng et al. 2019). Figure 4.3 demonstrates the changes in expected normalized information potential for different normalized rockfill dam failure cost.

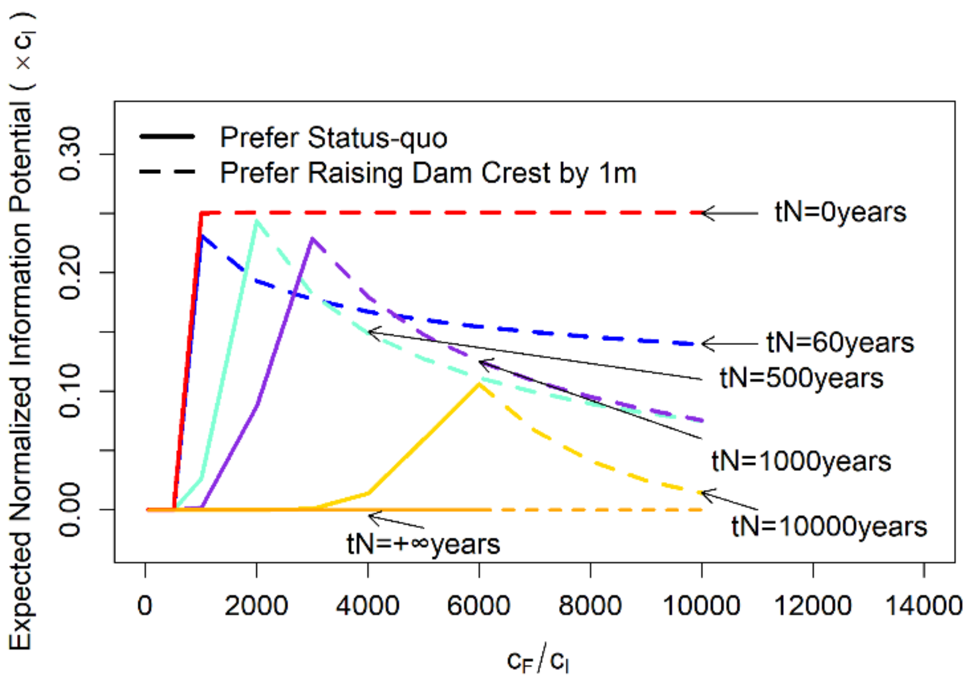

Figure 4.3. Expected normalized information potential for different normalized rockfill dam failure cost (Feng et al., 2019)

4.4 Evaluating the Objectivity, Transparency, and Defensibility of DET

An objective of the research was to evaluate whether the assessments of DET can be characterized as objective, transparent, and defensible. The comparison between the three principles of DET, as stated in 4.3.2, with the seven general principles, stated in Chapter 3, reveals that DET is in compliance with the more fundamental principles for objective assessment of probabilities. Thus, we can conclude that DET leads to assessments that are objective, transparent, and defensible. The details of the compliance are as follows:

4.4.1 Evaluative Orientation Principle

By considering the following statement of the principle of evaluative orientation:

- Inductive analysis is a goal-oriented act and must be directed toward answering a set of target questions

We can conclude that DET satisfies "Evaluative Orientation," as it takes inductive analysis to be a goal-oriented act and directs the assessment toward answering a set of target questions.

4.4.2 Investigative Prioritization Principle

By considering the following statement of the principle of investigative prioritization:

- Target questions must be prioritized in order to direct the inductive analysis to approach questions according to their importance

We may conclude that DET performs “Investigative Prioritization,” as it prioritized the questions according to their importance to the investigation..

4.4.3 Explicative Sufficiency Principle

Based on the following explanation of the principle:

- A (limited) number of distinct explanations/ hypotheses must be formulated such that each provide definitive answers to target questions

We may conclude that DET is Explicatively Sufficient, as it designates a number of distinct processual explanations, such that each provide definitive answers to the target questions.

4.4.4 Evaluative Inclusion Principle

By considering the statement of the inclusion principle as below:

- When data is limited such that an explanation/hypothesis does not stand out, assessments based on (the validity of) different hypotheses must be combined in order to form an inclusive/ collective assessment.

We may conclude that DET is evaluatively inclusive, as it allows the formation of a hybrid and inclusive explanation by considering and combining collection of explanations.

4.4.5 Credibility Conception Principle

According the following description of credibility conception principle:

- To form an inclusive assessment, each explanation/ hypothesis must be differentiated with a degree of credibility/ plausibility relative to other explanations.

DET is based on credibility conception, as it uses credibility values to form an inclusive assessment, and to differentiates various explanations with a degree of plausibility relative to other explanations.

4.4.6 Artifact-Reality Division Principle

By taking into account the statement of the artifact-reality division principle:

- Credibility must be defined to explicitly account for (1) what we know, observations, and (2) what we arbitrarily devise to reason beyond what we know, explanation/ hypothesis, such that the contribution of (1) and (2) in the collective assessment be tractable.

We may conclude that DET divides artifacts from reality, as its assigned credibilities explicitly account for (1) what we know, observations, and (2) what we arbitrarily devise to reason beyond what we know, explanation/ hypothesis, such that the contribution of (1) and (2) in the collective assessment is tractable.

4.4.7 Unbiased Evaluation Principle

According to the following statement of the unbiased evaluation principle:

- Credibility values must not be established such that the inclusive assessment favors any of the possible answers to the target questions before incorporation of observations.

We may conclude that DET provides with unbiased evaluations, as its assigned credibility values are not such that the inclusive assessment favors any of the possible answers to the target questions, before incorporation of observations.

4.5 SUMMARY

A Black Swan is an infrequent but consequential processual outcome, whose realization may even become more impactful due to the lack of preparation for a deemed improbable event. The study of an example Black Swan, the 1998 failure of the financial derivative trading firm named Long-Term Capital Management (LTCM), reveals that (a) how our most elaborate and highly praised frequency models, which are devised to describe and predict processual variations, can fail to account for unprecedented outcomes, and (b) how our confidence and reliance on such models, often justified by their success in addressing frequent outcomes, can result in unforeseen colossal consequences.

Accounting for Black Swan Outcomes (BSO) is declared to be the practical motivation of the research and offering a defensible procedure for assigning non-informative Bayesian credebilities is declared to be the theoretical motivation, which ultimately serves the practical motivation. The failure to contemplate Black Swan Outcomes (BSO) is attributed to our naïve empirical approach in making processual assessments, resulting in us neglecting extreme events of the past by not conducting through historical investigations, or ignoring the realization of unprecedented events in the future by (completely) matching the assessment with historical records. While the former deficiency can be covered by collecting a more exhaustive dataset, the later can only be addressed by conceiving a future (trend) different from the past and justifying the conception on (purely) theoretical grounds.

A (set of) theoretical means with the potential to account for yet-to-be-realized BSOs are non-informative Bayesian priors, which provide with assessments of processual tendencies before incorporating (past) realizations. Nevertheless, using Bayesian approach faces conceptual and methodical challenges. Conceptually, the fact that fundamental questions of "what Bayesian probabilities are" and "whether they represented (observationally) measurable quantities" still do not have satisfactory answers, have given the opponents reasons to question and reject the Bayesian approach. Methodically, the fact that there is no consensus on the assignment of Bayesian probabilities for a specific problem given certain observations, has forced analysts to make an uncomfortable choice among priors and to legitimize their choice either by portraying it as a (direct) representation of the (historical) data, or by downplaying its effects on the final assessment.

This research tries to address both conceptual and methodical challenges facing the Bayesian approach. While Chapters 2 and 3 are aimed at developing a philosophical framework to provide with conceptual clarity and precision, this chapter is aimed at introducing a computational theory to defensibly and transparently assign exact values to Bayesian probabilities for specific problems with certain available information.

A computational apparatus, *Decision Entropy Theory* (DET), is being introduced to assess unknown aspects of a process in an objective and defensible manner, in the sense described in Chapters 2 and 3, meaning that (a) its assessment procedures are transparent and imitable and (b) the constraints specifying the process are set in compliance with a set of intuitively logical guiding principles. DET considers a hierarchy of variables and assigns non-informative Bayesian priors to minimally inform the variables, whose exact values are the subjects of the inductive inquiry. Since Shannon's entropy is the adopted measure of information, the least informative set of distributions are those with the highest entropy values. According to the three principles of DET, (I) "decision outcome," (II) "difference in preference," and (III) "learning potential/information value," constitute the hierarchy of variables, which assigned priors must minimally inform. For risky decisions among a number of distinct alternatives, decision outcome is a categorial variable.

To evaluate objectivity, transparency, and defensibility of DET assessments, a comparison is being made between the three principles of DET, stated in this chapter and the seven general principles, stated in Chapter 3, namely (I) Evaluative Orientation, (II) Investigative Prioritization, (III) Explicative Sufficiency, (IV) Evaluative Inclusion, (V) Credibility Conception, (VI) Artifact-Reality Division, (VII) Unbiased Evaluation. Since we found DET is in compliance with the more fundamental principles for objective assessment of probabilities, we can conclude that DET delivers assessments that are objective, transparent, and defensible.

Chapter 5 Assessing Value of Test Wells in Developing an Unconventional Play

This chapter presents the research material published in the article (Habibi et al, 2014), which aims at providing inductive answers to questions related to proportion of object with certain characteristics, namely economic profitability.

5.1 INTRODUCTION

Unconventional plays are production opportunities with the potential for considerable uncertainty in production and economics. While trying to reduce production uncertainty with well tests may be valuable, there is a tradeoff between the quantity and quality of the information and its potential to improve development decisions. This paper proposes a new approach to represent uncertainty in an objective and defensible way, such that rare but important possibilities are not neglected in making decisions. This approach is called Decision Entropy Theory.

The motivation to develop Decision Entropy Theory is to account for the possibility of events outside our range of experience. Probabilities for possibilities beyond our experience cannot be assessed via statistics. Taleb (2007) refers to these possibilities as "Black Swan" events: "*Before the discovery of Australia, people in the Old World were convinced that all swans were white, an unassailable belief as it seemed completely confirmed by empirical evidence… [The sighting of the first black swan] illustrates a severe limitation to our learning from observations or experience and the fragility of our knowledge. One single observation can invalidate a general statement* *derived from millennia of confirmatory sightings of millions of white swan*s." The United States Secretary of Defense, Donald Rumsfeld, referred to these possibilities "unknown unknowns" in a famous press conference (DOD 2002): *"There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know."* Ten years ago, no one would have predicted that the United States would be contemplating exporting oil in 2014 because this possibility was beyond our range of experience at that time.

In the following, we will describe Decision Entropy Theory and illustrate its application and what can be learned from this theory in assessing the value of test wells in an unconventional play.

5.2 DECISION ENTROPY THEORY

Decision Entropy Theory is developed from three premises (axioms) that describe the case of no information or maximum uncertainty in making a decision between two alternatives.

Axiom 1: An alternative compared to another alternative is equally probable to be preferred or not to be preferred.

Axiom 2: The possible gains or losses for one alternative compared to another alternative are equally probable.

Axiom 3: The possibilities of learning about the preference of one alternative compared to another with new information are equally probable.

This theory characterizes uncertainty in making a decision; the case of maximum uncertainty corresponds to the maximum entropy for the possible outcomes of the decision. Therefore, the starting point for assessing probabilities, i.e., the non-informative prior probabilities of possibilities before information is included, depends on the decision at hand. Mathematical details laying out the development of the theory can be found in Gilbert et al. (2012). This paper will focus on the practical implications of this theory to the development of unconventional resources.

5.3 Go - No Go Decision

Consider a generic decision about whether or not to develop a play with unconventional resources (Figure 5.1). The profit for the play (p) can be calculated using Eq. (5-1) where c is the cost of developing the play, f is the frequency of productive ("good") wells in the development, r/c is the return on investment ratio for a good well, and there is no return on investment for bad wells. The frequency of good wells in the play is uncertain and denoted as a random variable, F.

$$p = -c + fr = r[-c/r + f] + c[-1 + (r/c)f] \qquad (5\text{-}1)$$

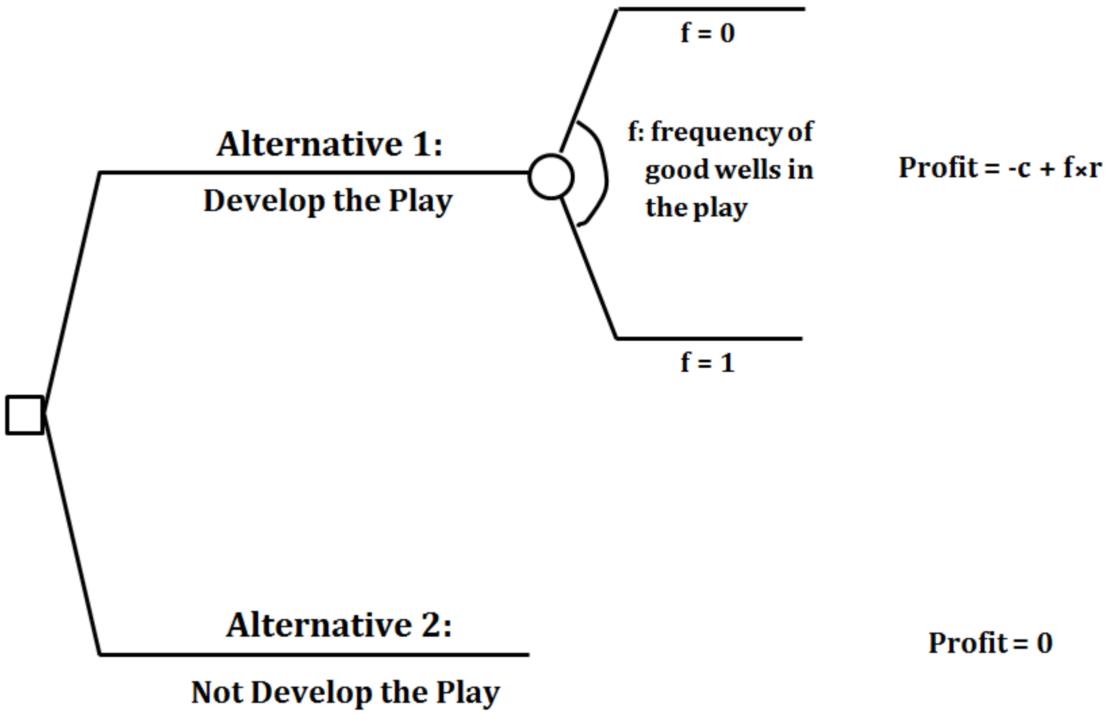

Figure 5.1. Decision tree for a development decision

Figure 5.2 shows how the non-informative prior probability distribution for the frequency of good wells is established based on Axioms 1 and 2 of Decision Entropy Theory. The probabilities that different outcomes of the decision will be realized are set such that there is a probability of 50 percent that the decision to develop the play will be correct and an equal probability for different possible net profits if the decision is or is not correct (Figure 5.2). This probability distribution is then mapped onto the probability distribution for the frequency of good wells (Figure 5.3).

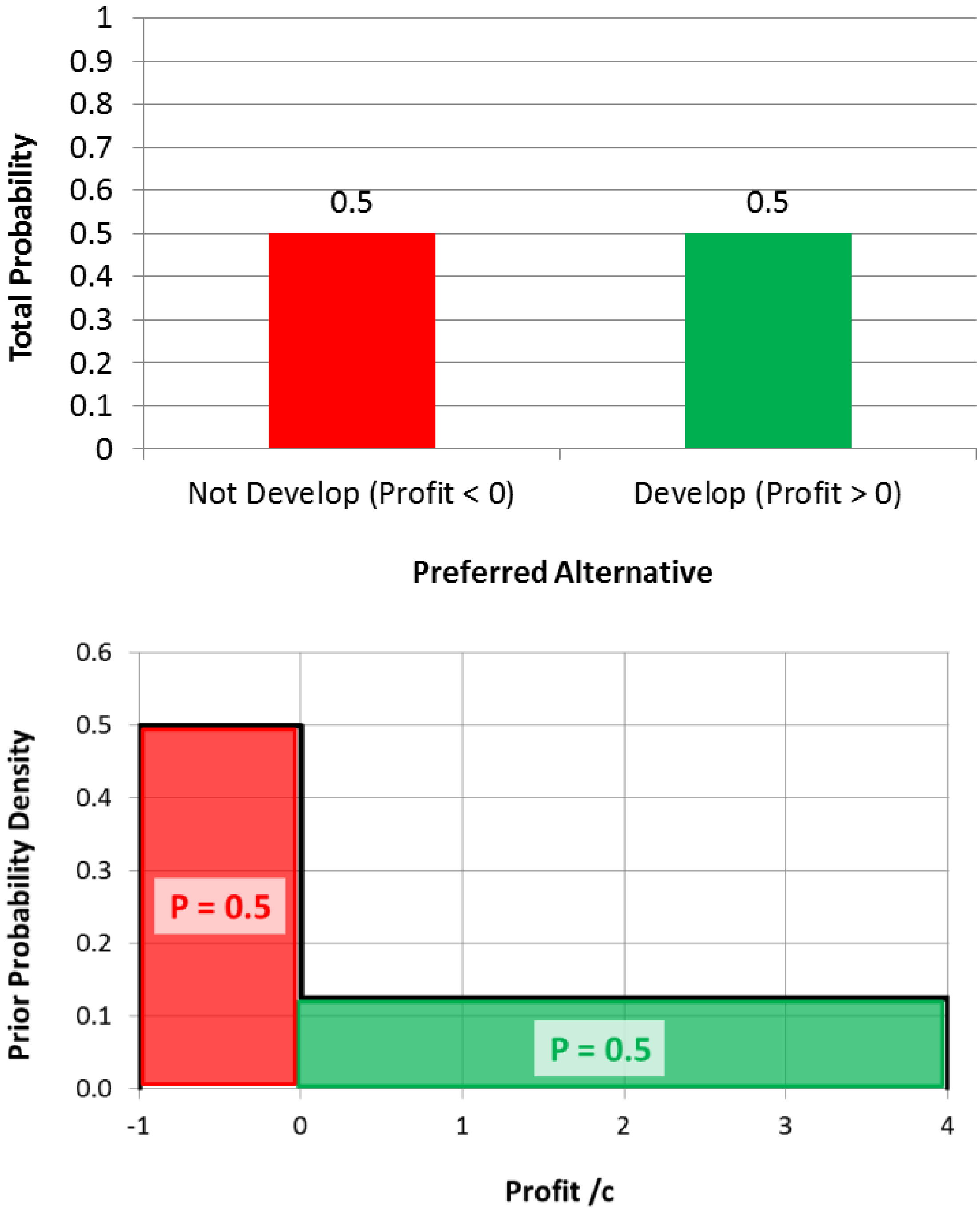

Figure 5.2. non-informative prior probability distribution for profit (p) using Decision Entropy Theory; Top: Applying axiom (1) to decision alternatives, Bottom: Applying axiom (2) to possible gains/losses

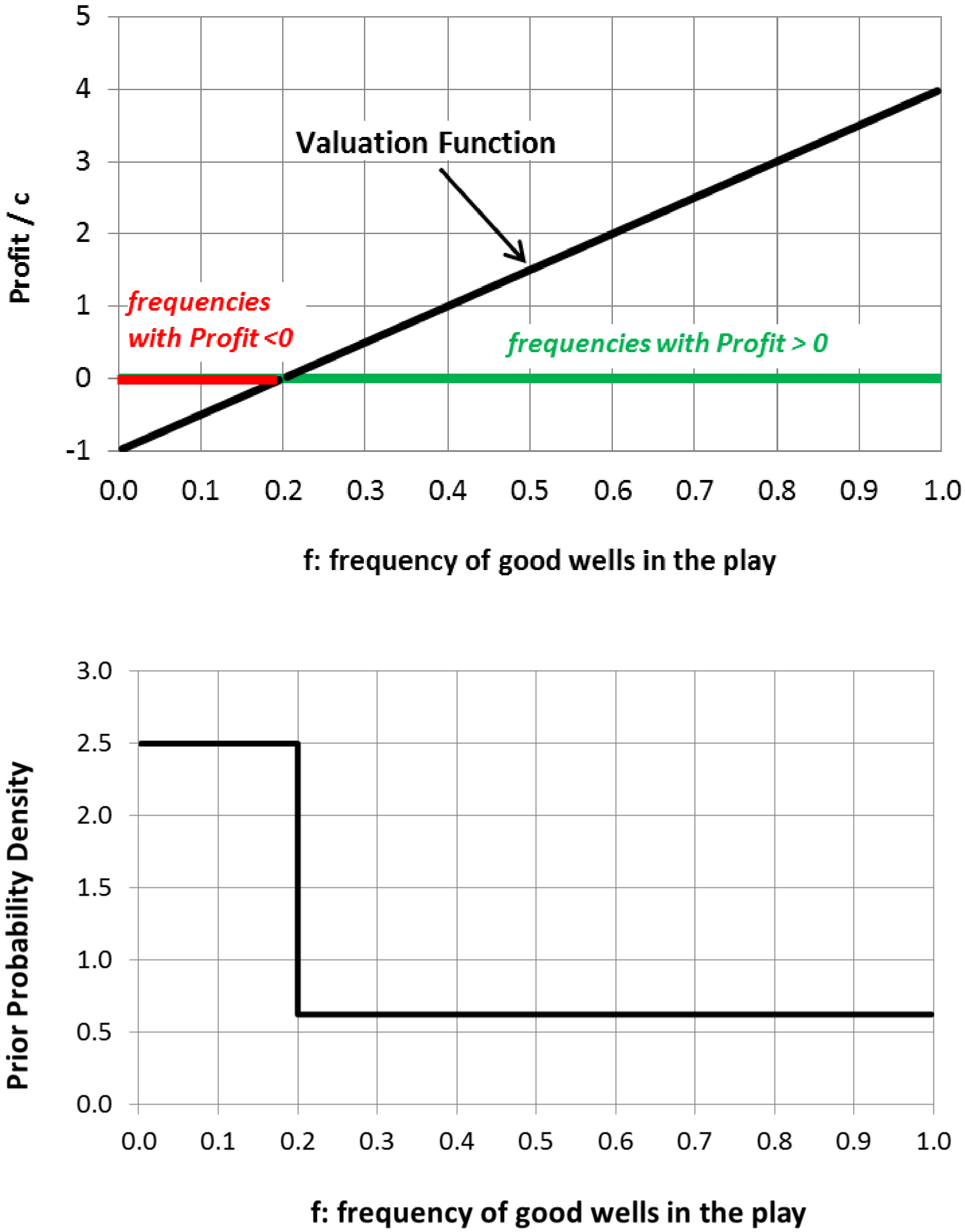

Figure 5.3. Establishing non-informative prior for frequency of good wells using Decision Entropy Theory; Top: relating profit (p) to frequency of good wells, Bottom: resulting non-informative prior for frequency of good wells

The non-informative prior probability distribution for the frequency of good wells depends on the specifics of the decision, specifically the return-on-investment ratio *r/c* for good wells since the break-even frequency of good wells for the play is equal to $f* = 1/(r/c)$. Figure 5.4 illustrates non-informative prior probability distributions for different break-even frequencies of good wells.

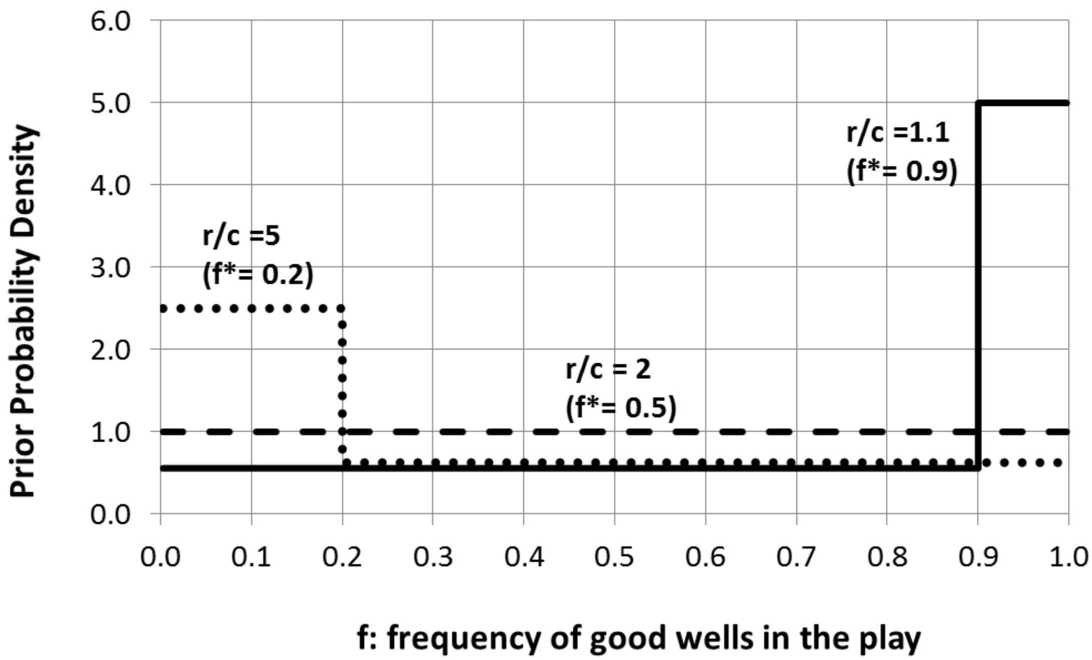

Figure 5.4. Establishing non-informative prior for frequency of good wells

The non-informative prior probability distribution can be used to make a decision about whether or not to develop the play in the event that no information is available on the frequency of good wells. While having no information is not necessarily a practical case (i.e. in most cases some information is available), it provides a limiting case as guidance. Figure 5.5 shows the expected profit of developing the play, $E[p]/c =$

$[(r/c)E(f) - 1]$, versus the return-on-investment ratio for good wells for the case of no information about the frequency of good wells.

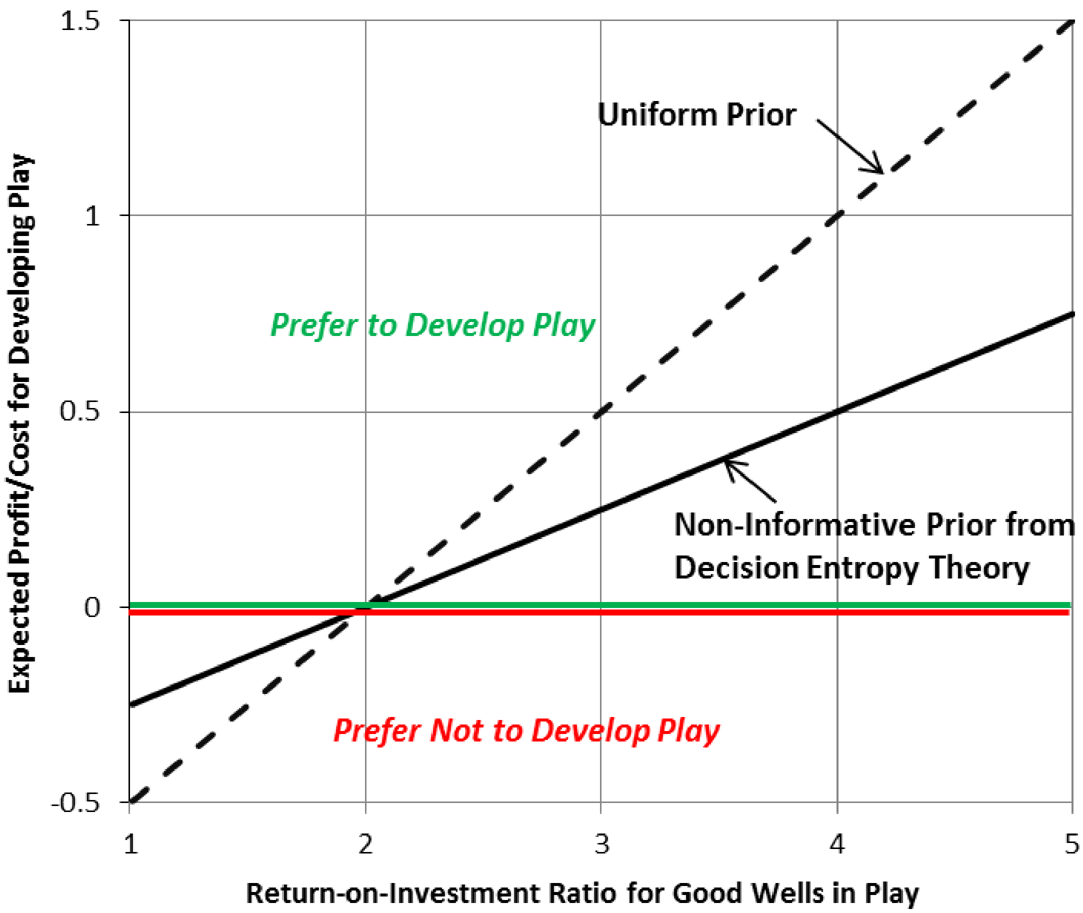

Figure 5.5. Establishing non-informative prior for frequency of good wells

Several alternatives for establishing non-informative prior probability distributions have been proposed over the years. Information theorists, such as Jaynes (1957) and Tribus (1967), propose that a uniform distribution for the frequency of good wells is a non-informative probability distribution because it maximizes the entropy for this variable. The challenge with this approach is that there is no consistent or defensible definition of the “variable” for which to maximize the entropy. For example, the non-informative prior probability distribution is different if you maximize the entropy of the logarithm of the frequency or the inverse of the frequency. Since Decision Entropy Theory maximizes the entropy in the decision rather than that in individual variables that affect the decision in different ways, it provides a consistent result for a particular decision. Bayesian decision analysts, such as Luce and Raiffa (1957), propose that there is always some information and therefore not a practical need for a non-informative prior probability distribution. The challenge with this approach is that it does not address most realistic situations where the available information is limited (as it always is) and does not allow for possibilities that have not been observed (i.e. the possibility of “Black Swans.”) Classical statisticians, such as Fisher (1935), propose that there is no defensible basis for establishing a non-informative prior probability distribution, meaning that probabilities can only be assessed when data are available. The challenge with this approach is that it is not practical; there is constantly a need to make decisions where we do not have “enough” relevant data because we are doing something new, such as with unconventional resources.

For comparison, Figure 5.5 includes the resulting expected profit of developing the play if a uniform distribution is assumed for the non-informative prior probability distribution of the frequency of good wells, regardless of the decision (i.e., regardless of the return-on-investment ratio for a good well). The threshold for obtaining a positive expected profit is the same with both approaches; a return-on-investment ratio greater than 2 leads to a positive expected profit. However, the rate of increase of the expected profit with increasing return-on-investment ratio is smaller when applying Decision Entropy Theory (Figure 5.5). For example, if the cost of developing the play is $100,000,000 and the return-on-investment ratio for good wells is 2.5, then the expected profit for the play is $12,500,000 with Decision Entropy Theory and $25,000,000 with a uniform prior distribution assumed for the frequency of good wells. Decision Entropy Theory pulls the expected profit toward zero by maintaining a 50-50 probability that the play alternative will be preferred with any return-on-investment ratio for good wells (Figure 5.4).

5.4 Value of Test Wells

A decision tree to assess the value of test wells is shown in Figure 5.6. We will assume that the test wells are representative of the wells in the play and that results between test wells are independent. The outcome of a test well program with n wells will be the number of good wells, x. The likelihood of such outcome can be calculated using binomial theorem. For any possible outcome (value of x) with a test program of n wells, the probability distribution for the frequency of good wells is updated with Bayes' Theorem:

$$P(F = f_i | x \text{ out of } n \text{ good wells}) = \frac{\left\{\left[\frac{n!}{x!\,(n-x)!}\right] {f_i}^{x} (1-f_i)^{n-x}\right\} P(F = f_i)}{\sum_{all\ i} \left\{\left[\frac{n!}{x!\,(n-x)!}\right] {f_i}^{x} (1-f_i)^{n-x}\right\} P(F = f_i)} \quad (5\text{-}2)$$

Where $P(F = f_i | x \text{ out of } n \text{ good wells})$ is the updated probability that the frequency is equal to a particular value and $P(F = f_i)$ is the prior probability for the frequency. An example of the prior and updated probability distributions for the frequency of good wells is shown in Figure 5.7 for one possible outcome of a well test program.

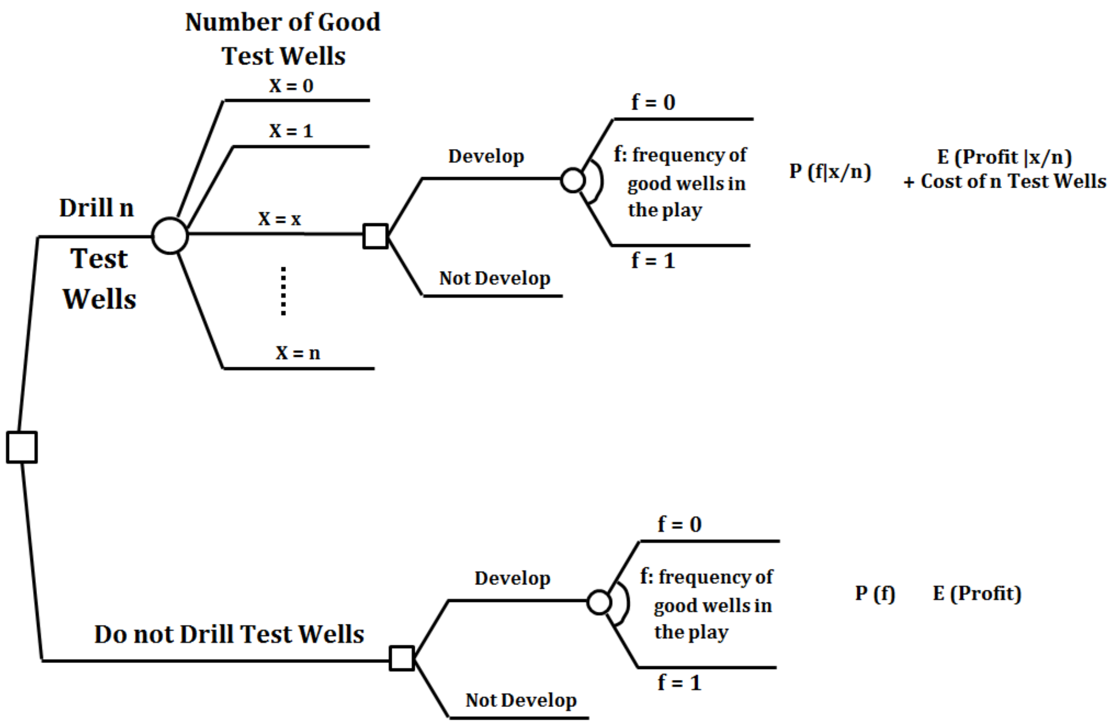

Figure 5.6. Decision tree for evaluating the value of test wells before deciding to develop the play

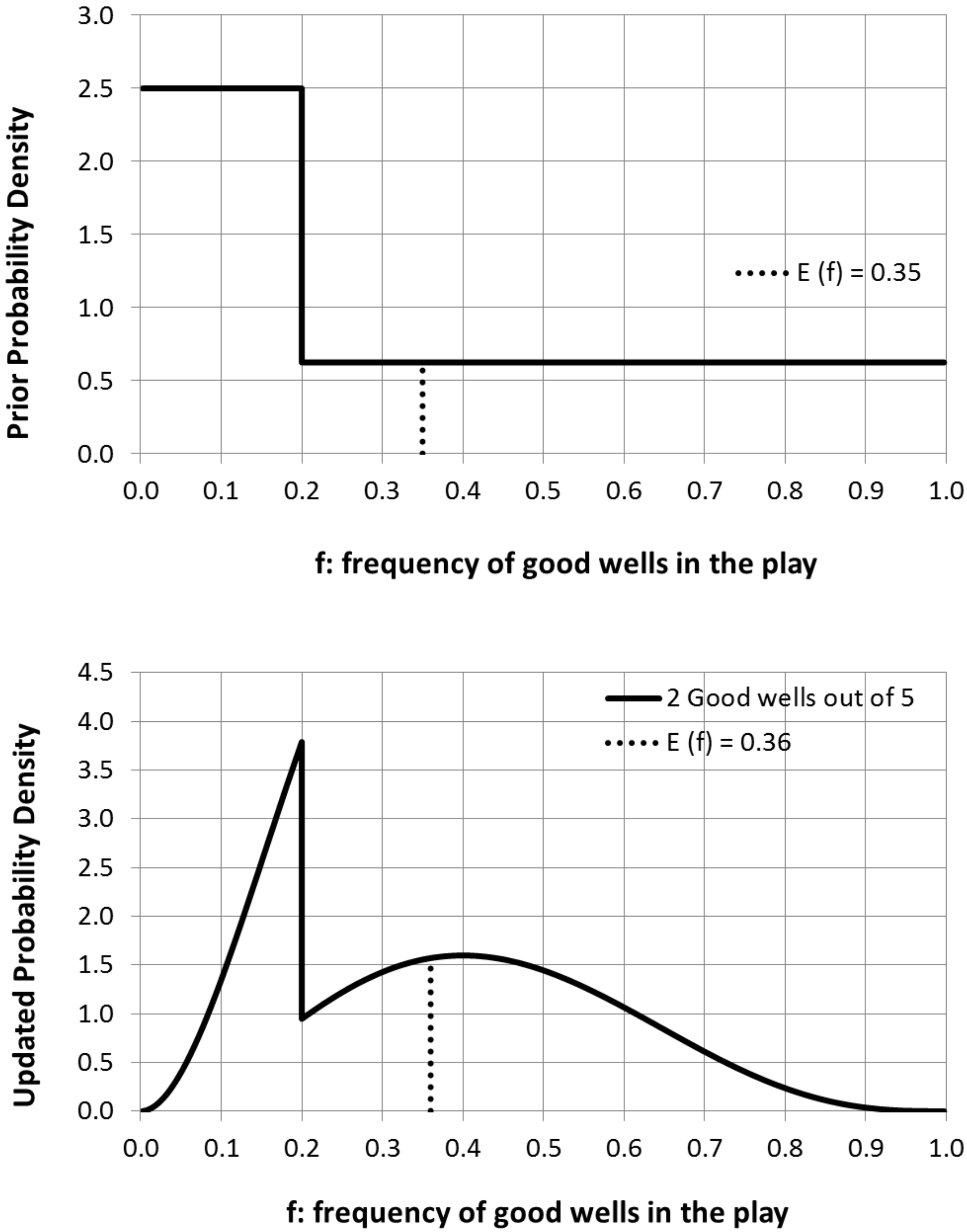

Figure 5.7. Probability distributions of frequency of good wells in a play with r/c = 5; Top: Prior distribution, Bottom: Posterior distribution from Decision Entropy Theory based on one result for test wells from the play

A limiting case for information is perfect information or an infinite number of test wells. The value of perfect information about the frequency of good wells is shown in Figure 8 as a function of the return-on-investment ratio for good wells in Figure 8. To facilitate showing a range of possibilities, the inverse of the return-on-investment ratio $1/(r/c)$, which is the break-even frequency for good wells, is used. To generalize the results, the value of perfect information is normalized by the maximum possible revenue for all good wells, r_{max}. The value of perfect information is the greatest when the break-even frequency is near 0.5 or the return-on-investment ratio is near 2. Information about the frequency of good wells is of greatest value in this case because the decision alternatives are balanced (the expected profit for developing the play equals zero – Figure 5.5) and there is the greatest potential for information to tip the decision either way.

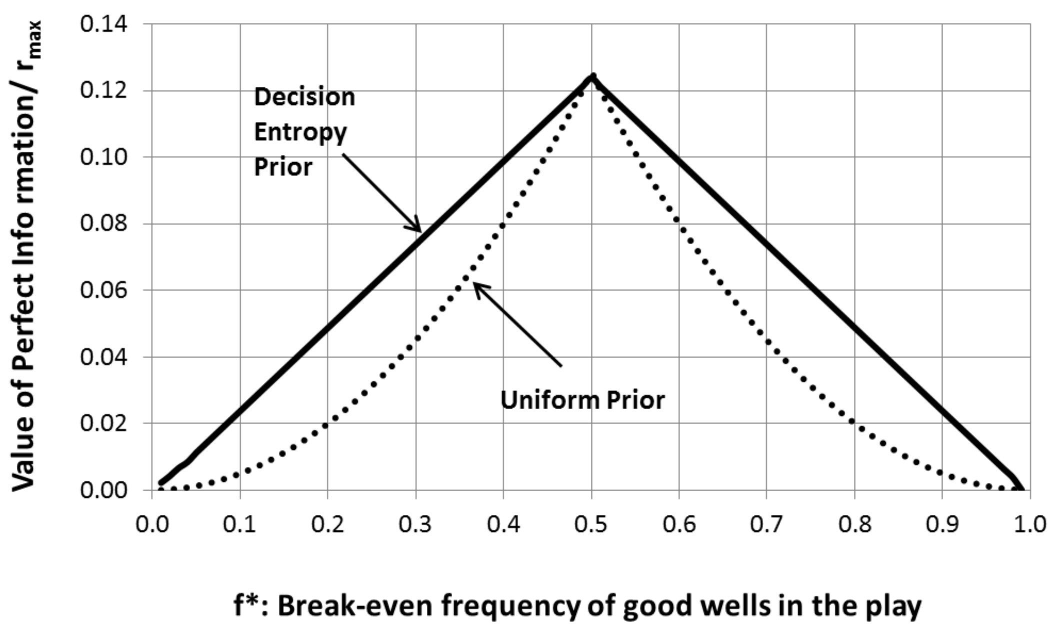

Figure 5.8. Value of perfect information for non-informative prior probability distributions for frequency of good wells in play based on Decision Entropy Theory compared with a uniform prior probability distribution for frequency of good wells

For comparison purposes, Figure 5.8 also shows the value of perfect information if a uniform distribution is assumed for the non-informative prior probability distribution of the frequency of good wells. The value of perfect information can be significantly greater when applying Decision Entropy Theory than when assuming a uniform prior probability distribution for the frequency of good wells. For example, if the return-on-investment ratio for good wells is equal to three (or the break-even frequency is 1/3) for a development costing $100,000,000, then the value of perfect information from a well test program is $24,000,000 with Decision Entropy Theory and $16,000,000 with a uniform prior probability distribution. This result reflects maximizing the entropy in the decision outcomes with Decision Entropy Theory; there is more potential to learn with information when entropy in the decision outcomes is greatest at the outset.

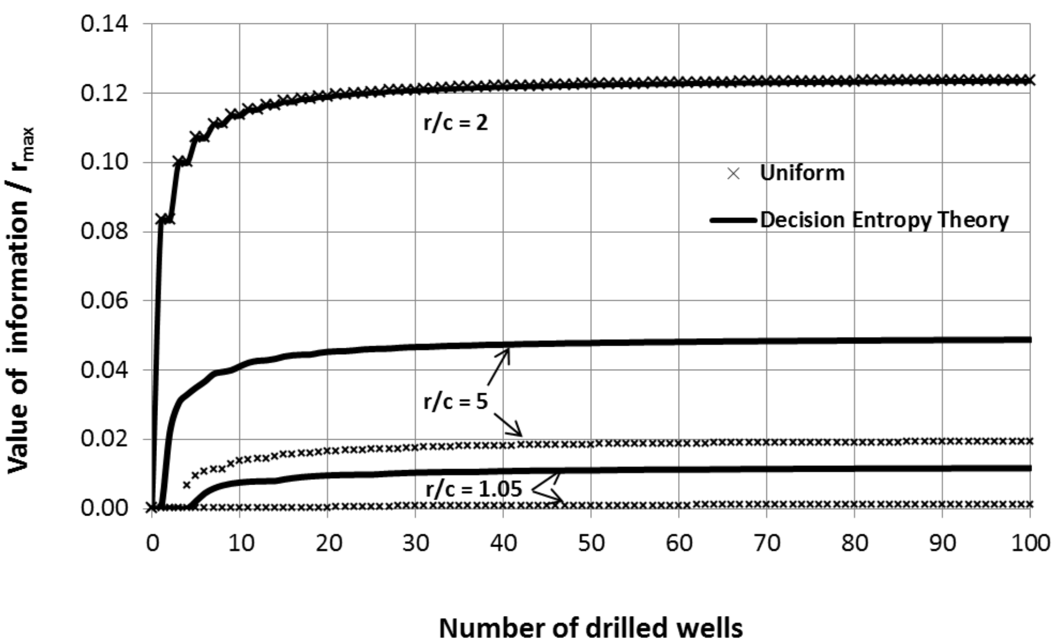

Figure 5.9. Value of information versus number of test wells for plays with different return-on-investment ratios

The value of information for a test program with a finite number of wells is shown in Figure 5.9. In some cases, multiple test wells are needed before any value is realized because too few wells will not be enough to change the preferred decision alternative even if all of them produce the same outcome. These results further highlight how the value of information is affected by the prior probability distribution for the frequency of good wells. For example, if the return-on-investment ratio for good wells is equal to 1.05, there is a positive value associated with drilling 10 test wells with Decision Entropy Theory but no value with a uniform prior probability distribution for the frequency of good wells (Figure 5.9). The difference between the two approaches is most pronounced in the extremes where the break-even frequency is either very small (a high return-on-investment ratio) or very large (a small return-on-investment ratio). These cases represent the "Black Swan" types of events, such as a play with a small return-on-investment ratio that turns out to be profitable because there is a high frequency of good wells or a play with a large return-on-investment ratio that turns out to be not profitable because there is a small frequency of good wells. Therefore, this example demonstrates how Decision Entropy Theory might change the perspective of a decision-maker to obtain more information before walking away from a play or committing to a play.

5.5 Relevancy of Information from Analogous Fields

The decision about whether or not to develop a new play will typically be informed by historical information from analogous plays. Consider a case where there is a Field A that may have the same frequency of good wells as in the play that is being considered, Field B. The use of the word "may" is significant; we will not know if the frequencies are the same until Field B is completely developed. A decision tree to assess the value of test wells in the new play is shown in Figure 10.

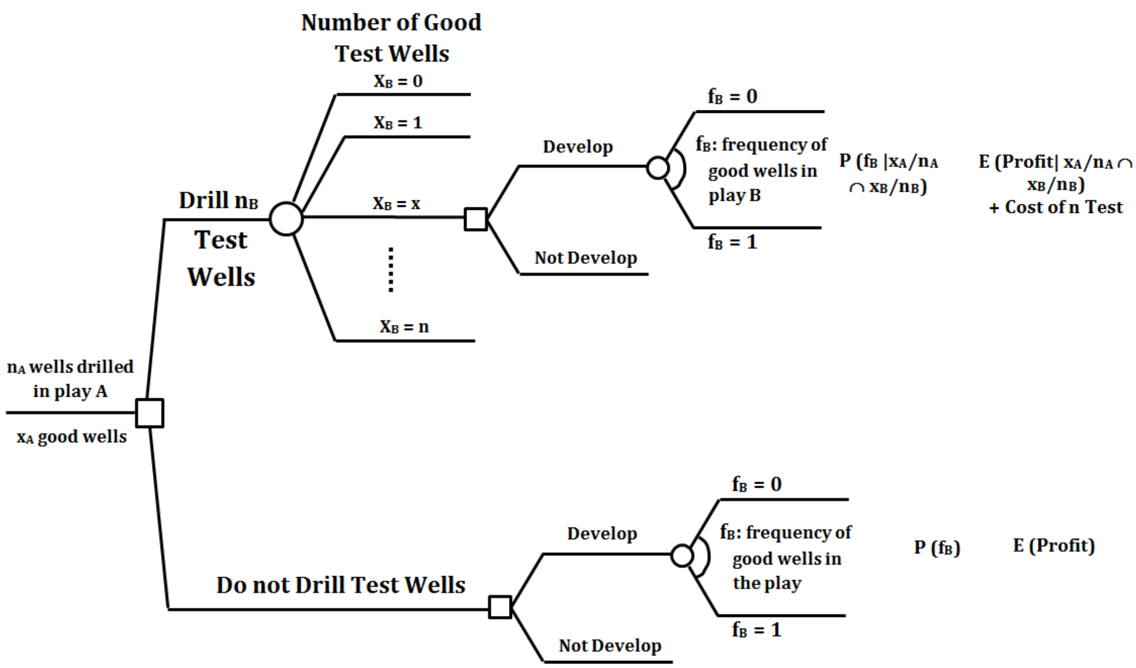

Figure 5.10. Decision tree for evaluating the value of information obtained from play B

The prior probability distribution for the frequency of wells in Field B without any information is the same (Figure 5.2 and 4.3). This prior probability distribution will be updated based on the information from Field A through Bayes' Theorem as follows:

$$P(F_B = f_{Bi}|\mathrm{x}_A \text{ out of } \mathrm{n}_A \text{ good wells in Field A})$$

$$= \frac{\begin{Bmatrix}[\mathrm{P}(\mathrm{x}_A/\mathrm{n}_A|\mathrm{f}_{Bi} \cap A = B)P(F_B = f_{Bi})P(A = B)] + \\ [\mathrm{P}(\mathrm{x}_A/\mathrm{n}_A|\mathrm{f}_{Bi} \cap A \neq B)P(F_B = f_{Bi})P(A \neq B)]\end{Bmatrix}}{\sum_{all\ f_{Bi}}\begin{Bmatrix}[\mathrm{P}(\mathrm{x}_A/\mathrm{n}_A|\mathrm{f}_{Bi} \cap A = B)P(F_B = f_{Bi})P(A = B)] + \\ [\mathrm{P}(\mathrm{x}_A/\mathrm{n}_A|\mathrm{f}_{Bi} \cap A \neq B)P(F_B = f_{Bi})P(A \neq B)]\end{Bmatrix}} \quad (5\text{-}3)$$

where x_A is the number of good wells out of n_A total wells in Field A, $P(A = B)$ is the prior probability that the new field is the same as the old field meaning the frequencies of good wells in the two fields are the same, and $P(A \neq B) = 1 - P(A = B)$ is the prior probability that the new field is different than the old field. If we assume the results between wells in each play are independent, then the likelihood of obtaining the set of data from Field A is

$$P(\mathrm{x}_A \,/\, \mathrm{n}_A | \mathrm{f}_{Bi} \cap A = B) = P(\mathrm{x}_A \,/\, \mathrm{n}_A | \mathrm{f}_A = \mathrm{f}_{Bi})$$

$$= \left\{ \left[\frac{\mathrm{n}!}{x!\,(n - x!)} \right] \mathrm{f}_{Bi}^{x_A} (1 - f_{Bi})^{n - x_A} \right\} \qquad \text{(5-4)}$$

and

$$P(\mathrm{x}_A \,/\, \mathrm{n}_A | \mathrm{f}_{Bi} \cap A \neq B)$$

$$= \sum_{all\, f_{Aj} \neq f_{Bi}} \left\{ \left[\frac{\mathrm{n}_A!}{\mathrm{x}_A!\,(\mathrm{n}_A - \mathrm{x}_A!)} \right] \mathrm{f}_{Aj}^{x_A} \left(1 - f_{Aj}\right)^{n_A - x_A} \right.$$

$$\left. \times\ P\left(F_A = f_{Aj} | A \neq B\right) \right\} \qquad \text{(5-5)}$$

Where $P(F_A = f_{Aj} | A \neq B) = P(F_A = f_{Aj} | A = B) = P(F_A = f_{Aj})$ since the prior probability for the frequency in Field A does not depend on whether Fields A and B are the same (i.e., the prior probability of obtaining a particular set of data from Field A is the same whether or not Fields A and B are the same and the probability of obtaining a particular set of data from Field A does not depend on the frequency in Field B if the two fields are different).

The prior probability that the two fields are the same is established based on Axiom 3 of Decision Entropy Theory. If the two fields are the same, then we can learn from Field A; if the two fields are different, then we cannot learn from Field A. Setting $P(A = B) = 1/2$ makes these two possibilities equally probable.

If n_B test wells are drilled in Field B, then the updated probability for the frequency of good wells in Field B is obtained from Bayes' Theorem as follows:

$$P\left(F_B = f_{Bi} \middle| \begin{matrix} \mathrm{x}_A \text{ out of } \mathrm{n}_A \text{ good wells in Field A} \\ \mathrm{x}_B \text{ out of } \mathrm{n}_B \text{ good wells in Field B} \end{matrix}\right)$$

$$= \frac{\begin{Bmatrix} [\mathrm{P}(\mathrm{x}_A/\mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A = B) P(F_B = f_{Bi}) P(A = B)] + \\ [\mathrm{P}(\mathrm{x}_A \,/\, \mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A \neq B) P(F_B = f_{Bi}) P(A \neq B)] \end{Bmatrix}}{\sum_{all\ f_{Bi}} \begin{Bmatrix} [\mathrm{P}(\mathrm{x}_A \,/\, \mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A = B) P(F_B = f_{Bi}) P(A = B)] + \\ [\mathrm{P}(\mathrm{x}_A \,/\, \mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A \neq B) P(F_B = f_{Bi}) P(A \neq B)] \end{Bmatrix}} \tag{5-6}$$

where the likelihood of obtaining the combined set of data if the two fields are the same is

$$P(\mathrm{x}_A/\mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A = B)$$

$$= \left\{ \left[\frac{\mathrm{n}_A!}{\mathrm{x}_A!\,(\mathrm{n}_A - \mathrm{x}_A!)} \right] \mathrm{f}_{Aj}^{x_A} \left(1 - f_{Aj}\right)^{n_A - x_A} \right\}$$

$$\times \left\{ \left[\frac{\mathrm{n}_B!}{\mathrm{x}_B!\,(\mathrm{n}_B - \mathrm{x}_B!)} \right] \mathrm{f}_{Bi}^{x_B} \left(1 - f_{Bi}\right)^{n_B - x_B} \right\} \tag{5-7}$$

and the likelihood of obtaining the individual data sets for each field if the two fields are different is

$$P(\mathrm{x}_A/\mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A \neq B)$$

$$= \mathrm{P}(\mathrm{x}_A/\mathrm{n}_A | \mathrm{f}_{Bi} \cap A \neq B)$$

$$\times \left\{ \left[\frac{\mathrm{n}_B!}{\mathrm{x}_B!\,(\mathrm{n}_B - \mathrm{x}_B!)} \right] \mathrm{f}_{Bi}^{x_B} \left(1 - f_{Bi}\right)^{n_B - x_B} \right\} \tag{5-8}$$

Figure 5.11 shows the value of perfect information about the frequency of good wells in the new play (Field B) versus the break-even frequency for an example case with a large quantity of data from an analogous field (Field A). The data from Field A comprise 25 good wells in a field of 100 total wells.

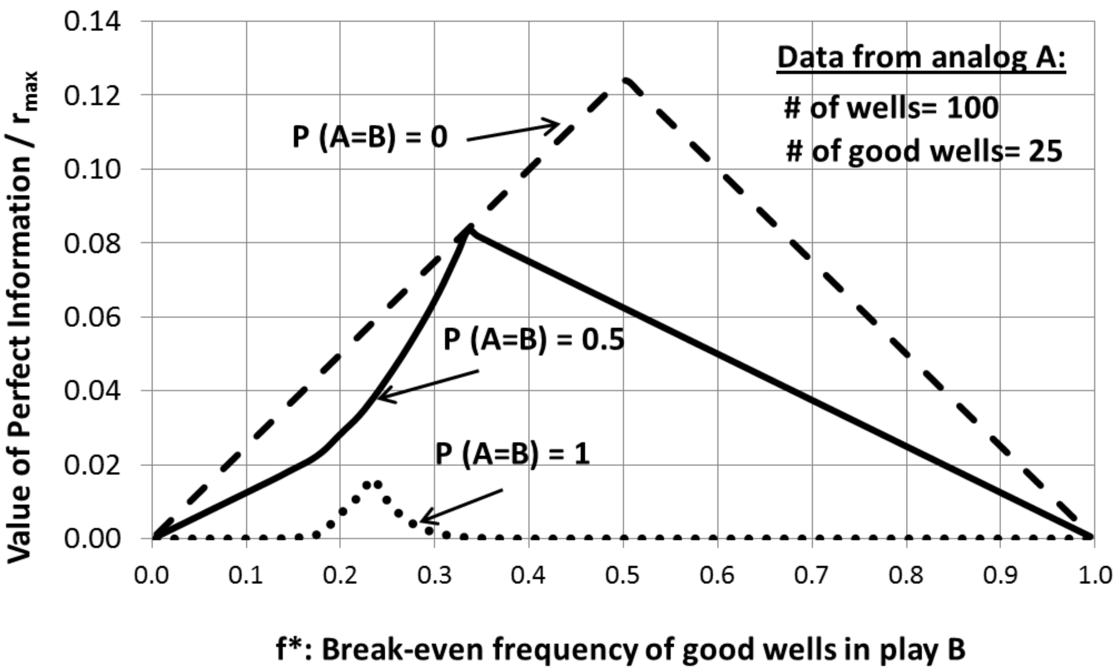

Figure 5.11. Value of perfect information about frequency for new play Field B versus the break-even frequency for Field B for different degrees of relevancy of data from analog Field A

If the data from analogous Field A are relevant, or $P(A = B) = 1$, then frequency of good wells in the new play Field B is very unlikely to be significantly different than 25/100 or 0.25 and there is essentially no value from additional information in the new play because the new data re not likely to change the development decision. Decision Entropy Theory, $P(A = B) = 1/2$, strikes a balance between relying entirely on the historical data, $P(A = B) = 1$, versus not relying on it at all, $P(A = B) = 0$ (Figure

5.11). This balance is particularly significant for cases where the new data from Field B contradict the analogous data with frequencies that are significantly greater than that from Field A (0.25). Decision Entropy Theory incorporates the previous experience while keeping the door open for possibilities that the new field will be something different than previous experience.

Figure 5.13 shows examples of the prior and updated probability distributions for the frequency of good wells in Field B after a number of test wells have been drilled in Field B for different results. These results demonstrate how the information from the analogous Field A takes on greater weight when the information from play Field B is consistent, while it takes on less weight when the information from Field B is inconsistent. To emphasize this point, the updated probability that the two fields are the same is included in Figure 5.13; this probability is obtained from Bayes' Theorem as follows

$$P\left(A = B \middle| \begin{matrix} \mathrm{x}_A \text{ out of } \mathrm{n}_A \text{ good wells in Field A} \\ \mathrm{x}_B \text{ out of } \mathrm{n}_B \text{ good wells in Field B} \end{matrix}\right)$$

$$= \frac{\sum_{all\ f_{Bi}} \mathrm{P}(\mathrm{x}_A/\mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A = B) P(F_B = f_{Bi}) P(A = B)}{\sum_{all\ f_{Bi}} \left\{ \begin{matrix} [\mathrm{P}(\mathrm{x}_A\ /\ \mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A = B) P(F_B = f_{Bi}) P(A = B)] + \\ [\mathrm{P}(\mathrm{x}_A\ /\ \mathrm{n}_A, \mathrm{x}_B/\mathrm{n}_B | \mathrm{f}_{Bi} \cap A \neq B) P(F_B = f_{Bi}) P(A \neq B)] \end{matrix} \right\}} \tag{5-9}$$

In the case when there are numerous data from Field B that are inconsistent with those from Field A (e.g., the lower left plot in Figure 5.13 where 50 of 100 wells in Field B are good but 25 of 100 wells in Field A were good), the updated probability that the fields are the same is reduced significantly and the data from Field A play very little role in the updated probability distribution for the frequency of good wells in Field B.

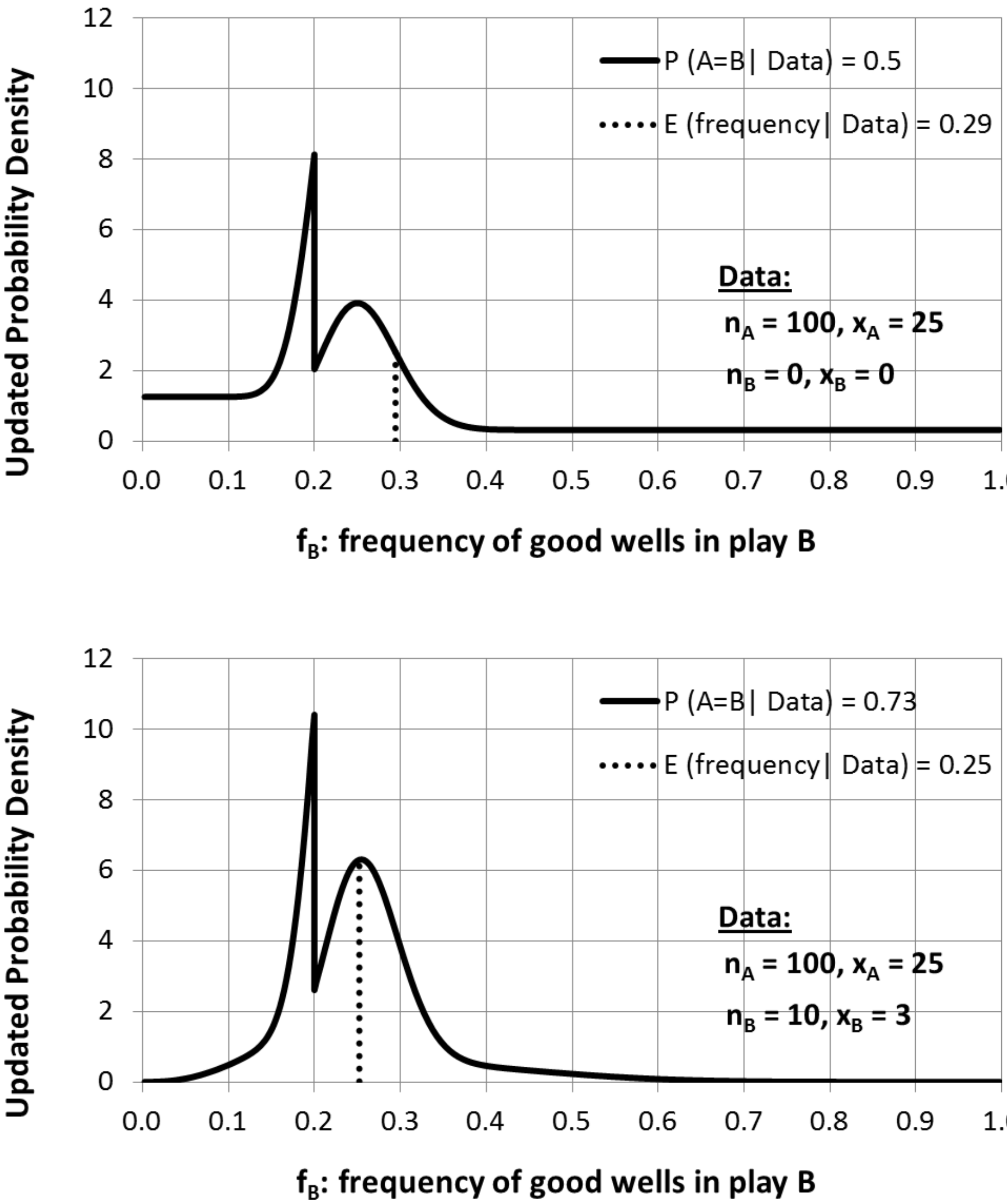

Figure 5.12. Distribution for frequency of good wells updated with data from analog Field A and from play Field B; Top: Prior distribution from analog data alone; Bottom: Posterior distribution with small number of test wells from play and results that are consistent with the analog

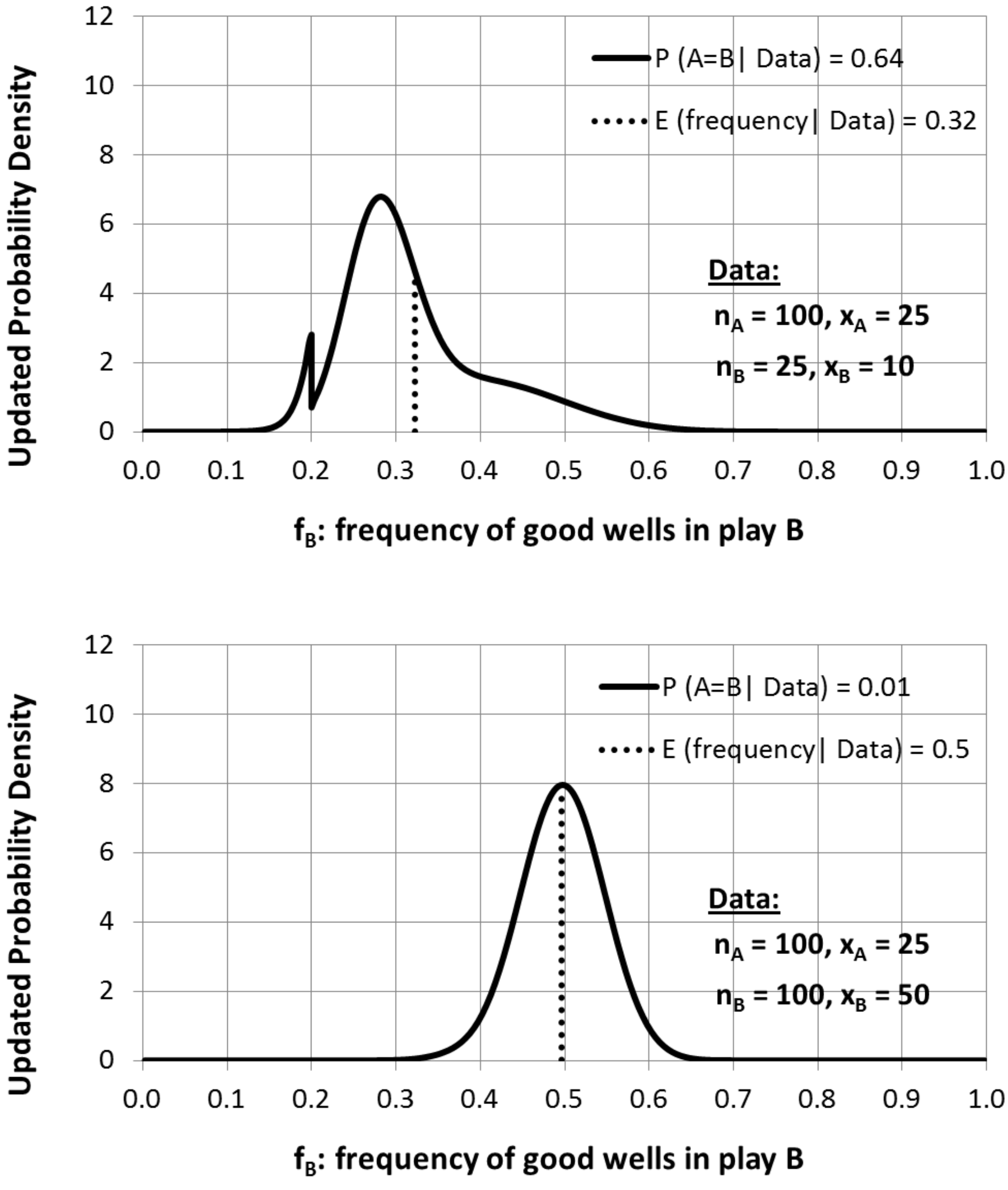

Figure 5.13. Distribution for frequency of good wells updated with data from analog Field A and from play Field B; Top: Posterior distribution with larger number of test wells from play and results that are not consistent with the analog; Bottom: Posterior distribution with equal number of wells from play as from analog and results that are not consistent with the analog

Figure 5.14 shows examples of the value of information versus the number of test wells in Field B. This value depends on the break-even frequency of good wells for Field B and the available data from Field A.

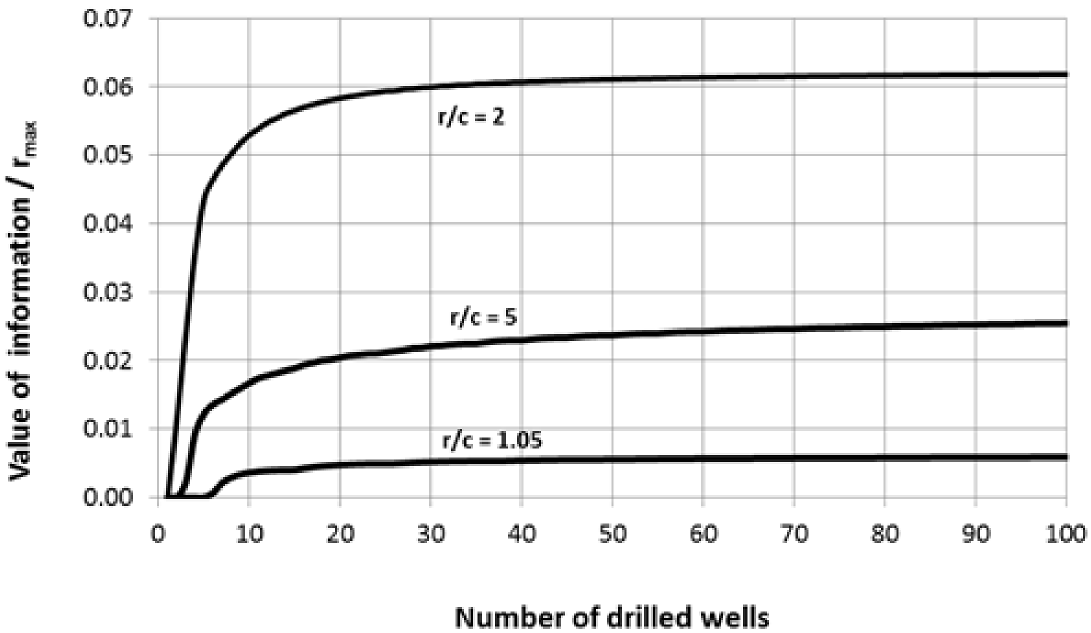

Figure 5.14. Value of information versus number of test wells in plays with different return-on-investment ratios

5.6 DISCUSSION

A practical principle arising from Decision Entropy Theory is that it is helpful to consider a decision from the top-down concerning the possibilities of preferring one decision alternative to another, also the possible gains or losses associated with each alternative. In this way, the set of possibilities in a decision is well defined and not constrained by experience. We will either prefer one alternative to another or not. It is not necessary to develop from the bottom up all the possible scenarios of variables affecting the decision that lead to one alternative being preferred. In this way, we can accommodate possibilities that are beyond our experience.

The value of a top-down perspective is illustrated in Figure 5.15. The profit for an unconventional play with gas wells is shown in Figure 5.15 as a function of the median value for the initial well production and a normalized metric for the commodity price of natural gas. The combinations of well performance and gas price leading to a profitable play are colored green, while those leading to an unprofitable play are colored red. Figure 14b shows the joint prior probability distribution for initial production and normalized gas price obtained from Decision Entropy Theory. The ridge in this joint distribution corresponds to the break-even combinations of these variables. The non-uniform shape of the prior joint probability distribution for initial well production and gas price reflects the non-linear relationship between these variables and the decision outcome. For example, the profit for the play reaches an asymptote as the median initial production rate increases because the potential profits become constrained by the distribution system and market demand. Possibilities for relatively high gas prices beyond experience are captured in this prior probability distribution without having to develop and assess probabilities for unprecedented scenarios that lead to those high prices.

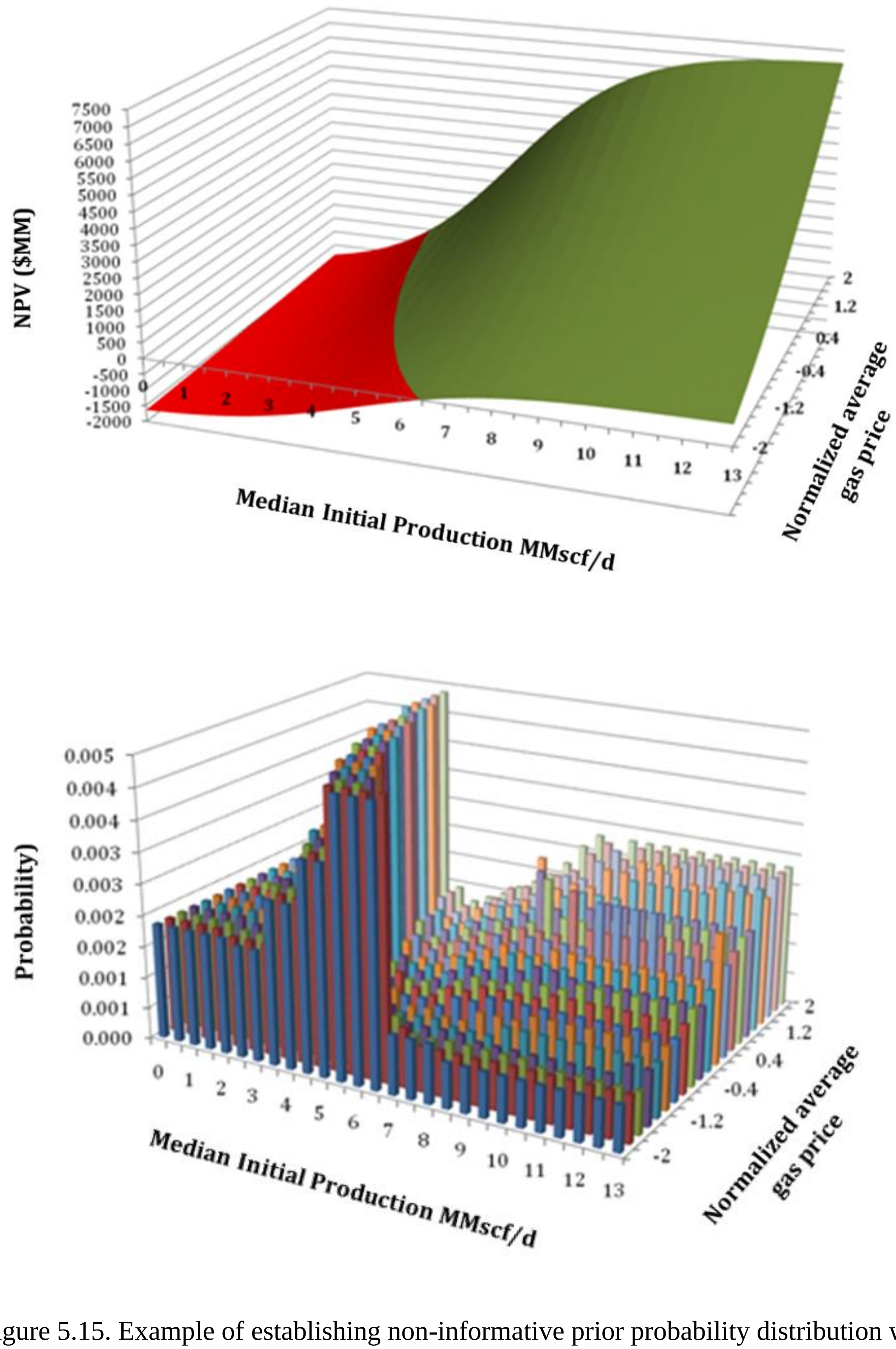

Figure 5.15. Example of establishing non-informative prior probability distribution with Decision Entropy Theory for unconventional play with gas wells

Another practical principle arising from Decision Entropy Theory is to be careful with the starting point in assessing probabilities for the input variables in a decision. In theory, the prior state should represent the case of having no information; information can then be used to update this prior state via Bayes' theorem. One reasonably common approach of maximizing uncertainty in the input variables, i.e., maximizing the entropy of their probability distribution, is not rational, consistent or even practical. It is not rational because different variables affect a decision in different ways. Journel and Deutsch (1993) demonstrate the lack of rationality very clearly with the spatial variability of permeability in a reservoir: the maximum entropy in the probability distribution for the permeability field (i.e., lots of randomness with very little spatial structure or correlation in extreme high or low permeabilities) produces relatively small variation and therefore uncertainty in well production across the field, while little entropy in the probability distribution for the permeability field (i.e., lots of structure with channels of high permeabilties and zones of low permeabilities) produces relatively large uncertainty in well production. This approach of uniform prior probability distributions for input variables is inconsistent; a uniform probability distribution for the permeability is different than one for the logarithm of permeability. This approach also is not practical when dealing with multiple input variables; if we extended the example in this paper to include three possible outcomes for the wells, "good, mediocre and bad," then a uniform probability distribution for one of these frequencies precludes a uniform probability distribution for the other two frequencies. Finally, the example in this paper illustrates how a uniform prior probability distribution can undervalue information (e.g., Figure 5.8 and Figure 5.9).

The examples presented in this paper were highly simplified to be clear. The examples would be more realistic if there were:

- More than two possible outcomes for each well, such as a range of cumulative production rates;
- Uncertainty in the investment cost and return;
- Different and multiple types of information about reservoir productivity, such as exploration data and short-term and localized well tests as well as production well data;
- Spatial and temporal correlations between wells, such as sweet spots or learning curves; and
- More than two possible cases of analog data being relevant or not, such as accommodating different correlative strengths between the analog data and the new field.

5.7 Conclusions

Unconventional plays have considerable uncertainty in production and economics. This paper proposes Decision Entropy Theory to represent uncertainty in an objective and defensible way, such that rare but important possibilities are not neglected in making decisions. This theory characterizes uncertainty in the context of making a decision; the case of maximum uncertainty corresponds to the maximum entropy for the possible outcomes of the decision.

The theory is applied to assessing the value of information from test wells for an unconventional play. The decision to develop a play is posed in terms of the the cost of development, the ratio of the return-on-investment for productive ("good") wells in the play, and the uncertain frequency of good wells in the play. Results for the value of tests wells are obtained and provided graphically in non-dimensional relationships for a variety of cases, including a finite and infinite number of test wells and incorporating data from an analogous play. The following conclusions are drawn from this application:

1. The probability distribution for the uncertain frequency of productive wells is important both for deciding whether or not to develop a play and for the value of information from test wells.

2. It is unreasonable to assume that the probability distribution for the frequency of productive wells can be established based entirely on experience because that precludes the possibility of events beyond our experience, which can be particularly important with unconventional plays.

3. The value of test wells is enhanced when leaving open the possibilities for profit on a play with a small return-on-investment ratio and the possibility for loss on a play with a large return-on-investment ratio.

4. There is a balance between relying entirely on historical data from analogous plays versus not relying on them at all. Within this balance, information for test wells at a new play can be combined with information from analogous plays to inform the decision about developing the new play.

5. The value test of wells at a new play depends on the break-even frequency of good wells in the play; in the case of no information, this value is a maximum when the break-even frequency is 50 percent.

There is significant room to extend this simplified example to more realistic cases.

Chapter 6 Achieving Reliability in the Face of Extreme Uncertainty

This chapter presents the research material published in 5APSSRA, Symposium on Structural Reliability and its Applications, which aims at providing a theoretical approach for recommending a rate of occurrence of a natural hazard for designing a structure to reliably withstand the recommended magnitude of the hazard. The article is the first published document, in which we label the collection of axioms we have used in our analysis "Decision Entropy Theory."

6.1 INTRODUCTION

Extreme uncertainty is the possibility of events that are beyond the range of experience. In probabilistic terms, it is the possibility of an event that is considered extremely unlikely or even impossible in establishing the sample space of possibilities.

In the popular literature, Taleb (2007) refers to possibilities of extreme uncertainty as "Black Swan" events: "*Before the discovery of Australia, people in the Old World were convinced that all swans were white, an unassailable belief as it seemed completely confirmed by empirical evidence… [The sighting of the first black swan] illustrates a severe limitation to our learning from observations or experience and the fragility of our knowledge. One single observation can invalidate a general statement derived from millennia of confirmatory sightings of millions of white swans*." The United States Secretary of Defense, Donald Rumsfeld, referred to extreme uncertainty as "unknown unknowns" in a famous press conference during the lead up to the invasion of

Iraq (DOD 2002): "*There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know*."

The world is filled with recent examples of extreme uncertainty. The storm surge in 2005 during Hurricane Katrina was almost 50-percent larger than the maximum storm surge that had ever been recorded along the Gulf Coast of the United States, it was well above the "extreme" conditions that were used to design the levee system, and it resulted in nearly 2,000 deaths. The magnitude of the 2011 Japanese earthquake was larger than anything in the historical record for that region, it was larger than what was expected to be the maximum possible, and it resulted in nearly 20,000 deaths. The 2008 banking crisis in the United States was unprecedented, it was well outside the range of possibilities considered in public policy, and it resulted in the worst economic conditions in the United States since the 1930's.

A specific example of extreme uncertainty is illustrated in Figure 6.1. For the first 50 years of offshore energy production in the Gulf of Mexico, from approximately 1950 to 2003, the largest hurricane waves impacting the offshore facilities were less than 26 m. Based on these data, the "100-year" design wave height was about 22 m. Then, five major hurricanes occurred between 2004 and 2008, and four of them had wave heights greater than the largest height that had ever occurred before (Figure 6.1). The impact of these unprecedented hurricanes was extensive. Figure 6.2 shows how five times as many offshore structures were destroyed in the past five years as in the entire previous history of operations. Based on a simple extrapolation of these new data for wave heights, the new “100-year” wave height is now over 26 m (Figure 6.1). Since the force on a fixed offshore structure is approximately proportional to the square of the wave height, the new “100-year” loads are nearly 50-percent greater than before based on just five years-worth of data.

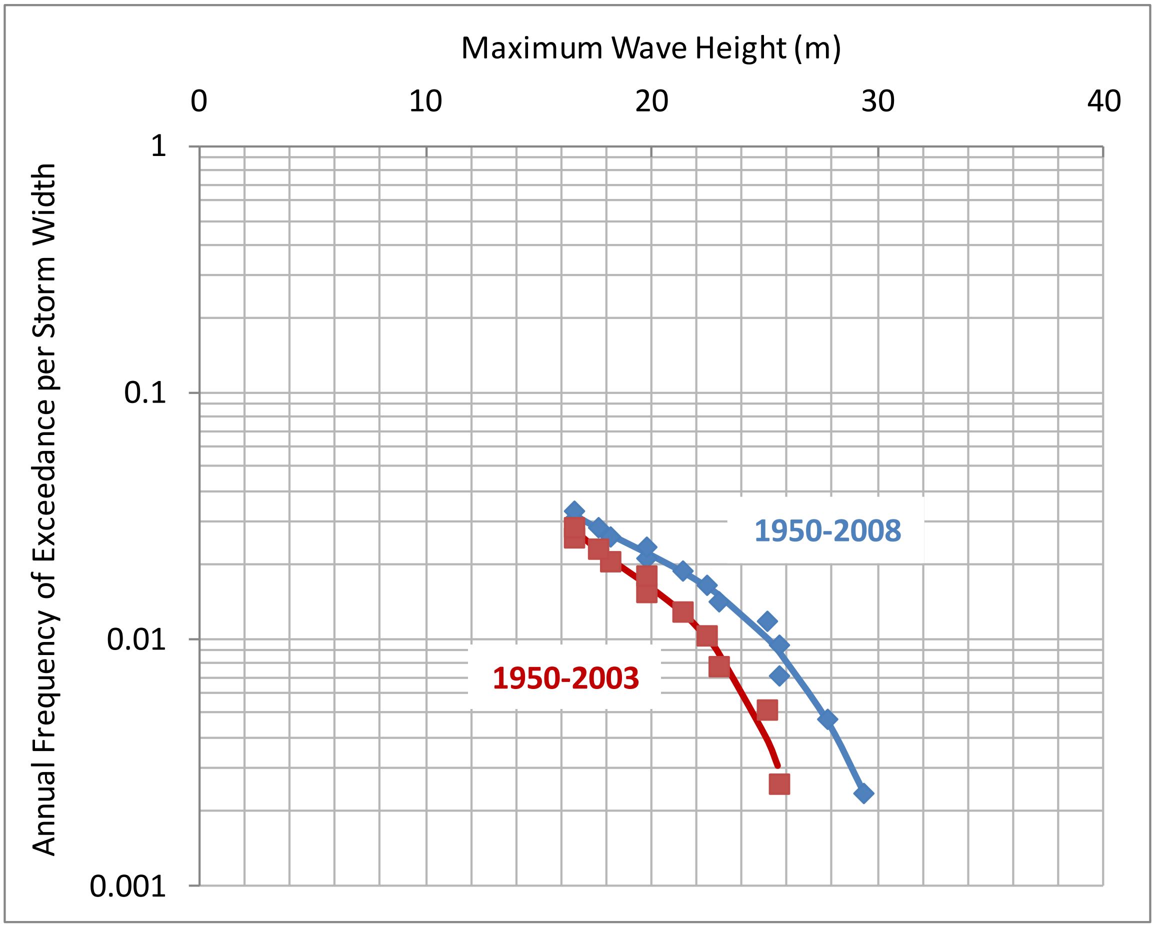

Figure 6.1. Data for wave heights from hurricanes in the central Gulf of Mexico

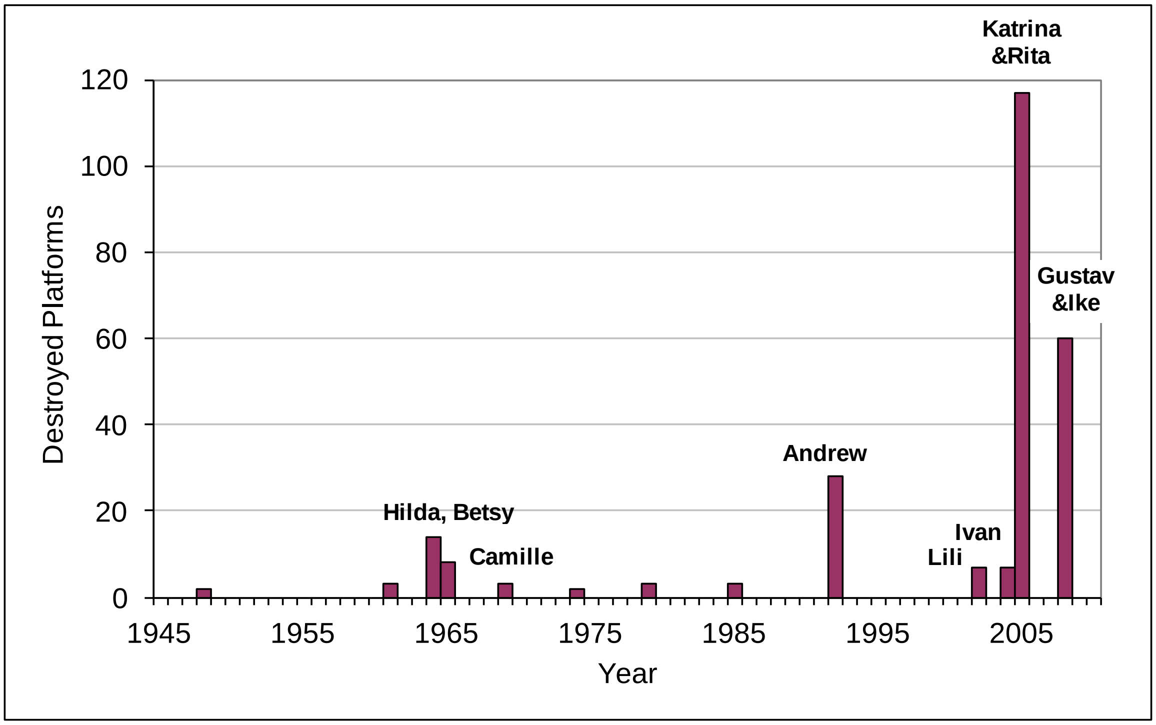

Figure 6.2. Impact of hurricanes on energy production facilities in Gulf of Mexico

The purpose of this paper is to propose a mathematical theory, Decision Entropy Theory, as a logical and defensible means to address extreme uncertainty. The basis for the theory is first described, and then an illustrative example is presented to demonstrate its use and significance.

plays are production opportunities with the potential for considerable uncertainty in production and economics. While trying to reduce production uncertainty with well tests may be valuable, there is a tradeoff between the quantity and quality of the information and its potential to improve development decisions. This paper proposes a new approach to represent uncertainty in an objective and defensible way, such that rare but important possibilities are not neglected in making decisions. This approach is called Decision Entropy Theory.

6.2 DEFINITION OF PROBLEM

The motivation for developing Decision Entropy Theory is that it is impossible to construct a sample space (the set of all possibilities) without knowing the set of all possibilities. That is, we cannot explicitly account for a black swan if we do not know it is possible. To demonstrate mathematically, consider an event of interest, F = failure of a civil engineering system, and an initial or prior sample space S that contains the set of all possible loading conditions on our system based on our experience. Let's say that subsequently a loading condition, E, occurs that is outside our prior sample space because it was outside our base of experience (Figure 6.3). There are two implications of this situation. First, the updated probability of E given that it has occurred is not defined since $P(E|ES) = \frac{P(E|ES)P(E|S)}{P(E|S)} = \frac{0}{0}$, which is nonsensical since E has occurred. Second, the probability of failure would have been under assessed if $P(F|S) < P(F|E)$ (Figure 6.3).

Figure 6.3. Example Venn diagram for sample space, S, that neglects the possibility of event E.

The prior sample space can be updated with available information, denoted the event I, using Bayes' Theorem:

$$P(E|IS) = \frac{P(I|ES)P(E|S)}{P(I|S)} \tag{6-1}$$

where $P(I|ES)$ is the likelihood of observing the available information given E and $P(E|S)$ is the prior probability of the event E. The prior probability of the event E is not conditioned on information. While these probabilities are non-informative, they are important because they affect the updated probabilities as information is incorporated.

To demonstrate the significance of prior probabilities, consider assessing the annual frequency that the wave height will exceed 30 m based on the data in Figure 6.1. If wave height exceedances follow a Poisson Process, then the likelihood of observing the available information (no exceedances in the past 60 years) is shown as a function of the mean rate of occurrence, ν_{30}, in Figure 6.4. Two alternative prior probability distributions for ν_{30} are shown in Figure 6.5: a uniform distribution for ν_{30} from zero to infinity (a diffuse prior) and a uniform distribution for log (ν_{30}) from zero to infinity. Figure 6.6 shows the significance of the prior probability distribution in the updated or posterior probability distribution; two very different decisions might result from one versus the other updated probability distribution.

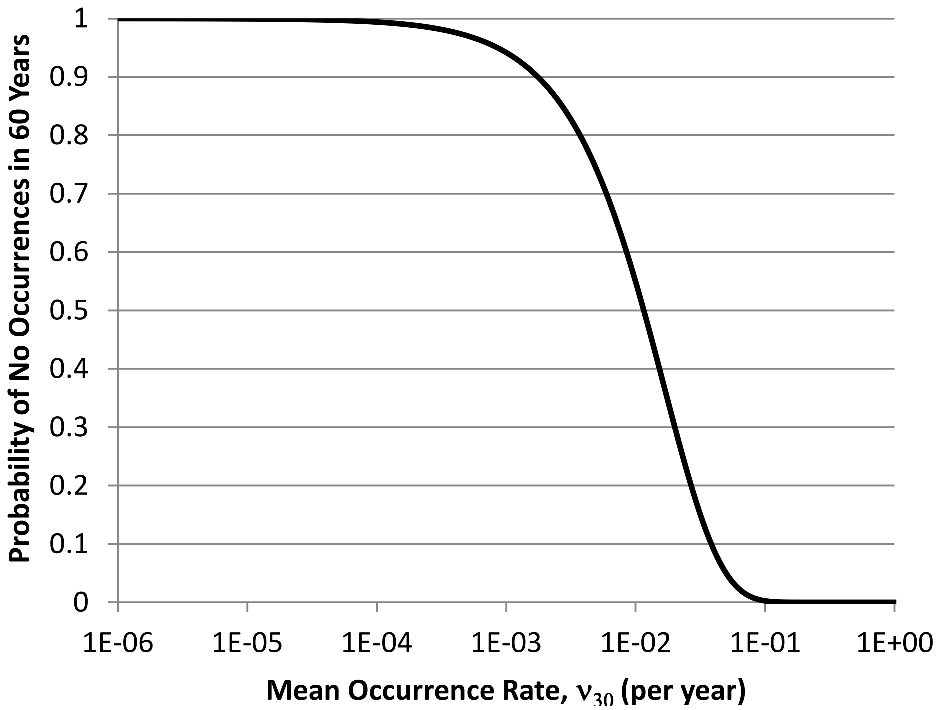

Figure 6.4. Likelihood function for mean occurrence rate of maximum wave height exceeding 30 m given no occurrences in 60 years

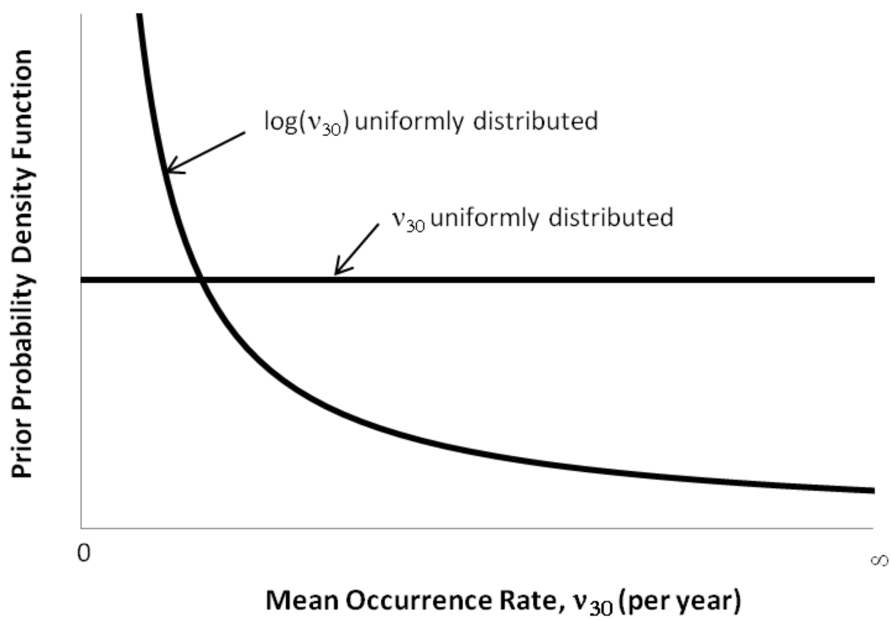

Figure 6.5. Alternative prior probability density functions for mean occurrence rate of maximum wave height exceeding 30 m

There is no theoretical, consistent and rational basis for either of the assumed prior probability distributions in Fig. 5 (e.g., Keynes 1921). Information theorists, such as Jaynes (1957), argue that the uniform distribution for ν_{30} maximizes the entropy for this variable. However, that argument does not address why we would maximize the entropy of ν_{30} versus $\log(\nu_{30})$, since the mean occurrence rate varies over many orders of magnitude and the probability of no occurrences in an interval t is $e^{-\nu_{30}t}$. Bayesian decision analysts, such as Luce and Raiffa (1957), argue that there is always some information, so the starting point in Bayes' theorem is a subjective prior probability distribution that implicitly contains some information. However, that argument does not address that our "information" is limited and may inadvertently neglect the Black Swan. Classical statisticians, such as Fisher (1935), argue that there is no defensible basis for a non-informative prior, that Bayes' theorem is not of practical use, and that we can only establish the likelihood function and not the updated probability distribution for ν_{30}. However, that argument is not helpful because information about the probabilities for different values of the mean occurrence rate is needed to make design decisions.

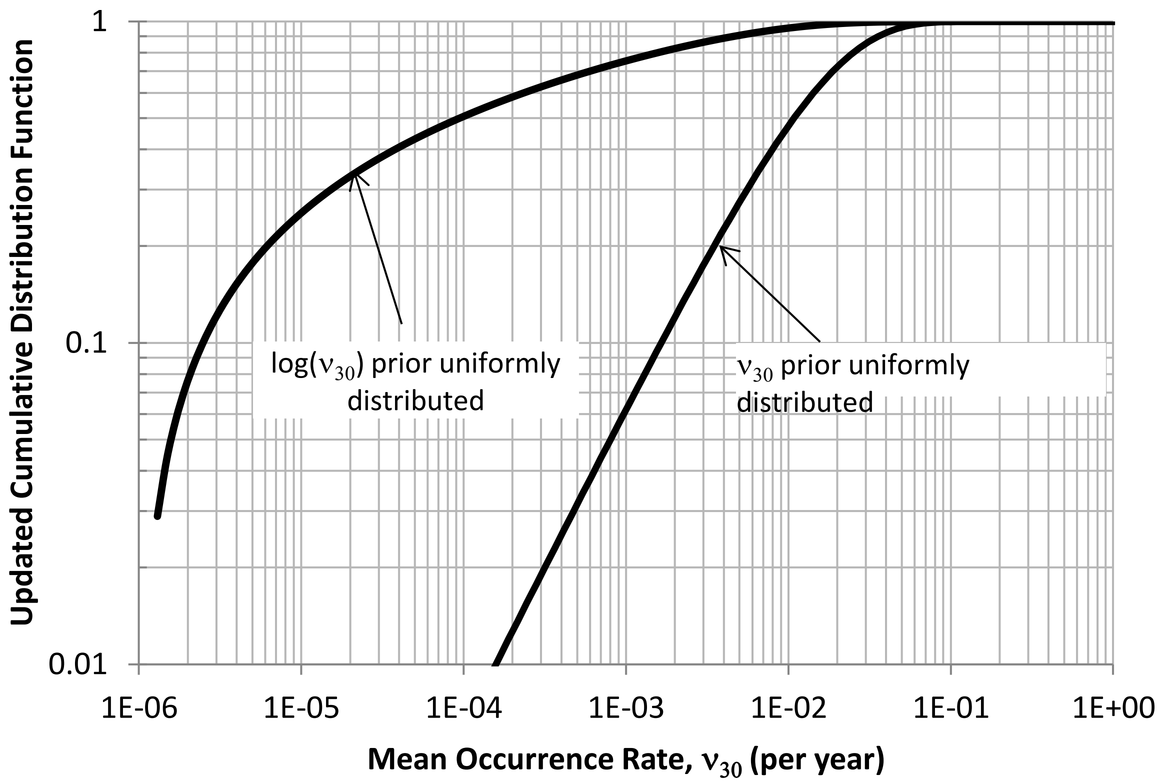

Figure 6.6. Alternative updated cumulative distribution functions for mean occurrence rate of maximum wave height exceeding 30 m given no occurrences in 60 years

6.3 DECISION ENTROPY THEORY

The objective of Decision Entropy Theory is to establish a prior sample space that is inclusive of all possibilities and reflects non-informative probabilities.

6.3.1 Set of all Possibilities

An inclusive prior sample space, S, is obtained in Decision Entropy Theory by characterizing it with respect to a decision and the possible outcomes from that decision. In this way, the set of all possibilities is known a priori. For a decision problem with two alternatives, the set of all possibilities is as follows:

1) The decision can be right or it can be wrong. If decision alternative A_1 is selected and compared to another alternative that could have been selected, A_2, then there are two possibilities: either A_1 is preferred to A_2 ($A_1 \succ A_2$) or A_1 is not preferred to A_2 ($A_1 \preccurlyeq A_2$).

2) If a decision alternative is selected, a particular value of utility can be gained or lost in comparison to having selected another alternative. If A_1 is selected and preferred to A_2, then all possible positive values of utility difference between A_1 and A_2, $\Delta u_{1,2}$ where $\Delta u_{1,2} > 0$, characterize the set of possibilities in this subset of the sample space. Likewise, if A_1 is selected and not preferred to A_2, then all possible non-positive values of utility difference between A_1 and A_2, $\Delta u_{1,2}$ where $\Delta u_{1,2} \leq 0$, characterize the set of possibilities in this subset of the sample space.

6.3.2 Non-Informative Probabilities

Non-informative probabilities for a decision comparing two alternatives, A_1 and A_2, are established in Decision Entropy Theory from two axioms that characterize the case of "no information:"

1) *The selected alternative is equally probable to be preferred or not compared to another alternative*: $P(A_1 \succ A_2) = P\left(A_i \preccurlyeq A_j\right) = 1/2$.

2) *If the selected alternative is preferred to another alternative, the possible gains in utility are equally probable:* $P\left(\Delta u_{1,2} \middle| A_1 \succ A_2\right) = 1/\mathrm{n}_{\Delta u_{1,2}|A_1 \succ A_2}$, *where* $\mathrm{n}_{\Delta u_{1,2}|A_1 \succ A_2}$ *are the number of possible values for* $\Delta u_{1,2}$ *where* $\Delta u_{1,2} > 0$.

Likewise, if the selected alternative is not preferred to another alternative, the possible non-gains in utility are equally probable: $P(\Delta u_{1,2} | A_1 \preccurlyeq A_2) = 1/\mathrm{n}_{\Delta u_{1,2}|A_1 \preccurlyeq A_2}$ *where* $\mathrm{n}_{\Delta u_{1,2}|A_1 \preccurlyeq A_2}$ *are the number of possible values for* $\Delta u_{1,2}$ *where* $\Delta u_{1,2} \leq 0$.

These axioms maximize the entropy for the possible states of a decision, meaning that a lack of information is equivalent to an equal or uniform probability of realizing any possible decision outcome.

6.3.3 Implementation of Utility Theory

Decision Entropy Theory establishes non-informative prior probabilities for differences in utility between possible decision alternatives, $P(\Delta u_{1,2})$. The expected utility difference between the alternatives is calculated as follows:

$$E(\Delta u_{1,2}) = \sum_{\mathrm{n}_{\Delta u_{\mathrm{i,j}}}} \Delta u_{1,2} P(\Delta u_{1,2}) \tag{6-2}$$

where $\mathrm{n}_{\Delta u_{1,2}} = \mathrm{n}_{\Delta u_{1,2}|A_1 \succ A_{j2}} + \mathrm{n}_{\Delta u_{1,2}|A_1 \preccurlyeq A_2}$. Following Utility Theory, the decision alternative that is preferred will have the largest expected utility difference [note that $E(\Delta u_{1,2}) = -E(\Delta u_{2,1})$. For a decision with n_A alternatives, preference is measured by the expected utility difference between each alternative and all other possible alternatives:

$$E(\Delta u_{\mathrm{i,all\ j}}) = \sum_{j=1}^{n_A} E(\Delta u_{\mathrm{i,j}})(1/n_A) \tag{6-3}$$

where $E(\Delta u_{\mathrm{i,i}}) = 0$. The preferred Alternative i has the largest value of $E(\Delta u_{\mathrm{i,all\ j}})$.

6.4 RELIABILITY-BASED DESIGN EXAMPLE

The current design basis for offshore platforms in the Gulf of Mexico is in question given the recent data from hurricanes (Figure 6.1). The maximum wave height used for design is designated h_D. The platforms are designed with a factor of safety of 1.5. Since the load is approximately proportional to the square of the wave height, the capacity of the platform corresponds to a wave height that is $\sqrt{1.5}$ times the design value. For simplicity, it will be assumed that failure will occur if the maximum wave height exceeds the ultimate design capacity, $h > \sqrt{1.5}h_D$, where $\sqrt{1.5}h_D$ will be referred to as the ultimate design capacity wave height. The average annual cost for operating a platform, $C(h_D)$, is simplistically given by the following:

$$C(h_D) = C_I(h_D) + (1 - e^{-\nu_D})C_F \tag{6-4}$$

where $C_I(h_D)$ is the annualized implementation cost, C_F is the cost if the maximum wave height exceeds the capacity, and ν_D is the annual rate that the wave height will exceed the ultimate design capacity, i.e., the annual rate for $h > \sqrt{1.5}h_D$.

A decision tree for this problem is shown in Figure 6.7. The annualized implementation cost is assumed proportional to the design load or to the square of the design wave height, $C_I(h_D) \propto h_D^2$. The existing state of practice corresponds to a design maximum wave height of 22 m, which will be designated h_{D*} where $h_{D*} = 22\,m$. The cost of any alternative design capacity wave height can then be normalized by the annualized implementation cost for the existing state of practice:

$$\frac{C(h_D)}{C_I(h_{D*})} = \left(\frac{h_D}{h_{D*}}\right)^2 + (1 - e^{-\nu_D})\frac{C_F}{C_I(h_{D*})} \quad (6\text{-}5)$$

The utility for each alternative ultimate design capacity wave height is set equal to the negative of the normalized annualized cost, $-\frac{C(h_D)}{C_I(h_{D*})}$.

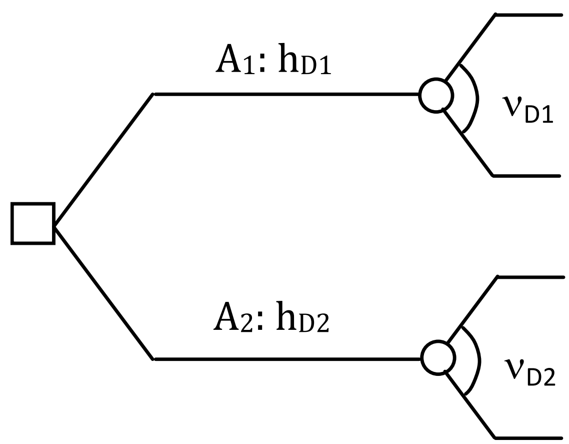

Figure 6.7. Decision tree comparing alternatives for design wave heights

6.4.1 Non-Informative Sample Space for Wave Heights

The non-informative sample space for wave heights, specifically the annual rate that the ultimate capacity wave height will be exceeded or ν_D, will depend on the decision alternatives being compared. Consider a comparison between maintaining the existing practice, $h_{D1} = h_{D*} = 22\,m$, versus increasing the design wave height to 26 m, $h_{D2} = 26\,m$. The difference in utility values between these two alternatives, $\frac{-C(h_{D1})}{C_I(h_{D*})} -$

$\frac{-C(h_{D2})}{C_I(h_{D*})}$, is plotted versus ν_{D1} and ν_{D2} in Figure 6.8 for a normalized failure cost of $\frac{C_F}{C_I(h_{D*})} = 20$.

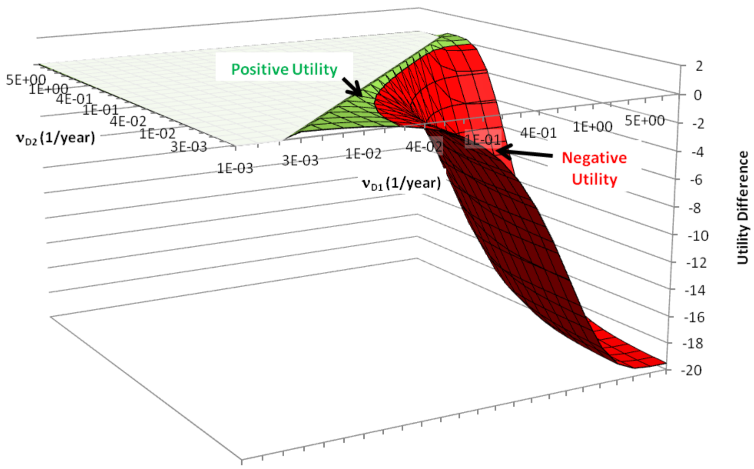

Figure 6.8. Difference in utility values for two alternative design wave heights, h_{D1} = 22 m and h_{D2} = 26 m, versus the annual frequencies that the platform capacity is exceeded for each design wave height

There are several important features in Figure 6.8. First, combinations where ν_{D2} is greater than ν_{D1} are not possible since $h_{D2} > h_{D1}$. Second, the combinations of (ν_{D1}, ν_{D2}) where the utility difference is greater than zero are combinations for which design wave height h_{D1} is preferred to h_{D2}, while combinations of (ν_{D1}, ν_{D2}) where the utility difference is less than or equal to zero are combinations for which h_{D1} is not preferred to h_{D2}. Third, the utility difference approaches asymptotic values at the corners: the utility difference approaches $-\left(\frac{h_{D1}}{h_{D*}}\right)^2 + \left(\frac{h_{D2}}{h_{D*}}\right)^2$ when $\nu_{D1} = \nu_{D2}$, while it approaches $-\left(\frac{h_{D1}}{h_{D*}}\right)^2 + \left(\frac{h_{D2}}{h_{D*}}\right)^2 - \frac{C_F}{C_I(h_{D*})}$ when ν_{D1} approaches infinity and ν_{D2} approaches zero.

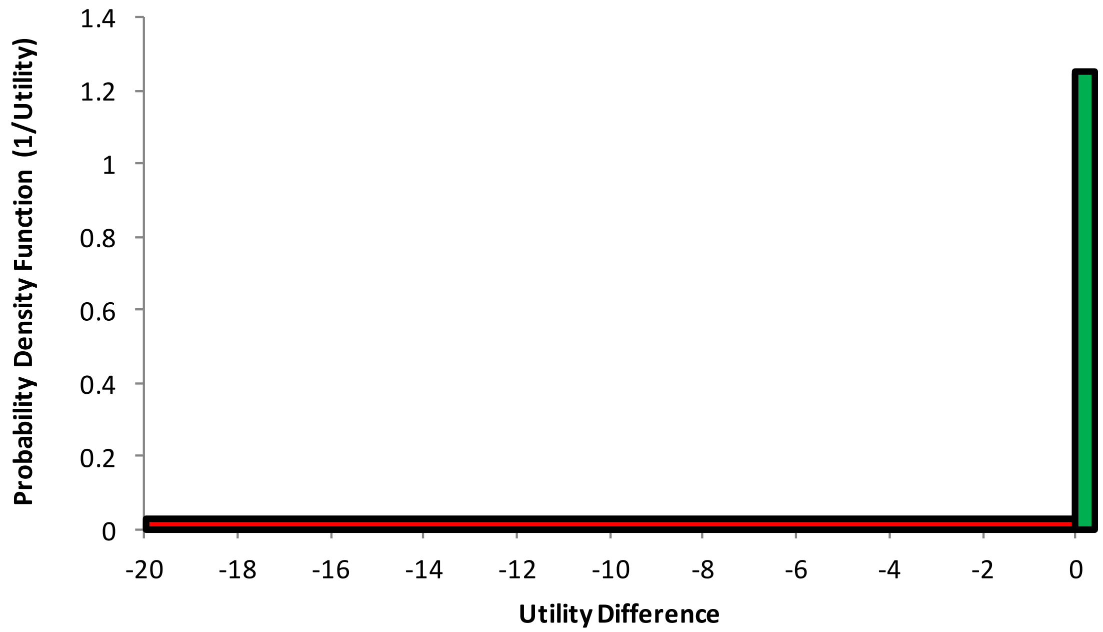

Figure 6.9. Non-informative sample space for utility differences in decision between two alternative design wave heights, h_{D1} = 22 m and h_{D2} = 26 m

The non-informative sample space for all possible combinations of (ν_{D1}, ν_{D2}) is constructed following the axioms of Decision Entropy Theory such that the probability that an alternative will be preferred is equal (50 percent for each alternative) and the possible values for the utility differences are equally probable (Figure 6.9). The resulting sample space is shown in Figure 10. Due to the complexity of the relationship between utility difference and (ν_{D1}, ν_{D2}), the non-informative prior probabilities were calculated by discretizing (ν_{D1}, ν_{D2}) into small areas. This non-informative sample space reflects the joint utility function versus (ν_{D1}, ν_{D2}) (Figure 6.8) in that there is a ridge dividing the combinations of (ν_{D1}, ν_{D2}) where one alternative would be preferred over the other (Figure 6.10) and the slope of the prior is inversely proportional to the slope of the utility difference versus (ν_{D1}, ν_{D2}), leading to spikes in the prior probability distribution where there are plateaus in the utility difference versus (ν_{D1}, ν_{D2}) (Figure 6.10). The non-informative prior probability distribution is not uniform and it contains complex, high-order relationships between ν_{D1} and ν_{D2}.

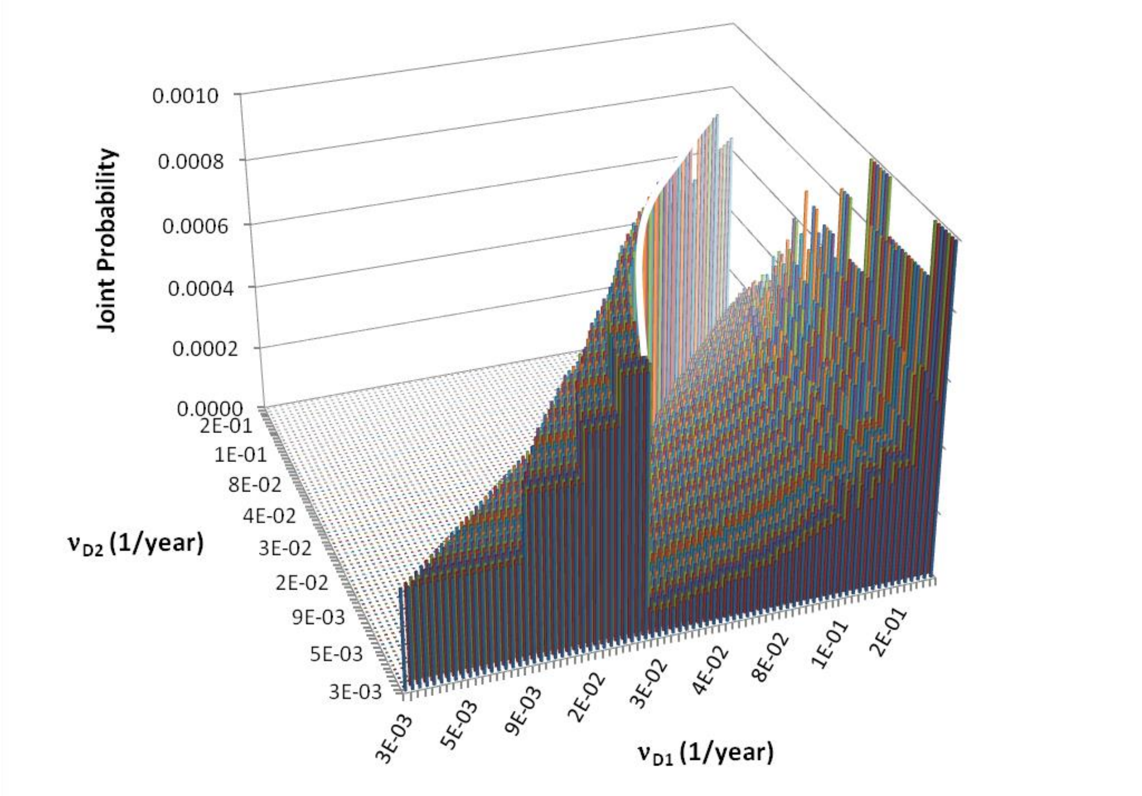

Figure 6.10. Non-informative sample space (prior probability distribution) for annual frequencies that the platform capacity is exceeded, (ν_{D1}, ν_{D2}), in decision between two alternative design wave heights, h_{D1} = 22 m and h_{D2} = 26 m

6.4.2 Updated Sample Space with Historical Wave Height Data

The likelihood function for combinations of (ν_{D1}, ν_{D2}) is shown in Figure 6.11 based on the 60-year long historical record (Figure 6.1). This likelihood function is developed assuming that exceedances of h_{D1} and h_{D2} each follow a Poisson process with a mean occurrence rate of ν_{D1} and ν_{D2}, respectively, and that the events of exceeding h_{D2} given that h_{D1} has been exceeded are independent:

$$P(x_1\ exceedances\ of\ h_{D1}\ and\ x_2\ exceedances\ of\ h_{D2}\ in\ 60\ years)$$
$$\cong \left[\frac{60!}{x_1!\,(60-x_1)!}{\nu_{D1}}^{x_1}(1-\nu_{D1})^{60-x_1}\right]$$
$$\times \left[\frac{x_1!}{x_2!\,(x_1-x_2)!}{\nu_2}^{x_2}(1-\nu_{D2})^{x_1-x_2}\right] \quad for\ h_{D1}$$
$$< h_{D2}\ and\ \ \nu_{D2} \le \nu_{D1} \tag{6-6}$$

where this approximation is most reasonable for small values of ν_{D1} and ν_{D2}, less than about 0.1 per year.

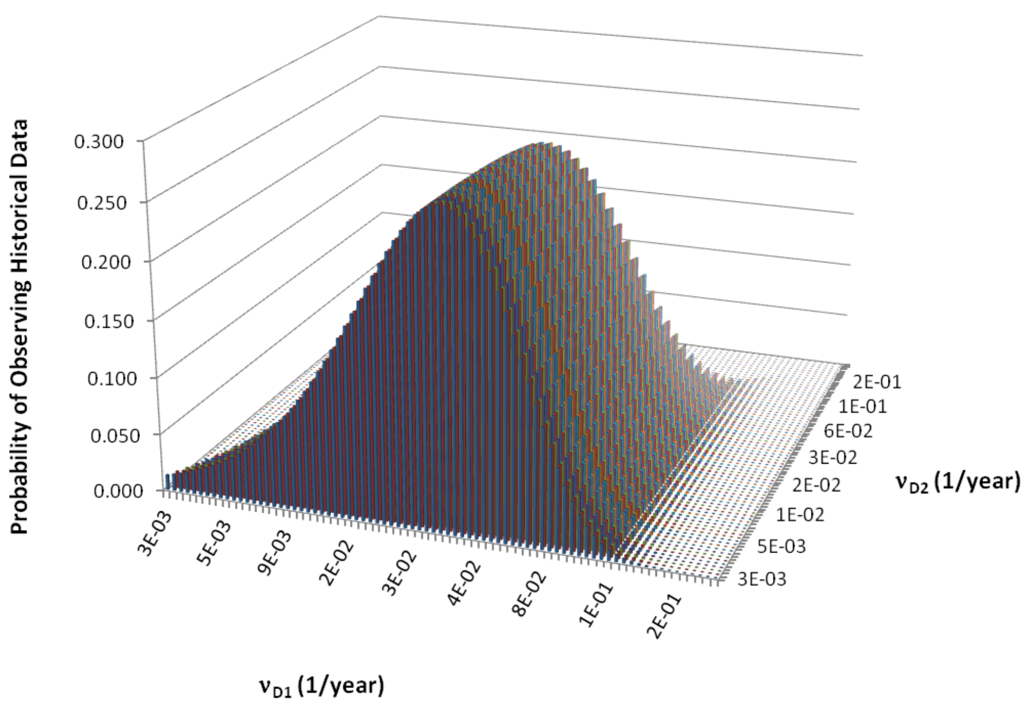

Figure 6.11. Likelihood function for combinations of (ν_{D1}, ν_{D2}) based on historical record for the number of exceedances of h_{D1} = 22 m and h_{D2} = 26 m in the past 60 years.

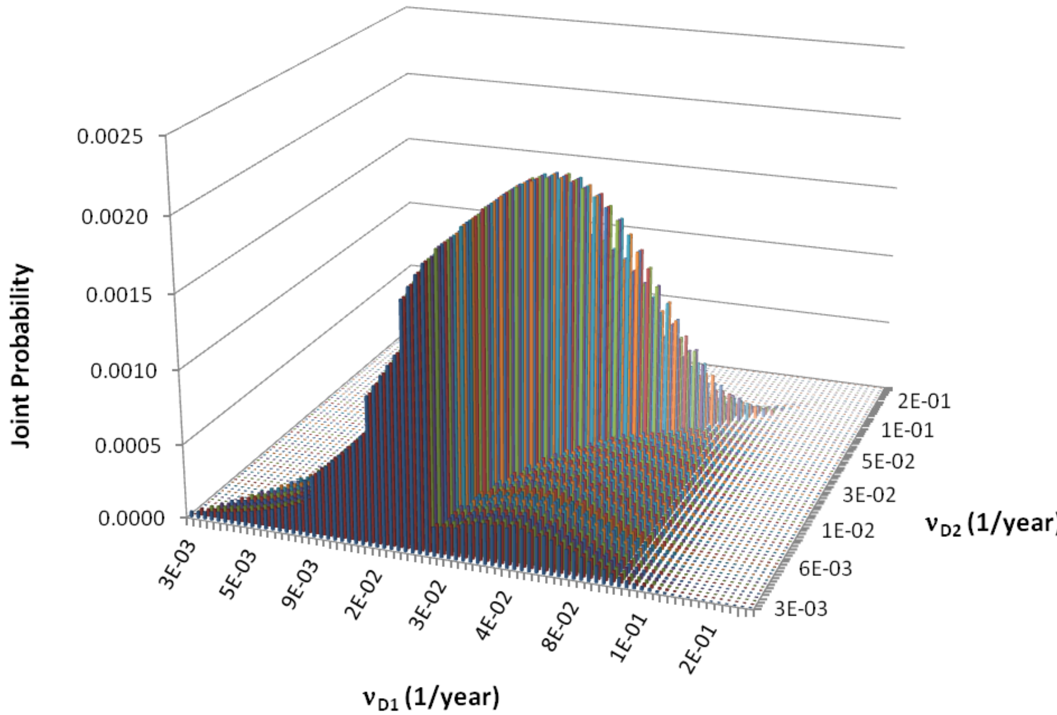

Figure 6.12. Updated sample space (posterior probability distribution) for annual frequencies that the platform capacity is exceeded, (ν_{D1}, ν_{D2}), given historical data in decision between two alternative design wave heights, h_{D1} = 22 m and h_{D2} = 26 m.

The updated sample space (posterior probability distribution) for all possible combinations of (ν_{D1}, ν_{D2}) given the historical data, obtained through Bayes' theorem, Eq. (6-1), is shown in Figure 6.12. This updated distribution reflects both the non-informative prior sample space (Figure 6.10) as well as the historical data (Figure 6.11). The ridge in the prior probability distribution dividing the regions where one alternative is preferred over the other is still clearly visible in the posterior probability distribution (Figure 6.12). In addition, the posterior probability distribution emphasizes the possibilities where the rate of exceedance for the conventional design, ν_{D1}, is similar to that for the alternative design, ν_{D2}, (in which case the conventional design is preferred) and possibilities where ν_{D1} is substantially larger than ν_{D2} (in which case the alternative design is preferred).

For the updated sample space (Figure 6.12), the expected utility difference between h_{D1} = 22 m and h_{D2} = 26 m is +0.05, indicating that the ultimate design capacity wave height of 22 m is preferred compared to the ultimate design capacity wave height of 26 m for this normalized failure cost of $\frac{C_F}{C_I(h_{D*})} = 20$.

6.4.3 Results of Decision Analysis for Design Wave Height

The expected utility differences between the standard of practice, h_{D1} = 22 m, and a variety of alternatives for h_{D2} are shown for different normalized failure costs in Figure 6.13. Note that each possible combination of a failure cost and an alternative design wave height corresponds to a different prior sample space for of ν_{D1} and ν_{D2} (e.g., Figure 6.10 shows it for $\frac{C_F}{C_I(h_{D*})} = 20$ and h_{D2}=26 m). When the expected utility difference between the standard of practice and an alternative design is positive, the standard of practice is preferred; when the expected utility difference is negative, the alternative is preferred. The shapes of the curves for expected utility difference versus the alternative design wave height reflect the effect of the historical data; the numbers of historical events where different design ultimate capacity wave heights have been exceeded, , $h > \sqrt{1.5}h_D$, are identified in Figure 6.13.

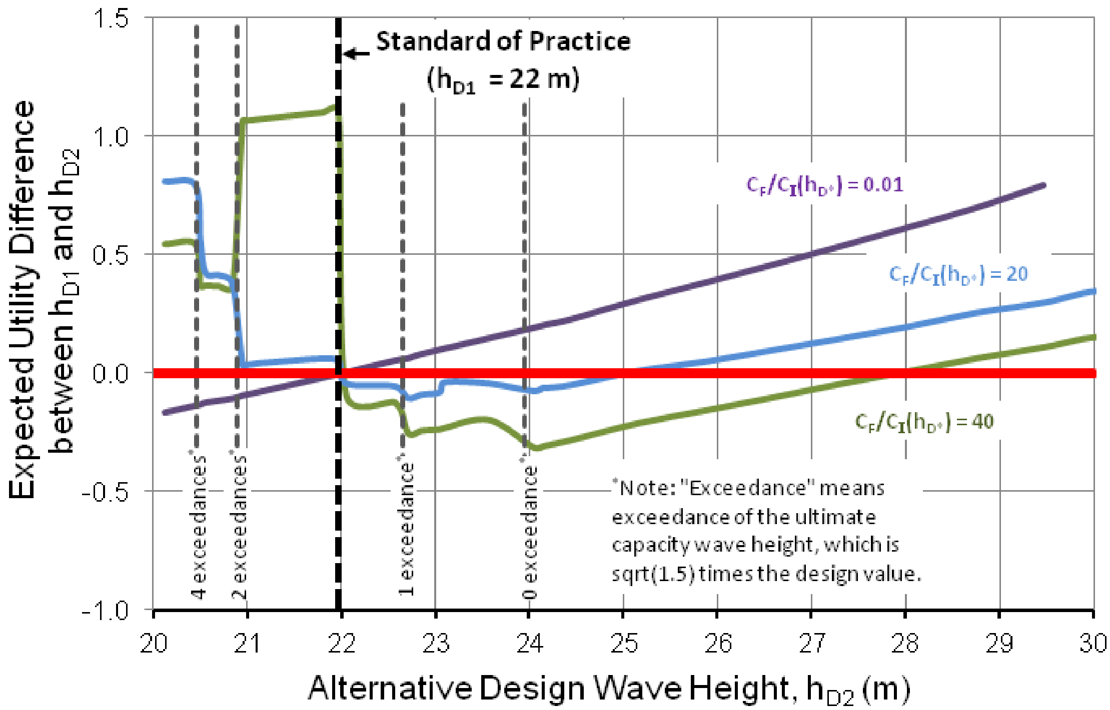

Figure 6.13. Expected utility difference for standard of practice (h_{D1} = 22 m) compared to a range of alternative values for the ultimate design capacity wave height, h_{D2}.

For a small failure cost, $\frac{C_F}{C_I(h_{D*})} = 0.01$, the alternative with the lower implementation cost is always preferred because the expected cost of failure is negligible for all alternatives (i.e., the expected utility difference is negative for $h_{D2} < h_{D1}$ and positive for $h_{D2} > h_{D1}$ in Figure 6.13). For the larger costs of failure, $\frac{C_F}{C_I(h_{D*})} = 20$ and $\frac{C_F}{C_I(h_{D*})} = 40$, the standard of practice is preferred to alternatives with lower design wave heights from 20 to 22 m, while the alternative design wave heights greater than the standard are preferred up to a point (Figure 6.13) due to the contribution of the risk of failure to the expected utility. However, as the alternative design wave height continues to increase, the expected utility difference eventually crosses zero and the standard of practice is preferred to the alternative due to the increasing cost of implementation for higher design wave heights. This threshold corresponds to an alternative design wave height of about 25 m for $\frac{C_F}{C_I(h_{D*})} = 20$ and to an alternative of about 28 m for $\frac{C_F}{C_I(h_{D*})} = 40$ (Figure 6.13).

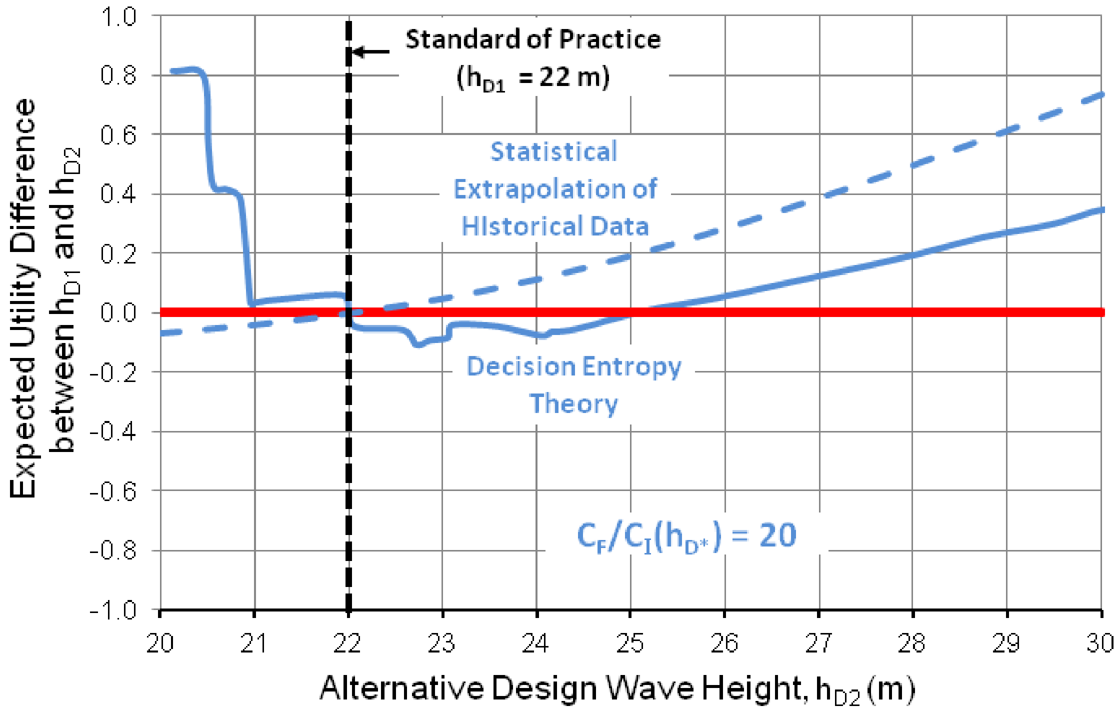

Figure 6.14. Comparison of results from Decision Entropy theory with those from a statistical extrapolation of the historical data.

The results from Decision Entropy Theory are compared in Figure 6.14 with those from a statistical extrapolation of the historical data (i.e., estimating ν_D as a function of $\sqrt{1.5}h_D$ from the trend-line in Figure 6.1). The two approaches give significantly different results for this cost ratio, $\frac{C_F}{C_I(h_{D*})} = 20$. Decision Entropy Theory leads to using a more conservative design wave height (i.e., h_{D1} is preferred for $h_{D1} > h_{D2}$ and h_{D2} is preferred for $22\,m < h_{D2} < 25\,m$), while a statistical extrapolation of the data leads to the opposite conclusion (Figure 6.14). Therefore, in this particular example, the establishment of a non-informative prior probability distribution for the mean occurrence rates of wave heights using Decision Entropy Theory puts greater emphasis on the possibility of large waves (e.g., see Figure 6.10 and Figure 6.12) that have not (yet) occurred in the historical record and results in a more conservative design decision. Note that the two approaches will ultimately lead the same result when an infinite amount of data is available to estimate wave height frequencies.

In order to identify the optimal value for the ultimate design capacity wave height, all of the various alternatives (including and excluding the standard of practice) are compared in a pair-wise fashion, see Eq. (6-3). The results for this design example indicate that for large costs of failure, $\frac{C_F}{C_I(h_{D*})}$ greater than about 50, the preferred value for the ultimate design wave height is just above 24 m or about 2 m greater than the standard of practice. The significance of 24 m is that it corresponds to an ultimate capacity wave height of $\sqrt{1.5}(24\,m) = 29.5\,m$, which is the largest event that has occurred in the historical record (see Figure 6.1). This result highlights the importance of an adaptable design approach in the face of extreme uncertainty. The preferred design wave height is sensitive to the historical data because it is governed by the largest observed wave height, meaning that the preferred design wave height will change with time and should not be considered as a static value in design codes. In addition, existing structures designed using lower wave heights may need to be upgraded over their design lives to increase their capacity based on new information.

6.5 CONCLUSIONS

Decision Entropy Theory is proposed here as a logical and defensible means to address extreme uncertainty, the possibility of events that are beyond the range of experience. The proposed theory provides a basis to establish a non-informative prior sample space for uncertain variables in a decision. In the non-informative prior sample space, there is an equal probability of being right or wrong in making a decision, and there is an equal probability of realizing the range of possible losses or gains from that decision. This non-informative sample space is then updated with any available information through Bayes' theorem.

An illustrative example for selecting the appropriate design wave height for offshore structures highlights the following points:

- The prior sample space depends on the decision, meaning that the importance of extreme uncertainty depends on its consequences to the decision and the availability of feasible decision alternatives to compensate for these consequences.
- The prior sample space emphasizes the possibilities that distinguish two alternatives from one another, such as differences in the mean occurrence rates for the wave heights that will cause failure for different design alternatives.
- The prior sample space can affect a design decision, even when substantive data are available to inform (update) this sample space. In this example, a statistical extrapolation from 60 years of data for wave heights would lead to a less

conservative decision than is obtained by incorporating these data within the context provided by Decision Entropy Theory.

- Adaptability is a key principle in achieving reliability in the face of extreme uncertainty. This example demonstrates that the optimal value for the design wave height can be sensitive to the available data, meaning that design codes and design products that can be readily adapted based on new information or data are needed.

Chapter 7 Conclusions and Suggestions

After presenting the conclusions from the work, a set of suggestions are being offered to further develop the investigated topics in the future.

7.1 CONCLUSIONS

The following summarizes the outcomes of the study in relation to the intended objectives of the investigation, as described in the introduction to the dissertation.

7.1.1 Philosophical Foundations of Probability

The first objective of the study was to establish the philosophical grounds for the probabilistic computational theory under development, and to provide the theory with clear and precise conceptual foundations. The importance of conceptual foundations is due to the irreconcilable divergences among opposing assessment methods, who all use measures termed as probability, but in different senses. Clarifying the sense of probability in a procedure not only makes the method more meaningful to its users, but more relevant and beneficial to theoreticians of other procedures, who may then see the method as plausible and incorporate the procedure and/or its technical innovations in their own developments.

The review and the evaluation of the existing philosophical positions regarding probability revealed that while philosophers have been able to develop a number of senses for the term probability, they have not succeeded in presenting a comprehensive account of probability and in assisting others to benefit from the assortment of probabilistic methods without compromising consistency and meaningfulness. Since various probability interpretations, including classic, frequency, propensity, logical, and subjective, have remained unconnected, many users who apply procedures developed by various schools of thought, usually do so in an instrumentalist approach and without considerations for the consistency of the senses implicit in those methods.

To present a comprehensive sense for the concept of probability, I embarked in a philosophical inquiry aimed at finding a more fundamental concept that can connect the diverging senses. While the sought-for concept may not yield a univocal definition of probability, it may clarify the commonalities and the differences among various senses in its own terms. As a result of the inquiry, I came to find "process" as the most fundamental concept, which can play the desired role of conceptual integration by seeing the senses of probability from a process-centered perspective.

I divided probability senses into three types of propensity, possibility, and credibility, and gave a process-centered definition for each type. While propensity and possibility values make empirically-refutable assertions about certain natures of a process, as observed or postulated to be so, credibility values are (means for) assessment of the first two, and result from a humanly devised inductive process, established to reflect our lack of knowledge. I distinguished Bayesian priors as prime examples of credibilities.

As the focus of the probabilistic computational theory under development is Bayesian priors, I directed the investigation toward the nature and the character of credibilities. By declaring credibility values to be the outcomes of human imagination and non-perceivable, I concluded that no subjective credibility assignment can be refuted by empirical observations. Nevertheless, I warned that credibilities assigned by a subject might not be acceptable to others and if the aim of the induction is to provide with admissible assessments, the subject must use an agreed-upon assessment process including approved credibilities.

7.1.2 Philosophical Foundations of Objective Mathematical Assessment of Actions

While the study was not originally aimed at developing a general epistemological theory that elaborates on the subjects of human understanding, along the natural progression of my inquiry, I came to the conclusion that it is very difficult to establish clear and precise philosophical foundations for probability without addressing a suite of concepts more fundamental to human understanding. To explain the philosophers' failure to clarify the concept of probability, attested by declarations such as Bertrand Russell's "*Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means*," I entertained the possibility that the failure might be due to the absence of a more general philosophical theory. If philosophers could see the subjects of human knowledge from a different perspective, they could also see specific concepts such is probability differently.

My efforts to develop a new epistemological perspective was centered around the concept of process, which I had found to be of immense explanatory power but not much of recognition among philosophers. The explicative reliance on the process concept resulted in the gradual development of a number of positions on fundamental issues in philosophy, and provided support for my hypothesis that process can be the central pillar of a philosophical perspective. Due to the role of the process, I labeled the perspective *Philosophical processualism*, and I aimed it at advancing the foundations of mathematics, as it is being applied to assist human cognition and action. To highlight the role of math in action, and to distinguish it from its general application to non-mathematical problems, I collectively labeled such uses as *pragmatic mathematics*, and I defined the collection as a system of linguistically manifested mental constructs aimed at guiding purposive actions with precision.

By analyzing the fundamental contents of human mind, percepts and concepts, through the lenses of processes, I came to oppose the common Platonist position in cognition and in the philosophy of mathematics, which takes concepts in general, and mathematical concepts in specific, to be abstract non-spaciotemporal entities that are directly accessible to human minds in an objective and absolute manner. I proposed that a process-centered ontology of concepts puts them inside the mind and sees their formation only attainable through earthly spaciotemporal processes.

I developed a processual answer toward the metaphysical problem of universals, as I see it central to both science and mathematics, where science aims at finding relationships among universal entities (in either categories of objects and properties), and math aims at describing the presumed relationships in terms of imaginary mental entities. I explained the relationship between the instances of a universal, by introducing the concept of *processual code*, as the container of information on the specifics of objects and their properties, whose presence trigger and advance the process resulting in the exemplification of the universal phenomena. I also declared that ignoring the process underlying a universal and taking instances as mere observations results in a nominalist account, where the universal is just a name to a set. I attributed the nominalist denial of universals on their exclusive reliance on *experiencing-organizing* mind rather than *questioning-hypothesizing* mind. By highlighting the role of the mind in recognition of universals, I extended the concept of the code to percepts and saw them as outcomes of a joint two-world process, one outside the mind in the world and the other inside the mind. The processual perspective on percepts turned out to be in opposition of both Aristotelian realists, who believe that types are out there in the world independent of our minds, and conceptualists, who believed that universals do not exist outside of the perceiving mind.

Since the processual ontology of concepts, including the mathematical ones, takes them to be spaciotemporal, it cannot take conceptual objectivity as default and has to elaborate if and when the objectivity is present. To specify the conditions, I developed the processual perspective toward the manifestation of concepts, especially through linguistic descriptions. On the platform of manifestation, I offered a communal verification-based account of conceptual objectivity, in which a conceived entity by members of a community is considered to be objective, if the subjects consistently deliver similar manifestations of the concept in any number of agreed-upon verification processes.

Last but not least, I presented a processual account of a few fundamental mathematical concepts including numbers, sets, and infinity, and also presented and discussed the hypothesis that mathematical entities are grounded in both human percepts and cognitive processes. For the concept of infinite set, which is also central to probability as applied to processes, I declared the extensional definition, as the collection of its tokens, imprecise. Since the infinite set containing yet-to-be-generated tokens lacks accessibility due to the inclusion of entities without spaciotemporal specification, I suggested that tie define an infinite set to the processual code that can generate its tokens.

7.1.3 Principles for Objective Assessment of Unknown Probabilities

Another objective for the research was to define the philosophical and logical foundations within which, the characters “objective, transparent, and defensible” can be clarified or even defined for Bayesian priors. Based on the results of the inquiry presented in 7.1.1, I classified Bayesian probabilities as credibilties; a set of imaginary values assigned by the evaluating subject as a means of coming up with an assessment of the unknown outcome of a process. Due to imaginary nature of Bayesian priors, I declared that subjects can assign any set of self-consistent values as priors without being refuted by experiments. Nevertheless, feasibility does not bring legitimacy, and when the subject’s goal from assessing the uncertainty is to deliver evaluations that are admissible to others, he must use an agreed-upon assessment process using authorized credibility values.

The concept of objectivity for imaginary entities were investigated in 7.1.2, and resulted in a general definition, where objectivity is tied to the verifiability of the entity’s manifestation in accordance with a set of established processes. So, objectivity can be considered to be transparency plus community acceptance. In mathematical terms, objectivity translates into having a perceivable mathematical form as the code for the assignment process. Simply put, an objective set of credibility values is the one assigned with a perceptually transparent formulation.

Defensibility is another characteristic that along with transparency, can assist an imaginary entity to be considered objective by a community. While transparency is a necessary condition for the objectivity of an imaginary entity, it is not sufficient and must be accompanied with community acceptance. Although any agreed-upon communal rule does have a contractual nature, which is independent of the justifications supporting it, the presence of justification is a crucial factor in making the procedure defensible, if the community is deciding to adopt any of proposed process as the rule.

I proposed a set of theoretical principles to evaluate the defensibility of an inductive system. While the assertions are presented as principles and are accepted without relying on other assertions, they are consistent with a presumably common reasoning system among humans, which way may come to label as *intuitive logic*. The seven principles are named as (I) Evaluative Orientation, (II) Investigative Prioritization, (III) Explicative Sufficiency, (IV) Evaluative Inclusion, (V) Credibility Conception, (VI) Artifact-Reality Division, and (VII) Unbiased Evaluation.

7.1.4 Theoretical Decision-Based Assignment of Bayesian Probabilities

Decision Entropy Theory (DET) is being offered as a computational machinery aimed at assessing unknown aspects of a process in the light of the available information, where sought-for aspects depend on the processual parameters; (structural) values specifying collective or relative individual features of processual input and/or output and/or the (deterministic) input-output relation. The theory is intended to be objective and defensible in the sense described in 7.1.2 and 7.1.3, meaning that (a) its assessment procedures are transparent and imitable and (b) the constraints specifying the process are set in compliance with a set of intuitively logical guiding principles.

The theory assigns non-informative Bayesian priors to sets of possibilities, representing the aspects of the process that are unknown to us. The priors are assigned to possibilities by considering a hierarchy of variables and such that prior credibilities minimally inform the variables, whose exact values are the subjects of the inductive inquiry. The adopted measure for the information contained in a distribution is Shannon's entropy, where the minimally informative distribution is the one with the maximum entropy. In the simplest case of the hierarchy composed of a single variable, whose only constraint is its boundaries, the uniform distribution is the one with the maximum entropy.

The following principles specify the hierarchy of variables, as the targets of the inquiry, and elaborate the distributions that are minimally informative:

1. If no information is available about the probabilities of decision outcomes, then a selected alternative is equally probable to be or not to be the preferred alternative.

2. If no information is available about the probabilities of decision outcomes, then the possible differences in preference between a selected alternative and the preferred alternative are equally probable.

3. If no information is available about the probabilities of decision outcomes, then the possibilities of learning with new information about the selected alternative compared to the preferred alternative are equally probable.

7.1.5 Implications for the Application of the Theory to Practical Problems

Another objective for the research was to identify the practical consequences of applying the theory of Decision Entropy. To do so, the following two examples from engineering practice were being worked out:

(a) selecting the appropriate design wave height for offshore structures, and

(b) assessing the value of test wells before committing to developing a hydrocarbon resource play.

Application of DET to example problems highlighted that the theory results in a rational design, as it prioritizes answering to the questions that matter more to the engineering design. Since the main goal of the design is to find the specifics of the engineering system that can provide the intended service with the minimum life-cycle cost under the to-be-encountered environmental conditions, systemic variables are the most important ones, followed by cost variables, and eventually the environmental variables. By considering design objectives, DET prioritizes learning about the systemic variables over the cost variables over the environmental variables, and tries to inform the systemic variables the least, and then the cost variables, and the last the environmental variables. DET's approach is in contrast to the inadvertently irrational approach that starts with a biased design recommendation due to prioritizing the environmental conditions that matter least to the design.

The example applications also illustrated the adaptability of credibility-based inductive methods and the justifiability of DET, in accounting for the situations, in which the available information is not only partial, but also dubious. For example, when the information is gathered from the analogous processes who share some features with the target process, but also have some irreconcilable differences, the assessment procedure must determine whether and how to consider the data from analogs. While Credibility-based procedures are able to incorporate the dubious date into the assessment by assigning a degree of credence, they face the obstacle of defending their assigned value, as credibilities are imaginary by nature and cannot be justified on empirical ground. Nonetheless, DET can both assign a degree of credence, as well as defending its assignment based on principles grounded in intuitive logic.

7.1.6 Evaluation of the Compliance of the Decision-Based Theory with Established Principles

The final objective of the research was to evaluate whether the assessments of DET can be characterized as objective, transparent, and defensible. The comparison between the three principles of DET, stated in 7.1.3, with the seven general principles, stated in 7.1.4, reveals that DET is in compliance with the more fundamental principles for objective assessment of probabilities. Thus, we can conclude that DET leads to assessments that are objective, transparent, and defensible.

The details of the compliance are as follows: (I) DET satisfies “Evaluative Orientation,” as it takes inductive analysis to be a goal-oriented act and directs the assessment toward answering a set of target questions. (II) DET performs “Investigative Prioritization,” as it prioritized the questions according to their importance to the investigation. (III) DET is Explicatively Sufficient, as it designates a number of distinct processual explanations, such that each provide definitive answers to the target questions. (IV) DET is evaluatively inclusive, as it allows the formation of a hybrid and inclusive explanation by considering and combining collection of explanations. (V) DET is based on credibility conception, as it uses credibility values to form an inclusive assessment, and to differentiates various explanations with a degree of plausibility relative to other explanations. (VI) DET divides artifacts from reality, as its assigned credibilities explicitly account for (1) what we know, observations, and (2) what we arbitrarily devise to reason beyond what we know, explanation/ hypothesis, such that the contribution of (1) and (2) in the collective assessment is tractable. (VII) Finally, DET provides with unbiased evaluations, as its assigned credibility values are not such that the inclusive assessment favors any of the possible answers to the target questions, before incorporation of observations.

7.2 SUGGESTIONS

The following suggestions are being offered to advance the developments presented in the dissertation.

7.2.1 Developing Objective Methods to Incorporate Subjective Experts' Opinion

An objective method, in the sense described in the dissertation, can be developed to capture the subjective opinion of experts regarding the variability of the outcome of a process, within the framework of the Bayesian approach. Since we have already reserved one of the two components of the framework, namely Bayesian priors, to account for our ignorance regarding the hierarchy of variables we intend to assess, we must use the remaining component, namely the likelihood function, to capture all kinds of information including opinions expressed by experts. While the information obtained from empirical observations can be represented by means of standard frequency functions, no standard method currently exists for reflecting experts' opinion in likelihood function. The challenge will be to develop norms to grant precision to subjective opinions, which by nature are vague and imprecise.

7.2.2 Illustrating Probability Interpretations by Examples

The proposed collective sufficiency of the three interpretations of probability, namely possibility, propensity, and credibility can be illustrated through a number of examples. The exemplary works not only demonstrates the practicality of the developed theoretical methods in accounting for all available knowledge in every possible format, but also the sufficiency of the offered set of interpretations in making sense of the quantities incoming to and outcoming from the assessment process. In addition, examples of the use of methods other than the ones advocated here can be reworked to pinpoint the fact that all numeric measures of uncertainty can be interpreted based on the highlighted types.

7.2.3 Reevaluating the Multi-alternative Decisions using Information Potential

Since Decision Entropy Theory has evolved since the time the examples in the dissertation has been worked out, especially in regards to analyzing multiple alternative decisions and assessing the value of information, it is desirable to evaluate whether the theoretical evolutions do influence the results of the analyses. In specific, it is recommended that the assessments presented in Chapters 5 and 6 of the dissertation, which are based on the concept of the utility difference and the method of integrating pairwise assessments, reworked based on the new concept of information potential and the method of reference alternative.

7.2.4 Evaluating the Conditions for Increasing the Community Acceptance of DET

It is recommended to find out the conditions under which the theory of decision entropy may gain communitywide acceptance. A reason for the latency of the acceptance conditions is the fundamental nature of DET, which places it beyond the territory of applied statistics and at the core of every discipline where consequential risky decisions ought to be made based on limited data. In other words, since there is no established community for the kind of discourse Decision Entropy Theory presents, a strategy must to develop to select and to prioritize communities, among which DET can be introduced and popularized.

APPENDIX A: DEFINITIONS

1. INTRODUCTION

Definition of a term provides a description elaborating the meaning, nature, and scope for the word/phrase in a formal fashion. Since the use of technical terms are inevitable in any theoretical development, providing with a glossary of relevant terms immensely facilitates the communication of arguments and results. Below, are a number of fundamental definitions that I believe are essential to the developments in this dissertation.

The importance of definitions to philosophy is such that some philosophers believe that the core of a subjects ought to be found in its definitions (Baggini and Fossi, 2010). Historically, many significant philosophical questions in the form of "what is concept x?" were posed due to the lack of a clear definition and it has been the quest to find answers to such questions that eventually has resulted in providing a definition for that concept. Such processes imply that a definition for a (core) concept of a subject matter can only be written after (a full) investigation of the subject, when the investigator has reached the level of mastery.

The glossary is arranged in two section. While the first section gives definitions of various terms, independent of their pertinent topics, the second section provide definitions of terms given specific topics. The order of appearance for the terms in the first section is alphabetic.

2. Non-Categorized Definitions

Action: is a physical process, triggered and/or sustained by an agent.

Action, (Purposeful) Human: is a planned/thoughtful action taken by man in order to achieve a certain outcome.

A**ctor/Agent, Purposeful**: is an agent, who initiates some physical processes in order for agent's desired objects/states to be realized.

Algebraic Process: an arithmetic process, where some of the numbers involved are variables, representing specific numerical sets. The numerical outcome of an algebraic process can itself be represented by a variable as, each combination of the values for numerical variables in the algebraic process yield a different outcome.

Algebraic Parameter: a numerical component of an algebraic process that is fixed/constant; in contrast to a variable, which can be any of the members of a specific numerical set.

Algebraic Expression/Form: a visual representation of an algebraic process composed of different symbols as visual representations of variables, parameters, and numerical operations.

Algebraic Function: another term for an algebraic process or its outcome.

The set of numerical values, which involving (input) variables of the process can take their values from, are called the "domain" of the function of the set of output values are called the "range."

Asset: is a resource owned/possessed by an agent. The most immediate asset of a biological agent is his body and its organs.

Characteristic/Specification: is a distinguished trait, quality, or property of an object. The reliance of the given definition on synonyms coveys that the term is a primitive notion; i.e. humans, given their innate abilities, are able to distinguish certain aspects/features of worldly objects around them and are able to make distinction between different features. For example, humans can recognize color, as a visual feature of objects.

(Characteristic) Value: is the condition/state/level/magnitude/amount a characteristic may take. Variation of a characteristic can be described by the change in its value.

Characteristic, Erratic (Final): is an unpredictable (final) attribute of the outcome of a process. While some (final) aspects of a process may be erratic, others could be predictable.

Characteristic, Random (Final): is a (set of final) attribute(s) of an outcome whose state is unpredictable in run by run operations of the process, but (collectively) predictable in the long run.

Choice (in a Uncertain Choice) , Right: is the act/process that the agent would choose, if it had access to a perfectly informative program.

Classification: is the act of fragmenting/partitioning/organizing/subdividing/classifying a population into a number of fragments/partitions/groups/divisions/classes, based on the value of a characteristics/specification. Classification can be done based on either qualitative or quantitative measures.

The roughest/coarsest/crudest classification/distinction can be made in a binary/ dichotomous fashion by evaluating whether objects possess or lack thereof a characteristics. If no characteristics is being used for the purpose of classification, each member can be considered as a class.

The act of classification, is an organizer mind's effort to systematically capture and make sense of the variation (of the characteristics) inherent in the population.

Classification, Joint: is the act of classifying a population based on more than one characteristics/specification. A joint class represents the fraction of the population that jointly possess certain/specific degrees of each of the polarizing characteristics.

Class size: is the number of members belonging to that class.

Class Size, Relative: is the ratio/quotient of the number of members belonging to that class to the population size.

Counting: is the process of assigning a number to a set of distinct objects as the set's size by establishing a one-to-one correspondence between each member (of the set) and the number "one," the unit for counting, and then using the arithmetic operation of addition to calculate the sum (of units).

Device/Means/Tool: is a (material) agent, with the ability to make/prevent changes in its surrounding world/environment. A device needs to types of ability, serviceability and survivability.

Distribution, Propensity/ Relative Frequency: for a random (set of) attribute(s) of a process, is a function informing the rate at which each possible (set of) state(s) is being adopted if the process is being run many times.

Event/outcome: is the result of a physical process in a system, characterized by the set of all the objects of the physical system and their corresponding states.

Function, Deterministic: is a relation between a set of input (algebraic) variables and a set of output (algebraic) variables, in which each (possible) input value(s) is associated with only one (of the possible) output value(s).

Function, Expression of: is the algebraic form/formula for the function, where input variables and parameters are combined through mathematical operations.

Function, Argument: is any of the input variables for the function.

Function, Parameter: is any of non-variable quantities present in the (algebraic) expression of a function.

Hypothesis: is a proposed, i.e. yet to be accepted, explanation for a phenomenon.

Hypothesis, Quantitative: is a plausible/likely value (number) for a (physical) quantity or a (mathematical) function for a quantitative regularity/ structure.

Hypothesis, Statistical: is a plausible/likely value (number) for a relative frequency or a (mathematical) function for a quantitative regularity/ structure.

A statistical hypothesis is one of the possible probability distribution functions, characterized by its functional family and a set of values for the parameters of the functional form, which informs the probabilities (either relative frequencies or degrees of plausibility/ credibility) for each possibility in the space of possibilities (either a sample space or a hypothesis space).

The relative frequency of event informed by a (statistical) hypothesis is also called the Likelihood of the event. From the standpoint of the mechanics of A Likelihood is a physical probability

Indexing: is the act of classification without consideration of any specific characteristic/specification. The purpose of indexing is differentiation among various members of a population.

Informative Program (regarding an Uncertain Process), A Perfectly/Completely: is a program providing with the information that enables (a previously ignorant) agent to (positively) predict the (future) outcome of every run of an uncertain process.

An informative program itself is an uncertain process, as the agent does not know the outcome foreseen by the program until he/she runs the informative procedure for each run of the (original) physical process, he/she was uncertain about. The same credibility values assigned by the agent to the outcomes of the (original) physical process extends to the outcomes of corresponding perfectly informative process.

Informative Program (regarding an Uncertain Choice), A Perfectly/Completely: is a program providing with the information that enables (a previously ignorant) agent to (positively) predict (a) the "right choice," as well as (b) the (future) outcome of every run of each of the uncertain processes, if that process is being taken.

Interaction: between two objects is their mutual effects on certain physical states of the objects.

Language: is a means of creating and communicating information composed of a collection of symbols and a set of rules/ structures that can be used to connect symbols and to create complex constructs. Linguistic descriptions can be communicated by physical means, including visual signs, sound waves and electrical signals. Spoken and written languages as well as mathematics, are example of languages.

Among other functions, a language can be used to create constructs in order to represent certain objects/states of the world and their connections. Linguistic representations/ descriptions may vary in terms of their preciseness/perfection. The assertion that a specific (linguistic) description precisely/ perfectly represents the real-world is a postulation, whose validity relies on the premise that the objects/states of the world and their changes follow the (arbitrary) forms and the rules of that language. The premise implies that the God/nature has been creating/changing the (preceding) objects/states of the world according to the (arbitrary) rules of that language, which have organically evolved and developed later in the minds of humans.

Load/Demand: an abstract concept as a representation of what is imposed on a device by external environment, due to device's existence and service, as well as its interaction with the environment.

Machine: is a complex object, composed of multiple components, with the ability to sense and/or to change some physical states of the world.

Machine, Biological: is a machine, whose components are living organisms. Humans are examples of advanced biological machines.

Machine/Observer, Sensing: is a machine with the ability to sense/observe and differentiate some physical states of physical systems.

Machine, Preferring: is a sensing machine that prefers some states of the world to the others, if given the option.

Machine/Agent, Acting: is a machine with the ability to trigger/initiate a physical process in a system and to provide with what is necessary for the completion of the process.

Mathematics: is a collection of branches of knowledge, each studying a specific group of (abstract/mental) mathematical objects/constructs and the structures/patterns relating them.

For example, arithmetic studies mathematical objects termed numbers and the operations for their manipulation, or algebra studies unknown (numerical) quantities, represented by symbols, the relationships between these quantities, and the rules for manipulating them.

The most essential tool for development of mathematical knowledge is logical reasoning, in which conclusions (end points/results) can be drawn from premises (starting points) in a ruled-governed (mental) process.

Measurement: is the process in which a number is being assigned to the level of a certain characteristic of an object by comparing it with the unit level at a designated reference object.

Measurement unit: for a certain characteristic, of continuous nature, is the level of the characteristic in a reference object.

Model: is an entity in representation relation with another entity, i.e. either representing or represented by the other, where the relation is being established due to perceivable similarities, possibly created by partial imitation, or due to conceivable meanings, possibly created by symbolization or signification.

An entity can be purposefully used as a model by (1) finding an existing or creating a new entity with some similarities with another entity, or/and by (2) attributing a meaning to an entity in relation to another entity. In (1), what differentiates the representation from the represented (composite) object is the specifics of the set of objects, their properties, and the "pattern of relations" between the attributes of the (parts of the composite) object; namely the different "systemic characteristics" of the two.

A model is an (system of/compound) object of certain attributes and relations, either internally and/or externally, with the potential to be partially copied/imitated in “forming” another object of similar attributes and relations; a replica. The specifics of to-be-copied attributes and the relations of a model system of objects, may be called the “pattern” of the system; thus, a model can be characterized based on its distinct pattern. Due to the similarity of patterns between a model and a replica, the latter me be seen as a representation of the former, at least in terms of the copied attributes. For an object perceivable to humans, an example feature is the visual look and an example internal relation is the dimensional proportions of various parts of the object.

A model is an example embodiment of the kind of objects a certain processual code can generate; i.e. is a token of a type. The code of the modeling process determines the specifics of the attributes and the relations to be imitated. An example code for a natural modeling process is the genetic code used by biological agents in reproduction, and an example (partial) code for a purposive modeling process by humans is the blueprint used in construction of an artefact. Within the philosophical framework of universals and/or types, a model can provide a concrete replacement for the abstract notion of the pertaining universal or type, where all instances or tokens can be seen as copies of a model instance or token. In other words, the concept of model grounds the concept of universal or type. In regard to the concept of representation, a model can be seen as a representation of either the processual/typical code and/or all the tokens the code can generate.

What distinguishes a model from another are the same things distinguishing a system of objects from another, namely the member objects and their individual attributes, as well as the systemic relations between them. Models whose systemic specifics are close, can be conceived to form a “model family,” where common (systemic) attributes characterize the family and the specifics of differentiating (Systemic) attributes characterize family members. While the proposition that a system "follows" a certain model means that we take the specifics of the system to be identical to those of the model system, the proposition that a system "follows" a model family means that we take a certain subset of the specifics of the system to be identical to common attributes of the system. By associating a system to a model or a model family, we take the system to be a complete copy of the model, or a partial copy of the members of the model family.

Model, cognitive: is a subject's mental representation of an entity that was originally outside one's mind. All understandings of things are formed by relying on cognitive models and to some degrees are representations of entities from outside one's mind.

While a model is imitable, itself could be a representation of or modeled/imitated from another entity. This is specially the case for a mental understanding, as it is being formed (in the mind) to cognize an outside entity; thus, a (mental) representation of the entity, but can also be used as a model for understanding another outside entities, where the features of the mental understanding are being projected onto the other outside entity. In other words, a mental representation of an object is also a model to see other objects in its lights. If we denote the original/first outside entity as O_1, the mental understanding as M, and any other outside entity as O_X, the directional relation between the three can be denoted by $O_1 \rightarrow M \rightarrow O_X$, where M is modeled from and a (mental) representation of O_1 , as well as a (cognitive) model for O_X. In this sense, a mental construct being a model, can be attributed to either of the two representation processes it participates in, the one where the construct represents an outside entity or the other where another outside entity is seen alike the (previously formed) mental representation.

Model, perceptual: is a mental entity, partially constructed from sensory input, whose specifics determine the way in which mind cognizes a certain object. Putting it differently, a perceptual model is a mental representation of a perceivable entity and will be consistently evoked upon receiving the same sensory stimuli. For example, the color red is the perceptual model human minds forms based on sensory signals from lights with the wavelength between 620 and 750 nanometers and the frequency between 400 and 484 THz.

The reason for naming an output of perceptual construction processes as a "model" is that (a) the features of the output are the essence what we understand as a "form," i.e. a cognizable object of certain attributes and relations and (2) upon receiving sensory inputs of the same character at different times, the mind consistently generates mental output of the same character, as if the mind uses a token output as a prototype/model to imitate from and to "form" similar output.

Model, conceptual: is a mental entity, formed by relating specific attributes of a number of conceived objects, which provides a kind of way for seeing similar conceptual instances; i.e. a model for understanding.

While examples of conceptual models are omnipresent, an example from physical science is Newton's model for the gravitation of heavenly bodies, which associates the specifics of bodies' movements with a number of (inherent) attributes of the bodies, in a certain way. A conceptual model can also be taken as a man-made idealization/simplification of a real-world object/phenomenon constructed in order to imitate/ simulate/ demonstrate/ describe/ explain/ elaborate/ predict certain aspects/ characteristics of the reality.

Model, Mathematical: is a mental/ cognitive/ intellectual model, where certain aspects of an object/ phenomenon (or its behavior) are represented by (abstract) mathematical objects.

Model, Quantitative: is a mathematical model, where certain aspects of an object/ phenomenon (or its behavior) are represented by interrelated quantities (that are in a state of equilibrium).

Motivation: factors that internally, inside biological agent's mind, arouse and direct agent's action to preserve/change certain (external) states/objects of the world.

The ultimate motivation of a biological agent can be synthetized as the "satisfaction of its needs." Very complex (sustained) needs, such as a constant need to achievement or the (extreme) need to control one's living environment, may be inferred from behavior exhibited from some humans. Success in meeting a need, might internally, inside biological agent's brain, result in a state of pleasure/comfort or failure to do so in a state of pain/discomfort.

Number: is a mathematical object, an abstract construct, developed (by humans) as a (mental) representation of a characteristic of a collection/ set of (physically) existing or imaginary objects/ states. When used regarding the physical world, numbers are the essential component for quantification processes, e.g. by counting, measuring, or calculating. Numbers can also be used as labels, which facilitates organizing processes such as sorting.

Numbers were originally developed to represent the size/ magnitude of a set of discrete/ distinct objects (natural numbers). The concept was later extended to establish a representation, where units could be partitioned into smaller fractions (rational numbers). Further extensions resulted in development of new categories of numbers such as negative, irrational, and complex numbers.

Number is a fundamental mathematical object, invented as a means for precisely differentiating between the size/largeness of the sets containing similar members.

The first (type of) numbers created (by humans) were “natural numbers,” as (mental) representations of the size of the sets of discrete objects obtained by the (simple) act of counting. In other words, there has been close relationship between the development of natural numbers and the basic arithmetic operation of addition. Later, the concept of number were (mentally) expanded to create more abstract types of numbers, such as integers, rational numbers, real numbers, and imaginary numbers. Some types of numbers are without a real-world analog/map and remain only as (abstract) objects (of the mind) suited for (mental) manipulations, according to (humanly) defined rules for manipulating such abstract entities.

Numerical/Arithmetic Operation: any (mental) operation (on numbers), where a set of numbers interact, according to the rules of that operation, and yield a numerical result. Four basic arithmetic operations are addition, subtraction, multiplication, and division.

Numerical/Arithmetic Process: a process composed of a set of arithmetic operations, arranged in series and/or in parallel, which (eventually) results in a numerical output.

(Numerical) Variable: a representative for every member of a heterogeneous set (of numbers).

The term “variable” reflects the variation of the numerical character/value of members of a (heterogeneous) set of numbers. Oftentimes, a variable is visually represented by a symbol such as an alphabetic letter, e.g. x, y in English, α, β in Greek.

Numeral: is a symbol invented to represent a number, as a mathematical object, in a non-verbal mode of (human) communication, i.e. written/ pictorial.

Observation: is the act of learning/receiving information regarding the pertaining class of a member of the population.

Operation, Mathematical: is a basic/ core component of a mathematical process that takes a number of input mathematical objects, arguments, and delivers some output mathematical object. For example, the main mathematical processes in arithmetic are numerical calculations, which can use the basic (arithmetic) operations of addition, subtraction, multiplication, and division.

Operation, Cartesian Production: is a mathematical operation that takes two sets and return a (product) set composed of all possible ordered pairs resulting from the association of each members of the first of two sets with every member of the second set. It is possible to extend the operation to take more than two input sets and return a n-fold/ n-dimensional product set.

Physical Process: is a group of steps/ tasks/ stages, triggered/initiated by a change in certain aspects of (a number of) worldly objects, that results in either the preservation of the old status, the creation of new objects or a change in characteristics of (the same or different) worldly objects. The new reality, in either form, is the outcome of the process.

Physical process is a set of interactions in a physical system, where the initial physical states of the objects change and results in the creation of new objects or the alteration of the physical state of the system.

Physical Process, Repeatable: is a physical process that if be triggered and completed at a different time and/or space results in the same outcomes.

(Physical Process) Realization Reference Set: for a physical process is the imaginary set containing all possible realizations of the process (past, present, and future).

Physical Object: is any being that its existence, at any instance of time, can be further described by its physical characteristics, e.g. location, mass, volume, chemical composition, etc.

Physical quantity: is the magnitude of a physical characteristics (mentally) mapped to the mathematical object, variable. The value of a physical quantity can be directly measured using a (standard) measurement device or be estimated/predicted using some (idealized) mathematical/algebraic formulas/expressions.

Physical State: is the level of any physical characteristic of an object, at any instance of time, which can be described qualitatively using a linguistic descriptor or quantitatively by assigning a number that quantifies the level in comparison with a (standard) unit of measurement.

While the perception of a certain qualitative descriptor may differ across the human population, the perception of the quantitative measure for the same level of the characteristic is the same as it relies on a (arbitrarily set) external measurement unit. In other words, qualitative assessments are subjective, i.e. depends on the subject who evaluates, and quantitative assessments are objective, i.e. depends on an external reference object. For example, the length of an x-foot long object can be qualitatively described by a subject as long and by another subject very long, however, the quantitative description of the length characteristics remains the same, x, and is independent of both subjects. The uniqueness of the quantitative description is the result of human invention of two different set of tools. First, is the development of a precise language, in human mind, for expressing, manipulating, and relating abstract/mental objects termed numbers. Second, is the designation of a specific level of the characteristic, embodied in an externally existing reference object, as the measurement unit. The set unit, a real-world existence, can then be mapped to the number “one”, an abstract/mental existence, which allows for a description of the reality in a precise language.

Physical System: is any set of physical objects that may get involved in interactions resulting in changes in their initial states of the objects. The state of a physical system is the set of the physical states of all its objects.

Prediction, Structural: is the forecast of a (set of) attribute(s) of an outcome of the process by using the process' structural relationship and the (knowledge of the) states adopted by the initial and circumstantial attributes of the process.

Prediction, Eventual: is the forecast of a (set of) attribute(s) of an outcome of the process by using the process' eventual relationship and the (knowledge of the) states adopted by (other) final attributes of the process, other than the final attribute(s) we aspire to predict.

Eventual prediction may not introduce a single state as eventual relationships may associate a specific set of states of (the rest of) final attributes to more than a single state for the (final) attribute of interest.

Probability, Objective: a consensual relative degree of plausibility/credibility for the occurrence of an event or the validity/truth of a hypothesis with a given set of information and a given set of presuppositions.

One reasonable rule for assigning probabilities to possible outcomes of a simple random sampling process is to equally divide the total probability of 1.0 among members of the population. As a result, the probability for the (compound) event/possibility that a random sample realization be associated with a given class is equal to the relative size of that class. This is the earliest assignment rule historically used in mathematical probability and is referred as classical interpretation of probability, where probability for a possibility is proportional to the number of ways through which the possibility may occur. Classical interpretation of probability is closely connected with the act of sampling with/without replacement.

To physically verify the assumption that a given sampling process is random, we have to sample using the process to see if all members of the population are being selected with the same frequency. It is only after this physical experimentation that we can make an assertion regarding the legitimacy of the nondiscriminatory assumption for a random sampling process.

An answer to a fundamental question regarding how to find the (joint) class an object belongs to before examining it, is that while we cannot say with certainty, but given the information at hand, we can differentiate a number of classes that the object is more likely to be associated with.

Probability, Physical: is the degree by which a repeatable (random) sampling/selection/ realization process/procedure/mechanism differentiates a possibility relative to other possibilities.

The relative degree of differentiation can be measured by repeating the action infinite number of times and counting the number/frequency of realizations for each possibility and calculating the relative frequency/proportion of occurrences for each possibility. The given definition makes physical probability an attribute of the random act/realization mechanism, rather than the outcomes.

Physical probability of an outcome is the degree by which the (underlying) physical process tends to generate the outcome, relative to all other outcomes of the process.

Physical probability distribution is being measured by running the process many times and finding the proportion of replications, where each outcome does occur. Such proportions are also called stable/long-run relative frequencies.

Probability Distribution: is a function/rule that associates a probability to each (possible) value of a random variable. The domain for a distribution is the range of the random variable and the range for the distribution is the set of assigned probabilities.

A probability distribution, when associated with the outcomes of a (physical) random process, represents an assessment/estimate of the relative frequencies/ propensities of the process, i.e. becomes an estimate of a frequency distribution.

Probability (of an Event/ Outcome), Bayesian: is the Bayesian estimate of the physical probability of the event calculated by averaging the likelihood values of the event, given various (statistical) hypotheses, weighted by the plausibility/ credibility of each hypothesis. This Bayesian estimate can also be called *credible physical probability*. Also, because a Bayesian estimate is the result of an averaging (arithmetic) procedure, some has named it a *Total Probability* of the event.

Plausibility/ Credibility of a Hypothesis, (Non-Informative) Prior: is a degree of plausibility/ credibility defined/ assigned/ established for a statistical hypothesis relative to other (competing) hypotheses, before incorporating the existing data on the outcome realized on various repetitions of the random act.

Plausibility/ Credibility of an Event/ Outcome, Prior: is the average of likelihood values of the event, given various (statistical) hypotheses, weighted by the prior plausibility/ credibility value for each hypothesis.

Uninformed prior credibility values for a statistical hypothesis relative to other (competing) hypotheses, are defined/ assigned/ established before incorporating the existing data on the outcome realized on various repetitions of the random act.

Prevalence of a risky act: is the degree to which a risky act is credible as the preferred choice, relative to (all other) alternative acts. Mathematically, is the sum of the credibilities for every outcomes of the act, where the act is preferred to all other alternative acts.

Prevalence of an act, in the light of a (perceived) perfectly informative program is the credibility of an event, in which the act turns out to be the "right choice." Probability of an events, is the sum of the probabilities for the (basic/simple mutually exclusive) outcomes defining the event.

Propensity: of a generating process toward an outcome is the physical probability of the outcome.

Population: is the real (i.e. physically existing) set of a specific number of objects, each a realization of a physical process.

Population/Reference set: is a set/group comprising all objects of interest with a number of common/shared characteristics/features.

(Population) Variable: is the codomain/range/output of a function/rule associating a number/value to each of the members/classes resulting from an act of classification (of a population). The domain for a (population) variable could be the population, the set of all members, or a set of population classes. Any function of a (population) variable is also a (population) variable.

I coined the term "population variable" to distinguish a variable used to describe the variation across a population from a variable used in mathematics or theoretical physics.

Population Class Size Distribution: is a function/rule that maps each class in the population to its size/relative size, the ratio of the number of members in that class to total population members.

Process, Known: is a physical process that is completely known to an observer, i.e the underlying system, its initial state, and the outcomes of the process.

Process, Controlled: is a purposefully triggered process that its advancement and its final outcomes are controlled by the agent.

Process, Deterministic: is a physical process that if being run consecutively under the same initial and circumstantial states, it consistently delivers the same outcome. Every physical process is deterministic as long as humans keep their deep seated believes in causality.

Process, Random: is a repeatable process that has more than a single set of outcomes. The initial state of the system and its evolution are such that the observer does not know the exact set of outcome to be realized before the commencement of each repetition. The randomness of a purposefully triggered process indicates that the agent is not in full control of the process, i.e. there are aspects of the process, which he cannot influence.

Process, Independent Random: is a random process, where the realization of a set of outcomes in one repetition does not change the observer's ability to predict the outcome of the process in coming realizations.

Process, Stationary: is a physical process whose structural relationship does not change along the time. Since the available knowledge of the physical world is constructed upon the past events, any prediction of the realizations in the future requires the assumption that processes remain stationary along the time.

Process, Predictable: in respect to a (set of) attribute(s) of its outcomes, is a label for a physical process, for which the changes in the state of the corresponding (set of) attribute(s) in consecutive runs of the process is predictable.

Process, Erratic: in respect to a (set of) attribute(s) of its outcomes, is the label for a physical process, for which the changes in the state of the corresponding (set of) attribute(s) in consecutive runs of the process are not predictable, either structurally or eventually.

Predictability and erraticism are not inherent characteristics of a physical process, but of the (level of) knowledge available to the agent who is interested in the outcome of the process. Also, a physical process may be labeled erratic in respect to a specific (final) attribute, but it may be partially erratic in respect to another (final) attribute, and be predictable in respect to other (final) attributes.

Process, Random/ Stochastic: in respect to a (set of) attribute(s) of its outcomes, is an erratic process that if being repeated many times, it exhibits a deterministic/ certain/ stable tendency/ propensity to adopt various states (relative to each other).

A natural random process is yet to be discovered, but it is possible to develop a man-made process that (almost) exhibit predictability requirements, namely, (1) a (limited) state space, and (2) stable long-run behavior, but still remains unpredictable to an agent observing the realization of the subsequent outcome (of the process). However, if the observer learns the (hidden) pattern according to which outcomes are generated, it can rename the process from random to predictable.

Process, Uncertain: is a (member of a class of) process, with some shared features, that (1) (collectively) yield different outcomes in various runs of the process, and (2) the agent, who intends to assess to the process, is unable to predict the outcome to-be-realized in each run of the process.

Relation/Association: is a connection between a number variables. Relation/association is a primitive notion. In a relation, one variable is the reference/original variable and another is associate variable.

Relation/Association rule: is a way by which (values of) a number of variables, are connected/ associated. In a relation, one variable is the reference/original variable and another is associate variable.

Relationship, Structural: is a group of relationships connecting/ associating/ mapping/ corresponding the adoption of the each possible set of states for the initial and the circumstantial attributes of a physical process to the realization of a specific set of states (of the various attributes) of the outcome of a physical process.

Relationship, Eventual: is a subset of structural relationships connecting/ associating/ mapping/ corresponding the adoption of an attribute of an outcome of a process with the set of all possible states for the other attributes of the outcome that may concurrently realize in the final outcome of the process. Eventual relationships can be obtained deductively by canceling out the (states of) initial and circumstantial attributes from the structural relationships

Resistance/Supply: the ability of the device to withstand an imposed load.

Resource: an agent's means/tool to preserve/change certain (external) states/objects of the world.

Sample: is a limited-size subset of the (real/ imaginary) set containing all possible members/ realizations of a mixed population/ physical process.

Sampling Process: is process through which a member of the population/ reference set, is being chosen to be included in a sample set. Sampling is a “selective observational process,” where the observer can only learn the character of a select number of objects rather than every object.

Sample, Random: is a sample collected such that the observer/ conductor of the sampling process is ignorant regarding the outcome to be observed before the completion of the process (i.e. he/she cannot predict the outcome in advance).

Sample, Simple Random: is the product of an unbiased sampling process, where each member of the population/ reference is equally likely to be chosen relative to all other members. In such an unbiased selective observational procedure, if the process is being repeated many times, relative to the size of the population, each member (of the population) is being observed as many times as every other member.

Sampling/Experimentation: is the act of selecting and inspecting/evaluating a member of the population in order to learn its class.

Sampling/Experimentation, Simple Random: is sampling from a population in a way that the member to be selected is not predetermined and each member is equally likely to be selected.

In other words, the process of simple random sampling does not discriminate across the population and each member is a possible selectee. The indiscrimination, however, creates uncertainty regarding the outcome of the selection. To quantify the uncertainty, probability measure is being introduced to numerically differentiate among various possibilities that may be selected.

Sample Space: is the set of all concrete outcomes for a random act, among which is being realized every time the act is being taken. A concrete outcome is the one that has a direct manifestation/ representation in the real-world; thus, its realization can (temporally and spatially) be verified. A frequency distribution function characterizes the sample space by informing the relative frequencies for all its outcomes, as they do realize in the real-world.

Space, (Statistical) Hypothesis: is the set of all frequency distribution functions (hypothesized by the analyst) that one of which may give the true/ real relative frequencies of a specific sample space. A plausibility/ credibility distribution function characterizes the hypothesis space by defining/ assigning/ establishing the relative degrees of plausibility/ credibility for all its hypotheses. A plausibility value does not exist in the real-world, and is an artifact of human analysis resulting from an approach, Bayesian, man has developed to systematically address his lack of knowledge regarding (the true/ real) relative frequencies in a sample space. Since prior plausibility/ credibility values are arbitrary they do not have truth values, i.e. being true or false, and they cannot be compared/checked against an (externally) measurable quantity.

Selector: is an agent/ device/ machine, either biological or non-biological that is able to sense and differentiate between certain (natural) objects/ states and favor some over the other.

The (observation of the) behavior of the agent reveals its preference among a set of available objects/ states. Many organisms, including plants and animals, can sense, differentiate, and select among certain objects/ states available to them.

Serviceability/Usability/Employability: the ability of a device to maintain its function, as an agent of change/preservation.

Survivability/Durability/Reliability: the ability of a device to maintain its integrity, as a prerequisite for its serviceability, in the face of various sources of threat; external threats such as (gradual or sudden) changes in environment, as well as internal threats such as wear and depreciation.

Set, Mathematical: a well-defined/ described collection of distinct (mathematical) objects, called members. A set can be defined/ described either by intention, i.e. by using a rule or semantic description, or by extension, i.e. by listing/ enumerating every member of the set.

Set (of two), Ordered: a pair of two mathematical object, where the order of the appearance of object does matter. An ordered pair is also called a sequence of length two, or a 2-dimensional vector is the mathematical objects are numbers.

State /Object of interest/desire: is a subset of the outcome of a process that a preferring machine wants to be realized.

Theory, Probability: a branch of mathematical study aimed at quantitative analysis of uncertainty/risk/randomness.

Theory, Mathematical: a collection of axioms/principles and theorems/propositions aimed at a systematic study of combinations of certain (mathematical) objects.

In the case of probability theory, objects of interest include simple random experiments/ trials and their outcomes, simple events, as well as compound events/outcomes in the form of sequences/n-tuples recording/capturing (simple) outcomes realized in consecutive repetitions of a simple random trial.

Variable: is an abstract algebraic object, represented by a symbol/alphabetic character, constructed to refer to an unknown/ unspecified number, the "value" of the variable.

Variables allow for formation of algebraic calculations, which are generalizations/ extensions of arithmetic calculations on a set of specific numbers to a situation, where the numbers are unknown/ unspecified.

Variable is a theoretical construct, denoted by a symbol, e.g. X, in reference to the value of a characteristic. Variable is a surrogate/substitute for all possible conditions/states/levels/ magnitudes/amounts a characteristic may take. The concept of variable significantly simplifies and facilitates description and analysis of changes in a characteristics.

Variable, Index: is an (integer) variable resulting from the act of indexing population members.

Variable, Numerical/quantitative: is a variable whose possible values are number representing the degree to which a characteristic is possessed.

Variable, Random/stochastic/aleatory: is any variable, associated with the population, whose values stochastically realize due to aleatory realization of population members in a random sample. Any random variable is described by probabilities assigned to its values.

Variation: is a change/difference in condition, level, or amount. Closely related to the concept of characteristic/specification, variation is also a primitive notion; i.e. humans are innately able to comprehend changes in a certain characters of objects. For example, humans can recognize variation in color of objects.

3. CATEGORIZED DEFINITIONS

Topic A: Naming:

Name/Label: is (one or more) words designating an individual entity.

Address: is the description of particulars of a location.

Nominal Address: is an address (created) based on the name of the object.

Characteristic Address: is an address (created) based on the (particular) character of the object.

Index: is an indicator for a member of a set in an ordered sequence of set members. An object's index can act as its address in a sequence.

Character-based Naming: is a naming method, where objects are being named according to the degree/intensity/type, by which they possess a character.

Character-based Sorting: is a sequencing method, where members of a set are being ordered according to the degree/intensity/type, by which they possess a character.

Position-based Indexing: is an indicating method, where sequence members are being indexed according to their position/address in the sequence.

Combinatorial Process: is a process involving two (or more) sets, where every member of one set is being combined with every member of other set(s), producing a new object of multiple components, an n-tuple.

The results of a combinatorial process are called combinatorial outcomes, which each are composed of various components, each a member of the (original) comprising sets.

Sequential Combinatorial Process: is a combinatorial process that combines members of a number of sequences, rather than unordered sets.

Multi-Attribute Indexing: is method for indicating at outcomes of a combinatorial process, where members of a multitude of ordered/sequenced sets are being combined.

The method requires establishment of a sorting criterion as a means of prioritizing one combinatorial outcome over the others.

Directional Index: is the index specifying the rank of members in one of the sequences combined through a combinatorial process.

Sorting Variable: is a function of all attributes/dimensions characterizing members of a combinatorial space and is defined to assess the priority of each combinatorial outcome over others.

The definition of a sorting variable can be based on a real or an arbitrary preference among characters characterizing components of combinatorial outcomes. Geometrically, a sorting variable provides a preferred movement path in the multi-dimensional combinatorial space, where each dimension is the characters, or a character-based index, characterizing one of the source sequences.

Combinatorial/Collective/Integrative Index: is an index specifying the position of a combinatorial outcome in a sequence of the members of a combinatorial space, ordered based on a sorting variable.

Arbitrary Combinatorial Index: is a combinatorial index, whose sorting variable does not reflect any real preference among combinatorial outcomes and is just a way of systematically sorting multi-dimensional outcomes.

Topic B: Discretization and Numerical Calculations:

Domain: is a (continuous) collection of points, an analogy for members of the domain, defined by one or more (quantitative) characteristics, the number of which specifies the dimension of the domain

Boundary: is a (continuous) collection of points, which divide the domain into two distinct subsets. A boundary is usually of one dimension less than that the number of dimensions defining a domain.

Non-intersecting Boundaries: are a number of boundaries, which do not intersect inside a specific domain.

Layer: is a subset of a domain bounded by a number of boundaries.

(Discretized) Element: is a small subset of a (numerical) domain, with the same characteristics/ dimensions defining the domain.

A discretized element in a continuous domain is specified by a range of values for each of the characters/dimensions of the domain.

Single-Layer Element: is an element not crossed with any boundary; thus, only confined with external boundaries of the element.

Multi-Layer Element: is an element crossed with one or more boundaries.

Cross-Boundary Element: is a multi-layer element, in which two or more boundaries cross each other.

Topic C: Calculus:

Number: is a mathematical object abstracted by human mind from its perception of the size of a crowd of discrete (similar) physical objects or the volume/ heft of a bulk matter or the magnitude/ intensity of a sensible quality, such as ambient heat or humidity.

Numbers provides a means for quantifying various aspects of the physical world, where a number specifies the (precise) size/ magnitude of a certain aspect of a physical object. In other words, a quantified size/ magnitude of a physical aspect provides a concrete representation for the pertinent number, which is an abstract construct of the human mind.

Number one, which is abstracted from human observation of a single physical object (of a kind), is central to the development of other numbers by using arithmetic operations.

Quantification: is the process of assigning a number as the size/ magnitude/ intensity of certain physical aspect of a target object, by using a (standard) means of quantification.

Quantification requires a physical representation for the number "one" as the unit of the physical attribute to be quantified. When the target physical attribute is the size of a crowd of discrete objects (who share certain features), each object is being mapped to number one and the size of the crowd is being found by the act of counting. When the target physical attribute is not comprised of discrete (countable) units, the process of quantification requires defining/ establishing a (standard) unit (amount) for the target (physical) aspect and a measurement device/ procedure to determine the magnitude of the aspect for a specific object.

Character/Aspect/Property/Quality: is a feature of an object as perceivable by sensors capable of sensing the feature.

State (of a character): is one of the possible distinct forms/variations of a specific character. For example, the character "color," as perceived by human senses, can take various states such as red, green, and blue.

Category/Class/Type: is the label for a set of objects, who share a number of (states of) specific characters.

The most basic categories are the ones encompassing every object that shares a specific state of a character. For example, the category "red" includes every (real and imaginary) object, whose color is red.

(Non-numeric) Variable: is a concept, constructed by human mind, as a means of reference to every possible state of a character, where the label for the state, that is realized (in an object), is set as the "value" of the variable. For example, if an object is sensed as red by a color detection sensor, the value of the color variable becomes "red."

(Numeric) Variable: is a mathematical object with the ability to take/ adopt any (numerical) value from a set of numbers.

The concept of variable is being abstracted by human mind from the capacity of physical objects for evolution/ adoption, characterized by change in the magnitude/ intensity of some of their (physical) characters. A variable provides a means for capturing every possible quantity a physical attribute may take. In other words, the (changing) quantity of a physical character provides a concrete representation for the pertinent variable, which is an abstract construct of human mind.

Associated Variables: are a number of variables, who are tied together such that each value of one of the variables is being associated with specific values of other variables.

Rule of Association: is a pattern, according to which variables become associated.

The rule informs the "how" of association by specifying that each value of one variable is associated with what values of other variables. The specification requires an exhaustive list of induvial associations between values, provided either by a (detailed) description/enumeration or by a (general) formulation with the ability to generate the detailed list at will.

Addition/ summation: is an arithmetic operation on two numbers resulting on a third number, larger than both input numbers.

The mathematical operation of addition is being abstracted from the situations of physical growth in the size/ magnitude of a physical attribute. The growth in the size of a crowd of objects occurs due to arrival/joining of new masses to the crowd, either through the process of birth from within or through inflow from outside. The growth in the magnitude of a physical attribute results from the evolution of the object in response to changes in other attributes of the object and/or of object's environment. Thus, physical processes such as birth/ generation, inflow, and evolution provide concrete representations for the abstract mathematical operation of addition.

Deduction/ subtraction: is an arithmetic operation on two numbers resulting on a third number.

The mathematical operation of addition is being abstracted from the situations of physical shrinkage/ decline in the size/ magnitude of a physical attribute. The decline in the size of a crowd of objects occurs due to departure/leaving of existing masses (from the crowd), either through the process of death (from within) or through outflow (from inside). The decline in the magnitude of a physical attribute results from the evolution of the object in response to changes in other attributes of the object and/or of object's environment. Thus, physical processes such as death/ decay, outflow, and evolution provide concrete representations for the abstract mathematical operation of addition.

Algebraic Expression: is a mathematical object composed of a set of constant numbers and variables combined through numeral operations.

The outcome of an algebraic expression is a variable whose value is determined by the numerical process delineated by the algebraic expression. An algebraic expression, as an abstract object, can be concretely represented a physical process where a series of growth/ decline, arrival/ departure, and evolution processes occur in series or in parallel. Thus, the value of the variable defined by the expression can be concretely represented by the magnitude of the target attribute at the end of the pertinent physical process.

Algebraic Equation: is a mathematical object informing the equality (of the values) of two (variables defined by) algebraic expressions.

An algebraic equation, as an abstract object, can be concretely represented by a physical event, which only occurs when the magnitudes of two physical attributes balance each other, i.e. be at equilibrium.

Function: is a mathematical device/machine, which uses a numerical process/rule to assign/map to each possible (numeric) value of a (independent) set, another value from a (dependent) set.

The (independent) set (of values), which the function takes/maps from/assigns to, is called the "domain" of the function, and the (dependent) set (of values), which the function gives/maps to/assigns, is called the "range" of the function. Every/all possible (numeric) member of domain, and range as two sets, can be represented by (numeric) variables and denoted by symbols such as x and y, respectively. In a different terminology, x-values are "assignee" and y-values are "assigned."

The terms "independent" and "dependent" used to label variables, as mathematical objects, are not in reference to a concrete physical reality. A justification for the use character "dependency" for differentiating input and output variables is that in any process, it is (the state of the) input that determines (the state of) the output, not the other way around; thus, the output variable is "dependent" on the input variable.

While my definition classifies a function as a (mathematical) device capable of the act of assignment/mapping, it is more common to see definitions, where a function is classified as the rule/ process underlying its act of assigning/mapping. It is also common to see definitions of a function that drop any reference to the underlying process or the device employing that process, and simply define a function based on associated variables. In such definitions, the dependent variable, y, is simply a function of the independent variable, x.

Beyond immediate differences among these definitions, the mere existence of these variants, reveals more facts about the varying nature of "thinking processes" across human population. These variations become more evident, when human mind needs to apply fundamental abstract concepts such as device/machine, process/mechanism/rule, input, and outcome, to the abstract world of mathematics. In the absence of a "standard thought process," thinkers may fail to distinguish between a mathematical device/machine, in this case a function, the process that the (man-made) device employs in its operation, and the outcome of applying the device.

Rule of a Function: is the numeric operation used by the function, in which numeric operators act on an input (number) along with (other) constant numbers, to produce a (numeric) output.

A rule can be symbolically represented by a mathematical expression, where the (numeric) value of the input/independent variable is denoted by variable's symbol, e.g. x. Rule of a function, written in (symbolic) algebraic language, is the only way to inform the "how" of the corresponding association process. Other means of representing a function including (1) spoken/written human language, (2) numerical tabulation, and (3) graphics, only inform the associations made by the function, without informing "how" the associations have come to the realization; i.e. why x* is associated with y* and not y**.

The rule of a function informs the "type" of the (numerical) process the function uses for associating members of domain with members of the range. Replacing the symbol representing the independent variable with the (numeric) value of a specific member of the domain, yields a "token" of that type of process. The specific numeric value differentiates corresponding token of that type of a numerical (association) process from other token of that process.

Inverse function: is another function whose direction of assignment is inverse of an original function such that it assigns to each member of range a member of domain.

An inverse function can only be formed if the original function is one-to-one such that no two distinct members of domain are being mapped to a single member of range. The inversion (of the direction of the assignment process) results in a switch in the positions of the domain and the range as well as independent and dependent variables. The domain of the original function becomes the range of the inverse function and the dependent variable of the original function becomes the independent variable of the inverse function. The switch between the positions of dependent and independent variables are consistent with the fact that (the state of) "dependency" character of variables is determined by the direction of the (mathematical) process connecting variables, than a (physical) precedent between (the realizations of) the concrete representations of the two variables. So, it is more accurate to perceive input and output variables as two mathematical objects connected through a mathematical process and call each (variable) an "associate variable" and take dependency-based labels as an indicator of the direction of association.

Variable Transformation: is a mathematic process that takes a variable and gives an associate variable as its transform.

Transformation process relies on feeding the variable as an input to a function (of choice) and receiving a dependent variable as a transform of the (dependent) variable. There are infinite number of transforms for a variable as there are many functions that can take the (independent) variable as their input. Common functions used for transforming variables include first and second order polynomials, as well as logarithm. Variables are being transformed for a variety of reasons including making algebraic manipulations easier, or providing a more familiar visual representation of the relationship between variables.

Derivative of a function: is another function, which takes the independent variable of the original function as input and gives the "rate of change" in the magnitude of the dependent variable in response to a (infinitesimally small) change of the magnitude of the independent variable.

In addition to the term the (instantaneous) "rate of changes," the output of the derivative function can be labeled by terms including "interval width ratio," "ratio of the differences" (at the limit), "sensitivity of output to input," or simply the "derivative." Although the use of the label "derivative" for the output variable of the "derivative function" might be confusing, it is in line with similar confusing labels such as the use of the label "function" for both the (mathematical) device/process and its output.

The rule for the derivative function, in its explicit algebraic representation as an equation, can be found through the process of "differentiation" and by manipulating the formula for the original function. The process involves development of the "difference quotient," in an algebraic form, which is a function of two variables, the independent variable, x, and another variable, called the increment of x, denoted by Δx. When the increment of x approaches zero, it is possible to calculate the "limit" of the difference quotient, by manipulating its formula and by canceling out the variable Δx in the numerator and the denominator, and to obtain the rule for the derivative function.

Rate of Change: (of two associated variables) is the ratio of the (induced) change in the magnitude of the dependent variable to the (inducing) change in the magnitude of the independent variable. Rate of change is the output of the derivative function.

The simplest concrete representation of the rate of change is the magnitude of the change realized in one of the characters of an evolving system in a unit time. For example, when the evolving system is comprised of an object moving toward a target, the character required for precise/ quantitative description of the evolution is the spatial position of the moving object at each instance of time. The rate of change in the position of the object in respect to time, called the "speed" of movement, gives the magnitude of displacement realized in a time unit.

The rate of change can also be used to mathematically describe a physical production/ consumption process over time, where by providing the required input, the process delivers its output. Since the generation/degeneration process occurs along the time, it is possible to describe the magnitude of (various aspects of) input or output in respect with time. While the rate of consumption gives the magnitude of (a certain character of) the input taken by the process during a unit time, the rate of production gives the magnitude of (a certain character of) the output given by the process. It is also possible to bypass the timeframe and directly associate the input and the output of the process. For example, in economics, the marginal cost of production is the cost, of input to a production process, for delivering a unit of the output.

Anti-derivative of a function: is a function, whose derivative is the original function. Thus, the output of the original function gives the rate of the change for the anti-derivative function.

The process to find the rule of anti-derivative function is called “integration,” which is the inverse of the algebraic process of “differentiation.” The input of the process of integration, i.e. the original function whose anti-derivative is sought for, is called the “integrand” function. A mistake similar to

A consequence of the inverse relationship between differentiation and integration processes is the non-existence of a unique anti-derivative for an original (integrand) function, where adding a (constant) number to the rule of an anti-derivative results in the rule for another anti-derivative of the original function. The reason is that any original function can be perceived as the function plus the number zero, and the anti-derivative of zero could be any other (constant) number.

A concrete representation of the output of an anti-derivative function is the total amount of input/output being generated/degenerated in a production/consumption process, after a specific time. To label the mathematical objects consistent with the physical analog, the output of the original function gives the “rate of generation” of the process and the output of the anti-derivative gives the “cumulative amount” of the input/output being taken/given by the process.

Derivative of a function: is another function, which takes the independent variable of the original function as input and gives the “rate of change” in the magnitude of the dependent variable in response to a (infinitesimally small) change of the magnitude of the independent variable.

The rule for the derivative function, in its explicit algebraic representation as an equation, can be found through the process of “differentiation” and by manipulating the formula for the original function.

Topic D: Philosophy:

Type/ Category/ Class: is an abstract construct developed to label groups of objects based on their common characters. Based on the defining character of a type, all members of a class are the "same."

Token: any of the members of a class of objects, who share a number of features. A token acts as a concrete representation of an abstract type. Tokens of the same type are "type-identical."

A Priori: is the (theoretical) knowledge obtained by means of deduction rather than empirical observation.

A Posteriori: is the knowledge obtained by using empirical observations as the basis for reasoning.

Topic E: Mathematics:

Formal: presented in an arranged/ordered fashion according to specific rules pertaining the appearance (of an object)

Recipe: is a set of instructions developed to inform "how to" take certain acts including the act of creating/building a new system, or a token/instance of an (already known) system.

Formula: is a statement informing the relationship between (various) characters of a (typical) system.

Law/Rule: is a statement developed to inform the acceptability/legitimacy of an object, act, or a recipe, in a system of thought.

Form: the appearance of a system or of a (linguistic) statement about the system, including pertinent rules and formulas.

Formal/Axiomatic system: is a system of thought consisting of forms/symbols representing (1) a number of objects with certain characteristics, (2) rules for (acceptable) interactions between the objects, and (3) formulas relating (various) characters of (basic and compound) objects after certain interactions.

Objects of a formal system are either basic/primitive or compound. While primitive objects are accepted (by users of the formal system) without a definition, compound objects are well-defined in terms of the outcomes of interactions involving primitive objects. Formulas of a formal system are either axioms or theorems/propositions. While axioms are the result of an (initial) establishment/assignment (by developers of a formal system) and are accepted as they are, theorems are the result of applying a set of accepted rules of reasoning, as means of an analytical process, to (already established) axioms.

Examples of formal mathematical systems include Euclidean geometry, algebra, and set theory. In Euclidean geometry, primitive/basic/undefined objects include point, line, and plane. The first of the five axioms of Euclid asserts that it is possible to connect two points by a straight line-segment, which postulates formation of another object, a line-segment, as a consequence of the arrangement of other objects, namely the mere existence of two points. The first axiom also introduces a property/character for an arrangement consisting of two points, namely the length of the line-segment connecting two points. An exemplification of the interaction between objects is the crossing of three lines, which results in the formation of a new object; a triangle. An example of a proposition, is “triangle angle sum” theorem, which gives the magnitude of the sum of the three angles of a triangle as 180 degrees.

Topic F: Physics:

Random Physical Process: is a process, whose outcome changes in subsequent runs of the process and its observer cannot predict the (exact) outcome, before its realization. A random process may be (simply) selective or generative.

An observer's predictive inability is due to her/his lack of knowledge about the attributes influencing the outcome of the process; either by not knowing the exact input-output relationship or by not knowing the exact state of the influencing attributes due to unavailability of means of detection/sensation.

Selective Random Process: is a process, in which an agent selects a subset from a reference/source set of already produced objects.

A Selective process, by itself, does not result in generation of new objects or in modification of the (physical) attributes of existing objects, other than the objects' (physical or conventional) address. If an observer knows the "how of selection" by knowing the address of every object in the reference set and the address to which the selective agent reaches out, she/he can predict the outcome of the process. For such knowledgeable observer, the selection process is predictable and not random, even as the outcome varies along subsequent runs/trials. For example, if an observer knows (1) the entire sequence of (varying/mixed) numbers a "random function" uses to generate a random number, and (2) the address (of the position), where the function reaches, at the beginning of a new (selective) run, she/he can predict the sequence to be selected and exposed by the process.

Generative Random Process: (to an observer) is a process, whose exact outcome, in each run of the process, is not predictable by the observer.

A generative process results in the production of new physical objects or modification of physical attributes of existing objects.

Topic G: Modeling (of Reality):

Class: is a collection of objects, who are grouped based on affinities between a number of their features; the defining characteristics of the class.

Class members can be distinguished by slight differences in their defining characters, as well as (potential) differences in their non-defining attributes.

Unit Magnitude: (of a defining attribute) is the (numeric) amount of the physical character defined as one, usually based on an agreed standard measurement system. The magnitude of the character in a class member informs the intensity of the attribute relative to the unit (of the measurement).

Attribute Range: (in a class) is the difference between the highest and lowest magnitudes across the class. There is no unique range for a class as range is attribute dependent and class members can be (potentially) described with reference to (either of their) attributes.

Class Size: is the number of the members of a class, which is different than the range of the magnitudes for class characteristics. While class size is an integer (number), the range can be a real number.

Indexing: (of a class) is labeling class members by assigning numbers.

While numerical labels can be assigned arbitrarily and without any physical basis, it is also possible to use defining physical attributes of a class, either individually or jointly, as a basis for indexing. An attribute-based indexing does not differentiate between the members with the same magnitude of the defining attribute(s).

Law of the Class Size Conservation: is the rule asserting that the attribute(s) chosen for indexing a class must not influence the total size of the class. In other words, while a change in the indexing attribute(s) changes the member indices, the total class size remains attribute indifferent.

Abstract Class: is a class, whose defining characters do not include spatial and/or temporal attributes. An abstract class is a "type" (of a class) with an unspecified size, which can be assumed as infinite.

Concrete Class: is a class, whose defining characters do include spatial and/or temporal attributes. A concrete class is a "token" of a class type and has a finite size, due to its spatiotemporal constraints.

Neighboring Class: is a class, in which the magnitudes of the defining characteristics of the class members are within a bounded range, a neighborhood (of numeric values).

Concentrated Neighboring Class: is a neighboring class, whose neighborhood (of the magnitudes of the defining characters) is very narrow.

In a visual representation of the magnitudes of the class, members are concentrated within an infinitesimal (numeric) range.

Spread Neighboring Class: is a neighboring class, whose neighborhood (of the magnitudes of the defining characters) is not narrow.

In a visual representation of the magnitudes of the class, members are not concentrated around a single point. A spread class can be parted into a (finite) sequence of concentrated neighboring classes. The sequence members inform the "composition" of the spread class.

Prevalence: (of a concentrated neighboring class in a spread class) is the proportion of the concentrated class size to the size of the spread class encompassing that concentrated class. The term prevalence can be substituted with terms such as relative frequency and (class) size proportion.

Prevalence Distribution: (of a spread neighboring class) is the sequence of the prevalence values for the concentrated neighboring classes composing a spread class.

Prevalence Density: (of a concentrated neighboring class in a spread class) is the prevalence of the class relative to the class range, i.e. the width of the neighborhood. Prevalence density at a point, defined by a specific magnitude of an indexing attribute, gives the prevalence of a concentrated neighboring class of unit width encompassing the point.

Topic H: Fundamental Concepts (of Thought):

Arrangement/Organization: is a way, in which a number of objects are positioned relative to each other. The position of a physical object can be spatial and/or temporal. Example arrangements include serial/ sequential/ chain-wise/ back-to-back/ in-succession and parallel/ concurrent/ side-by-side/ coexistent.

Plan: is an arrangement of the components/parts making a whole.

Plan of an action: is a (spatiotemporal) arrangement of the tasks (smaller actions) needed to be done, to achieve the end of the (purposive) action. The plan specifies the "what" and the "when" of the designated set of tasks.

Method of an action: is (the general and the specifics) of the process needed to be initiated and maintained, to achieve the intended end of an action. The method details the "how" of an action/task.

Planning: is the mental process/activity aimed at developing a plan.

APPENDIX B: FRAMEWORK FOR THE THEORY OF DECISION ENTROPY

1. INTRODUCTION

The Theory of Decision Entropy is intended to provide a basis to establish a non-informative sample space for the purpose of making decisions. The desired characteristics for this non-informative sample space are that it be rational and logical, consistent, and incorporate the deterministic properties of a decision (the set of alternatives, possible outcomes and associated consequences) but no information about the probabilities of outcomes. The Theory of Decision Entropy is derived from three principles that describe a non-informative sample space:

1. If no information is available about the probabilities of decision outcomes, then a selected alternative is equally probable to be or not to be the preferred alternative.

2. If no information is available about the probabilities of decision outcomes, then the possible differences in preference between a selected alternative and the preferred alternative are equally probable.

3. If no information is available about the probabilities of decision outcomes, then the possibilities of learning with new information about the selected alternative compared to the preferred alternative are equally probable.

2. FORMULATION OF DECISION AND UTILITY THEORY USING INFORMATION POTENTIAL

The Theory of Decision Entropy is derived from classical Decision Analysis and Utility Theory. This section summarizes Decision Analysis and Utility Theory and formulates it in a manner that provides the basis for the Theory of Decision Entropy.

The following notation will be used (Figure B.1):

- A_i is decision alternative i among n_A alternatives;
- $\vec{\theta}_k$ is a set of possible decision outcomes that affect the preference or utility associated with the decision alternatives amongst n_θ sets;
- $u(\vec{\theta}_k, A_i\ Selected)$ is the utility value that will be realized given that alternative A_i has been selected; and
- $P(\vec{\theta}_k | A_i\ Selected \cap S_{n_A})$ is the probability that decision outcome $\vec{\theta}_k$ will occur given that alternative A_i has been selected and S_{n_A} is the sample space (the subscript n_A is included with the sample space to denote that it is the sample space for making a decision between these n_A alternatives.

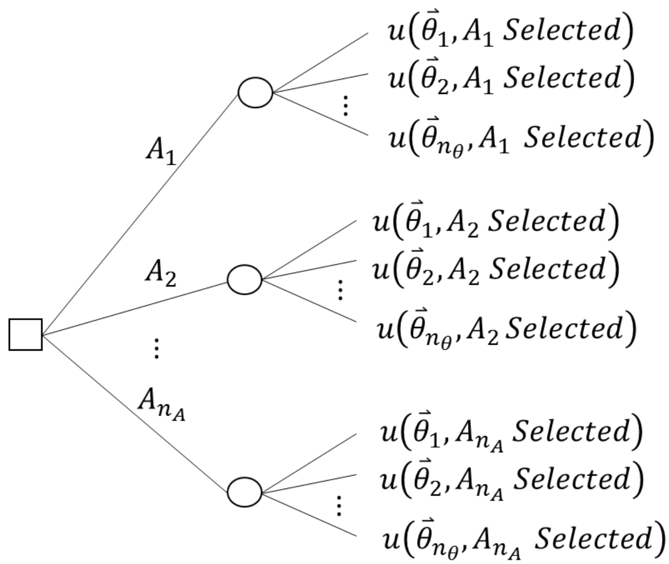

Figure B.1. Prior Decision Tree

In accordance with Utility Theory, the selected alternative with the maximum expected utility value is the preferred alternative:

$$\begin{aligned} &Preferred\ alternative\ A_j\ is\ such\ that\ E_\theta\left(u|A_j\ Selected \cap S_{n_A}\right) \\ &= max\left[E_\theta\left(u|A_i\ Selected \cap S_{n_A}\right)\ for\ all\ i\right] \end{aligned}$$

where

$$E_\theta\left(u|A_i\ Selected \cap S_{n_A}\right) = \sum_{\text{all } n_x} u\left(\vec{\theta}_k, A_i\ Selected\right) P\left(\vec{\theta}_k \middle| A_i\ Selected \cap S_{n_A}\right)$$

and $E_\theta\left(u|A_i\ Selected \cap S_{n_A}\right)$ is the expected utility value given that alternative Ai has been selected. Note that multiple alternatives are preferred if they all have the same expected utility value equal to the maximum expected utility value.

The maximum potential effect of information on the decision for each decision outcome is related to which alternative would be preferred if $\vec{\theta}_k$ occurred (Figure B.2):

$$\begin{aligned} &Given\ each\ \vec{\theta}_k, Preferred\ alternative\ A_{j|\vec{\theta}_k}\ is\ such\ that\ u\left(\vec{\theta}_k, A_j\ Selected\right) \\ &= max\left[u\left(\vec{\theta}_k, A_i\ Selected\right)\ for\ all\ i\right] \end{aligned}$$

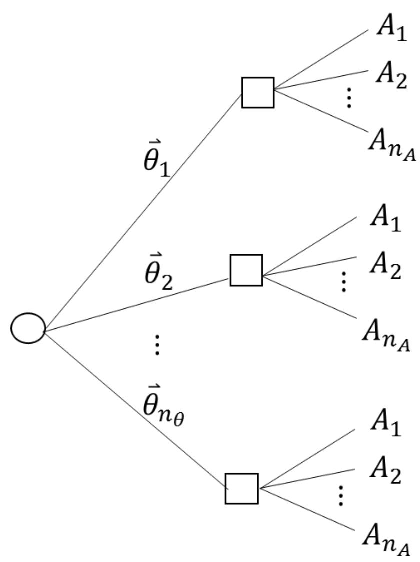

Figure B.2. Posterior Decision Tree

A quantitative measure of the maximum potential effect of information for each decision outcome is the difference between the utility value if alternative A_i is selected, $u(\vec{\theta}_k, A_j\ Selected)$, and the maximum utility value, $max[u(\vec{\theta}_k, A_i\ Selected)\ for\ all\ i]$:

$$\Delta u(\vec{\theta}_k, A_j\ Selected) = u(\vec{\theta}_k, A_j\ Selected) - max[u(\vec{\theta}_k, A_i\ Selected)\ for\ all\ i]$$

where $\Delta u(\vec{\theta}_k, A_j\ Selected)$ will be designated the information potential. The information potential is less than or equal to zero: $\Delta u(\vec{\theta}_k, A_j\ Selected) = 0$ if A_j is the preferred alternative and $\Delta u(\vec{\theta}_k, A_j\ Selected) < 0$ if A_j is not the preferred alternative. The more negative the value of $\Delta u(\vec{\theta}_k, A_j\ Selected)$, the greater the difference in utility values between the selected alternative and the preferred alternative for that decision outcome.

In conventional terms, the preferred alternative in the prior decision where $\vec{\theta}_k$ is uncertain is equivalent to the alternative with the maximum expected value of the information potential since subtracting the same value, $max[u(\vec{\theta}_k, A_i\ Selected)\ for\ all\ i]$, from the utility values for all alternatives, $u(\vec{\theta}_k, A_j\ Selected) for\ all\ j$, does not change the ordering of the utility values:

$$Preferred\ alternative\ A_j\ is\ such\ that\ E_\theta(\Delta u|A_j\ Selected \cap S_{n_A})$$
$$= max[E_\theta(\Delta u|A_i\ Selected \cap S_{n_A})\ for\ all\ i]$$

where

$$E_\theta(\Delta u|A_i\ Selected \cap S_{n_A}) = \sum_{\text{all } n_\theta} \Delta u(\vec{\theta}_k, A_i\ Selected) P(\vec{\theta}_k|A_i\ Selected \cap S_{n_A})$$

In addition, the value of perfect information about $\vec{\theta}$ given that alternative A_j has been selected (Figure B.3) is equal to the negative of $E_\theta(\Delta u|A_j\ Selected \cap S_{n_A})$:

$$Value\ of\ Perfect\ Information\ about\ \vec{\theta}\ if\ A_j\ Selected = -E_\theta(\Delta u|A_j\ Selected \cap S_{n_A})$$
$$= \sum_{\text{all } n_\theta} \{max[u(\vec{\theta}_k, A_i\ Selected)\ for\ all\ i]$$
$$- u(\vec{\theta}_k, A_j\ Selected)\} P(\vec{\theta}_k|A_j\ Selected \cap S_{n_A})$$

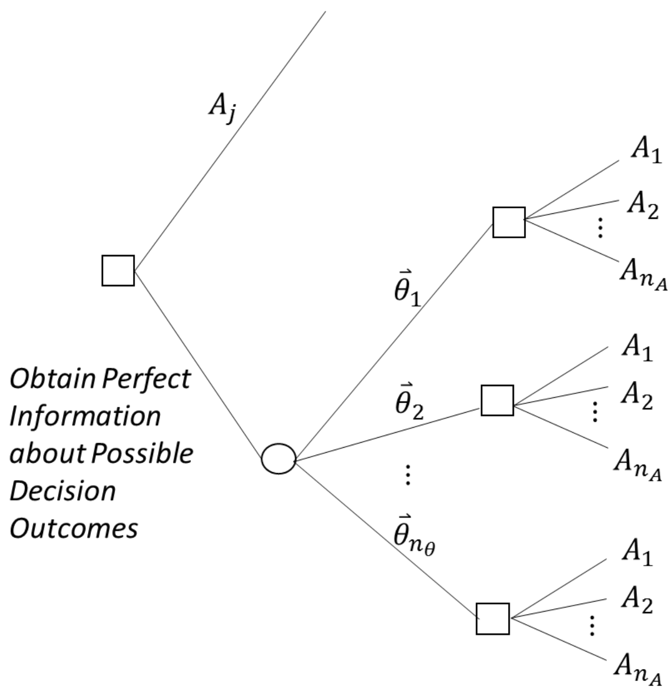

Figure B.3. Pre-Posterior Decision Tree for Perfect Information

When information is available, represented by $\vec{\varepsilon}_r$, the probabilities for different decision outcomes are updated through Bayes' Theorem:

$$P\left(\vec{\theta}_k \middle| \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right) = \frac{P\left(\vec{\varepsilon}_r \middle| \vec{\theta}_k \cap A_i\ Selected \cap S_{n_A}\right) \times P\left(\vec{\theta}_k \middle| A_i\ Selected \cap S_{n_A}\right)}{P\left(\vec{\varepsilon}_r \middle| A_i\ Selected \cap S_{n_A}\right)}$$

where

$$P\left(\vec{\varepsilon}_r \middle| A_i\ Selected \cap S_{n_A}\right) = \sum_{\text{all } n_\theta} P\left(\vec{\varepsilon}_r \middle| \vec{\theta}_k \cap A_i\ Selected \cap S_{n_A}\right) \times P\left(\vec{\theta}_k \middle| A_i\ Selected \cap S_{n_A}\right)$$

Again, the selected alternative with the maximum expected value for the information potential is the preferred alternative:

$$\begin{gathered} Preferred\ alternative\ A_k\ is\ such\ that\ E_\theta\left(\Delta u | \vec{\varepsilon}_r \cap A_k\ Selected \cap S_{n_A}\right) \\ = max\left[E_\theta\left(\Delta u | \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right) for\ all\ i\right] \end{gathered}$$

where $E_\theta\left(\Delta u | \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right)$ is the expected value of the information potential

$$E_\theta\left(\Delta u | \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right) = \sum_{\text{all } n_\theta} \Delta u\left(\vec{\theta}_k, A_i\ Selected\right) P\left(\vec{\theta}_k | \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right)$$

The conventional value of information for an information scheme $\boldsymbol{E} = \left(\vec{\varepsilon}_1, \vec{\varepsilon}_2, \ldots, \vec{\varepsilon}_{n_\varepsilon}\right)$ given that alternative A_j is selected in the initial sample space S_{n_A} (Figure B.4) is related to the expected information potential as follows:

$$\begin{aligned} &Value\ of\ Information\ for\ Set\ of\ Possible\ Information\ Outcomes\ \boldsymbol{E}\ if\ A_j\ Selected \\ &= \sum_{\text{all } n_\varepsilon} \begin{Bmatrix} max\left[E_\theta\left(\Delta u | \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right) for\ all\ i\right] \\ -E_\theta\left(\Delta u | \vec{\varepsilon}_r \cap A_j\ Selected \cap S_{n_A}\right) \end{Bmatrix} \\ &\times P\left(\vec{\varepsilon}_r \middle| A_j\ Selected \cap S_{n_A}\right) \end{aligned}$$

Note that while the probability for each possible information outcome, $P\left(\vec{\varepsilon}_r \middle| A_i\ Selected \cap S_{n_A}\right)$, depends on the decision alternative selected amongst the n_A alternatives, the probability for each possible information outcome in assessing the value of information, $P\left(\vec{\varepsilon}_r \middle| A_j\ Selected \cap\right.$

S_{n_A}), corresponds to having selected alternative A_j in the initial sample space. In other words, the value of information depends the information scheme, $\boldsymbol{E} = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \ldots, \vec{\varepsilon}_{n_\varepsilon})$, the effect of information on each of the decision alternatives being considered, $P(\vec{\theta}_k | \vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A})$, and the alternative that is preferred in the original sample space, alternative A_j such that $E_\theta(\Delta u | A_j\ Selected \cap S_{n_A}) = max[E_\theta(\Delta u | A_i\ Selected \cap S_{n_A})\ for\ all\ i]$.

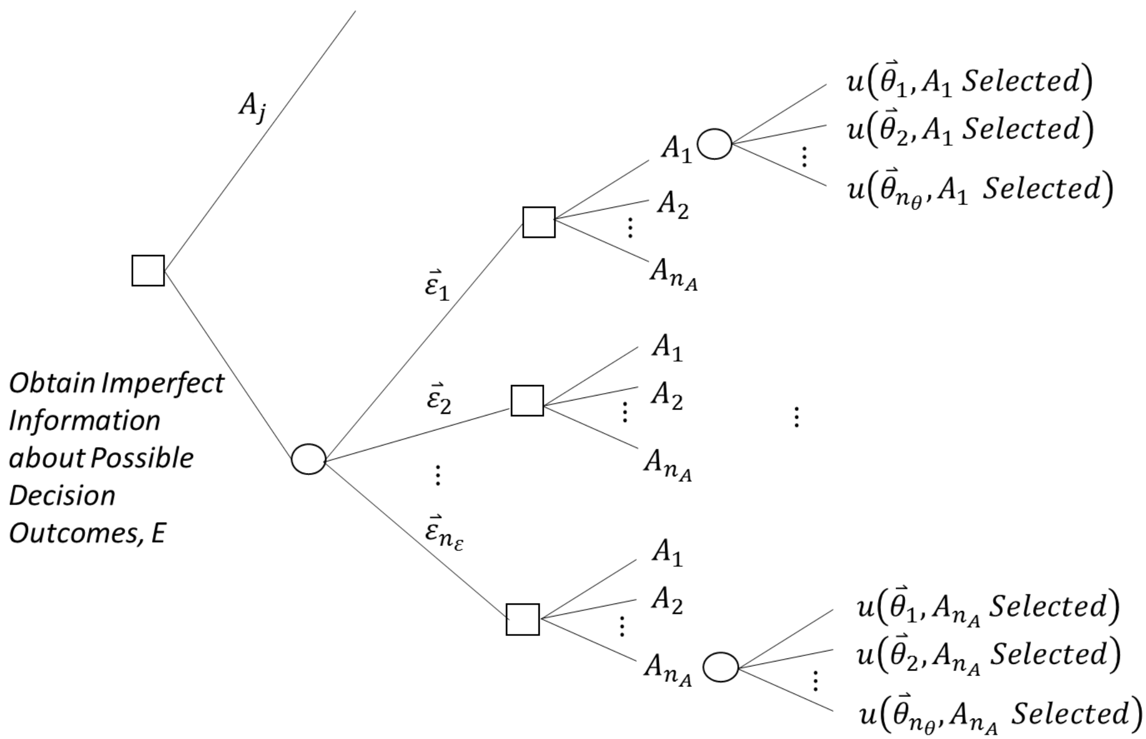

Figure B.4. Pre-Posterior Decision Tree for Imperfect Information

3. Basis for Non-Informative Sample Space for a Decision Alternative

The objective of the Theory of Decision Entropy is to establish a non-informative sample space for the decision outcomes, i.e., the n_θ values of $P\left(\vec{\theta}_k | A_i\ Selected \cap S_{n_A}\right)$ for each of the nA alternatives. The theory is developed from the following three principles to establish the maximum lack of information about the probabilities of decision outcomes in a decision:

1. If no information is available about the probabilities of decision outcomes, then a selected alternative is equally probable to be or not to be the preferred alternative.

2. If no information is available about the probabilities of decision outcomes, then the possible differences in preference between a selected alternative and the preferred alternative are equally probable.

3. If no information is available about the probabilities of decision outcomes, then the possibilities of learning with new information about the selected alternative compared to the preferred alternative are equally probable.

These principles can be expressed mathematically using three axioms that establish a non-informative sample space for the decision outcomes.

Principle Number 1: If no information is available about the probabilities of decision outcomes, then a selected alternative is equally probable to be or not to be the preferred alternative.

Axiom Number 1: Given that alternative A_i is selected, the probabilities for possible decision outcomes are those that maximize the relative entropy (see Appendix I for a description of relative entropy) for the events that this alternative is and is not preferred:

$$\begin{aligned}
&Maximize\ H_{rel}\left(Preference\ Outcome|\ A_i\ Selected \cap S_{n_A}\right) \\
&\quad = -P\left(A_i\ Preferred|A_i\ Selected \cap S_{n_A}\right) \\
&\quad \times ln\left[P\left(A_i\ Preferred|A_i\ Selected \cap S_{n_A}\right)\right] \\
&\quad - P\left(A_i\ \overline{Preferred}|A_i\ Selected \cap S_{n_A}\right) \\
&\quad \times ln\left[P\left(A_i\ \overline{Preferred}|A_i\ Selected \cap S_{n_A}\right)\right] - \ln(2) \\
&\quad = -\Big\{P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ |A_i\ Selected \cap S_{n_A}\right] \\
&\quad \times ln\left\{2P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0|A_i\ Selected \cap S_{n_A}\right]\right\} \\
&\quad + \left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ |A_i\ Selected \cap S_{n_A}\right]\right\} \\
&\quad \times ln\Big\{2\left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ |A_i\ Selected \cap S_{n_A}\right]\right\}\Big\}\Big\}
\end{aligned}$$

where $\overline{Preferred}$ denotes the complement of $Preferred$, or that A_i is not preferred.

For two states, $A_i\ Preferred$ and $A_i\ \overline{Preferred}$, the relative entropy $H_{rel}\left(Preference\ Outcome|\ A_i\ Selected \cap S_{n_A}\right)$ is maximized in the ideal case where

$$P\left(A_i\ Preferred|A_i\ Selected \cap S_{n_A}\right) = P\left(A_i\ \overline{Preferred}|A_i\ Selected \cap S_{n_A}\right) = 1/2$$

or

$$\begin{aligned}
&P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0|A_i\ Selected \cap S_{n_A}\right] \\
&\quad = 1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ |A_i\ Selected \cap S_{n_A}\right] = 1/2
\end{aligned}$$

If there are no possible decision outcomes, i.e., sets of $\vec{x}_k$, in which A_i is preferred or in which A_i is not preferred, then the probabilities for $P\left(A_i\ Preferred|A_i\ Selected \cap S_{n_A}\right)$ are 0 or 1 and the probabilities for $P\left(A_i\ \overline{Preferred}|A_i\ Selected \cap S_{n_A}\right)$ are 1 or 0 and the relative entropy becomes equal to its minimum possible value, $-\ln(2)$, because there is no uncertainty in the preference [i.e., $p\ln(p) \rightarrow 0$ as $p \rightarrow 0$ or $p \rightarrow 1$].

Principle Number 2: If no information is available about the probabilities of decision outcomes, then the possible differences in preference between a selected alternative and the preferred alternative are equally probable.

Axiom Number 2: Given that alternative A_i is selected and not preferred, the probabilities for possible decision outcomes are those that maximize the conditional relative entropy for the possible non-zero values of the information potential, $\Delta u(\vec{\theta}_k | A_i \; Selected)$:

$$Maximize \; H_{rel}(Information \; Potential | \; A_i \; \overline{Preferred} \cap A_i \; Selected \cap S_{n_A})$$
$$= \sum_{p=1}^{\mathrm{n}_{\Delta u_{\mathrm{i}<0}}} \begin{matrix} -P[\Delta u(\vec{\theta}_k, A_i \; Selected) = (\Delta u_{\mathrm{i}})_p | A_i \; \overline{Preferred} \cap A_i \; Selected \cap S_{n_A}] \\ \times \; ln\{P[\Delta u(\vec{\theta}_k, A_i \; Selected) = (\Delta u_{\mathrm{i}})_p | A_i \; \overline{Preferred} \cap A_i \; Selected \cap S_{n_A}]\} \end{matrix}$$
$$- \; ln(\mathrm{n}_{\Delta u_{\mathrm{i}<0}})$$
$$= -\sum_{p=1}^{\mathrm{n}_{\Delta u_{\mathrm{i}<0}}} \begin{matrix} P[\Delta u(\vec{\theta}_k, A_i \; Selected) = (\Delta u_{\mathrm{i}})_p | A_i \; \overline{Preferred} \cap A_i \; Selected \cap S_{n_A}] \\ \times \; ln\{\mathrm{n}_{\Delta u_{\mathrm{i}<0}} P[\Delta u(\vec{\theta}_k, A_i \; Selected) = (\Delta u_{\mathrm{i}})_p | A_i \; \overline{Preferred} \cap A_i \; Selected \cap S_{n_A}]\} \end{matrix}$$

where the set of possibilities for $\Delta u(\vec{\theta}_k, A_i \; Selected) < 0$ are divided into $\mathrm{n}_{\Delta u_{\mathrm{i}<0}}$ possible values, each designated $(\Delta u_{\mathrm{i}})_p$.

The maximum value for $H_{rel}(Information \; Potential | \; A_i \; \overline{Preferred} \cap A_i \; Selected \cap S_{n_A})$ is zero and it is realized when the possible values of $\Delta u(\vec{\theta}_k, A_i \; Selected) < 0$ are uniformly distributed between the minimum and the maximum possible values. Note that the number of sub-states for values or intervals of $(\Delta u_{\mathrm{i}})_p$, $\mathrm{n}_{\Delta u_{\mathrm{i}<0}}$, is not important in maximizing the relative entropy; the entropy is maximized when the possible intervals (however many there are) are as equally probable as possible.

The first and second axioms can be combined using the joint relative entropy of information where $(A_i\ Preferred)$ and $\left(A_i\ \overline{Preferred}\right)$ are the two main states and $\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_\mathrm{i})_p\right]$ are the $\mathrm{n}_{\Delta u_{\mathrm{i}<0}}$ sub-states within the state $\left(A_i\ \overline{Preferred}\right)$. Given that alternative A_i is selected, the probabilities for all possible decision outcomes are those that maximize the total relative entropy for possible preferences and information potentials:

$$
\begin{aligned}
&Maximize\ H_{rel}\left(Decision\ Outcome|\ A_i\ Selected \cap S_{n_A}\right) \\
&= H_{rel}\left(Preference\ Outcome|\ A_i\ Selected \cap S_{n_A}\right) \\
&+ H_{rel}\left(Information\ Potential|\ A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A}\right) \\
&\times P\left(A_i\ \overline{Preferred}\middle|A_i\ Selected \cap S_{n_A}\right) \\
&= -\left\{P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle|A_i\ Selected \cap S_{n_A}\right]\right. \\
&\times \ln\left\{2P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\middle|A_i\ Selected \cap S_{n_A}\right]\right\} \\
&+ \left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle|A_i\ Selected \cap S_{n_A}\right]\right\} \\
&\times \ln\left.\left\{2\left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle|A_i\ Selected \cap S_{n_A}\right]\right\}\right\}\right\} \\
&+ \left\{- \sum_{p=1}^{\mathrm{n}_{\Delta u_{\mathrm{i}<0}}} \begin{matrix} P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_\mathrm{i})_p\middle|A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A}\right] \\ \times \ln\left\{\mathrm{n}_{\Delta u_{\mathrm{i}<0}} P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = (\Delta u_\mathrm{i})_p\middle|A_i\ \overline{Preferred} \cap A_i\ Selected \cap S_{n_A}\right]\right\} \end{matrix} \right\} \\
&\times \left\{1 - P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle|A_i\ Selected \cap S_{n_A}\right]\right\}
\end{aligned}
$$

The maximum value for $H_{rel}\left(Decision\ Outcome|\ A_i\ Selected \cap S_{n_A}\right)$ is zero and it is realized when it is equally probable that the selected alternative will be preferred, i.e., $P\left[\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) = 0\ \middle|A_i\ Selected \cap S_{n_A}\right] = 1/2$, and the possible values of information potential when the selected alternative is not preferred, $\Delta u\left(\vec{\theta}_k, A_i\ Selected\right) < 0$, are uniformly distributed between the minimum and the maximum possible values.

Principle Number 3: If no information is available about the probabilities of decision outcomes, then the possibilities of learning with new information about the selected alternative compared to the preferred alternative are equally probable.

Mathematical Formulation: Given that alternative A_j is selected in the initial sample space and the set of possible information outcomes is $\boldsymbol{E} = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \dots, \vec{\varepsilon}_{n_\varepsilon})$, the probabilities for possible information outcomes are those that maximize the relative entropy for the possible Values of Information,

$$VI_{\boldsymbol{E},A_j,S_{n_A}} = \sum\nolimits_{\text{all } n_\varepsilon} \left\{ \begin{matrix} max\left[E_\theta\left(\Delta u|\vec{\varepsilon}_r \cap A_i\ Selected \cap S_{n_A}\right) for\ all\ i\right] \\ -E_\theta\left(\Delta u|\vec{\varepsilon}_r \cap A_j\ Selected \cap S_{n_A}\right) \end{matrix} \right\} \times$$

$P\left(\vec{\varepsilon}_r \middle| A_j\ Selected \cap S_{n_A}\right)$:

$$\begin{aligned} &Maximize\ H_{rel}\left(Value\ of\ Information|\boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right) = \\ &\quad = -\left\{P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right. \\ &\quad \times \ln\left\{2P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\} \\ &\quad + \left\{1 - P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\} \\ &\quad \left.\times \ln\left\{2\left\{1 - P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\}\right\}\right\} \\ &\quad + \left\{-\sum_{q=1}^{\mathrm{n_{vi>0}}} \begin{matrix} P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = (\mathrm{vi})_q\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right] \\ \times \ln\left\{\mathrm{n_{vi>0}}P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = (\mathrm{vi})_q\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\} \end{matrix}\right\} \\ &\quad \times \left\{1 - P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0\middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right]\right\} \end{aligned}$$

where the set of possibilities for $VI_{\boldsymbol{E},A_j,S_{n_A}} > 0$ are divided into $\mathrm{n}_{VI>0}$ possible values, each designated $(\mathrm{vi})_q$.

The maximum value for $H_{rel}\left(Value\ of\ Information|\boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right)$ is zero and it is realized when it is equally probable that the information will or will not have value, i.e., $P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} = 0 \middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right] = P\left[VI_{\boldsymbol{E},A_j,S_{n_A}} > 0 \middle| \boldsymbol{E} \cap A_j\ Selected \cap S_{n_A}\right] = 1/2$, and the possible values of the value of information when it has value, $VI_{\boldsymbol{E},A_j,S_{n_A}} > 0$, are uniformly distributed between the minimum and maximum possible values.

Note that in the case in which the set of information outcomes is equivalent to perfect information about the decision outcomes, Principles 1 and 2 are consistent with Principle 3. When no information is available, Principle 1 means that there is an equal probability that the selected alternative will or will not be the preferred alternative (i.e., there is an equal probability that perfect information will nor will not lead to changing the selected alternative and therefore realizing information potential). Furthermore when no information is available, Principle 2 means that the possible values of information potential are uniformly distributed outcomes in which the selected alternative is not preferred (i.e., the possible changes in utility when the selected alternative is not preferred are uniformly distributed between the minimum and maximum values).

Bibliography

Aleksandrov, A.D., Kolmogorov, A.N., Lavrent'ev, M.A. (1956), Математика; Ее содержание, методы и значение, Moscow, USSR, Translation as (1963), *Mathematics; Its Content, Methods, and Meaning*, translated and introduced by Gould, S.H. and Bartha, T., Cambridge MA, The M.I.T. press.

Baggini, J., and Fossi, P. S. (2010), *The Philosopher's Toolkit: A Compendium of Philosophical Concepts and Methods*, Second Ed., Wiley-Blackwell, West Sussex.

Balaguer, M. (1998), *Platonism and Anti-Platonism in Mathematics*, Oxford University Press, New York City NY.

Baker, T. J. (2005), "Mesh Generation: Art or Science." *Progress in Aerospace Sciences*, Vol. 41, 29-63.

Bayes, T. R. (1763), "An Essay Towards Solving a Problem in the Doctrine of Chances," *Philosophical Transactions of the Royal Society*, 53, 370-418. Reprinted in *Biometrika*, 45, 243-315, 1958.

Berger, J. O. (2006), "The Case for Objective Bayesian Analysis." *Bayesian Analysis*, Vol. 1, 385-402 and 457-464 (discussions).

Berkeley, G. (1734), *The Analyst: A Discourse addressed to an Infidel Mathematician*, J. Tonson, London.

Bernardo, J. M. (1997), "Non-Informative Priors Do Not Exist: A Dialogue with Jose Bernardo." *Journal of Statistical Planning and Inference*, Vol. 65, 159-189.

Bernoulli, J. (1713), *Ars Conjectandi*, 1, Opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola Gallicè scripta De ludo pilae reticularis, Thurnisiorum Fratrum, Basel, Switzerland, Translation as (2006), *The Art of Conjecturing*, translated and introduced by Sylla E.D., Johns Hopkins University Press, Baltimore MD.

Box, G. E. P., and Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading MA.

Bunge, M., (2006), *Chasing Reality: Strife over Realism*, University of Toronto Press, Toronto ON.

Burge, T., (2009), *Origins of Objectivity*, Oxford University Press, Oxford UK.

Carey, S. E., (2009), *The Origin of Concepts*, Oxford University Press, New York City NY.

Carnap, R. (1937), *Logical Syntax of Language*, Kegan Paul, Trench, Trubner & Co Ltd.

Carnap, R. (1950), *Logical Foundations of Probability*, University of Chicago Press, 2nd Ed. (1962).

Carnap, R. (1956), "The Methodological Character of Theoretical Concepts." In Feigl, H., and Scriven, M. (eds.), *Minnesota Studies in the Philosophy of Science; Vol I*, 38-76, University of Minnesota Press, Minneapolis MN.

Carnap, R. (1966), *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*, Basic Books Inc, New York City NY.

Cartwright, N. (1983), *How the Laws of Physics Lie*, Oxford University Press, New York City NY.

Caws, P. (1959) "The Functions of Definition in Science," *Philosophy of Science*, 26: 3, pp. 201-228.

Chernoff, H. (1954), "Rational Selection of Decision Functions," *Econometrica*, 22, 422-443.

Christian, B. (2011), *The Most Human Human*, Doubleday, New York City NY.

Christian, B. (2012) "Mind vs Machine," *The Atlantic*, March Issue.

Clarke, R.D. (1946) "An Application of the Poisson Distribution," *Journal of the Institute of Actuaries*, 72, pp. 481.

Conniffe, D. (1992) "Keynes on Probability and Statistical Inference and the Links to Fisher," *Cambridge Journal of Economics*, Vol. 16, pp. 475-489.

Cosmides, L., Barrett, H. C., and Tooby, J. (2010), "Adaptive Specializations, Social Exchange, and the Evolution of Human Intelligence," *Proceedings of the National Academy of Science USA*, Vol. 107, Supplement 2, 9007-9014.

Cox, R.T. (1946) "Probability, Frequency and Reasonable Expectation," *American Journal of Physics*, 14, pp. 1–13.

Cox, R. T. (1961), *The Algebra of Probable Inference*, Johns Hopkins University, Baltimore MD.

Crosby, A. W. (1997), *The Measure of Reality; Quantification and Western Society, 1250-1600*, Cambridge University Press, Cambridge UK.

Dale, A. I. (1999), *A History of Inverse Probability: From Thomas Bayes to Karl Pearson*, Second Ed., Springer-Verlag, New York City NY.

Damerow, P. (1996), A*bstraction and Representation: Essays on the Cultural Evolution of Thinking*, Springer Science + Business Media, Dordrecht, Netherlands.

Davis, P. J., Reuben, H., and Marchisotto, E.A. (1995), *The Mathematical Experience*, Birkhauser, Boston,

de Finetti, B. (1931), "Sul Significato Soggettivo Della Probabilita," *Fundamenta Mathematicae*, XVII, 298-329, Translated in de Finetti, B. (1993), *Induction and Probability*, ed. by Monari, P. and Cocchi, D., Bologna, CLUEB (supplement to *Statistica*, Vol. 52, No. 3), pp. 291-321.

de Finetti, B. (1937), "La Prevision: Ses Lois Logiques, Ses Sources Subjectives." *Annales de l'Institut Henri Poincare*, 7:1–68. Translated as "Foresight: Its Logical Laws, Its Subjective Sources," in H. E. Kyburg, H.E. and Smokler, H.E. eds., (1964), *Studies in Subjective Probability*. Wiley, New York, pp. 91–158.

de Boer, J. (1995), "On the History of Quantity Calculus and the International System" *Metrologia*, Vol. 31, 405-429.

de Finetti, B. (1972), *Probability, Induction, and Statistics*, John Wiley, New York City NY.

de Finetti, B. (1974, 1975), *Theory of Probability*, Vols. 1 and 2, John Wiley, New York City NY.

DeGroot, M.H. (1974), "Reaching a Consensus," *Journal of the American Statistical Association*, 69: 345, 118-121.

Dehaene, S. (2011), *The Number Sense: How the Mind Creates Mathematics*, Revised and Expanded Ed., Oxford University Press, New York City NY.

Dellantonio, S., and Pastore, L. (2017), *Internal Perception; the Role of Bodily Information in Concepts and Word Mastery*, Springer-Verlag, Berlin.

Deming, W. E. (1942), "On a Classification of the Problems of Statistical Inference," *Journal of the American Statistical Association*, Vol. 37, No. 218, 173-1852.

Deming, W. E. (1950), *Some Theory of Sampling*, John Wiley & Sons, New York City NY.

Deming, W. E. (1975), "On Probability as a Basis for Action," *The American Statistician*, Vol. 29, No. 4, 146-152.

Diamond, J. M. (2005), *Guns, Germs, and Steel: The Fates of Human Societies*, Expanded Ed., W. W. Norton & Company, New York City NY.

Doob, J. L. (1996), "The Development of Rigor in Mathematical Probability." *The American Mathematical Monthly*, Vol. 103, No. 7, 586-595.

Duhem, P. (1914), *La Theorie Physique: Son Objet, Sa Structure*, Paris, France, Marcel Riviere & Cie, Translation as (1954), *The Aim and Structure of Physical Theory*, 2nd Edition, translated by Wiener P.P., Princeton University Press, Princeton NJ.

Eaton, M. L. (1989), "Group Invariance Applications in Statistics." *Regional Conference Series in Probability and Statistics*, Vol. 1, I–133.

Eaton, M. L. and Freedman, D. A. (2004), "Dutch book against some 'Objective' priors." *Bernoulli*, 10, 861–872.

Eaton, M. L. and Sudderth, W. D. (2010), "Invariance of Posterior Distributions under Reparametrization." *Sankhyā: The Indian Journal of Statistics; Series A (2008-)*, 72:1, 101–118.

Espland, W. M. and Stevens, M. L. (2008), "A Sociology of Quantification." *European Journal of Sociology*, 49:3, 401-436.

Edelsbrunner, H. (2001), *Geometry and Topology for Mesh Generation*, Cambridge University Press, New York City NY.

Edwards, A. W. F. (1992), *Likelihood*, Expanded Ed., Johns Hopkins University Press, Baltimore MD.

Emerson, W.H. (2008), "On Quantity Calculus and Units of Measurement" *Metrologia*, Vol. 45, 134-138.

Ernest, P. (1998), *Social Constructivism as a Philosophy of Mathematics*, State University of New York Press, Albany NY.

Energo, (2010), *Assessment of Damage and Failure Mechanisms for Offshore Structures and Pipelines in Hurricanes Gustav and Ike*, Final Report to U.S. Mineral Management Services prepared by Energo Engineering, MMS TAR No: 642, 132 p.

Euler, L. (1748), *Introductio in analysin infinitorum*, Lausanne, apud Marcum-Michaelem Bousquet & socios, as Euler, L. (1988), *Introduction to Analysis of the Infinite*, translated by Blanton J.D., Springer-Verlag Inc., New York City NY.

Evans, J. B. T. and Frankish, K., (2009), *In Two Minds: Dual Processes and Beyond*, Oxford University Press, Oxford UK.

Fauconnier, G. (1997), *Mapping in Thought and Language*, Cambridge University Press, Cambridge UK.

Fauconnier, G., and Turner, M. (2002), *The Way We Think; Conceptual Blending and the Mind's Hidden Complexities*, Basic Books, New York City NY.

Feng, K., Mostofi, A., Habibi M., Gilbert R.B, and Nadim F. (2019), "Decision-Based Approach to Account for Uncertainty in Estimating the Overtopping Hazard to Manage Risk for Dams", *Proceedings of ISGSR 2019*, Dec 11-13, Taipei, Taiwan.

Ferguson, N. C. (2011), *Civilization: The West and the Rest*, Penguin Books Ltd., London UK.

Fienberg, S.E. (2006a), "When Did Bayesian Inference Become "Bayesian"?" *Bayesian Analysis*, Vol. 1, No. 1, 1–41.

Fienberg, S.E. (2006b), "Does it Make Sense to be an "Objective Bayesian"? (Comment on Articles by Berger and by Goldstein)" *Bayesian Analysis*, Vol 1, No. 3, 429–432.

Fisher, R. A. (1922), "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transitions of the Royal Society of London*, Ser. A, 222, 309-368.

Fisher, R. A. (1925a), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh UK.

Fisher, R. A. (1925b), "Theory of Statistical Estimation," *Proceedings of the Cambridge Philosophical Society*, 22, 700-725.

Fisher, R. A. (1950), *Contributions to Mathematical Statistics*, John Wiley & Sons, New York City NY.

Fisher, R. A. (1936), "Uncertain Inference," *Proceedings of American Academy of Arts and Science*, 71, 245-58. Reprinted in Fisher (1950).

Fisher, R. A. (1955), "Statistical Methods and Scientific Induction," *Journal of Royal Statistical Society B*, 17, 69-78.

Frege, G. (1884), *Die Grundlagen der Arithmetik; Eine Logisch-mathematische Untersuchung über den Begriff der Zahl*, Breslau, Verlage Wilhelm Koebner, as Frege, G. (1960), *The Foundations of Arithmetic; a Logico-Mathematical Inquiry into the Concept of Number*, translated by Austin J.L., 2nd Edition, Harper & Brothers, New York City NY.

Frankish, K. (2010), "Dual Process and Dual-System Theories of Reasoning," *Philosophy Compass*, 5/10, 914-926.

Freudenthal, H. (1960), "Models in Applied Probability," *Synthese*, 12, 204-210.

Frey, P. J. and George, P.L. (2008), *Mesh Generation: Application to Finite Elements*, 2nd Ed., John Wiley & Sons Inc., Hoboken NJ.

GEER (2014), "The 22 March 2014 Oso Landslide, Snohomish County, Washington," Contributing Authors: Keaton, J.R., Wartman, J., Anderson, S., Benoit, J. deLa Chapelle, J.

Gelman, A. and Henning, C. (2017), “Beyond Subjective and Objective in Statistics,” *J. of Royal Statistical Society; Series A: Statistics in Society*, 180: 4, 967-1033.

Gilbert R.B, Habibi M., and Nadim F. (2016), “Chapter 18: Accounting for Unknown Unknowns in Managing Multi-Hazard Risks", In: *Multi-hazard Approaches to Infrastructure Engineering*, Gardoni P., and LaFave J.M. (Editors), Springer Int., Switzerland, ISBN 978-3-319-29711-8.

Gilbert R.B, and Habibi M. (2015), "Chapter 13: Assessing the Value of Information to Design Site Investigation and Construction Quality Assurance Programs,” In: *Risk and Reliability in Geotechnical Engineering*, Phoon K.K. and Ching J. (Editors), CRC Press, Taylor and Francis Group, Boca Raton FL, ISBN 9781482227215

Gilbert R.B, Habibi M., and Min N. (2012), "Achieving Reliability in the Face of Extreme Uncertaity", *Fifth Asian-Pacific Symposium on Structural Reliability and its Applications* (5APSSRA), My 23-25, Singapore.

Gillies, D. A. (1973), *An Objective Theory of Probability*, Methuen & Co., London UK.

Gillies, D. A. (2000), *Philosophical Theories of Probability*, Routledge, London UK.

Goldblatt, R. I. (1984), *Topoi: The Categorical Analysis of Logic*, Revised Ed., Elsevier Science Publishing Co., Amsterdam, Netherlands.

Goldstein, M. (2006), “Subjective Bayesian Analysis: Principles and Practice.” *Bayesian Analysis*, 1, 403-420.

Good, I. J. (1950), *Probability and the Weighing of Evidence*, Griffin, London UK.

Good, I. J. (1959), “Kinds of Probability.” *Science*, Vol. 129, No. 3347, 443-447.

Good, I. J. (1962), “Subjective Probability as the Measure of a Non-Measurable Set.” in *Logic, Methodology and Philosophy of Science*, Stanford University Press, Stanford CA.

Good, I. J. (1965), *The Estimation of Probabilities*, M.I.T. Press, Cambridge MA.

Good, I. J. (1972), “46,656 varieties of Bayesians.” *American Statistician*, 25, 62–63.

Good, I. J. (1983), *Good Thinking: The Foundations of Probability and its Applications*. University of Minnesota Press, Minneapolis MN.

Good, I. (1976), “The Bayesian Influence, or How to Sweep Subjectivism Under the Carpet.” In Hooker, C. and Harper, W. (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, volume 2, 125–174, Dordrecht, Netherlands.

Good, I. J. (2006), “Belief, Degrees of.” In Kotz, S., Balakrishnan, N., Read, C. B., Vidakovic, B. and Johnson, N. L. (eds.), *Encyclopedia of Statistical Science*, Second Ed., 457-463, Wiley-Interscience, Hoboken N.J.

Habibi M., Gilbert R.B, Lake L.W., and McIntosh P. (2014), “Assessing Value of Test Wells in Developing an Unconventional Play – Decision Entropy Theory", *Proceedings of SPE HEES*, May 19-20, Houston, TX.

Halmos, P. R. (1974), *Naïve Set Theory*, Springer-Verlag, New York City NY.

Hand, D. J. (2004), *Measurement Theory and Practice*, Arnold, London UK.

Hand, D. J. (2016), *Measurement; a Very Short Introduction*, Oxford University Press, Oxford UK.

Harbaugh, J. W., Davis, J. C., and Wendebourg J. (1995), *Computing Risk for Oil Prospects: Principles and Programs*, Pergamon Press, Oxford UK.

Hardy, G. H. (1940), *A Mathematician's Apology*, Cambridge University Press, Cambridge UK.

Hauser, M. D., Chomsky, A. N., and Tecumseh Fitch, W. (2002), “The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?” *Science*, Vol. 298, Issue 5598, pp. 1569-1579.

Hausman, D. M. (2012), *Preference, Value, Choice, and Welfare*, Cambridge University Press, New York City NY.

Hempel, C. G. (1964), *Fundamentals of Concept Formation in Empirical Science*, University of Chicago Press, Chicago IL.

Hersh, R. (1997a), “Math Lingo vs. Plain English: Double Entendre,” *The American Mathematical Monthly*, 104: 1, 48-51.

Hersh, R. (1997b), *What Is Mathematics, Really?*, Oxford University Press, New York City NY.

Hilbert, D. (1993), *Theory of Algebraic Invariants*, Translated by R.C. Laubenbacher, Cambridge University Press, Cambridge UK.

Holger, A. (2017), “Theoretical Terms in Science.” In Zalta, E. N. (ed), *The Stanford Encyclopedia of Philosophy*, Stanford University.

Hopf, E. (1934) “On causality, Statistics and Probability,” *Journal of Mathematics and Physics*, 13, pp. 51-102.

Hossenfelder, S. (2018), *Lost in Math: How Beauty Leads Physics Astray*, Basic Books, New York City NY.

Howie, D. (2002), *Interpreting Probability; Controversies and Developments in the Early Twentieth Century*, Cambridge University Press, Cambridge UK.

Howson, C., and Urbach, P. (2006), *Scientific Reasoning: The Bayesian Approach*, 3rd Edition, Open Court, Chicago IL.

Hume, D. (1748), *An Enquiry Concerning Human Understanding*, A Millar, London UK.

Hurley, P. J. (2015), *A Concise Introduction to Logic*, 12th Edition, Cengage Learning, Stamford CT.

Ioannidis, J. P. A. (2005), "Why Most Published Research Findings are False," *PLoS Medicine*, 2: 8, 696-701.

IPET, (2009), *Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System*, Final Report, Interagency Performance Evaluation Task Force, U.S. Army Corps of Engineers.

Jackendoff, R. (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford University Press, Oxford UK.

Jackendoff, R. (2011), "Conceptual Semantics." In Maienborn, C., von Maienborn, K. and Portner, P. (eds.), *Semantics: An International Handbook of Natural Language Meaning; Vol I*, 688-709, De Gruyter Mouton, Berlin, Germany.

Jackendoff, R. (2012), *A User's Guide to Thought and Meaning*, Oxford University Press, Oxford UK.

Jaynes, E. T. (1957), "Information Theory and Statistical Mechanics I, II," *Physical Review*, 106, 620-630; 108, 171-190.

Jaynes, E. T. (1968), "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, SSC—4, 227-241.

Jaynes, E. T. (2003), *Probability Theory: The Logic of Science*, (ed. Bretthorst, G. L.), Cambridge University Press, Cambridge UK.

Jeffreys, H. (1946), "An Invariant Form for the Prior Probability in Estimation Problems," *Proceedings of the Royal Society of London*, Ser. A, 186, 453-161.

Jeffreys, H. (1955), "The Present Position in Probability Theory," British Journal for Philosophy of Science, 5, 275-289.

Jeffreys, H. (1957), *Scientific Inference*, 2nd Ed., Cambridge University Press, Cambridge U.K.

Jeffreys, H. (1961), *Theory of Probability*, 3rd ed., Oxford University Press, London UK.

Jeffreys, H. (1963), “Review of The Foundations of Statistical Inference”, by L. J. Savage, M. S. Bartlett, G. A. Barnard, D. R. Cox, E. S. Pearson, and C. A. B. Smith, *Teclinornetrics*, 5, 407-410.

Johansson, I. (2010), “Metrological Thinking Needs the Notions of Parametric Quantities, Units and Dimensions"” *Metrologia*, Vol. 47, 219-230.

Johnson, N. L. and Kotz, S. (1977), *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory*, John Wiley & Sons, New York City NY.

Kahneman, D. (2011), *Thinking Fast and Slow*, Farrar, Straus, and Giroux, New York City NY.

Kass, R. E. and Wasserman, L. (1996), “The Selection of Prior Distributions by Formal Rules,” *J. Amer. Statist. Assoc*., 91, 1343-1370.

Kemble, E. C. (1942), “Is the Frequency Theory of Probability Adequate for All Scientific Purposes?,” *American Journal of Physics*, 10, 6-16.

Kendall, M. G. (1973), “Entropy, Probability, and Information,” *International Statistical Review*, Vol 41, No 1, 59-68.

Keynes, J. M. (1921), *A Treatise on Probability*, volume 8. St Martin’s.

Kitcher, P. S. (1981), “Mathematical Rigor- Who needs it?” *Noûs*, Vol 15, No 4, 469-493.

Kitcher, P. S. (1984), *The Nature of Mathematical Knowledge*, Oxford University Press, New York City NY.

Kleiner, I. (1989), “Evolution of the Function Concept: A Brief Survey,” *The College Mathematics Journal*, Vol 20, No 4, 282-300.

Kline, M. (1959), *Mathematics and the Physical World*, Thomas Y. Crowell Co., New York City NY.

Kline, M. (1977), *Calculus: An Intuitive and Physical Approach*, 2nd Ed, John Wiley & Sons, Inc., New York City NY.

Knight, F. H. (1921), *Risk, Uncertainty, and Profit*, Haughton Mifflin Company, Boston MA.

Kolmogorov, A. N. (1933), *Grundbegriﬀe der Wahrscheinlichkeitsrechnung*, Berlin, Springer. Translation as (1956), *Foundations of the Theory of Probability*, 2nd Edition, translation edited by Morrison N., Chelsea Publishing Company, New York City NY.

Kullback, S. (1959), *Information Theory and Statistics*, John Wiley & Sons, New York City NY.

Kullback, S. and Leibler R.A. (1951), “On Information and Sufficiency,” *The Annals of Mathematical Statistics*, Vol 22, No 1, 79-86.

La Mettrie, J.O. (1747), *L'homme Machine*, Paris, Translation as (1912), *Man a Machine*, translated and introduced by Bussey, G. C., The Open Court Publishing Co., Chicago IL.

Lakoff, G. P., (1993), “The Contemporary Theory of Metaphor", In: *Metaphor and Thought*, Ortony A. (Editor), Cambridge University Press, Cambridge UK.

Lakoff, G. P., and Johnson, M. L. (1980), *Metaphors We Live By*, 2nd Ed., The University of Chicago Press, Chicago IL.

Lakoff, G. P., and Nunez, R. E. (2000), *Where Mathematics Come From; How the Embodied Mind Bring Mathematics into Being*, Basic Books, New York City NY.

Laplace, P.-S. (1774), “Mémoire sur la probabilité des causes par les événements,” *Mémoires de mathématique et de physique, présentés à l'Académie royale*, Par Divers Savans, & Lˆus dans ses Assembl´ees, 6, 621–656, Translated by Stigler, M. S., as (1986), “Memoir on the Probability of the Causes of Events,” *Statistical Science,* Vol. 1, No. 3, pp. 364-378

Laplace, P.-S. (1812), *Théorie Analytique des Probabilités,* Courcier, Paris.

Laplace, P.-S. (1814), *Essai Philosophique sur les Probabilités*, Courcier, Paris, Translated from the 6th Ed. by Truscott, F. W., and Emory, F., L. as (1902), *Philosophical Essays on Probabilities*, John Wiley and Sons, New York City NY.

Lewis, C. I. (1929), *Mind and the World-Order*, Charles Scribner’s Sons, New York City NY.

Lo, S. H. (2015), *Finite Element Mesh Generation*, CRC Press, Boca Raton FL.

Logan, R. K. (2007), *The Extended Mind: The Emergence of Language, the Human Mind*, Toronto Studies in Semiotics and Communication, University of Toronto Press, Toronto ON.

Luce, R. D. and Raiffa, H. (1957), *Games and Decisions*, John Wiley & Sons, New York City NY.

MacKinnon, E. Jr. (2011), *Interpreting Physics: Language and the Classical/Quantum Divide*, Boston Studies in the Philosophy of Science: Vol. 289, Springer Science + Business Media, Dordrecht, Netherlands.

MacKinnon, E. Jr. (2013), “The Origin of Quantification,” *Open Journal of Philosophy*, Vol 3, No. 4A, 6-9.

Manheim, J. (1961), “Word Problems or Problems with Words,” *The Mathematics Teacher*, Vol. 54, No. 4, 234-238.

Mahmoud H. M. (2009), *Polya Urn Models*, CRC Press, Boca Raton FL.

Marconi, D. (1997), *Lexical Competence,* M.I.T. Press, Cambridge MA.

Martin, F. G. and England, J.W. (1984), *Mathematical Theory of Entropy*, Cambridge University Press, Cambridge UK.

Mazur, J. C. (2008), *Zeno's Paradox: Unraveling the Ancient Mystery behind the Science of Space and Time*, Plume, New York City NY.

Menger, K. (1954), "On Variables in Mathematics and in Natural Science," *The British Journal for the Philosophy of Science*, Vol. 5, No. 18, 134-142.

Michalowicz, J. V., Nichols, J. M., and Bucholtz, F. (2014), *Handbook of Differential Entropy*, CRC Press, Boca Raton FL.

Mill J. S. (1863), *A System of Logic; Ratiocinative and Inductive*, Longsman Green, London UK.

Mill J. S. (1843), *Utilitarianism*, Parker, Son, & Bourn, West Strand, London UK.

Min, N. (2008), *A Method to Establish Non-Informative Prior Probabilities for Risk-Based Decision Analysis*, Ph.D. Dissertation, the University of Texas at Austin.

Mostofi, A. (2018), *Development of a New Decision-Based Framework for Risk Assessment and Management of Landslides*, Ph.D. Dissertation, the University of Texas at Austin.

Mostofi, A., Gilbert R.B., Montgomery, D.R., and Wartman J., (2019), "Assessing Recurrence Probability for Oso 2014 Landslide in Order to Manage Risk." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, Vol. 13, No. 4, 333-340, DOI: 10.1080/17499518.2019.1623408

Nadim F. (2007), "Tools and Strategies for Dealing with Uncertainty in Geotechnics", In: *Probabilistic Methods in Geotechnical Engineering*, Griffiths D. V., and Fenton G. A. (Editors), Springer, Vienna, ISBN 978-3-211-73365-3.

Newendorp, P. D. and Schuyler, J. R. (2000), *Decision Analysis for Petroleum Exploration*, 2nd Ed., Planning Press, Aurora CO.

Neyman, J. (1937), "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability," *Philosophical Transactions of the Royal Society of London; Series A: Mathematical and Physical Sciences*, Vol. 236, No. 767, 333-380.

Neyman, J. (1942), "Basic Ideas and Some Recent Results of the Theory of Testing Statistical Hypotheses," *Journal of the Royal Statistical Society*, Vol. 105, No. 4, 292-327.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E. and Rakow, T. (2006), *Uncertain Judgements: Eliciting Expert Probabilities*, John Wiley & Sons, Chichester UK.

OGJ, (1967), *Oil and Gas Journal*, Tulsa, PennWell Corporation, Vol. 65, Dec. 11.

OGJ, (1968), *Oil and Gas Journal*, Tulsa, PennWell Corporation, Vol. 66, Sep. 23.

Pap, A. (1964), "Theory of Definition," *Philosophy of Science*, Vol. 31, 49-54.

Pavlov, I. P. (1927), *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*, Oxford University Press, London UK.

Pearl, J., and Mackenzie, D. (2018), *The Book of Why; The New Science of Cause and Effect*, Basic Books, New York NY.

Piaget, J. (1977), *Recherches sur l'Abstraction Réfléchissante*, Presses Universitaires de France, Paris, Translated by Campbell, R. L. as (2001), *Studies in Reflecting Abstraction*, Psychology Press, Sussex, UK.

Pinker, S. A. (2007), *The Stuff of Thought: Language as a Window into Human Nature*, Penguin Group, New York City NY.

Pinker, S. A. (2010), "The Cognitive Niche: Coevolution of Intelligence, Sociality, and Language," *Proceedings of the National Academy of Science USA*, Vol. 107, Supplement 2, 8893-8999.

Poisson, S. (1837), *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile, Précédées des Règles Générales du Calcul des Probabilité*, Bachelier, Paris. Translated as (2013), *Researches into the Probabilities of Judgements in Criminal and Civil Cases, Precedents of the General Rules of Probability Calculation*, Sheynin, O., Berlin, ISBN 978-3-942944-27-4.

Popper, K. R. (1959), "Propensity Interpretation of Probability," *The British Journal for the Philosophy of Science*, Vol. 10, No. 34, 25-42.

Popper, K. R. (1963), *Conjectures and Refutations; The Growth of Scientific Knowledge*, Routledge & Kegan Paul, London UK.

Popper, K. R. (1979), *Objective Knowledge; An Evolutionary Approach*, Revised Ed., Oxford University Press, Oxford UK.

Popper, K. R. (1990), *A World of Propensities*, Thoemmes Press, Bristol UK.

Popper, K. R., and Eccles, J. C. (1977), *The Self and its Brain; An Argument for Interactionism*, Springer-Verlag, Berlin, Germany.

Porter, T. M. (1995), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton University Press, Princeton NJ.

Ramsey, F. P. (1926), "Truth and Probability," Published in 1931 as *Foundations of Mathematics and Other Logical Essays*, Ch. VII, pp. 156–198. Edited by R.B. Braithwaite. Kegan, Paul, Trench, Trubner & Co., London. Reprinted in Kyburg, H.E. and Smokler, H.E. eds. (1964), *Studies in Subjective Probability*, John Wiley and Sons, New York City NY, 61–92.

Redlich, O. (1970), "Intensive and Extensive Properties," *Journal of Chemical Education*, Vol. 47, No. 2, 154-156.

Rescher, N. (2001), *Paradoxes: Their Roots, Range, and Resolution*, Open Court, Chicago IL.

Ribot, T. (1899), *the Evolution of General Ideas*, Translated by Welby F.A., Open Court, Chicago IL.

Robert, C. P. (2007), *The Bayesian Choice; from Decision-Theoretic Foundations to Computational Implementation*, Second Ed., Springer, New York City NY.

Robert, C. P., (2011), "Reading Keynes' Treatise on Probability," *International Statistical Review*, Vol. 79, No. 1, pp. 1-15.

Robert, C. P., Chopin, N., and Rousseau, J. (2009), "Harold Jeffrey's Theory of Probability Revisited," *Statistical Science*, Vol. 24, No. 2, pp. 141-172.

Robinson, R. (1950), *Definition*, Oxford University Press, Oxford UK.

Rossi, R. J. (2006), *Theorems, Corollaries, Lemmas, and Methods of Proof*, John Wiley & Sons, Inc., Hoboken NJ.

Russell, B. A. W. (1946), *A History of Western Philosophy*, George Allen & Unwin Ltd., London UK.

Russell, B. A. W. (1948), *Human Knowledge; Its Scope and Limits*, George Allen & Unwin Ltd., London UK.

Sainsbury, R. M. (2009), *Paradoxes*, 3rd Ed., Cambridge University Press, Cambridge U.K.

Sarewitz, D., Pielke Jr. R. A., and Byerly, R. (2000), *Prediction: Science, Decision Making, and the Future of Nature*, Island Press, Washington D.C.

Savage, L. J. (1962), "Bayesian Statistics," in *Recent Developments in Information and Decision Theory*, eds. R. F. Machol and P. Gray, New York: Macmillan, reprinted in *The Writings of Leonard Jimmie Savage-A Memorial Selection* (1981), Washington, DC: American Statistical Association and Institute of Mathematical Statistics.

Savage, L. J. (1972), *The Foundations of Statistics*, Second Revised Ed., Dover Publications Inc, New York City NY.

Savage, L. J. (1976), "On Reading R. A. Fisher," *The Annals of Statistics*, Vol. 4, No. 3, 441-500.

Savage, L. J., Bartlett, M. S., Barnard, G. A., Cox, D. R., Pearson, E. S., and Smith, C. A. B. (1962), *The Foundations of Statistical Inference,* Methuen & Co. Ltd., London UK.

Sawyer, W. W. (1962), *What is Calculus About?*, Washington D.C., The Mathematical Association of America.

Skilling, J. (1985), "Prior Probabilities," *Synthese*, 63: 1, 1-34.

Stanovich, K. E. (2004), *The Robot's Rebellion; Finding Meaning in the Age of Darwin*, University of Chicago Press, Chicago IL.

Stevens, S.S. (1946), "On the Theory of Scales of Measurement," *Science*, 103:2684, 677-680.

Stigler, S. M. (1980), "Stigler's Law of Eponymy," *Transactions of the New York Academy of Sciences*, Vol. 39, 147-158.

Stigler, S. M. (1982), "Poisson on the Poisson Distribution," *Statistics & Probability Letters*, 1, 33-35.

Stigler, S. M. (1999), *Statistics on the Table: The History of Statistical Concepts and Methods*, Harvard University Press, Cambridge MA.

Schultz, W. (2015), "Neuronal Reward and Decision Signals: From Theories to Data," *Physiological Reviews*, 95:3, 853-951.

Tabak, J. (2004), *Numbers: Computers, Philosophers, and the Search for Meaning,* Facts on File Inc., New York City NY.

Taleb, N. N. (2007), *The Black Swan: The Impact of the Highly Improbable*, Random House Inc., New York City NY.

Toffoli, T. (2004), "Honesty in inference," *American Scientist,* Vol. 92, 182–185.

Tolman, R. C. (1917), "The Measurable Quantities of Physics," *Physical Review*, Vol. 9, No. 3, 237-253.

Tribus, M. (1969), *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York City NY.

Turnbell, C. (2017), *A History of British Actuarial Thought*, Palgrave Macmillan, Switzerland.

Turner, M., (1988), "Categories and Analogies", In: *Analogical Reasoning; Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy,* Helman D. H. (Editor), Springer Science + Business Media, Dordrecht, Netherlands.

von Mises, R. (1957), *Probability, Statistics, and Truth*, 2nd Revised English Ed., George Allen & Unwin Ltd., New York City NY.

von Mises, L. H. E. (1966), *Human Action: A Treatise on Economics*, Fourth Revised Ed., Fox and Wilkes, San Francisco CA.

von Neumann, J. and Morgenstern, O., (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton NJ.

Wald, A. (1950), *Statistical Decision Functions*, John Wiley & Sons, New York City NY.

Wapner, L. M. (2012), *Unexpected Expectations: The Curiosities of a Mathematical Crystal Ball*, CRC Press, Boca Raton FL.

Westfall, R. S. (1971), *Force in Newton's Physics: The Science of Dynamics in the Seventeenth Century*, Neale Watson Academic Publications, New York City NY.

Wilson, E. B. (1927), "Probable Inference, the Law of Succession, and Statistical Inference," *Journal of the American Statistical Association*, 22: 158, 209-212.

Winkler, R. L. (1962), "The Assessment of Prior Distributions in Bayesian Analysis," *Journal of the American Statistical Association*, 62, 776-800.

Wittgenstein, L.J.J. (1974), *Tractatus Logico-Philosophicus*, Revised Ed., Routledge & Kegan Paul, New York City NY.

World Bank (2010), *Natural Hazards, UnNatural Disasters: The Economics of Effective Prevention*, The World Bank, Washington DC.

Wuppuluri, R. I. and Doria, F. A. (2018), *The Map and the Territory; Exploring the Foundations of Science, Thought and Reality*, Springer Int. Publishing, Switzerland.

Vita

Mahdi Habibi was born in Tehran, Iran. After receiving his high school diploma in mathematics and physics, he entered the University of Tehran to study civil engineering. He followed his graduate studies in geotechnical engineering, a branch of civil engineering focused on solving ground-related engineering problems, and received his Master of Science from the University of Tehran; the same institution where he earned his B.Sc. Upon completion of his masters, he worked as a consulting geotechnical engineer, working on a variety of marine port infrastructure projects. On the side, he also lectured undergraduate students on courses ranging from soil mechanics to pavement engineering. Eventually, Mahdi left consulting practice to begin his Ph.D. studies at the University of Texas at Austin, where he worked under the supervision of professors Gilbert and Lake on developing a theory for objective assignment of Bayesian probabilities for the purpose of making risky economic decisions.

Email address: mhabibi@utexas.edu

This dissertation was typed by Mahdi Habibi.