University of Strathclyde, Glasgow
Department of Civil and Environmental Engineering

SUPPORTING RATIONAL DECISION-MAKING IN CIVIL ENGINEERING

Andrea Verzobio

Supervisors:
Enrico Tubaldi
John Quigley
Daniele Zonta

A thesis presented in fulfilment of the requirements for the degree of
Doctor of Philosophy

August, 2020

Declaration of authenticity and author's rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: Andrea Verzobio

Date: 08/08/2020

Previously published work

- Research outputs present in the thesis (published or submitted to journal papers)

1. Chapter 3:
   Consequences of representativeness bias on SHM-based decision-making.
   Verzobio, Bolognani, Quigley, Zonta.
   Structure & Infrastructure Engineering Journal.
   Submitted May 2020, revised August 2020 (currently under review, manuscript ID NSIE-2020-0269).
   Contribution to this work: I developed, with the help of Bolognani, the literature review, the mathematical framework and the application to the case study. All the authors provided feedback and reviewed the manuscript before the submission to the journal.
2. Chapter 4:
   An elicitation process to quantify Bayesian Networks for dam failure analysis.
   Verzobio, El-Awady, Ponnambalam, Quigley, Zonta.
   Canadian Journal of Civil Engineering.
   Submitted February 2020, revised June 2020 (currently under review, manuscript ID cjce-2020-0089.R1).
   Contribution to this work: I developed the literature review and subsequently the methodology about the elicitation process. I carried out, together with El-Awady, the application to the case study: in particular, El-Awady developed the Bayesian Network while I focused on all the aspects of the elicitation process. All the authors provided feedback and reviewed the manuscript before the submission to the journal.
3. Chapter 5:
   Quantifying the benefit of Structural Health Monitoring: what if the manager is not the owner?

Bolognani, Verzobio, Tonelli, Cappello, Glisic, Zonta, Quigley (2018). Structural Health Monitoring, 17(6): 1393-1409.

Submitted January 2018, revised June 2018, published August 2018.

Contribution to this work: Bolognani and I developed the mathematical formulation about the Value of Information and the application to the case study. Tonelli and Cappello gave some feedback as concern the mathematical formulation. Glisic allowed us to work on this real-life case study. All the authors provided feedback and reviewed the manuscript before the submission to the journal.

4. Chapter 6: Quantifying the benefit of Structural Health Monitoring: can the Value of Information be negative?

   Verzobio, Bolognani, Quigley, Zonta.

   Structure & Infrastructure Engineering Journal.

   Submitted May 2020 (currently under review, manuscript ID NSIE-2020-0304).

   Contribution to this work: I developed the literature review and the mathematical framework about the negative Value of Information. I carried out, together with Bolognani, the application to the case study. All the authors provided feedback and reviewed the manuscript before the submission to the journal.

- Other research outputs (conference papers)

1. Quantifying the benefit of Structural Health Monitoring: can the Value of Information be negative?

   Verzobio, Bolognani, Quigley, Zonta.

   Proc. 12th International Workshop on Structural Health Monitoring. Stanford University, California, 2019.

2. Elicitation process to generate Bayesian Networks with application to dams safety.

   Verzobio, El-Awady, Ponnambalam, Quigley, Zonta.

Proc. 8th International Conference on Water Resources and Environmental Research. Nanjing, China, 2019.

3. The consequences of heuristic distortions on SHM-based decision problems.
   Verzobio, Bolognani, Quigley, Zonta.
   Proc. 9th European Workshop on Structural Health Monitoring, Manchester, United Kingdom, 2018.
4. Monitoring-based Decision Support System for optimal management of Colle Isarco Viaduct.
   Verzobio, Tonelli, Bolognani, Cappello, Bursi, Zonta.
   Proc. SPIE. Denver, Colorado, 2018.
5. Bayesian Network approach for failure prediction of Mountain Chute dam and generating station.
   El-Awady, Ponnambalam, Bennett, Zielinsky, Verzobio.
   Proc. International Commission On Large Dams. Ottawa, Canada, 2019.
6. Quantifying the benefit of SHM: what if the manager is not the owner?
   Bolognani, Verzobio, Tonelli, Cappello, Glisic, Zonta, Quigley.
   Proc. 11th International Workshop on Structural Health Monitoring. Stanford University, California, 2017.
7. An application of Prospect Theory to a SHM-based decision problem.
   Bolognani, Verzobio, Tonelli, Cappello, Zonta.
   Proc. SPIE. Portland, Oregon, 2017.
8. Non Gaussian multi-parameter Bayesian estimation through the mechanical equivalent of logical inference.
   Bolognani, Verzobio, Cappello, Tonelli, Zonta.
   Proc. IABSE TU1406. Zagreb, Croatia, 2017.
9. Solving Bayesian multi-parameter estimation problems using the mechanical equivalent of logical inference.
   Bolognani, Verzobio, Tonelli, Cappello, Zonta.
   Proc. 12th International Conference on Structural and Safety & Reliability. Vienna, Austria, 2017.

Abstract

The management of civil engineering structures, such as bridges and dams, is fundamental for ensuring their continued safe and economical operation, where decisions such as whether or not to suspend operations are based on uncertain knowledge concerning the state of the structure. Modern development in technology has made available several accurate monitoring devices providing Structural Health Information, that can be used to support such decision-making through more informed assessment of the structural state.

The wide-spread adoption of these devices has led to the development of Structural Health Monitoring (SHM) decision support system, to identify appropriate courses of action based on observed data from the monitoring system. Essentially, it is a two-step process, which includes the judgment of the structural state based on the SHM information, and the decision about the optimal action based on the knowledge of the structural state. When engineering knowledge concerning the state of the system is uncertain, as the monitoring system does not directly observe the state of the structure, Bayesian inference and Expected Utility Theory provide the only consistent way to judge and to make decisions, respectively, as all alternative inferential methods for decision support are susceptible to logical inconsistency. However, we must recognize that in the real world the process followed by decision makers may be distorted. The goal of the research proposed in this contribution is twofold: we investigate how heuristic behaviours may affect human judgment and decision-making in civil engineering, and also how decision-making can be distorted when multiple agents, even rational but with different appetites for risk, are involved in the decision chain.

Firstly, most agents in everyday life apply heuristic approaches rather than a formal Bayesian procedure in order to make inference to support decisions. In particular, without the use of formal algorithms to support rational interpretation of data, humans apply simple strategies or mental processes to interpret data, which are prone to systemic errors. This may happen with data that come from various data sources, such

as SHM but also engineering expert knowledge. Innovative frameworks to support rational decision-making are then required, in order to minimize the risk of biased judgments or decisions. For instance, being able to predict the behavior of an irrational manager is necessary when we set a general policy for bridge management, and we know that someone else who is going to enact the policy may behave irrationally. In this doctoral thesis, we start reviewing the literature of heuristics and cognitive biases in order to identify the most relevant as regards human judgment and decision-making for civil engineering structures. We identify Kahneman and Tversky's *representativeness* as a heuristic for which SHM-based decision-making is particularly susceptible, where simplified rules for updating probabilities can distort the decision maker's perception of risk. Therefore, we reproduce mathematically this observed irrational behavior to investigate how it distorts human judgment. In addition, it is recognized that heuristic behaviors may affect expert knowledge. Consequently, we propose a method for eliciting engineering expert knowledge in order to assess civil engineering structures: the process is required in order to support the collection of valid and reliable data, by minimizing the adverse impact of cognitive biases.

Secondly, the decision process can be distorted when multiple agents are involved, not only in the case of irrational behaviors, where the distortion is expected, but even in the case of rational behaviors. Indeed, decision makers may differ in their decisions under uncertainty according to their different appetites for risk. Again, predicting the behavior of managers is required for instance when there is a management policy for which the final decision of an agent has to consider the opinion of other decision makers, who may behave differently. In this thesis, we formalize an innovative rational method for quantifying the value of information (*VoI*) of SHM when two different agents are involved in the decision chain: this framework allows one to investigate how decisions may be distorted due to the different appetites for risk of decision makers. In addition, we understand that the interaction between rational agents with different appetites for risk may lead to a negative *VoI*, which is unexpected since it means that the monitoring information may be perceived as damaging. Therefore, we develop a mathematical formulation to investigate under which specific circumstances it is possible to achieve this unexpected outcome.

Finally, all the studied theories and proposed frameworks are applied respectively to various civil engineering case studies. In summary: we evaluate the structural safety of a common type of bridge of the Autonomous Province of Trento stock, in Italy; we investigate the system reliability of the Mountain Chute dam and generating station in Ontario, Canada; we analyse the management of a pedestrian bridge in Princeton University campus equipped with a monitoring system, in USA. These applications allow us to demonstrate the operationalizability of the methods developed in this thesis, and to prove their relevance in various civil engineering case studies.

Acknowledgements

I would like to acknowledge the professors and researchers that have contributed during the three years of my PhD. First of all, I want to thank my main supervisor, Prof. Daniele Zonta, for his support, mentorship and encouragement throughout all my research. Then, I wish to thank my second supervisor, Prof. John Quigley, who has been fundamental for the development of my PhD especially for his contribution as regards the behavioural economics aspects of my doctoral activity. Many thanks also to Dr. Enrico Tubaldi, who has been my official supervisor during the last months of my PhD, and to Prof. Kumaraswamy Ponnambalam, who has been my supervisor during the period that I spent at the University of Waterloo (Ontario, Canada).

Among the many researchers with whom I have collaborated, I would like to thank Denise Bolognani and Carlo Cappello, from University of Trento, and Ahmed El-Awady, from University of Waterloo: your collaboration and your support have been fundamental for the development of my research studies.

Contents

Declaration of authenticity and author's rights 3
Previously published work 4
Abstract 7
Acknowledgements 10
1. Introduction 13
1.1 Decision theory based on Structural Health Information 14
1.2 Motivation 16
1.3 Aims and objectives 18
1.4 Overview 19
2. Literature review 22
2.1 SHM-based decision process rational framework 22
2.2 Irrational behaviours based on heuristics and biases 30
2.3 Value of Information of SHM 44
2.4 Elicitation process 49
3. Consequences of representativeness bias on SHM-based decision-making 54
Summary of the paper 54
3.1 Introduction 54
3.2 SHM-based decision-making rational framework 57
3.3 Heuristics and biases 60
3.4 The representativeness heuristic 61
3.5 A classical representativeness problem 66
3.6 Case study 71
3.7 Conclusions 81
4. An elicitation process to quantify Bayesian networks for dam failure analysis 84
Summary of the paper 84

4.1 Introduction 85
4.2 Bayesian network 87
4.3 Elicitation Process for Bayesian networks 89
4.4 The Mountain Chute Dam and GS case study 94
4.5 Conclusions 106
5. Quantifying the benefit of Structural Health Monitoring: what if the manager is not the owner? 108
Summary of the paper 108
5.1 Introduction 109
5.2 SHM-based decision 111
5.3 Two individuals, two decisions 111
5.4 The Streicker Bridge case study 116
5.5 Concluding remarks 132
6. Quantifying the benefit of Structural Health Monitoring: can the Value of Information be negative? 135
Summary of the paper 135
6.1 Introduction 135
6.2 Value of information for SHM-based decision 138
6.3 When does the *VoI* become negative? 141
6.4 The Streicker Bridge case study 154
6.5 Conclusions 167
7. Conclusions 170
References 175

1. Introduction

This thesis is concerned with the effect of management of civil engineering structures. Civil structural engineering includes all structural engineering related to the built environment, including bridges, dams, tunnels, buildings, and towers. This type of structures is often exposed to extreme forces, for instance static and dynamic loads, temperature variations and high pressures. When we are managing these structures, we are dealing with their stability, strength, and rigidity. Moreover, a large number of these critical structures are getting old and therefore they increasingly require maintenance, inspection and replacement. It is then fundamental to guarantee their continued safe and economical operation, as well as to prevent potential catastrophic structural failures. Unfortunately, catastrophic events have happened recently: an example is the collapse of the Morandi Bridge (Calvi, et al., 2019), in Italy. The bridge, also known as Polcevera viaduct, designed in the early 1960s and opened to traffic in 1967, collapsed in August 2018 killing 43 people and causing an invaluable economic damage, i.e. around €422 million as said by the Genoa Chamber of Commerce. Another example is the Brumadinho dam disaster (Cambridge & Darren, 2019), in Brazil: on January 2019, the tailings dam suffered a catastrophic failure releasing 12 million cubic meters of tailings slurry, which caused the death of 248 people along with the devastation of the surrounding environment.

In order to avoid similar catastrophic events and to assure the continued safe and economical operation of the structure, it is evident that the management of these structures is crucial and in particular decision-making plays an important role. Generally, decision-making is the process of making choices by identifying a decision, gathering information, and assessing alternative resolutions. In the case of engineering structures, decision-making is about how to make the optimal decision based for instance on the knowledge of the structural state. The managers of these structures deal with decision-making problems every day: for instance, a bridge manager has to decide whether or not to close the bridge to traffic after an accident that can question the structural safety of the bridge; similarly, the manager of a dam has to make decisions to optimize the long-term productivity of the asset, subject to safety conditions.

In order to make the best decision, it is fundamental for the decision maker to have an optimal understanding about the condition of the structure, i.e. the actual structural state as well as what it is expected to be in the future. Fortunately, in the 21st century we are getting access to much better data than we had before: with the development of monitoring equipment, the structural state can be evaluated based on accurate data referred as Structural Health Information (SHI). In detail, SHI may be provided by various sources such as Structural Health Monitoring (SHM) and other digital technologies and networks. In particular, SHM is commonly seen as a powerful tool as regards the management of critical civil structures, especially in supporting decisions concerning maintenance, reconstruction and repairs of their assets through reducing uncertainty about the state of the structure. The main purpose of SHM is to provide accurate and real-time information about the state of the structure, which can be used subsequently as inputs for decision-making regarding its management. SHM can be useful both for obtaining information about a structure just after an extreme event, e.g. earthquake, and for monitoring the long-term structural behaviour of the structure. SHM technology offers significant economic and life-safety benefits: for example, the accurate information about the structural state obtained with the monitoring system allow to use the available resources effectively, resulting in clear economic benefits. SHM for civil structures has been investigated in the literature since the early 2000s (Farrar & Worden, 2007) (Chen & Ni, 2018), and it has been applied in various structures such as bridges, dams, tunnels, towers, buildings and offshore installations (Brownjohn, 2007). For a more in-depth reading about SHM applied to civil engineering structures please see (Chen, 2018).

1.1 Decision theory based on Structural Health Information

In order to study these engineering decision problems a generic abstract representation is needed: the management of these critical engineering structures can be formalized with a two-step process, which includes a judgement and a decision, as presented in Figure 1.1. This process comprises the acquisition and subsequent interpretation of data to support informed decision-making. For instance, in the case of a monitoring system, we can call it *SHM-based decision process*, i.e. deciding based on the information from a SHM system. It is fundamental to highlight that the two steps, although linked, are clearly two separate processes: decision-making is a complex process that must not be confused with the judgment of the structural state, which instead is precisely the starting point of decision-making itself. In addition, decision-making based on SHI is challenging as it requires the decision maker to choose appropriate actions that maximizes benefits while minimizing cost: we can refer to it as a trade-off between risk and benefits to prioritize activities.

The first step of the process, i.e. the judgement, allows to judge the structural state of a structure based on the observations, through an interpretation model. The interpretation model, which allows to judge the state of the structure, can be a numerical or analytic function, with a mechanical or heuristic background. Generally, uncertainties are present in the relationship between states and observations: aleatory uncertainties are caused by an intrinsic randomness of the observed phenomenon, e.g. sensor noise, while epistemic uncertainties are due to lack of knowledge, e.g. error of the structural model. As regards the state, it refers to the condition of the structure involved in the analysis, for example in the case of a bridge it can be *safe* or *failure*, or other classes describing the severity of its damage. Within the scope of this dissertation, we define observation to be any information acquired on site which is suitable to infer the state of the structure. For instance, it may be data collected by sensors temporarily or permanently installed on the structure, in the case of a monitoring system, but also data apprised through visual inspections or site tests.

The second step of the process, i.e. the decision, starts after the assessment of the state of the structure, and is about choosing the optimal action based on the knowledge of the state, through a decision model. For example, in the case of a bridge the decision maker has to choose between alternative actions such as *close the bridge*, *limit the traffic*, *do nothing*. Taking an action produces measurable consequences, and the consequences of an action can be mathematically described by several parameters, encoded in an outcome vector: it measures the direct and indirect consequences of the possible combination between an action and the structural state. For instance, it allows to consider the economical, e.g. direct and indirect costs, and the safety aspects needed to develop the trade-off between risk and benefits. Finally, the choice of the decision model, which allows to identify the optimal action based on its outcome, depends on the specific behaviour of the decision maker.

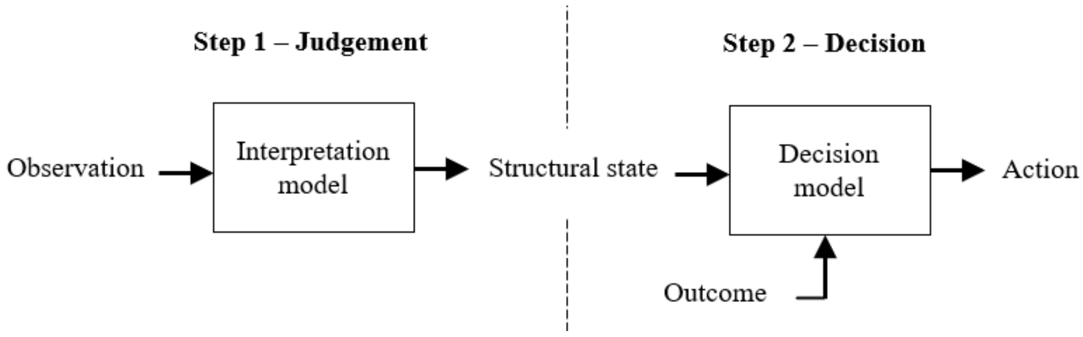


Figure 1.1. The generic framework of the two-step process.

As we will describe in detail in the literature review, the logic of making decision based on SHI is formally stated under the assumption that the decision maker is an ideal rational agent. Consequently, the logical inference process followed by a rational agent as regards the judgment step is mathematically developed in the Bayes' rule (Bolstad, 2010): the posterior knowledge of the structural state is evaluated based on the observation and on the prior knowledge (Sivia & Skilling, 2006). Moreover, as concerns the decision step, Expected Utility Theory (EUT) (Neumann & Morgenstern, 1944) (Raiffa & Schlaifer, 1961) describes the analysis of decision-making under risk and is considered as a normative model of rational choice (Parmigiani & Inoue, 2009).

1.2 Motivation

The analysis of the current state of art, presented in the literature review of chapter 2, highlights the necessity to analyse deeper these decision-making processes since there are engineering cases in real-life where this framework is somehow distorted: the main goal of this thesis is to support rational decision makers through understanding the consequences of distorted human judgment and decision-making.

To start, Bayes' theorem and EUT provide the only consistent way to judge and to make decision under uncertainties, while any alternative judgment or decision model may produce a logical inconsistency (Kahneman & Tversky, 1979) (Pope, 1986) (Lindley, 2006) (Parmigiani & Inoue, 2009). However, we must recognize that in the real world the process followed by decision makers may be different. For instance, we observe that most people in everyday life favour heuristic approaches to this rational framework in order to judge or make decision (Gilovich, et al., 2002). In particular, real-life decision makers often depart from this ideal model of rationality, but judge and decide using common sense and privileging fast and frugal heuristics to rational analytic thinking. Therefore, in order to predict the choices of a real word manager, we have to accept that their behaviour may not be necessarily fully rational. Biased judgment and decision-making have been widely reported and systematically investigated starting 1970s in the fields of cognitive sciences, social sciences and behavioural economics (Gilovich, et al., 2002), but rarely implemented in engineering applications. An effective corrective strategy is then needed for real-world engineering decisions because it is here that a biased judgment or decision has potentially high costs. Consequently in this thesis, to support the management of civil engineering structures, we will focus on discrepancies in behaviour between idealized rational actors and real people. As we will describe in the next chapters, cognitive biases may distort decision-making processes based not only on monitoring information, but also on different data sources such as engineering expert knowledge (Dias, et al., 2018).

Moreover, even in the case of rational decision makers we must recognize that the decision process may be distorted. In particular, we notice that real-life decision problems are typically complex with more individuals involved in the decision chain. In this case with various rational decision makers, they may differ in their choices under uncertainty, even when they have the same information, for instance because of their different appetites for risk. This means that innovative frameworks are needed in order to study how the interaction between different rational agents can influence decisions-making and its context. As regards the main data source analysed in this contribution, i.e. the monitoring system, a clear example can be found in the evaluation of the benefit of SHM, that is formally quantified by the so-called Value of Information (*VoI*) (Straub, et al., 2017). By analysing the *VoI*, we also observe that in the case of multiple decision makers, data may have a negative value (which is caused by the interaction between agents with different appetites for risk), meaning that it is not true that having more data automatically lead to a better decision process. This concept is very innovative since it is in contrast with the acknowledged principle (in the case of a single decision maker) that "*information can't hurt*" (Cover & Thomas, 2006).

Consequently, to understand how the interaction between different decision makers can distort decision-making, in the next chapters we will focus on the quantification of the *VoI* of SHM.

1.3 Aims and objectives

The main goal of this thesis is then to support the rational management of critical structures in civil engineering, especially as regards *SHM-based decision problems*, through investigating the consequences of distorted human judgment and decision-making. Specifically, we have four aims:

i) To investigate how heuristic behaviours affect human judgment and decision-making.
ii) To develop a process to elicit engineering expert knowledge, by minimizing the risk of biased judgments.
iii) To investigate how decision-making can be distorted when multiple decision makers are involved in the decision chain.
iv) To demonstrate that the value of information may be negative in the case of multiple rational decision makers with different appetites for risk.

These aims will be achieved through the following objectives for this research:

- To identify at least one heuristic behaviour which is relevant as regards judgment and decision-making for civil engineering structures.
- To mathematically reproduce the impact of this biased behaviour.
- To apply the heuristic mathematical framework to a real-life SHM-based decision problem, in order to prove how the cognitive bias affects decision-making by distorting the representation of information provided by SHM.
- To develop a process for eliciting engineering expert knowledge that addresses key biases, in order to assess civil engineering structures.
- To apply the developed elicitation process to a real-life case study where it is necessary to rely on expert judgment due to a luck of data, in order to assess its validity and usefulness as an accurate and reliable data source.

- To assess the impact of the interaction between multiple decision makers, by developing an appropriate mathematical formulation that takes into account the possibility that they may act differently due to different appetites for risk.
- To validate the developed framework by applying it to a SHM-based decision problem.
- To develop a mathematical formulation that allows to demonstrate under which conditions the value of information of SHM may become negative in the case of multiple rational decision makers.

1.4 Overview

The thesis comprises 7 chapters, including this introduction. In the second chapter, a critical review of both the current literature as well as key theories that underpin our research is presented. To start, the current state of art about SHM-based decision problems is reviewed, focusing especially on the rational methods that are commonly used to solve the two steps of the process, i.e. respectively Bayesian inference and Expected Utility Theory. In addition, the criticisms about these rational methods, which have motivated the research presented in this thesis, are described: they are principally based on the presence of heuristics and irrational behaviours in human judgments and decision-making. Therefore, a review of their current state of art is presented. Finally, we introduce the state of art about two applications that we have developed in our research: the quantification of the benefit of SHM and the elicitation process.

In chapter 3, we investigate the consequences of heuristic distortions on SHM-based decision problems. Based on the developed literature review as regards heuristic behaviours in human judgment and decision-making, we identify one of these behaviours that is frequently observed in bridge management: the confusion between condition state and bridge safety. This biased judgment can be described by Kahneman and Tversky's *representativeness* heuristic. The aim is then to describe this bias from a mathematical perspective, in order to reproduce it and to understand how it distorts the final judgment of the manager, in contrast with the one achieved rationally following Bayesian logic. To validate the developed heuristic framework, we apply it to a real-life case study concerning the evaluation of the safety of a bridge based on visual inspections: the outcomes properly demonstrate how the judgment of a biased inspector is clearly distorted.

In chapter 4, we investigate the role of engineering expert knowledge as a useful data source. Even if our research in this case is motivated by a specific real-life case study regarding a dam failure analysis, we aim to develop an elicitation process based on a structured methodology that can be apply to many different engineering structures, in the case that the mathematical model used is the Bayesian network. A four-stage structured elicitation process is then proposed, based on the literature review and paying close attention to all the biases that can influence the process, such as the *anchoring*. Finally, the application of the developed methodology to the case study allows us to prove its validity, along with making us learn some lessons that can be useful to improve the procedure for future similar engineering applications.

In chapter 5, we study how the interaction between multiple decision makers can distort decision-making and its context. We decide to analyse it as regards the quantification of the benefit of information, such as that coming from a monitoring system, through an index called Value of Information (*VoI*). After a review of the basis of *VoI*, where it is assumed that all decisions are making by only one decision maker, we formalize a new method for quantifying the *VoI* when two different agents are involved in the decision chain, as we often observe in the real world. To illustrate how this framework works, a hypothetical *VoI* for a pedestrian bridge equipped with a monitoring system is evaluated: the outcomes, evidently different from the case of a single decision maker, show how decisions may be distorted depending on the different appetites for risk of the two agents.

In chapter 6, starting from the method for quantifying the *VoI* of a monitoring system proposed in chapter 5, we aim to demonstrate that in the case of multiple rational decision makers the *VoI* may be negative due to their different appetites for risk. In particular, we develop a mathematical formulation that allows to understand when, under specific assumptions, the *VoI* becomes negative. To validate this framework we apply it to the same case study as in chapter 5, the Streicker bridge at Princeton campus. The achieved results prove our statement, which is very innovative because a negative *VoI* means that the monitoring information are perceived as damaging, in contrast with the acknowledged principle that "*information can't hurt*".

Finally, the last chapter presents concluding remarks, based on a discussion about the outcomes of the research in order to demonstrate the achievement of the predetermined aims. In addition, the limitations of the developed research along with related future works are provided.

2. Literature review

In this chapter, we present a critical review of both the current literature as well as key theories that underpin our research. In the first section, we present the general framework which provides the basis of our study, i.e. the SHM-based decision process, and we describe how rational methods such as Bayesian inference and Expected Utility Theory (EUT) are used to support such decision problems. Subsequently, the state of art about heuristics and cognitive biases is presented: we review their definitions and then we introduce the main heuristic behaviours that have been investigated in the literature, focusing principally on the ones that have been applied to our research. Finally, the state of art about two key theories that underpin our research is introduced: the first concerns the quantification of the benefit of SHM, based on the so-called Value of Information (*VoI*); the second is about elicitation processes that are required to elicit meaningful expert engineering knowledge, in order to support rational decision-making.

2.1 SHM-based decision process rational framework

In this section, we present the concepts of SHM-based decision process, i.e. deciding based on the information from a SHM system. The main purpose of SHM is to provide accurate and real-time information about the state of the structure, which can be used as objective inputs for decision-making regarding its management. SHM-based decision problems have been recently investigated in the literature: Flynn and Todd (Flynn & Todd, 2010a) (Flynn & Todd, 2010b) proposed innovative Bayesian approaches to optimal sensor placement for SHM applications, focusing on the example of active sensing and implementing an appropriate statistical model of the wave propagation and feature extraction process; Zonta et al. (Zonta, et al., 2014) evaluated a rational framework for assessing the impact of SHM on decision-making, with application to a pedestrian bridge on Princeton University campus (USA) equipped with a fiber optic sensing system; Cappello et al. (Cappello, et al., 2016) proposed a decision framework based on EUT for civil engineering decision problems, in the case that agents have to act based on the information of a SHM system; Tonelli et al. (Tonelli, et al., 2017) proposed a decision support system that interprets the data coming from a monitoring system and consequently suggests the optimal decision to undertake: this framework is based on rational methods and is applied to the Colle Isarco Viaduct in Italy.

As observed in (Cappello, et al., 2016), SHM-based decision-making is properly a two-step process which, following the framework introduced in section 1.1, includes a judgement and a decision: first, based on the information $\mathbf{y}$ from the sensors, we infer the state S of the structure; next, based on our knowledge of the state S of the structure we choose the optimal action a_{opt} to take. Figure 2.1 shows the process. In the following, each component of the framework is properly described:

- Structural state: it represents the condition of the structure involved in the analysis. For instance the structure, e.g. a bridge, can be in one out of N mutually exclusive and exhaustive states S_1, S_2, …, S_i, …, S_{N} (e.g.: S_1 = 'severely damaged', S_2 = 'moderately damaged', S_3 = 'not damaged', …). The state of the structure is generally not deterministically known, so it can be only described in probabilistic terms.
- Observation: within the scope of this dissertation, we define observation $\mathbf{y}$ to be any information acquired on site which is suitable to infer the state of the structure. It may be measurements acquired by sensors installed on the monitored structure, documents and reports containing results of tests or inspections performed on the monitored structure.
- Interpretation model: it allows one to judge the state S of the structure, based on the observation $\mathbf{y}$, by considering all the aleatory (e.g. sensor noise) and epistemic (e.g. error of the structural model) uncertainties that are present in the relationship between state and observation. Generically, the model may be a numerical or analytic function, with a mechanical or heuristic behaviour.
- Action: it is an option that the decision maker has to make at the decision step. The set of actions can be discrete or continuous. In the case of discrete variables, the agent can choose between a set of M alternative actions a_1, a_2, …, a_j, …, a_{M} (e.g.: a_1 = 'do nothing', a_2 = 'limit traffic', a_3 = 'close the bridge

to traffic', …). In the case of a continuous set, the action may be for instance the choice of the frequency with which the structure needs maintenance.

- Outcome: taking an action produces measurable consequences (e.g.: a monetary gain or loss, a temporary downtime of the structure, in some case causalities); the consequences of an action can be mathematically described by several parameters (e.g.: the amount of money lost, the number of day of downtime, the number of casualties), encoded in an outcome vector **z**. The outcome **z** of an action depends on the state of the structure, thus it is a function of both action *a* and state *S*: $\mathbf{z}(a, S)$; when the state is certain the consequence of an action is also deterministically known; therefore, the only uncertainty in the decision process is the state of the structure *S*.
- Decision model: it is a framework that allows one to identify the optimal action a_{opt} to undertake based on the outcome **z** and on the knowledge of structural state *S*. The choice of the decision model is very important and depends on the specific behaviour of the decision maker.

The logic of decision-making under uncertainties based on SHM is formally stated under the assumption that the decision maker is a rational agent, who judges using Bayes' theorem (Bolstad, 2010) and decides consistently with Neumann-Morgenstern's Expected Utility Theory (Neumann & Morgenstern, 1944), as presented in Figure 2.1. In the next subsection, the state of art about the two rational steps of the process is presented.

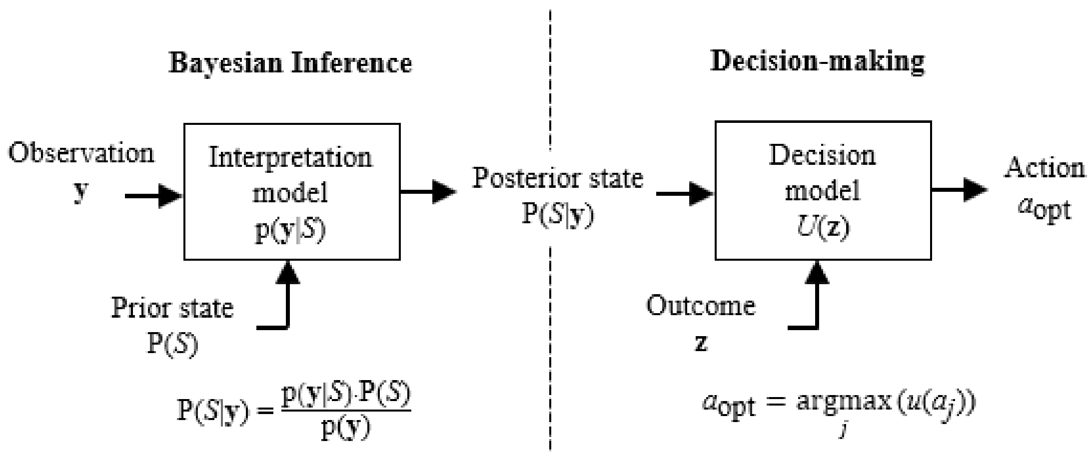


Figure 2.1. The rational process of SHM-based decision-making.

2.1.1 Rational judgment: Bayesian inference

Judgment is about understanding the state of the structure based on the observation, which is exactly what SHM is about from a logical standpoint. The logical inference process followed by a rational agent is mathematically encoded in Bayes' rule (Parmigiani & Inoue, 2009): the structural state is assessed using a structural model along with the prior knowledge.

Many modern textbooks offer a critical review and applications of this theory to data analysis, see for instance Gregory (Gregory, 2005), Sivia and Skilling (Sivia & Skilling, 2006), Murphy (Murphy, 2012). In addition, Bishop (Bishop, 2006) explained the principles of Bayesian trend fitting, MacKay (MacKay, 2003) presented Bayesian model updating, while Beck and Yuen (Beck & Yuen, 2004) applied Bayesian model updating to structural identification, and Yuen and Kuok (Yuen & Kuok, 2011) proposed some applications to dynamic models. Nowadays, Bayesian logic is consistently implemented in SHM techniques: among the many, Beck and Au (Beck & Au, 2002) proposed a Markov Chain Monte Carlo (MCMC) to perform Bayesian inference based on SHM data; Mthembu et al. (Mthembu, et al., 2011) proposed the use of Bayesian logic in the context of model selection; Memarzadeh et al. (Memarzadeh, et al., 2014) used Bayes' rule to propose a novel learning and planning method; Zonta et al. (Zonta, et al., 2014) suggested a SHM framework, based on Bayesian inference, for quantifying the benefit of SHM in bridge management; Cappello et al. (Cappello, et al., 2015) used Bayes' theorem to judge the structure behaviour of a bridge based on data coming from a SHM system. Other examples can be found in Sohn and Law (Sohn & Law, 1997), Enright and Frangopol (Enright & Frangopol, 1999) and Vanik et al. (Vanik, et al., 2000).

In summary, Bayesian inference is based on the observation, the prior knowledge of the structural state and the structural model that approximates the behaviour of the structure: in this way it is possible to achieve an estimation of the structural state. In addition, the consequent Bayes' theorem depends on the type of state variables: they can be discrete, e.g. $\mathbf{S} = (S_1, S_2, \ldots, S_i, \ldots, S_\mathrm{N})$, or continuous, i.e. composed by an infinite number of discrete variables. In the first case, Bayes' theorem in the presence of uncertainty states that the state of the structure S_i after observing the sensors data $\mathbf{y}$

= (y_1, y_2, …, y_K) is probabilistically described by the posterior information P(S_i|**y**) (Sivia & Skilling, 2006) (Bolstad, 2007). In formula:

$$\mathrm{P}(S_i|\mathbf{y}) = \frac{\mathrm{P}(\mathbf{y}|S_i)\ \mathrm{P}(S_i)}{\mathrm{P}(\mathbf{y})}. \tag{2.1}$$

Eq. (2.1) says that the posterior knowledge of the ith structural state P(S_i|**y**) depends on the prior knowledge P(S_i), i.e. what we expect the state of the structure to be before reading any monitoring data (Cappello, et al., 2015), and the likelihood P(**y**|S_i), i.e. the probability of observing the data given the state of the structure. P(**y**) is instead simply a normalization constant, referred to as evidence, calculated as:

$$\mathrm{P}(\mathbf{y}) = \sum_{i=1}^{N} \mathrm{P}(\mathbf{y}|S_i)\ \mathrm{P}(S_i). \tag{2.2}$$

It is interesting to notice that, if we neglect this normalization constant, it is evident that the posterior probability of the state of the structure *S* after the acquisition of the observation **y** is simply proportional to the product between the probability of observing the observation **y**, given the state *S*, and the probability of the state *S* before acquiring the observation **y**. In formula:

$$\mathrm{P}(S|\mathbf{y}) \propto \mathrm{P}(\mathbf{y}|S) \cdot \mathrm{P}(S). \tag{2.3}$$

On the other hand, in the case of structural states described by continuous variables, the definition of the posterior information is different. In this case, rather than a discrete structural state *S*, a state parameter θ has to be inferred, which can be for example some damage features of the structure, e.g. material properties or structural stiffness. Consequently, Bayes' theorem becomes:

$$\mathrm{p}(\theta|\mathbf{y}) = \frac{\mathrm{p}(\mathbf{y}|\theta) \cdot \mathrm{p}(\theta)}{\mathrm{p}(\mathbf{y})}, \tag{2.4}$$

where p(θ|**y**) is the posterior probability function, which tell us how the state parameter θ is distributed, p(**y**|θ) is the likelihood probability function, p(θ) is the prior probability function and p(**y**) is the evidence. If the observation **y** = (y_1, y_2, …, y_K) are uncorrelated, the likelihood can be calculated as follows:

$$p(\mathbf{y}|\theta) = \prod_{i=1}^{K} p(y_i|\theta). \tag{2.5}$$

In addition, the evidence becomes:

$$p(\mathbf{y}) = \int_{D\theta} p(\mathbf{y}|\theta) \cdot p(\theta)\, d\theta. \tag{2.6}$$

Finally, the likelihood can be seen as a representation of the interpretation model, chosen by the individual who judges. It is important to highlight that Bayesian logic does not suggest itself which interpretation model should be used, it just provides a logical way to calculate the posterior once we have assumed the prior and the interpretation model. In any case, the choice of the prior and of the interpretation model has an obvious impact on the calculation of the posterior: elicitation processes, introduced in section 2.4, can be useful to inform the choice of this interpretation model, as well as to provide prior distributions.

In conclusion, although Bayesian inference is the only consistent way to judge under uncertainties and any alternative inference model may produce a logical inconsistency (Kahneman & Tversky, 1979) (Pope, 1986) (Lindley, 2006) (Parmigiani & Inoue, 2009), nowadays the research community recognizes that most people favour heuristic approaches to this rational framework (Gilovich, et al., 2002): heuristics can be seen as simplified algorithms that approximate the solutions in comparison to Bayesian inference. Consequently, in section 2.2 the state of art about heuristics and biases will be provided, in order to understand their main features and how they may affect human judgment.

2.1.2 Rational decision-making: Expected Utility Theory

Decision-making is about choosing the best action based on the knowledge of the state, i.e. once the posterior probability of the structural state has been assessed. The basis of rational decision-making under uncertainty is encoded in the so-called Expected Utility Theory (EUT), which was first introduced by Von Neumann an Morgenstern in 1944 (Neumann & Morgenstern, 1944), and later developed in the form that we currently know by Raiffa and Schlaifer (Raiffa & Schlaifer, 1961) in

1961. EUT is largely covered by a number of modern textbooks, among the many we recommend Parmigiani and Inoue (Parmigiani & Inoue, 2009). As regards the application of EUT to SHM, the Reader can find an extensive recent reference in the doctoral thesis of Cappello (Cappello, 2017).

In order to choose the optimal action by implementing EUT, three terms have to be considered: the probability of the structural states evaluated using Bayesian inference P(S_i); the quantification of the outcome **z**, which usually includes direct and indirect costs, i.e. the consequences; the appetite for risk of the decision maker, which is usually described by the so-called utility function U. In detail, when the structural state S is deterministically known, the rational decision maker ranks an action a based on the consequences **z** through a utility function $U(\mathbf{z})$. Mathematically, the utility function is a transformation that converts the vector **z**, which describes the outcome of an action in its entire complexity, into a scalar U, which indicates the agent's order of subjective preference for any possible outcome. When the state of the system is uncertain, and therefore the consequences of an action are only probabilistically known, Expected Utility Theory (EUT) says that decision makers rank their preferences based on the expected utility u, defined as:

$$u(\mathrm{a}) = \mathrm{E}_S\big[U\big(\mathbf{z}(a, S)\big)\big], \tag{2.7}$$

where E_S is the expected value operator of random variable S, which we have assumed be the only uncertainty into the problem. To clarify the notation, note that U indicates the utility function, while u denotes an expected utility. In other words, if each state S has a probability P(S_i) and has an outcome $\mathbf{z}(a_j, S_i)$ when action a_j is taken, the expected utility $u(a_j)$ of a particular action a_j is evaluated as follows:

$$u(a_j) = \sum_{i=1}^{N} U\Big(\mathbf{z}\big(a_j, S_i\big)\Big)\,\mathrm{P}(S_i). \tag{2.8}$$

Consequently, the decision maker, consistent with EUT, will choose that action a_{opt} which carries the maximum expected utility payoff u:

$$u = \max_j u(a_j), \qquad a_{\mathrm{opt}} = \arg\max_j u(a_j). \tag{2.9a,b}$$

As regards the utility function U required to describe the behaviour of decision makers, it can be very different based on their particular individual risk aversion (Bernoulli, 1954) (Kahneman & Twersky, 1984), as presented in Figure 2.2 and explained in detail in chapter 5. For instance, an agent is risk neutral if his or her utility function U is linear with the loss or gain $\mathbf{z}$, i.e. $U(\mathbf{z}) \propto \mathbf{z}$. Since the expected utility is proportional to the probability of realization, as shown in Eq. (2.8), risk neutrality implies indifference to a gamble with an expected value of zero. So, for example, to a risk neutral agent, a 1% probability of losing US$100 is equivalent to a certain loss of US$1. Moreover, a decision maker can also be risk adverse or risk seeking. An agent has a risk adverse behaviour if he or she tends to reject gambles with a neutral expected payoff: this condition can be graphically represented with a concave, i.e. with negative second derivate, utility function. The condition of risk aversion is consistent with the observation that the marginal utility of most goods, including money, diminishes with the amount of goods, or the wealth of decision maker, as observed since Bernoulli (Bernoulli, 1954). Conversely, an agent with a risk seeking behaviour is clearly the opposite: the utility function is convex, i.e. with a positive second derivate.

In conclusion, even if the basis of rational decision-making is encoded in EUT, choices are sometimes made based on heuristics: not rational decision strategies, but decisions leaded by shortcuts or emotions, which are influenced by systematic cognitive biases. Consequently, in section 2.2, the state of the art about heuristics and biases is also discussed to understand how they can influence the decision-making process, leading to different models that take into account irrational behaviours and heuristic biases in contrast with EUT.

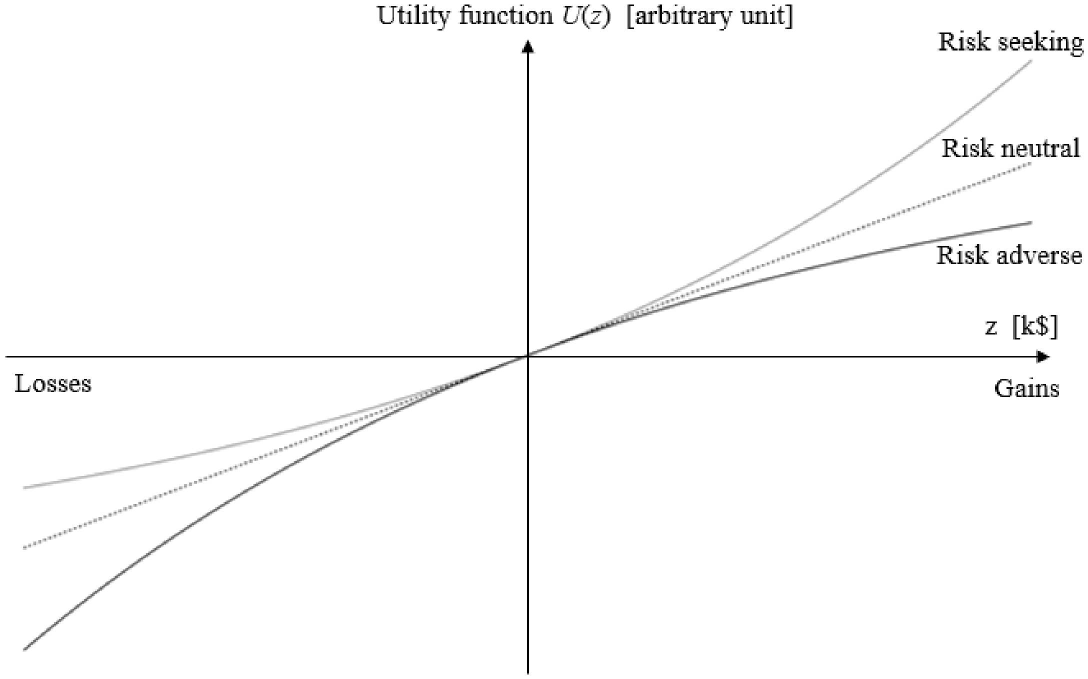


Figure 2.2. Different utility functions according to the risk aversion of the decision maker (Bolognani, et al., 2018).

2.2 Irrational behaviours based on heuristics and biases

Heuristics represent a simplified method to judge or make decision which is based on rules of thumb, logical simplifications or shortcuts rather than a proper rational method. Even if heuristics have some advantages such as saving time, information and energy, the natural consequence of using them is that they can lead to biases. In general, depending on their nature, a heuristic may affect the process outlined in section 2.1 in the inference step, in the decision step, or in both cases.

Biased judgement and decision-making have been widely reported and systematically investigated since the 1970s in the fields of cognitive sciences, social sciences and behavioural economics. The most important contribution to the formal characterization of the heuristic behaviour is the work that Kahneman and Tversky carried out in the early 1970s (Kahneman & Tversky, 1973) (Tversky & Kahneman, 1974) (Kahneman & Tversky, 1979) (Tversky & Kahneman, 1983), which had a significant impact to the understanding and description of the human behaviour and represents the basis of a new discipline we currently refer to as *behavioural economics*. They developed the so-called *heuristics and biases approach*, challenging the dominance of strictly rational models. Textbooks such as (Kahneman, et al., 1982) and (Gilovich, et al., 2002) are extensive references for those approaching the topic for the first time.

The relevance of heuristics has been recently investigated in various engineering applications: for instance, Nikolova evaluated the practical performance of heuristic approaches in stochastic traffic engineering problems (Nikolova, 2010); Fortz introduced a heuristic approach for internet traffic engineering, by considering changing demands and robustness issues with respect to network failures (Fortz, 2011); Martinelli investigated the nature of technological changes using engineering heuristics in the telecommunications switching industry (Martinelli, 2012); Daly et al. studied how engineering students and practitioners generate ideas, using a methodology that identifies design heuristics (Daly, et al., 2012); Elms and Brown investigated the presence of heuristics in engineering complex systems, in order to provide quality control against bias and error (Elms & Brown, 2013); Yakovis and Chechurin studied the relationship between heuristics and standard tool application in the design of process control systems (Yakovis & Chechurin, 2015). Other examples can be found in engineering decision-making (Leonard, 2014), engineering design problems (Studer, et al., 2016), optimization engineering algorithms (Larijani & Ahmadinia, 2018), and mathematical engineering problems (Rodriguez, et al., 2018).

In the following, we start reviewing the definition of heuristics and biases during history to properly understand their meaning, since everyone who made use of the term seemed obliged to give his own interpretation of it. Subsequently, the state of art about Kahneman and Tversky's heuristics and biases approach is provided, since we have decided to focus our research principally on their work, in order to understand which among the heuristics and biases they introduced can mainly affect the process of judgment or decision as regards our engineering applications.

2.2.1 Heuristic definition during history

The term *heuristic* is of Greek origin: *εὑρίσκω* means *to find out*, *to discover*. In this sense, Whewell stated that "*if you will not let me treat the Art of Discovery as a kind of Logic, I must take a new name for it, Heuristic, for example*" (Todhunter, 1876). Similarly, Immanuel Kant affirmed that "*the ideas of reason are heuristic not* *ostensive: they enable us to ask a question, not to give the answer*" (Caird, 1877). The concept of heuristic has been subject to several definitions and everyone who made use of the term seemed obliged to give his own interpretation of it. Thus, in this section, an overview about different meanings and the development of such expression is presented.

To start, according to Holton (Holton, 1988), Albert Einstein included this term in the title of one of his Nobel Prize winning papers in 1905 on quantum physics, indicating that the view he presented was incomplete, due to the limits of our knowledge, but highly useful. Subsequently, as stated by Gigerenzer and Gaissmaier (Gigerenzer & Gaissmaier, 2011), the mathematician George Polya distinguished heuristics from analytical methods, and he was listed by Minsky as the earliest reference to heuristic in the artificial intelligence (AI) literature (Minsky, 1961). Polya's explanation went as follows (Polya, 1945): "*The aim of heuristic is to study the methods and rules of discovery and invention…Heuristic reasoning is reasoning not regarded as final and strict but as provisional and plausible only, whose purpose is to discover the solution of the present problem…We shall attain complete certainty when we shall have obtained the complete solution, but before obtaining certainty we must often be satisfied with a more or less plausible guess. We may need the provisional before we attain the final. We need heuristic reasoning when we construct a strict proof as we need scaffolding when we erect a building*." Besides, Polya wished to revive heuristic in a "*modest and modern form*", explaining that: "*Modern heuristic endeavors to understand the process of solving problems, especially the mental operations typically useful in this process*." Therefore, as pointed out by Romanycia and Pelletier (Romanycia & Pelletier, 1985), Polya's idea was that heuristic is a science of problem-solving behavior that focuses on plausible, provisional, useful, but fallible, mental operations for discovering solutions. Moreover, these authors claimed that Gelernter was also one of the first to point out that heuristics work in effect by eliminating options from an impractically large set of possibilities (Romanycia & Pelletier, 1985). Gelernter's opinion can be summarized as follows: "*A heuristic is, in a very real sense, a filter that is interposed between the solution generator and the solution evaluator*" (Feigenbaum & Feldman, 1963).

Similarly, in Tonge's discussion of his heuristic program, he emphasized efficiency and effort reduction in achieving a satisfactory solution, introducing the meaning of heuristic as any arbitrary "*device*", which provides "*shortcuts*" and employs "*simplifications*" (Tonge, 1960). His official definition was: "*…by heuristics we mean principles or devices that contribute, on the average, to reduction of search in problem-solving activity…Heuristic problem-solving procedures are procedures organized around such effort-saving devices*" (Feigenbaum & Feldman, 1963). Recalling Polya's idea, Minsky underlined also that a heuristic must be applicable to more than just a restricted set of problems and an effort–saving method that worked on only one problem would be more properly called a specific tool rather than a heuristic method (Minsky, 1961). Thus, summarizing several opinions and concepts shown above, Feigenbaum and Feldman gave their definition as follows (Feigenbaum & Feldman, 1963): "*A heuristic (heuristic rule, heuristic method) is a rule of thumb, strategy, trick, simplification, or any other kind of device which drastically limits search for solutions in large problem spaces. Heuristics do not guarantee optimal solutions; in fact, they do not guarantee any solution at all; all that can be said for a useful heuristic is that it offers solutions which are good enough most of the time.*"

Finally, according to Romanycia and Pelletier, during the early AI period the concept of heuristic was transformed, starting with Polya from a vague psychological groping for a solution, into the notion of an exploration guided along paths in a formal problem-solving structure or space (Romanycia & Pelletier, 1985). We can say that such transformation was strictly connected to the advent of computer programming, when it became clear that most problems of any importance are computationally intractable, and the optimal solution is unknown. Differently by AI researchers, as explained by Gigerenzer and Gaissmaier, psychologists became interested in demonstrating human reasoning errors (Gigerenzer & Gaissmaier, 2011).

On the other hand, Gilovich and Griffin claimed that, in the late 1960s and early 1970s, after the early AI era, a series of papers by Amos Tversky and Daniel Kahneman revolutionized the academic research on human judgment. They developed the so-called *heuristics and biases* approach, challenging the dominance of strictly rational models. Their work highlighted the reflexive mental operations used to make complex problems manageable, pointing how the same processes can lead both to accurate and to dangerously flawed judgments. Their central idea of the heuristics and biases program was that (Gilovich, et al., 2002): *"...judgment under uncertainty is often based on a limited number of simplifying heuristics rather than more formal and extensive algorithmic processing. These heuristics typically yield accurate judgments but can give rise to systematic error*." The main innovation lays in the analysis of the descriptive adequacy of ideal models of judgment and in the proposal of a cognitive alternative that explained human error without invoking motivated irrationality. According to the classical model of rational choice, the rational actor chooses the optimal action by assessing the probability of each possible outcome, detecting their utility and combining them together. The theory of rational choice assumes that people judge these two aspects and judge them well. However, evidence displays that people's assessments of likelihood and risk do not conform to the laws of probability. In this field, Simon developed the concept of *bounded rationality,* that is, people reason and choose rationally, but only within the constraints imposed by their limited search and computational capacities (Simon, 1956). Kahneman and Tversky developed instead their own perspective on bounded rationality, asserting that the processes of intuitive judgment are not merely simpler than rational models demanded, but are categorically different in kind. In fact, they suggested that heuristics are simple and efficient because they exploit evolved or learned capacities (Gilovich, et al., 2002).

As regards to the adequacy of probability theory as a descriptive theory, Angner (Angner, 2012) highlighted that the former theory was never designed to capture the precise cognitive processes people use when forming judgments, indeed there appears to be a wide range of circumstances under which people's intuitive probability judgments differ substantially, systematically and predictably from the demands of the theory. In this sense, the heuristics and biases program was a prominent effort to develop a descriptively adequate theory of probabilistic judgment, that is, to capture the manner in which people actually make judgments (Angner, 2012). Another interesting definition was provided by Pearl, who claimed that (Pearl, 1984): "*Heuristics are criteria, methods, or principles for deciding which among several alternative courses of action promises to be the most effective in order to achieve some goal. They represent compromises between two requirements: the need to make such* *criteria simple and, at the same time, the desire to see them discriminate correctly between good and bad choices*."

In the 1990s, Gigerenzer introduced the *fast and frugal* concept, developing and testing quantitative models of heuristics that, "*...when compared to standard benchmark strategies…can be faster, more frugal, and more accurate at the same time*" (Gigerenzer, et al., 1999). He defined also three qualities that a heuristic should embody: *a*) heuristics are simple relative to the evolved or learned capacities of an organism; *b*) heuristics exploit structures of environments, that is, their rationality is not logical but ecological, thus it suggests that a heuristic is not good or bad, rational or irrational per se, but only relative to an environment; *c*) heuristics are distinct from "*as–if*" optimization models (Gigerenzer, 2004). Similarly, Shah and Oppenheimer proposed that all heuristics rely on effort reduction by one or more of the following (Shah & Oppenheimer, 2008): *a*) examining fewer cues, *b*) reducing the effort of retrieving cue values, *c*) simplifying the weighting of cues, *d*) integrating less information, and *e*) examining fewer alternatives. Finally, as Tversky, Kahneman and Gigerenzer did, Katsikopoulos considered heuristics not from a computational short-cuts point of view, but as psychological basis. In particular, by psychological heuristics he meant models for making decisions that (Katsikopoulos, 2011): *a*) rely heavily on core human capacities; *b*) do not necessarily use all available information and process the information they use by simple computations; *c*) are easy to understand, apply, and explain.

2.2.2 Cognitive biases linked to heuristics

The term *bias* originates in mid-16th century from French *biais* and it is probably of Greek origin: *ἐπικάρσιος* means *crosswise, oblique*. Gilovich and Griffin explained that in the work of Tversky and Kahneman each heuristic was associated with a set of *biases*, that is, "*departures from the normative rational theory that serves as markers or signatures of the underlying heuristics*" (Gilovich, et al., 2002). In this sense, Tversky and Kahneman identified positive and negative agendas for the heuristics and biases program, where positive agenda illustrates the processes through which people make a variety of important and difficult real world judgments, whereas negative agenda shows the conditions under which intuitive judgments are likely to depart from the rules of probability (Gilovich, et al., 2002). Several aspects of this approach led to a comparison between the rules of logic and statistics and the heuristics. As pointed out by Gigerenzer and Gaissmaier, the former has been linked to rational reasoning, whereas the latter to error-prone intuitions or even irrationality (Gigerenzer & Gaissmaier, 2011).

However, Simon, who is the father of bounded rationality, emphasized a different point of view, asking: "*How do human beings reason when the conditions for rationality postulated by the model of neoclassical economics are not met*?" (Simon, 1989). According to Gigerenzer and Gaissmaier, the research formalized recently that in a number of large worlds (i.e. "*situations in which some relevant information is unknown or must be estimated from samples, and the future is uncertain, violating the conditions for rational decision theory*" (Gigerenzer & Gaissmaier, 2011)), simple heuristics were more accurate than standard statistical methods that have the same or more information, placing heuristics at the same importance level than standard statistical models of rational cognition (Gigerenzer & Gaissmaier, 2011).Therefore, as summarized by Angner (Angner, 2012), heuristics are rules of thumb that can be used when forming judgments under uncertainty to reduce the time and effort required to solve everyday problems, but they are not assumed to be perfect, thus they can lead to answers that are systematically and predictably wrong, that is, they can lead to bias. The awareness of the conditions under which this may happen might limit the likelihood that they do.

2.2.3 Kahneman and Tversky's heuristics and biases approach

As we have seen in the above introduction about heuristics and biases, various definitions and approaches have been developed during history. We think that the most important contribution is the work of Kahneman and Tversky, and consequently we have decided to focus on their *heuristics and biases approach*, which challenged the dominance of strictly rational models and therefore it is clearly relevant for the aims of our research.

Kahneman and Tversky revolutionized the academic research on human judgment in the early 1970s (Kahneman & Tversky, 1972) (Kahneman & Tversky, 1973) (Tversky & Kahneman, 1974). The main innovation lies in the analysis of the descriptive adequacy of ideal models of judgment and in the proposal of a cognitive alternative that explained human error without invoking motivated irrationality. Evidence displays that people's assessments of likelihood and risk do not conform to the laws of probability. Additionally, this approach has demonstrated a large number of cognitive biases, i.e. systematic errors in human judgment and decision-making. In the following, we present the state of art about the main heuristics presented by them in 1974, in their paper titled "Judgment under Uncertainty: Heuristics and Biases" (Tversky & Kahneman, 1974), as well as the decision theory based on these irrational behaviours that they introduced in 1979, i.e. the Prospect Theory (Kahneman & Tversky, 1979).

2.2.3.1 Representativeness

Representativeness is commonly used to describe the level of how well or how accurately something reflects upon a sample. Citing Kahneman and Tversky (Kahneman & Tversky, 1972), an individual who follows the representativeness heuristic "*evaluates the probability of an uncertain event, or a sample, by the degree to which it is: (i) similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated*". This means that, a hypothesis, or event A, is judged more probable than a hypothesis, or event B, whenever A appears more representative than B: in other words, the ordering of events by their subjective probabilities (i.e. any estimate of the probability of an event, which is given by a subject, or inferred from his behavior, without demanding to satisfy any axioms or consistency requirements) coincides with their ordering by representativeness.

To illustrate such heuristic, Kahneman and Tversky presented an intuitive example (Tversky & Kahneman, 1974), in which people had to assess the probability of Steve's employment from a list of possibilities (e.g. farmer, salesman, airline pilot, librarian or physician), simply considering Steve's description provided by a former neighbour. Such description went as follows (Tversky & Kahneman, 1974): "*Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.*" The representativeness heuristic leads to assess the probability that Steve is a librarian, for example, by the degree to which he is representative of the stereotype of a librarian. Therefore, to be representative an uncertain event should not only be similar to its parent population, but it should also reflect the properties of the uncertain process by which it is generated, i.e. it should reflect the idea of randomness. In summary, events are ranked according to their representativeness; people consistently judge the more representative event to be the more likely, whether it is or not. In addition, representativeness is not affected by several factors that affect rational judgments instead and this leads to relevant biases, such as: insensitivity to prior probability, insensitivity to sample size, misconceptions of chance, insensitivity to predictability, illusion of validity and misconceptions of regression.

This heuristic has been widely analysed in the literature from a descriptive point of view, principally with empirical research based on experiments, with application to many different fields such as cognitive sciences, social sciences, behavioural economics, finance, health care, psychology and gambling. For instance, Johnson (Johnson, 1983) demonstrated the representativeness behaviour when assessing risk of bankruptcy; Brannon and Carson (Brannon & Carson, 2003) showed how this heuristic affects health care; Luo (Luo, 2013) investigated the effect of representativeness in financial markets data; Woodland and Woodland (Woodland & Woodland, 2015) showed the impact of this heuristic in sport gambling. As regards the impact of representativeness on decision-making and its context, an extensive recent literature review can be found in Bilek et al. (Bilek, et al., 2018).

However, if we wish to investigate this heuristic as regards the engineering cases of our interests, the work of the above-mentioned authors is not sufficient because they do not propose mathematical models that allow us to reproduce such behaviour. In the literature, there are only a few models attempting to explain this heuristic from a mathematical perspective, see for instance Edward (Edward, 1968), Grether (Grether, 1992) (Grether, 1980), Gigerenzer (Gigerenzer, 1995), Barberis et al. (Barberis, et al., 1998), Tenenbaum and Griffiths (Tenenbaum & Griffiths, 2001), Bordalo et al. (Bordalo, et al., 2016). These models will be studied in detail in the paper presented in chapter 3.

Based on our studies about this heuristic, we think that this behaviour can affect our engineering applications: for instance, it perfectly describes the cognitive bias frequently observed in bridge management about the confusion between condition state and bridge safety, as reported in (Zonta, et al., 2007). Consequently, in chapter 3, we develop a mathematical framework that allows us to reproduce this biased behaviour, based on the literature review about representativeness presented in this section, and we validate it with an application to a real-life case study concerning the structural safety of a bridge.

2.2.3.2 Adjustment or anchoring

The anchoring heuristic was first proposed by Slovic and Lichtenstein in 1971 (Slovic & Lichtenstein, 1971), and later developed by Kahneman and Tversky in 1974 (Tversky & Kahneman, 1974). According to Kahneman and Tversky, people make estimates by starting from an initial value (which may be suggested by the formulation of the problem, or it may be the results of a partial computation), that is adjusted to yield the definitive answer. However, adjustments are typically insufficient, that is, different starting points yield different estimates, which are biased toward the initial values, and this phenomenon is called anchoring. In addition, this heuristic leads to the biases in the evaluation of conjunctive and disjunctive events and also in the assessment of subjective probability distribution (Tversky & Kahneman, 1974).

The anchoring heuristic has been investigated in the literature, an extensive recent literature review about the anchoring effect in decision-making processes can be found in Furnham and Boo (Furnham & Boo, 2011). Moreover, there are many papers where this bias is observed and analysed, principally based on experiments: for instance, Wegener et al. (Wegener, et al., 2010) investigated the anchoring effect in judgment and decision-making; Welsh et al. (Welsh, et al., 2014) examined anchoring in simulated poker-like card games; Mochon and Frederick (Mochon & Frederick, 2013) studied the influence of this heuristic in sequential judgments; Meub and Proeger (Meub & Proeger, 2015) showed how anchoring can be present in social context; Jetter and Walker (Jetter & Walker, 2017) demonstrated the substantial role that can be played by this bias in financial decision-making; De Wilde et al. (De Wilde, et al., 2018) investigated the anchoring effect when decision-making is developed in groups. Other examples can be found in Cohen et al. (Cohen, et al., 1972), Epley and Gilovich (Epley & Gilovich, 2001) (Epley & Gilovich, 2005), Russo (Russo, 2010).

As in the previous case of the representativeness heuristic, also for the anchoring it is not easy to find in the literature mathematical models that allow the reproduction of this behaviour: some authors proposed mathematical equations to describe it, see Hogarth and Einhorn (Hogarth & Einhorn, 1992), Birnbaum and Zimmermann (Birnbaum & Zimmermann, 1998), and Kusev et al. (Kusev, et al., 2018); other authors suggested instead that this bias can be reproduced with existing models, e.g. cognitive process model based on Metropolis-Hastings algorithm (Lieder, 2012), Selective Accessibility model (Mussweiler, et al., 2000), Bidirectional Associative Memory network (Bhatia & Chaudhry, 2013). The most interesting work for our research is the one of Turner and Schley (Turner & Schley, 2016), because they proposed a model that has a clear analogy with Bayes' theorem: the model, called Anchor Integration Model, is a descriptive tool for the measurements and quantification of this bias, and is based on three components, i.e. the prior representation, the influence of the anchor and the posterior representation. It is then evident that in this model the anchoring effect replaces mathematically the function of the likelihood of Bayes' theorem.

As regards the application to our research, the anchoring effect is probably the most influential bias in elicitation processes: for instance, an expert making a series of assessments provides an initial assessment for the first quantity of interest and all subsequent assessments may be adjustments; consequently, it is important to check for trends in order to understand if there is any indicators of the anchoring bias. Consequently, in the paper presented in chapter 4, while proposing a structured methodology for eliciting engineering expert knowledge, we develop each stage of the process in a way that the effect of this bias is minimized.

2.2.3.3 Availability

The last heuristic investigated by Kahneman and Tversky is the availability: an individual evaluates the frequency of classes or the probability of events by availability, i.e. by the ease with which relevant instances come to mind (Kahneman & Tversky, 1973) (Tversky & Kahneman, 1974). Thus, a person could estimate the numerosity of a class, the likelihood of an event or the frequency of co-occurrences by assessing the ease with which the relevant mental operation of retrieval, construction or association can be conducted. This heuristic leads to predictable biases, e.g.: biases due to the retrievability of instances, biases due to the effectiveness of a search set, biases of imaginability and bias in the judgment of the frequency with which two events co-occur, i.e. illusory correlation (Tversky & Kahneman, 1974).

The effect of the availability heuristic in judgment and decision-making has been investigated in the literature, see for instance Taylor and Thompson (Taylor & Thompson, 1982), Jacoby et al. (Jacoby, et al., 1989), Watkins and LeCompte (Watkins & LeCompte, 1991), Betsch and Pohl (Betsch & Pohl, 2002), Oppenheimer (Oppenheimer, 2004). Moreover, many papers examined the presence of this bias in different fields based on experiments: for instance, Maniset et al. (Manis, et al., 1993) analysed the role of availability in judgments of category size and frequency of occurrence; Geurten et al. (Geurten, et al., 2015) investigated the presence of this bias in early childhood; Chen et al. (Chen, et al., 2017) studied the role of availability in investor behaviour; Kudryavtsev (Kudryavtsev, 2018) showed its effect on large daily stock price changes. Other examples can be found in McKelvie and Drumheller (McKelvie & Drumheller, 2001) and Pachur et al. (Pachur, et al., 2012). Conversely, how to reproduce mathematically this behaviour is not studied in the literature: we can just find few papers, such as Shrum and O'Guinn (Shrum & O'Guinn, 1993) and Shrum (Shrum, 1996), where a descriptive model is proposed but without using mathematical equations.

As regards our research, the availability is one of the biases that can adversely influence elicitation processes, therefore, in the same way as for the anchoring, in the development of the methodology presented in chapter 4 we pay close attention to it in order to minimize the risk of biased judgments.

2.2.3.4 Prospect theory

Kahneman and Tversky developed an alternative model as regards the decision step, i.e. the second step of the process presented in section 2.1, which takes into account irrational behaviours and heuristic biases: the prospect theory (PT), which was first introduced in 1979 (Kahneman & Tversky, 1979) and later further developed in 1992 (Kahneman & Tversky, 1992).

PT is an alternative account of individual decision-making under risk and it is developed for simple prospects with monetary outcomes and stated probabilities, but it can be extended to more involved choices. This theory is mainly based on two effects: the certainty effect and the isolation effect. The first means that people underestimate the utility of uncertain scenarios compared to outcomes obtained with certainty, in other words it supports risk aversion in choices involving sure gains and risk seeking in choices involving sure losses. The second is instead based on the fact that people generally discard components that are shared by all prospects under consideration and this leads to inconsistent preferences when the same choice is presented in different forms.

While in EUT utilities of outcomes are weighted by their probabilities, PT suggests an alternative approach, in which the analysis of outcomes should be applied to gains and losses rather than to final assets, and in which probabilities are replaced by decision weights. In particular, risk aversion and risk seeking are determined solely by the utility function when considering EUT. In the PT instead, risk aversion and risk seeking are determined jointly by the utility function and by the capacities, which can be called cumulative weighting functions, or weighting functions for short.

In the original PT introduced in 1979 (Kahneman & Tversky, 1979), the authors proposed the utility function U as a two-part function, in formula:

$$U(z) = \begin{cases} z^{\alpha} & \text{if } z \geq 0 \\ -\lambda(-z)^{\beta} & \text{if } z < 0 \end{cases}, \tag{2.10}$$

where α and β are subjective decision parameters, whereas λ is a coefficient that represents the different slope of the utility function for losses in comparison with the one for gains, i.e. it indicates the degree of loss aversion. Figure 2.3 shows a hypothetical utility function, that can be compared to the one presented in Figure 2.2 about EUT.

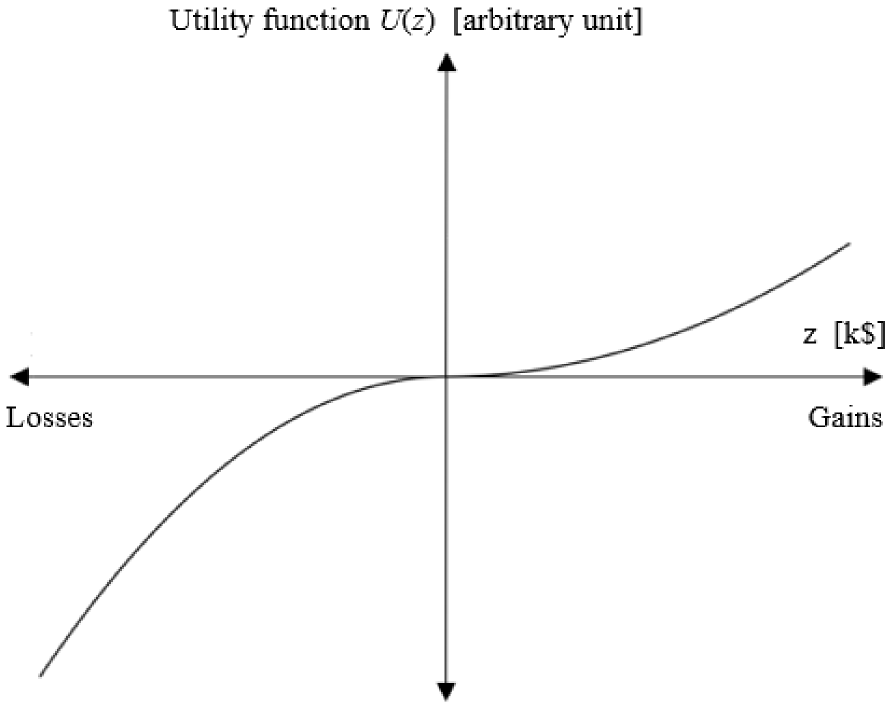


Figure 2.3. Prospect theory: a hypothetical utility function.

Subsequently, the advanced PT (Kahneman & Tversky, 1992) relates the observed nonlinearity of preferences to the shape of the weighting function. In particular, it transforms cumulative rather than individual probabilities and extends the theory to uncertain as well to risky prospects with any number of outcomes. It allows also different weighting functions for gains and for losses and it provides a unified treatment of both risk and uncertainty. The authors derived from different experiments that the cumulative weighting function *w* is concave near the origin. In addition, the conjunction of the above inequalities implies that, in accord with diminishing sensitivity, *w* has an inverted S-shape: it is steepest near the endpoints and shallower in the middle of the range. Figure 2.4 plots *w*, while Eq. (2.11) shows its functional form for gains (+) and losses (-), respectively:

$$w^{+}(p) = \frac{p^{\gamma}}{\left(p^{\gamma} + (1 - p)^{\gamma}\right)^{1/\gamma}}, \qquad w^{-}(p) = \frac{p^{\delta}}{\left(p^{\delta} + (1 - p)^{\delta}\right)^{1/\delta}}, \qquad \text{(2.11a,b)}$$

where γ and δ are subjective decision parameters, while p is the probability.

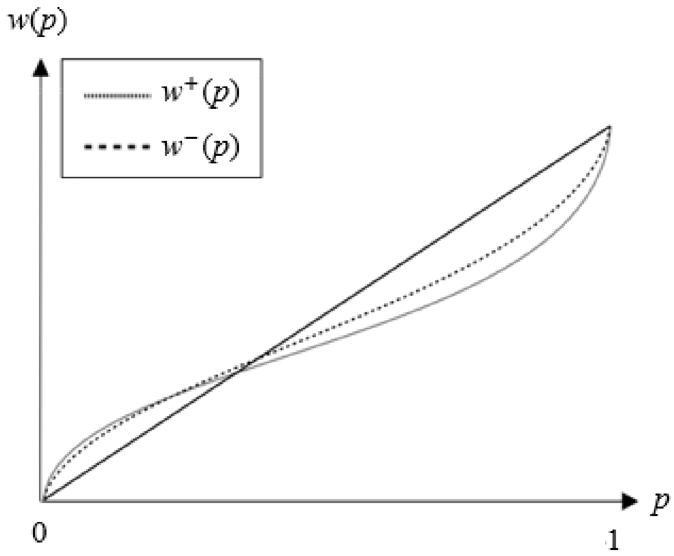


Figure 2.4. Weighting function based on subjective parameters γ and δ.

Following the ideas introduced by the PT, other theories have been proposed, for instance the support theory (ST) (Tversky & Koehler, 1994), introduced by Tversky and Koehler in 1994. This is a theory of subjective probability that proves how various descriptions of the same event may lead to different judgments, due to heuristic behaviours, e.g. the unpacking principle. Unlike the PT, which focuses on the decision, the ST focuses on the judgment.

In conclusion, these theories model the behaviour of decision makers when they are biased by some heuristics and cognitive biases. In the literature, it is possible to find a few papers where a comparison between rational EUT and irrational PT is developed as regards engineering case studies: see for instance our paper (Bolognani, et al., 2017) and also the one of Gong and Frangopol (Gong & Frangopol, 2020), where the behaviour of a bridge manager is in both cases investigated.

2.3 Value of Information of SHM

In this section, we present the state of the art concerning one of the studies that we have developed in our research, i.e. how to evaluate the benefit of SHM. The framework used for the achievement of this aim, called Value of Information (*VoI*), is very interesting for our research since it can be seen precisely as an application of EUT based on a preposterior analysis.

The utility of SHM has rarely been questioned in our community, however only recently a few published papers (Thons & Faber, 2013) (Zonta, et al., 2014) have clarified how to evaluate it. The benefit of information is formally quantified by the so-called Value of Information (*VoI*), a concept anything but new: it was first introduced by Lindley (Lindley, 1956) in 1956, as a measure of the information provided by an experiment, and later formalized by Raiffa and Schlaifer (Raiffa & Schlaifer, 1961) and DeGroot (DeGroot, 1984). Since its introduction, it has been continuously applied in manifold fields, including statistics, reliability (Goulet, et al., 2015), and operational research (Wagner, 1969) (Sahin & Robinson, 2002) (Ketzenberg, et al., 2007) (Quigley, et al., 2017). Its first appearance in the SHM community was implicitly in the 1980s (Thoft-Christensen & Sorensen, 1987), while explicitly it is much more recent and dates back, in our best knowledge, to a paper published in 2005 by Straub and Faber (Straub & Faber, 2005), where it is applied to risk based inspection planning for engineering system. Subsequently, many papers studied the *VoI* for SHM: Bernal et al. proposed a *VoI* framework based on Bayesian decision-making with application to damage detection (Bernal, et al., 2009); Pozzi et al. provided a framework to evaluate the impact of SHM in bridge management based on *VoI* (Pozzi, et al., 2010); Pozzi and Der Kiureghian proposed a framework to evaluate the *VoI* for long-term SHM systems based on Monte Carlo simulations (Pozzi & Der Kiureghian, 2011); Thöns & Faber evaluated the *VoI* for SHM based on a life-cycle cost analysis (Thons & Faber, 2013); Zonta et al. suggested a SHM framework for quantifying the benefit of SHM in bridge management (Zonta, et al., 2014); Limongelli et al. proposed a framework based on the concept of *VoI* for the case of emergency management of road bridges subjected to seismic risk (Limongelli, et al., 2017); Giordano et al. proposed a framework for assessing the *VoI* for SHM of scoured bridges (Giordano, et al., 2020). A recent state of the art can be found in Straub et al. (Straub, et al., 2017) and Thöns (Thons, 2017).

In the last few years, quantifying the value of SHM has known a renewed popularity thanks to the activity of the EU-funded COST action TU1402 (Thons, et al., 2017). In addition, special sessions about *VoI* have been recently organized in International Conferences, e.g. IWSHM and ICASP. In the following, the most recent contributions are summarized: Iannacone et al. proposed a framework for the evaluation of *VoI* of selected inspection procedures based on the information from a SHM system (Iannacone, et al., 2019); Cantero-Chinchilla et al. used the concept of *VoI* as a rational index to provide optimal ultrasonic sensor configuration (Cantero-Chinchilla, et al., 2019); Zhang et al. applied the concept of *VoI* to understand the optimal SHM strategy decision as concerns risk-based inspection planning (Zhang, et al., 2019); Honfi applied the concept of *VoI* for SHM to rational decision-making in bridge management (Honfi, 2019); Valkonen et al. analysed the *VoI* of individual risk preference for decision-making in SHM (Valkonen & Glisic, 2019); Geroulas et al. proposed a process for SHI in the context of maintenance strategy and investment decisions based on *VoI* (Geroulas, et al., 2019).

We introduce in the following the concept of *VoI* as regards the topic of our research. According to (Zonta, et al., 2014), the value of a SHM system can be simply defined as the difference between the benefit, or expected utility u^*, of operating the structure *with* the monitoring system and the benefit, or expect utility u, of operating the structure *without* the system. In formula:

$$VoI = u^* - u. \tag{2.12}$$

Both u^* and u are expected utilities calculated *a priori*, i.e. *before* actually receiving any information from the monitoring system. While in u it is assumed that the knowledge of the manager is his a priori knowledge, u^* is calculated assuming the decision maker has access to the monitoring information and is sometimes referred as to *preposterior utility*. The difference between these values measures the value of the information to the decision maker. Clearly, if the monitoring does not provide any useful information, the preposterior u^* is equal to the prior u, and the value of monitoring information is zero.

In the case of a structure not equipped with a monitoring system, the decision maker decides without accessing any SHM data. In this case, the manager's prior expected utility $u(a_j)$ of a particular action a_j, depends on their prior probabilistic knowledge $\mathrm{P}(S_i)$ of each possible state S_i:

$$u(a_j) = \sum_{i=1}^{N} U\left(\mathbf{z}(a_j, S_i)\right) \mathrm{P}(S_i). \tag{2.13}$$

Consistent with EUT, the rational manager will then choose that actions a_{opt} which carries the maximum expected utility payoff u:

$$u = \max_j u(a_j), \qquad a_{\text{opt}} = \arg\max_j u(a_j). \tag{2.14a,b}$$

Conversely, if a monitoring system is installed, and data are accessible by the agent, the monitoring observation **y** affects the state knowledge, and therefore indirectly their decision. This time, the posterior expected utility $u(a_j,\mathbf{y})$ of actions a_j depends on the posterior probabilities $\mathrm{P}(S_i|\mathbf{y})$, which are now functions of the observation **y**:

$$u(a_j,\mathbf{y}) = \sum_{i=1}^{N} U\Big(\mathbf{z}\big(a_j, S_i\big)\Big)\,\mathrm{P}(S_i|\mathbf{y}). \tag{2.15}$$

Since the posterior probability depends on the particular observation **y**, in the posterior situation the expected utility is a function of **y** as well, and so are the maximum expected utility and the optimal choice:

$$u(\mathbf{y}) = \max_j u(a_j,\mathbf{y}), \qquad a_{\text{opt}} = \arg\max_j u(a_j,\mathbf{y}). \tag{2.16a,b}$$

Eq. (2.14a) and (2.16a) are the utilities calculates before and after a monitoring system is interrogated. Note that, in order to evaluate the posterior utility of an action $u(a_j,\mathbf{y})$, it is required to know the particular realization of observation **y**, so it is not possible to evaluate the posterior utility until the monitoring system is installed and its readings are available. How does the utility change if we have decided to install a monitoring system, but we have still to observe the sensors' readings? Technically, what we should do is to evaluate a priori (i.e. now that the system is not installed yet) the expected value of the utility a posteriori (i.e. at the time when the system will be installed and operating). This quantity is called *preposterior utility*, u^*, to separate it both from the prior and posterior utilities introduced above. The preposterior utility u^* is independent on the particular realization and can be derived from the posterior expected utility $u(\mathbf{y})$ by marginalizing out the variable **y**, (Zonta, et al., 2014) (Cappello, et al., 2016):

$$u^* = \mathrm{E}_{\mathbf{y}}\left[\max_j u(a_j,\mathbf{y})\right] = \int_{D\mathbf{y}} \max_j u(a_j,\mathbf{y}) \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y}, \tag{2.17}$$

where distribution $\mathrm{p}(\mathbf{y})$ is the same *evidence* defined by Eq. (2.2). The preposterior expected utility encodes the total expected utility of a decision process, based on the information provided by the monitoring system, but evaluated before the monitoring system is actually installed.

In conclusion, the Value of Information of the monitoring system is calculated as follows (Zonta, et al., 2014):

$$VoI = u^* - u = \int_{D\mathbf{y}} \max_j u(a_j,\mathbf{y}) \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y} - \max_j u(a_j). \tag{2.18}$$

In other words, the *VoI* is the difference between the *expected maximum* utility and the *maximum expected* utility. It is easily mathematically verified that u^* is always greater or equal than u, and therefore the *VoI* as formulated above can only be positive. This is to say that under the assumption above SHM is always useful, consistently with the principle that "*information can't hurt*" (Cover & Thomas, 2006) (Pozzi, et al., 2017). It is worth reminding that these assumptions are performed before acquiring the data. That means that the value of those data is anticipated by the decision maker, even if the realized value, once the decision is made, may be quite different. As well, it may be that the cost of data exceeds its value, but this would be reflected in the calculation as we assess the utility associated with the cost of obtaining the data.

In addition, this process of deciding on the monitoring system installation can be graphically represented as a two-stage decision tree, as shown in Figure 2.5. At the first stage the agent decides on whether to go or not with the SHM system, while at the second stage he decides on the action $a_1, \ldots, a_j$ to undertake on the structure. The realization of the state occurs at the following chance node and the outcome $\mathbf{z}$ depends on the action and the state. On the 'without SHM' branch of the tree, the state is determined by the prior information and the expected utility corresponds to u in Eq. (2.13). On the 'with SHM' branch of the tree instead, the second stage action is decided based on the information $\mathbf{y}$ from the monitoring system and the final outcome includes the cost $\mathbf{z}_{SHM}$ of the monitoring system. The best choice of stage one is the one that provides maximum utility, and this can be calculated by solving the two-stage tree by backward induction (Parmigiani & Inoue, 2009).

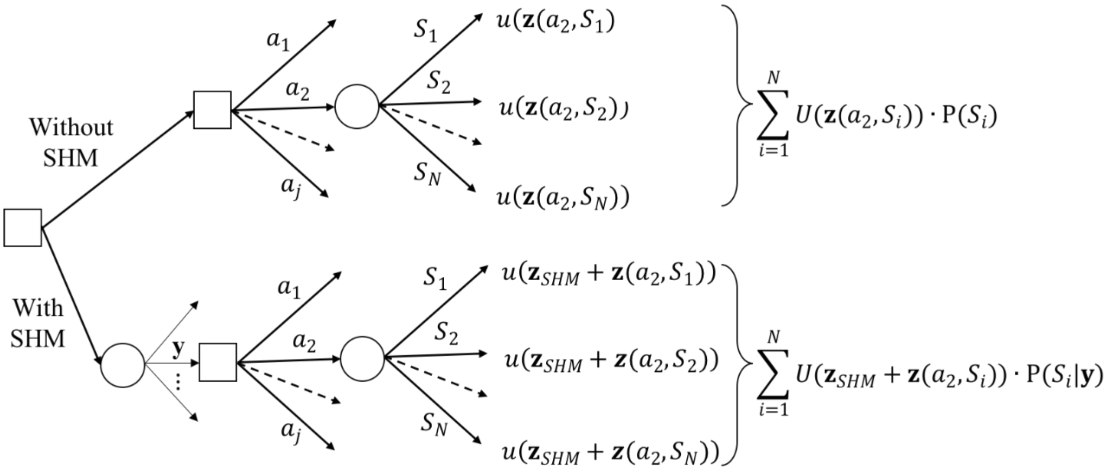


Figure 2.5. Graphical representation of the decision problem of whether or not to install a monitoring system (SHM) (Bolognani, et al., 2018).

In the classical literature, i.e. the formulation presented above, it is assumed that the decision is taken at any stage by the same rational individual. However, we must recognize that in the real world the process whereby a decision maker makes decision is typically more complex, with more individuals involved in the decision chain. Even oversimplifying, we always have at least two different decision stages. Consequently, in chapter 5 and chapter 6, we propose and demonstrate an innovative rational method for the quantification of the *VoI* in the case where two different individuals are involved in the decision process. This framework allows one to investigate how decisions may be distorted due to the difference appetite for risk of decision makers, for instance leading to a negative *VoI*, which is not consistent with the principle that "*information can't hurt*" (Cover & Thomas, 2006) (Pozzi, et al., 2017) introduced above.

2.4 Elicitation process

In this section, we introduce the state of art as concerns another study of our research, i.e. the elicitation process. Elicitation processes are fundamental in order to use expert engineering knowledge as an accurate and reliable data source, and then to support rational decision-making. Expert judgments are useful not only when observed data are not sufficient, but also when they are abundant since the significance of the past to the future can be evaluated also with expertise (Hora, 2007) (Quigley & Walls, 2018). Nowadays, in the research community there is a general agreement about the role of expert knowledge as a useful data source: some researchers even think that observed data are history while expert judgement is the future (Quigley & Walls, 2020).

For the purpose of this contribution, we are interested in eliciting expert judgment in the form of subjective probabilities: it is a socio-technical activity and requires a structured and facilitated process to extract meaningful judgments because people, even experts, are unable to provide accurate and reliable data simply on request (Ferrell, 1994) (Vick, 2002). Indeed, simply asking a person for their best estimate results in poor data due to the plethora of biases in human judgment, such as the ones introduced in section 2.2. Elicitation processes are then designed to minimize the influence of these biases (Quigley & Walls, 2020).Textbooks such as (Cooke, 1991), (Meyer & Booker, 1991) and (Dias, et al., 2018) are extensive references for general aspects of elicitation. For instance, Cooke introduced some generic principles that should be followed by expert judgment processes (Cooke, 1991): accountability, i.e. all data should be available and the results reproducible; empirical control of expert assessments; neutrality, i.e. the elicitation should lead the experts to provide their true beliefs; fairness, i.e. before the process all experts should be considered at the same level.

Even if there is not a protocol for probability elicitation that is universally accepted, the most used protocols that we can find in the literature are three: the Stanford Research Institute (SRI) protocol (Ferrell, 1985) (Spetzler & Stael Von Holstein, 1985) (Merkhofer, 1987), the Morgan and Henrion's protocol and the Wallsten/EPA protocol (Morgan, et al., 1990). They are similar and based on them it is possible to follow seven stages:

- *Motivating*, i.e. to motivate the experts by explaining the aims of the elicitation process and why their expertise is fundamental. In the development of this stage it is important to identify and address motivational biases, such as: management bias, i.e. when experts provide goals rather than judgments, e.g. "*the dam will not fail*"; expert bias, i.e. when experts become overly confident because they have been labelled as "*experts*".

- *Structuring*, i.e. to structure the uncertain quantities of the problem in an unambiguous way. This stage aims also to manage cognitive biases, for instance by disaggregating the quantity of interest into more elemental variables.
- *Conditioning*, i.e. to discuss relevant information in order to condition the expert's judgement to avoid cognitive biases, such as: anchoring bias, i.e. when the evaluation is conditioned by an initial assessment; availability bias, i.e. when the evaluation is based on the ease with which relevant instances come to mind.
- *Encoding*, i.e. to encode the probability distributions of the uncertain quantities using one of the available procedures, such as: direct assessment of probabilities, fractile method, graphical techniques.
- *Verifying*, i.e. to verify the consistency of the elicited distributions by checking that the experts have provided a reflection of their true beliefs and that the elicited probabilities have no indicator of the possible biases.
- *Aggregating*, i.e. in the case of multiple experts, to aggregate the elicited probabilities from different experts to obtain one single final result.
- *Discretizing*, i.e. in the case of continuous variables, to discretize the continuous probability distributions.

The Wallesten/EPA protocol suggests, in addition, to write a document about the aims of the process, the possible heuristics and biases that can influence the process, and other pertinent issues.

To conduct an elicitation process at least two characters are necessary:

- A subject, i.e. the expert, that is who provides expertise, in other words "*a person with substantive knowledge about the events whose uncertainty is to be assessed*" (Ferrell, 1985).
- An analyst, i.e. the interviewer, that is who takes responsibility for designing, developing and executing the process as well as evaluating the procedures. They are also called facilitators, and it is common to have at least one person who is very knowledgeable in elicitation practice and can manage the process,

and another one with wide expertise in the area of the design project.

Recent studies about elicitation processes have introduced other characteristics required to achieve reliable and accurate data from expert knowledge (Quigley & Walls, 2020). For instance, fundamental is the pre-elicitation phase, i.e. a phase of the process before the seven stages introduced above, when the experts have to be carefully selected. In particular, the analysts should identify the essential and desired characteristic of experts and build up profiles of experts who may be able to answer questions concerning the quantities of interest. Constructing a profile matrix can be useful (Bolger, 2018), which matches the knowledge requirements with the expert roles: it supports the identification of expertise needed as well as justification for the choice of experts. The number of required experts depends then on the variability of expertise per domain. Adding as many experts as possible seems beneficial, however, practically it may be difficult to manage many experts and there will be a diminishing return on adding more experts. In addition, it is important to be aware that in real-world it is not so easy to have the availability of many experts.

Elicitation processes have been applied to different fields such as health economic and medicine, ecology, aerospace, nuclear safety, investment banking, business planning and environmental sciences. As regards our engineering field, in the literature it is possible to find only few existing processes for eliciting expert knowledge with engineering applications. For instance, Bubniz et al. (Bubniz, et al., 1998) proposed a process of multiple-expert elicitation and aggregation for a probabilistic seismic hazard analysis. Hodge et al. (Hodge, et al., 2001) suggested an elicitation process where engineering knowledge is used to understand and estimate the reliability performance of complex system. Astfalck et al. (Astfalck, et al., 2018) used engineering expert knowledge to design and operational decision-making in offshore engineering. In summary, while expert knowledge is widely recognized as an important data source in many other fields, in engineering its potential has not yet been fully understood. In addition, it is evident that each specific situation requires a particular elicitation process.

As regards our research, in chapter 4 we have to deal with an engineering application based on a Bayesian network (BN). Eliciting expert knowledge as concerns a BN is particularly challenging, since BNs have a lot of interdependency between the variables and require multiple subsequent assessments. A deep analysis of the state of art shows that very little has been reported about elicitation processes based on BN, especially for civil engineering applications. An example is the paper of Norrington et al. (Norrington, et al., 2008), where an elicitation process aimed specifically to develop the qualitative aspect of a BN is proposed in order to model the reliability of Search and Rescue (SAR) operations in UK. Papers which introduced the possibility to use the elicitation process in order to elicit the conditional probabilities of a BN are (Sigurdsson, et al., 2001) and (Christophersen, et al., 2018). Nevertheless, a structured methodology for eliciting expert knowledge to support the collection of valid and reliable data in order to quantify a BN does not exist, so in chapter 4 we propose a four-stage structured elicitation process with this specific aim, based on the literature review and on the specific requirements of the case study.

3. Consequences of representativeness bias on SHM-based decision-making

Submitted for publication to *Structure and Infrastructure Engineering*
(Under revision)

Andrea Verzobio[a], Denise Bolognani[b], John Quigley[c], Daniele Zonta[d]

Position and affiliation:
[a]PhD Student, Department of Civil and Environmental Engineering, University of Strathclyde, UK
[b]Postdoc, Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy
[c]Professor, Department Management Science, University of Strathclyde, UK
[d]Professor, Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy

Summary of the paper

This paper investigates how heuristic behaviours may affect human judgment and decision-making in civil engineering. In particular, we identify Kahneman and Tversky's representativeness (introduced in section 2.2.3.1), as a heuristic for which SHM-based decision-making is particularly susceptible, where simplified rules for updating probabilities can distort the decision maker's perception of risk. In this contribution, we describe mathematically this specific heuristic in order to understand how it affects the interpretation of data, providing a deeper understanding of the differences between a heuristic method affected by cognitive biases and the classical rational approach (introduced in section 2.1). Our study is conducted both theoretically through comparison with formal Bayesian methods as well as empirically through the application to a real-life case study about the evaluation of the safety of a bridge.

3.1 Introduction

Structural health monitoring (SHM) is commonly recognized as a powerful tool that allows bridge managers to make decisions on maintenance, reconstruction and repair of their assets. The logic of making decision based on SHM is formally stated in Cappello et al. (Cappello, et al., 2016), under the assumption that the decision maker is an ideal rational agent, who judges using Bayes' theorem (Bolstad, 2010) and decides consistently with Neumann-Morgenstern's Expected Utility Theory (EUT) (Neumann & Morgenstern, 1944). Not that surprisingly, we often observe real-life decision makers departing from this ideal model of rationality, judging and deciding using common sense and privileging fast and frugal heuristics to rational analytic thinking. Hence, if we wish to describe mathematically and to predict the choices of real-world bridge managers, we have to accept that their behaviour may not be necessarily fully rational. Biased judgement and decision-making have been widely reported and systematically investigated starting the 1970s in the fields of cognitive sciences, social sciences and behavioural economics: key papers include the fundamental works by Kahneman and Tversky (Kahneman & Tversky, 1973) (Tversky & Kahneman, 1974) (Tversky & Kahneman, 1983) (Kahneman & Tversky, 1979); Kahneman's famous textbook (Gilovich, et al., 2002) is an extensive reference for those approaching the topic for the first time.

As regards SHM-based bridge management, apparent irrational behaviours are reported in (Zonta, et al., 2014) (Bolognani, et al., 2018) (Bolognani, et al., 2017), and also suggested in (Cappello, et al., 2016). In particular, a typical example of cognitive bias frequently observed in bridge management is the confusion between condition state and safety of a bridge, as reported for instance in (Zonta, et al., 2007). We remind here for clarity that safety is about the capacity of a bridge to withstand the traffic loads and the other external actions without collapsing, while the condition state expresses the degree of deterioration of a bridge, or bridge element, respect to its design state. The condition state is usually apprised through a combination of routing visual inspections, non-destructive evaluation and SHM. It is expressed in the form of a condition index that depends on the particular management system. For example, bridge management systems based on AASHTO (American Ass. State Highway and Transportation Off, 1997) Commonly Recognized (CoRe) Standard Element System, such as PONTIS, BRIDGIT and the APT-BMS reported in (Zonta, et al., 2007), classify the state of an element on a scale from 1 to 5, where 1 means 'as per design' and 5 corresponds to the most severe observable deterioration state. On the contrary, the safety of a bridge is typically encoded in its probability of failure P_F, reliability index β, or safety factor γ, evaluated through formal structural analysis. Condition state and safety are obviously correlated (logically, the load-carrying capacity of a deteriorated bridge is equal or lower than that of the same bridge in undamaged condition) but are not the same thing. For example, an old bridge can be unsafe, regardless its preservation state, simply because it was designed to an old code, which does not comply with the current load demand. As a counterexample, we may have the case of bridge, severely deteriorated, but still with enough capacity to safely withstand all the external loads, either because overdesigned or simply because its deterioration does not affect its load-carrying capacity. In principle, rational bridge management should target the safety of the bridge stock, and therefore prioritize retrofit of unsafe bridges, regardless of their degree of deterioration. In practice, we frequently observe that bridge managers tend to delay retrofit of substandard bridges which do not show sign of deterioration, while repair promptly deteriorated bridges as soon as the damage is observed, regardless of the actual residual load-carrying capacity. The biased rationale behind this apparent behaviour is that undamaged bridges 'look' safe, while damaged bridges 'look' unsafe, simply because, generally speaking, we know that deterioration negatively affects safety.

The aim of this paper is to describe mathematically this observed biased judgement, a condition that, we will show, is broadly described by Kahneman and Tversky's representativeness heuristic (Kahneman & Tversky, 1972). We clarify that it is not objective of this paper to suggest that it is correct to use representativeness to judge the state of a bridge, and we reiterate that the only rational way to judge in presence of uncertainties is to use Bayesian logic. We just wish to verify whether the irrational judgment sometimes observed in bridge managers' behavior could be described and possibly predicted using Kahneman and Tversky's representativeness heuristic. Being able to predict the behavior of an irrational manager is necessary when we set a general policy for bridge maintenance and we know that someone else who is going to enact the policy may behave irrationally. As an example, Gong and Frangopol (Gong & Frangopol, 2020) discuss a case where modelling the irrational behavior of a manager is instrumental to an optimization process in bridge maintenance.

We begin reminding, in section 3.2, the formal framework of rational decision based on SHM information. We discuss in section 3.3 various classical judgmental heuristics and the consequential biases, while, in section 3.4, the mathematical model of representativeness is developed to appropriately reproduce the heuristic behaviour. Then, in section 3.5, we develop a classical representativeness problem to assess the model. Finally, section 3.6 presents the engineering application where the model is used to reproduce the biased evaluation about the safety of a bridge, based on the condition state apprised through visual inspections. Some concluding remarks are presented at the end of the paper.

3.2 SHM-based decision-making rational framework

We refer to the problem of optimal decision based on data provided by SHM. As shown in Figure 3.1, SHM-based decision-making is properly a two-step process, which includes the judgement of the state of the structure h based on the observations $\mathbf{y}$, and the decision of the optimal action a_{opt} based on the uncertain knowledge of the state. Within the scope of this paper, we define observation to be any information acquired on site which is suitable to infer the state of the structure. Sources of observation, in the broad sense, could be visual inspections, site tests, sensors temporarily or permanently installed on the structure.

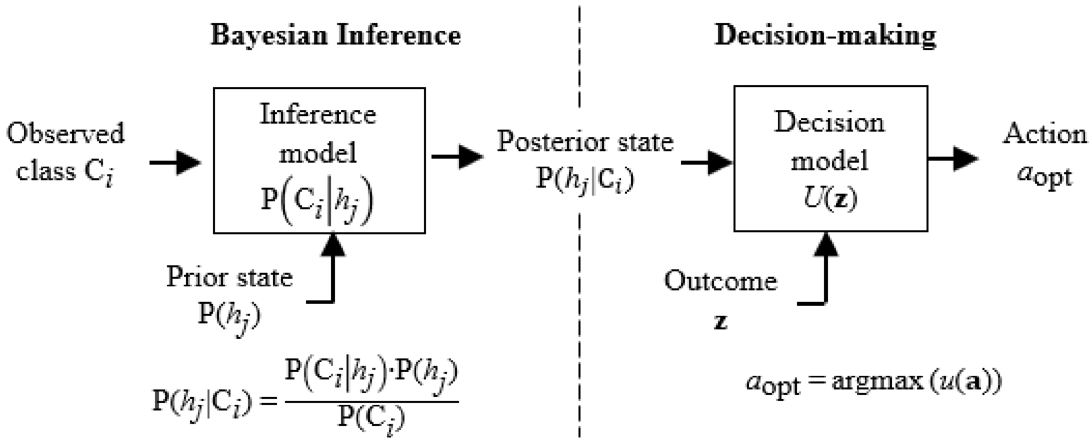


Figure 3.1. The rational process of SHM-based decision-making.

Assume that the safety state of the bridge is described by one of n mutually exclusive and exhaustive state hypothesis $\mathcal{H} = \{h_1, h_2, \ldots, h_j, \ldots, h_n\}$ (e.g.: h_1 = 'safe', …, h_n = 'failure'). Further assume that observing the bridge, or bridge element, either through visual inspection or SHM, ultimately consists of assessing its condition out of a number of m possible classes $C_1, C_2, \ldots, C_i, \ldots, C_m$ which express its degree of damage or deterioration (e.g.: C_1 = 'not damaged', C_2 = 'moderately damaged', C_3 = 'severely damaged', …). Therefore, the value of an observation y_i is one of the possible condition classes: $y_i \in \{C_1, C_2, C_3, C_4, C_5\}$. Multiple independent observations on the same bridge may occur because of repeated inspections by different inspectors, or redundant independent measurements by the monitoring system. We indicate with vector **y** the full set of observations $\mathbf{y} = \{y_1, y_2, \ldots, y_k \ldots, y_N\}$. The likelihood of condition C_i for a bridge, or bridge element, in state h_j is then encoded in the probabilistic distribution $P(C_i|h_j)$.

If we restrict the problem to a single-observation case, the first step of the process consists of judging the state of a structure h_j based on the i-th class observed C_i. In the presence of uncertainty, the state of the structure after observing the class C_i is probabilistically described by the posterior probability $P(h_j|C_i)$, and the inference process followed by a rational agent is mathematically developed in Bayes' rule (Sivia & Skilling, 2006) (Bolstad, 2010):

$$P(h_j|C_i) = \frac{P(C_i|h_j)\, P(h_j)}{P(C_i)}, \tag{3.1}$$

where $P(h_j|C_i)$ is the posterior knowledge of the structural state and represents the best estimation after the acquisition of SHM observation. It depends on the likelihood $P(C_i|h_j)$ and the prior knowledge $P(h_j)$, which is our estimate of the structural state h_j before the acquisition of the observation. $P(C_i)$ is simply a normalization constant, referred to as evidence, calculated as:

$$P(C_i) = \sum_{j=1}^{n} P(C_i|h_j)\, P(h_j). \tag{3.2}$$

The second step of the process starts after the assessment of the posterior probability of the structure, and concerns choosing the 'best' action. The decision maker can choose between a set of M alternative actions a_1, a_2, …, a_M (e.g.: a_1 = 'do nothing', a_2 = 'limit traffic', a_3 = 'close the bridge to traffic', …). Taking an action produces measurable consequences (e.g.: a monetary gain or loss, a temporary downtime of the structure, in some case causalities) and the consequences of an action can be mathematically described by several parameters (e.g.: the amount of money lost, the number of days of downtime, the number of casualties), encoded in an outcome vector **z**. The outcome **z** of an action depends on the state of the structure; thus, it is a function of both action a and state h_j, i.e. $\mathbf{z}(a, h_j)$. When the state is certain the consequence of an action is also deterministically known; therefore, the only uncertainty in the decision process is the state of the structure h_j. The rational decision maker ranks actions based on the consequences **z** through a utility function $U(\mathbf{z})$, which can vary among different individuals with different behaviors. According to the different risk appetite of the decision maker, the utility function can be risk neutral, risk adverse or risk seeking. Expected utility theory (EUT) describes the analysis of decision-making under risk and is considered as a normative model of rational choice (Parmigiani & Inoue, 2009). EUT was introduced by von Neumann and Morgenstern in 1944 (Neumann & Morgenstern, 1944) and later developed in the form that we currently know by Raiffa and Schlaifer in 1961 (Raiffa & Schlaifer, 1961). Its axioms state that the decision maker ranks their preferences based on the expected utility u, defined as:

$$u(a) = E_{h_j}\left[U\left(\mathbf{z}\left(a, h_j\right)\right)\right], \tag{3.3}$$

where E_{h_j} is the expected value operator of the random variable h_j, while U indicates the utility function. The latter is very important and represents the evaluation of a decision maker's beliefs about the outcome **z**. The decision maker then chooses the action that maximizes the expected utility.

In summary, the rational way to decide based on observation in presence of uncertainties goes through a judgment based on Bayes' theorem and a proper decision based on EUT.

3.3 Heuristics and biases

It is possible to demonstrate that Bayes' theorem and EUT provide the only consistent way to judge and make decision under uncertainties, respectively, while any alternative inference or decision model may produce a logical inconsistency. Nevertheless, it is frequently observed that most people in everyday life favor heuristic approaches (Gilovich, et al., 2002) (Kahneman & Tversky, 1979) to this rational framework in order to judge or make decisions.

The concept of heuristic has been defined in different ways in the scientific literature, depending on the discipline and the scope of application, see for instance (Tonge, 1960) (Feigenbaum & Feldman, 1963) (Romanycia & Pelletier, 1985) (Gigerenzer & Gaissmaier, 2011). For the purpose of this paper, we define a heuristic, together with Feigenbaum and Feldman (Feigenbaum & Feldman, 1963), as any approach to judgement or decision based on rules of thumb, logical simplifications or shortcuts rather that the proper rational process, as described in section 3.2. Possibly, the most important contribution to the formal characterization of the heuristic behavior is the work that Kahneman and Tversky carried out in the early 1970s (Kahneman & Tversky, 1972) (Kahneman & Tversky, 1973) (Tversky & Kahneman, 1974), which had a significant impact to the understanding and description of the human behavior and represents the basis of a discipline we currently refer to as *behavioral economics*. They developed the so-called *heuristics and biases approach*, challenging the dominance of strictly rational models. The main innovation lays in the analysis of the descriptive adequacy of ideal models of judgment and in the proposal of a cognitive alternative that explained human error without invoking motivated irrationality. Evidence displays that people's assessments of likelihood and risk do not conform to the laws of probability. They offer in (Tversky & Kahneman, 1974) a list of frequently observed heuristics which include:

(1) Representativeness. Events are ranked according to their representativeness; people consistently judge the more representative event to be the more likely, whether it is or not (Kahneman & Tversky, 1972). Representativeness is not affected by several factors that affect rational judgments instead and this leads to relevant biases, such as: insensitivity to prior probability, insensitivity to

sample size, misconceptions of chance, insensitivity to predictability, illusion of validity and misconceptions of regression (Tversky & Kahneman, 1974).

(2) Availability. An individual evaluates the frequency of classes or the probability of events by availability, i.e. by the ease with which relevant instances come to mind (Kahneman & Tversky, 1973) (Tversky & Kahneman, 1974). Thus, a person could estimate the numerosity of a class, the likelihood of an event or the frequency of co-occurrences by assessing the ease with which the relevant mental operation of retrieval, construction or association can be conducted. It leads to predictable biases, e.g.: biases due to the retrievability of instances, biases due to the effectiveness of a search set, biases of imaginability and biases in the judgment of the frequency with which two events co-occur, i.e. illusory correlation.

(3) Adjustment or anchoring. People make estimates by starting from an initial value (which may be suggested by the formulation of the problem, or it may be the results of a partial computation), that is adjusted to yield the definitive answer. However, adjustments are typically insufficient, that is, different starting points yield different estimates, which are biased toward the initial values, and this phenomenon is called anchoring (Tversky & Kahneman, 1974).

Depending on their nature, a heuristic can affect the process outlined in section 3.2 in the inference step, in the decision step, or in both cases. In the rest of the paper we will focus on the representativeness, that seems the heuristic that better reproduces the irrational behaviour introduced in section 3.1. This specific heuristic affects the inference step of the process, i.e. the judgment.

3.4 The representativeness heuristic

Representativeness is commonly intended as the level of how well or how accurately something reflects upon a sample. A judgment is biased by the representativeness heuristic when the ordering of hypotheses h_j by subjective perceived probabilities coincides with their ordering by representativeness, rather than by Bayes' posterior probability (Kahneman & Tversky, 1972). In other words, a hypothesis, or event A, is judged more probable than a hypothesis, or event B, whenever A appears more representative than B. Citing Kahneman and Tversky (Kahneman & Tversky, 1972), an individual who follows the representativeness heuristic "*evaluates the probability of an uncertain event, or a sample, by the degree to which it is: (i) similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated*".

The literature illustrates numerous cases of behavioral experiments where representativeness bias is observed. For example, in a classic experiment reported in (Tversky & Kahneman, 1974), the interviewee is asked to assess the probability of Steve's employment from a list of possibilities (e.g. farmer, salesman, airline pilot, librarian or physician), simply based on this description: "*Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail*." It is observed that most interviewees tend to judge highly likely that Steve is a librarian, simply because the description provided is representative of the stereotype of a librarian, and with complete disregard for the proportion of the population that are librarians compared with the other employments. This example also clarifies that to be representative an uncertain event should not only be similar to its parent population, but it should also reflect the properties of the uncertain process by which it is generated. This agreement on the representativeness formulation is in line with the definition in (Tversky & Kahneman, 1983); they write that: "*an attribute is representative of a class if it is very diagnostic; that is, the relative frequency of this attribute is much higher in that class than in the relevant reference class.*"

While representativeness heuristic has been widely analysed from a descriptive point of view, in the literature there are only few models attempting to describe this heuristic from a mathematical perspective, see for instance Edward (Edward, 1968), Grether (Grether, 1992) (Grether, 1980), Gigerenzer (Gigerenzer, 1995), Barberis et al. (Barberis, et al., 1998), Tenenbaum and Griffiths (Tenenbaum & Griffiths, 2001), Bordalo et al. (Bordalo, et al., 2016). While introducing these models, we want to point and analyse, case by case, two main aspects regarding the definition of representativeness and its application, which are:

(1) What is the mathematical formulation of representativeness proposed by the different authors?
(2) To what extent and how does the representativeness bias the final judgment in comparison to Bayes' rule?

3.4.1 Formulation of Representativeness

In the literature mentioned above there is a general agreement whereby the degree of representativeness of an observable class C_i for a reference hypothesis h_j is in some way related to odds of observable C_i, which are the ratio between its likelihood $P(C_i|h_j)$ and the likelihood of its negation $P(C_i|-h_j)$, where $-h_j$ denotes the set of alternative hypotheses.

Edward (Edward, 1968), Gigerenzer (Gigerenzer, 1995) and Bordalo et al. (Bordalo, et al., 2016) all define the quantity representativeness $R(C_i, h_j)$ of a class C_i for the reference hypothesis h_j, exactly as the odds of class C_i:

$$R(C_i, h_j) = \frac{P(C_i|h_j)}{P(C_i|-h_j)}. \tag{3.4}$$

Therefore, they assume that a class C_i is representative for a hypothesis h_j, relative to an alternative hypothesis $-h_j$, if it scores high on the likelihood ratio described by Eq. (3.4).

Similarly, Tenenbaum and Griffiths (Tenenbaum & Griffiths, 2001) define representativeness with the likelihood ratio described by Eq. (3.4), but using a logarithm scale, apparently to provide a more natural measure of how good a class C_i is in representing a hypothesis h_j:

$$R(C_i, h_j) = \log \frac{P(C_i|h_j)}{P(C_i|-h_j)}. \tag{3.5}$$

Grether (Grether, 1980) (Grether, 1992) agrees on Eq. (3.5) for a problem with two possible hypotheses. In the case of more alternative hypotheses, Tenenbaum and Griffiths (Tenenbaum & Griffiths, 2001) suggest the following expression:

$$\mathrm{R}\left(\mathrm{C}_i, h_j\right) = \log \frac{\mathrm{P}\left(\mathrm{C}_i \middle| h_j\right)}{\sum_{h_k,\, \mathrm{k}\neq \mathrm{j}} \mathrm{P}(\mathrm{C}_i|h_k)\, \mathrm{P}\left(h_k \middle| -h_j\right)}, \tag{3.6}$$

where $\mathrm{P}\left(h_k \middle| -h_j\right)$ is the prior probability of the k-th hypothesis, given that the reference hypothesis h_j is not the true explanation of C_i: 0 when $j = k$ and $\mathrm{P}(h_k)/(1-\mathrm{P}(h_j))$ otherwise. Eq. (3.6) effectively says that C_i is representative of h_j to the extent that its likelihood under h_j exceeds its average likelihood under alternative hypotheses.

3.4.2 Representativeness in judgment

Before revising the mathematical models proposed to reproduce the representativeness bias in judgment, we remind, for maximum clarity, that the rational way to judge the probability of a hypothesis h_j based on an observation class C_i is to calculate its posterior probability $\mathrm{P}\left(h_j \middle| \mathrm{C}_i\right)$ in Bayesian sense, using Eq. (3.1). When judging using representativeness heuristic, an individual ranks the hypothesis h_j by a subjective perceived probability which departs from standard Bayesian posterior. In analogy with Bordalo et al. (Bordalo, et al., 2016), we define this subjective perceived probability as distorted posterior $\mathrm{P}(h_j|\mathrm{C}_i)^{st}$. While all authors agree on that representativeness distorts judgment, there is not a general agreement on the cognitive mechanism whereby representativeness affects the distorted posterior probability, i.e. how the standard Bayes' rule, which reflects the judgment of a rational thinker, must be adjusted to consider representativeness instead. Most authors do not provide an explicit expression for the distorted posterior, but understand the vanilla statement that ordering hypotheses by perceived probability follows representativeness rather than Bayesian posterior. From a strict mathematical standpoint, we can define different models of distorted posterior that satisfy this statement. The simpler is assuming that (i) representativeness is used instead of likelihood and (ii) the prior information is neglected. In this case, judgment by representativeness should be consistent with the following expression:

$$\mathrm{P}\left(h_j \middle| \mathrm{C}_i\right)^{st} = \frac{\mathrm{R}(\mathrm{C}_i, h_j)}{\mathrm{R}(\mathrm{C}_i, h_j) + \mathrm{R}(\mathrm{C}_i, -h_j)}. \tag{3.7}$$

Some of the authors introduced above provide more refined models. Bordalo et al. (Bordalo, et al., 2016) suggest that representativeness $\mathrm{R}(\mathrm{C}_i, h_j)$ distorts Bayesian likelihood $\mathrm{P}(\mathrm{C}_i, h_j)$ as follows:

$$\mathrm{P}(\mathrm{C}_i | h_j)^{st} = \mathrm{P}(\mathrm{C}_i | h_j) \cdot (\mathrm{R}(\mathrm{C}_i, h_j))^{\theta}, \tag{3.8}$$

where $\theta \geq 0$ is a subjective parameter that describes how heavily representativeness biases the likelihood. According to the same authors, this parameter should be calibrated with cognitive tests and could vary considerably among different people. A biased posterior is therefore inferred, using this distorted likelihood into Bayes' theorem:

$$\mathrm{P}(h_j | \mathrm{C}_i)^{st} = \frac{\mathrm{P}(\mathrm{C}_i | h_j)^{st} \, \mathrm{P}(h_j)}{\mathrm{P}(\mathrm{C}_i)^{st}}, \tag{3.9}$$

where $\mathrm{P}(\mathrm{C}_i)^{st}$ is the distorted evidence, calculated as:

$$\mathrm{P}(\mathrm{C}_i)^{st} = \sum_{j=1}^{n} \mathrm{P}(\mathrm{C}_i | h_j)^{st} \, \mathrm{P}(h_j). \tag{3.10}$$

It is easily noticed that Eq. (3.9) is exactly Bayes' theorem when $\theta = 0$, while it collapses to Eq. (3.7) when θ tends to infinity.

A different approach is provided by Grether (Grether, 1980) (Grether, 1992). The author suggests a model that provides the final judgment of h_j, by considering the representativeness heuristic:

$$\log O(h_j | \mathrm{C}_i) = \alpha + \beta_1 \cdot \mathrm{R}(\mathrm{C}_i, h_j) + \beta_2 \cdot \log O(h_j), \tag{3.11}$$

where $O(h_j|\mathrm{C}_i)$ is the posterior odds, $\mathrm{R}(\mathrm{C}_i, h_j)$ is the representativeness calculated as in Eq. (3.5), $O(h_j)$ is the prior odds, while α, β_1 and β_2 are subjective parameters that must be calibrated. Thus, the interpretation of Kahneman and Tversky's representativeness heuristic suggested by the author is that individuals place greater weight on the likelihood ratio than on the prior odds. Consequently, the author proposed $\beta_1 > \beta_2 \geq 0$ for this inference model, in contrast with $\alpha = 0, \beta_1 = \beta_2 > 0$ of

Bayes' rule.

With the aim to compare these last two judgement models, we express Bordalo et al.'s model, stated in Eq. (3.8), in its logarithmic posterior odds:

$$\log O(h_j|\mathrm{C}_i) = (2\theta + 1)\cdot\mathrm{R}(\mathrm{C}_i, h_j) + \log O(h_j), \tag{3.12}$$

where $\mathrm{R}(\mathrm{C}_i, h_j)$ is, in the same way as in Eq. (3.11), the representativeness calculated as in Eq. (3.5). It is possible to notice that this final equation agrees with the one proposed by Grether, i.e. Eq. (3.11), if we assume $\alpha = 0$, $\beta_1 = (2\theta + 1)$ and $\beta_2 = 1$. This means that the two models are based on the same mathematical formulation, they only differ in the representation of the subjective parameters.

In summary, while there is a general agreement on the definition and the mathematical formulation of the representativeness, different inference models are proposed or understood to describe the biased judgment. Moreover, some of these models account for a number of subjective parameters that have to be properly calibrated on the individual who judges.

3.5 A classical representativeness problem

Before developing the bridge engineering problem that is motivating our research, we discuss in this section how the models introduced in section 3.4 apply to a classical representativeness problem, reported in different forms in (Tversky & Kahneman, 1974) (Griffin & Tversky, 1992) (Tenenbaum & Griffiths, 2001) (Griffiths & Tenenbaum, 2007).

Consider two coin-flip sequences, C_1 = HHHHH and C_2 = HTTHT, where H is for *Head* and T for *Tail*. To start, we would like to clarify the difference between representativeness and likelihood. The first question we ask ourselves is: which of the two sequences is more *representative* for a fair coin? We presume that most of the readers would answer sequence C_2. Actually, we expect that a fair coin would generally produce a random sequence of H and T, as in C_2, while a sequence of H only, as in C_1, looks intuitively peculiar from a genuinely fair coin. Intuitively, we conclude that the representativeness of sequence C_2 is greater than the representativeness of sequence C_1 in the case of a fair coin h_{FC}. In formula:

$$\mathrm{R}(\mathrm{C}_2, h_{\mathrm{FC}}) > \mathrm{R}(\mathrm{C}_1, h_{\mathrm{FC}}). \tag{3.13}$$

However, which of the two sequences is more likely to occur for a fair coin? In this case we can simply calculate the likelihood of a sequence, i.e. the probability of obtaining that particular sequence C_i conditional to the assumption of fair coin $\mathrm{P}(\mathrm{C}_i|h_{\mathrm{FC}})$, by computing the possible combinations. If our coin is fair, for each toss we have equal probability of p = 1/2 of H or T. Therefore, the particular sequence C_2, which is the result of 5-coin tosses, has the following likelihood:

$$\mathrm{P}(\mathrm{C}_2{=}\mathrm{HTTHT}|h_{\mathrm{FC}}) = \left(\frac{1}{2}\right)^5 = 0.0313. \tag{3.14}$$

Notice now that even sequence C_1 is a possible output of 5-coin tosses, and therefore its likelihood is exactly the same as C_2:

$$\mathrm{P}(\mathrm{C}_1{=}\mathrm{HHHHH}|h_{\mathrm{FC}}) = \left(\frac{1}{2}\right)^5 = 0.0313. \tag{3.15}$$

Let's now ask to the layman the following question: a coin has produced sequence C_1; based on this sequence, do you believe this coin is most likely fair or has a prevalence of H? Most of the interviewees, and possibly even the reader, answer that the coin is most likely unfair, i.e. with a prevalence of H. Let's tackle the problem in logical terms using Bayes' theorem. As regards the coin with a prevalence of H, we refer to a coin that mostly comes up heads with the term h_{MH}. In essence, we have to calculate the following posterior probability:

$$\mathrm{P}(h_{\mathrm{FC}}|\mathrm{C}_1) = \frac{\mathrm{P}(\mathrm{C}_1|h_{\mathrm{FC}})\ \mathrm{P}(h_{\mathrm{FC}})}{\mathrm{P}(\mathrm{C}_1|h_{\mathrm{FC}})\ \mathrm{P}(h_{\mathrm{FC}}) + \mathrm{P}(\mathrm{C}_1|h_{\mathrm{MH}})\ \mathrm{P}(h_{\mathrm{MH}})}. \tag{3.16}$$

The condition whereby it is more probable that the coin is fair is that the posterior probability of h_{FC} is greater than 0.5, or:

$$\frac{\mathrm{P}(h_{\mathrm{FC}})}{\mathrm{P}(h_{\mathrm{MH}})} > \frac{\mathrm{P}(\mathrm{C}_1|h_{\mathrm{MH}})}{\mathrm{P}(\mathrm{C}_1|h_{\mathrm{FC}})}, \tag{3.17}$$

which means that the ratio between the priors has to be greater than the ratio between the opposite likelihoods. If p is the probability of occurrence of H in a toss, n the number of coin tosses and k the number of H achieved in a sequence C_i, the likelihood of this sequence can be calculated as follows:

$$P(C_i|\ h_{FC}) = p^k\ (1 - p)^{n - k}. \tag{3.18}$$

In the case of a fair coin we have already observed that p = 1/2 and therefore $P(C_1|\ h_{FC}) = P(C_2|\ h_{FC}) = 0.0313$. On the other hand, coin h_{MH} is the one that mostly comes up heads, and therefore, since C_1 is a sequence with a prevalence of H, the only thing that we can conclude is that:

$$P(C_1|\ h_{MH}) > P(C_1|\ h_{FC}). \tag{3.19}$$

Therefore, the only logical conclusion we can draw is that the ratio between the two likelihood, useful to solve Eq. (3.17), is strictly greater than 1. In any case, it states nothing about the posterior because it depends also on the prior rate, i.e. how likely is a priori, before observing the sequence, that the coin is fair. There is always a value of prior $P(h_{FC})$ whereby it is more probable that, given the sequence C_1, the coin is fair:

$$P(h_{FC}|\ C_1) > P(h_{MH}|\ C_1). \tag{3.20}$$

In conclusion, from a strict logical standpoint, the coin could be fair or nor fair depending on the prior information.

Let's make a numerical example: we suppose to have a fair coin h_{FC}, and a coin that mostly comes up heads h_{MH}, assuming that the probability of occurrence of H with this coin is p = 0.85. For concreteness, we choose the following prior probabilities for the two hypotheses: $P(h_{FC}) = 0.95$ and $P(h_{MH}) = 0.05$. First of all, we have to calculate all the likelihood and the representativeness values. Using respectively Eq. (3.18) and Eq. (3.4), we obtain:

$$P(C_1|h_{FC}) = 0.0313, \qquad P(C_2|h_{FC}) = 0.0313. \tag{3.21a,b}$$

$$P(C_1|h_{MH}) = 0.4437, \qquad P(C_2|h_{MH}) = 0.0024. \tag{3.22a,b}$$

$$R(C_1|h_{FC}) = 0.07, \qquad R(C_2|h_{FC}) = 13.04. \tag{3.23a,b}$$

$$\mathrm{R}(\mathrm{C}_1|h_{\mathrm{MH}}) = 14.18, \qquad \mathrm{R}(\mathrm{C}_2|h_{\mathrm{MH}}) = 0.08. \qquad (3.24\text{a,b})$$

These results confirm what we were presuming and clearly show the difference between representativeness and likelihood: while sequences HTTHT and HHHHH are equally likely for a fair coin, i.e. $\mathrm{P}(\mathrm{C}_1|h_{\mathrm{FC}}) = \mathrm{P}(\mathrm{C}_2|h_{\mathrm{FC}})$, the representativeness model shows that sequence HTTHT is clearly more representative for a fair coin than sequence HHHHH, i.e. $\mathrm{R}(\mathrm{C}_2, h_{\mathrm{FC}}) > \mathrm{R}(\mathrm{C}_1, h_{\mathrm{FC}})$. This outcome reflects the effect of such heuristic bias, because most people judge the sequence HTTHT to be more likely for a fair coin than sequence HHHHH, which does not appear random, even if the two sequences have the same probability of occurrence. The second column of Table 3.1 presents the achieved results.

Let's now calculate the Bayesian posterior probabilities that, given sequence C_1, the coin is fair, or it is the one that mostly comes up heads. Using Bayes' theorem, as in Eq. (3.16), we achieve:

$$\mathrm{P}(h_{\mathrm{FC}}|\mathrm{C}_1) = 57.27\% > \mathrm{P}(h_{\mathrm{MH}}|\mathrm{C}_1) = 42.73\%. \qquad (3.25)$$

Notice that, with the prior assumptions made, the rational conclusion is that the coin is most probably fair, even if it has yielded a sequence of 5 heads in a row. This result may sound counterintuitive to the layman, unfamiliar with formal logic, who tends to judge heuristically driven by the representativeness of the observed result.

We can reproduce this heuristic behaviour using, for instance, the vanilla inference model of Eq. (3.7), which indeed yields the following distorted posterior judgments:

$$\mathrm{P}(h_{\mathrm{FC}}|\mathrm{C}_1)^{st} = 0.49\% < \mathrm{P}(h_{\mathrm{MH}}|\mathrm{C}_1)^{st} = 99.51\%, \qquad (3.26)$$

which is to say that to the individual biased by representativeness the coin looks most likely the one that mostly comes up heads. We find a similar result using the other inference model of Bordalo et al., as in Eq. (3.9). With a subjective parameter $\theta = 0.8$, we obtain:

$$\mathrm{P}(h_{\mathrm{FC}}|\mathrm{C}_1)^{st} = 1.88\% < \mathrm{P}(h_{\mathrm{MH}}|\mathrm{C}_1)^{st} = 98.12\%, \qquad (3.27)$$

which again shows that a sequence of five heads heuristically (but mistakenly) suggests that the coin is the one that mostly comes up heads. Clearly, in this case the perceived posterior probability depends on the parameter θ, as will be discussed in detail in section 3.6.3.

Table 3.1 presents the outcomes from all the inference models reviewed in section 3.4, including Grether's, i.e. Eq. (3.11), evaluated with $\alpha = 0$, $\beta_1 = 0.8$ and $\beta_2 = 0.2$. It is evident that all the heuristic inference models agree on judging most likely that the coin is the one that mostly comes up heads h_{MH}, in contrast to the rational conclusion inferred through Bayes' theorem.

In conclusion, with this numerical example we have clarified the substantial difference between likelihood and representativeness. We have also shown how the representativeness bias may alter the posterior judgment to the point of suggesting conclusions opposite to those consistent with rational inference.

Table 3.1. Achieved results for each model.

	Likelihood $\mathrm{P}(\mathrm{C}_i\|h_j)$ or Representativeness $\mathrm{R}(\mathrm{C}_i\|h_j)$	Posterior probability $\mathrm{P}(h_j\|\mathrm{C}_i)$	Posterior odds $\mathrm{P}(h_j\|\mathrm{C}_i)/\mathrm{P}(-h_j\|\mathrm{C}_i)$
Bayes	$\mathrm{P}(\mathrm{C}_1\|h_{\mathrm{FC}}) = 0.0313$ $\mathrm{P}(\mathrm{C}_2\|h_{\mathrm{FC}}) = 0.0313$ $\mathrm{P}(\mathrm{C}_1\|h_{\mathrm{MH}}) = 0.4437$ $\mathrm{P}(\mathrm{C}_2\|h_{\mathrm{MH}}) = 0.0024$	$\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1) = 57.27\%$ $\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1) = 42.73\%$	$\frac{\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)}{\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)} = 1.34$
Vanilla model (Eq. (3.4) and Eq. (3.7))	$\mathrm{R}(\mathrm{C}_1\|h_{\mathrm{FC}}) = 0.07$ $\mathrm{R}(\mathrm{C}_2\|h_{\mathrm{FC}}) = 13.04$ $\mathrm{R}(\mathrm{C}_1\|h_{\mathrm{MH}}) = 14.18$ $\mathrm{R}(\mathrm{C}_2\|h_{\mathrm{MH}}) = 0.08$	$\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)^{st} = 0.49\%$ $\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)^{st} = 99.51\%$	$\frac{\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)^{st}}{\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)^{st}} = 0.01$
Bordalo et al. (θ=0.8)	$\mathrm{R}(\mathrm{C}_1\|h_{\mathrm{FC}}) = 0.07$ $\mathrm{R}(\mathrm{C}_2\|h_{\mathrm{FC}}) = 13.04$ $\mathrm{R}(\mathrm{C}_1\|h_{\mathrm{MH}}) = 14.18$ $\mathrm{R}(\mathrm{C}_2\|h_{\mathrm{MH}}) = 0.08$	$\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)^{st} = 1.88\%$ $\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)^{st} = 98.12\%$	$\frac{\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)^{st}}{\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)^{st}} = 0.02$
Grether (α=0; β_1=0.8; β_2=0.2)	$\mathrm{R}(\mathrm{C}_1\|h_{\mathrm{FC}}) = -2.65$ $\mathrm{R}(\mathrm{C}_2\|h_{\mathrm{FC}}) = 2.57$ $\mathrm{R}(\mathrm{C}_1\|h_{\mathrm{MH}}) = 2.65$ $\mathrm{R}(\mathrm{C}_2\|h_{\mathrm{MH}}) = -2.57$	$\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)^{st} = 18.05\%$ $\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)^{st} = 81.95\%$	$\frac{\mathrm{P}(h_{\mathrm{FC}}\|\mathrm{C}_1)^{st}}{\mathrm{P}(h_{\mathrm{MH}}\|\mathrm{C}_1)^{st}} = 0.22$

3.6 Case study

In this section we wish to verify whether the judgment models reviewed in section 3.4 are suitable to describe the typical confusion between condition state and safety of a bridge frequently observed in bridge management. As described in section 3.1, bridge managers often tend to delay retrofit of substandard bridges which do not show sign of deterioration, while repair promptly deteriorated bridges as soon as the damage is observed, regardless their actual residual load-carrying capacity. We have already observed that the biased rationale behind this apparent behaviour is that undamaged bridges ‘look’ safe, while damaged bridges ‘look’ unsafe, simply because, generally speaking, we know that deterioration negatively affects safety.

We discuss this bias with reference to one of the case studies reported in (Zonta, et al., 2007), i.e. the SP65 bridge on the Maso River, which is operated by the Autonomous Province of Trento (APT). The bridge, shown in Figure 3.2(a), is a common type of bridge in the APT stock. The structure has two simple spans of 19.0 m and 22.0 m, and a total length of 43.0 m. Each span has four girders spaced at 2.1 m, 2.4 m and 2.1 m respectively. The cross-section of the girders is shown in Figure 3.2(b). The deck slab consists of 22–27 cm of reinforced concrete and a 15 cm surface layer of asphalt. The roadway width is 7 m with 0.70 m pedestrian pavements and hand railing on each side.

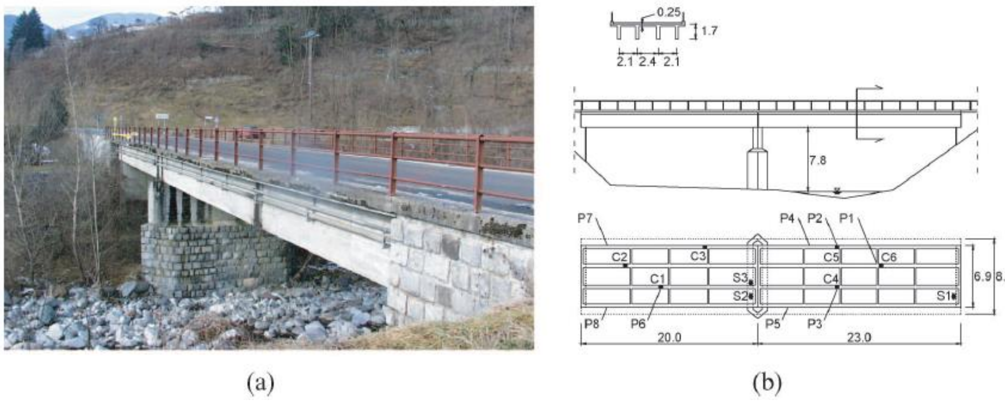


Figure 3.2. SP65 bridge on Maso River: (a) overview; (b) plan view, elevation and cross section of the deck (Zonta, et al., 2007).

Managing its bridges, APT uses an inventory model and condition state appraisal system consistent with the AASHTO (1997) Commonly Recognized (CoRe) Standard Element System (American Ass. State Highway and Transportation Off, 1997). The CoRe element standard has been adopted since 1995 by FHWA and AASHTO as broadly accepted way to represent bridges condition on a uniform scale (American Ass. State Highway and Transportation Off, 1997). The CoRe element standard inventories a bridge into a set of Standard Elements (SE), each specified in term of quantity (surface, length or number). For example, the bridge deck of the SP65 bridge includes the following SE: slab, beam, pavement, sidewalk, guard rail and railing.

The state of deterioration of each element is appraised through routine visual inspections. The inspector classifies the state of deterioration of an element choosing among five possible deterioration levels, called Condition States (CS), specified, for each element type, in the inspector manual. Table 3.2 reports, as an example, the definition of the five CS of a concrete slab, or CoRe standard element #12, as reported in APT inspection manual available from the website of the APT (Autonomous Province of Trento , 2018). As a general rule, Condition State 1 (CS_1) always means 'as per design', or 'no deterioration', while CS_5 corresponds to the most severe observable deterioration state.

Table 3.2. SE #12 concrete slab: state description for each Condition State (CS).

CS	State description of the slab surface
1	No delamination, spalling or water infiltration.
2	Possible delamination, spalling or water infiltration. Possible segregation and consequently reinforcement exposure.
3	Previously repaired or subjected to delamination or spalling. Segregation and consequently reinforcement exposure. Limited water infiltration.
4	Extended parts previously repaired or subject to delamination or spalling; deep segregation phenomena with extended exposure of reinforcement. Extended water infiltration.
5	Deep deterioration or anomalies. Reinforcement corrosion and cross-section loss require a deep analysis to verify the structural safety of the element.

While the deterioration condition is apprised through visual inspection, its safety level is evaluated separately, through a five-step formal assessment procedure (Zonta, et al., 2007), whose ultimate objective is to calculate the bridge reliability index β. We have already observed in the Introduction that condition state and safety are obviously correlated, but not the same thing, and that we can well have a severely deteriorated bridge which is perfectly safe or an intact bridge which is not safe. We have also noticed that a rational bridge manager should address safety above all, while in practice the intervention priority is often biased by the apparent state of deterioration of the bridges, regardless their actual residual load-carrying capacity.

In this section, we want to numerically analyse and describe the following case:

- As far as its safety is considered, the bridge could be in two possible states: SAFE (h_S) or FAIL (h_F). SAFE means that, following to a formal safety assessment carried out by an expert structural engineer, the bridge load-carrying capacity is judged sufficient for the bridge to operate without restrictions. On the other hand, FAIL means that the bridge is not found to have sufficient load-carrying capacity and should be closed to traffic.
- Based on a frequentist analysis of the load-carrying capacity formally assessed for similar bridges of the same type and age, it is estimated that only one bridge out of one thousand is found to be in the FAIL state. We formalize this information assuming prior base rates $P(h_F) = 0.001$ for the state hypothesis FAIL, and therefore $P(h_S) = 0.999$ for the state hypothesis SAFE.
- Based on the last visual inspection, the bridge exhibits no or minimal deterioration, except for the concrete slab, which is classified in the most severe condition state, or CS_5.
- Based on the condition state assessed via visual inspection, the bridge manager judges the bridge in FAIL state.

This case study effectively describes a prototypical situation where the bridge manager judges the state of safety of the bridge based on the condition state of one of its elements, and disregarding any information on its actual residual load-carrying capacity. The manager implicitly assumes that a severe deterioration of an element automatically implies that the bridge load-carrying capacity is insufficient, simply because deterioration is representative for a reduced capacity. We hypothesize this situation could be described as a case of the representativeness bias, where the safety is improperly judged based on how much deterioration is representative of loss in capacity.

In order to verify this conjecture, we will answer quantitatively to the following questions:

(1) What is the likelihood $\mathrm{P}(\mathrm{CS}_5|h_\mathrm{F})$ of an unsafe bridge to be in CS_5?
(2) How much CS_5 is representative of a bridge in FAIL state?
(3) What is the proper posterior probability of this bridge to be in FAIL state?
(4) How does representativeness bias distort the manager judgment as to the bridge safety?

3.6.1 Likelihood and representativeness

To start, we have to define a proper likelihood distribution for each hypothesis, i.e. $\mathrm{P}(\mathrm{CS}_i|h_\mathrm{F})$ and $\mathrm{P}(\mathrm{CS}_i|h_\mathrm{S})$. In the following, the procedure used for the definition of the likelihood is the same as in (Zonta, et al., 2007).

According to (Melchers, 1999), we employ II level probabilistic methods, which allows to calculate the reliability index $\beta = -\Phi^{-1}(P_{h_\mathrm{F}})$, where Φ is the cumulative normal distribution function. Two stochastic variables are considered: the loads effect S and the starting resistance R_0 of the bridge, both supposed to be Normal distributions (Norm), with their mean μ and coefficient of variation CoV. In formula:

$$f_{R_0}(r) = \mathrm{Norm}(r, \mu_{R_0}, CoV_{R_0}), \qquad f_S(s) = \mathrm{Norm}(s, \mu_S, CoV_S). \qquad (3.28a,b)$$

Because of the prioritization approach, we assume that the structure will not maintain its mechanical characteristics in the years, i.e. we have to take into accounts the deterioration of construction material through the following probabilistic degradation model (Zonta, et al., 2007):

$$R = R_0(1 - \delta(\mathrm{CS}_i)), \qquad (3.29)$$

where $\delta(\mathrm{CS}_i)$ is a probabilistic capacity degradation function, depending only on the CS_i of the SE that controls the capacity of the structural unit at the limit state. Its density function δ_i is the probability density function of the loss in capacity when the element is in the *i*-th CS. Reminding that the elements are rated based on visual inspections, δ_i represents the likelihood of a certain loss in capacity when the element has been rated into the *i*-th reference state. Typically, low values of CS, i.e. CS_1 CS_2 and CS_3, are not associated with any loss of capacity: in this case δ_i coincides with a Dirac delta function and therefore $R = R_0$. On the other hand, higher CS_i are associated with distributions that reflect the uncertainty of the system in correlating the actual loss in capacity, with the verbal description of the reference state proposed by the inspection manual. CS_4 is associated with a uniform distribution δ_4 of loss in capacity, for values of δ included in [0, 5%]. In the same way, the system associates the reference state 5 with a triangular distribution, for values of δ included in [5%, 70%], as Figure 3.3 shows.

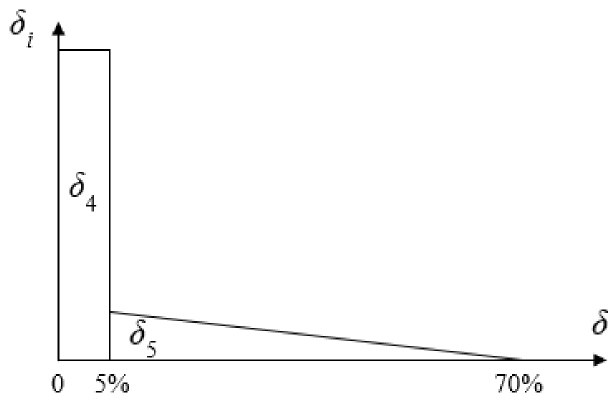


Figure 3.3. Capacity degradation function $\delta(\mathrm{CS}_i)$.

Because most of the information required to define the distribution of capacity R and actions S are not explicitly contained in the system database, therefore a simplified approach must be adopted. It is convenient to define a normalized capacity $r = R/\mu_S$, with mean value $\mu_r = \mu_{R_0}/\mu_S$, equal to the central safety factor γ_0, associated with the limit state Z, and a normalized demand $s = S/\mu_S$ with mean value $\mu_S = 1$. The coefficients of variations of the normalized variables γ and s are equal to those of R and S. The Normal distribution of the capacity and actions become:

$$f_{\gamma_0}(r) = \text{Norm}(r, \gamma_0, CoV_R), \qquad f_S(s) = \text{Norm}(s, 1, CoV_S), \tag{3.30a,b}$$

where the reliability index is related to the central safety factor γ_0 through the expression:

$$\beta = \frac{\gamma_0 - 1}{\sqrt{{CoV_R}^2 \cdot {\gamma_0}^2 + {CoV_S}^2}}. \tag{3.31}$$

Finally, the normalized limit state function is $z = r - s$, and the probability of failure P_{h_F} associated with the limit state Z coincides with that of z:

$$P_{h_F}(\text{CS}_i) = P(Z < 0) = P(z < 0). \tag{3.32}$$

According to Eurocode 0, if we employ II level probabilistic methods, the target reliability index β for Class RC2 structural member in the Ultimate Limit State and with a reference time of 1 year is equal to $\beta = 4.75$. Assuming $V_R = 0.05$ and $V_S = 0.10$, from Eq. (3.31) we can obtain $\gamma_0 = 1.96$. Once we know γ_0, the probability of failure $P_{h_\text{F}}(\text{CS}_i)$ is then calculated through Monte Carlo by computing the cumulative-time failure probability of the normalized limit state z, by using a normalized Gaussian distribution for the demand $f_S(r)$ and a normalized non-Gaussian distribution for the reduced capacity $r = \gamma_0(1\text{-}f_\delta)$, which depends on CS_i:

$$f_r(r, \text{CS}_i) = f_{\gamma_0}(r)\big(1 - f_\delta\big). \tag{3.33}$$

Consequently, we obtain the following failure probabilities for each CS_i:

$$\begin{gathered}\big[P_{h_\text{F}}(\text{CS}_1);\ P_{h_\text{F}}(\text{CS}_2);\ P_{h_\text{F}}(\text{CS}_3);\ P_{h_\text{F}}(\text{CS}_4);\ P_{h_\text{F}}(\text{CS}_5)\big] = \\ [6.12 \cdot 10^{-5};\ 2.68 \cdot 10^{-6};\ 6.47 \cdot 10^{-6};\ 6.61 \cdot 10^{-4};\ 2.04 \cdot 10^{-1}].\end{gathered} \tag{3.34}$$

Assuming the following a priori distributions for CS, i.e. $\text{P(CS)} =$ [50%, 20%, 15%, 10%, 5%], we can calculate the probability for each hypothesis, i.e. "FAIL = h_F" and "SAFE = h_S" respectively:

$$P_{h_{\mathrm{F}}} = \sum_{i=1}^{\mathrm{CS}_{max}=5} P_{h_{\mathrm{F}}}(\mathrm{CS}_i) \cdot P(\mathrm{CS}_i) = 0.0103, \tag{3.35}$$

$$P_{h_{\mathrm{S}}} = 1 - P_{h_{\mathrm{F}}} = 0.9897. \tag{3.36}$$

Then, according to Bayes' rule, for both hypothesis "S = SAFE" and "F = FAIL" we can evaluate the relative likelihood distributions for each Condition State CS_i, as follows:

$$\begin{aligned} \mathrm{P}(\mathrm{CS}_i|h_{\mathrm{F}}) &= \frac{P_{h_{\mathrm{F}}}(\mathrm{CS}_i) \cdot P(\mathrm{CS}_i)}{P_{h_{\mathrm{F}}}}, \\ \mathrm{P}(\mathrm{CS}_i|h_{\mathrm{S}}) &= \frac{(1 - P_{h_{\mathrm{F}}}(\mathrm{CS}_i)) \cdot P(\mathrm{CS}_i)}{P_{h_{\mathrm{S}}}}. \end{aligned} \tag{3.37a,b}$$

Figure 3.4 shows the results for each CS_i.

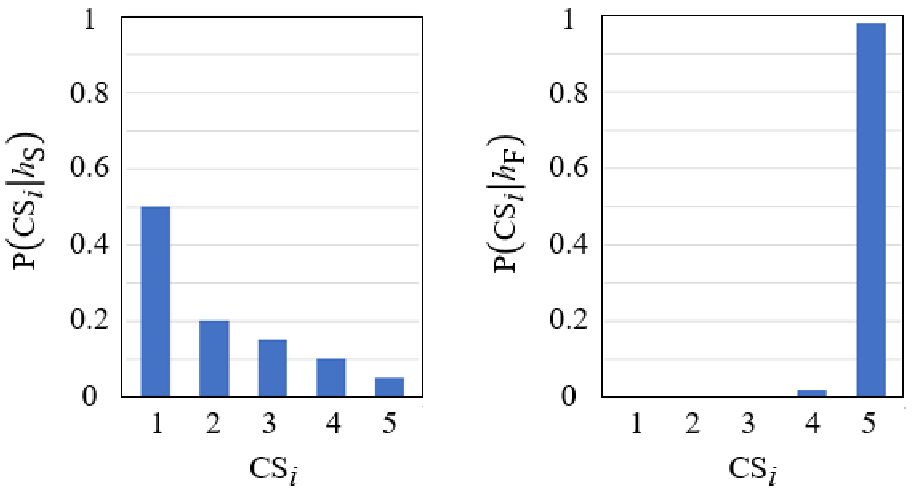


Figure 3.4. Likelihood distributions for each state hypothesis.

To be consistent with the outcomes of Figure 3.4, we choose the following likelihood distributions:

$$\begin{aligned} \mathrm{P}(\mathrm{CS}_i|h_{\mathrm{S}}) &= [50\%, 20\%, 15\%, 10\%, 5\%], \\ \mathrm{P}(\mathrm{CS}_i|h_{\mathrm{F}}) &= [0, 0, 0, 2\%, 98\%]. \end{aligned} \tag{3.38a,b}$$

After the evaluation of the likelihoods, we are interested in understanding how much CS_5 is representative of the bridge in FAIL state. We can calculate it according to Eq. (3.4):

$$\mathrm{R}(\mathrm{CS}_5|h_\mathrm{F}) = 19.6, \qquad \mathrm{R}(\mathrm{CS}_5|h_\mathrm{S}) = 0.05. \tag{3.39a,b}$$

These outcomes show that, as expected, CS_5 is very representative of the failure state of the bridge, with an enormous difference in comparison to the safe state of the bridge, i.e. $\mathrm{R}(\mathrm{CS}_5|h_\mathrm{F}) \gg \mathrm{R}(\mathrm{CS}_5|h_\mathrm{S})$: this is very important because we have learnt that this can be the reason of a distorted final judgment.

3.6.2 Posterior judgment

Let's now evaluate the posterior judgment of the manager, in the case that the bridge is classified in CS_5. The proper posterior probabilities, computed using the rational framework provided by Bayes' theorem, results:

$$\mathrm{P}(h_\mathrm{F}|\mathrm{CS}_5) = 1.92\% < \mathrm{P}(h_\mathrm{S}|\mathrm{CS}_5) = 98.08\%. \tag{3.40}$$

This means that rational managers, in line with Bayes' rule and after observing CS_5, would judge the possibility that the bridge could be in the FAIL state as very unlikely.

However, we have introduced before that, based on the condition state assessed via visual inspection, the bridge manager has judged the bridge in FAIL state. It is possible to explain this judgment by evaluating the distorted posterior probability. Using the vanilla inference model of Eq. (3.7), we achieve:

$$\mathrm{P}(h_\mathrm{F}|\mathrm{CS}_5)^{st} = 99.75\% > \mathrm{P}(h_\mathrm{S}|\mathrm{CS}_5)^{st} = 0.25\%. \tag{3.41}$$

Similarly, accepting the inference model of Bordalo et al., the distorted posterior probability is:

$$\mathrm{P}(h_\mathrm{F}|\mathrm{CS}_5)^{st} = 69.97\% > \mathrm{P}(h_\mathrm{S}|\mathrm{CS}_5)^{st} = 30.03\%. \tag{3.42}$$

In both cases the failure state turns out to be the most likely, and this outcome allows to explain the judgment of the manager, which is biased since CS_5 is very representative of a fault bridge.

Table 3.3 reports all the achieved results; the last row of the table presents again the results that come from the inference model of Grether, which agree with those obtained with the other biased models, i.e. the FAIL state is the most likely, in contrast to the rational conclusion inferred through Bayes' theorem.

In summary, we have demonstrated that when an inspector judges the safety state of a bridge by only accounting for the observed condition state CS, they are biased by representativeness: in their posterior judgments they tend to neglect the prior probability of the failure condition, which is typically very low, $\mathrm{P}(h_{\mathrm{F}}) = 0.001$ in this specific case study, and to weight too much the ratio between the likelihood of the observations, which actually is the representativeness itself. Therefore, their final judgment results distorted in comparison to the one achieved by rational managers who follow Bayes' theorem.

Table 3.3. Achieved results for each model.

	Likelihood $\mathrm{P}(\mathrm{C}_i \mid h_j)$ or Representativeness $\mathrm{R}(\mathrm{C}_i \mid h_j)$	Posterior probability $\mathrm{P}(h_j \mid \mathrm{C}_i)$	Posterior odds $\mathrm{P}(h_j \mid \mathrm{C}_i)/\mathrm{P}(-h_j \mid \mathrm{C}_i)$
Bayes	$\mathrm{P}(\mathrm{CS}_5 \mid h_{\mathrm{F}}) = 0.98$ $\mathrm{P}(\mathrm{CS}_5 \mid h_{\mathrm{S}}) = 0.05$	$\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5) = 1.92\%$ $\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5) = 98.08\%$	$\frac{\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)}{\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)} = 0.02$
Vanilla model (Eq. (3.4) and Eq. (3.7))	$\mathrm{R}(\mathrm{CS}_5 \mid h_{\mathrm{F}}) = 19.6$ $\mathrm{R}(\mathrm{CS}_5 \mid h_{\mathrm{S}}) = 0.05$	$\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)^{st} = 99.75\%$ $\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)^{st} = 0.25\%$	$\frac{\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)^{st}}{\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)^{st}} = 399$
Bordalo et al. (θ=0.8)	$\mathrm{R}(\mathrm{CS}_5 \mid h_{\mathrm{F}}) = 19.6$ $\mathrm{R}(\mathrm{CS}_5 \mid h_{\mathrm{S}}) = 0.05$	$\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)^{st} = 69.97\%$ $\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)^{st} = 30.03\%$	$\frac{\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)^{st}}{\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)^{st}} = 2.33$
Grether (α=0; β_1=0.8; β_2=0.2)	$\mathrm{R}(\mathrm{CS}_5 \mid h_{\mathrm{F}}) = 2.98$ $\mathrm{R}(\mathrm{CS}_5 \mid h_{\mathrm{S}}) = -2.98$	$\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)^{st} = 73.19\%$ $\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)^{st} = 26.81\%$	$\frac{\mathrm{P}(h_{\mathrm{F}} \mid \mathrm{CS}_5)^{st}}{\mathrm{P}(h_{\mathrm{S}} \mid \mathrm{CS}_5)^{st}} = 2.73$

3.6.3 Discussion about inference models

To develop the numerical calculations in the previous sections, we had to suppose some specific values for the subjective parameters of the inference models introduced in section 3.4.2, i.e. $\theta = 0.8$, $\alpha = 0$, $\beta_1 = 0.8$, $\beta_2 = 0.2$: these values correspond to a high level of the representativeness heuristic since they maximize the importance of R and minimise the contribute of the prior information. Since these parameters depend on the different behaviour of the people and could vary considerably, it is interesting to develop a sensitivity analysis in order to understand how they affect the model and then the final results.

Let's take for instance the model of Bordalo et al.: as we can see from Eq. (3.8), it depends only on one subjective parameter, i.e. $\theta \geq 0$. Figure (3.5) shows how the posterior failure probability of the bridge, after observing CS_5, varies according to θ: even if θ can be also bigger than 1, we study just the interval $0 \leq \theta \leq 1$ since this is sufficient to understand how the results change. While with our previous assumption of $\theta = 0.8$ the result is $P(h_F|CS_5) = 69.97\%$, we can notice that the outcome is highly sensitive to the choice of θ: it changes from $P(h_F|CS_5) = 1.92\%$ if $\theta = 0$, i.e. in line with a rational manager who follows Bayes' rule, to $P(h_F|CS_5) = 88.49\%$ if $\theta = 1$, i.e. in line with an irrational manager biased with a high level of the representativeness heuristic. Furthermore, we can observe that the posterior failure probability $P(h_F|CS_5)$ is bigger than the posterior safe probability $P(h_S|CS_5)$ when $\theta > 0.67$. These results allow to demonstrate the truthfulness of this inference model, and explain the importance to calibrate properly the subjective parameters according to the specific inspector. The same generic conclusions can be extended to the model of Grether, since we have demonstrated that it is based on the same mathematical formulation.

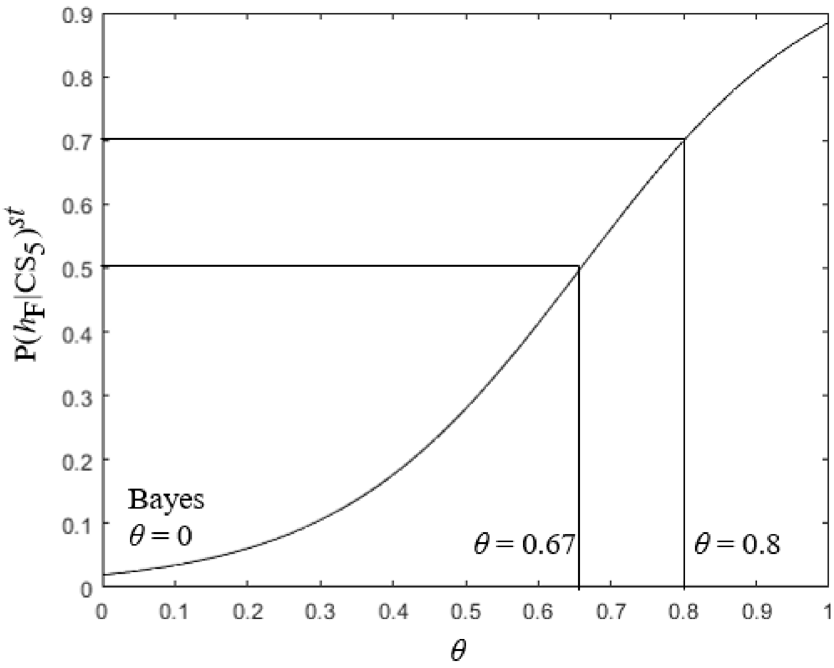


Figure 3.5. How the distorted posterior probability $P(h_F|CS_5)^{st}$ varies according to the subjective parameter θ.

Conversely, the vanilla model introduced in Eq. (3.7) is less sophisticated because it does not depend on a subjective parameter. Even if this may seem like a shortcoming, the results obtained in both section 3.5 and section 3.6 demonstrate the correctness of the vanilla model in reproducing the distorted judgment based on the representativeness bias. In detail, it is evident that its outcomes are very similar to those that can be obtained assuming the maximum level of representativeness in the subjective parameters of the other inference models, meaning that the vanilla model reproduces the behaviour of an inspector completely biased by this heuristic. This conclusion is consistent with the mathematical formulation of the model itself, since it overlooks the contribution of the prior and it completely replaces the likelihood with the representativeness.

3.7 Conclusions

Judging the state of a bridge based on SHM observations is an inference process which should be rationally carried out using Bayesian logic. However, we often observe that real-life decision makers depart from this ideal model of rationality, judge and decide using common sense, and privilege fast and frugal heuristics to rational analytic thinking. For instance, confusion between condition state and safety of a bridge is one of the most frequently observed examples in bridge management. In this contribution, we have demonstrated that this bias can be described by Kahneman and Tversky's representativeness heuristic.

A review of the technical literature shows that representativeness heuristic has been widely analysed from a descriptive point of view, while only few models have been proposed to describe this bias from a mathematical perspective. In the literature there is a general agreement on that the degree of representativeness of an observable class for a reference hypothesis is in some way related to odds of observable quantities. Instead, there is not a general agreement on how the standard Bayes' rule, which reflects the judgment of a rational thinker, should be adjusted to consider representativeness. Most authors do not provide an explicit expression for the distorted posterior, but understand the statement that ordering hypotheses by perceived probabilities follows representativeness rather than Bayesian posterior. This is consistent with a distorted judgement model, here referred to as 'vanilla', whereby (i) representativeness is used instead of likelihood and (ii) the prior information is neglected. Bordalo et al. and Grether provide more refined models for reproducing the subjective distorted judgement, which allow to blend more flexibly likelihood, representativeness and prior information, through a number of subjective parameters, in order to better reproduce the distorted perception of a particular subject.

We have first applied these mathematical models to a classical literature representativeness problem, to better appreciate the difference among the various formulations of representativeness and heuristic judgement models. Next, we have applied the same models to the case of a transportation manager who wrongly judges a particular bridge unsafe simply because deteriorated, regardless its actual residual load-carrying capacity. Their judgment is biased due to the apparent behaviour that damaged bridges 'look' unsafe, in contrast with undamaged bridges which 'look' safe.

In the particular case study, we have demonstrated that Bayes' theorem correctly identifies the bridge as safe, while application of the three judgment models analysed (vanilla, Bordalo et al.'s and Grether's) all predict the manager will mistakenly judge the bridge as unsafe based on the observed condition state. Given the simplicity of the case study, which is essentially a two hypotheses inference problem where the individual distorted behaviour is characterized by the ordering of the two hypotheses by subjective probabilities, the three models are equivalent in this particular instance, as they reproduce equally well the observed distorted perception. The main difference between these three inference models is that 'vanilla' model reproduces the behaviour of an individual whose judgement is blatantly driven by representativeness, while the other two models allow to describe more subtle forms of distorted judgment, whose limit cases are rational Bayesian inference on one side and the vanilla representativeness bias on the other. The three models may not be equivalent in a more complex setting, where the vanilla inference model may fail to reproduce the observed representativeness bias. Bordalo et al.'s and Grether's model are clearly more flexible, but at the same time very sensitive to a number of subjective parameters, which have to be accurately calibrated, typically with cognitive tests, on the particular individual whose distorted judgment is to be described.

To conclude, to clear the Reader's mind from any possible equivocal interpretation, we reiterate once again that the only rational way to judge under uncertainties is to use Bayesian logic, and here we are not suggesting in any way that representativeness should be used instead of Bayes' theorem. At the same time, predicting the actual behavior of managers is required when setting a general policy for bridge maintenance, acknowledging that the managers who are going to enact the policy may behave irrationally.

4. An elicitation process to quantify Bayesian networks for dam failure analysis

Submitted for publication to *Canadian Journal of Civil Engineering*
(Under revision)

Andrea Verzobio[a], Ahmed El-Awady[b], Kumaraswamy Ponnambalam[c], John Quigley[d], Daniele Zonta[e]

Position and affiliation:
[a]PhD Student, Department of Civil and Environmental Engineering, University of Strathclyde, UK
[b]Postdoc, Department of Environmental Engineering Science, Ain Shams University, Egypt
[c]Professor, Department Systems Design Engineering, University of Waterloo, Canada
[d]Professor, Department Management Science, University of Strathclyde, UK
[e]Professor, Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy

Summary of the paper

This paper proposes a process to elicit engineering expert knowledge with the specific aim of quantifying a Bayesian Network, while minimizing the adverse impact of biases to which judgment is commonly subjected. In the development of the methodology, each stage of the process is proposed by highlighting all the potential biases that may influence the process, such as anchoring and availability (introduced respectively in section 2.2.3.2 and 2.2.3.3), as well as by proposing appropriate actions in order to minimize the risk of biased judgments. The developed elicitation process is applied to a real-life case study regarding the safety of the Mountain Chute Dam and Generating Station (Ontario, Canada). This contribution provides a demonstration of the usefulness of eliciting engineering expertise with regard to system reliability analysis.

4.1 Introduction

Dams fail due to a combination of more frequent load and reduced resistance to the load exceeding the facility's capacity, design problems, unexpected flood events or inappropriate decisions in managing dams. Such failures, including breaches, may lead to catastrophic events which affect both properties and lives of people. Maintaining dams is challenging, as resources are limited, facilities are remote and usage profiles are uncertain. Global weather patterns have been changing, causing periods of flooding, which have resulted in an increase in operating the dams. Understanding and anticipating the environment in which the dams will operate is vital for maintaining the availability of the asset. Effectively maintaining the asset requires a mathematical model to explicate the relationship between environment, usage, hazards and management decisions, and to support the optimal long-term productivity of the asset.

While several examples of mathematical and probabilistic approaches used to evaluate the safety of dams can be found in the literature (Yanmaz & Gunindi, 2008) (Li, et al., 2011) (Goodarzi, et al., 2012) (Su, et al., 2015), in this contribution we decide to use the Bayesian Network (BN) since it has many advantages and it is an increasingly popular method for reasoning under uncertainty and modelling uncertain domains. For instance, in comparison with two most commonly used approaches, i.e. the Event Tree Analysis (ETA) and the Fault Tree Analysis (FTA), BNs can more succinctly represent the dependency relationship between a large number of variables, permit variables to be described in multiple states not just binary, i.e. true or false, describe and represent multiple initiating events, and explicitly integrate different types of data, e.g. technical, environmental and social, in a single unified representation. Comparisons between BN and ETA or FTA in safety analysis can be found in (Khakzad, et al., 2011) (Jong & Leu, 2013) (Zerrouki & Tamrabet, 2015a) (Zerruki & Tamrabet, 2015b).

BNs provide a powerful framework for reasoning under uncertainty, and consequently have been recently applied to various engineering problems, e.g. earthquake risk management (Bayraktarli, et al., 2005) (Bensi, et al., 2011) (Liu & Nadim, 2013), avalanche risk assessment (Gret-Regamey & Straub, 2006), landslide hazard mitigation (Medina-Cetina & Nadim, 2008), reliability analysis (Langsetha &

Portinaleb, 2007), climate change assessment (Peter, et al., 2009), risk assessment in maritime engineering (Kelangath, et al., 2011), environmental modelling and management (Aguilera, et al., 2011), risk assessment for fatigue damage (Sankararaman, et al., 2011) (Ling & Mahadevan, 2012), scour management (Maroni, et al., 2019). In addition, as regards the topic of this paper, in the literature we can find many papers in which BNs are used to develop dam safety analysis, among the many we recommend (Smith, 2006) (Xu, et al., 2011) (Zhang, et al., 2011) (Miroslaw-Swiatek, et al., 2012) (Peng & Zhang, 2013) (Ahmadi, et al., 2015) (Gang, et al., 2016) (Eldosouky, et al., 2017) (Liu, et al., 2017) (Briseno-Ramiro, et al., 2019) (Dassanavake & Mousa, 2020).

Specifically, BNs are probabilistic graphical models that use directed acyclic graph to represent a set of uncertain variables and their conditional dependencies (Charniak, 1991) (Ben Gal, 2007) (Jensen & Nielsen, 2007). In detail, nodes represent the collection of random variables, while edges represent the interrelationship between these variables. While the topology of the BN provides the causal structuring of the problem under study, the quantitative strength of the interrelationships among variables is measured using conditional probability distributions, which can be updated when new data become available. Typically, the quantification of the probabilities may be obtained from statistical and historical data, existing physical or empirical models and logic inference. However, these quantification sources and methodologies are often not easy to be conducted and not sufficient to quantify the entire BN, due to the lack of sufficient models that interpret the interrelationships among system variables and due to the lack of data and information. Consequently, we decide to rely on expert judgments to quantify these dependencies: engineering knowledge and experience can be an important data source for estimating these probabilities (Dias, et al., 2018).

Eliciting expert judgment in the form of subjective probabilities is a socio-technical activity. As such it requires a structured and facilitated process to extract meaningful judgments because people, even experts, are unable to provide accurate and reliable data simply on request (Ferrell, 1994) (Vick, 2002). An example about discrepancies between experts in risk assessment can be found in (Rizak & Hrudey, 2005). In addition, since the work of Tversky and Kahneman in the early 1970s (Tversky & Kahneman, 1974), there has been awareness of the biases and heuristics people apply in decision-making under uncertainty that can result in poor probability assessments. Elicitation processes are designed to minimize the influence of these biases (Quigley & Walls, 2020). In the literature, there are a variety of existing processes for eliciting expert knowledge with engineering applications, see for instance (Bubniz, et al., 1998), (Hodge, et al., 2001) and (Astfalck, et al., 2018). Textbooks such as (Cooke, 1991), (Meyer & Booker, 1991) and (Dias, et al., 2018) are references for general aspects of elicitation. However, very little has been reported about elicitation processes aimed specifically at quantifying BNs using expert judgment (Sigurdsson, et al., 2001) (Norrington, et al., 2008) (Christophersen, et al., 2018), especially for civil engineering applications, where we require experts to assess a variety of dependent variables, each of which is in one of several possible states.

In this paper, the aim is to develop a methodology for eliciting expert knowledge in the specific case where the model is described by a BN. We start with an introduction of the fundamentals of BNs in section 4.2. In section 4.3, a four-stage structured elicitation process is developed generically so that it can be applied to many civil engineering structures, e.g. dams and bridges. Section 4.4 presents an implementation of this methodology, with its application to a real-life case study regarding the safety of the Mountain Chute Dam and Generating Station, which is situated on the Madawaska River in Ontario, Canada. Concluding remarks, along with the explanation of the lessons learnt from the application, are presented at the end of the paper.

4.2 Bayesian network

Bayesian Networks (BNs), also known as Bayes networks, belief networks or decision networks, are probabilistic graphical models used to represent knowledge about an uncertain domain using a combination of principles from graph theory, probability theory, computer science, and statistics (Charniak, 1991) (Ben Gal, 2007) (Jensen & Nielsen, 2007). In the graph, nodes represent the collection of random variables, while edges represent the interrelationship between these variables. In addition, each node is associated with conditional probability values that model the uncertain relationship between the node and its parents. BNs can model the quantitative strength of the interrelationships among variables, i.e. the nodes, allowing their probabilities to be updated using any new available data and information. They are mathematically rigorous, understandable, and efficient in computing joint probability distribution over a set of random variables, and consequently very useful in supporting risk analysis of complex systems.

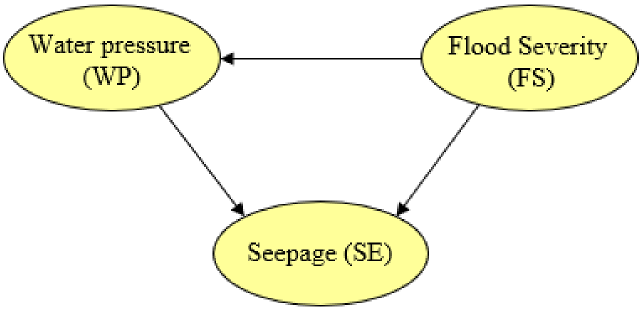
BNs are probabilistic graphical models that use directed acyclic graph (DAG): this means that a set of directed edges are used to connect the set of nodes, where these edges represent direct statistical dependencies among variables, with the constraint of not having any directed cycles. Let $X = (X_1, \dots, X_i, \dots, X_n)$ represent the set of nodes, i.e. the uncertain variables. A node X_j is called parent of a child node X_i if there is a directed edge from node X_j to node X_i, meaning that X_i depends on X_j. Each node can have many parents nodes, while nodes with no parent are called root nodes and nodes with no child are called leaf nodes. In addition, each root node is associated with a basic probability table (BPT), while each child node with a conditional probability table (CPT). The joint probability function of random variables in a BN can be expressed as follows:

$$P(X) = \prod_{i=1}^{n} P[X_i \mid Pa(X_i)], \tag{4.1}$$

where $P(X)$ is the joint probability and $Pa(X_i)$ is the parent set of node X_i. If X_i has no parents, i.e. it is a root node, then the function reduces to the unconditional probability of $P(X_i)$. A simple example of BN with three variables as regards dam safety analysis is shown in Figure 4.1: both the severity of the flood and a high-water pressure can cause the presence of seepage in the dam; in addition, the flood severity has a direct effect on the level of water pressure. The table related to the flood severity, that is a root node, represent an example of BPT, while the tables of the other two child nodes are examples of CPT.

Generically, in BNs there are two main types of reasoning: predictive reasoning, i.e. top-down or forward reasoning, in which evidence nodes are connected through parent nodes (cause to effect), and diagnostic reasoning, i.e. bottom-up or backward reasoning, in which evidence nodes are connected through child nodes (effect to cause).

	WP	
FS	Low	High
Low	0.9	0.1
High	0.6	0.4



FS	
Low	High
0.9	0.1

		SE	
WP	FS	Yes	No
Low	Low	0.1	0.9
High	Low	0.5	0.5
Low	High	0.3	0.7
High	High	0.7	0.3

Figure 4.1. An example of BN with three variables.

Finally, we can summarize how to build and use a BN with three steps: structuring the problem, defining the conditional probabilities, and making the final inference. The first step aims to define the topology of the BN: first, the relevant variables of the problem are identified and expressed as statistical variables, discrete or continuous; then, the network is created by joining the variables according to their dependency. The second step is about quantifying the interrelationship among connected nodes, i.e. defining the CPTs, as well as the BPTs in the case of root nodes. They may be obtained from statistical and historical data, existing physical or empirical models, logic inference or they may be elicited from experts. Lastly, the inference step concerns entering the evidence in the BN, updating the probabilities, and interpreting the final results.

4.3 Elicitation Process for Bayesian networks

In this paper, the aim is to support the collection of valid and reliable data in order to quantify a BN, by developing a methodology for the specific case where the topology of the BN has already been defined, i.e. with the problem already structured.

In this case, the elicitation process is then required to extract and quantify the subjective judgments about the uncertain quantities, which are the conditional probabilities that represent the interrelationships among connected nodes.

There are various protocols for probability elicitation (Morgan, et al., 1990), for a recent review see (Quigley & Walls, 2020). The methodology proposed in this contribution is adapted from the Stanford Research Institute (SRI) model (Ferrell, 1985) (Spetzler & Stael Von Holstein, 1985) (Merkhofer, 1987). Accordingly, the process for eliciting expert judgment is based on seven possible stages: motivating the experts with the aims of the elicitation process, structuring the uncertain quantities in an unambiguous way, conditioning the expert's judgement to avoid cognitive biases, encoding the probability distributions, verifying the consistency of the elicited distributions, aggregating probabilities from different experts and discretizing continuous probability distributions. Moreover, to conduct an elicitation process at least two characters are necessary: a subject, i.e. the expert, and an analyst, i.e. the interviewer. The first one provides expertise, i.e. he/she is "*a person with substantive knowledge about the events whose uncertainty is to be assessed*" (Ferrell, 1985), while the second one has responsibility for designing, developing and executing the process as well as evaluating the procedures. For the rule of analyst, also called facilitator, it is common to have at least one person who is very knowledgeable in elicitation practice and can manage the process, and another one with wide expertise in the area of the design project.

Starting from the SRI protocol and according to the specific requirements of a BN, we develop a four-stage structured methodology to support the elicitation meaningfully. In the next subsection each stage is extensively presented by defining each phase of the process, presenting the roles of the key personnel and highlighting all the potential biases that may influence the process, while proposing appropriate actions in order to minimize the risk of a biased judgment.

4.3.1 The four-stage structured elicitation process

In the following, each stage of the process is presented in detail; the flowchart in Figure 4.2 shows the proposed elicitation process.

Stage 1: *Selecting*. To start, the analysts have to study carefully the project and the proposed BN, to understand which kind of expertise is required: it is fundamental to ensure coverage of all the different aspects of the problem, so more than one expert is usually necessary. This is even more important in civil engineering applications, because in this field experts are usually very specialized. Therefore, the analysts should identify the essential and desired characteristic of experts and build up profiles of experts who may be able to answer questions concerning the quantities of interest, i.e. the values required to be quantified in the BN. Constructing a profile matrix can be useful (Bolger, 2018), which matches the knowledge requirements with the expert roles: it supports the identification of expertise needed as well as justification for the choice of experts. The number of required experts depends then on the variability of expertise per domain. Adding as many experts as possible seems beneficial, however, practically it may be difficult to manage many experts and there will be a diminishing return on adding more experts. In addition, we have to be aware that in real-world it is not so easy to have the availability of many experts. Once the experts have been selected, the analysts have to arrange meetings to conduct interviews. Prior to the meetings, it is recommended to give to the experts an outline about the project and where their knowledge will be useful, so that they have the opportunity to reflect upon the events.

Stage 2: *Structuring*. Individual interviews between the analysts and the selected experts are conducted. The initial part of the interview has two purposes: to introduce the expert to the encoding task as well as identifying and addressing motivational biases (Fischhoff, 1989), such as management bias and expert bias. Management bias occurs when experts provide goals rather than judgments, e.g. "*the dam will not fail*", while expert bias comes when experts become overly confident because they have been labelled as "*experts*". During this initial part of the interview, the BN should be explained, indicating the uncertainty variables that will be elicited and explaining how this process can be useful as regards the resolution of the overall problem. The second part of this stage is concerned with structuring the variables: each quantity of interest that will be quantified needs to be specified so that a measurement scale can be determined. Even if the topology of the BN has already been defined, it is fundamental to review with the experts the definitions of the variables and their states, in order to structure the uncertain quantities in an unambiguous and meaningful way, before starting with the encoding phase. Each variable must have a clear definition that will be understood without any possibility of misunderstanding by the expert. In addition, the states of every variable have to be determined in order to make unambiguous the final estimation of the expert. It is common for a BN to represent the nodes with discrete states: we suggest keeping them binary if possible, to minimize the number of variables to quantify. Depending on the experience and mental models of the experts, it may be appropriate to disaggregate the variable into more elemental variables. This can be very useful in the case of the BN, because each node might depend on several aspects and it can be easier for the experts to evaluate these secondary probabilities. This technique also allows the analysts to combat the motivational biases introduced at the beginning of this stage, i.e. the so-called management bias and expert bias, and also some cognitive biases, e.g. the conjunctive bias, by increasing the level of detail. The conjunctive bias is one of the biases associated to the anchoring heuristic (Tversky & Kahneman, 1974), which states that the overall probability is overestimated in conjunctive problems and underestimated in disjunctive problems.

Stage 3: *Encoding*. This stage is concerned with encoding the expert's uncertainty on the quantity of interest as a probability. Prior to eliciting these quantities training experts on probability and providing relevant information for discussion should be conducted to minimize the presence of potential biases (Tversky & Kahneman, 1974) (Armstrong, et al., 1975). In particular, this can address biases such as anchoring (Tversky & Kahneman, 1974), i.e. when the evaluation is conditioned by an initial assessment, and availability (Kahneman & Tversky, 1973), i.e. when the evaluation is based on the ease with which relevant instances come to mind. Probability training should be provided to calibrate the experts: a brief review of basic probability concepts may be helpful, along with some training questions which can help the experts to become familiar with the elicitation process itself. Experts should be trained on problems relevant to the questions on which they will be providing judgement. When the training is completed, the encoding stage commences. There are many available approaches to elicit probabilities, including direct assessments of probabilities; for a review of methods see (O'Hagan, et al., 2006). A popular encoding procedure for distributions is the fractile method (Cooke 1991), where the expert assesses the median value of their subjective probability distribution along with the (25^{th},75^{th}) and the (5^{th}, 95^{th}) percentiles. Once the initial values have been elicited a parametric distribution can be investigated and assessed for fit with the elicited values. The order in which these quantities are elicited should start with the extreme values first and progress towards the central values, in order to avoid the so-called central bias, i.e. the tendency to give an answer that is closer to the centre of opinions, and to not give an extreme answer. If the expert is uncomfortable with percentiles, questions can be rephrased using qualitative bands, such as "*highly likely*" or "*highly unlikely*", but the percentiles associated with these qualitative terms must be discussed and understood by both expert and analysts. Alternately, graphical techniques (Chaloner, et al., 1993) may be useful to improve the quality of the results. We recommend using the technique which makes the expert more comfortable. In the case that there are a lot of probabilities to be elicited for the same node, we suggest that the expert first ranks the factors from the most to the least influential and subsequently quantifying the relationship, for instance following the swing weight method to elicitation used with multi-attribute decision analysis (Belton & Steward, 2001). Moreover, sometimes it is not possible to elicit data for all the BN components, especially when it is composed by a huge number of nodes or due to a limited time available. In this case, we recommend identifying the quantities of interest that make the most significant contribution to the assessment of the structure, for example through a sensitivity analysis (Li & Mahadevan, 2018). Finally, during the encoding phase, asking the same question in several ways can be a useful way of identifying potential inconsistencies with expert assessments. If this occur the expert should be confronted and encouraged to reflect and respond on the assessments.

Stage 4: *Verifying*. This final stage starts by verifying the consistency of the elicited probabilities. First of all, the analysts should verify that each expert has provided a reflection of their true beliefs. Moreover, it is important to check for trends across the elicited probabilities to determine if there are any indicators of anchoring bias or availability bias. If the results are not satisfactory or biased, the previous stage should be repeated. In the case that the same conditional probabilities have been elicited from different experts, the analysts should then develop an aggregation technique to obtain one single final result: see (Quigley, et al., 2018) for a performance-based approach or, if a consensus amongst experts is desired, see (Gosling, 2018) for a behavioural based approach. Since the proposed methodology is based on discrete states, the final stage of the SRI model, i.e. discretizing continuous probability distributions, is not needed. Once each elicited probability has been verified and, if necessary, aggregated, the analysts should solve the overall BN to achieve the final results. We suggest discussing with the experts also these final outcomes in order to have a further validation of the developed process. After that, the interview ends and the process can be considered concluded.

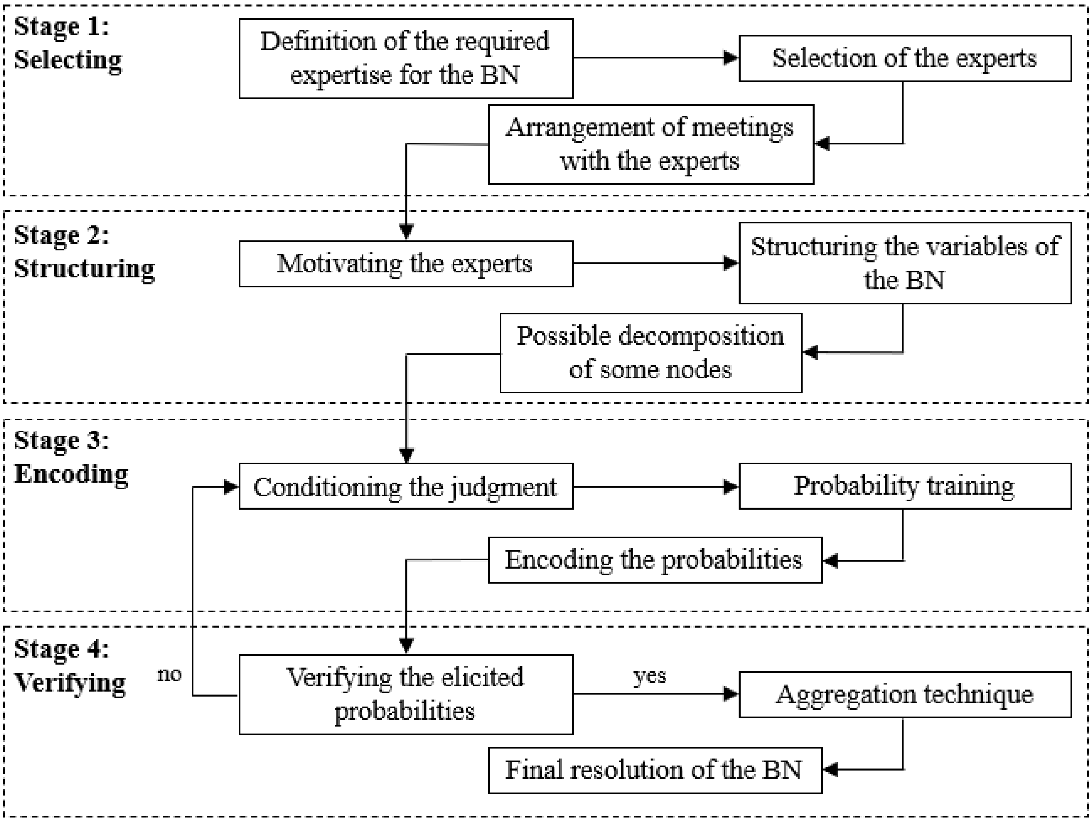


Figure 4.2. Flowchart of the proposed elicitation process.

4.4 The Mountain Chute Dam and GS case study

The case study motivating our research is the Mountain Chute Dam and Generating Station (GS), which is operated by Ontario Power Generation (OPG). Mountain Chute Dam and GS, presented in Figure 4.3, is located in Greater Madawaska Township in Renfrew County (Ontario, Canada): it has an electric power generation capacity of 170 megawatts of clean, renewable electricity. It is situated on the Madawaska River, 64 km upstream from its confluence with the Ottawa River, and it is in the upstream of four other hydroelectric facilities on the Madawaska River: Barrett Chute GS, Calabogie GS, Stewartville GS and Arnprior GS. The construction started in 1965 and was completed in December 1967. Three dams are located at the Mountain Chute GS: one main concrete dam and two earthen block dams, i.e. the north block dam and the whitefish draw dam. The main dam, shown in Figure 4.3(a), consists of the north and the south concrete gravity walls, the sluiceway and the headworks. It is 436 m long and 55 m above the rock foundations at the deepest section; the elevation of the top of the concrete structure is 249.9 m. The north block dam, which is an embankment structure constructed across a shallow depression about 300 m north east of the north abutment of the main dam north, is about 125 m long and has a maximum height of 12 m. Finally, the whitefish draw dam is a block dam preventing the reservoir from flowing out via a side valley, it is located about 2.5 km north of the main dam, it is 204 m long and it has a maximum height of 18 m. More details about Mountain Chute GS and its case study are provided in (El-Awady, et al., 2019) and (Verzobio, et al., 2019).

a) b) c)

Figure 4.3. Mountain Chute Dam and GS: a) the main dam and the sluice gates; b) the downstream of the dam; c) the upstream of the dam with the reservoir.

The main scope of this project is about the general safety of Mountain Chute, with the final aim to estimate the probability of failure of the dams, intended as failure to perform at least one of the required operations, according to the interrelated dam components. In the next subsections, we describe the developed BN and successively the application of the proposed elicitation process, which allows for improving the estimation of the failure probability of the dams, thanks to the acquisition of valid and reliable data from expert knowledge.

4.4.1 Bayesian Network of Mountain Chute Dam and GS

Mountain Chute station includes different kinds of system components. For the purpose of analyzing the failure of this system, all system components should be defined, explained and analyzed. Specifically, components such as rain precipitation, ice loading, earthquake and seismic actions, water pressure, geology and rock type, flood severity, adequacy of discharge capacity, sluice gate, drainage, vegetation control and other secondary components have to be considered. A BN was constructed based on these components and based on the factors that can lead to the failure of the dams, e.g. overtopping, seepage, sliding, stability issues and any operational failure, such as problems related to the head gates or to the electromechanical equipment. The resultant BN is presented in Figure 4.4.

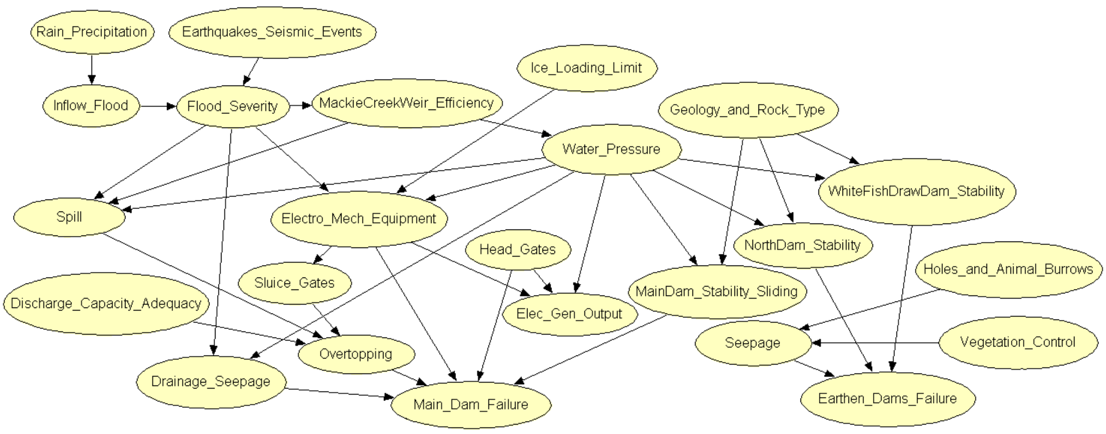


Figure 4.4. Bayesian Network of Mountain Chute Dam and GS showing all the primary variables.

The main purpose of the developed BN, which is represented by 24 different nodes, i.e. the yellow ovals in Figure 4.4, is predicting the probability of failure of the main dam from overtopping, seepage, sliding or any operational failures. Moreover, it estimates the probability of failure of the earthen block dams resulting from the threats of seepage and sliding. In the following, we analyze in detail the BN.

The basic events are rain precipitation, ice loading limits, earthquakes, geological and rock stability, vegetation control and control of animal burrows. It can be seen from the BN that the amount of rain affects the inflow to Mountain Chute dam; this inflow is considered a flood if it exceeds a certain limit. If a flood takes place, it may be normal or severe. Flood severity is also affected by seismic actions and earthquakes. The inflow rate and the severity level of the flood are controlled by the Mackie Creek weir. Controlling the inflow is about preventing severe floods from reaching the dam reservoir. The weir may be efficient or not, depending on the flood severity. After passing the weir, the water in the reservoir, blocked by two earthen block dams and the main concrete dam, is ready to be controlled by the dam head gates; this means that there is water pressure behind the dams that may affect their stability. The geological and rock stability for the structure of the three dams have been considered as it affects the sliding of the dam; sliding is one of the causes of dam breach failure.

In addition, ice loading, water pressure and flood severity are connected to the electromechanical equipment, including turbines; for instance, ice loading affects the failure of the mechanical equipment and at the time of a severe flood and high-water pressure could result in dam failure from maloperations of gates. As regards the electric power generation, the head gates are opened to let the water flow through the penstock to generate electricity from hydropower turbines. If the head gates fail to open, this is considered a failure of the main dam, especially if the water pressure is high in the upstream side of the dam; this may affect the dam stability and also the amount of power generated by the turbines.

Moreover, the flood severity, the weir efficiency in controlling the inflow to the reservoir and the water pressure are all affecting the probability to have spill in the main dam; the spill is the amount of water that exceeds the reservoir maximum capacity limit after considering various controlled outflows. This amount should be released from the upstream side to the downstream side through the spillway (sluiceway) gates or an overtopping failure could take place. The amount of water spill is also related to the capacity of sluiceway, which may not be adequate for that amount of water to be discharged, and to the condition of the sluice gate, i.e. open or failed to open due to electromechanical failure. If the water spill is not released from behind the main dam because of the inadequate capacity of the sluiceway, or because the sluice gate fails to open, there is an increasing probability, i.e. risk, of overtopping failure.

As concerns the main dam, severe floods with increased water pressure increases the possibility to have seepage in the body of the main dam. If the seepage is not completely controlled and monitored through a drain system which may include drain inspection tunnel, this would result in an increasing risk that reduces the remaining lifetime of the dam. Finally, as regards the earthen dams in Mountain Chute GS, seepage may take place because of uncontrolled vegetation and due to animal burrows and holes in the vicinity of the dams. Seepage in the earthen block dams is then an increasing risk for seepage piping and dam breach failure.

After the development of the topology of the BN, the corresponding states have been defined. It was clear that defining more than two states for every component of the BN would have turned the system into a more complex network. On the other hand, more states would have allowed to get more accurate results. Following the proposed methodology of the elicitation process, due to the considerable number of nodes, it has been decided to keep the states of the nodes binary, e.g. fail/no fail, safe/not safe, controlled/not controlled, efficient/not efficient. Table 4.1 presents the defined states for each node. In addition, each state has been associated with a detailed definition or a numerical value, so as to make them quantifiable. As an example, according to the available data, the threshold according to which the rain precipitation passes from the state *low* to the state *high* is when the rain depth reaches 60 mm.

Once the BN structure is completely defined, the conditional probability distributions were determined based on logical inference and limited historical data; these probabilities are defined to represent 100 years of operation for the Mountain Chute Dam and GS. Nevertheless, the available data were not enough, and they did not allow to cover all the nodes of the BN. Then, it was necessary to rely on expert judgment to provide subjective probabilities in order to populate completely the model.

Table 4.1. States of the BN variables.

Variable	States	
Rain precipitation	Low	High
Earthquakes seismic events	Normal	Severe
Ice loading limit	Safe	Not safe
Geology & rock type	Stable	Unstable
Discharge capacity adequacy	Adequate	Not adequate
Head gates main dam	Open	Close/Fail to open
Holes and animal burrows	Controlled	Not controlled
Vegetation control	Controlled	Not controlled
Inflow flood	Low	High
Flood severity	Normal	Severe
Mackie Creek weir efficiency	Efficient	Not efficient
Water pressure	Normal	High
Spill	Yes	No
Electromechanical equipment main dam	Efficient	Not efficient
Sluice gates main dam	Open	Close/Fail to open
Electric generation output	Low	High
Overtopping	Yes	No
Drainage main dam seepage	Leakage	No leakage
Main dam stability sliding	Stable	Unstable
Main dam failure	Fail	No fail
North dam stability	Stable	Unstable
White fish drawn dam stability	Stable	Unstable
Seepage	Exist	Not exist
Earthen dams failure	Fail	No fail

4.4.2 Elicitation Process

By following the methodology proposed in section 4.3, we implemented each stage of the process as follows.

Stage 1: *Selecting*. There were two analysts: one with knowledge in elicitation practice and another with experience in the specific engineering area of failure analysis. After studying the project and the defined model, we identified three areas of expertise from which we sought to elicit expert judgment: structural stability expertise, environmental expertise and system design expertise. While finding one expert per each area was desirable, due to availability constraints we were given access to only one expert, who had a reasonable expertise in all the three areas: he was an engineer of the Ontario Power Generation who was responsible for monitoring the operations of this specific GS. We were aware about the possible difficulty in finding available experts, but managed to satisfy an essential coverage of expertise in all relevant area. A meeting was then arranged at the site of the dam, in order to develop the interview. In preparation, the expert was informed by email about the project and the specific aims of the interview.

Stage 2: *Structuring.* At the beginning of the interview the expert was motivated by explaining the importance of the project, his fundamental rule and how the results will be used. Moreover, the possible presence of motivational biases was investigated, especially the expert bias: it was carefully pointed out to the expert that the goal is not to measure his personal expertise, but to measure his knowledge about the events. Successively, we moved to the second part of this stage: we reviewed the topology of the BN and the states of the variables together with the expert, to ensure that there was no misunderstanding about their definition before moving to the encoding phase. The expert therefore had the opportunity to review the topology of the BN but he decided not to modify it, probably because we arrived at the meeting with a too refined model; he also agreed with the proposed variables, refusing the possibility to disaggregate some nodes too. In addition, we spent more time explaining meticulously to the expert the meaning of each variable and the corresponding states: after this discussion, and based on his opinions, we agreed to change the definition as well as the threshold of some states.

Stage 3: *Encoding*. The encoding phase started by conditioning the expert's judgment in order to avoid the possible presence of some cognitive biases. In particular, we focused mainly on the anchoring, which is of particular concern with BNs given the large number of variables being quantified: after the first assessment, the expert must avoid linking the subsequent assessments with the previous one, as it would result in a biased adjustment. Following this discussion, a probability training was carried out: we reviewed some probability concepts and trained the expert with some specific questions similar to those that we would be asked in the encoding phase, trying for instance to clarify the difference between a frequent event and a very rare event. In addition, the probability scale presented in Figure 4.5 was introduced, that we had established in order to help the expert during this stage of the process. This led to the encoding phase, which was the most important and the longest, i.e. around 1 hour. It was developed by asking questions in several ways, e.g. direct assessment of probabilities but also rephrasing the questions using qualitative bands, to find potential inconsistencies in the answers and also to reduce the influence of the explained biases. We chose these types of questions because we had noticed that the expert was not completely comfortable using the percentiles. For example, we asked the following questions: "*What is the probability of a high inflow if the state of the rain precipitation is low?*"; "*How frequently does it occur that the head gates of the main dam fail to open?*"; "*How many days per year is it highly likely to have an inadequate capacity of sluiceway?*" During this phase it is important that the questions are very clear: for instance, we had to pay attention to the reference time of each question in order to avoid misunderstanding with interpreting the expert data, for example caused by the difference between the design time of a dam and the real-life time of the dam.

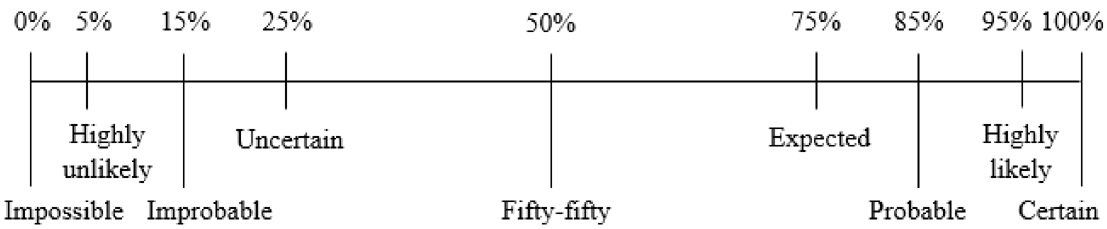


Figure 4.5. Probability scale used during the elicitation process.

Stage 4: *Verifying.* Finally, a verification of the individual elicited probabilities was developed: the results were satisfactory because the numerical outcomes seemed to coincide appropriately with the true beliefs of the expert. Since we had the availability of just one expert, no aggregation technique was necessary. Due to a limited time available the interview ended without the time to solve the overall BN and to discuss the resulting outcomes, which would have been useful also as an additional verification. In the end the interview lasted approximately two hours.

4.4.3 Case study results

In conclusion, after updating the probability distributions with data inferred from expert engineering judgment as presented in the previous subsection, the overall BN was solved in order to estimate the failure probabilities, which we remember are intended as failure to perform at least one of the required operations. Figure 4.6 shows the results, achieved using the software Hugin (http://hugin.sourceforge.net/): the Bayesian inference results in a failure probability $p_F = 0.0135$ for the main dam and $p_F = 0.0133$ for the earthen block dams, both evaluated over the lifetime of the dams, i.e. 100 years. It is evident that adding expert engineering judgments helps in reducing the uncertainties in the network, and gives better estimates for the operation of the dam in comparison with those obtained using only the limited available data and logical inference (El-Awady, et al., 2019): Table 4.2 shows the improvement achieved in the results using engineering expert knowledge. These final results about the failure probability are satisfactory as they are close to those expected when considering these kind of systems design components: it provides approximately a failure of 1 in 10000 at any year or equivalent to designing a dam for failure due to the so-called ten thousand years flood.

Table 4.2. Comparison of results.

	p_F main dam	p_F earthen block dams
Our results using engineering expert knowledge	1.35%	1.33%
Results without engineering expert knowledge (El-Awady, et al., 2019)	77.15%	68.77%

In addition, a BN is useful because explicates the cause-effect relationship, that is essential for a better understanding of the dam safety. For instance, it is possible to understand the main contributors to the failure of the main dam. Figure 4.7 shows the conditional probabilities of each node given the main dam has failed. The most influential variables and the associated probabilities are: seepage, i.e. 0.46 leakage, electromechanical equipment, i.e. 0.25 fail, sliding stability, i.e. 0.24 unstable, head gates, i.e. 0.23 failed to open. On the other hand, overtopping has just a probability of the 0.08.

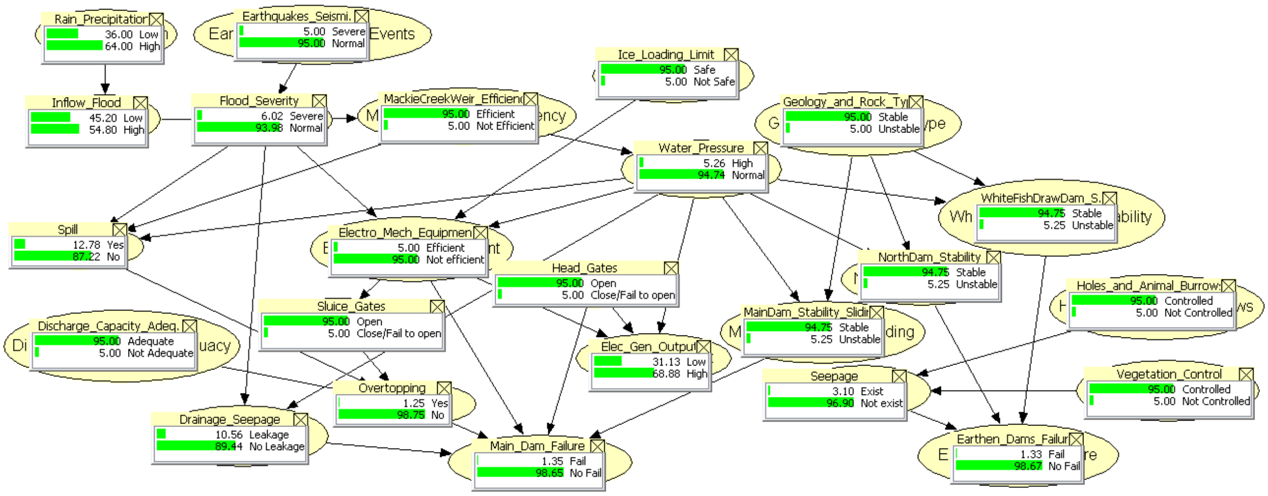


Figure 4.6. The quantified BN of Mountain Chute Dam and GS (note that the numerical values are percentage).

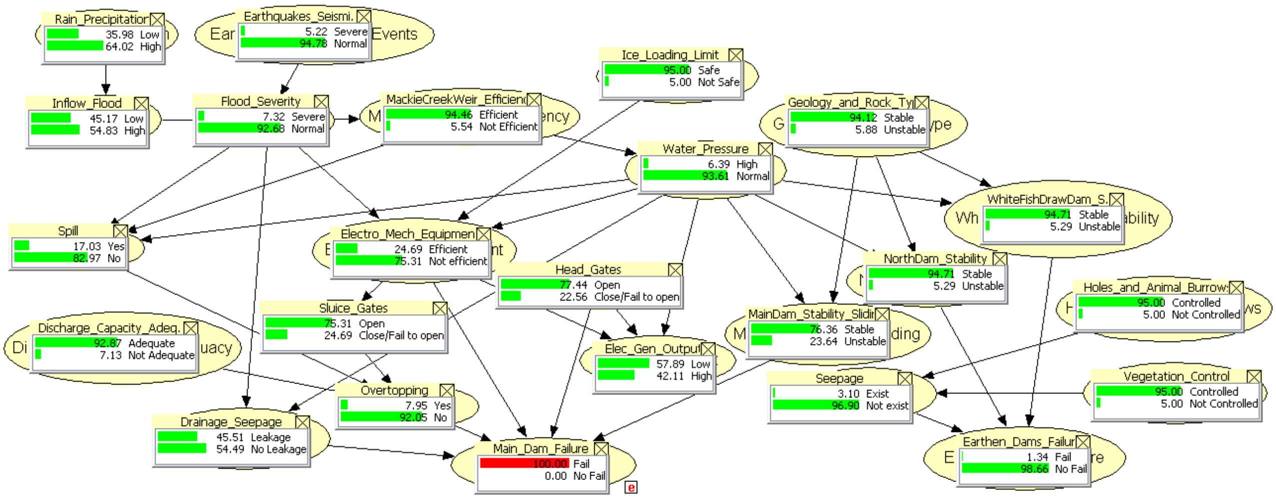


Figure 4.7. BN of Mountain Chute Dam and GS given the evidence that the main dam fails (note that the numerical values are percentage).

4.4.4 Evaluation of the process

After discussing the implementation of the planned elicitation process to this specific engineering case study and the consequent results, in this section we propose a critical discussion about the main steps of the process, based on what happened during its application, in order to understand how to improve the process and to give practical guidelines that can be used for similar applications in the future.

- The selection of the experts is fundamental and should not be underestimated. In particular, working in a field where experts have narrow areas of expertise rather than generalists requires more experts to be involved in the elicitation to ensure sufficient coverage of the relevant issues. It is worthwhile reflecting on expertise that is desirable for the study or essential. In our case, even if we had the availability of only one expert, we managed to satisfy an essential coverage of expertise in all relevant area. For larger projects, expert profile matrices can be useful at structuring this reflection (European Food Safety Authority, 2014) (Bolger, 2018).
- As regards the number of interviewers, the choice of two analysts with different competences seemed appropriate: it is essential to have at least one facilitator with the expertise in the elicitation practice that have to lead all the process, and at least another one with engineering knowledge that have to make his contribution regarding the technical aspects of the specific design project.
- The interview was conducted at the dam site: this choice has proved to be suitable because it allowed us also to understand better some practical aspects of the dam operation. As regards the available time for the interview, we had scheduled a two-hour meeting but in the end we realised that it was not enough to properly complete all the planned elicitation process. During the scheduling phase we had probably underestimated some aspects of the interview that can lead to a delay, so we suggest detailed planning of the interview to identify an appropriate time.
- As concerns the *structuring* phase, we started with a very refined model, which can have some disadvantages, as it was evident that the expert did not propose many changes to the structure and agreed almost completely with our proposal;

if the model had been less refined then the expert would have been more empowered to create a different model. Since this phase is fundamental in order to achieve accurate results during the encoding, we recommend involving the experts in the creation of the model and its variables.

- The training phase is fundamental to get accurate and reliable data from the experts. Unfortunately, the time that we spent on training was too little, both because of the limited available time and because the expert did not seem too convinced about the importance of this phase. Consequently, we suggest adding a motivational phase at the beginning of this stage, i.e. *encoding*, in the same way as in the *structuring* stage, with the aim to explain to the expert why it is necessary its development in order to calibrate him before encoding.
- There is a trade-off between the level of detail in a model and the time required to populate with probabilities. The model structure needs to be flexible and adapt during the *encoding*, as experts may not be comfortable expressing uncertainties on variables and require an elaboration of the node.
- As concerns the *encoding* techniques, the choice to ask the questions with direct assessment of probabilities and rephrasing the question using qualitative bands was made according to the specific features of our expert: it was clear to us for instance that he was not comfortable with the use of the percentiles. A good idea is then to prepare the questions in different ways before the meeting, and to choose which ones to use only during the interview, so as to make the expert as comfortable as possible.
- As regards the *verifying* stage, the limited available time did not let us to carry out it completely. This is a problem that we have already highlighted and should be considered properly during the scheduling phase. In particular, it would have been important to have more time available in order to verify with the expert also the final resolution of the BN, based on the elicited variables.
- Finally, during the implementation of all the stages we have paid close attention to the possible presence of heuristics and biases, by following the appropriate actions suggested in the methodology in order to minimize the risk of biased judgments. The achieved results allow us to confirm the suitability of our four-stage elicitation process.

4.5 Conclusions

BNs allow for analysing complex systems like dams in order to develop a safety analysis based on probabilistic estimates of failure. Due to the lack of data, in this paper we proposed a methodology for an elicitation process aimed specifically at quantifying BNs, with the final goal of collecting reliable data from engineering knowledge. The elicitation exercise we carried out for this specific case study regarding the safety of the Mountain Chute Dam and GS, even if developed in a simplified way, demonstrated the potential and the usefulness of the engineering expertise, and allowed us to learn many lessons that are useful for improving the methodology, which we intend to address in future for similar applications. In summary, we can conclude as follows:

- While the elicitation process has been applied in many fields, in civil engineering there is little experience of applying formal elicitation processes to quantify models. This paper demonstrates that engineering knowledge and experience can be very useful to solve appropriately also this type of analysis.
- It is undeniable that the elicitation requires a structured and facilitated process in order to achieve accurate and reliable data, by avoiding the adverse impact of biases. However, there is no perfect elicitation process: it has to be planned according to the particular context and to the specific aims. Consequently, we proposed a detailed methodology for the precise aim to quantify a BN.
- Our four-stage structured elicitation process works properly according to the results achieved in the case study. However, this has been just our first experience in implementing an elicitation process and indeed, during the application, we have noticed some aspects that need to be improved in order to make the process even more successful and reliable.
- As regards to future work, we aim to improve this structured methodology based on what we have learnt from this first application, and to apply it to other civil engineering structures, e.g. bridges.

Acknowledgements

The authors would like to thank Ontario Power Generation, especially Tony Bennett and Andy Zielinski, for letting us to work on such an important real-life case study, sharing all the required material and allowing us to develop the interview with one of their engineers, whose availability is greatly acknowledged.

5. Quantifying the benefit of Structural Health Monitoring: what if the manager is not the owner?

Structural Health Monitoring, 2018, 17(6): 1393-1409

Denise Bolognani[a], Andrea Verzobio[b], Daniel Tonelli[c], Carlo Cappello[d], Branko Glisic[e], Daniele Zonta[f], John Quigley[g]

Position and affiliation:
[a]Postdoc, Department of Civil, Environmental & Mechanical Engineering, University of Trento, Italy
[b]PhD Student, Department of Civil and Environmental Engineering, University of Strathclyde, UK
[c]PhD Student, Department of Civil, Environmental & Mechanical Engineering, University of Trento, Italy
[d]Postdoc, Department of Civil, Environmental & Mechanical Engineering, University of Trento, Italy
[e]Professor, Department of Civil and Environmental Engineering, University of Princeton, USA
[f]Professor, Department of Civil, Environmental & Mechanical Engineering, University of Trento, Italy
[g]Professor, Department Management Science, University of Strathclyde, UK

Summary of the paper

This paper investigates how decision-making can be distorted when multiple rational decision makers are involved in the decision chain. In particular, we focus on the quantification of the benefit of Structural Health Monitoring (SHM), using the concept of Value of Information (*VoI*) (introduced in section 2.3): we formalize a rational method for quantifying the *VoI* when two different actors are involved in the decision chain, i.e. the owner and the manager of the structure. The two decision makers, even if both rational and exposed to the same background information, may still act differently because of their different appetites for risk (discussed in section 2.1.2). To illustrate how this framework works, we evaluate a hypothetical *VoI* for the Streicker Bridge, a pedestrian bridge in Princeton University campus equipped with a fiber optic sensing system. This contribution demonstrates that, with the developed

methodology, decision-making is distorted and the *VoI* results different, in comparison to the case where all the decisions are made by the same individual.

5.1 Introduction

Although the utility of structural health monitoring (SHM) has rarely been questioned in our community, very recently a few published papers (Thons & Faber, 2013) (Zonta, et al., 2014) have clarified the way that the benefit of monitoring can be properly quantified. Indeed, seen from a mere structural engineering perspective, the utility of monitoring may not be immediately evident. Wear for a minute the hat of the manager of a Department of Transportation (DoT), responsible for the safety of a bridge: would you invest your limited budget on a reinforcing work or on a monitoring system? A retrofit work will increase the bridge load-carrying capacity and therefore its safety. On the contrary, sensors do not change the bridge capacity, nor reduce the external loads. So how can monitoring affect the safety of the bridge? The answer to this legitimate question goes roughly along these lines: monitoring does not provide structural capacity, rather better information on the state of a structure; based on this information, the manager can make better decisions on the management of the structure, minimizing the chances of wrong choices, and eventually increasing the safety of the bridge over its lifespan. Therefore, to appreciate the benefit of SHM, we need to account for how the structure is expected to be operated and eventually recast the monitoring problem into a formal economic decision framework.

The basis of the rational decision-making is encoded in axiomatic Expected Utility Theory (EUT), first introduced by Von Neumann and Morgenstern (Neumann & Morgenstern, 1944) in 1944, and later developed in the form that we currently know by Raiffa and Schlaifer (Raiffa & Schlaifer, 1961) in 1961. EUT is largely covered by a number of modern textbooks (among the many, we recommend Parmigiani and Inoue (Parmigiani & Inoue, 2009) to the Reader of SHM who is approaching the topic for the first time). Within the framework of EUT, the benefit of information, such as that coming from a monitoring system, is formally quantified by the so-called Value of Information (*VoI*). For the state of art about the *VoI* see section 2.3.

Broadly speaking, the value of a SHM system can be simply defined as the difference between the benefit, or expected utility u^*, of operating the structure *with* the monitoring system and the benefit, or expect utility *u*, of operating the structure *without* the system. Both *u** and *u* are expected utilities calculated *a priori*, i.e. *before* actually receiving any information from the monitoring system. While in *u* we assume the knowledge of the manager is his a priori knowledge, *u** is calculated assuming the decision maker has access to the monitoring information and is sometimes referred as to *preposterior utility*. The difference between these values measures the value of the information to the decision maker. Clearly, if the monitoring does not provide any useful information, the preposterior *u** is equal to the prior *u*, and the value of monitoring information is zero.

Typically, it is assumed that there is one decision maker for all decisions, i.e. deciding on both investments as well as operations. This individual could be for example an idealized manager of a DoT, as the fictitious character 'Tom' who appears in (Zonta, et al., 2014). We must recognize that in the real world the process whereby a DoT makes decision over its stock is typically more complex, with more individuals involved in the decision chain. Even oversimplifying, we always have at least two different decision stages. First a decision is made on whether or not to buy and install the monitoring system on the structure; this is a problem of long-term planning and investment of financial resources. This decision is typically carried out by a high-level manager, that in this paper we will conventionally refer to as *owner*, whose key performance measure is return on investment. The second stage concerns the day-to-day operation of the structure which includes for example maintenance, repair, retrofit or enforcing traffic limitations, once the monitoring system is installed; if installed these decisions may be informed by the monitoring system. Most of the time, the manager and the owner of the structure are different individuals. Both decision makers are motivated to maintain a high level of long-term availability for the structure, which is challenging as the state of the structure is never known precisely while in operation. Operators balance two types of errors, either removing a structure from operation prematurely for maintenance or operating too long resulting in a failure; both of which are based on imperfect information concerning the state of the structure. Decision makers will differ in their choices under uncertainty even when they have access to the same information if they have different appetites for risk. As such, the owner needs to consider the operators appetite for risk when deciding whether to install a monitoring system, as this will indicate how the system will inference the operators decision-making and as such the value of this information.

The aim of this work is to formalize a rational method for quantifying the Value of Information when two different actors are involved in the decision chain: the *manager*, who makes decisions regarding the structure, based on monitoring data; and the *owner*, who chooses whether to install the monitoring system or not, before having access to these data. We start explaining why and how two different individuals, both rational and provided with the same background information, can end up with different decisions. Next, we review the basis of the *VoI*, which illustrates a method for evaluating the *VoI* in SHM-based decision-problems, and revise the framework of Zonta et al. (Zonta, et al., 2014), to include the difference between the manager and the owner. To illustrate how this framework works we apply it to the same decision problem reported in (Zonta, et al., 2014): the Streicker Bridge case study. This is a pedestrian bridge at Princeton University campus, which is equipped with a continuous monitoring system. Some concluding remarks are reported at the end of the paper.

5.2 SHM-based decision

This section has already been presented in the literature review of chapter 2: the assumptions and the framework of a rational SHM-based decision process are presented in section 2.1; the rational judgment, i.e. using Bayesian inference, is illustrated in section 2.1.1; the rational decision, i.e. using EUT, is presented in section 2.1.2; the consequent calculation of the *VoI* is formulated in section 2.3.

5.3 Two individuals, two decisions

In the *classical* formulation of the *VoI* stated above, we have implicitly assumed that the decision is taken at any stage by the same rational individual, characterized by a defined background information and utility function. We address now the problem of quantifying the *VoI* when two separate individuals are involved in the decision chain. We conventionally denote *manager* (M), the one who makes decisions on the day-to-day operation of the structure, and *owner* (O), the one who is in charge of the strategic investments on the asset and decide on whether to install the monitoring system or not. Referring to Figure 2.5, the manager is the one who takes decisions at stage two, while the owner decides at stage one. We will refer to the *classical* formulation of *VoI*, as stated in the previous section, as to *unconditional* - in contrast with the *conditional VoI* which we are about to introduce.

A common misunderstanding, not only in our community, is that two individuals, if both rational and exposed to the same observation, should always end up with the same decision. In the real world, there are a number of components in the SHM-based decision process that are inherently subjective, so different decisions by different individuals should not be necessarily be seen as an inconsistency. This concept needs a deeper explanation: with reference to Figure 2.1, the reasons whereby two individuals, both rational, can take a different decision based on the same observation include the following:

a) The two have a different prior knowledge of the problem – i.e. they use different priors P(S).
b) They interpret differently the observation – i.e. they use different interpretation models, which are encoded in the likelihood function P($\mathbf{y}|S$).
c) They have a different expectation or knowledge of the possible outcome of an action – i.e. they assume different outcome vectors $\mathbf{z}$.
d) They weight differently the importance of an outcome - i.e. they use different utility functions $U(\mathbf{z})$.

Differences in (a) (b) and (c) are merely about background knowledge and may actually occur in the real world; however, we expect that two individuals with similar experience and education should generally agree on any of that. For example, two structural engineers with common background will probably agree on the limited importance of a bending crack visible on an unprestressed reinforced concrete beam, while a non-expert could be over-concerned. In this paper, we will assume that the two agents fully agree on (a), (b) and (c), while they only differ in the way how they weight outcomes (d), through their utility function. The utility function is not a matter of background knowledge, rather it reflects the value of the individual as to the consequence of an action. Therefore, there is no logical argument to judge one utility function better than another one, as long as it does not violate the axioms of the expected utility theory.

Even limiting our discussion to the case where the outcome z is just a monetary loss or gain, the utility function adopted by different people can be very different based on their particular individual risk aversion (Bernoulli, 1954) (Kahneman & Twersky, 1984). For instance, an agent is risk neutral if his or her utility function U is linear with the loss or gain z, as shown in Figure 2.2. Since the expected utility is proportional to the probability of realization, as shown in Eq. (2.8), risk neutrality implies indifference to a gamble with an expected value of zero. So, for example, to a risk neutral agent a 1% probability of losing $100 is equivalent to a certain loss of $1.

In practice it is commonly observed that individuals tend to reject gambles with a neutral expected payoff: in the example above individuals often prefer to pay $1 off the pocket rather than taking the risk of losing $100. This condition is referred to as risk aversion and can be graphically represented with a concave (i.e., with negative second derivative) utility function, as shown in Figure 2.2. The condition of risk aversion is consistent with the observation that the marginal utility of most goods, including money, diminish with the amount of goods, or the wealth of the decision maker, as observed since Bernoulli (Bernoulli, 1954).

Dealing with losses, risk aversion respect to a loss depends on the amount of the loss with respect to the decision maker's own wealth or the extent of his or her own asset: when the loss is much smaller than the whole value of the asset, the agent tends to be risk neutral, while they became risk averse when the loss is a significant fraction of their asset. In our situation, the owner, who is in charge of the strategic development of the agency, typically manages a large stock of structures, and the loss corresponding to an individual structure is a much smaller than the overall asset value. In this case, it is likely that the owner is risk neutral with respect to the loss compared to the value of a single structure. In contrast, the manager is responsible for the safety of a single structure: in this case the value of the structure corresponds to the value of the asset, and their behaviour is likely to be risk adverse respect to the loss of that particular structure.

To proceed with the mathematical formulation, we have to acknowledge that the two agents involved in the decision chain, the owner and the manager, may have different utility functions. We are going to use indices $^{(M)}$ or $^{(O)}$ to indicate that a quantity is intended from one or the other perspective. The expected utility of the manager is calculated as:

$$^{(\mathrm{M})}u(a_j) = \sum_{i=1}^{N} {}^{(\mathrm{M})}U\left(\mathbf{z}(a_j, S_i)\right) \mathrm{P}(S_i), \tag{5.1}$$

and we may calculate the optimal action and the maximum utility from the manager perspective as in the following:

$$^{(\mathrm{M})}u = \max_j {}^{(\mathrm{M})}u(a_j), \qquad {}^{(\mathrm{M})}a_{\mathrm{opt}} = \arg\max_j {}^{(\mathrm{M})}u(a_j). \tag{5.2a,b}$$

If the owner was in charge of the entire decision chain, we would end up with analogous expressions of optimal action $^{(\mathrm{O})}a_{\mathrm{opt}}$ and maximum expected utility $^{(\mathrm{O})}u_{\max}$, this time from the owner perspective. Observe that the optimal choice of the owner does not necessarily coincide with that of the manager, meaning that if the owner was in charge of the full decision chain, they would behave differently from the manager. Continuing on this rationale, we can reformulate the expression of posterior utilities, preposterior utilities and *VoI* from the owner or the manager perspective.

However, the situation we are discussing is different: the owner is the one who decides on the monitoring system installation, but the manager is the one who decides which is the optimal action at the second stage. Therefore, all utilities are from the owner perspective, but should be evaluated accounting for the action that the manager, not the owner, is expected to choose. In other words, the utility of the owner is *conditional* to the action chosen by the manager $^{(\mathrm{M})}a_{\mathrm{opt}}$. For example, the prior utility of the owner conditional to the decision expected by the manager reads:

$$^{(\mathrm{O|M})}u = {}^{(\mathrm{O})}u\left({}^{(\mathrm{M})}a_{\mathrm{opt}}\right) = {}^{(\mathrm{O})}u\left\{\arg\max_j {}^{(\mathrm{M})}u(a_j)\right\}, \tag{5.3}$$

where the index $^{(\mathrm{O|M})}$ on the utility $^{(\mathrm{O|M})}u$ indicates that this utility is *conditional* to the manager's choice, in opposition to the *unconditional* utility $^{(\mathrm{O})}u$ calculated assuming the owner in charge of the full decision chain. We can proceed accordingly to formulate the *posterior conditional utility* (the utility of the owner after the manager has observed the monitoring response):

$$ {}^{(\mathrm{O|M})}u = {}^{(\mathrm{O})}u\left({}^{(\mathrm{M})}a_{\mathrm{opt}}(\mathbf{y})\right) = {}^{(\mathrm{O})}u\left\{\arg\max_j {}^{(\mathrm{M})}u\left(a_j,\mathbf{y}\right)\right\}, \tag{5.4} $$

and similarly the *preposterior conditional utility* (the utility of the owner in the expectation of what the manager would decide if a monitoring system was installed):

$$ {}^{(\mathrm{O|M})}u* = \int_{D_y} {}^{(\mathrm{O})}u\left\{\arg\max_j {}^{(\mathrm{M})}u\left(a_j,\mathbf{y}\right)\right\} \cdot \mathrm{p}(\mathbf{y})\, dy. \tag{5.5} $$

Eventually, the *conditional VoI* is the difference between the preposterior and the prior *conditional* utilities:

$$ \begin{aligned} VoI = {}^{(\mathrm{O|M})}u* - {}^{(\mathrm{O|M})}u &= \\ &= \int_{\mathrm{D_y}} {}^{(\mathrm{O})}u\left\{\arg\max_j {}^{(\mathrm{M})}u\left(a_j,\mathbf{y}\right)\right\} \cdot \mathrm{p}(\mathbf{y})\, \mathrm{dy} \\ &- {}^{(\mathrm{O})}u\left\{\arg\max_j {}^{(\mathrm{M})}u\left(a_j\right)\right\}. \end{aligned} \tag{5.6} $$

The unconditional and conditional formulations are summarized and compared in Table 5.1. At this point, it is interesting to compare the unconditional and the conditional utilities, and also the value of information. The unconditional utility, prior or preposterior, is basically the owner's utility of their favourite choice, while the conditional utility is the owner's utility of the choice of someone else. If the two choices coincide, the conditional utility is equal to the unconditional prior utility. If they do not coincide, the manager's choice can only be suboptimal from the owner's perspective, and therefore the conditional utility must be equal or lower than the unconditional. Therefore, the following relationships must hold:

$$ {}^{(\mathrm{O|M})}u \le {}^{(\mathrm{O})}u, \qquad {}^{(\mathrm{O|M})}u^* \le {}^{(\mathrm{O})}u^*. \tag{5.7a,b} $$

In a situation with one decision maker the *VoI* cannot be negative; if the decision maker anticipated misleading data it would be optimal to discard it resulting in a *VoI* of 0. However, we consider a situation of two decision makers and demonstrate that the *VoI* can be negative to one, i.e. the owner, as the new information is resulting in the other, i.e. the manager, making decisions that are less preferred by the owner than with no information.

Table 5.1. Value of Information of a monitoring system in the unconditional and conditional formulation.

Unconditional formulation Manager (M) = Owner (O)	Conditional formulation Manager (M) ≠ Owner (O)
Prior utility without monitoring	
${}^{(O)}u = \max_j {}^{(O)}u(a_j)$ ${}^{(O)}a_{\text{opt}} = \arg\max_j {}^{(O)}u(a_j)$	${}^{(O\vert M)}u = {}^{(O)}u\left({}^{(M)}a_{\text{opt}}\right) = {}^{(O)}u\left\{\arg\max_j {}^{(M)}u(a_j)\right\}$
Posterior utility with monitoring	
${}^{(O)}u(\mathbf{y}) = \max_j {}^{(O)}u(a_j,\mathbf{y})$ ${}^{(O)}a_{opt}(\mathbf{y}) = \arg\max_j {}^{(O)}u(a_j,\mathbf{y})$	${}^{(O\vert M)}u(\mathbf{y}) = {}^{(O)}u\left({}^{(M)}a_{opt}(\mathbf{y})\right)$ ${}^{(O\vert M)}u(\mathbf{y}) = {}^{(O)}u\left\{\arg\max_j {}^{(M)}u(a_j,\mathbf{y})\right\}$
Preposterior utility with monitoring	
${}^{(O)}u^* = \int_{D_\mathbf{y}} \max_j {}^{(O)}u(a_j,\mathbf{y}) \cdot \mathrm{p}(\mathbf{y})\, dy$	${}^{(O\vert M)}u^* = \int_{D_\mathbf{y}} {}^{(O)}u\left\{\arg\max_j {}^{(M)}u(a_j,\mathbf{y})\right\} \cdot \mathrm{p}(\mathbf{y})\, dy$
Value of information of the monitoring system	
$VoI = {}^{(O)}u^* - {}^{(O)}u$	$VoI = {}^{(O\vert M)}u^* - {}^{(O\vert M)}u$

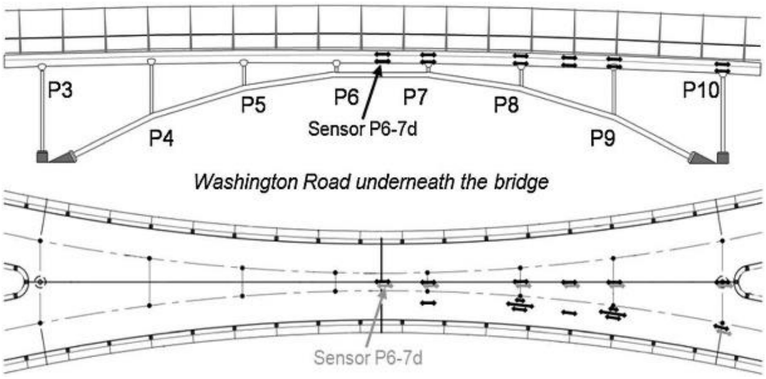
5.4 The Streicker Bridge case study

To illustrate how the presence of two different decision makers in the decision chain affects the way how the *VoI* is evaluated, we consider the case of Malcolm, the fictitious manager of an imaginary Office of Design and Construction at Princeton University, protagonist in (Zonta, et al., 2014) and (Cappello, et al., 2016). Malcolm is responsible for the Streicker Bridge, a pedestrian bridge located on Princeton University campus. The bridge and its monitoring system are illustrated in much detail in a number of past publications (Glisic & Adriaenssens, 2010) (Glisic & Inaudi, 2012) (Glisic, et al., 2011), we summarise the main structural features, for clarity. The deck of the bridge is a continuous thin concrete posttensioned deck featuring a characteristic X-shape connecting four different sectors of Princeton Campus. From the structural point of view, it consists of a thin post-tensioned supported by a high resistance steel lattice. The main span of the bridge overpasses Washington road, a busy public road the campus (see Figure 5.1(a) and Figure 5.1(b)).

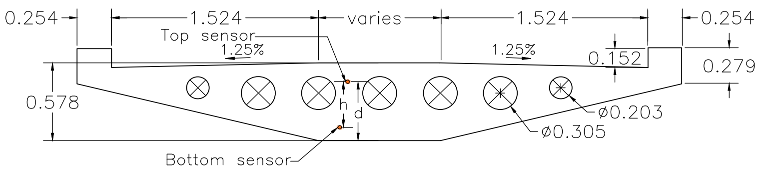
The SHM-*lab* of Princeton University instrumented the bridge with two SHM systems: (i) global structural monitoring using discrete long-gauge strain Fiber Optic Sensors (FOS), based on fiber Bragg-grating (FBG) (Kang, et al., 2007), and (ii) integrity monitoring, using truly distributed FOS based on Brillouin Optical Time Domain Analysis (BOTDA) (Nikles, et al., 1996). These two approaches are complementary: discrete sensors monitor an average strain at discrete points, while the distributed sensors monitor one-dimensional strain field. Discrete FOS embedded in the bridge deck have gauge length 60 cm and feature excellent measurement properties with error limits of ± 4 με. Thus, they are excellent for assessment of global structural behaviour and for structural identification. Instead, distributed FOS have accuracy an order of magnitude lower than discrete sensors and so cannot be used for accurate structural identification; they are used for damage detection and localization. Figure 5.1(c) shows the sensors map in the main span, while Figure 5.1(d) its cross section.

a)

b)



c)



d)

Figure 5.1. The Streicker bridge: view of the bridge (a)(b), location of sensors in the main span (c), main cross section (d).

5.4.1 Agents

To make the case study easier to understand, we imagine the bridge managed by two agents with distinct roles:

- Ophelia (O) is the *owner* responsible for Princeton's estate; she is Malcolm's supervisor and decides on whether to install the monitoring system or not.
- Malcolm (M) is the *manager* responsible for the bridge operation and maintenance, graduated in civil engineering and registered as a professional engineer, who has to take decisions on the state of the bridge based on monitoring data, exactly as in Zonta et al. (Zonta, et al., 2014).

We assume that Ophelia and Malcolm are both rational individuals and that have the same knowledge background as for possible damage scenarios S of the bridge, prior information, and they have the same knowledge of the consequence of a bridge failure. They only differ in the way how to weight the seriousness of the consequences of a failure. It is probably unnecessary to remind that, while the Streicker Bridge is a real structure, the two characters, Ophelia and Malcolm, are merely fictitious and do not reflect in any instance the way how asset maintenance and operation is performed at Princeton University.

5.4.2 States and likelihoods

As part of this fictitious story, we suppose that both Ophelia and Malcolm are concerned by a single specific scenario: a truck, maneuvering or driving along Washington road, could collide with the steel arch supporting the concrete deck of the bridge. In this oversimplified example, we will assume that after an incident the bridge will be in one of the following two states:

- *No Damage* (U): the structure has either no damage or some minor damage, with negligible loss of structural capacity.
- *Damage* (D): the bridge is still standing but has suffered major damage; consequently, Malcolm estimates that there is a chance of collapse of the entire bridge.

Similar to the assumptions in (Thons & Faber, 2013), we assume Malcolm (and similarly Ophelia) focuses on the sensor installed at the bottom of the middle cross-section between P6 and P7 (called Sensor P6-7d, see Figure 5.1(c)).

We understand that for both Ophelia and Malcolm the two states represent a set of mutually exclusive and exhaustive possibilities, which is to say that P(D) + P(U) = 1. On the basis of their experience, they both agree that scenario U is more likely than scenario D, with prior probabilities P(D) = 30% and P(U) = 70%, respectively.

We can also assume that both use the same interpretation model, i.e. they interpret identically the data from the monitoring system. As Malcolm will pay attention only to the changes at the midspan sensor (labelled P6-7d in Figure 5.1(c)), we presume, in the same way as in (Zonta, et al., 2014), that he expects the bridge to be undamaged if the change in strain will be close to zero. However, he is also aware of the natural fluctuation of the strain, due to thermal effects, and to a certain extent due to creep and shrinkage: he estimates this fluctuation to be in the order of ±300 με (Zonta, et al., 2014). We can represent this quantity with a probability density function pdf(ε|U), with zero mean value and standard deviation $\sigma = 300$ με, which describes Malcolm's expectation of the system response in the undamaged (U) state, i.e. this is the likelihood of no damage. On the other hand, if the bridge is heavily damaged (D) but still standing. Malcolm expects a significant change in strain; we can model the likelihood of damage pdf(ε|D) as a distribution with mean value 1000 με and standard deviation of $\sigma = 600$ με, which reflects Malcolm's uncertainty of expectation (Zonta, et al., 2014). Before the data are available, he can also predict the distribution of ε, which is practically the so-called evidence in classical Bayesian theory, through the following formula:

$$\mathrm{pdf}(\varepsilon) = \mathrm{pdf}(\varepsilon|\mathrm{D})\cdot\mathrm{P(D)} + \mathrm{pdf}(\varepsilon|\mathrm{U})\cdot\mathrm{P(U)}. \tag{5.8}$$

When the measurement ε is available, both update their estimation of the probability of damage consistently with Bayes' theorem:

$$\mathrm{pdf}(\mathrm{D}|\varepsilon) = \frac{\mathrm{pdf}(\varepsilon|\mathrm{D})\cdot\mathrm{P(D)}}{\mathrm{pdf}(\varepsilon)}, \tag{5.9}$$

where pdf(D|ε) is the posterior probability of damage. Figure 5.2(a) shows the two unnormalized posterior distributions along with the evidence. Note that the posterior probability of damage starts exceeding the posterior of no-damage when the measurement ε exceeds the threshold $\bar{\varepsilon}_p = 569$ με.

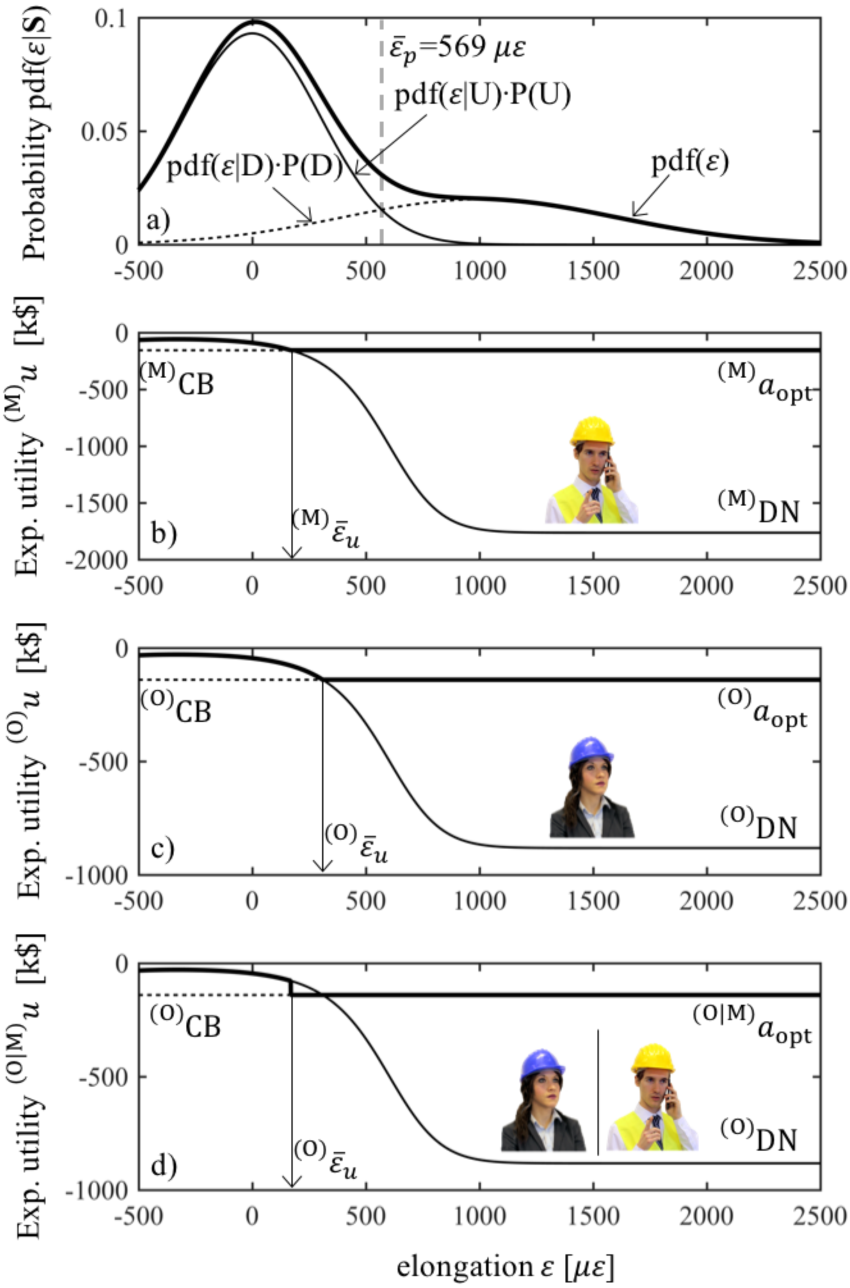


Figure 5.2. Representation of Malcolm's estimation of the state of the bridge a priori (a), Malcolm's decision model with monitoring data (b), Ophelia's decision model with monitoring data (c), Ophelia's decision model based on Malcolm's own (d).

5.4.3 Decision model

After he assesses the state of the bridge, we assume that Malcolm can decide between the two following actions:

- *Do nothing* (DN): no special restriction is applied to the pedestrian traffic over the bridge or to road traffic under the bridge.
- *Close Bridge* (CB): both Streicker Bridge and Washington Road are closed to pedestrians and road traffic, respectively; access to the nearby area is restricted for the time needed for a thorough inspection, which both Ophelia and Malcolm estimates to be 1 month.

Ophelia and Malcolm agree that the costs related to each action, for each scenario, are the same as estimated in Glisic and Adriaenssens (Glisic & Adriaenssens, 2010), and reported in Table 5.2.

Table 5.2. Costs per action and state (Glisic & Adriaenssens, 2010).

	Scenario U (no damage)	Scenario D (bridge fails)
Action DN (do nothing)	nothing happens you pay nothing	failure cost z_{F} = \$881,600
Action CB (close bridge)	1-month downtime cost z_{DT} = \$139,800	1-month downtime cost z_{DT} = \$139,800

However, Ophelia and Malcolm differ in their utility functions, which is the weight they apply to the possible economic losses. Ophelia is risk neutral, meaning that according to her a negative utility is linear with the incurred loss, as illustrated in Figure 5.3. Strictly speaking, a utility function is defined except for a multiplicative factor, therefore it should be expressed in an arbitrary unit sometime referred to as *util* (McConnell, 1966). Since Ophelia's utility is linear with loss, for the sake of clarity we will deliberately confuse negative utility with loss, and therefore we will measure Ophelia's utility in k$.

Unlike Ophelia, Malcolm is likely to behave risk adversely, i.e. his negative utility increases more than proportionally with the loss. We can describe mathematically the risk aversion classically defined in Arrow-Pratt theory (Pratt, 1964) (Arrow, 1965), where the level of risk aversion of an agent is encoded in the coefficient of Absolute Risk Aversion (ARA), defined as the rate of the second derivative (curvature) to the fist derivative (slope):

$$A(z) = \frac{U''(z)}{U'(z)}. \tag{5.10}$$

To state Malcolm's utility function, we can make the following assumptions:

- Malcolm's and Ophelia's reaction are virtually identical for a small amount of loss, while their way of weighting the losses departs for bigger losses.
- For small losses, therefore, the two-utility function may be confused, and we will adopt for Malcolm's the same conventional unit (call it equivalent k$) for measuring utility. Malcolm's utility function derivative for zero loss is equal to 1.
- We assume that Malcolm's utility has constant ARA; it is easily demonstrated that a function with constant ARA and unitary derivative at zero (Wakker, 2008) takes the form of an exponential:

$$^{(\mathrm{M})}U(z) = \frac{1 - e^{-z \cdot \theta}}{\theta}, \tag{5.11}$$

  where θ is the constant ARA coefficient: $A(z) = \theta$.
- To calibrate θ, we assume that for a loss equal to the failure cost, Malcolm's negative utility is twice that of Ophelia's. This results in a constant ARA coefficient $\theta = -1.425$ M^{-1}.

Using these assumptions, the resulting Malcolm's utility function is plotted in Figure 5.3.

We wish now to verify how the different utility functions affect the decision of the two a priori and a posteriori.

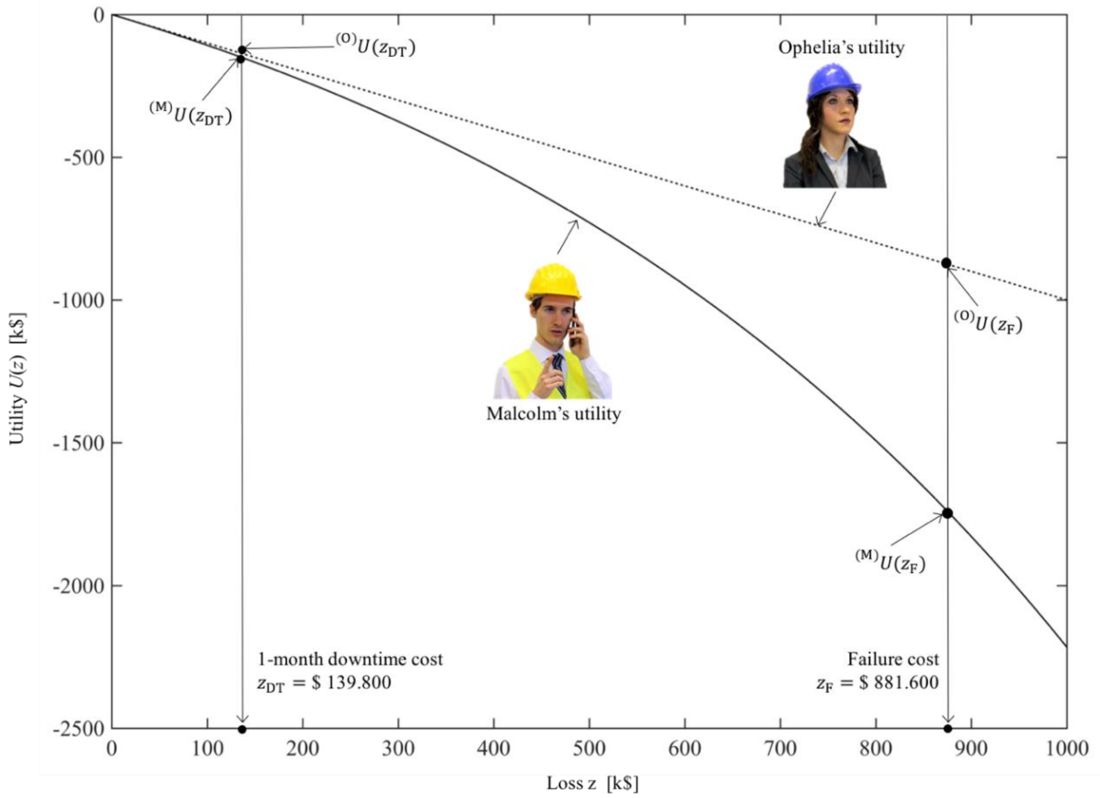


Figure 5.3. Representation of Ophelia's and Malcolm's utility functions.

5.4.4 Prior utility

Consider the case where Malcolm has no monitoring information. Based on his utility, Malcolm estimates the utilities involved in each action. Action CB depends only on the downtime cost z_{DT}, while action DN depends also on his estimate of the state of the bridge:

$$^{(M)}u_{DN} = {}^{(M)}U(z_F)\cdot P(D) = -528.883 \text{ k\$},\ {}^{(M)}u_{CB} = {}^{(M)}U(z_{DT}) = -154.940 \text{ k\$}. \quad (5.12)$$

Since the utility of action CB is clearly less negative than the utility of action DN, Malcolm would always choose to close the bridge after an incident if he has no better information from the monitoring system. Therefore, Malcolm's maximum expected utility without the monitoring system is $^{(M)}u = {}^{(M)}u_{CB}$ = -154.940 k$.

Now imagine Ophelia in charge of the decision; her prior utilities are different from Malcolm's and their values are somewhat closer:

$$^{(0)}u_{\mathrm{DN}} = {}^{(0)}U(z_{\mathrm{F}}) \cdot \mathrm{P(D)} = \text{-264.480 k\$}, \quad {}^{(0)}u_{\mathrm{CB}} = {}^{(0)}U(z_{\mathrm{DT}}) = \text{-139.800 k\$}, \quad (5.13)$$

but in the end, in this particular case, her optimal action would be again 'close the bridge'.

5.4.5 Posterior utility

Now imagine that the monitoring system is installed and let us go back to Malcolm. Since now Malcolm can rely on the monitoring reading, in this case the expected utility of an action is calculated using the posterior probability of damage pdf(D|ε) rather than the prior:

$$^{(\mathrm{M})}u_{CB|\varepsilon} = u(z_{\mathrm{DT}}), \qquad {}^{(\mathrm{M})}u_{\mathrm{DN}|\varepsilon} = u(z_{\mathrm{F}}) \cdot \mathrm{pdf(D}|\varepsilon). \qquad (5.14a,b)$$

Note that since the cost of closing the bridge is independent on the bridge state, the monitoring observation ε does not affect the posterior utility of closing the bridge (CB), which is always equal to -154.940 k$ as in the prior case. On the contrary, the expected utility of doing nothing (DN) does depend on the probability of having the bridge damaged, and this probability, in turn, depends on the monitoring observation through Eq. (5.14b). Malcolm's posterior expected utilities (i.e. after observing data from the monitoring system) for actions DN and CB are plotted in the graph of Figure 5.2(b) as functions of the observation ε. As a rational agent, Malcolm will always take the decision that maximizes his utility. For very small values of ε, suggesting a small probability of collapse, Malcolm's utility of DN is bigger than the utility of CB, and therefore Malcolm will keep the bridge open. Malcolm's utility of closing the bridge starts exceeding the utility of doing nothing above a threshold of strain of $^{(\mathrm{M})}\bar{\varepsilon}_{\mathrm{u}}$ = 170 με, and therefore Malcolm will always close the bridge above this threshold.

Note that this threshold is much smaller than the threshold $\bar{\varepsilon}_p$ whereby Malcolm would judge the damage more likely, so there is a range of values whereby Malcolm, in consideration of the possible consequences, will still prefer to close the bridge even if it is more likely the bridge is not damaged. Malcolm's maximum expected utility is plotted in bold in the graph of Figure 5.2(b).

Assume now that Ophelia is in charge of the decision. Since she weights the losses differently, her utility curves as functions of ε are different from Malcolm's, and are plotted in the graph of Figure 5.2(c). For the same reason, the threshold above which she would close the bridge, ${}^{(O)}\bar{\varepsilon}_u = 310\ \mu\varepsilon$, is different and much higher than Malcolm's, reflecting Ophelia's risk neutrality in contrast to Malcolm's risk aversion. Therefore, there is a range of values of measurements, from 170 με to 310 με, where the two decision makers, both rational, behave differently under the same information, simply because of their different level of risk aversion.

5.4.6 Posterior utility and Value of Information

In this scenario, Ophelia and Malcolm are both involved in the decision chain. Malcolm is the operational *manager* who decides whether or not to close the bridge in the occurrence of an incident. Ophelia is the *owner* who decides on the purchase of the monitoring system. This is illustrated as a decision tree in Figure 5.4. We seek the *VoI* as anticipated by Ophelia (she has to decide), which explicitly accounts from Malcolm reacting to the signals from the monitoring system.

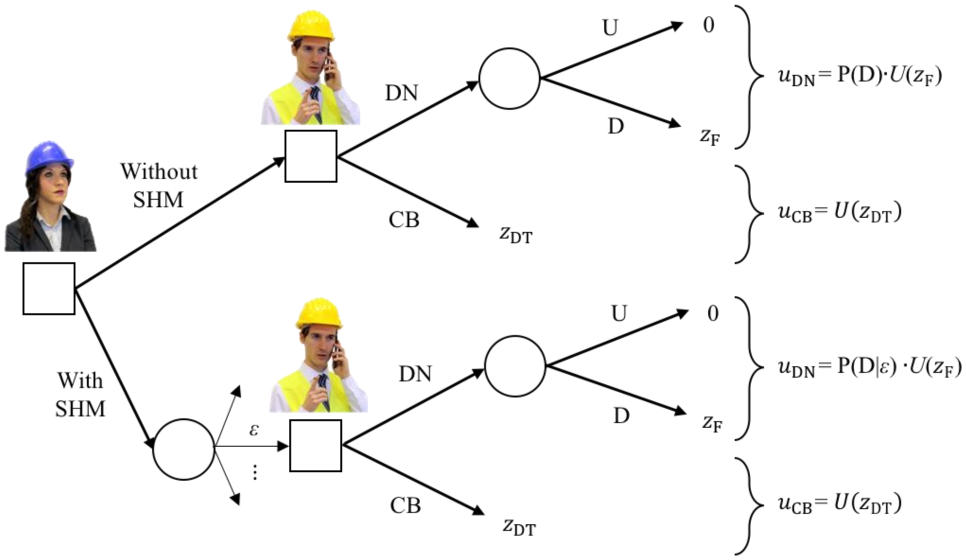


Figure 5.4. Decision tree for the Streicker Bridge case study.

Before attacking this problem, let us first see what happens if the decision chain was in the hands of a single individual. We start, for example, with Malcolm. His preposterior utility (i.e., the prior utility of operating the bridge with the monitoring system) can be calculated with the equation:

$$ {}^{(M)}u^{*} = \int_{D_\varepsilon} {}^{(M)}u\left\{ \underset{j}{argmax}\, {}^{(M)}u(a_j,\varepsilon) \right\} \cdot p(\varepsilon)\; d\varepsilon = -\,88.504 \text{ k\$}, \qquad (5.15) $$

where the index (M) indicates that all the utilities are calculated from Malcolm's perspective. Malcolm's *VoI* is simply the difference between the preposterior utility (i.e. the prior utility of operating the bridge with the monitoring system) and the prior utility (i.e. the utility of operating the bridge without the monitoring system):

$$ VoI = {}^{(M)}u^{*} - {}^{(M)}u = -88.505 \text{ k\$} + 154.940 \text{ k\$} = 66.435 \text{ k\$}. \qquad (5.16) $$

Note that the *VoI* is a utility, not an actual amount of money, and is measured in Malcolm's utility unit, which in our case is Malcolm's dollar-equivalent as defined above.

Now we can calculate the *VoI* from Ophelia's perspective, assuming that she takes decisions at any stage of the decision chain. In this case being Ophelia less risk adverse than Malcolm, her utilities will be ${}^{(O)}u^{*}$ = -84.600 k$ and ${}^{(O)}u$ = -139.800 k$, so eventually Ophelia's *VoI* would be:

$$ VoI = {}^{(O)}u^{*} - {}^{(O)}u = -84{,}600 \text{ k\$} + 139{,}800 \text{ k\$} = 55.200 \text{ k\$}. \qquad (5.17) $$

This practically means that, if Ophelia was in charge of all the decisions, she would be willing to spend up to 55.200 k$ for the information from the monitoring system.

In reality, Ophelia is only in charge of the purchase of the monitoring system, while the one who is going to use it is her colleague Malcolm. So, in taking her decision, Ophelia has to figure out how Malcolm is going to behave both with and without the monitoring system. In other words, we have to calculate the prior and preposterior utility from Ophelia's perspective, but conditional to the action that Malcolm will undertake.

For example, to calculate the prior (i.e. the utility of Ophelia of operating the bridge without the monitoring system, conditioned to Malcolm's actions) conditional utility, Ophelia thinks: *what will Malcolm do after an accident if no monitoring system is* *installed? I know Malcolm, and I know he will close the bridge right away (I would do the same, but that's irrelevant). My utility, if he closes the bridge, is:*

$$^{(O|M)}u = {}^{(O)}u\left\{\underset{j}{argmax}\ {}^{(M)}u(a_j)\right\} = {}^{(O)}u_{\mathrm{CB}} = \text{-139.800 k\$}, \tag{5.18}$$

which in this case is the same as the unconditional. *And what* – Ophelia continues to think – *would Malcolm do if a monitoring system was installed. I know that he would look at the strain ε and he would close the bridge if ε* >170 με *and keep the bridge open otherwise. I personally would NOT do the same, but that's it, I have to live with Malcolm's decision!*

The way Ophelia evaluates the utility on Malcolm's decisions is explained in Figure 5.2(d): her utilities for each possible Malcolm's choice are calculated using *her* utility function, hence all individual curves are identical to those of Figure 5.2(c), However, the threshold whereby she expects the bridge is closed is Malcolm's threshold, i.e. the same as in Figure 5.2(d). Ophelia's utility of Malcolm's choice is, for any value of ε:

$$^{(O)}a_{opt} = {}^{(O)}u\left\{\underset{j}{argmax}\ {}^{(M)}u(a_j,\varepsilon)\right\} = \text{-154.94 k\$}, \tag{5.19}$$

and therefore the preposterior utility conditional to Malcolm is:

$$^{(O|M)}u^{*} = \int_{D_\varepsilon} {}^{(O)}u\left\{\underset{j}{argmax}\ {}^{(M)}u(a_j,\varepsilon)\right\} \cdot p(\varepsilon)\ d\varepsilon = \text{-88.504 k\$}. \tag{5.20}$$

Eventually, Ophelia's *VoI, conditional* on Malcolm's decision, is:

$$^{(O|M)}VoI = {}^{(O|M)}u^{*} - {}^{(O|M)}u = \text{-88.505 k\$} + \text{139.800 k\$} = \text{51.295 k\$}. \tag{5.21}$$

Again, this quantity is the money Ophelia believe is worth spending on a monitoring system, having accepted that Malcolm, not her, is going to use it. The conditional $^{(O|M)}VoI = 51.495\ \text{k\$}$ is slightly lower than the unconditional $^{(O)}VoI = 55.200\ \text{k\$}$. Generally, it is clear from Ophelia perspective, that when Malcolm's decision is different from hers it is always suboptimal. Therefore, the conditional prior and pre-posteriors are always smaller than the corresponding unconditional: $^{(O|M)}u \leq {}^{(O)}u$,

${}^{(O|M)}u^* \leq {}^{(O)}u^*$. In the present example, Ophelia and Malcolm agree on what to do a priori ${}^{(O|M)}u = {}^{(O)}u$, the conditional ${}^{(O|M)}VoI$ is necessarily smaller than the conditional ${}^{(O)}VoI$. In simple words, Ophelia's rationale goes along these lines: *I can exploit the monitoring system better than Malcolm, therefore the benefit of the monitoring system would be greater if I was using the monitoring system rather than Malcolm.*

However, this is not the most general case. Assume for example the prior probability of damage P(D) is 10%: Ophelia's prior utility of action DN ${}^{(O)}u_{\mathrm{DN}}$= -88.160 k$, small enough for Ophelia to keep the bridge open; on the contrary Malcolm's prior utility${}^{(M)}u_{\mathrm{DN}}$ = -176.294 k$, is still big enough for Malcolm to close it. In this case the unconditional prior is much bigger than the conditional one, since Ophelia doesn't agree with Malcolm's choice, and the conditional ${}^{(O|M)}VoI$ = 103.670 k$ is much bigger than the unconditional ${}^{(O)}VoI$ = 53.217 k$, meaning that monitoring is much more useful in this case. We can almost hear Ophelia commenting: *This Malcolm can't make the right decision alone, hopefully some monitoring will help him! For sure a monitoring system is more useful to him rather than me!*

5.4.7 Negative Value of Information?

We noted above that in the unconditional case (i.e. when Ophelia is both *owner* and *manager*), the preposterior utility *u** is always greater or equal than the prior *u*, hence the *VoI* cannot be negative. In simpler words, if a monitoring system if offered to Ophelia at no cost, she has no reason not to accept it. Of course, if at any time Ophelia realizes that the monitoring system yields junk data, she can always decide to disregard this information, but she has no economic reason to refuse a priori to see the data ('*Take each man's censure, but reserve thy judgment*').

We also noted that in the unconditional case (i.e. when Ophelia is the *owner* but someone else, Malcolm, is the *manager* who decide based on the SHM data) there is no logical necessity whereby Ophelia's preposterior utility must be greater than her prior. So in principle we can always find a combination of prior probabilities and utility functions which ultimately yield a negative conditional *VoI*. We illustrate this concept with an example.

Imagine that Malcolm, instead of being risk adverse, is risk seeking. This is to say that his utility function is convex (i.e., with positive second derivative), as shown in Figure 5.5: for this exercise we can again assume an Arrow-Pratt's utility model, as in Eq. (5.11), but this time with a positive ARA coefficient $\theta = 5.234$ M^{-1}. Also, assume, both for Ophelia and Malcolm, a high prior probability of damage, say P(D) = 55%.

Using these assumptions, Ophelia's prior utilities for doing nothing (DN) and closing the bridge (CB) are ${}^{(O)}u_{DN}$= -484.88 k$ and ${}^{(O)}u_{CB}$= -139.800 k$ respectively, while Malcolm's are ${}^{(M)}u_{DN}$ = -108.660 k$ and ${}^{(M)}u_{CB}$ = -100.680 k$. For both, closing the bridge (CB) is the action that yields the maximum expected utility a priori: so they both agree that, without a monitoring system, the best thing to do is to close the bridge.

Their decisions start departing after receiving data from the monitoring system. Figure 5.6 shows how Ophelia's and Malcolm's decision models change based on the new assumptions. We note that:

- Because of the high prior risk of collapse, risk-neutral Ophelia is very conservative and thinks it is a good idea to close the bridge as soon as the elongation recorded is greater than ${}^{(O)}\bar{\varepsilon}_u = 70$ με;
- Risk-seeking Malcolm does not take a collapse so seriously and he would rather keep the bridge open unless the sensor reads an elongation greater than ${}^{(M)}\bar{\varepsilon}_u$ = 423 με.

So, there is a very wide range of values, from 70 με to 423 με, whereby Malcolm would keep the bridge open in disagreement with Ophelia, who believes this is a dangerous practice which can potentially result in a big loss. Based on these premises, Ophelia's conditional preposterior (i.e. Ophelia expected utility conditional to Malcolm's decision) is calculated, using Eq. (5.20), in ${}^{(O|M)}u^* = -150.362$ k$, and eventually her conditional value of information is:

$$ {}^{(O|M)}VoI = {}^{(O|M)}u^* - {}^{(O|M)}u = -150.362 \text{ k\$} + 139.800 \text{ k\$} = -10.562 \text{ k\$}. \quad (5.22) $$

Contrary to the example above, now the conditional value of information is negative, meaning that Ophelia's perceives the monitoring information as damaging: Ophelia thinks that, in observing the monitoring data, Malcolm may wrongly decide to keep the bridge open even when, in her opinion, it should be closed; she concludes that, after all, it is better not to install the monitoring system at all. In Ophelia's own words: *Malcolm is an irresponsible and should not use the monitoring system! I would rather pay money than letting him use the system!* Indeed, the negative value of information is exactly the amount of money Ophelia is willing to pay to prevent Malcolm using the monitoring system.

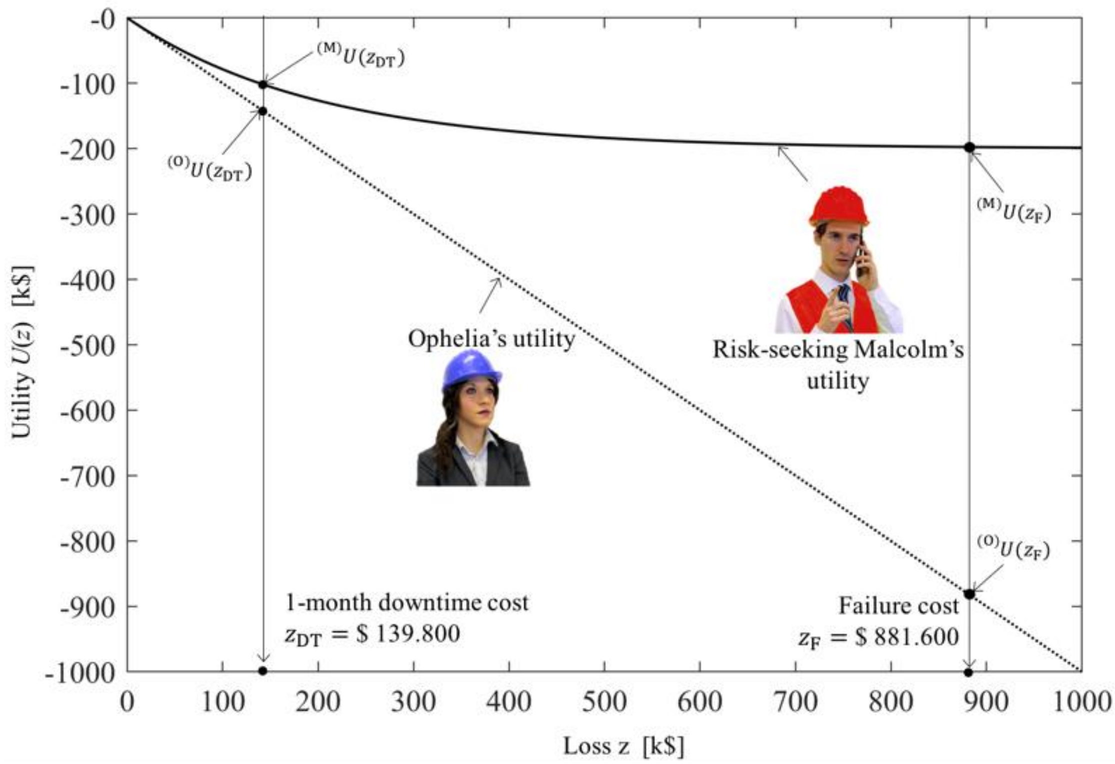


Figure 5.5. Representation of Ophelia's and risk-seeking Malcolm's utility functions.

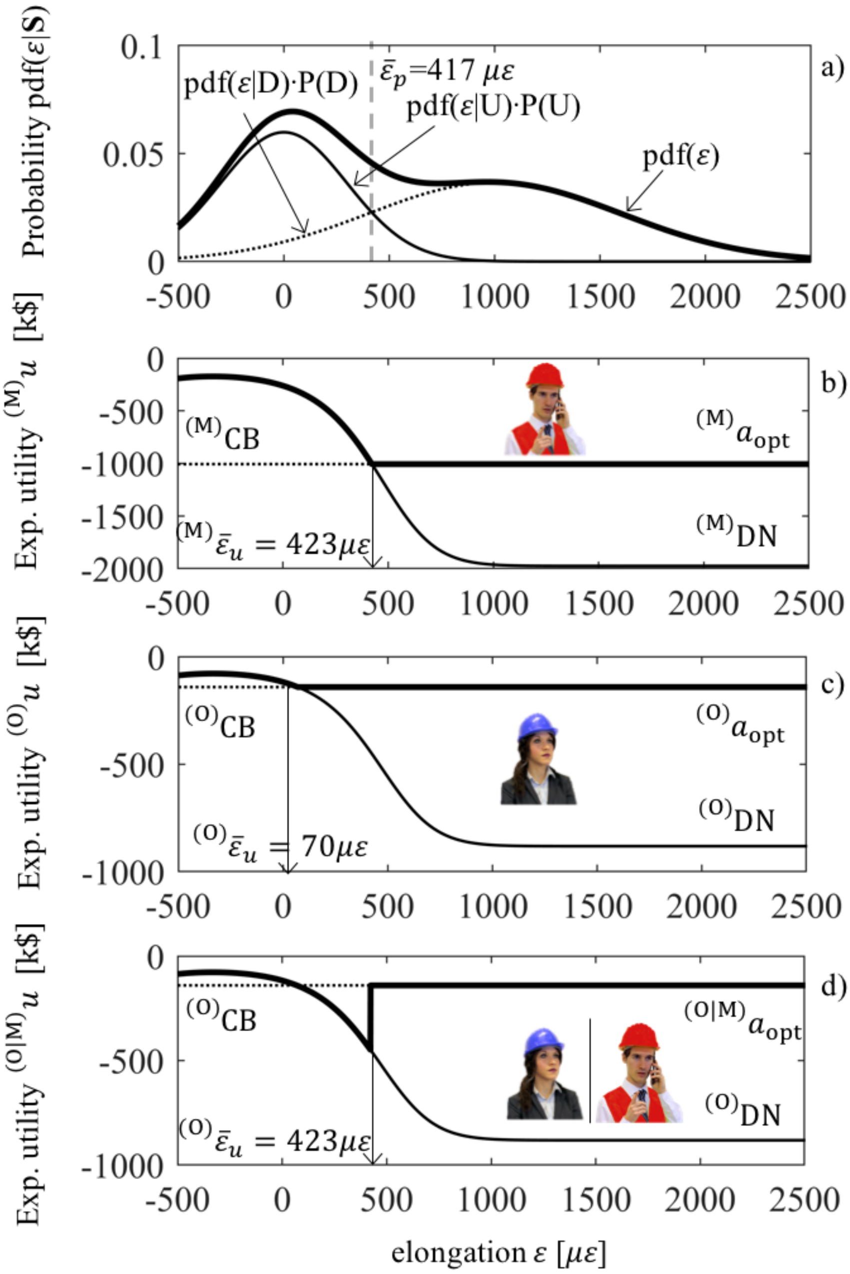


Figure 5.6. Representation of risk-seeking Malcolm's estimation of the state of the bridge a priori (a), risk-seeking Malcolm's decision model with monitoring data (b), Ophelia's decision model with monitoring data (c), Ophelia's decision model based on risk-seeking Malcolm's own (d).

5.5 Concluding remarks

The benefit of SHM can be quantified using the concept of Value of Information. This is the difference between the anticipated utilities of operating the structure with the monitoring system (the *preposterior* utility) and without the monitoring system (the *prior* utility). *Preposterior* utility, *Prior* utility and *Value of Information* are all subjective quantities: they depend on the particular background information and risk appetite of the individual in charge of the decision. In calculating the *VoI*, a commonly understood assumption is that the individual who decide on the installation of the monitoring system is the same rational agent who will later use it.

In the real world, these could be two separate subjects. We labelled conventionally *owner* the individual who decides on buying a monitoring system and *manager* the one who is going to use it, once the system has been installed. The two decision makers, even if both rational and exposed to the same background information, may still act differently because of their different appetites for risk.

We developed a formulation to properly evaluate the *VoI* from the owner perspective, when the manager is a different individual. The rationale of the formulation is that the owner, in evaluating the benefit of the monitoring system, must anticipate the way how the manager will actually react to the monitoring information. The calculation requires the definition of the owner's prior and preposterior utilities *conditional* to the manager anticipated behaviour. For convenience, we defined the *VoI conditional* in the case when the manager is not the owner, and *unconditional* when manager and owner coincide.

To illustrate how this framework works, we have evaluated a hypothetical *VoI* for the Streicker Bridge, a pedestrian bridge in Princeton University campus equipped with a fiber optic sensing system, assuming that two fictional characters, Ophelia the *owner* and Malcolm the *manager,* are involved in the decision chain. In the example, Malcolm is the manager who decide whether to keep the bridge open or close it, following to an incident that could potentially jeopardize its safety. Ophelia is the owner who decide whether to purchase a monitoring system to help Malcolm making the right decision in that event. We noted that:

- Seen from the owner's perspective, the choices of the manager are always suboptimal: Malcolm's decisions do not necessarily coincide with what Ophelia would have made in the same situation.

- In the prior situation (i.e. without SHM), the conditional utility (i.e. when the manager is not the owner) is always equal or lower than the unconditional one (i.e. when manager and owner coincide).
- The conditional (i.e. manager is not owner) *VoI* could be bigger or smaller than the unconditional (i.e. manager is owner); if Ophelia agree on how Malcolm makes decision without the monitoring system, the conditional value of monitoring is always lower than the unconditional.
- If Ophelia does not agree with Malcolm, the conditional value of information may be bigger than the unconditional: Ophelia would strongly support the purchase of the monitoring system in the hope it will help Malcolm to make the right decision.
- While the unconditional *VoI* is never negative, we demonstrate that under appropriate combination of prior information and utility functions, the conditional value of information could be negative: this can happen when Ophelia believe than the monitoring system can seriously mislead Malcolm's decision.

Acknowledgments

The case study reported in this paper is based on the Streicker Bridge monitoring project, developed with the kind help and collaboration of several professionals and companies. We would like to thank S. Hancock and his construction team, Turner Construction Company; R. Woodward and T. Zoli, HNTB Corporation; D. Lee and his team, A. G. Construction Corporation; S. Mancini and T. R. Wintermute, Vollers Excavating & Construction, Inc.; SMARTEC SA; and Micron Optics, Inc. In addition, the following personnel from Princeton University supported and helped the realization of the project: G. Gettelfinger, J. P. Wallace, M. Hersey, S. Weber, P. Prucnal, Y. Deng, M. Fok; faculty and staff of CEE; Princeton University students: D. Sigurdardottir, Y. Yao, M. Wachter, J. Hsu, G. Lederman, J. Chen, K. Liew, C. Chen, A. Halpern, D. Hubbell, M. Neal, D. Reynolds and D. Schiffner.

6. Quantifying the benefit of Structural Health Monitoring: can the Value of Information be negative?

Submitted for publication to *Structure and Infrastructure Engineering*
(Under revision)

Andrea Verzobio[a], Denise Bolognani[b], John Quigley[c], Daniele Zonta[d]

Position and affiliation:
[a]PhD Student, Department of Civil and Environmental Engineering, University of Strathclyde, UK
[b]Postdoc, Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy
[c]Professor, Department Management Science, University of Strathclyde, UK
[d]Professor, Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy

Summary of the paper

This paper is the continuation of the paper presented in chapter 5, where a rational method for quantifying the *VoI* when two different actors are involved in the decision chain is proposed. Since one of the conclusions of chapter 5 is that the *VoI* may become negative using this innovative method, in this contribution we propose a mathematical formulation which allows to assess when and under which circumstances it is possible to achieve a negative *VoI*, which is a clear proof of the relevance of the consequences that can be caused by distorted decisions. In the same way as in chapter 5, in order to verify the developed formulation, we apply it to the Streicker Bridge case study.

6.1 Introduction

Structural Health Monitoring (SHM) is a powerful tool for bridge management that support decisions concerning maintenance, reconstruction and repairs of assets through reducing uncertainties on the state of the structure. Uncertainty increases the likelihood of unwelcome outcomes such as neglecting necessary repairs while engaging in unnecessary ones. Such decision-making is challenging as it requires the decision maker to trade-off between anticipated risk and benefits to prioritize activities. The prioritization of activities will be determined in part by uncertainty and in part by the appetite for risk of the decision maker, which varies across individuals such that, given the same alternatives with the same state of uncertainty, two rational decision makers may take a different course of action. Through reduction in uncertainty such decision makers become more aligned in their choices. However, monitoring systems are costly and with limited budgets the anticipated value of the information provided towards the safety of the structure must be considered.

Although the utility of SHM has rarely been questioned in our community, very recently a few published papers (Thons & Faber, 2013) (Zonta, et al., 2014) have clarified how to evaluate it. The benefit of information is formally quantified by the so-called Value of Information (*VoI*), for its state of art see section 2.3.

In summary, the value of a SHM system can be simply defined as the difference between the benefit, or expected utility u_{pp}, of operating the structure *with* the monitoring system and the benefit, or expect utility u_0, of operating the structure *without* the system. Both u_{pp} and u_0 are expected utilities calculated *a priori*, i.e. before actually receiving any information from the monitoring system. While in u_0 we assume the knowledge of the manager is his/her a priori knowledge, u_{pp} is calculated assuming the decision maker has access to the monitoring information and is sometimes referred as to *preposterior utility*. In classical decision theory, one of the main assumptions is that all decisions concerning system installation and operation are taken by the same rational agent. In this case, it is easily proved that the *VoI* can only be positive, consistently with the principle that "*information can't hurt*", as first introduced by Cover and Thomas (Cover & Thomas, 2012) and later by Pozzi (Pozzi, et al., 2017).

However, there are several cases in the literature where a negative *VoI* is observed: we think that these cases, regardless of the field of application, can be classified in three different classes. The first one relates to non-cooperative games and decisions against nature: in summary, when agents compete against each other, an information can produce a negative value to some of them, precisely because we are in the area of competitive decisions. In the literature, we can find some examples principally in the field of financial markets, see for instance (Baiman, 1975) (Schredelseker, 2001) (Pfeifer, et al., 2009). The second case is instead about the presence of constraints in the decision process, which can lead the decision maker to take irrational decisions. A clear example is reported in (Pozzi, et al., 2017) (Pozzi, et al., 2020), where a system's maintenance agent has to blindly follow the prescriptions of codes and regulations, regardless their inherent rationality: in this case, the decision maker, in order to bypass these society's constraints, may find it convenient to avoid information. As demonstrated in the papers, the constraint extremely affects the *VoI*, which may consequently result negative. Finally, the last case relates to the presence of multiple rational decision makers that have to take decisions at different levels, which are somehow connected. This case is presented for instance in (Bolognani, et al., 2018), where two individuals are involved in the decision chain as regards a SHM-based decision process. In the paper it is proved that, because of the different appetite for risk of the two rational agents, the *VoI* may become negative. While in the literature it is easily demonstrated the reason why in the first two cases introduced above it is possible to find a negative *VoI*, respectively because of competitive decisions and because of irrational constrains, it is not so immediate to understand why it may happen in the third case, which is based only on rational behaviours. Consequently, in this contribution we will focus on this specific third case, which is also the one that mainly affects the field of our interest, i.e. SHM.

The case we analyse has been then introduced in (Bolognani, et al., 2018); we summarize the main assumptions in the following. Two decision makers are involved in the decision chain, and they have to take decisions at two different decision stages. Firstly, a decision is made on whether or not to buy and install the monitoring system on the structure; typically, this decision is carried out by a high-level manager, who we conventionally refer to as *owner*. The second stage concerns the day-to-day operation of the structure, which includes for example maintenance, repair, retrofit or enforcing traffic limitations, once the monitoring system is installed; if installed these decisions may be informed by the monitoring system. Typically, this decision is carried out by an engineer, who we will refer to as *manager*. The two agents are both rational and with the same background knowledge, they only differ in the weight they apply to the possible economic losses, meaning that they have different utility functions. Therefore, the two decision makers may differ in their choices under uncertainty: for instance, the owner needs to consider the operator's appetite for risk when deciding whether to install a monitoring system, as this will indicate how the system will influence the operator's decision-making and as such the value of this information. As proved in (Bolognani, et al., 2018), these assumptions can lead to a negative *VoI* because, even if the two agents have the same prior knowledge of the problem, their optimal actions can diverge after the installation of the monitoring system, due to their different attitudes towards risk. While in the paper it is showed that the *VoI* can become negative, it is not proved under which generic mathematical conditions this is true.

The aim of this contribution is to demonstrate under which conditions, e.g. appropriate combination of prior information and utility functions, it is possible to find a negative *VoI* in this specific case, by developing a mathematical formulation. In section 6.2 we start reviewing the formulation for the quantification of the *VoI* in a SHM-based decision process, both in the classical case of a singular rational agent and in the case of two different individuals, needed for a better understanding of how it is possible to achieve a negative *VoI* only in the second case. Next, in section 6.3, we introduce a prototype decision problem and we develop a mathematical formulation to investigate under which circumstances the *VoI* becomes negative. Finally, in section 6.4, to illustrate how this framework works, we apply it to the same decision problem reported in (Zonta, et al., 2014) and (Bolognani, et al., 2018), i.e. the Streicker Bridge case study: it is a pedestrian bridge at Princeton University campus equipped with a continuous monitoring system. Some concluding remarks are presented at the end of the article.

6.2 Value of information for SHM-based decision

In this section, we review the concept of *VoI* for SHM-based decision problems, following a similar path as in (Zonta, et al., 2014) and (Bolognani, et al., 2018). The assumptions and the framework of a rational SHM-based decision process have already been presented in section 2.1, we summarize in the following the formulation of the *VoI* (using a different notation that is necessary for the development of this chapter).

In the classical formulation of *VoI* (Zonta, et al., 2014), which we will refer to as *unconditional*, i.e. assuming all decisions concerning system installation and operation taken by the same rational agent, the *VoI* of a monitoring system is simply the difference between the expected utility with the monitoring system u_{pp}, and the corresponding utility without the monitoring system u_0:

$$VoI = u_{pp} - u_0. \tag{6.1}$$

In the case of a structure not equipped with a monitoring system, the rational manager decides without accessing any SHM data, and they will choose the action a that maximize the expected utility u_0. Consequently, the utility without monitoring, also called *prior utility*, is calculated as follows:

$$u_0 = \max_i u(a_i), \qquad a_{\text{opt}} = \arg\max_i u(a_i), \tag{6.2a,b}$$

where a_{opt} is the action which carries the maximum expected utility u. Conversely, if a monitoring system is installed and the data are available for the agent, the monitoring observation $\mathbf{y}$ affects the state knowledge, and therefore indirectly their decisions. In this case, the expected utility u_{pp}, also called *preposterior utility*, can be derived from the posterior expected utility $u(\mathbf{y})$ by marginalizing out the variable $\mathbf{y}$ (Zonta, et al., 2014) (Cappello, et al., 2016):

$$u_{pp} = \mathrm{E}_{\mathbf{y}}\left[\max_i u(a_i,\mathbf{y})\right] = \int_{D\mathbf{y}} \max_i u(a_i,\mathbf{y}) \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y}, \tag{6.3}$$

where $\mathrm{E}_{\mathbf{y}}$ is the expected value operator of $\mathbf{y}$**,** while distribution $\mathrm{p}(\mathbf{y})$ is the so-called evidence in classical Bayesian theory (Sivia & Skilling, 2006). In conclusion, the *unconditional VoI* of a monitoring system is calculated as follows:

$$VoI = u_{pp} - u_0 = \int_{D\mathbf{y}} \max_i u(a_i,\mathbf{y}) \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y} - \max_i u(a_i). \tag{6.4}$$

In other words, the *VoI* is the difference between the expected maximum utility and the maximum expected utility. It is easily mathematically verified that u_{pp} is always greater than or equal to u_0, and therefore the *VoI* as formulated above can only be positive. This is to say that under these assumptions we would never prefer not to have the data if they were available, which is consistent with the principle "*information can't hurt*" (Cover & Thomas, 2012).

Bolognani et al. (Bolognani, et al., 2018) have investigated a variant of the decision problem above where two different rational individuals, rather than one, are involved in the decision chain. In particular, there is an *owner* who decides whether or not to install a monitoring system, and a *manager* who decides which is the optimal action once the monitoring system is installed or not. Therefore, all utilities are from the *owner* perspective, but should be evaluated accounting for the action that the *manager*, not the owner, is expected to choose. The prior expected utility of Eq. (6.2), in the case of a structure without the monitoring system, changes to:

$$u_0 = u\left(a^*_{\mathrm{opt}}\right) = u\left\{\arg\max_i u^*(a_i)\right\}, \tag{6.5}$$

where the star * indicates the optimal action or the utility from the *manager* perspective. Similarly, the expected utility of the owner in the expectation of what the manager would decide if a monitoring system was installed turns into:

$$u_{pp} = \int_{D\mathbf{y}} u\left\{\arg\max_i u^*(a_i,\mathbf{y})\right\} \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y}. \tag{6.6}$$

The *VoI* of a monitoring system calculated under these assumptions is labelled *conditional*, to remind that the utility of the owner is conditional to the action chosen upstream by the manager, and reads (Bolognani, et al., 2018):

$$\begin{aligned} VoI = u_{pp} - u_0 \\ &= \int_{D\mathbf{y}} u\left\{\arg\max_i u^*(a_i,\mathbf{y})\right\} \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y} \\ &- u\left\{\arg\max_i u^*(a_i)\right\}. \end{aligned} \tag{6.7}$$

Table 6.1 summarizes the *unconditional* and *conditional* formulations. As observed in (Bolognani, et al., 2018), in the *conditional* case it is no longer automatically verified that the owner's *preposterior* utility u_{pp} is always greater than or equal to the *prior* utility u_0. Therefore, unlike the *unconditional* case, we could find a combination of prior probabilities, likelihood and utility functions which yield a negative *conditional VoI*. The aim of this contribution is then to demonstrate under which mathematical conditions it is possible to find a negative *VoI*.

Table 6.1. Formulation of *VoI* for SHM in the *unconditional* and *conditional* case.

	Unconditional formulation Manager = Owner	*Conditional* formulation Manager ≠ Owner
Prior utility without monitoring u_0	$\max_i u(a_i)$	$u\left\{\arg\max_i u^*(a_i)\right\}$
Preposterior utility with monitoring u_{pp}	$\int_{D\mathbf{y}} \max_i u(a_i,\mathbf{y}) \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y}$	$\int_{D\mathbf{y}} u\left\{\arg\max_i u^*(a_i,\mathbf{y})\right\} \cdot \mathrm{p}(\mathbf{y})\ \mathrm{d}\mathbf{y}$

6.3 When does the *VoI* become negative?

In the previous section, we have introduced the concept of conditional value of information and we have noticed that, ‘under certain conditions’, it could become negative. In this section, we wish to clarify which exactly are the conditions whereby the *conditional VoI* becomes negative. To do so, we will focus our analysis on a 2-state 2-action prototype decision problem, graphically illustrated in the decision tree of Figure 6.1, which is representative of a number of binary decision setting that can be found in the literature (Raiffa & Schlaifer, 1961) (Parmigiani & Inoue, 2009). Particularly, we make the following assumptions:

- The structure can be in one of two mutually exclusive and exhaustive states S_1 and S_2 (e.g.: S_1 = *the bridge is damaged*; S_2 = *the bridge is not damaged*).
- The decision maker can choose between two alternative decisions a_1 and a_2 (e.g.: a_1 = *do nothing*; a_2 = *close the bridge*).
- Both actions may have consequences, depending on the (uncertain) state: we indicate with $\mathbf{z}(a_i, S_j)$ the set of consequences of action a_i on the realization of state S_j. Both manager and owner are equally aware of these consequences.

- We indicate with $U(\mathbf{z})$ the utility function of the *owner*, where the argument $\mathbf{z}$ is a particular set of consequences. To simplify the notation, we label $U_{ij} = U(\mathbf{z}(a_i, S_j))$ the utility of the consequences of action a_i on the realization of state S_j.
- Regardless the complexity of the monitoring system, its ultimate output is represented by a single parameter y, defined in the domain $[0, y_{\max}]$. Parameter y could be, for instance, a compensated measurement, or a synthetic damage index calculated using the full dataset recorded to date. The *manager* makes decision solely based on parameter y.
- The two agents, *owner* and *manager*, have the same prior knowledge of the problem, i.e. their prior probability $\mathrm{P}(S_j)$ of being in one of the two states is identical. They also interpret the data from the SHM system using the same interpretation model, which is encoded in the two likelihood functions $\mathrm{p}(y|S_j)$. They are both rational and judge consistently with Bayes' rule: therefore, their judgement on the state of the structure, prior or posterior, is always identical.
- Similarly, the two agents decide rationally consistently with EUT. However, their utility functions are generally different, thus their decisions, in the same situation, could differ.
- Parameter y is defined in such a way that the values of y whereby the owner chooses an action rather than the other are separated by a single threshold $\bar{y}$. Without losing generality, we can assume here that when $y < \bar{y}$ the owner chooses action a_1. The same applies to the manager, except that their threshold, labelled $\bar{y}^*$, could be different.

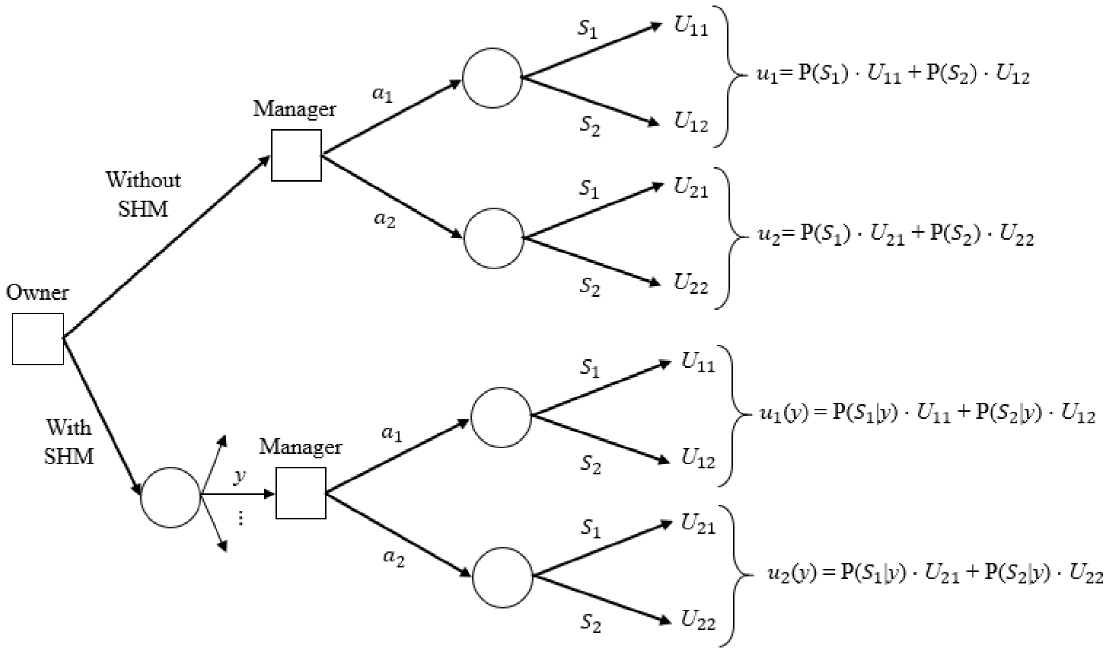


Figure 6.1. Decision tree of the prototype decision problem.

With the above assumptions, we will establish the conditions whereby the *VoI* becomes negative. Before tackling the problem in full, we start with the assumption that $U_{12} = U_{21} = 0$ and $U_{11} < U_{22} < 0$. This simplifying hypothesis makes the solution much more intuitive and easier to understand and will be released at the end of this section. To further help picturing the problem, imagine we are dealing with a bridge that may be in damaged, i.e. S_1, or undamaged, i.e. S_2, condition. The manager can decide to keep the bridge open or close it. If the bridge is left open and is damaged, the bridge fails producing a negative utility U_{11}. If the bridge is unnecessarily closed when not damaged, the manager is sanctioned with a penalty U_{22}. The loss for a failure is in absolute value much greater than the penalty for closing the bridge without necessity, i.e. $|U_{11}| > |U_{22}|$ or $U_{11} < U_{22}$, reminding that we are dealing with negative utilities. More generally, this is the prototype of any problem where an agent is faced with a binary decision, and each decision can be right or wrong depending on the unknown state. If the agent makes the right choice, nothing happens, otherwise they are sanctioned with a penalty. This situation is illustrated in the decision tree of Figure 6.2.

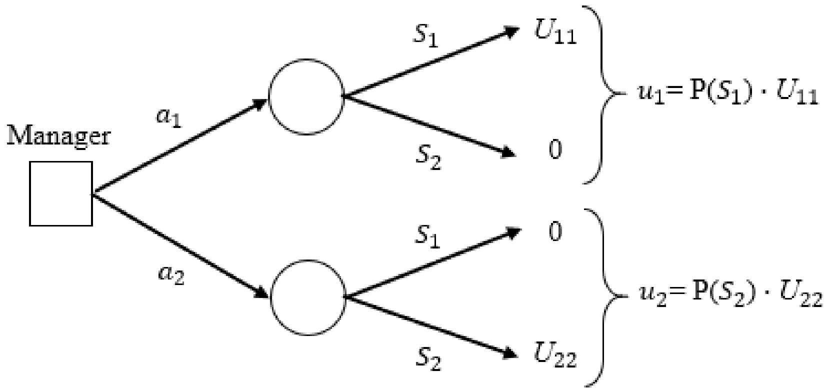


Figure 6.2. Decision tree of a simplified version of the prototype decision problem.

6.3.1 Decision a priori

We analyse the problem of decision a priori from the *owner* perspective. The owner will favour action a_2 over a_1 when the prior expected utility u_2 is greater than the prior expected utility u_1, i.e.:

$$u_1 < u_2. \tag{6.8}$$

Recall $u_1 = \mathrm{P}(S_1)U_{11}$, $u_2 = \mathrm{P}(S_2)U_{22}$, and both utilities are negative, we can rewrite the inequality of Eq. (6.8) as:

$$R = \frac{\mathrm{P}(S_2)}{\mathrm{P}(S_1)} \frac{U_{22}}{U_{11}} < 1, \tag{6.9}$$

where R is a discriminant ratio which expresses the optimal action a priori from the owner perspective. We observe that, by definition, $R = 1$ corresponds to the indifference in the choice a priori between the two actions a_1 and a_2, i.e. $u_1 = u_2$, while it is preferred to choose action a_1 if $R > 1$, i.e. $u_1 > u_2$, or action a_2 if $R < 1$, i.e. $u_1 < u_2$. It is convenient to express the discriminant R as:

$$R = \frac{r}{q}, \tag{6.10}$$

where:

$$q = \frac{\mathrm{P}(S_1)}{\mathrm{P}(S_2)}, \qquad r = \frac{U_{22}}{U_{11}}. \tag{6.11a,b}$$

Index q is the prior odds of state S_1 respect to S_2, while index r is an indicator of the subjective risk appetite of the decision maker: the more the agent is risk seeking, the bigger is index r. The risk seeking index r is subjective and changes with the actor. So, for the manager in general we may have a different value r^* and therefore a different value of the discriminant ratio R^* a priori. We assume in the following that manager and owner agree on that the optimal action a priori is, for example, a_2, thus both ratios R and R^* are smaller than one.

6.3.2 Decision a posteriori

We start the analysis a posteriori from the owner's perspective. After observing a particular output y from the monitoring system, the owner updates their knowledge of the structural state from prior $\mathrm{P}(S_j)$ to posterior $\mathrm{P}(S_j|y)$. Similar to the prior case, the owner decides a posteriori by comparing the expected utilities of the two actions a posteriori, i.e. $u_1(y) = \mathrm{P}(S_1|y)U_{11}$ and $u_2(y) = \mathrm{P}(S_2|y)U_{22}$, as shown in Figure 6.3(d) in the example case of Gaussian likelihood distributions (Figure 6.3(a), 6.3(b), 6.3(c)). The threshold $\bar{y}$ is the value of y for which a posteriori the expected utilities are the same, i.e. $u_1(\bar{y}) = u_2(\bar{y})$, which we express in the following:

$$\bar{y}\colon\ u_1(\bar{y}) = u_2(\bar{y}). \tag{6.12}$$

Recall we have assumed the owner's choice a priori is action a_2, and that y is defined in such a way that the optimal action a posteriori is a_1 for $y < \bar{y}$. Therefore, a posteriori the owner will change their decision when $y < \bar{y}$ and confirm the prior decision otherwise. Using Bayes' theorem, Eq. (6.12) can be rearranged in the form:

$$\bar{y}:\ \frac{\mathrm{p}(\bar{y}|S_1)}{\mathrm{p}(\bar{y}|S_2)} = \frac{\mathrm{P}(S_2)}{\mathrm{P}(S_1)}\frac{U_{22}}{U_{11}}. \tag{6.13}$$

We immediately recognize that the right-hand term of Eq. (6.13) is the same ratio R a priori introduced in Eq. (6.9). Further, we define the function $g(y)$ as the ratio between the likelihoods of the two states:

$$g(y) = \frac{\mathrm{p}(y|S_1)}{\mathrm{p}(y|S_2)}. \tag{6.14}$$

Therefore, the owner threshold $\bar{y}$ is determined by the following simple equation:

$$\bar{y}: \; g(\bar{y}) = R. \tag{6.15}$$

As such function g, which depends only on the likelihood distributions, equals R when evaluated in the threshold, as shown in Figure 6.3(e). We observe that the threshold effectively depends on ratio R, which in turn depends on the risk apatite of the owner.

In a similar manner, the manager threshold $\bar{y}^*$ is such that $g(\bar{y}^*) = R^*$, as illustrated again in Figure 6.3(e). The manager threshold $\bar{y}^*$ can be bigger or smaller than the owner threshold $\bar{y}$ depending on whether the manager is respectively more or less risk seeking than the owner. Because the two thresholds in general do not coincide, we can have essentially three situations a posteriori following to a monitoring observation y:

- If observation y is smaller than the two thresholds, both manager and owner agree to change their decision to a_1.
- If observation y is bigger than the two threshold, manager and owner agree to keep the prior decision a_2.
- if observation y is included between the two thresholds, manager and owner disagree on the decision to be made.

6.3.3 Preposterior analysis

We define $\Delta u(y) = u_1(y) - u_2(y)$ the utility gain resulting from changing decision a posteriori. Evidently, changing their mind is convenient to the owner when the monitoring system yield value smaller than their threshold. The *conditional VoI*, introduced in Eq. (6.7), based on the developed assumptions can be calculate as follows:

$$VoI = \int_0^{\bar{y}^*} \Delta u(y) \cdot \mathrm{p}(y)\ \mathrm{d}y, \tag{6.16}$$

where $\Delta u(y) \cdot \mathrm{p}(y)$ can be seen as an expected utility density function (EUDF), plotted in Figure 6.3(f). The figure shows that the *VoI* is effectively the area under the expected utility function up to the threshold of the manager $\bar{y}^*$. We also observe that:

- Because the EUDF is greater than zero under the threshold of the owner $\bar{y}$, evidently the *VoI* is maximum and always positive when the two thresholds coincide; this is the case of the unconditional value of information *uVoI*.
- When the manager is less risk seeking than the owner, i.e. $\bar{y}^* < \bar{y}$, the *conditional VoI* is smaller than the unconditional, but can never be negative – could be at least zero when $\bar{y}^* = 0$.
- When the manager is more risk seeking than the owner, i.e. $\bar{y}^* > \bar{y}$, the negative integral of the EUDF between the two thresholds can be interpreted as a *Loss for Disagreement (LfD)* of the two decision makers. If the *LfD* equals the *uVoI*, then the *conditional VoI* results negative.

In order to better clarify the condition whereby the *VoI* is negative, note that in our particular case the EUDF can be written as:

$$\begin{aligned} \Delta u(y) \cdot \mathrm{p}(y) &= (\mathrm{P}(S_1|y)U_{11} - \mathrm{P}(S_2|y)U_{22}) \cdot \mathrm{p}(y) = \\ &= \mathrm{P}(y|S_1)\mathrm{P}(S_1)U_{11} - \mathrm{P}(y|S_2)\mathrm{P}(S_2)U_{22}. \end{aligned} \tag{6.17}$$

Therefore, the *conditional VoI* becomes:

$$VoI = \int_0^{\bar{y}^*} \mathrm{p}(y|S_1)\mathrm{P}(S_1)U_{11}\ \mathrm{d}y - \int_0^{\bar{y}^*} \mathrm{p}(y|S_2)\mathrm{P}(S_2)U_{22}\ \mathrm{d}y. \tag{6.18}$$

The *VoI* is equal to zero either if $\bar{y}^* = 0$ or:

$$\bar{y}^*: \quad \frac{\int_0^{\bar{y}^*} \mathrm{p}(y|S_1)\ \mathrm{d}y}{\int_0^{\bar{y}^*} \mathrm{p}(y|S_2)\ \mathrm{d}y} = \frac{\mathrm{P}(S_2)}{\mathrm{P}(S_1)} \frac{U_{22}}{U_{11}}. \tag{6.19}$$

Notice that the format of Eq. (6.19) is strikingly similar to Eq. (6.13), with the only difference that the left-hand term is the ratio between the cumulative distributions of the two likelihoods, rather than the two mass density functions. Therefore, we define another function $G(y)$, as the ratio between the cumulative distributions of the two likelihoods:

$$G(\mathrm{y}) = \frac{\mathrm{F}(y|S_1)}{\mathrm{F}(y|S_2)} = \frac{\int_0^{y} \mathrm{p}(y|S_1)\,\mathrm{d}y}{\int_0^{\bar{y}} \mathrm{p}(y|S_2)\,\mathrm{d}y}. \tag{6.20}$$

Consequently, the minimum manager threshold $\bar{y}^*$ that makes the *VoI* negative is determined by the following simple equation:

$$\bar{y}^*:\quad G(\bar{y}^*) = R. \tag{6.21}$$

This outcome, together with Eq. (6.15), explicates how the threshold $\bar{y}$ and the index R of the manager, i.e. $\bar{y}^*$ and R^*, must be, in comparison to the ones of the owner, in order to achieve a null *conditional VoI*:

$$\bar{y}^* = G^{-1}\big(g(\bar{y})\big), \qquad R^* = g\left(G^{-1}(R)\right). \tag{6.22a,b}$$

In other words, in order to have a null *VoI*, the ratio between the thresholds and between indexes r of the two agents are:

$$\frac{r^*}{r} = \frac{R^*}{R} = \frac{g\left(G^{-1}(R)\right)}{R}, \qquad \frac{\bar{y}^*}{\bar{y}} = \frac{G^{-1}(R)}{g^{-1}(R)}. \tag{6.23a,b}$$

6.3.4 Generalization and summary

We have derived these formulations under the very stringent assumption that $U_{12} = U_{21} = 0$. Now let us release this assumption: the condition whereby the owner will favour action a_2 over a_1, which was previously encoded into Eq. (6.9), now reads:

$$\frac{\mathrm{P}(S_2)}{\mathrm{P}(S_1)}\frac{U_{22} - U_{12}}{U_{11} - U_{21}} < 1, \tag{6.24}$$

so it suffices to redefine the risk seeking factor r as:

$$r = \frac{U_{22} - U_{12}}{U_{11} - U_{21}}, \tag{6.25}$$

and the rest of the formulation is completely identical. Index r, and consequently also R, is an indicator about the risk appetite of the decision maker based on the definition of the four utilities: even in this general case, the more the agent is risk seeking, the bigger is index r.

In summary, the necessary and sufficient condition to have a negative *VoI* is:

$$R^* > g\left(G^{-1}(R)\right), \qquad \bar{y}^* > G^{-1}\left(g(\bar{y})\right), \tag{6.26a,b}$$

where the ratio R depends on the prior odds q and on the risk seeking ratio r, defined in Eq. (6.25). We can conclude that, in order to achieve a negative *VoI*, the manager has to be more risk seeking than the owner, i.e. $r^* > r$, so that their threshold $\bar{y}^*$ is bigger of an amount that only depends on the choice of the likelihood distributions.

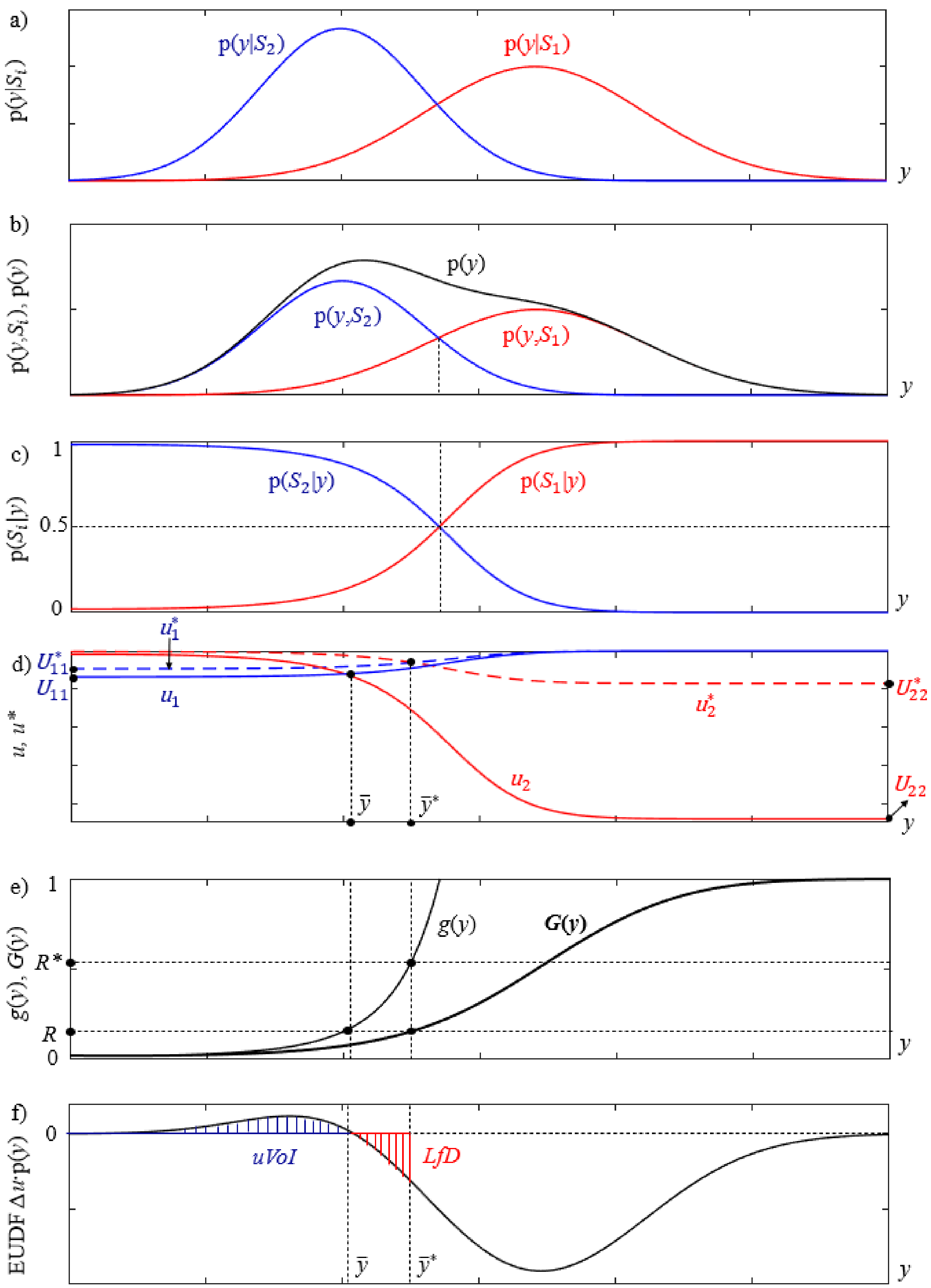


Figure 6.3. Graphical representation of how the *conditional VoI* may become negative: likelihood distributions (a), joint probabilities and evidence (b), posterior probabilities (c), expected utilities (d), indexes g and G (e), EUDF (f).

6.3.5 Notable case

Eq. (6.26a) shows that the ratio R^* that produces a negative *VoI* depends only on the choice of the likelihood distributions and on the owner ratio R. In order to calculate R^*, we need to express the functions $g(y)$ and $G(y)$, and to calculate their inverse functions $g^{-1}(R)$ and $G^{-1}(R)$. Unfortunately, in most cases it is not easy, and sometime not even possible, to express the inverse functions in closed form. A notable exception is when we describe the likelihood distributions with polynomial functions, as follows:

$$\mathrm{p}(y|S_1) = (\mathrm{n}+1)\, y^{\mathrm{n}}, \quad \mathrm{p}(y|S_2) = \frac{\mathrm{n}+1}{\mathrm{n}}(1 - y^{\mathrm{n}}), \quad \text{with } y \in [0,1]. \tag{6.27}$$

These likelihoods are presented in Figure 6.4, as an example, for the polynomial degree n varying from 1 to 4.

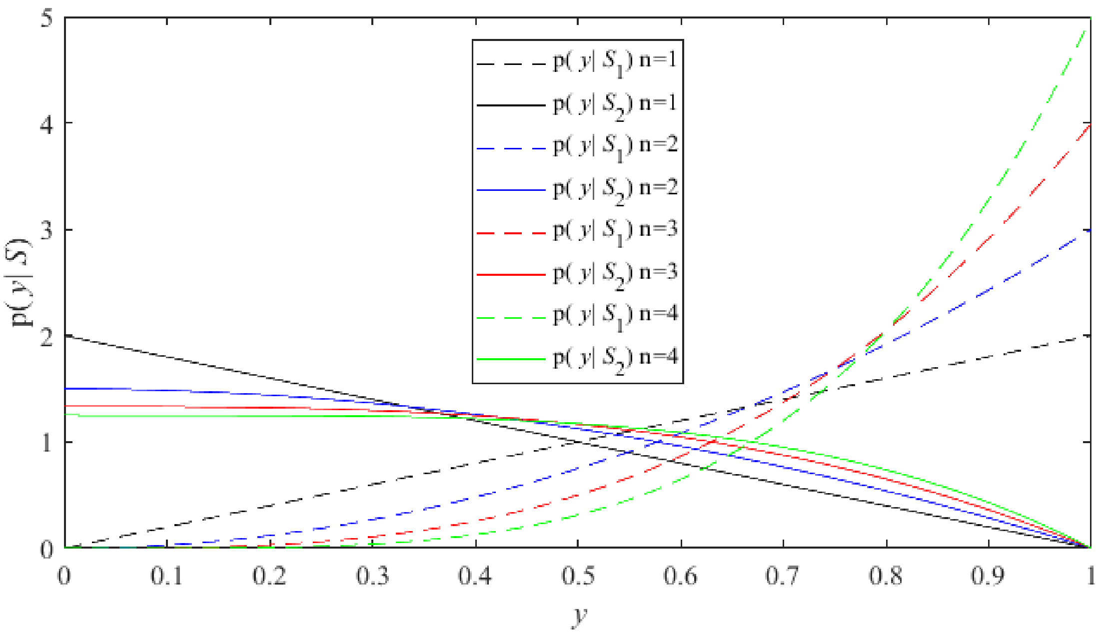


Figure 6.4. Likelihood distributions according to the polynomial degree n.

In this case, functions $g(y)$, $G(y)$ and their inverse are:

$$g(y) = \mathrm{n}\,\frac{y^{\mathrm{n}}}{1 - y^{\mathrm{n}}}, \qquad g^{-1}(R) = \sqrt[\mathrm{n}]{\frac{R}{\mathrm{n}+R}}. \tag{6.28a,b}$$

$$G(y) = \mathrm{n}\,\frac{y^{\mathrm{n}}}{(\mathrm{n}+1) - y^{\mathrm{n}}}, \qquad G^{-1}(R) = \sqrt[\mathrm{n}]{\frac{R(\mathrm{n}+1)}{\mathrm{n}+R}}. \tag{6.29a,b}$$

An interesting feature of this class of likelihood functions is that the rate between the manager and owner threshold is constant and equal to:

$$\frac{\bar{y}^*}{\bar{y}} = \sqrt[\mathrm{n}]{\mathrm{n}+1}. \tag{6.30}$$

This means that, to achieve a null *conditional VoI*, the threshold of the manager has to be bigger than the one of the owner of a quantity that depends only on the polynomial degree n. For instance, in the linear case, i.e. n = 1, it results that $\bar{y}^*$ has to be double of $\bar{y}$. Table 6.2 reports the results for n from 1 to 4.

Table 6.2. How the ratio between $\bar{y}^*$ and $\bar{y}$ varies according to n to achieve *VoI* = 0.

n	1	2	3	4
$\frac{\bar{y}^*}{\bar{y}}$	2	$\sqrt{3}$	$\sqrt[4]{3}$	$\sqrt[5]{4}$

It is evident that, as n increases, it decreases how much $\bar{y}^*$ has to be bigger than $\bar{y}$ in order to have $^{(O|M)}VoI = 0$, and consequently a negative *conditional VoI*. We can easily understand the reason of this outcome by analysing it graphically: in the linear case, presented in Figure 6.5(a), the threshold of the manager has to be clearly double of the one of the owner, because *uVoI* and *LfD* are two triangles. Conversely, with n > 1, as for example Figure 6.5(b) shows for n = 2, it is evident that, in order to have the area of *uVoI* and the one of *LfD* equal, $\bar{y}^*$ has to be bigger than $\bar{y}$, but less than double.

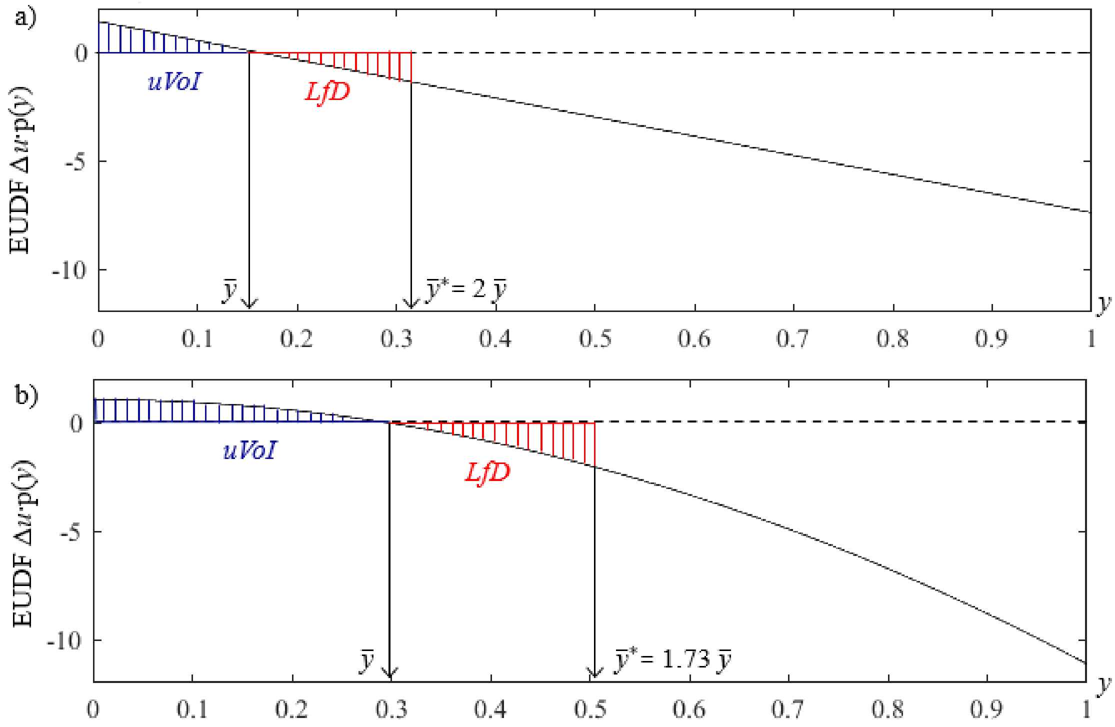


Figure 6.5. Expected utility density function (EUDF) for n = 1 (a), and n = 2 (b).

In addition, we have already anticipated that a bigger threshold corresponds to a bigger index R, meaning that the manager has to be more risk seeking than the owner in order to have a null *conditional VoI*, and consequently a negative one. We can verify this sentence by developing Eq. (6.23a), in this case of polynomial likelihood distributions:

$$\frac{r^*}{r} = \frac{R^*}{R} = \frac{g\left(G^{-1}(R)\right)}{R} = \frac{\mathrm{n}+1}{1-R} = \frac{\mathrm{n}+1}{1-r/q}. \tag{6.31}$$

This means that, in order to have a null *conditional VoI*, the manager has to be more risk seeking than the owner, i.e. $r^* > r$, by an amount that increases as the polynomial degree n rises, and which depends also on r itself. In conclusion, while it is clear that in real-life the likelihood distributions may have various different shapes, e.g. Gaussian as in the case study of section 6.4, defining them with polynomial functions allows us to achieve results in closed form, which is useful to understand better the practical meaning of the developed formulation.

6.4 The Streicker Bridge case study

To illustrate how the developed framework works, we consider the same case study as in (Zonta, et al., 2014) (Bolognani, et al., 2018), i.e. the Streicker bridge, since it respects all the assumptions introduced in the previous sections. The bridge has already been introduced in section 5.4.

6.4.1 Introduction of the SHM-decision problem

The SHM-based decision problem, the main assumptions and the individuals involved are the same as in (Bolognani, et al., 2018). The bridge is managed by two fictitious agents with distinct roles:

- Ophelia (O) is the *owner* responsible for Princeton's estate, who has to decide on whether or not to install the monitoring system; she is Malcolm's supervisor.
- Malcom (M) is the *manager* responsible for the bridge operation and maintenance, who has to take decisions on the state of the bridge based on monitoring data.

They are both rational individuals and they have the same background knowledge, they only differ in the way how to weight the seriousness of the consequences of a failure. They are concerned by a single specific scenario: a truck, driving along Washington road, could collide with the steel arch of the bridge. After the incident, the bridge will be in one of the following two states:

- S_1 = *damaged* (D), i.e. the bridge is still standing but has suffered major damage, and there is a change of collapse of the entire bridge.
- S_2 = *undamaged* (U), i.e. the structure has either no damage or some minor damage.

According to Malcolm and Ophelia, the two states are mutually exclusive and exhaustive, i.e. P(D) + P(U) = 1. We assume that they focus on the sensor installed at the bottom of the middle cross-section between P6 and P7 (called Sensor P6-7d, see Figure 5.1(c)). The output of the monitoring system is then represented by the strain ε of this specific fiber optic sensor. We can also assume that the two agents use the same interpretation model, i.e. they interpret identically the data from the monitoring system, as we will present in section 6.4.2.

After Malcolm the manager estimates the state of the bridge, he may decide between two alternative actions:

- a_1 = *do nothing* (DN), i.e. no special restrictions to traffic under and over the bridge.
- a_2 = *close bridge* (CB), i.e. both Streicker Bridge and Washington Road are closed to traffic for the time needed for a thorough inspection, estimated to be 1 month.

Finally, Ophelia and Malcolm agree that the costs, denoted by z, related to each action, for each state, are the same as estimated in (Glisic & Adriaenssens, 2010) and reported in Table 6.3. The resultant decision tree of this case study is illustrated in Figure 6.6.

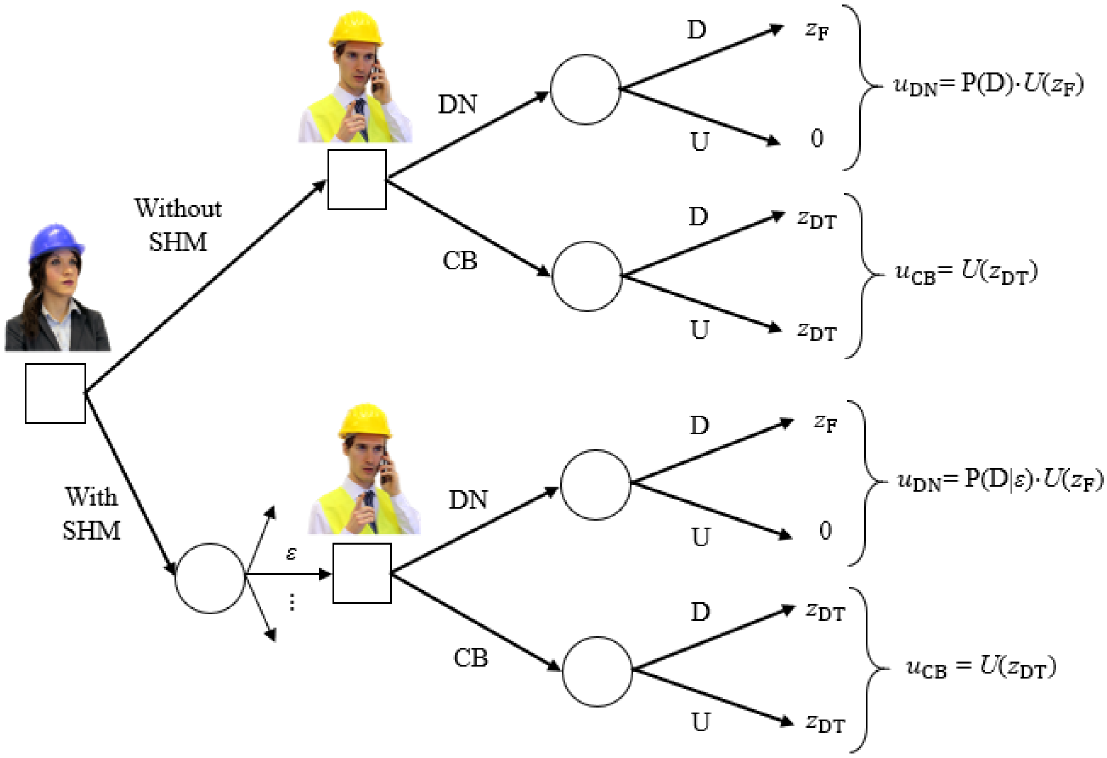


Figure 6.6. Decision tree for the Streicker bridge case study.

Table 6.3. Costs per action and state (Glisic & Adriaenssens, 2010).

	State S_1 = D	State S_2 = U
Action a_1 = DN	$z_{11} = z_F$ = 881.60 k$	z_{12} = 0.00 k$
Action a_2 = CB	$z_{21} = z_{DT}$ = 139.80 k$	$z_{22} = z_{DT}$ = 139.80 k$

In order to apply the formulation about the negative *VoI* introduced in section 6.3, we need to analyse indexes $g(\varepsilon)$ and $G(\varepsilon)$, which depend only on the likelihood distributions, and index R, which instead depends on the appetite for risk of the decision maker and on the choice of prior probabilities.

6.4.2 Analysis of likelihood distributions

In this subsection, firstly we introduce the likelihood distributions of the case study, then we evaluate the resultant indexes $g(\varepsilon)$ and $G(\varepsilon)$.

Similar to (Zonta, et al., 2014), and in the same way of chapter 5, the likelihoods of the two states are described by Gaussian distributions: $p(\varepsilon|U)$ is the likelihood of no damage, defined with mean value $\mu = 0$ με and standard deviation $\sigma = 300$ με, since Malcolm and Ophelia expect the bridge to be undamaged if the change in strain will be close to zero, along with a natural fluctuation of the strain due to thermal effects and to a certain extent due to creep and shrinkage; $p(\varepsilon|D)$ is instead the likelihood of damage, defined with mean value $\mu = 1000$ με and standard deviation $\sigma = 600$ με, since in this case they expect a significant change in strain.

Before the data are available, Malcolm and Ophelia can predict the distribution of ε, which is practically the so-called evidence in classical Bayesian theory, through the formula:

$$p(\varepsilon) = p(\varepsilon|D) \cdot P(D) + p(\varepsilon|U) \cdot P(U). \tag{6.32}$$

When the measurement ε is instead available, both the agents update their estimation of the probability of damage consistently with Bayes' theorem:

$$p(D|\varepsilon) = \frac{p(\varepsilon|D) \cdot P(D)}{p(\varepsilon)}, \tag{6.33}$$

where $p(D|\varepsilon)$ is the posterior probability of damage.

The defined likelihood distributions, illustrated in Figure 6.7(a), allow us to calculate the resulting indexes $g(\varepsilon)$ and $G(\varepsilon)$, which are presented in Figure 6.7(b). Note that the plot of $g(\varepsilon)$ has been cut at $g(\varepsilon) = 2$, since for $\varepsilon > 500$ με $g(\varepsilon)$ tends to infinity.

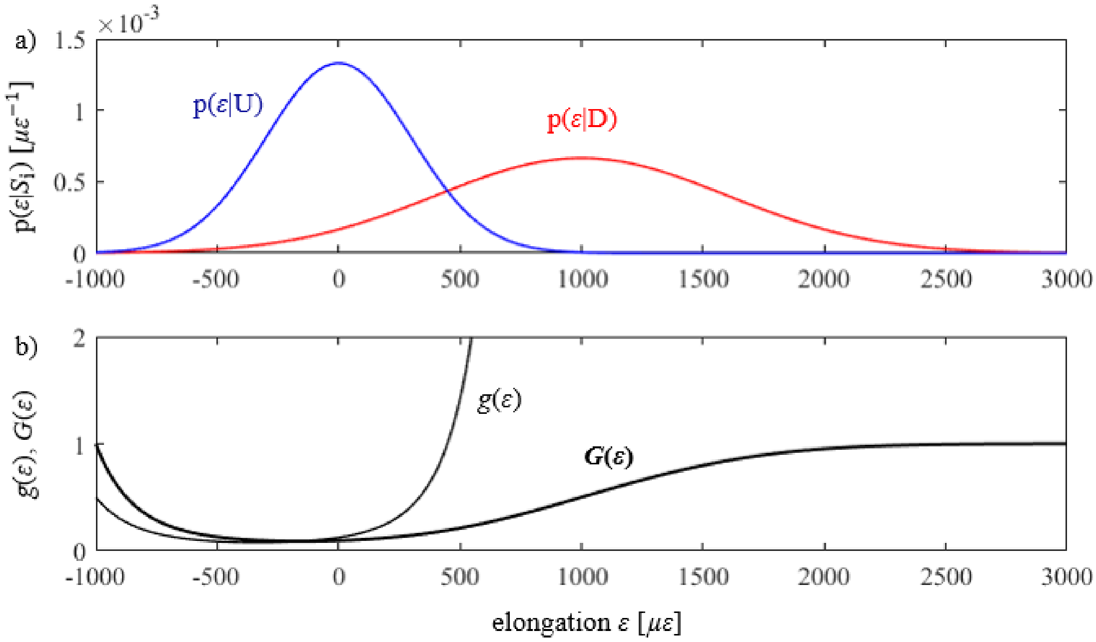


Figure 6.7. Analysis of the likelihood distributions: likelihoods (a), index $g(\varepsilon)$ and $G(\varepsilon)$ (b).

6.4.3 Analysis of appetite for risk of decision makers

The index R varies according to the appetite for risk of the decision maker, i.e. index r, and to the choice of prior probabilities, i.e. index q. While we will analyse the prior probabilities later, we introduce here the different utility functions of the two agents.

As introduced before, Ophelia and Malcolm differ in their utility functions, which is the weight they apply to the possible economic losses. In the following, we use the indices (M) and (O) to indicate that a quantity is intended respectively from Malcom the manager's perspective and Ophelia the owner's perspective. According to

(Bolognani, et al., 2018), we define Ophelia the owner risk neutral with respect to the loss compared to the value of a single structure, since she is in charge of a large stock of structures. This means that, according to her behaviour, a negative utility is linear with the incurred loss. Strictly speaking, a utility function is defined except for a multiplicative factor, therefore it should be expressed in an arbitrary unit sometime referred to as *util* (McConnell, 1966). Since Ophelia's utility is linear with loss, for the sake of clarity we will deliberately confuse negative utility with loss, and therefore we will measure Ophelia's utility in k$.

Unlike Ophelia, in order to demonstrate the formulation introduced in section 6.3 about the negative *VoI*, we suppose that the behaviour of Malcolm the manager can be risk adverse or risk seeking: in this way, since we assume the owner to be always risk neutral, we can analyse both the situations of a manager more risk seeking and more risk adverse than the owner. It is possible to describe mathematically these behaviours using the Arrow-Pratt's utility model (Pratt, 1964) (Arrow, 1965), where the different aptitude of an agent is encoded in the coefficient of Absolute Risk Aversion (ARA) θ. Similarly to (Bolognani, et al., 2018), we assume that the manager's utility has constant ARA, and then the utility function takes the form of an exponential:

$$ {}^{(\mathrm{M})}U(z) > \frac{1 - \mathrm{e}^{-z \cdot \theta}}{\theta}, \tag{6.34} $$

where θ is the constant ARA coefficient. Figure 6.8 shows the linear utility function of Ophelia's behaviour and both Malcom's utility functions, which depend on his particular behaviour:

- Risk adverse, i.e. his negative utility increases more than proportionally with the loss, using $\theta = -1.423$ M^{-1}.
- Risk seeking, i.e. his negative utility increases less than proportionally with the loss, using $\theta = 3.034$ M^{-1}.

Based on these utility functions, we may calculate the utilities U of the costs related to each action, for each state, and consequently the index r, which in this case turns into:

$$r = \frac{U(z_{DT})}{U(z_F) - U(z_{DT})}. \qquad (6.35)$$

All the outcomes are reported in Table 6.4. As expected, we can notice that the more the decision maker is risk seeking, the bigger is the index r: ${}^{(O)}r = 0.189 > {}^{(M)}r = 0.096$ if the manager is risk adverse and then he is less risk seeking than the owner, ${}^{(M)}r = 0.590 > {}^{(O)}r = 0.189$ if the manager is risk seeking and then he is more risk seeking than the owner, who we remember to be considered always risk neutral.

In summary, we have fixed index r, while R will depend on the choice of the prior probabilities. In the next subsections, to understand how the *VoI* varies depending on the different appetite for risk of the two agents, and when it may consequently become negative, we evaluate it both in the case of the manager risk adverse and risk seeking.

Table 6.4. Ophelia's and Malcolm's loss perception.

Ophelia the owner RISK NEUTRAL			
	State D	State U	${}^{(O)}r$
Action DN	${}^{(O)}U(z_F)$ = -881.60 k$	${}^{(O)}U(z)$ = 0.00 k$	0.189
Action CB	${}^{(O)}U(z_{DT})$ = -139.80 k$	${}^{(O)}U(z_{DT})$ = -139.80 k$	
Malcolm the manager RISK ADVERSE			
	State D	State U	${}^{(M)}r$
Action DN	${}^{(M)}U(z_F)$ = -1762.94 k$	${}^{(M)}U(z)$ = 0.00 k$	0.096
Action CB	${}^{(M)}U(z_{DT})$ = -154.94 k$	${}^{(M)}U(z_{DT})$ = -154.94 k$	
Malcolm the manager RISK SEEKING			
	State D	State U	${}^{(M)}r$
Action DN	${}^{(M)}\mathrm{U}(z_F)$ = -306.88 k$	${}^{(M)}\mathrm{U}(z)$ = 0.00 k$	0.590
Action CB	${}^{(M)}\mathrm{U}(z_{DT})$ = -113.93 k$	${}^{(M)}\mathrm{U}(z_{DT})$ = -113.93 k$	

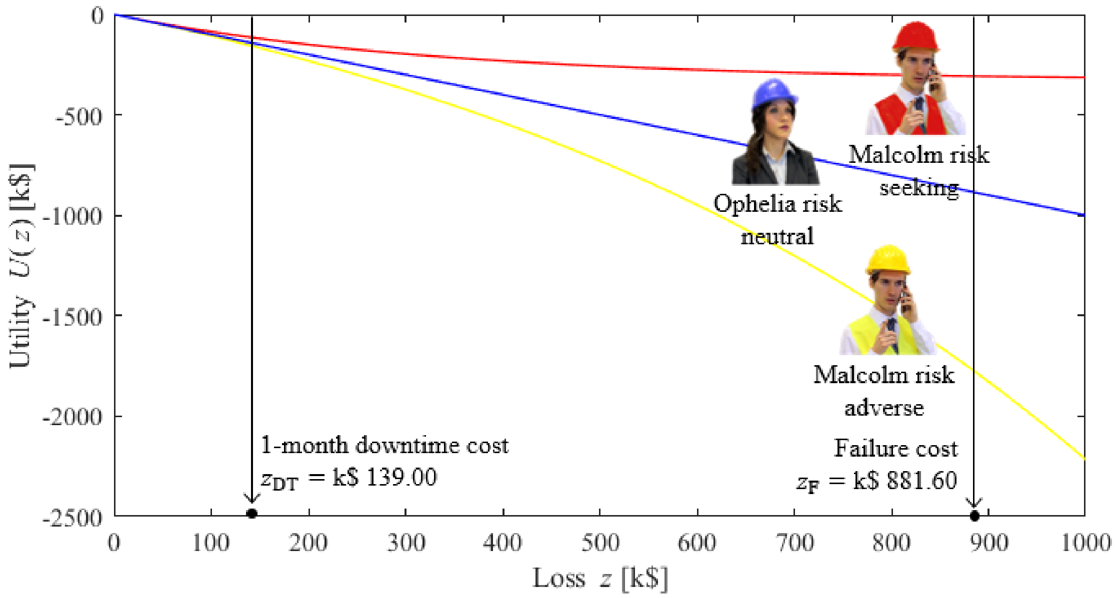


Figure 6.8. Representation of the utility functions for Malcolm the manager and Ophelia the owner.

6.4.4 Case 1: Malcolm the manager risk adverse

In this first case, Ophelia is risk neutral while Malcolm is risk adverse, meaning that the manager is less risk seeking than the owner. According to the formulation introduced in section 6.3, we want to verify that in this case it is impossible to find a negative *conditional VoI*. Since we have defined all the indexes about the formulation except q, in the following we analyse everything in term of P(D), i.e. in term of q.

To start, we evaluate the expected utilities u_0 a priori, i.e. if the monitoring system is not installed. We know that a decision maker would always choose to close the bridge when their utility related to the action CB is less negative than the utility of DN:

$$u_{\mathrm{CB}} \geq u_{\mathrm{DN}}, \qquad U(z_{\mathrm{DT}}) \geq U(z_{\mathrm{F}}) \cdot \mathrm{P(D)}. \tag{6.36a,b}$$

Consequently, we achieve that for Ophelia it is always convenient a priori to close the bridge if P(D) > 0.16, while for Malcolm if P(D) > 0.09, that is smaller because of his risk adverse behaviour. The outcomes are presented in Figure 6.9(a), along with the conditional expected utility ${}^{(\mathrm{O|M})}u_0$ calculated as in Eq. (6.5), which is what we really need in order to evaluate the *conditional VoI*. Note that ${}^{(\mathrm{O|M})}u_0$ has a discontinuity for P(D) = 0.09, since this is the value of P(D) for which a priori the manager changes the decision from action DN to action CB. In addition, we remind that the formulation introduced in section 6.3 is based on the assumption that a priori it is always convenient to choose action a_2, i.e. CB for this case study: this corresponds to having P(D) > 0.16, since in this way both the agents agree on choosing action CB a priori, i.e. their index R is < 1.

Consider the case of the monitoring system installed. In this case the decision maker can rely on the monitoring data ε, and then we can evaluate the preposterior expected utilities, in formula:

$$u_{\mathrm{CB}|\varepsilon} = u(\mathrm{z_{DT}}), \qquad u_{\mathrm{DN}|\varepsilon} = u(\mathrm{z_F}) \cdot \mathrm{p}(\mathrm{D}|\varepsilon). \tag{6.37a,b}$$

Note that the preposterior expected utilities of action DN depends on the posterior probability of having the bridge damaged p(D|**ε**), that can be calculate as in Eq. (6.33). The resultant preposterior expected utilities u_{pp} are presented in Figure 6.9(b), in term of P(D), both in the unconditional and conditional form. It is possible to notice that the conditional outcome, i.e. $^{(\mathrm{O|M})}u_{pp}$, has again a discontinuity, this time for P(D) = 0.53, which corresponds to the value of P(D) for which a posteriori the manager changes his decision from action DN to action CB.

Finally, the *VoI* is simply the difference between the preposterior expected utility and the prior expected utility. We can then calculate both the *unconditional* and *conditional VoI*, according respectively to Eq. (6.4) and Eq. (6.7). Figure 6.9(c) shows the results, always in term of P(D). As regards the *unconditional VoI*, i.e. $^{(\mathrm{O})}uVoI$ and $^{(\mathrm{M})}uVoI$, we can observe that they are maximum exactly at the value of P(D) for which it becomes convenient a priori to close the bridge, i.e. if P(D) = 0.16 for the owner and P(D) = 0.09 for the manager; we remind that these are the values which corresponds to having R = 1. In addition, it is possible to verify that it is never possible to find a negative *unconditional VoI*, according to the principle introduced in section 6.2 that "*information can't hurt*". In addition, we can notice that, as expected since in this case the manager is less risk seeking than the owner, we cannot find any value of P(D) for which the *conditional VoI*, i.e. $^{(\mathrm{O|M})}VoI$, becomes negative. This happens because, due to Malcolm's risk adverse behaviour, he would always choose to close the bridge a posteriori sooner than Ophelia, i.e. $^{(\mathrm{M})}\,\bar{\varepsilon} < {}^{(\mathrm{O})}\,\bar{\varepsilon}$, and then we may obtain a smaller positive *uVoI*, but we can't achieve what we have defined as *Loss for Disagreement* (*LfD*): consequently, it is impossible to get a negative *VoI*, as expected.

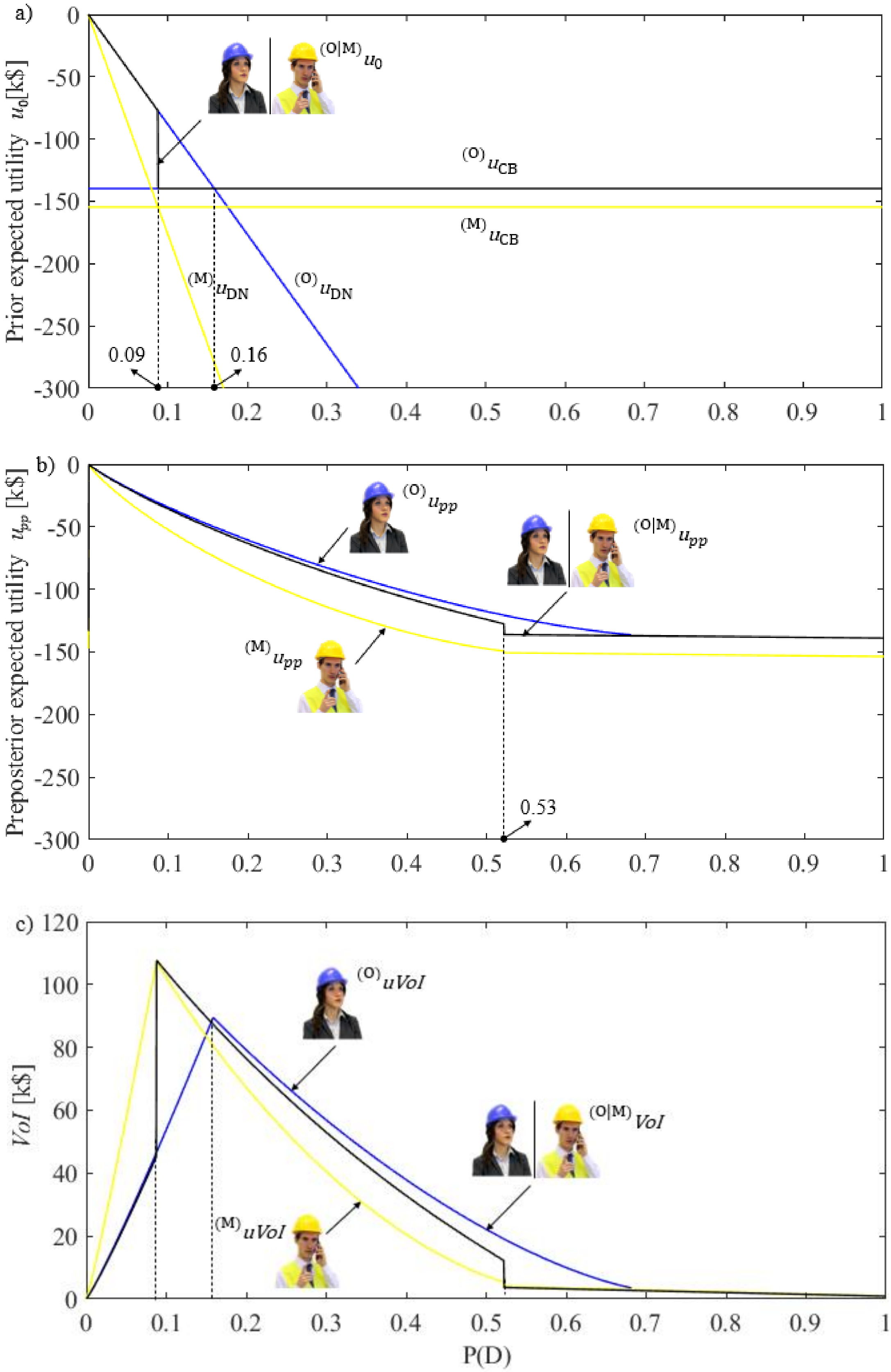


Figure 6.9. Prior expected utilities u_0 (a), preposterior expected utilities u_{pp} (b), *VoI* (c).

6.4.5 Case 2: Malcolm the manager risk seeking

In this second case, we consider Ophelia the owner still risk neutral, while Malcolm the manager is now risk seeking. This corresponds to the case where, according to the developed formulation of section 6.3, it should be possible to find a negative *conditional VoI*. We want then to find out for which values of P(D), and consequently of the term q, this happens. The procedure followed is the same as in section 6.4.4.

We evaluate the expected utilities u_0 a priori, i.e. in the case of a monitoring system not installed. In this case, as shown in Figure 6.10(a), for Ophelia is again convenient a priori to close the bridge if P(D) > 0.16, since she is still risk neutral, while for Malcolm it becomes P(D) > 0.37, that is clearly higher because of his risk seeking behaviour. As a consequence, the two agents agree on choosing a priori action CB if P(D) > 0.37.

Figure 6.10(b) presents the unconditional and conditional preposterior expected utilities, needed in order to evaluate the *VoI*, which is instead illustrated in Figure 6.10(c). As regards the *unconditional VoI*, i.e. $^{(O)}uVoI$ and $^{(M)}uVoI$, we can again observe that they are maximum exactly at the value of P(D) for which it becomes convenient a priori to close the bridge, i.e. if P(D) = 0.16 for the owner and P(D) = 0.37 for the manager, and that it is never possible to find a negative *unconditional VoI*. Conversely, it is clearly possible to find some values of P(D) for which the *conditional VoI*, i.e. $^{(O|M)}VoI$, becomes negative: 0.58 < P(D) < 0.87. This happens because, due to his risk seeking behaviour, Malcolm would always choose to close the bridge a posteriori later than Ophelia, i.e. $^{(M)}\bar{\varepsilon} > {}^{(O)}\bar{\varepsilon}$, and therefore we achieve what we have defined *LfD*. Since there are some values of P(D), i.e. 0.58 < P(D) < 0.87, for which this the *LfD* is bigger than the *uVoI*, the consequence is that we achieve a negative *conditional VoI*.

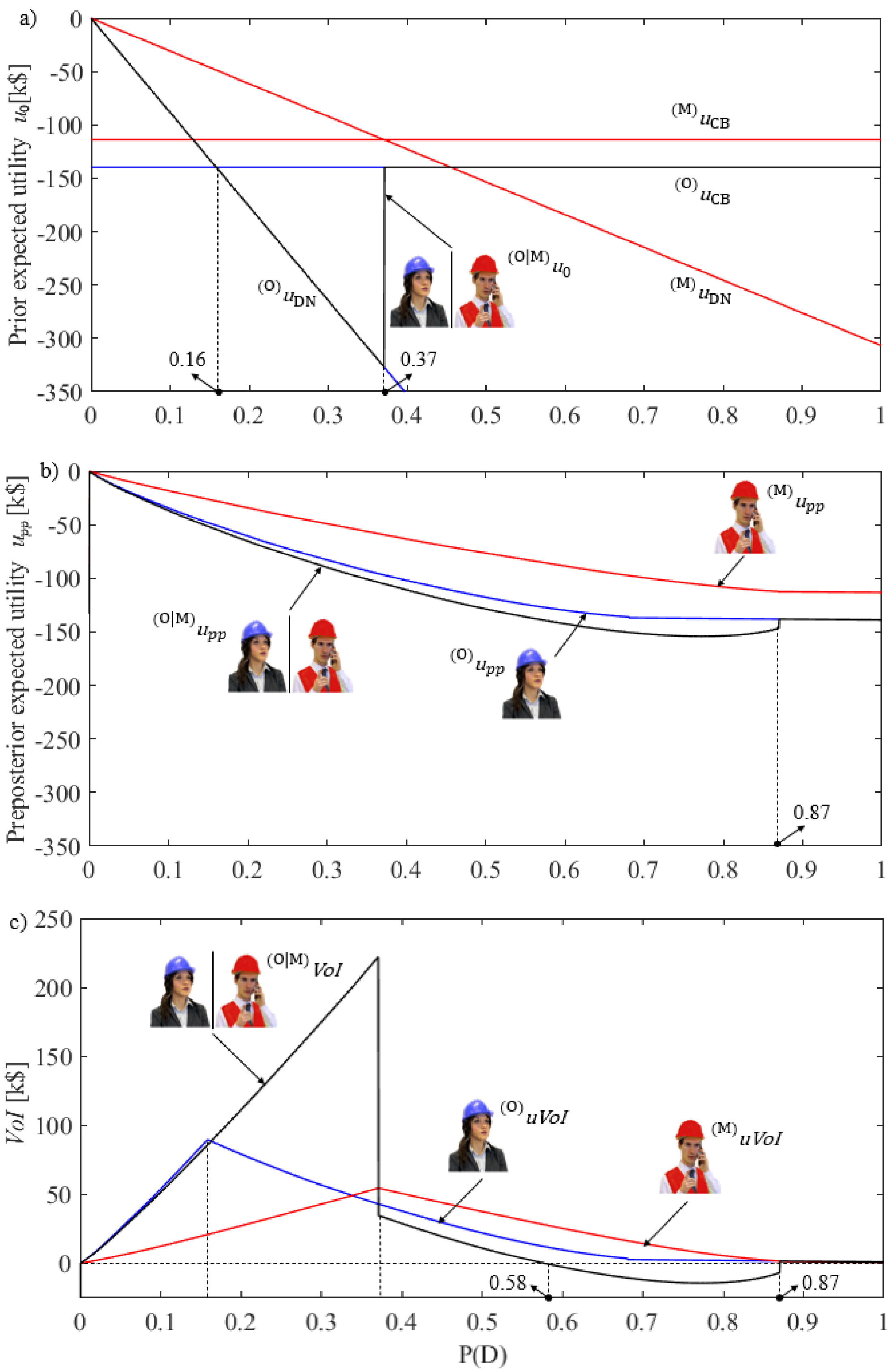


Figure 6.10. Prior expected utilities u_0 (a), preposterior expected utilities u_{pp} (b), *VoI* (c).

6.4.6 Discussion about negative conditional *VoI*

In the previous subsections we have demonstrated that, as expected, it is possible to achieve a negative *conditional VoI* only when the manager is more risk seeking than the owner, which agrees with the conclusions obtained theoretically in section 6.3. In this specific case study, it happens when 0.58 < P(D) < 0.87, which corresponds to 1.38 < q < 6.69.

We analyse one specific case in this range: we choose for instance P(D) = 0.65, i.e. q = 1.86. In this case, indexes R and the thresholds for the two agents are:

$$^{(\mathrm{M})}R = 0.32 > {}^{(\mathrm{O})}R = 0.10. \tag{6.38}$$

$$^{(\mathrm{M})}\bar{\varepsilon} = 247\ \mu\varepsilon > {}^{(\mathrm{O})}\bar{\varepsilon} = -84\ \mu\varepsilon. \tag{6.39}$$

As expected, the threshold of Malcolm the manager is bigger than the one of Ophelia the owner, since Malcolm is more risk seeking than Ophelia. Consequently, there is a very wide range of values, from -84 με to 247 με, whereby Malcolm would keep the bridge open in disagreement with Ophelia, who instead believes this is a dangerous practice which can potentially result in a big loss. She is then forced to keep the bridge open for ε = [-84 με ÷ 247 με], even if it would be more convenient for her to close it: this causes a *LfD*, as shown in Figure 6.11(b). Since this negative area is bigger than the one of *uVoI*, the resultant *conditional VoI* is negative:

$$^{(\mathrm{O|M})}VoI = -11.61\ k\$. \tag{6.40}$$

This means that in this case Ophelia perceives the monitoring information as damaging: in summary, a negative *VoI* corresponds exactly to the amount of money Ophelia the owner is willing to pay to prevent Malcolm the manager using the monitoring system.

In conclusion, we have proved that, for the prototype decision problem analysed in this contribution, the achievement of a negative *conditional VoI* depends on a combination between how much more seeking is the manager in comparison to the owner, and the choice of prior probabilities. While the development of our case study has been conditioned by the choice of specific risk appetites of the agents, i.e. fixed θ, in order to have a final verification of our conclusions, it is interesting to investigate how the *conditional VoI* varies according to both the prior damage probability P(D), i.e. in term of index q, and the ARA coefficient θ of the manager, i.e. in term of his appetite for risk. Figure 6.12 shows graphically the results, with both a top view and a 3D view: it is clear that we achieve a negative *conditional VoI*, i.e. the dark blue area, only for some specific combinations of high P(D) and positive θ, which indeed corresponds to a manager who is more risk seeking in comparison to the owner, who instead we remind to be defined as risk neutral (θ = 0). In the top view of Figure 6.12(a), we have highlighted the specific case analysed in this section, i.e. θ =3.034 M$$^{-1}$ for the manager and P(D) = 0.65: it allows us to verify that, as calculated in Eq. (6.40), this case falls into the area where we achieve a negative *conditional VoI*.

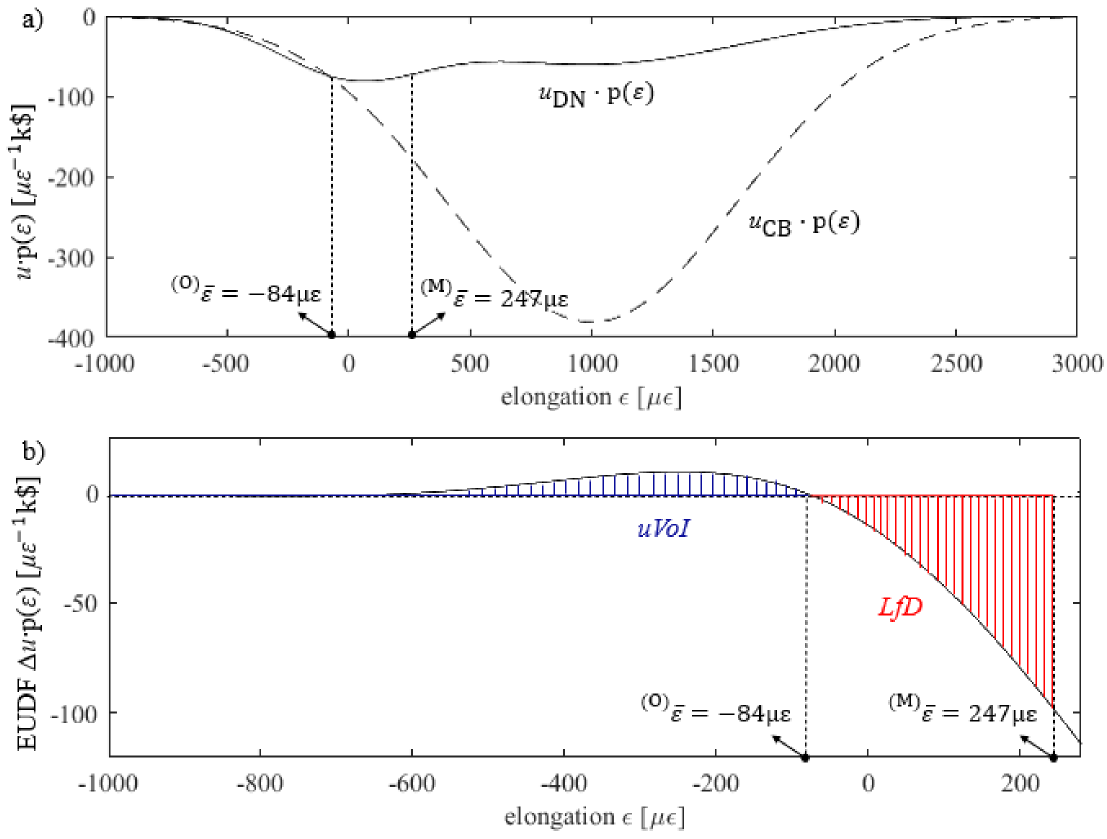


Figure 6.11. Analysis of the *conditional VoI*: density function of the two expected utilities (a); EUDF, zoom in the values of interest (b).

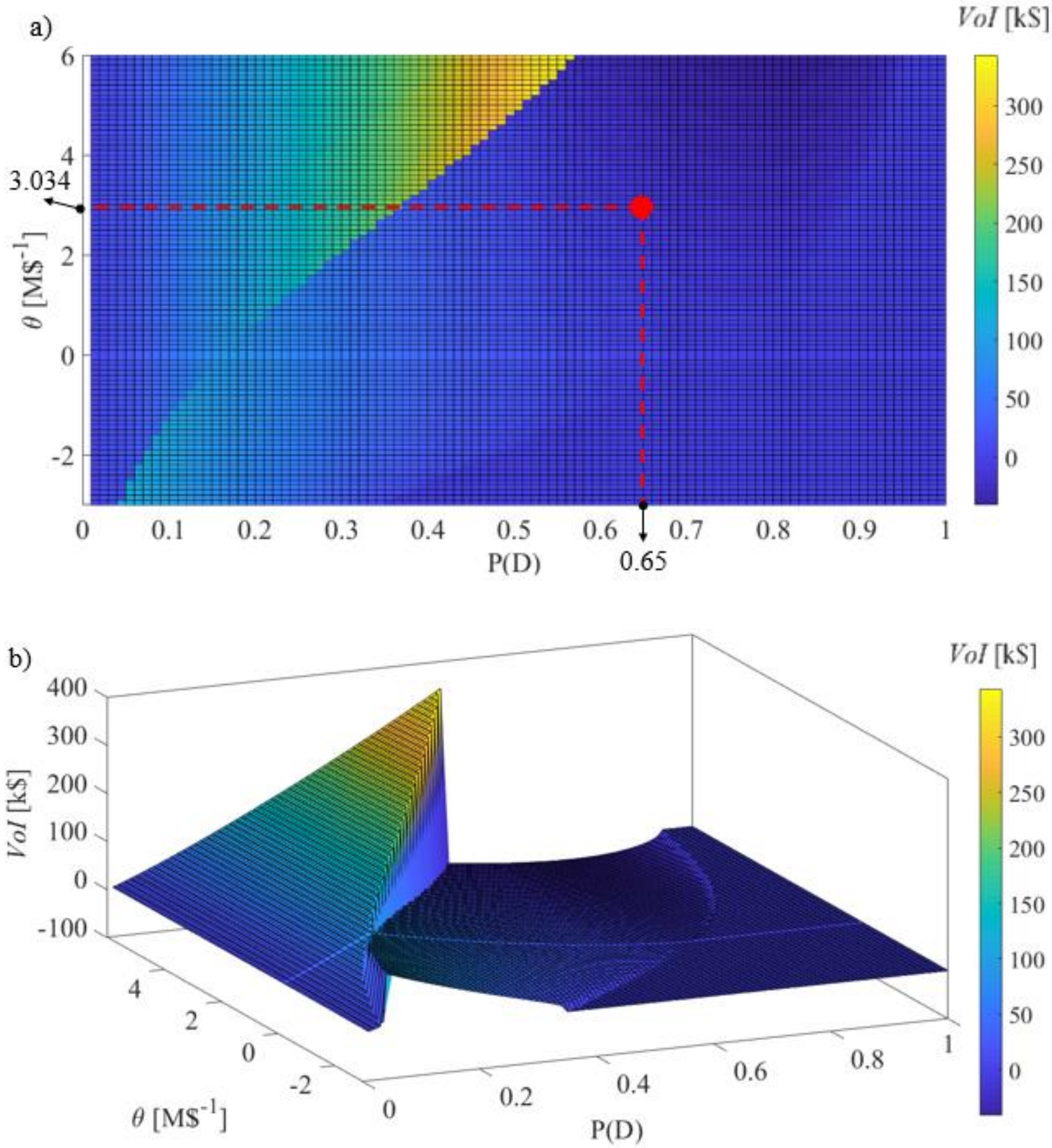


Figure 6.12. Graphical representation of the *conditional VoI* in function of both the prior damage probability P(D) and the ARA coefficient θ: top view(a) and 3D view (b).

6.5 Conclusions

The benefit of SHM can be quantified using the concept of the *VoI*. In its calculation, a commonly understood assumption is that the individual who decide on the installation of the monitoring system, i.e. the owner, is the same rational agent who will later use it, i.e. the manager. With this assumption, the so-called *unconditional*

VoI is never negative, according to the assumption that "*information can't hurt*". On the other hand, we must recognize that in the real word these two agents involved in the decision chain are usually different individuals. In this case, it has been demonstrated that the so-called *conditional VoI* may become negative, meaning that the monitoring information are perceived as damaging. The aim of this contribution has been to investigate, for a specific prototype decision problem, which are the conditions, necessary and/or sufficient, for which it is possible to obtain a negative *conditional VoI*.

We have first summarized the two mathematical formulations for the evaluation of the *VoI*, i.e. in the unconditional and conditional case. Then, we have investigated theoretically how it is possible to achieve a negative *conditional VoI* based on the assumptions of our prototype decision problem: we have understood that the predominant factor is the different risk appetite of the two decision makers, in particular how much more risk seeking is the manager in comparison to the owner. Secondly, other influential factors are the shape of the likelihood distributions and the values of prior probabilities. We have proved these theoretical conclusions by describing the likelihood distributions with polynomial functions, which have allowed to get results in closed form.

To verify the developed framework, we have applied it to a SHM-based decision problem regarding the Streicker Bridge, which is a pedestrian bridge at Princeton University campus equipped with a continuous monitoring system. The achieved results allow us to verify that it is never possible to find a negative *unconditional VoI*, while it is possible to find a negative *conditional VoI* in the case of a manager more risk seeking than the owner. Indeed, in this case the owner is forced to keep the bridge open even if it would be more convenient for her to close it: this *Loss for Disagreement LfD* between the manager and the owner, due to their different risk appetite, may ultimately lead to a negative *conditional VoI*. This outcome means that the owner perceives the monitoring information as damaging, because she believes that the monitoring system can seriously mislead the decision of the manager: a negative *conditional VoI* corresponds exactly to the amount of money the owner is willing to pay to prevent the manager using the monitoring system.

Acknowledgements

The case study reported in this paper is based on the Streicker Bridge monitoring project, as illustrated in references (Zonta, et al., 2014) (Bolognani, et al., 2018). The authors wish to particularly thank Prof. Branko Glisic and Prof. Sigrid Adriaenssens, Princeton University, for sharing the information on this monitoring project. Fictional owner Ophelia and manager Malcom, who appear in section 6.4, are impersonated by Denise Bolognani and Daniel Tonelli, University of Trento; their contribution is greatly acknowledged.

7. Conclusions

In this final chapter, we present a summary and a discussion about the outcomes of the research presented in this thesis, along with its limitations and the related future works.

The main goal of this contribution is to support the rational management of critical structures in civil engineering. We have concentrated especially on decision-making processes based on structural information, since their research interest has grown increasingly in the last decades. The analysis of the current state of art highlighted the necessity to analyse deeper these engineering decision-making processes, since in real-life decision makers often distort the rational framework that can be found in the literature, presented in section 2.1. We remind that, being able to predict the behaviour of a real-word agent is fundamental in various real-life engineering situations, for instance when the management of the structure is based on a policy that depends on decisions of several decision makers that may behave differently, even irrationally.

In the following, we discuss the achievement of the 4 aims introduced in section 1.3.

i) To investigate how heuristic behaviours affect human judgment and decision-making.

We have investigated this first aim in chapter 3. In general, people use heuristics as efficient rules to simplify complex problems and overcome the limits in rationality and computation of the human brain. Even though the results are typically satisfactory, they can differ from those derived from a rational process; psychologists call these differences cognitive biases. Many heuristic behaviours have been identified and investigated in the literature, with applications in various fields such as psychology, cognitive science, economics and finance, but rarely in SHM-based decision problems. In particular, we have identified Kahneman and Tversky's *representativeness* as a heuristic for which SHM-based decision-making is particularly susceptible, where simplified rules for updating probabilities can distort the decision maker's perception of risk. This biased behaviour is frequently observed in bridge management as the confusion between condition state and safety of the bridge. Therefore, based on the available literature, we have proposed a mathematical framework in order to reproduce the impact of this biased behaviour, and to compare it with the rational one which is instead based on Bayesian methods. We have tested the developed framework on a classical representativeness problem. Then, we have applied it to a real-life case study concerning the evaluation of the safety of a bridge, based on visual inspection. The results of the case study demonstrate that, while rational Bayes' theorem correctly identifies the bridge as safe, the application of the proposed representativeness framework predicts that the manager mistakenly judges the bridge as unsafe, based on the observed condition state. This outcome supports our hypothesis that this irrational judgment sometimes observed in bridge managers' behaviour can be described using Kahneman and Tversky's representativeness heuristic, and that our proposed mathematical framework reproduces it appropriately. In conclusion, this work opens a new research path in bridge management since it demonstrates the influence of heuristics in managers' behaviour. As regards to future work, we think that real-life bridge managers may be biased by other heuristic behaviours, and as such it would be of interest to study these as we have done in this thesis with the representativeness. In particular, we aim to study the other two main heuristics of Kahneman and Tversky, i.e. availability and anchoring, which we have started to analyse in this thesis.

ii) To develop a process to elicit engineering expert knowledge, by minimizing the risk of biased judgments.

This second aim is developed in chapter 4. In general, eliciting expert judgment in the form of subjective probabilities is a socio-technical activity and requires a structured and facilitated process to extract meaningful judgments because people, even experts, are unable to provide accurate and reliable data simply on request: indeed, simply asking a person for their best estimate results in poor data due to the plethora of biases in human judgment. The case study motivating this part of our research concerned the system reliability of a dam in Ontario (Canada), with the final aim being to predict the probability of failure of the dam. We chose the Bayesian Network (BN) as the mathematical model to explicate the relationship between environment, usage, hazard and management decisions, and to support the optimal long-term productivity of the asset. Due to the lack of data and information, we decided to rely on expert judgment to quantify the BN. Even if a variety of existing processes for eliciting expert knowledge with engineering applications are available in the literature, very little has been reported about elicitation processes aimed specifically at quantifying BNs. Consequently, we have developed a four-stage structured elicitation process to support the collection of valid and reliable data with the specific aim to quantify a BN. In the development of the methodology, each stage is proposed by highlighting all the potential biases that may influence the process as well as by proposing appropriate actions in order to minimize the risk of biased judgments. In particular, we have highlighted the following six biases: management, expert, conjunctive, anchoring, availability and central. Our process was applied in a case study where the numerical outcomes demonstrated the operationalizability of the four-stage structured elicitation process. Upon reflection following the implementation of the process, we have identified aspects of the process for improvement in order to enhance the reliability of the results. For instance, a possible limitation of these processes in real-life is due to the difficulty in finding available experts, and therefore, we need to consider carefully which expertise is essential and which is desirable. In conclusion, this work demonstrates the validity and the usefulness of engineering expert knowledge as an accurate and reliable data source, if an appropriate methodology is developed in order to minimize the risk of biased judgments. As regards to future work, we aim to improve this structured methodology based on what we have learnt from this first application, such as about expert selection, and to apply it to other civil engineering structures, e.g. bridges.

iii) To investigate how decision-making can be distorted when multiple decision makers are involved in the decision chain.

This aim is investigated in chapter 5. In particular, we recognize that real world decision-making processes are based on the interactions between more individuals, that can take different decisions according to their appetites for risk. We chose to study this interaction between rational decision makers by analysing the benefit of SHM, using the concept of *VoI*. In classical decision theory, one of the main assumptions is that all decisions concerning system installation and operation are taken by the same rational agent, so a methodology for evaluating the *VoI* in this case is available in the literature. However, nothing is available as regards more individuals involved in the decision chain. Therefore, we have developed an innovative rational method for quantifying the *VoI* when two different agents are involved in the decision chain: we have called them respectively manager, i.e. who makes decisions regarding the structure, and owner, i.e. who chooses whether to install the monitoring system or not. This new methodology, which takes into account the possibility that the two decision makers may act differently due to different appetites for risk, allows one to evaluate what we have called *conditional VoI*, i.e. the *VoI* achieved using the utility of the owner, but conditional to the action chosen by the manager. To illustrate how this framework works, we have evaluated a hypothetical *VoI* for a pedestrian bridge at Princeton campus, which is equipped with a monitoring system. The achieved results demonstrate how decisions may be distorted using this new formulation, in comparison to the classical one where all decisions are taken by the same rational agent, which we have referred to as *unconditional*. In conclusion, the benefit of monitoring in the conditional case could be greater or smaller in comparison to the unconditional case, according to the different risk appetites of decision makers. In addition, this work has allowed us to note that, in contrast with the unconditional formulation, it is possible to find a negative *VoI* with the conditional formulation. This is an unexpected outcome since it means that the monitoring information may be perceived as damaging. This final outcome has been the motivation for our last contribution of this thesis, presented in chapter 6.

iv) To demonstrate that the value of information may be negative in the case of multiple rational decision makers with different appetites for risk.

This final aim is discussed in chapter 6. In particular, starting from the conditional formulation of *VoI* that we proposed in the contribution of chapter 5, we have developed a mathematical formulation that aims to understand the circumstances for which, under specific assumptions, it is possible to find a negative conditional *VoI*. The formulation shows that the predominant factor is the different risk appetites of the two decision makers, since this may lead to a big disagreement between them, and consequently in some cases even to a negative conditional *VoI*. After proving the outcomes achieved theoretically using specific likelihood functions, we have applied it to the same decision problem as in chapter 5. The outcomes of the case study demonstrate the validity and the usefulness of the developed methodology. In conclusion, we claim a negative conditional *VoI* corresponds exactly to the amount of money the owner is willing to pay to prevent the manager using the monitoring system. The possibility to achieve a negative conditional *VoI* is a further proof of the relevance of the consequences that can be caused by distorted decisions. As regards to future work, we aim to investigate the negative *VoI* in decision problems based on different assumptions, for instance as regards the number of decision makers involved, the level of uncertainty, the likelihood distributions.

In conclusion, the four contributions presented in this thesis prove the relevance of investigating the consequences of distorted human judgments and decision-making in civil engineering applications, which was the main goal of this thesis. In particular, the outcomes achieved allow us to demonstrate the operationalizability of the methods developed in this thesis, and to prove their relevance in various civil engineering case studies.

References

Aguilera, P. A. et al., 2011. Bayesian networks in environmental modeling. *Environmental Modelling & Software*, 26(12), pp. 1376-1388.

Ahmadi, A., Moridi, A. & Han, D., 2015. Uncertainty Assessment in Environmental Risk through Bayesian Networks. *Journal of Environmental Informatics,* 25(1).

American Ass. State Highway and Transportation Off, 1997. *AASHTO guide for commonly recognized (CoRe) structural elements.* Washington D.C.: AASHTO.

Angner, E., 2012. *A course in behavioral economics.* New York: Palgrave Macmillan.

Armstrong, J. S., Denniston, W. B. & Gordon, M. M., 1975. The use of decomposition principle in making judgements. *Organizational Behavior and Human Performance,* 14(2), pp. 257-263.

Arrow, K. J., 1965. *Aspects of the Theory of Risk Bearing.* Helsinki: Yrjo Jahnssonin Saatio.

Astfalck, L. et al., 2018. Expert elicitation of directional metocean parameters. *Ocean Engineering,* Volume 161, pp. 268-276.

Autonomous Province of Trento , 2018. *www.bms.provincia.tn.it/bms.* [Online].

Baiman, S., 1975. The evaluation and choice of internal information systems within a multiperson world. *Journal of Accounting Research,* Volume 8, pp. 1-15.

Barberis, N., Shleifer, A. & Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics,* 49(3), pp. 307-343.

Bayraktarli, Y., Ulfkjaer, J., Yazgan, U. & Faber , M., 2005. *On the application of Bayesian probabilistic networks for earthquake risk management.* Rome, Proc. 9th International Conference on Structural Safety and Reliability.

Beck , J. L. & Au, S., 2002. Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *Journal of Engineering Mechanics,* 128(4), pp. 380-391.

Beck , J. L. & Yuen, K. V., 2004. Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics,* 130(2), pp. 192-203.

Belton, V. & Steward, T., 2001. *Multiple criteria decision analysis: an integrated approach.* Dordrecht: Kluwer Academic Publishers.

Ben Gal, I., 2007. Bayesian Networks. In: Ruggeri, Faltin & Kennett, eds. *Encyclopedia of Statistics in Quality & Reliability.* Chester: Wiley.

Bensi, M. T., Der Kiureghian, A. & Straub, D., 2011. *A Bayesian network methodology for infrastructure seismic risk assessment and decision support.* PEER Report 2011/02.

Bernal, D., Zonta, D. & Pozzi, M., 2009. *An examination of the ARX as a residual generator for damage detection.* San Diego, CA, In Proceedings of SPIE: health monitoring of structural and biological systems.

Bernoulli, D., 1954. Exposition of a new theory on the measurement of risk. *Econometria,* 22(1), pp. 23-36.

Betsch, T. & Pohl, D., 2002. Tversky and Kahneman's availability approach to frequency judgment: A critical analysis. In: P. Sedlmeier & T. Betsch, eds. *Frequency processing and cognition.* Oxford, UK: Oxford University Press, pp. 109-119.

Bhatia, S. & Chaudhry, S., 2013. The Dynamics of Anchoring in Bidirectional Associative Memory Networks. *Psychology, Computer Science.*

Bilek, J., Nedoma, J. & Jirasek, M., 2018. Representativeness heuristics: a literature review of its impact on the quality of decision-making. *Journal of Civil Engineering,* 26(2), pp. 29-38.

Birnbaum, M. H. & Zimmermann, J. M., 1998. Buying and Selling Price of Investments: Configural Weight Model of Interactions Predicts Violations of Joint Independence. *Organizational behavior and human decision processes,* 74(2), pp. 145-187.

Bishop, C., 2006. Pattern recognition and machine learning. Springer.

Bolger, F., 2018. The selection of Experts for (Probabilistic) Expert Knowledge Elicitation. In: L. C. Dias, A. Morton & J. Quigley, eds. *Elicitation.* Springer, pp. 393-443.

Bolognani, D. et al., 2017. *An application of Prospect Theory to a SHM-based decision problem.* Portland, Proceedings of SPIE.

Bolognani, D. et al., 2018. Quantifying the benefit of strutural health monitoring: what if the manager is not the owner?. *Structural Health Monitoring,* 17(6), pp. 1393-1409.

Bolstad, W. M., 2007. *Introduction to Bayesian Statistics.* New York City: Wiley.

Bolstad, W. M., 2010. *Understanding Computational Bayesian Statistics.* NJ, USA: John Wiley & Sons.

Bordalo, P., Coffman, K., Gennaioli, N. & Shleifer, A., 2016. Stereotypes. *Quarterly Journal of Economics,* 131(4), pp. 1753-1794.

Brannon, L. & Carson, K., 2003. Research belief: The representativeness heuristic. *Applied Nursing Research,* 16(3), pp. 201-204.

Briseno-Ramiro, R. A. et al., 2019. *Dam risk assessment using the event tree analysis and Bayesian Networks.* Panama City, Proc. 38th IAHR World Congress.

Brownjohn, J. M., 2007. Structural health monitoring of civil engineering. *Philosophical Transactions of the Royal Society A,* Volume 365, pp. 589-622.

Bubniz, R. J. et al., 1998. Use of technical expert panels: Application to probabilistic seismic hazard analysis. *Risk assessment,* 18(4), pp. 463-469.

Caird, E., 1877. *A critical account of the philosophy of Kant, with an historical introduction.* Glasgow: James Maclehose, St. Vincent street.

Calvi, G. M. et al., 2019. Once upon a Time in Italy: The Tale of the Morandi Bridge. *Structural Engineering International,* 29(2), pp. 198-217.

Cambridge, M. & Darren, S., 2019. Preliminary reflections o the failure of the Brumadinho tailings dam in January 2019. *Dams and Reservoirs*, 29(3), pp. 113-123.

Cantero-Chinchilla, S. et al., 2019. *Optimal Ultrasonic Sensor Configuration for Plate-Like Structures Using the Value of Information.* Stanford, California, Proc. 12th International Workshop on Structural Health Monitoring (IWSHM).

Cappello, C., 2017. *Theory of Decision Based on Structural Health Monitoring,* Trento: Doctoral theis.

Cappello, C., Zonta, D. & Glisic, B., 2016. Expected utility theory for monitoring-based decision-making. *Proceedings of the IEEE,* 104(8), pp. 1647-1661.

Cappello, C. et al., 2015. Impact of prior perception on bridge health diagnosis. *Journal of Civil Structural Health Monitoring,* 5(4), pp. 509-525.

Chaloner, K., Church, T., Lois, T. & Mattis, J., 1993. Graphical Elicitation of a Prior Distribution for a Clinical Trial. *The Statistician,* Issue 42, pp. 341-353.

Charniak, E., 1991. Bayesian networks without tears. *AI magazine,* 12(4), pp. 50-63.

Chen, C.-S., Cheng, J.-C., Lin, F.-C. & Peng, C., 2017. The role of house money effect and availability heuristic in investor behavior. *Management Decision,* 55(8).

Chen, H. P., 2018. *Structural Health Monitoring of Large Engineering Structure.* First Edition ed. Chichester, UK: John Wiley & Sons.

Chen, H. P. & Ni, Y. Q., 2018. Introduction to Structural Health Monitoring. In: *Structural Health Monitoring of Large Civil Engineering Structure.* Chichester, UK: John Wiley & Sons, pp. 1-14.

Christophersen, A. et al., 2018. Bayesian network modeling and expert elicitation for probabilistic eruption forecasting: pilot study for Whakaari/White island, New Zeland. *Frontiers in Earth Science,* Volume 6.

Cohen, J., Chesnick, E. I. & Haran, D., 1972. A confirmation of the inertial-ψ effect in sequential choice and decision. *British Journal of Psychology,* Volume 63, pp. 41-46.

Cooke, R. M., 1991. *Experts in uncertainty.* Oxford: Oxford University Press.

Cover, T. M. & Thomas, J. A., 2006. *Elements of information theory.* 2nd edition ed. John Wiley & Sons.

Cover, T. M. & Thomas, J. A., 2012. *Elements of information theory.* Hoboken, NJ: John Wiley & Sons.

Daly, S. R., Yilmaz, S. & Christian, J. L., 2012. Design heuristics in engineering concept generation. *Journal of Engineering Education,* 101(4), pp. 601-629.

Dassanavake, S. M. & Mousa, A., 2020. Probabilistic stability evaluation for wildlife-damaged earth dams: a Bayesian approach. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards ,* 14(1), pp. 41-55.

De Wilde, T. R. W., Ten Velden, F. S. & De Dreu, C. K. W., 2018. The anchoring-bias in groups. *Journal of Experimental Social Psychology,* Volume 76, pp. 116-126.

DeGroot, M. H., 1984. Changes in utility as information. *Theory and Decision,* 17(3), pp. 287-303.

Dias, L. C., Morton, A. & Quigley, J., 2018. *Elicitation - The Science and Art of Strcturing Judgment.* Springer.

Edward, W., 1968. Conservativism in Human Information Processing.. In: *Formal representation of Human Judgment.* New York: Wiley, pp. 17-52.

El-Awady, A. et al., 2019. *Bayesian Network Approach for Failure Prediction of Mountain Chute Dam and Generating Station.* Ottawa, Ontario, Canada, ICOLD 87th annual meeting and symposium.

Eldosouky, A., Saad, W. & Mandayam, N., 2017. Resilient Critical Infrastructure: Bayesian Network Analysis and Contract-Based Optimization. *arXiv: 1709.00303.v1*.

Elms, D. G. & Brown, C. B., 2013. Intuitive decisions and heuristics - an alternative rationality. *Civil Engineering and Environmental Systems,* Volume 30, pp. 274-284.

Enright, M. P. & Frangopol, D. M., 1999. Condition prediction of deteriorating concrete bridges using Bayesian updating. *Journal of Structural Engineering,* 125(10), pp. 1118-1125.

Epley, N. & Gilovich, T., 2001. Putting Adjustment Back in the Anchoring and Adjustment Heuristic: Differential Processing of Self-Generated and Experimenter-Provided Anchors. *Psychological Science*, 12(5), pp. 391-396.

Epley, N. & Gilovich, T., 2005. When Effortful Thinking Influences Judgmental Anchoring: Differential Effects of Forewarning and Incentives on Self-generated and Externally Provided Anchors. *Journal of Behavioral Decision Making,* Volume 18, pp. 199-212.

European Food Safety Authority, 2014. Guidance on expert knowledge elicitation in food and feed safety risk assessment. *EFSA journal,* 12(6).

Farrar, C. R. & Worden, K., 2007. An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A,* Volume 365, pp. 303-315.

Feigenbaum, E. A. & Feldman, J., 1963. *Computers and thought.* New York: McGraw-Hill Inc..

Ferrell, W., 1985. *Combining individual judgements. In: Wright G, editor. Behaviour decision making.* New York: Plenum Press.

Ferrell, W., 1994. *Discrete subjective probabilites and decision analysis: elicitation, calibration and combination. In: Wright G, Ayton P, editors. Subjective probability.* New York: Wiley.

Fischhoff, B., 1989. Eliciting knowledge for analytical representation. *IEEE Trans Syst, Man Cybernet*, 19(3), pp. 448-461.

Flynn, E. & Todd, M., 2010a. A Bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. *Mechanical Systems and Signal Processing,* 24(4), pp. 891-903.

Flynn, E. & Todd, M., 2010b. Optimal Placement of piezoelectric actuators and sensors for detecting damage in plate sensors. *Journal of Intelligent Material Systems and Structures,* 21(3), pp. 265-274.

Flynn, E. et al., 2011. Enhanced detection through low-order stochastic modeling for guided-wave structural health monitoring. *Structural Health Monitoring,* 11(2), pp. 149-160.

Fortz, B., 2011. Applications of meta-heuristics to traffic engineering in IP networks. *International transactions in operational research,* Volume 18, pp. 131-147.

Furnham, A. & Boo, H. C., 2011. A literature review of the anchoring effect. *The journal of Socio-Economics,* Volume 40, pp. 35-42.

Gang, W., Zhenyue, N., Jingjing, Q. & Guan, L., 2016. Risk assessment method of earth rock-fill dam break using multi-state and fuzzy basic events. *Journal of Hydroelectric Engineering,* 35(8), pp. 95-104.

Geroulas, V., Harrison, D., Gugaratshan, G. & Smith, M., 2019. *Method of generating economic value from navigation lock structural health information.* Stanford (California), Proc. 12th International Workshop on Structural Health Monitoring (IWSHM).

Geurten, M., Willems, S., Germain, S. & Meulemans, T., 2015. Less is more: The availability heuristic in early childood. *British Journal of Developmental Psychology,* Volume 33, pp. 405-410.

Gigerenzer, G., 1995. How to improve Bayesian Reasoning Without Instructuion: Frequency Formats. *Psychological Review,* 102(4), pp. 684-704.

Gigerenzer, G., 2004. Fast and Frugal Heuristics: The Tools of Bounded Rationality. In: *Blackwell handbook of judgment and decision making.* Oxford, UK: Blackwell, p. 62–88.

Gigerenzer, G. & Gaissmaier, W., 2011. Heuristic Decision Making. *Annual Review of Psychology,* pp. 451-482.

Gigerenzer, G. & Gaissmaier, W., 2011. Heuristic Decision Making. *Annual Review of Psychology,* Volume 62, pp. 451-482.

Gigerenzer, G., Todd, P. M. & ABC Research Group, 1999. *Simple Heuristics That Make Us Smart.* New York: Oxford University Press.

Gilovich, T., Griffin, D. W. & Kahneman, D., 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment.* New York: Cambridge University Press.

Giordano, P. F., Prendergast, L. J. & Limongelli, M. P., 2020. A framework for assessing the value of information for health monitoring of scoured bridges. *Journal of Civil Structural Health Monitoring,* 10(3), pp. 485-496.

Glisic, B. & Adriaenssens, S., 2010. *Streicker Bridge: initial evaluation of life-cycle cost benefits of various structural health monitoring approaches.* Philadelphia, PA, In: Proceedings of the 5th international conference on bridge maintenance, safety and management.

Glisic, B., Chen, J. & Hubbell, D., 2011. *Streicker Bridge: a comparison between Bragg-gratings long-gauge strain and temperature sensors and Brillouin scattering-based distributed strain and temperature sensors.* San Diego, CA, In: Proceedings of SPIE: health monitoring of structural and biological systems.

Glisic, B. & Inaudi, D., 2012. Development of method for in-service crack detection based on distributed fiber optic sensors. *Structural Health Monitoring,* 11(2), pp. 696-711.

Gong, C. & Frangopol, D. M., 2020. Condition-Based Multiobjective Maintenance Decision Making for Highway Bridges Considering Risk Perceptions. *Journal of Structural Engineering,* 146(5).

Goodarzi, E., Mirzaei, M. & Ziaei, M., 2012. Evaluation of dam overtopping risk based on univariate and bivariate flood frequency analyses. *Canadian Journal of Civil Engineering,* 39(4), pp. 374-387.

Gosling, J. P., 2018. SHELF: the Sheffield elicitation framework. In: M. Q. Dias, ed. *Elicitation: the science and art of structuring judgment.* New York: Springer.

Goulet, J. A., Der Kiureghian, A. & Li, B., 2015. Pre-posterior optimization of sequence of measurement and intervention actions under structural reliability constraints. *Structural Safety,* Volume 52, pp. 1-9.

Gregory, P., 2005. Bayesian logical data analysis for the physical sciences. Cambridge: Cambridge University Press.

Grether, D. M., 1980. Bayes Rule as a Descriptive Model: The Representativeness Heuristic. *Quarterly Journal of Economics,* 95(3), pp. 537-557.

Grether, D. M., 1992. Testing Bayes rule and the representativeness heuristic: some experimental evidence. *Journal of Economic Behavior and Organization,* 17(1), pp. 31-57.

Gret-Regamey, A. & Straub, D., 2006. Spatially explicit avalanche risk assessment linking Bayesian networks to a GIS. *Natural Hazards and Earth Science, Copernicus Publications on behalf of the European Geosciences Union,* 6(6), pp. 911-926.

Griffin, D. & Tversky, A., 1992. The weighing of evidence and the determinants of confidence. *Cognitive Psychology,* 24(3), pp. 411-435.

Griffiths, T. L. & Tenenbaum, J. B., 2007. From mere coincidences to meaningful discoveries. *Cognition,* 103(2), pp. 180-226.

Hodge, R. et al., 2001. Eliciting engineering knowledge about reliability during design-lessons learnt from implementation. *Quality and reliability engineering international,* 17(3), pp. 169-179.

Hogarth, R. M. & Einhorn, H. J., 1992. Order effects in belief updating: the belief-adjustment model. *Cognitive Psychology,* Volume 24, pp. 1-55.

Holton, G., 1988. *Thematic Origins of Scientific Thought.* Cambridge, MA: Harvard Univ. Press. 2nd ed..

Honfi, D., 2019. *Practical issues in collecting and evaluating structural health issues on bridges.* Stanford (California), Proc. 12th International Workshop on Structural Health Monitoring (IWSHM).

Hora, S. C., 2007. Eliciting probabilities from experts. In: W. Edwards, R. F. Miles & D. von Winterfeldt, eds. *Advances in decision analysis: From foundations to applications.* s.l.:Cambridge University Press, pp. 129-153.

Iannacone, L., Gardoni, P., Giordano, P. F. & Limongelli, M. P., 2019. *Decision making based on value of information of different inspection methods.* Stanford (California), Proc. 12th International Workshop on Structural Health Monitoring (IWSHM).

Jacoby, L. L., Kelley, C. M., Brown, J. & Jasachko, J., 1989. Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, Volume 56, pp. 326-338.

Jensen, F. V. & Nielsen, T. D., 2007. *Bayesian networks and decision graphs.* 2nd ed. Berlin: Springer.

Jetter, M. & Walker, J. K., 2017. Anchoring in financial decision-making: Evidence from Jeopardy!. *Journal of Economic Behavior & Organization*, Volume 141, pp. 164-176.

Johnson, W., 1983. 'Representativeness' in Judgmental Predictions of Corporate Bankruptcy. *The Accounting Review*, 58(1), pp. 78-97.

Jong, C. G. & Leu, S. S., 2013. Bayesian-network-based hydro-power fault diagnosis systemdevelopment by fault tree transformation. *Journal of Marine Science and Technology*, 21(4), pp. 367-379.

Kahneman, D., Slovic, P. & Tversky, A., 1982. *Judgment under uncertainty: Heuristic and biases.* New York: Cambridge University Press.

Kahneman, D. & Tversky, A., 1972. Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology,* Volume 3, pp. 430-454.

Kahneman, D. & Tversky, A., 1973. Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology 5,* 5(2), pp. 207-232.

Kahneman, D. & Tversky, A., 1973. On the psychology of prediction. *Psychological review,* 80(4), pp. 237-251.

Kahneman, D. & Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica,* 47(2), pp. 263-292.

Kahneman, D. & Tversky, A., 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty,* Volume 5, pp. 297-323.

Kahneman, D. & Twersky, A., 1984. Choices, values and frames. *American Psychologist*, 39(4), pp. 341-350.

Kang, D. H., Park, S. O., Hong, C. S. & Kim, C. G., 2007. Mechanical strength characteristics of fiber bragg gratings considering fabrication process and reflectivity. *Journal of Intelligent Material Systems and Structures,* 18(4), pp. 303-309.

Katsikopoulos, K. V., 2011. Psychological Heuristics for Making Inferences: Definition, Performance, and the Emerging Theory and Practice. *Decision Analysis 8,* pp. 10-29.

Kelangath, S., Purnendu, K. D., Quigley, J. & Hirdaris, S. E., 2011. Risk analysis of damaged ships - a data-driven Bayesian approach. *Ships and Offshore Structures,* pp. 1-15.

Ketzenberg, M. E., Rosenzweig, E. D., Marucheck, A. E. & Matters, R. D., 2007. A framework for the value of information in inventory replenishment. *European Journal of Operational Research,* Volume 182, pp. 1230-1250.

Khakzad, N., Khan, F. & Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety,* Volume 96, pp. 925-932.

Kudryavtsev, A., 2018. The availability heuristic and reversals following large stock price changes. *Journal of Behavioral Finance,* 19(2), pp. 159-176.

Kusev, P. et al., 2018. Adaptive Anchoring Model: How Static and Dynamic Presentations of Time Series Influence Judgments and Predictions. *Cognitive Science,* Volume 42, pp. 77-102.

Langsetha, H. & Portinaleb, L., 2007. Bayesian networks in reliability. *Reliability Engineering and System Safety,* 92(1), pp. 92-108.

Larijani, H. & Ahmadinia, A., 2018. Random neural network learning heuristics. *Probability in the Engineering and Informational Sciences,* 32(3).

Leonard, N. E., 2014. On human decision-making heuristics and algorithm design in engineering. *Proc. IEEE,* 102(4).

Li, C. & Mahadevan, S., 2018. Sensitivity analysis of a Bayesian Network. *Journal of Risk and Uncertainty in Engineering Systems,* Volume 4.

Lieder, F., 2012. Burn-in, bias, and the rationality of anchoring. *Advances in neural information processing systems.*

Li, H., Li, J. & Kang, F., 2011. Risk analysis of dam based on artificial bee colony algorithm with fuzzy c-means clustering. *Canadian Journal of Civil Engineering,* 38(5), pp. 483-492.

Limongelli, M. P., Omenzetter, P., Yazgan, U. & Soyoz, S., 2017. *Quantifying the value of monitoring for post-earthquake emergency management of bridges.* Vancouver, Proc. IABSE.

Lindley, D. V., 1956. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics,* 27(4), pp. 986-1005.

Lindley, D. V., 2006. *Understanding uncertainty.* Wiley.

Ling, Y. & Mahadevan, S., 2012. Integration of Structural Health Monitoring and Fatigue Damage Prognosis. *Mechanical Systems and Signal Processing,* Volume 28, pp. 89-104.

Liu, Z. Q. & Nadim, F., 2013. *Publications on Quantitative risk assessment for eartquake-triggered landslides using Bayesian network.* Paris, Proc. 18th International Conference on Soil Mechanics and Geotechnical Engineering.

Liu, Z. Q., Nadim, F., Eidsvig, U. K. & Lacasse, S., 2017. Reassessment of Dam Safety Using Bayesian Network.. *Geo-Risk*, pp. 168-177.

Luo, G., 2013. Can representativeness heuristic traders survive in a competitive securities market?. *Journal of Financial Markets,* 16(1), pp. 152-164.

MacKay, D. J. C., 2003. Information theory, inference and learning algorithms. Cambridge: Cambridge University Press.

Manis, M., Shedler, J., Jonides, J. & Nelson, T. E., 1993. Availability heuristic in judgments of set size and frequency of occurrence. *Journal of Personality and Social Psychology,* 65(3), pp. 448-457.

Maroni, A. et al., 2019. *A decision support systemfor scour management of road and railway bridges based on Bayesian networks.* Stanford, Proc. 12th International Workshop on Structural Health Monitoring.

Martinelli, A., 2012. An emerging paradigm or just another trajectory? Understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry. *Research Policy,* Volume 41, pp. 414-429.

McConnell, C. R., 1966. *Economics: Principles, Problems, and Policies - 3rd Edition.* New York: McGraw-Hill Book Company.

McKelvie, S. J. & Drumheller, A., 2001. The availability heuristic with famous names: a replication. *Perceptual and motor skills,* Volume 92, pp. 507-516.

Medina-Cetina, Z. & Nadim, F., 2008. Stochastic design of an early warning system. *Georisk: Assessment and Management of Risk for Engineerign Systems and Geohazards,* 2(4), pp. 223-236.

Melchers, R. E., 1999. *Structural reliability: analysis and prediction*. 2nd ed. Chichester: John Wiley & Sons.

Memarzadeh, M., Pozzi, M. & Zico Kolter, J., 2014. Optimal planning and learning in uncertain environments for the management of wind farms. *Journal Comp. Civ. Eng.*.

Merkhofer, M., 1987. Quantifying judgemental uncertainty: methodology, experiences, and insights. *IEEE Transactions on Systems, Man, and Cybernetics,* 17(5), pp. 741-52.

Meub, L. & Proeger, T. E., 2015. Anchoring in social context. *Journal of Behavioral and Experimental Economics,* Volume 55, pp. 29-39.

Meyer, M. & Booker, J., 1991. *Eliciting and analyzing expert judgment - a practical guide.* New York: Academic Press.

Minsky, M. L., 1961. A selected descriptor-indexed bibliography to the literature on artificial intelligence. *IRE Transactions on Human Factors in Electronics, 2,* pp. 39-55.

Minsky, M. L., 1961. Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers, 9,* pp. 8-30.

Miroslaw-Swiatek, D., Kemblowski, M. & Jankowski, W., 2012. Application of the Bayesian Belief Nets in dam safety monitoring. *Annals of Warsaw University of Life Sciences - SGGW. Land reclamation,* 44(1), pp. 25-33.

Mochon, D. & Frederick, S., 2013. Anchoring in sequential judgments. *Organizational Behavior and Human Decision Processes,* Volume 122, pp. 69-79.

Morgan, M. G., Small, M. J. & Henrion, M., 1990. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis.* New York: Cambridge University Press.

Murphy, K., 2012. Machine learning, a probabilistic perspective. MIT Press.

Mussweiler, T., Strack, F. & Pfeiffer, T., 2000. Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility. *Personality and Social Psychology Bulletin,* 26(9), pp. 1142-1150.

Mthembu, L., Marwala, T., Friswell, M. I. & Adhikarid, S., 2011. Model selection in finite element model updating using the Bayesian evidence statistic. *Mech. Syst. Signal Process,* 25(7), pp. 2399-2412.

Neumann, J. V. & Morgenstern, O., 1944. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Nikles, M., Thevenaz, L. & Robert, P., 1996. Simple distributed fiber sensor based on Brillouin gain spectrum analysis. *Optics Letters,* 21(10), pp. 758-760.

Nikolova, E., 2010. High-performance heuristics for optimization in stochastic traffic engineering problems. *Lecture Notes in Computer Science,* pp. 352-360.

Norrington, L., Quigley, J., Russell, A. & Van der Meer, R., 2008. Modelling the reliability of search and rescue operations with Bayesian Belief Networks. *Reliability Engineering and System Safety,* 93(7), pp. 940-949.

O'Hagan, A. et al., 2006. *Uncertainty Judgments: Eliciting Experts' Probabilities.* Wiley.

Oppenheimer, D. M., 2004. Spontaneus discounting of availability in frequency judgment tasks. *Psychological Science,* Volume 15, pp. 100-105.

Pachur, T., Steinmann, F. & Hertwig, R., 2012. How do people judge risks: availability heuristic, affect heuristc, or both?. *Journal of Experimental Psychology,* 18(3), pp. 314-330.

Parmigiani, G. & Inoue, L., 2009. *Decision Theory: Principles and Approaches.* Chichester: Wiley.

Pearl, J., 1984. *Heuristics: intelligent search strategies for computer problem solving.* London: Addison-Wesley Publ. Co..

Peng, M. & Zhang, L. M., 2013. Dynamic decision making for dam-break emergency management – Part 1: Theoretical framework. *Natural Hazards & Earth System Sciences*, Volume 13, pp. 425-437.

Peter, C. et al., 2009. Applying Bayesian modelling to assess climate change effects on biofuel production. *Climate research,* Volume 40, pp. 240-260.

Pfeifer, C., Schredelseker, K. & Seeber, G. U. H., 2009. On the negative value of information in informationally inefficient markets: Calculations for large number of traders. *European Journal of Operational Research*, Volume 195, pp. 117-126.

Polya, G., 1945. *How to solve it.* New York: Doubleday Anchor.

Pope, R., 1986. Consistency and Expected Utility Theory. In: L. Daboni, A. Montesano, M. Lines, eds. *Recent Developments in the Foundations of Utility and Risk Theory*. Theory and Decision Library, vol 47. Springer, Dordrecht.

Pozzi, M. & Der Kiureghian, A., 2011. *Assessing the value of information for long-term structural health monitoring.* San Diego, CA, In: Proceeding of SPIE: health monitoring of structural and biological systems.

Pozzi, M., Malings, C. & Minca, A. C., 2017. *Negative value of information in systems' maintenance.* Vienna, Austria, 12th International Conference on Structural Safety and Reliability.

Pozzi, M., Milangs, C. & Minca, A., 2020. Information avoidance and overvaluation under epistemic constraints: Principles and implications for regulatory policies. *Reliability Engineering and System Safety*, Volume 197.

Pozzi, M., Zonta, D., Wang, W. & Chen, G., 2010. *A framework for evaluating the impact of structural health monitoring on bridge management.* Philadelphia, PA, In: Proceeding of the 5th international conference on bridge maintenance, safety and management (IABMAS 2010).

Pratt, J. W., 1964. Risk Aversion in the Small and in the Large. *Econometrica*, 32(1-2), pp. 122-136.

Quigley, J., Colson, A., Aspinall, W. & Cooke, R., 2018. Elicitation in the Classical Model. In: Dias, Morton & Quigley, eds. *Elicitation: The Science and Art of Structuring Judgement.* Springer.

Quigley, J. & Walls, L., 2018. A methodology for constructing subjective probability distributions with data. In: L. C. Dias, A. Morton & J. Quigley, eds. *Elicitation.* pp. 141-170.

Quigley, J. & Walls, L., 2020. Characteristics of a Process for Subjective Probability Elicitation. In: French, Bedford, Nane & Hanea, eds. *Expert judgment in Risk and Decision Analysis.* Springer International Publishing.

Quigley, J. et al., 2017. Supplier quality improvement: The value of information under uncertainty. *European Journal of Operational Research,* Volume 264, p. 932:947.

Raiffa, H. & Schlaifer, R., 1961. *Applied Statistical Decision Theory.* Boston, MA: Clinton Press.

Rizak, S. & Hrudey, S. E., 2005. Interdisciplinary comparison of expert risk beliefs. *Journal of Environmental Engineering Science,* 4(3), pp. 173-185.

Rodriguez, N., Gupta, A., Zabala, P. L. & Cabrera-Guerrero, G., 2018. Optimization algorithms combining (Meta)heuristics and mathematical programming and its application in engineering. *Mathematical problems in engineering.*

Romanycia, M. H. J. & Pelletier, F. J., 1985. What is a heuristic?. *Computational Intelligence,* 1(1), pp. 47-58.

Russo, E. J., 2010. Understanding the effect of a numerical anchor. *Journal of Consumer Psychology,* Volume 20, pp. 25-27.

Sahin, F. & Robinson, E. P., 2002. Flow coordination and information sharing in supply chains: Review, implications, and directions for future research. *Decision Sciences,* 334(4), pp. 505-536.

Sankararaman, S., Ling, Y. & Mahadevan, S., 2011. Uncertainty Quantification and Model Validation of Fatigue Crack Growth Prediction. *Engineering Fracture Mechanics,* 78(7), pp. 1487-1504.

Schredelseker, K., 2001. Is the usefulness approach useful? Some reflections on the utility of public information. In: S. McLeay & A. Riccaboni, eds. *Contemporary Issues in Accounting Regulation.* Boston: Kluwer Academic Publishers, pp. 135-153.

Shah, A. K. & Oppenheimer, D. M., 2008. Heuristics made easy: an effort-reduction framework. *Psychol. Bull. 137,* p. 207–222.

Shrum, L. J., 1996. Psychological process underlying cultivation effects. *Human Communication Research,* Volume 22, pp. 482-509.

Shrum, L. & O'Guinn, T. C., 1993. Process and effects in the construction of social reality: Construct accessibility as an explanatory variable. *Communication Research,* Volume 20, pp. 436-471.

Sigurdsson, J. H., Walls, L. A. & Quigley, J., 2001. Bayesian belief nets for managing expert judgment and modelling reliability. *Quality and reliability engineering international,* Volume 17, pp. 181-190.

Simon, H. A., 1956. Rational choice and the structure of the environment. *Psychological Review, 63,* p. 129–138.

Simon, H. A., 1989. The scientist as problem solver. In: *Complex Information Processing: The Impact of Herbert A. Simon.* Hillsdale, NJ: Erlbaum, p. 373–398.

Sivia, D. & Skilling, J., 2006. *Data analysis: A Bayesian Tutorial.* Oxford: Oxford University Press.

Slovic, P. & Lichtenstein, S., 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance 6,* pp. 649-744.

Smith, M., 2006. *Dam Risk Analysis Using Bayesian Networks.* Lillehammer, Norway, Proc. Engineering Conferences International.

Sohn, H. & Law, K. H., 1997. A Bayesian probabilistic approach for structure damage detection. *Earthquake Engineering and Structural Dynamics,* 26(12), pp. 1259-1281.

Spetzler, C. & Stael Von Holstein, C. A., 1985. Probability Encoding in Decision Analysis. *TIMS,* Issue 22, pp. 340-358.

Straub, D. et al., 2017. *Value of Information: A roadmap to quantifying the benefit of structural health monitoring.* Vienna, Austria, 12th International Conference on Structural Safety & Reliability (ICOSSAR 2017).

Straub, D. & Faber, M., 2005. Risk based inspection planning for structural systems. *Structural safety,* 27(4), pp. 335-55.

Studer, J. A., Yilmaz, S., Daly, S. R. & Seifert, C. M., 2016. *Cognitive heuristics in defining engineering design problems.* Proc. ASME Design Engineering Technical Conference.

Su, H., Wen, Z., Xiaoran, S. & Yang, M., 2015. Time-varying identification model for dam behavior considering structural reinforcement. *Structural Safety,* Volume 57, pp. 1-7.

Taylor, S. E. & Thompson, S. C., 1982. Stalking the elusive "vividness" effect. *Psychological Review,* Volume 89, pp. 155-181.

Tenenbaum, J. B. & Griffiths, T., 2001. *The rational Basis of Representativeness.* Edinburgh, Proc. 23rd Annual Conf. of the Cognitive Science Society.

Thoft-Christensen, P. & Sorensen, J., 1987. Optimal strategy for inspection and repair of structural systems. *Civil Engineering Systems,* 4(2), pp. 94-100.

Thons, S., 2017. On the Value of Monitoring Information for the Structural Integrity and Risk Management. *Computer-Aided Civil and Infrastructure Engineering,* Volume 33, pp. 79-94.

Thons, S. & Faber, M. H., 2013. *Assessing the Value of Structural Health Monitoring.* New York, USA, In 11th International Conference on Structural Safety & Reliability (ICOSSAR 2013).

Thons, S. et al., 2017. *Progress of the COST Action TU1402 on the Quantification of the Value of Structural Health Monitoring.* Stanford, USA, 11th International Workshop on Structural Health Monitoring.

Todhunter, I., 1876. *William Whewell, D. D., master of Trinity college, Cambridge. An account of his writings with selections from his literary and scientific correspondence.* London: Macmillan & Co..

Tonelli, D. et al., 2017. *Expected utility theory for monitoring-based decision support system.* Stanford, USA, Proceedings of the 11th International Workshop on Structural Health Monitoring.

Tonge, F. M., 1960. Summary of a heuristic line balancing procedure. *Management Science,* 7(1), pp. 21-42.

Turner, B. M. & Schley, D. R., 2016. The anchor integration model: A descriptive model of anchoring effects. *Cognitive Psychology,* Volume 90, pp. 1-47.

Tversky, A. & Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. *Science, New Series*, Volume 185, pp. 1124-1131.

Tversky, A. & Kahneman, D., 1983. Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review,* 90(4), pp. 293-315.

Tversky, A. & Koehler, J., 1994. Support theory: a nonextensional representation of subjective probability. *Psychological review,* 101(4), pp. 547-567.

Valkonen, X. A. & Glisic, B., 2019. *Measurement of individual risk preference for decision-making in SHM.* Stanford (California), Proc. 12th International Workshop on Structural Health Monitoring (IWSHM).

Vanik, M. W., Beck, J. L. & Au, S. K., 2000. Bayesian probabilist approach to structural health monitoring. *Journal of Engineering Mechanics,* 126(7), pp. 738-745.

Verzobio, A., Bolognani, D., Quigley, J. & Zonta, D., 2018. *The consequences of Heuristic distortions on SHM-based decision problems.* Manchester, UK, Proceedings of the 9th European Workshop of Structural Health Monitoring.

Verzobio, A. et al., 2019. *Elicitation Process to Populate Bayesian Networks with Application to Dam Safety.* Nanjing, China, 8th International Conference on Water Resources and Environmental Research (ICWRER).

Vick, S. G., 2002. *Degrees of Belief: Subjective Probability and Engineering Judgment.* ASCE Press.

Wagner, H. M., 1969. *Principles of operations research.* Englewood Cliffs ed. NJ: Prentice-Hall.

Wakker, P. P., 2008. *Prospect Theory for Risk and Ambiguity.* Cambridge University Press.

Watkins, M. J. & LeCompte, D., 1991. Inadequacy of recall as a basis for frequency knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* Volume 17, pp. 1161-1176.

Wegener, D. T., Petty, R. E., Blankenship, K. L. & Detweiler-Bedell, B., 2010. Elaboration and numerical anchoring: Implications of attitude theories for consumer judgment and decision making. *Journal of Consumer Psychology,* Volume 20, pp. 5-16.

Welsh, M. B., Delfabbro, P. H., Burns, N. R. & Begg, S. H., 2014. Individual differences in anchoring: Traits and experience. *Learning and Individual Differences,* Volume 29, pp. 131-140.

Woodland, L. & Woodland, B., 2015. The National Football League season wins total betting market: The impact of heuristics on behavior. *Southern Economic Journal,* 82(1), pp. 38-54.

Xu, Y., Zhang, L. M. & Jia, J. S., 2011. Diagnosis of embankment dam distresses using Bayesian networks. Part II: Diagnosis of a specific distressed dam. *Canadian geotechnical journal,* pp. 1645-1657.

Yakovis, L. & Chechurin, L., 2015. Creativity and heuristics in process control engineering. *Chemical Engineering Research and Design,* Volume 103, pp. 40-49.

Yanmaz, A. M. & Gunindi, M. E., 2008. Assessment of overtopping reliability and benefits of a flood detention dam. *Canadian Journal of Civil Engineering,* 35(10), pp. 1177-1182.

Yuen, K. V. & Kouk, S. C., 2011. Bayesian methods for updating dynamics models. *Appl. Mech. Rev.,* 64(1).

Zerrouki, H. & Tamrabet, A., 2015a. *Fault tree and Bayesian Network in safety analysis: Comparative study.* Djelfa, Algeria, Proc. International Conference on Applied Automation and Industrial Diagnostics.

Zerrouki, H. & Tamrabet, A., 2015b. *Mapping Fault Tree into Bayesian Network in safety analysis of process system.* Boumerdes, Proc. 4th International Conference on Electrical Engineering.

Zhang, L. M., Xu, Y., Jia, J. S. & Zhao, C., 2011. Diagnosis of embankment dam distresses using Bayesian networks. Part I: Global-level characteristics based on a dam distress database. *Canadian geotechnical journal,* pp. 1630-1644.

Zhang, W. H., Lyu, D. G., Faber, M. H. & Thons, S., 2019. *Value of the SHM system in the context of risk-based inspection planning.* Stanford, California, Proc. 12th International Workshop on Structural Health Monitoring (IWSHM).

Zonta, D., Glisic, B. & Adriaenssens, S., 2014. Value of information: impact of monitoring on decision-making. *Structural Control Health Monitoring,* Volume 21, p. 1043:1056.

Zonta, D., Zandonini, R. & Bortot, F., 2007. A reliability-based bridge management concept. *Structure and Infrastructure Engineering,* 3(3), pp. 215-235.