Sequential decision making with adaptive utility

Houlding, Brett

Sequential Decision Making with Adaptive Utility

Brett Houlding

A Thesis presented for the degree of
Doctor of Philosophy

Department of Mathematical Sciences
Durham University
UK

May 2008

Dedication

To my parents, for their ever enduring support.

Sequential Decision Making with Adaptive Utility

Brett Houlding

Submitted for the degree of Doctor of Philosophy
May 2008

Abstract

Decision making with adaptive utility provides a generalisation to classical Bayesian decision theory, allowing the creation of a normative theory for decision selection when preferences are initially uncertain.

The theory of adaptive utility was introduced by Cyert & DeGroot [27], but had since received little attention or development. In particular, foundational issues had not been explored and no consideration had been given to the generalisation of traditional utility concepts such as value of information or risk aversion. This thesis addresses such issues.

An in-depth review of the decision theory literature is given, detailing differences in assumptions between various proposed normative theories and their possible generalisations. Motivation is provided for generalising expected utility theory to permit uncertain preferences, and it is argued that in such a situation, under the acceptance of traditional utility axioms, the decision maker should seek to select decisions so as to maximise expected adaptive utility. The possible applications of the theory for sequential decision making are illustrated by some small-scale examples, including examples of relevance within reliability theory.

Declaration

The work in this thesis is based on research carried out at the Department of Mathematical Sciences, Durham University, UK. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all the author's original work unless referenced to the contrary in the text.

Acknowledgements

First I would like to thank my supervisor, Prof. Frank Coolen, for his expert advice and guidance. The discussions we had on decision theory, and the resulting suggestions for progress, were very much appreciated. Also within the Durham Department of Mathematics Sciences, I would like to acknowledge the support of both Prof. Michael Goldstein and Dr. Peter Craig, and I am grateful for their interest and help in my research.

Amongst my postgraduate colleagues, I would like to thank Alicia, Anisee, Becky, Ric, Jonathan, Ian, Danny and Becca, all of whom made my work very enjoyable. I am especially appreciative of Alicia and Anisee, with whom I had the good fortune to share an office with for three and a half years.

I am also indebted to various members of Trevelyan College, and in particular I wish to acknowledge the support that was afforded me at various times by Zara, Kate, and Maggie.

I would also like to thank Pamela for her support. Not only does she possess the wonderful ability to find the funny side in any dilemma, but she also supplied the essential fuel of morning coffee.

Finally I acknowledge the support of my family, without which I would not have been able to embark on this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter offers explanation of the type of decision theory this thesis is concerned with, and introduces the fundamental concepts of probability and utility that are used to measure beliefs and preferences, respectively. Finally, we discuss the notational conventions that are to be employed throughout, and provide an outline to the focus of subsequent chapters.

1.1 Problem Description

This thesis is concerned with normative decision theory. Normative here is used to denote theories that describe the action a decision maker (DM) should take if she agrees with a small number of axioms of preference. That is to say, agreement with such axioms leads to a direct logical argument for detailing a set of possible decisions for selection. Such axioms are created through philosophical consideration and are assumed to be in agreement with the fundamental beliefs of a rational and coherent DM.

In contrast to normative decision theories, descriptive decision theories seek to explain real-world selection of decisions. Descriptive decision theories employ psychological analysis in an attempt to explain or predict the actual actions of real-world DMs. Such theories are not the focus of this research, but in Subsection 2.2.2 we mention some alterations that have been made to normative theories in order to increase their descriptive ability.

We consider a DM facing either a solitary decision, or a sequence of decisions. In the latter case, the DM observes results of previous choices before selection of the next, and throughout the thesis we permit uncertainty over the result of decision selection. The main focus will be to develop a theory that also permits initial uncertainty over preferences, but where a DM may learn about these through trial.

We begin by introducing the concepts of (subjective) probability and utility that are relevant when considering decision selection. Only a short summary is provided, and the interested reader is referred to more detailed accounts available in decision theory text books such as Clemen [24], French & Rios Insua [46] or Lindley [71]. The section concludes by formally introducing the decision problem under consideration.

1.1.1 Probability & Utility

Uncertainty over events will be modelled through the DM's degree of belief, *i.e.*, assuming the DM is acting in a rational manner (later discussed in Section 2.1) we work with the interpretation of *probability* as the DM's subjective belief given her personal background knowledge and experience (though we will not formally include this in notation). Objective probabilities still arise in discussion, but we assume that if events of interest refer to outcomes of fair chance mechanisms, *e.g.*, a roulette, then the DM's subjective probability will agree with that dictated by the classical theory of probability, as first developed following correspondence between Pascal and Fermat in the seventeenth century (see, *e.g.*, Hacking [52]).

The arguments of authors such as de Finetti [31] and Savage [89] state that the subjective probability of an event h occurring can be elicited as follows. The subjective probability of event h, denoted $P(h)$, is the fair price, as viewed by the DM, for entering the bet paying one unit if h occurs, and nothing otherwise. Assuming that the DM is acting rationally, it can be shown that such a definition agrees with the DM's beliefs, *e.g.*, the DM should specify a greater price if and only if she has a greater belief in the occurrence of h. It can also be shown that this definition satisfies all of Kolmogorov's [66] axioms of probability for the case of a finite set of possible events (again assuming the prerequisite of a rational DM).

In contrast to probabilities measuring degree of belief, a *utility* function measures the DM's subjective preferences over decisions and outcomes. The use of a non-linear function for determining the worth of a reward was first suggested by Bernoulli [18], and an axiomatization for the existence of such a function was later developed by von Neumann & Morgenstern [101]. A utility function u is formally defined as a function with domain the set of randomized decisions $\mathcal{D}$ and co-domain the set of reals $\mathbb{R}$, with the property that it is in agreement with the DM's preferences, *i.e.*, if the DM strictly prefers decision d_1 to decision d_2, then $u(d_1) > u(d_2)$. The utility of a specific reward or decision outcome is then determined by considering the utility of the degenerate decision that leads to that specific reward or outcome with certainty. In practice, however, utility functions are considered as having domain the set of all possible randomized rewards $\mathcal{R}$, and the utility of a decision is then determined by considering the expected utility that it will entail, *i.e.*, the utility of the decision d leading to reward r with probability $P(r|d)$ is determined by the expectation $\sum_{r\in\mathcal{R}} u(r)P(r|d)$ (the sum being replaced by an integral if beliefs are represented by a probability density function). In this thesis we will at times consider both the possibilities of $\mathcal{R}$ or $\mathcal{D}$ for the domain of a utility function, using the relationship $u(d) = \sum_{r\in\mathcal{R}} u(r)P(r|d)$ to interchange from one to the other.

In parts of the Economics literature, utility is seen as an ordinal concept (see Abdellaoui *et al.* [1]), hence to prevent any potential misunderstanding we will make the following distinction between a utility function and a value function. A utility function u, often referred to as a cardinal utility function, provides the 'moral' worth of an outcome. A value function v, often referred to as an ordinal utility function, is a more primitive concept that simply ranks outcomes in a manner consistent with preferences. In particular, value functions do not take into account relative strength of preference, and these will not be considered further (see Keeney

& Raiffa [64, Ch.3] for more information on value functions).

Fundamentally, subjective probabilities and utilities can be viewed as twin concepts (see, *e.g.*, the discussion in French [45]). Indeed, through the betting price interpretation of subjective probability it is difficult to formally define one without making explicit reference to the other, and this is also the case in elicitation. For example, when above we discussed the interpretation of subjective probability as a fair betting price, the return of the bet should actually have been expressed in utility units. Similarly, when eliciting utility values the knowledge of subjective probabilities is also required. Assuming that a most preferred reward r^* has a utility value of 1 and a least preferred reward r_* has a utility value of 0, the utility value of any other reward r is equal to the subjective probability p such that the DM is indifferent between obtaining r for certain, or risking the gamble that results in r^* with probability p and r_* otherwise. Furthermore, the work of Anscombe & Aumann [5] (discussed in Section 2.2) provides a method of defining subjective probability and utility simultaneously from a single preference relation. In this setting duality exists in so far that both subjective probabilities and utilities are measured by comparing preferences over gambles whose outcomes are determined by objective probabilities.

1.1.2 Decision Selection

Given her relevant utility and probability specifications, the problem of the DM is to select a decision d out of a set of possibilities $\mathcal{D}$. The problem of the decision analyst is to determine a logical system for explaining how and why a specific choice should be made. There are many different types of decision problem, but throughout this thesis we concern ourselves with the case of a single DM who is motivated to select the decision that is best for her (utility returns only accrue to the DM), and where the outcome of a decision is selected by an unconcerned Nature. This is in contrast to game theory where the DM faces an intelligent and motivated opponent (see, *e.g.*, Luce & Raiffa [72]).

We concern ourselves with two situations. The first is selection of a solitary decision, where the DM has stated belief and preference specifications before selection, and where the problem is completed as soon as the selection is made and return obtained. The second, more interesting, problem is when a sequence of decisions must be made. In this case, due to the extra information that may become available, or the relevant insights that may be made between one decision choice and the next, the DM can learn about likely decision outcome as she proceeds through the decision sequence.

1.2 Outline of Thesis

In comparison to alternative mathematical disciplines, the study of decision theory usually only requires a relatively low level of mathematical expertise. An undergraduate course in Probability and Bayesian Statistics should be sufficient to understand the majority of this thesis, and hence we assume that the reader has such knowledge. However, though of a relatively simple nature, it will become apparent that necessary calculations can be tedious and time consuming. When presented, the reader should be aware that numerical results were determined through use of the software package Maple 10. Nevertheless, although the technical level of the mathematics is low, the complexity of arguments and level of understanding necessary to produce solutions is high. In reading this thesis a general understanding of decision theory is useful, but not essential, and we will seek to explain the necessary decision theoretic terms that have been included. When it is deemed inappropriate to include full replication of standard results, the reader will be directed to relevant sources.

To ease explanation of the theory a number of examples will be provided. Whenever possible, the re-examination of previous examples will be employed when highlighting new aspects of the theory, and it is hoped that this will enable familiarity with these problems. However, in special situations, new and different examples will be considered when it is believed that these will either highlight the issues in the theory more clearly, or if the new example is deemed of interest itself. We will mark the end of an example, and the return to general discussion, through the use of a $\diamond$ symbol.

We have sought to use standard decision theory notation as much as possible within this thesis, however, unless mentioned otherwise, we employ a notational convention such that, in general, right-hand subscripts denote different values for a decision or reward *etc.*, or when placed beside an operator, denote the state variable that operator is connected with. In a sequential problem, right-hand superscripts will be used to denote the epoch a reward or decision is being considered in. For example, r_3^2 is a particular reward value to be received in the second period, whilst $E_X[Y]$ is the expected value of Y with respect to beliefs over X.

When required, we will also highlight that we are considering functions or operators within an adaptive utility setting by placing an a in the left-hand subscript position. For example, we differentiate a classical utility function from an adaptive utility function, a concept to be formally defined in Chapter 3, by denoting the latter as ${}_au$ (a glossary of the main mathematical notation employed is available at the end of this thesis in Appendix A). Furthermore, in keeping with the tradition of the decision theory literature, DMs will be referred to as being female, whilst experts or analysts will be referred to as being male.

The format for the remainder of this thesis is as follows. Chapter 2 offers a review of known decision theories for the situation of a single DM. Chapter 3 introduces the adaptive utility concept and offers motivation for its use, a literature review of adaptive utility and similar theories is also provided. The contribution of this thesis to the study of decision theory commences in Chapter 4, where the foundational implications of permitting uncertain utility are considered and it is shown that a traditional system of axioms of preference is sufficient to entail the use of maximising expected adaptive utility as the logical decision selection rule.

The focus of the thesis changes in Chapter 5, where extra results are examined under the assumption of optimal decision selection through maximisation of expected adaptive utility. In particular, Chapter 5 considers solutions to sequential problems and illustrates possible applications of the theory for reliability decision problems through a couple of small hypothetical examples. Chapter 6 focuses on two characteristics of an adaptive utility function that have in the past been overlooked in the literature, considering implications for risk aversion and value of sample information. Finally, Chapter 7 provides concluding comments and potential directions for further research.

Three appendices are given at the end of the thesis. Appendix A provides a glossary of mathematical notation employed, Appendix B provides further discussion to Example 5.3.1, and Appendix C introduces a conjugate class of utility functions that is relevant for the discussion in Section 5.1.

Chapter 2

Review of Decision Theory

This chapter offers a brief review of the literature on decision making under uncertainty. Section 2.1 considers the meaning of rational probability specification and decision selection, and also interpretations for conditional beliefs. Section 2.2 briefly examines some of the theories that have been suggested for solving solitary decision problems, and the chapter concludes with a discussion of issues concerning utility forms for sequential problems in Section 2.3.

2.1 Rational and Coherent Decision Selection

We begin this chapter by briefly reviewing the meaning of rational decision selection, essential in explaining why a specific decision should be selected. We also consider methods of specifying beliefs, and the interpretations that can be given to conditional probabilities.

We make the following distinction between decisions that are admissible, and those that are merely feasible. A *feasible* decision is any that the DM can identify as a possible course of action. As a subset of these feasible decisions, we follow the suggestion of Levi [68] for stating the definition of an admissible decision. Given some criteria of rationality, a decision will be said to be *admissible* if and only if its selection does not contradict these criteria.

Hence, admissibility is a concept that depends on the specific criteria under consideration, and in this chapter we will discuss various possibilities that have been suggested.

We consider a DM to be acting in a rational manner if she is acting in agreement with an accepted system of axioms of preference. The particular system of axioms considered will hence provide the meaning of rationality. Although many differing axiomatic systems entail the same decision selection procedure (*e.g.*, the systems of Anscombe & Aumann [5] and Savage [89]), there are nevertheless varying suggestions that entail different decision selection procedures. A few of these will be reviewed in Section 2.2.

2.1.1 Coherence

Another concept of rationality arises when we consider the DM's belief specification, and following arguments by Ramsey [86, Ch.7] and de Finetti [31], we assume a DM is specifying rational beliefs if they are coherent. By coherent we mean that the DM's subjective probabilities are specified in such a way that it is not possible for her to wish to enter into a bet, or a system of bets, such that regardless of which event takes place, the DM will lose (*i.e.*, a Dutch Book can not be made against her). Although in this thesis we will take it as granted that the DM is specifying coherent beliefs, this can follow automatically from acceptance of a collection of axioms regarding the set of bets a DM would accept (see, for example, Walley's axioms of desirability that are used to imply coherence [104]).

It can be shown (*e.g.*, Kaplan [61, p.155]) that coherence implies that the DM's belief specification, assuming it is a precise specification over a finite event space, satisfies the Kolmogorov [66] axioms of probability. Nevertheless, we should be aware that the argument for coherence implying agreement with the Kolmogorov axioms does require that the DM cares to win bets, regardless of the amount at stake, and that she has an indifference to gambling. This will not be true for most real-world DM's, and is a difference between normative and descriptive theories. For the purpose of this thesis, we will imagine the DM is specifying coherent subjective probabilities.

2.1.2 Imprecision

Frequently, decision theories make an *a priori* assumption that the DM is able to fully express her beliefs and preferences through precise probability and utility statements. However, this can sometimes be an ambitious and unreasonable assumption. Whilst the focus of research in this thesis is based on the assumption of precise belief specifications, it will nevertheless be beneficial to review the meaning of imprecise probabilities and utilities in order to comment on some of the decision theories that have been designed to incorporate them.

Theories that imprecisely quantify uncertainty, sometimes referred to as non-additive or generalised theories, generalise classical results by permitting the DM to remain vague or even ignorant about actual probabilities or utilities. Though consideration of such a problem appears as early as the work of Boole [20], recent works such as Augustin [7] and Walley [104] (in the case of imprecise probabilities), and Farrow & Goldstein [40] and Moskowitz *et al.* [80] (in the case of imprecise utilities) demonstrate that this is still an area of interesting and active research.

Taking the subjective definition for the probability of an event h, $P(h)$, as being the fair price for entry into the bet paying one unit if h occurs and nothing otherwise, an often used method of permitting imprecision is to accept that this fair price may be difficult to identify. Instead a DM may only be willing to fix a maximum price $\underline{P}(h)$ for which she is happy to buy into the bet, and a higher value $\overline{P}(h)$ representing the minimum price she would be happy to sell the bet for. For prices between $\underline{P}(h)$ and $\overline{P}(h)$ the DM may not wish to commit to any fixed strategy.

The quantities $\underline{P}(h)$ and $\overline{P}(h)$ are, respectively, interpreted as the lower and upper subjective probabilities for the occurrence of event h. Provided $\underline{P}(h) \neq \overline{P}(h)$, there will be a whole class $\mathcal{P}$ of distributions which satisfy the constraints set out by the DM's betting behaviour, and only in the case that $\underline{P}(h) = \overline{P}(h)$ for all events h will

$\mathcal{P}$ reduce to containing a single distribution. In the more general case a DM must consider how best to select a decision when she only has the information concerning the set of possible distributions $\mathcal{P}$ and nothing more.

Imprecise utilities may also originate in a similar way, where only known bounds are stated for the utility value of a relevant reward. Alternatively, and as mentioned in Subsection 1.1.1, we can note that fundamentally utilities and probabilities may be seen as twin concepts that are both derived from a stated preference ordering. Yet if only a partial ordering of preferences is used, only imprecise probabilities and utilities will be available (see, *e.g.*, Seidenfeld *et al.* [93]). That is not to say that the use of imprecise probabilities necessitates the requirement for imprecise utilities and, for example, the works of Boole [20], Walley [104], and Williams [107] all deal with imprecise probability and precise utility simultaneously.

Whilst from a decision analysis context it is ideal to have known and precise probability and utility specifications, there are several arguments as to why this is not always the case, and Walley [104, Ch.1] provides a good overview in the case of probabilities. Levi [69] claims that a bounded rationality prevents DMs from fully comprehending all that is necessary for precision, and often required calculations are beyond computational abilities. Indeed, if a prior analysis identifies a unique decision that is optimal under all possible distributions that satisfy the constraints of imprecision, then the extra effort required in identifying a precise specification is just not needed.

2.1.3 Conditional Probabilities

Before discussing solutions to the general decision problem under consideration, we should first review the possible interpretations of conditional probability. That is to say, what is the interpretation of the conditional probability of event h_2 being true given that event h_1 is true, to be denoted by $P(h_2|h_1)$.

Providing $P(h_1) > 0$, we formally define $P(h_2|h_1)$ to be the numerical quantity

$P(h_2 \cap h_1)/P(h_1)$, with $P(h_2 \cap h_1)$ being the probability of the compound event $h_2 \cap h_1$ that is true if and only if both events h_2 and h_1 are true. However, there are various suggestions for its meaning (see, *e.g.*, the discussion in Kadane *et al.* [59]). The 'called-off gamble' interpretation arises from extending the subjective theories that view probability as a fair betting price, and is present in works such as de Finetti [31] and Savage [89]. Here one views $P(h_2|h_1)$ as the fair price for the bet paying one unit if h_2 is true and nothing otherwise, but where the bet is cancelled if h_1 is not found to be true.

As Kadane *et al.* [59] note, an alternative 'temporal updating' view is a common Bayesian interpretation for when dealing with sequential problems. It assumes that either $P(h_2|h_1)$ is the probability the DM expects to assign to h_2 in the case she learns h_1 is true and nothing else, or that it is the probability the DM will assign to h_2 in the case she learns h_1 and nothing more. In this thesis we seek to develop a strategy for sequential decision making from the view of a DM who is about to select her first decision and will view $P(h_2|h_1)$ as meaning the former of these.

Finally, 'hypothetical reasoning' is the view taken by Kadane *et al.* [59], and considers $P(h_2|h_1)$ to be the DM's current hypothetical belief in h_2 if she were to place herself in the imagined world in which h_1 is true. This differs from the 'called-off gamble' interpretation by requiring the DM to hypothesise h_1 as certain. It differs from 'temporal updating' because the DM is not seeking to predict how she will update her beliefs at some future time.

Though we follow a 'temporal updating' view of conditional probability, there are arguments claiming it is equivalent to the 'called-off gamble' interpretation. Goldstein agrees with the 'called-off gamble' interpretation, and discusses a notion of 'temporal coherence' in [49,50]. In [49], Goldstein claims that, if a DM wishes to act coherently and avoid a 'temporal' sure loss, then it is irrational for her to propose that she now believes h has probability $P(h)$, but that at a well-defined future time t, her beliefs will change to $P^t(h)$ with $E[P^t(h)] \neq P(h)$.

This argument can then be extended to conditional proabilities. If a DM considers that an unobserved event h_1 is relevant for establishing beliefs over event h_2, and as such, states conditional probability $P(h_2|h_1)$, then Goldstein [50] argues that, at a future well-defined time t, the DM may revise beliefs to $P^t(h_2|h_1)$, but that current beliefs should be such that $P(h_2|h_1) = E[P^t(h_2|h_1)|h_1]$. The connection with the 'temporal updating' view arises when time t is considered to be the point when the DM will know whether h_1 is true or false.

2.2 Solitary Decision Problem

Having briefly outlined the meaning of rationality as acting in accordance with an agreed system of axioms of preference together with specification of coherent beliefs, we now focus attention on suggestions that have been given for solving the solitary decision problem.

2.2.1 Subjective Expected Utility Theory

The most popular and famous solution to the decision problem under consideration is that provided by Subjective Expected Utility (SEU) theory. This solution dictates that, given a set of feasible decisions $\mathcal{D}$ and a utility function u that is in agreement with the DM's preferences over $\mathcal{D}$, the DM should select that decision $d' = \arg\max_{d \in \mathcal{D}} u(d)$. However, and as discussed in Subsection 1.1.1, it is usual to consider a utility function u as representing preferences over a reward space $\mathcal{R}$. In this case, and given a probability distribution $P(r|d)$ capturing the DM's beliefs over the relevant outcome for each feasible decision d, the admissible decisions are those that maximise expected utility. In the case of a finite reward space the admissible decisions are thus those that maximise $\sum_{r \in \mathcal{R}} u(r)P(r|d)$.

The maximisation of SEU was first proposed as a selection technique by Bernoulli in 1738 [18], however, not until 1947 was an axiomatic formulation created by von Neumann & Morgenstern [101], who provided such an axiomatization for when decisions are equivalent to lotteries with objective probabilities, each with finite support (*i.e.*, the set of possible rewards $\mathcal{R}$ is finite).

In what follows we employ a notational convention in which '$\succeq$', '$\succ$' and '$\sim$' are binary relations used to represent the DM's preferences between two decisions or rewards. In particular, $d_1 \succeq d_2$ represents the situation in which the DM's preferences are such that decision d_1 is deemed at least as preferable as decision d_2. Similarly, $d_1 \succ d_2$ represents the situation in which the DM strictly prefers decision d_1 to decision d_2, and $d_1 \sim d_2$ represents the situation in which the DM is indifferent between d_1 and d_2. The notation $\alpha d_1 + (1-\alpha)d_2$, with $\alpha \in [0, 1]$, will be used to represent that decision which pays reward $r \in \mathcal{R}$ with probability $\alpha P(r|d_1) + (1-\alpha)P(r|d_2)$.

With this notation in mind a list of axioms concerning the DM's preference relations that is similar to von Neumann & Morgenstern's, but which is in fact that given by Jensen [57], is as follows:

- A1 Completeness: $\succeq$ is a complete semi-ordering and the set of feasible decisions $\mathcal{D}$ is a closed convex set of lotteries.
- A2 Transitivity: $\succeq$ is a transitive relation.
- A3 Archimedian: If $d_1, d_2, d_3 \in \mathcal{D}$ are such that $d_1 \succ d_2 \succ d_3$, then there is an $\alpha, \beta \in (0,1)$ such that $\alpha d_1 + (1-\alpha)d_3 \succ d_2 \succ \beta d_1 + (1-\beta)d_3$.
- A4 Independence: For all $d_1, d_2, d_3 \in \mathcal{D}$ and any $\alpha \in (0,1]$, $d_1 \succeq d_2 \Leftrightarrow \alpha d_1 + (1-\alpha)d_3 \succeq \alpha d_2 + (1-\alpha)d_3$.

This axiomatization leads to the same utility representation theorem over the closed convex set of feasible decisions $\mathcal{D}$ that was first derived by von Neumann & Morgenstern (though von Neumann & Morgenstern's result also holds for all possible finite-support lotteries over the reward set $\mathcal{R}$), and indeed there are additional alternative axiomatizations that also perform the same task (see Fishburn [42] for a more general review).

The Completeness axiom simply states that a comparison using the preference relation $\succeq$ can be made between any two decisions in the set $\mathcal{D}$, *i.e.*, for any two decisions $d', d'' \in \mathcal{D}$, at least one of $d' \succeq d''$ or $d'' \succeq d'$ is true, whilst the Transitivity axiom states that if $d_1 \succeq d_2$ and $d_2 \succeq d_3$, then $d_1 \succeq d_3$ for all $d_1, d_2, d_3 \in \mathcal{D}$. The Archimedian axiom works as a continuity axiom for preferences, and as with the Completeness and Transitivity axioms, draws little objection.

The Independence axiom, however, draws criticism in certain circles. It effectively claims that preferences between two decisions are unaffected if they are both combined in the same way with a third decision. Nevertheless, one should remember that von Neumann & Morgenstern's axiomatization is developed only for decisions that are equivalent to objective lotteries, and in this setting, Independence is simply claiming that preference relations between two decisions should remain constant if there is a chance that neither decision (lottery) will be played, but rather that some other lottery will be played instead. Even so, it is still the axiom that is altered most frequently when non-SEU theories are suggested.

If a DM agrees to a similar system of axioms as that given above, then von Neumann & Morgenstern proved that there exists a unique (up to a positive linear transformation) utility function u, with domain the convex set $\mathcal{D}^+$ of finite support lotteries over $\mathcal{R}$ and co-domain $\mathbb{R}$, satisfying the following two properties:

1. For all $d_1, d_2 \in \mathcal{D}^+$, $u(d_1) \geq u(d_2) \Leftrightarrow d_1 \succeq d_2$.

2. For all $d_1, d_2 \in \mathcal{D}^+$ and any $\alpha \in (0,1)$, $u(\alpha d_1 + (1-\alpha) d_2) = \alpha u(d_1) + (1-\alpha) u(d_2)$.

The first of these properties states that the utility function is in agreement with the DM's preferences and, in particular, a lottery will have the largest utility value if and only if the DM ranks it as her preferred choice. The second property explains why utilities have a cardinal meaning, and do not simply rank lotteries, hence differing them from value functions. It also explains how utilities for non-degenerate lotteries can be formed from utilities for degenerate lotteries, and hence gives rise to the expected utility representation when one considers a utility function as a function that takes its domain to be the set $\mathcal{R}$ of possible rewards. Of course the above properties also hold true for the closed convex subset of feasible decisions (lotteries) $\mathcal{D}$.

Unfortunately, von Neumann & Morgenstern's theory is unable to deal with situations where the outcomes of decisions are not determined by objective probabilities, *e.g.*, horse races. This situation was later resolved by Savage [89], whose list of seven postulates (axioms) of rational choice permitted subjectivity in beliefs. Savage considered a different setup where, given a set of possible states of nature (possible event outcomes) S and a set of consequences (rewards) F, decisions were seen to be arbitrary functions from S to F.

Axiomatizations that permitted subjective beliefs, but were instead based on developing the objective theory of von Neumann & Morgenstern, were also later developed. Hence, rather than reviewing the relatively complicated theory of Savage, we will instead briefly review the somewhat simpler theory of Anscombe & Aumann [5].

Anscombe & Aumann extended the von Neumann & Morgenstern axioms, and thus also required the presence of lotteries with objective probabilities. For this reason Anscombe & Aumann's theory should be seen as an intermediate theory between the fully objective setting of von Neumann & Morgenstern and the fully subjective setting of Savage. Anscombe & Aumann achieved the introduction of subjective beliefs by viewing decisions as functions that mapped event outcomes to the simple lotteries considered by von Neumann & Morgenstern. The DM could then have subjective beliefs over what would be the actual event outcome (*e.g.*, the horse that wins the horse race).

Anscombe & Aumann use the representation of von Neumann & Morgenstern in two ways, matching up the two systems of preferences. The first way is to consider a utility function over roulette (objective) lotteries that pay rewards in the form of horse (subjective) lotteries which then pay out another roulette lottery. The second way is to consider standard von Neumann & Morgenstern roulette lotteries. Using the notation whereby $[R(1), \ldots, R(n)]$ represents the horse lottery paying roulette $R(i)$ if event i is true, and where $(p_1O_1, \ldots, p_mO_m)$ represents the roulette lottery paying the solitary outcome O_i with probability p_i, Anscombe & Aumann use the following two additional axioms to generate their required utility representation:

- A5 Monotonicity: If $R_1(i) \succeq R_2(i)$, then $[R(1), \ldots, R_1(i), \ldots, R(n)] \succeq [R(1), \ldots, R_2(i), \ldots, R(n)]$.
- A6 Reversal: $(p_1[R_1(1), \ldots, R_1(n)], \ldots, p_m[R_m(1), \ldots, R_m(n)]) \sim$
$[(p_1R_1(1), \ldots, p_mR_m(1)), \ldots, (p_1R_1(n), \ldots, p_mR_m(n))]$.

Monotonicity simply states that if two horse lotteries are identical except for the returns associated with one outcome, then preferences between these horse lotteries are dependent on preferences between the returns associated with that outcome. Reversal is an axiom stating that, if the return to be received depends on the outcome of both a horse lottery and a roulette lottery, then it makes no difference in which order these two types of lottery are played.

Anscombe & Aumann demonstrated that the logical implication of agreeing to all six Axioms A1-A6 is that, not only do subjective probabilities actually exist (though previous authors dating back to the work of Ramsey [86] have provided alternative arguments for this), but also there exists a unique (up to a positive linear transformation) utility function agreeing with the DM's preferences for the situation where probabilities of outcomes are subjective. Thus no longer does one require the assumption that probabilities are objective and imposed externally.

2.2.2 Alternatives to SEU

The use of maximising SEU as the normative theory in decision selection is not without criticism, as various authors criticise one or more of the axioms it is based upon. The first, and possibly most famous criticism, is that given by Allais [3]. Allais claims that perfectly rational people do make decision selections that are not in keeping with those dictated by the maximisation of SEU. Further studies by authors such as Kahneman & Tversky [60], Ellsberg [38] and Fellner [41] also expand upon Allais' objection.

An illustration of Allais' criticism, often referred to as the Allais paradox, can be given by considering the following pair of choices:

Choice 1: l_1 pays £4000 with probability 0.8, £0 otherwise.
l_2 pays £3000 for certain.

Choice 2: l_3 pays £4000 with probability 0.2, £0 otherwise.
l_4 pays £3000 with probability 0.25, £0 otherwise.

An investigation by Kahneman & Tversky [60] shows that DMs commonly hold a preference for l_2 over l_1, whilst simultaneously preferring l_3 over l_4. However, there is no possible utility function that can accommodate this.

Such a combination of preferences violates the Independence axiom of expected utility theory. Indeed, in this example the only difference between lotteries l_1 and l_3, or between l_2 and l_4, is a common increased chance of receiving £0. This is a descriptive shortcoming of what is deemed a normative theory. Nevertheless, Allais argues that the Independence axiom should not be seen as a normative axiom of rational choice. He claims it is not enough to consider the expected utility return of a decision, but that higher moments taking into account variation or dispersion should also be considered.

Allais claims that the utility of a lottery should be some functional of the probability density, and that the DM should have a preference for security in the neighbourhood of certainty. He proposed a system which concentrates on the dispersion of rewards around their mean, replacing the Independence axiom with Iso-Variation, an axiom that requires the DM to select decisions not only on the basis of maximising expected value, but also taking into account second and higher order moments of the distribution of possible rewards.

An alternative theory, also motivated by real-world observations and by similar contradictions to SEU as demonstrated by the Allais paradox, is that of Kahneman & Tversky's Prospect Theory [60]. Prospect Theory, like the theory of von Neumann & Morgenstern, is concerned with the selection of objective lotteries, but rather than using such objective probabilities to weight the utility of rewards, it uses some non-linear function of them.

Kahneman & Tversky argue that instead of maximising SEU, DMs are subject to a Certainty effect (where DMs overweight outcomes that are highly probable and underweight outcomes that are very unlikely) and an Isolation effect (were DMs ignore common elements of decisions), both of which are incompatible with the Independence axiom. Further developments can be found in Tversky & Kahneman [99] and Wakker & Tversky [102].

In its original form, Prospect Theory made a distinction between two phases of a DM's choice process. First a DM performs a preliminary analysis of the offered choices with the aim of yielding a simpler representation of the problem, a so-called 'editing phase'. Later the DM evaluates the edited choices and the one with the maximum valuation is selected. The editing phase will code (turn outcomes into gains or losses, rather than final states of wealth), cancel (ignore components shared in choices), simplify (values are rounded up or down), and finally remove dominated alternatives (even if they were not dominated before simplification).

Once the editing phase is complete, Prospect Theory evaluates a score for each decision that is determined through a weighted average of the utility of possible outcomes. However, instead of weighting by the probability of those outcomes, Prospect Theory uses a non-linear scale that reflects the psychological impact of the probability which, for example, will overweight high probabilities and underweight low ones (see Kilka & Webber [65] for an elicitation suggestion). As is to be expected, Prospect Theory's departures from SEU can lead to some objectionable consequences, and as Kahneman & Tversky [60] note, intransitivities and violations of dominance mean it should primarily be seen as a descriptive theory.

Prospect Theory is known as a rank-dependent model of choice under uncertainty. The defining property of such a model is that cumulative probabilities are transformed by a non-linear weight function in order to account for real-world inconsistencies to SEU theory. Further extensions for when decisions are not lotteries with specified objective probabilities are suggested by Schmeidler [92] (Choquet SEU Theory) and Wakker & Tversky [102] (Cumulative Prospect Theory).

Proponents of SEU theory, however, offer their own arguments as to why non-SEU theories should not be seen as normative, and how SEU can accommodate so-called paradoxes of the theory. De Finetti [32] and Amihud [4] argue against the claim that the dispersion of utility values should be considered, with de Finetti stating that utilities themselves were introduced to accommodate riskiness in extreme values.

Luce & von Winterfeldt [73] argue that DMs may be attempting to behave in accordance with SEU theory, even if they are likely to fail in more complex situations. Allais' objection is that SEU theory does not correspond to observed results, yet this may be due to a bounded rationality, as suggested by Levi [69].

Amihud [4] also notes that the Allais Paradox can be resolved by use of utility functions that are contingent upon the decision making history of the DM. Such history dependent utilities exempt the DM from consistency of preferences between periods, instead only requiring consistency within each period itself. Further solutions in agreement with SEU theory are provided by Morrison [79] and Markowitz [74, pp.220-223]. Indeed, Luce & von Winterfeldt [73] show that if the participants of Kahneman & Tversky's survey did not treat both choices simultaneously and independently, but instead conditioned utilities on the first choice before making the second, than a utility function agreeing with observed results can be found.

A final alternative theory that we will mention as an alternative to SEU is the Info-Gap theory of Ben-Haim [14]. Info-Gap decision theory is a non-probabilistic theory that suggests the DM seek to be robust against failure. Unlike SEU, it permits the DM to tackle decision problems without requiring a full probabilistic description of events. Instead a best estimate is provided and uncertainty is incorporated by accepting this best estimate could be incorrect by various degrees. A minimum required reward level is specified, and the decision is selected that maximises the chance of achieving this level, *i.e.*, the most robust decision is suggested.

2.2.3 Generalisations of SEU

The use of the maximisation of SEU as a decision selection technique requires that the DM can specify precise and correct beliefs and preferences. However, as mentioned in Subsection 2.1.2, this can be quite a difficult task. For this reason recent research has been focused on finding decision theories that remain in the spirit of maximising SEU, but which also permit the DM to be vague in elicitation.

Kadane *et al.* [58] provide an overview of how differing axiomatic formulations manage to cater for the situation in which only imprecise probability specifications are provided. Generally, such axiomatizations arise through weakening the Completeness axiom of von Neumann & Morgenstern. This axiom is sometimes deemed to be too restrictive and enforces the DM to state and commit to preference rankings between any two decisions, not permitting indecision or non-comparability between options. Instead, when wishing to deal with imprecise probabilities, the Completeness axiom is often weakened by replacing it with one that only calls for a strict partial ordering. Yet, if one makes such a replacement to the Completeness axiom, then no longer is it required that the DM rank all decisions, and so no longer is she necessarily able to determine which decision should be selected. There are, however, several suggested rules for selecting decisions when a complete ranking is not provided, and we now briefly review these.

The Γ-Maximin choice rule permits imprecise probabilities and its motivation for selection is similar in manner to the Maximin choice rule that was pioneered by Wald [103]. Under this rule, given a convex set $\mathcal{P}$ of probability distributions that satisfy the constraints of the DM's imprecise probability specifications, each feasible decision is ranked by considering the smallest SEU value that is possible when we are free to choose any element of $\mathcal{P}$. The decision that has the largest minimum value is then selected, and in the case of ties, rankings are considered by repeating the process, but where for each decision the 'worst' distribution is eliminated from $\mathcal{P}$ before again finding the smallest possible SEU, *etc.*

Obviously in the case of $\mathcal{P}$ containing just a single distribution, the Γ-Maximin choice rule returns to classical maximisation of SEU. However, when $\mathcal{P}$ contains more than one distribution, Γ-Maximin seeks to protect against worst possible outcomes, and as such, is considered a robust method of decision selection (similar to the Info-Gap theory discussed in Subsection 2.2.2).

An axiomatization of the Γ-Maximin choice rule is provided by Gilboa & Schmeidler [47]. Gilboa & Schmeidler use Axioms A1-A3, A5, and A6, however, the Independence axiom is kept only for decisions with certain consequences, and when decisions have uncertain outcomes, it is replaced by an axiom of Uncertainty Aversion:

- Uncertainty Aversion: For all $d_1, d_2 \in \mathcal{D}$ and $\alpha \in (0, 1)$,
  $d_1 \sim d_2 \Rightarrow \alpha d_1 + (1 - \alpha)d_2 \succeq d_1$.

Gilboa & Schmeidler claim that an intuitive objection to the Independence axiom is that it ignores the phenomenon of hedging (a preference for spreading bets), and Uncertainty Aversion specifically states that hedging is never less preferred to not hedging.

An alternative choice rule for when probabilities are imprecise is Maximality, which dates back to at least the work of Condorcet [30], and which has been further discussed, for example, in the works of Sen [94] and Walley [104]. Under this choice rule a decision is admissible if and only if there exists no other feasible decision that has a higher SEU value for every possible distribution in the set $\mathcal{P}$. Hence, unlike Γ-Maximin, Maximality does not guarantee a complete ranking of decisions, and often a DM will find that the set of admissible decisions is not much reduced from the set of feasible decisions, especially if beliefs are quite imprecise and vague.

Again Maximality will reduce to the classical maximisation of SEU if there is only one distribution in the set $\mathcal{P}$. When $\mathcal{P}$ contains more than one distribution, however, Maximality only seeks to reduce the set of feasible decisions to a set of admissible ones by removing those decisions where it is known that, regardless of which distribution in $\mathcal{P}$ is considered, there exists a decision that will always have greater SEU. An axiomatization of Maximality is offered by Seidenfeld *et al.* [93] who, unlike in the axiomatization of Γ-Maximin, retain the Independence axiom. Instead a slight alteration is made to the Archimedian axiom and the Completeness axiom is changed to a strict partial ordering axiom. Another, earlier, axiomatization is also provided by Walley [104].

The last choice rule we review for when probabilities are imprecise is Expectation Admissibility, or E-Admissibility. This rule was suggested by Levi [68] and, like Maximality, does not seek to provide an ordered ranking of the feasible decisions. Levi's suggestion is that only those decisions that maximise SEU for some distribution in $\mathcal{P}$ should be considered admissible, and nothing else can be stated to distinguish between admissible decisions. Again E-Admissibility reduces the set of feasible decisions to a set of admissible decisions, yet under E-Admissibility, the set of admissible decisions is a subset of the admissible decisions under the Maximality choice rule.

Like both the Γ-Maximin and Maximality choice rules, E-Admissibility reduces to the maximisation of SEU when there is only one distribution in $\mathcal{P}$. Its axiomatization again transforms the Completeness axiom to one of a strict partial ordering, and although it satisfies the property that if two decisions are both inadmissible then so is any convex combination of them, it does not fully satisfy the classical Independence axiom.

A final comment on the choice rules mentioned for imprecise probabilities is that, for all three of Γ-Maximin, Maximality, and E-Admissibility, Schervish *et al.* [91] show that, if a 'favourable' decision is defined as one that is uniquely admissible when considered in a pairwise comparison against the option of making no decision selection, then no finite combination of favourable decisions can result in a sure loss.

Further to the above generalisations which seek to incorporate imprecise probabilities in the choice rule of maximising SEU, there are also generalisations seeking to incorporate imprecise utilities, and on a foundational level this setting is considered by Seidenefeld *et al.* [93], who extract imprecise probability and utility statements from preference relations that only satisfy the properties of a partial order. Moskowitz *et al.* [80] also permit both imprecise probabilities and utilities, allowing imprecision over the certainty equivalence for a simple lottery (the sure amount which the DM holds in equal preference to the uncertain lottery) in order to introduce imprecise utility information. Imprecise probabilities are included as bounds over the probability that an event is indeed true.

Moskowitz *et al.* assume a parametric exponential form for the utility function, with information about probabilities of events and preferences between lotteries being used to create both a set of possible distributions, and a set of possible utility functions. Progressive questioning over relationships between probabilities and strict preferences between rewards is then used to reduce the number of possibilities for a precise distribution and a specific utility function. This questioning continues until, regardless of the possibilities that remain, there is a unique decision that will maximise SEU.

In their multi-attribute utility setting (where attribute values of rewards are combined to form an overall utility value for the outcome), Farrow & Goldstein [40] also permit imprecise preferences, and this is achieved by permitting imprecision in the trade-off values between attributes. A trade-off value is used to describe how a relative increase in one attribute in the reward is used to lead to an increase or decrease in the overall utility of the reward.

Taking a specific attribute, the DM is permitted to offer a strict, a weak, or an indifference preference relation over various possible rewards. Each such preference places constraints on the allowable choices for the trade-off value for that attribute, and a collection of possible trade-off values consistent with stated preferences may be considered. Hence, Farrow & Goldstein's use of imprecise trade-off parameters greatly eases what would be a very complicated problem of eliciting multi-attribute utilities.

2.3 Sequential Decision Solution

Having briefly examined a few of the theories seeking to provide rational methods for solving a solitary decision problem, we now consider some of the theories developed for sequential problems. In particular, we examine the various considerations for the form of the utility function in these theories.

Usually, sequential decision problems with a finite planning horizon are also solved through maximisation of SEU, with dynamic programming used to determine the optimal decision sequence (see, for example, DeGroot [34] or Berger [15]). This technique considers all the possible situations that the DM could find herself in by the time she selects her final decision, constructing a decision strategy for each possible situation. With knowledge of what the DM will do in the final period, next an optimal policy is determined for decision selection at the penultimate choice, considering again all the possible situations the DM could be in at that time. This procedure is continued backwards through decision choices until eventually an optimal first decision is determined. Again, if the DM is able to learn about probabilities for events of interest, then conditional updating is performed similar to that outlined in Subsection 2.1.3. Nevertheless, there are alternatives for applying this procedure, usually due to the form of utility function considered. We briefly review a few of these.

2.3.1 Discounting

Discounting the utility of rewards that are to be received in the future is often applied to model preference for early receivership. The utility a DM gains from knowledge that a reward will be received at some future time need not be the same as the utility for receiving it now, if for no other reason than that the DM will have access to the reward for a greater duration. However, relative to determining the utility for receiving a reward immediately, it is often difficult to elicit the current utility a DM attributes to knowledge that the same reward will be received in the future. However, if agreed with, discounting functions can provide a link between utility values for receiving the reward at various future times.

The discount model essentially multiplies the utility value for receiving a reward now by a function of the duration of time before it is to be received. The most common such discounting function is the Exponential Discounting Function (EDF), which as a function of time elapse t, is of the form λ^t, with $\lambda \in [0, 1]$ a parameter of the model (see, for example, Ahlbrecht & Weber [2]). A common alternative to the EDF is the Hyperbolic Discounting Function (HDF) of the form $1/(1+t)^\gamma$, with $\gamma > 0$ a parameter (see, for example, Harvey [53]). However, whilst the EDF is often used in normative models for discounting the utility of rewards to be received in the future, models that employ the HDF are primarily to be seen as descriptive. In particular, and as will be discussed below, the HDF does not satisfy the property of dynamic consistency.

The advantage of this approach is that, to find the desired solution to a decision problem in which rewards are felt at future time points, the DM simply has to discount each utility value by the appropriate amount and apply the result in a standard decision problem. When a stream of rewards is to be received, the discounting model discounts the utility of each reward and aggregates to form a single meaningful number. This is known as the Net Present Value, with the interpretation being that the DM is indifferent to receiving that amount of utility now and receiving the stream of future utility values (see Meyer [76, p.479]).

Arguments for discounting utilities of future rewards appear to have been first presented by von Böhm-Bawerk [100] and Fisher [44], both of whom use economic and psychological motivation. It is certainly mathematically convenient, and for when an infinite planning horizon is considered, offers a method for comparing reward streams. Nevertheless, there is little normative reason for discounting or agreement of an objective discount rate, though there do exist arguments detailing why certain discounting functions have more appealing properties than others.

Strotz [98] argues that only the EDF is a justifiable discounting function (see also Weller [105]). He claims that any such function should not change the utility of immediate rewards, that it be non-negative and decreasing in time delay, and that it be dynamically consistent. Dynamic consistency requires that preferences between future rewards should not be changed if receivership is to be hastened or delayed, *i.e.*, if one reward is deemed more desirable than another if they are to be received at time t_1, then the preference relation should remain unchanged if both rewards are to be received at time $t_2 \neq t_1$. Only the EDF satisfies all of these properties.

There are indeed axiomatizations for the use of maximising NPV and discounting through the EDF, and Meyer [76], Koopmans [67] and Fishburn & Rubinstein [43] have all offered similar suggestions. The most controversial axiom in Meyer's axiomatization is that of Successive Pairwise Preferential Independence (SPPI):

- SPPI: Trade-offs between utility (consumption) amounts in periods i and $i+1$ are not dependent on utility (consumption) amounts in alternative periods.

Treating a reward stream as a single multi-attributed reward implies that the DM's utility function could, in theory, be any arbitrary function of the entire collection of trade-off parameters and multi-period utilities. However, certain independence assumptions, if true, can reduce the complexity of this function to varying degrees of simplicity. SPPI is such an assumption that keeps the overall utility form tractable.

SPPI is similar to the Stationarity axiom of Koopmans [67], which claims that if two streams have identical first period reward, then preferences over the modified streams that are obtained by deleting the first period and advancing the timing of all subsequent rewards by one period, must be ordered in the same manner as the original unmodified streams were. Using this axiom, Koopmans establishes the existence of an additive utility function for determining the worth of reward streams.

Many philosophers, however, believe discounting to be irrational. Both Rawls [87] and Ramsey [85] criticise the action, with Ramsey claiming time discounting to be "a practice which is ethically indefensible and arises merely from the weakness of the imagination". Rawls [87, p.293] states that "the avoidance of pure time preference is a feature of being rational ... the mere difference of location in time ... is not in itself a rational ground for having more or less regard for it".

2.3.2 History Dependent Utility

One of the complaints of axioms like SPPI or Stationarity is that they do not allow previous reward realisation to affect preferences over future rewards. History dependent utilities instead explicitly permit this, though at the cost of a more complicated utility function and the requirement to elicit more trade-off parameters. Discounting is then only included to incorporate effects such as inflation or mortality rates (see Yaari [109]) and is no longer expected to agree to the principles of the EDF, *e.g.*, discount rates are no longer expected to be constant.

A simple extension suggested by Meyer [77] is to state that future preferences are independent of past rewards, but that ordering of future rewards is relevant. This assumption does not generally imply the existence of unconditional single period utilities (as SPPI or Stationarity allows), but is the most general assumption permitting a solution through dynamic programming. However, the most general situation is contained within the work of Bell [11], who permits all forms of dependencies and independencies in preferences (though at a cost of requiring a great amount of trade-off parameters, whose interpretations are difficult to understand).

One alternative to reduce complexity is to introduce state descriptors that record past reward realisations (Meyer [76, 77]). Preferences over future rewards are then conditioned upon these descriptors, permitting future preferences to depend on the decision history. To keep computation tractable, and for ease of elicitation, it may be that only influential summary statistics of the past, rather than a complete record, are used to condition future utilities on.

An axiomatization for the use of state descriptors was provided by Bodily & White [19], who considered an economic sequential decision problem. Bodily & White's DM must at each period i, for $i = 0, 1, \ldots, n-1$, select a consumption level r^i such that, if w^i is the level of the DM's wealth at time i, $r^i \leq w^i$. The DM's problem is then to decide upon investment and consumption levels to optimise the consumption stream $r^0, r^1, \ldots, r^n, w^{n+1}$, with w^{n+1} being terminal wealth.

As decisions are made, the DM's decisions will be contingent on the outcomes of previous choices. The DM is assumed to base consumption and investment decisions on beliefs concerning future returns on investment and preferences for alternative consumption streams. Bodily & White permit attitudes towards future consumption to depend upon current wealth and past consumption, and hence a summary descriptor is included for this purpose.

2.3.3 Evolving Utility

That utilities may evolve is an additional consideration that is not only used to permit a change in preferences as a result of consumption level experienced (such as history dependent utilities), but also to allow preferences to change following nothing but a passage of time and a change in tastes.

Witsenhausen [108] suggests a theory of Assumed Permanence in which it is accepted that future preferences may not be the same as those currently held. The DM still makes current decision selection under the assumption that future preferences will remain constant to what they currently are, but when coming to a future decision she may re-evaluate preferences and seek to select decisions that maximise utility return with respect to previous choices, and where again it is assumed that future preferences will remain the same as the now re-evaluated levels. Witsenhausen's theory has the great practical advantage of not requiring a model of how preferences will evolve. Its obvious disadvantage, however, is that early commitments may be made which are costly to reverse if preferences are found to have changed.

White [106] also considers a sequential decision problem in which the DM's future preferences are uncertain, but as opposed to the Assumed Permanence of Witsenhausen, attempts to model how preferences may change. White achieves this by assuming that the DM's preferences are modelled by a vector of trade-off weights which may change as the DM progresses through decision selection.

White considers a finite stage decision problem where preferences may change from stage to stage. Uncertainty over the result of decision selection is not considered, and hence the DM is assumed to know the result of any choice she makes. Thus her problem is, given knowledge of how preferences may evolve, which decision sequence should be selected. Three evolution mechanisms are considered, consisting of an optimistic scheme where preferences evolve in order to maximise utility, a pessimistic scheme where preferences evolve to minimise utility, and a scheme where preferences evolve randomly.

The final decision theory we mention that is based on evolving utilities is that discussed by Meyer [76]. Meyer extends his work with history dependent utilities by permitting preferences to be influenced by a time stream of extraneous events that are characterised by a sequence of parameter values. It is assumed that such a parameter will influence the DM's utility function and that it is independent of previous rewards realised (the history dependence already allows for this) with the parameter value evolving randomly according to some specified probability function.

Chapter 3

Adaptive Utility

This chapter provides motivation for the use of adaptive utility theory. A review of works that have either developed or made use of adaptive utilities is also included.

3.1 Motivation

The expected utility theory that was discussed in Chapter 2 proves that, provided the DM agrees to a certain collection of axioms, there will be a unique (up to a positive linear transformation) utility function that is in agreement with the DM's preferences. However, although we now know that this is the case, there is still the problem of determining what this function actually is. To determine the utility value of a particular reward, one can use the system of comparing the gamble which pays that particular reward with certainty, to a gamble which either pays the best reward or the worst reward. In practice though, it is common to simply assume that a utility function has a general form with particular properties, *e.g.*, the logarithmic function that was suggested by Bernoulli [18] for monetary rewards.

Nevertheless, this practice still assumes that a correct utility function representing the DM's true preferences for all possible outcomes can be identified. Furthermore, implicit within this is the assumption that the DM actually knows her true preferences. Yet in the real world this is not always the case, and it is perfectly natural for a DM to be unsure of her preferences.

A DM could, for example, be considering a reward that would not be received until some future time point, and then they need to consider how their preferences may have changed by that time point due to them being older and possibly also in a different situation. Another common example occurs when a DM is asked to consider rewards that are vague or unfamiliar to her. As Simon [97] states, "the consequences that the organism experiences may change the pay-off function ... it does not know how well it likes cheese until it has eaten cheese".

It is traditionally assumed, however, that the DM's utility function is fully known, and that it is even possible to make hypothetical choices. After a decision is made it is assumed that the utility realised is the same as was indicated by the utility function, and that there can be no surprises. Hence classical theory cannot account for uncertain preferences, with preferences over sure things being fixed. Nevertheless, in the real world a DM may learn about her likes and dislikes of new and novel rewards, a situation that classical theory cannot account for as it has no element of utility learning following new information.

There are plenty of examples in the real-world demonstrating that a DM will not always be sure about her preferences, but rather that she may be uncertain of these and that she is able to learn about them. A DM seeking to purchase a new car and who test drives a possible choice is one such situation, for if the DM knew her preference for the car, as would be assumed in classical utility theory, what would be the reason for test driving it?

Another example is that often companies offer a trial introductory price on a new product, or they may even offer free samples, but what would be the motivation for this if all potential customers knew precisely how much they liked or disliked the new product? Indeed, how many individuals have ever been disappointed with a result that was expected, or pleasantly surprised, for example, by how nice a new recipe is?

In order to illustrate this point, and in order to introduce a basic example that will be returned to throughout this thesis when demonstrating new aspects of adaptive utility theory, consider the following problem, which will be referred to as the Apple or Banana example.

Example 3.1.1

A DM faces the problem of deciding upon which fruit to purchase at lunch. The shop has on offer two choices, either an apple or a banana. The DM is experienced with eating apples, having done so many times before, but she has never previously consummed a banana, and as such, is unsure which fruit she would prefer. Nevertheless, she is able to look at the banana, to smell the banana, and to even ask the suggestion of friends. What she is not able to do, however, is taste the banana herself before making the decision to purchase it. How then should such a DM make her choice? ◇

3.2 Adaptive Utility

Instead of assuming that the DM's actual preferences and corresponding utility function is precisely known, the theory of adaptive utility allows the DM to be uncertain over her true preferences and permits her to learn about them. In this sense the theory of adaptive utility is a normative theory for rational decision making when one accepts that there is uncertainty over the DM's true utility function. It is assumed that the DM is able to envision possible preferences and to form expected preferences from these prior beliefs. Such expected preferences can then later change as a result of the DM's experiences, with some comparison being made between what she *a priori* expected and what she actually realised. In this manner the DM will be able to learn about her true utility function, update her beliefs, and adapt her decision making accordingly.

However, for a single decision this strategy is no different to classical theory, as decision selection would be the same as if it were assumed actual preferences equal prior expected preferences. Only in a sequential problem can a difference be observed, where the DM may learn about preferences and expected preferences may change. Indeed, a DM may make a different choice when faced with an identical but repeated problem, an act which is in contradiction to classical theory, but which is common in the real-world. Furthermore, in a sequential setting the optimal initial decision need not be that which appears to give greatest utility under current beliefs. This is because information gained following selection of the first decision may depend on that decision, and so it may be optimal to select one which is more informative of true preferences, enabling the DM to learn and select better decisions in the future. This first decision may then be different to what would be selected in a one-off problem, even if outcomes of decisions are known.

Uncertainty in the utility function is incorporated by conditioning the utility on a parameter θ, which we will refer to as the DM's *state of mind.* A particular state of mind represents a particular preference ranking and is included within notation in a similar way as conditioning is in probability, *i.e.*, the utility from a reward r when θ is the true state of mind will be represented as $u(r|\theta)$.

Definition 3.2.1

Given a set of possible classical utility functions $u(\cdot|\theta_1), \ldots, u(\cdot|\theta_n)$ that have been suitable scaled to ensure they are commensurable, and given a probability distribution P_θ representing the DM's beliefs over the correct value for the state of mind θ, an *adaptive utility function* ${}_au(\cdot)$ is defined[1] to have domain the the convex set $\mathcal{D}$ of decisions and co-domain the set of real numbers $\mathbb{R}$ and is such that it equals the expectation of $u(\cdot|\theta)$ with respect to beliefs over θ, *i.e.*, ${}_au(\cdot) = E_\theta[u(\cdot|\theta)]$. ◇

[1]This definition of an adaptive utility function is slightly different to that given by Cyert & DeGroot [27–29]. As will be mentioned in the following section, Cyert & DeGroot introduce the term adaptive utility, but consider it as a function of both r and θ. However, for the purpose of this thesis, we find it beneficial to use the slightly different definition of ${}_au(\cdot) = E_\theta[u(\cdot|\theta)]$.

Definition 3.2.2

Possible classical utility functions $u(\cdot|\theta_1), \ldots, u(\cdot|\theta_n)$ are said to be commensurable if they have been scaled in such a way that it is meaningful to compare the utility values of different rewards when conditioned under differing possibilities for θ. That is to say, they are considered commensurable if given any three reward and state of mind pairs (r_1, θ_1), (r_2, θ_2), and (r_3, θ_3), with the property such that $u(r_1|\theta_1) \neq 0$ and $u(r_1|\theta_1) \geq u(r_2|\theta_2) \geq u(r_3|\theta_3)$, then the DM would be indifferent between receiving r_2 when $\theta = \theta_2$ and playing the hypothetical gamble where r_1 is paid and $\theta = \theta_1$ with probability $u(r_2|\theta_2)/u(r_1|\theta_1)$, and where otherwise r_3 is paid and $\theta = \theta_3$. ◇

Notice that, although it has not been explicitly included in the notation, an adaptive utility function ${}_au(\cdot)$ is not only a function of the reward or decision under consideration, but also of the DM's beliefs, P_θ, over the state of mind θ. Furthermore, changes in these beliefs over θ cause the adaptive utility function to change, or rather, the adaptive utility function adapts to incorporate changes in beliefs.

Such changes in beliefs over θ may occur in many ways, *e.g.*, information from associates, advertising, or even actual reward realisation. It depends on the problem under consideration, but once a source of information and a likelihood function are identified, Bayesian updating will lead to posterior beliefs over preferences. We will return to the possibility of learning of θ in Chapter 5.

3.3 Review of Adaptive Utility

Adaptive utility was introduced by Cyert & DeGroot [27–29], who argued that paradoxes such as Allais' could be constructed because classical theory does not incorporate learning of preferences. Cyert & DeGroot also claimed that adaptive utility is consistent with casual empiricism, as fads in style and products can be observed. Their argument is that new products do not always have a genuine advantage over past options, but products may no longer be purchased as consumers change beliefs over preferences. Hence, although not directly seen as a descriptive theory, Cyert & DeGroot claim that adaptive utility is a concept that better models actual behaviour.

Although we agree with the normative concept of adaptive utility, we do not agree with these additional motivations. The Allais paradox is more connected with DMs not understanding the implications of their decisions on the axioms of rational choice, and adaptive utility can only offer a possible explanation if the DM can learn of her preferences by simply being offered a choice between two potential lotteries (and learning nothing else). This is similar to another example by Cyert & DeGroot in [27]. In this example a DM is in the process of selling a house and it is argued that it is enough for the DM to be offered her requested price for her to change her preference for selling at that price. Also, although we agree that trial of various products can lead to a change in preferences over those products, it is more likely that fads and fashions arise due to a change in actual utilities, not a change in beliefs over them, and as such, a theory of evolving utility would be more suitable.

Cyert & DeGroot primarily considered examples of the use of adaptive utility in implications for economic problems such as consumer demand and income allocation. They suggested several settings for uncertainty over a DM's utility form, *e.g.*, uncertainty over particular weights in the utility function when rewards are multi-attributed, with beliefs over correct values possibly changing as a result of directly experiencing particular rewards.

Erdem & Keane [39] considered uncertain preferences in an analysis of data[2] on the sale of liquid detergents in the US over a 3 year period. In their study it was assumed that consumers were uncertain about particular brand attributes (especially for new brands), and a product with uncertain brand attributes was further assumed to have uncertain utility value at the time of purchase. However, it is assumed that

[2]Daily panel data of laundry detergent purchases for 3,000 households from two test markets.

DMs can learn about the utility of products through usage and through external advertisements. As many purchases may be required for certain attributes to become apparent (*e.g.*, how effectively does the detergent prevent colour fading), or because certain attributes may only become known on specific occasions (*e.g.*, how effectively does the detergent clean a particular stain), a learning model is incorporated where noise corrupted observations of the true utility are observed.

Erdem & Keane suggest that, rather than being myopic and selecting detergents that appear optimal under current utility beliefs, consumers may recognise that current choices affect their information set, providing them with incentive to try new and unfamiliar brands. This model of decision making, where consumers consider the impact of their choice upon the expected present value of utility over the entire planning horizon (the same suggestion as Cyert & DeGroot), is compared to the myopic model of maximising immediate utility return. Thus Erdem & Keane seek to examine the descriptive validity for this decision selection technique, finding it to be slightly superior than the model which seeks immediate utility maximization.

Chajewska *et al.* [22] consider a sequential decision problem which they refer to as adaptive utility elicitation, the focus of which is solution to the problem an analyst faces when seeking to elicit a DM's utility function. As the authors note, the complexity of utility elicitation means that a decision must often be made when only partial utility information is available. Chajewska *et al.* consider utility as a random variable drawn from a specific prior distribution, and determine optimal strategies for sequentially asking the DM questions about her preferences. This process is continued until a unique optimal decision is identified.

Finally, Boutilier [21] considers foundational issues of adaptive utility, which he refers to as expected expected utility. As Boutilier notes, the decision that maximises adaptive utility is sensitive to the scaling of the possible utility functions, which are only unique up to a positive linear transformation. The term commensurable is used to represent those possible utility functions with which it is meaningful to make comparisons. Boutilier shows that under an assumption of extremum equivalence (where under each possible utility function there exists the same best and worst reward), utility functions can be scaled in a specific manner to make them commensurable. Further discussion on foundations of adaptive utility is included in Chapter 4, and in Section 4.3 we will return to discussing commensurable utilities, demonstrating that possible utility functions can be scaled to ensure commensurability without Boutilier's strong assumption of extremum equivalence.

As Cyert & DeGroot [28] note, "we have begun to investigate the concept in one area of economics, and we are aware that much remains to be done". The work of Cyert & DeGroot does not discuss foundational implications for uncertain utilities, nor what it might mean to compare different possible utility functions that are only unique up to a positive linear transformation. As such, the contribution of this thesis towards the subject of adaptive utility commences in the next chapter where we consider the precise interpretation of the utility parameter θ and consider how, through repeated use of the classical axioms of rational choice, we can accommodate the maximisation of adaptive utility as a decision selection technique. Furthermore, in Chapter 5 we also consider how to determine optimal decision strategies for a problem consisting of n sequential decisions (Cyert & DeGroot only provided a solution for up to 2 decisions), discuss utility forms that can simplify the computation of solution algorithms, and present possible applications for reliability problems. Finally, in Chapter 6 we consider implications for the meaning of risk aversion and of value of information. These are two diagnostics of the decision problem that have been ignored in previous literature on adaptive utility related works.

Chapter 4

Foundations

In this chapter we offer discussion on the interpretation of a state of mind. We also propose a method of employing the classical system of expected utility axioms to provide an argument for using the maximisation of adaptive utility as the logical decision selection rule. We conclude the chapter by discussing a method for creating commensurable utility functions that does not require Boutilier's [21] assumption of extremum equivalence (see Section 3.3).

4.1 State of Mind

As mentioned in Section 3.3, Cyert & DeGroot [27] introduce uncertainty in a DM's utility function by parameterising it with an unknown variable θ. However, their focus is primarily on introducing the uncertain utility concept and in discussing implications for economics. As such they do not address foundational issues, and examination of the precise ontological nature of θ is avoided. For this reason we now offer our own interpretation of this utility parameter which we have named the DM's state of mind.

The true value of the state of mind θ will be used to represent the DM's true preferences. Classical theory, under the Completeness Axiom A1, states that there is a true preference ranking over the convex set $\mathcal{D}$ of decisions, the state of mind is simply used to characterise this. This is in contrast to ω, the *state of nature*, which determines the actual reward that any given decision will lead to (*e.g.*, the horse that wins the race, when the decision is which horse to back).

In discussing the interpretation of a state of mind we return to the setting of Anscombe & Aumann [5] as described in Subsection 2.2.1. In this setting decisions can either be objective (roulette) lotteries or subjective (horse) lotteries, with possible rewards included in some finite set $\mathcal{R}$. A utility function is then a function, unique up to a positive linear transformation, that is in agreement with the DM's preference ranking $\succeq$ over the closed convex set of decisions $\mathcal{D}$. Hence, provided that changing θ in the parameterised utility function $u(d|\theta)$ does not entail a positive linear transformation, each value of θ will represent a different preference ranking $\succeq_\theta$. That is to say, a state of mind fundamentally represents a possible preference ranking over the set of decisions. Note also that, at the risk of entering a rather abstract philosophical debate, we assume that the DM is not free to affect the correct value of θ, *i.e.*, she can not select her true preferences. For example, we would say that the DM can not choose which of apples or bananas she truly prefers (or whether they are equally preferable). Instead we claim that the correct value of θ is pre-determined independently of the views or wishes of the DM (hence the reference to an hypothetical gamble in Definition 3.2.2).

Our assumption then is that uncertainty in the utility function is represented by uncertainty over an unknown parameter θ representing certain characteristics of the DM's preferences. The DM is assumed to hold prior beliefs about θ and these are used to create expected preferences. Once the DM receives information concerning θ she can update beliefs through Bayes' Theorem (simultaneously updating beliefs over preferences). True preferences are not assumed to change (unlike the evolving utility theories of Subsection 2.3.3), only the DM's beliefs over what these actual true preferences are can alter, and this is done through a change in beliefs over θ.

As discussed in Section 3.1, a DM may not necessarily know her true preferences and so may be uncertain of θ. By using de Finetti's classification of the term [31, p.11],

θ is a well-defined random quantity[1], and as such, a probability distribution P_θ may be specified. Furthermore, as preferences are inherently subjective, it seems natural to assume that P_θ is also subjective. In essence then, the use of adaptive utility is analogous to the use of a hierarchical prior in Bayesian analysis (see Berger [15]). A utility, when scaled to fall in the interval $[0, 1]$, corresponds to a probability value, with the utility of any reward r being that probability p which makes the DM indifferent between playing the gamble paying reward r for sure, and playing the gamble that pays best reward r^* with probability p and worst reward r_* otherwise. The use of an adaptive utility function simply states that this utility (probability) value p is uncertain, and that instead of assuming a specific value the DM may assign a probability distribution over it, which is analogous to the use of hierarchical priors.

Beliefs over θ can depend on beliefs over the state of nature ω (Cyert & DeGroot [27] also suggest this), but unlike ω, the actual value of θ will not affect the reward to be realised once a decision has been selected, *i.e.*, θ will not affect the actual physical outcome of the decision (although if probabilistic dependence exists between θ and ω, then it will also exist between θ and the reward r that will be realised following selection of any decision d). In order to demonstrate the role of the state of mind θ in the DM's decision problem, Figures 4.1 and 4.2 show influence diagrams for the classical situation and the adaptive utility situation, respectively (see, *e.g.*, Shachter [95] and the references therein for more details on influence diagrams). Note that in Figure 4.2 we have explicitly permitted the case of probabilistic dependence between θ and ω. Independence between θ and ω could be represented by deleting the arrow connecting these two nodes.

[1]This random quantity may be observable (learning apples are preferred to bananas by eating one of each) or unobservable (only a noisy observation may be available) and depends on the scenario under consideration. We consider how one may learn about θ in Chapter 5.

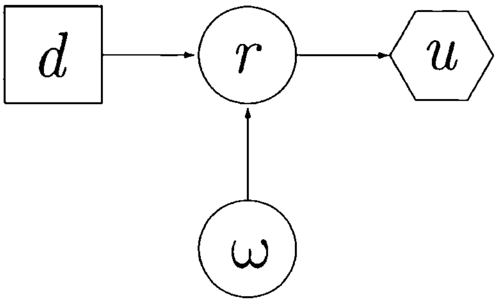

Figure 4.1: Influence diagram for certain utility problem.

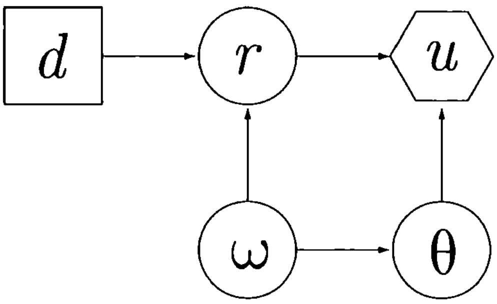

Figure 4.2: Influence diagram for adaptive utility problem.

As a further demonstration of the role of the state of mind in connection to the other decision components within an adaptive utility problem, we can consider the following two examples of situations in which it may be reasonable to assume probabilistic dependence between θ and ω. It should be noticed, however, that these examples are not intended to represent interesting adaptive utility problems (which, and as will be discussed in Chapter 5, are necessarily of a sequential nature), but that rather they simply aid in the explanation of our intended role for a state of mind.

For a first example of possible probabilistic dependence between θ and ω, we can consider the following situation. A DM must decide which of two cake shops she will visit. She can only visit one of them and must purchase the cake that is on offer in that shop. The first shop only sells chocolate cakes, whilst the second switches between selling carrot cakes and sponge cakes. In this situation we let θ represent the DM's true preference relation over lotteries paying one of these cake types as the return, whilst ω is used to represent the type of cake on sale in the second shop.

If the DM believed that the second shop decided each day what type of cake to sell through some random chance mechanism, then there would be probabilistic independence between θ and ω, with the DM's beliefs over θ not influencing her beliefs over what type of cake is on offer. However, the DM might believe that the owner of the second shop is quite good at determining the type of cake his customers on that day are likely to prefer, and selects which cake to sell accordingly. In this case the DM might believe that there exists probabilistic dependence between θ and ω, with the DM believing that if she is more likely to truly prefer carrot to sponge cakes, then the shop owner is more likely to have this type of cake on offer. Notice, however, that once a decision has been made, the true value of θ has no causal influence on the actual reward that is realised.

Another example of possible probabilistic dependence between θ and ω arises when we consider the following situation of a DM who is visiting a kiosk. On each day the kiosk sells only ice-cream or only hot dogs, with the item on sale, represented by the state of nature ω, being decided by the kiosk manager at the beginning of the day according to what he thinks will be the weather on that day. The DM's preferences between ice-cream and hot dogs, as characterised by state of mind θ, is also assumed to depend on the weather. In this situation beliefs over ω and θ are conditionally independent given knowledge of the weather, so the DM's beliefs over ω and θ would only be independent if she knew what the weather on that day would be like.

The adaptive utility concept is closely related to the concept of state-dependent utility[2]. A state dependent utility will alter the DM's preferences depending on the

[2]For further information on state dependent utilities see the discussions and references included in Drèze & Rustichini [37], Karni [62], or Schervish *et al.* [90], *etc.*

state of nature or situation that she finds herself in. For example, a state dependent utility would be appropriate if a return of £100 for a correct backing of Red Rum was preferred to the same win for a correct backing of Desert Orchid (for the DM may just like Red Rum more than Desert Orchid).

The assumption of state dependency is usually removed in classical utility theory (or at least the descriptions of the problem is altered to incorporate the state of nature as an additional attribute of the return from decision selection), and is explicitly removed in the theory of Anscombe & Aumann [5] who derive only state independent utilities (this is attributed to the Monotonicity Axiom A5, which is more commonly reworded as a state independence axiom, see, *e.g.*, Nau [81]). State dependency, however, is similar to the adaptive utility concept where the DM's preferences may change depending on the true state of mind θ. Indeed, in a one-off decision problem an adaptive utility could be considered a special type of state-dependent utility, as even though the true value of θ does not affect the physical reward following decision selection, it does affect the utility value for that reward.

Although the DM is not free to choose her state of mind, she is permitted to choose to act in a way that is inconsistent with her beliefs over it. However, under the requirement that the DM maximises expected adaptive utility, such an action is irrational. To see this assume that the joint distribution $P_{\omega,\theta}$ represents true beliefs over the correct states of nature and mind, and let $P'_{\omega,\theta}$ be any other distribution. Let d_1 and d_2 denote those decisions deemed optimal under $P_{\omega,\theta}$ and $P'_{\omega,\theta}$ respectively. Then, under $P_{\omega,\theta}$, either $d_1 \sim d_2$, in which case the DM achieves the same adaptive utility level by reporting either $P_{\omega,\theta}$ or $P'_{\omega,\theta}$, or $d_1 \succ d_2$, in which case the DM will be acting irrationally by reporting false distribution $P'_{\omega,\theta}$ and selecting decision d_2.

4.2 Axioms of Adaptive Utility

The system of classical axioms of rational choice leads to the maximisation of expected adaptive utility as the logical decision selection technique. To achieve this result in the case of independence between the state of mind and state of nature we use repeated application of the theory of Anscombe & Aumann, which states that, assuming the DM agrees to the relevant axioms outlined in Subsection 2.2.1, there exists a unique utility function (up to a positive linear transformation) that is in agreement with the DM's preferences, regardless of whether decisions are seen as horse lotteries or as roulette lotteries. Furthermore, the utility of a non-degenerate decision is equal to the probability weighted average of the utilities of the individual possible rewards. Formally, we wish to prove the following result:

Theorem 4.2.1
Acceptance of axioms A1-A6 of Subsection 2.2.1 implies that, for a given distribution P_θ, an adaptive utility function ${}_a u(\cdot)$ exists and is the unique function (up to a positive linear transformation) that satisfies the following two properties:

1. For all $d_1, d_2 \in \mathcal{D}$, $d_1 \succeq d_2 \Leftrightarrow {}_a u(d_1) \geq {}_a u(d_2)$.

2. For all $d_1, d_2 \in \mathcal{D}$ and $p \in (0,1)$, ${}_a u(pd_1 + (1-p)d_2) = p_a u(d_1) + (1-p)_a u(d_2) \diamond$.

Theorem 4.2.1 is of course a generalisation of the traditional von Neumann & Morgenstern expected utility result that was discussed in Subsection 2.2.1, returning to that case when P_θ is a degenerate distribution. In effect this means that the adaptive utility function is the DM's actual utility function for this setting. That is not to say the adaptive utility function represents the true underlying preferences of the DM, and only when P_θ is degenerate will this be the case. Rather, the adaptive utility function is representing the DM's preferences over decisions when it is accepted that true preferences are uncertain and when such uncertainty is modelled by distribution P_θ.

The proof of existence of the function is straight forward, using both the fact that possible classical utility functions $u(\cdot|\theta_1), \ldots, u(\cdot|\theta_n)$ exist and that the adaptive utility function is defined in Definition 3.2.1 for a given distribution P_θ. Also, that the function satisfies property 2 arises as a direct consequence of the analogous property of classical utilities and because expectation is a linear operator. Hence, what remains to be shown to prove Theorem 4.2.1 is that the adaptive utility function is the unique function for this setting that satisfies property 1.

In order to proceed we consider a decision outcome set ${}_a\mathcal{R}$ that consists of all possible classical utility values for all possible decisions. This means that, in addition to a decision being associated with a distribution over $\mathcal{R}$, it will also now be associated with a distribution over ${}_a\mathcal{R}$. Additional use of Anscombe & Aumann's result (see Section 2.2) then means that there exists a unique (up to a positive linear transformation) utility function $u^* : {}_a\mathcal{R} \to \mathbb{R}$ representing preferences over elements of ${}_a\mathcal{R}$. Such a utility function, presuming the set of possible classical utility functions have been scaled to ensure they are commensurable, can simply return the original value that was the element of ${}_a\mathcal{R}$. Property 2 of a classical utility function can then be used to extend the preference relation to cover the entire set of decisions that are equivalent to distributions over ${}_a\mathcal{R}$. However, this is exactly what an adaptive utility function does in taking the expectation of possible utility values, and so as the utility function u^* is unique, it must then be the adaptive utility function.

All that remains is to question whether or not considering the outcome of a decision as a utility value, and as such an element of ${}_a\mathcal{R}$, affects the rationality of any of Axioms A1-A6. This is a question the DM must consider herself. However, under the assumption that such utility values are scaled to ensure commensurability, and that it is reasonable to consider preferences over hypothetical gambles that can never actually be played (the DM is not permitted to affect the true value of θ), it would appear that there is no additional reason for not accepting them. The only possible query would be whether or not preferences over utility values depend on the state of mind that occurred. However, this is not the case as once it is assumed that it is meaningful to compare various classical utility values, decisions are being viewed as lotteries over numerical values, with preference always being in agreement with the size of the particular value considered.

4.3 Constructing Commensurable Utilities

In the previous section we argued that, under the assumption that Axioms A1-A6 are true, the optimal decision for selection is that which maximises the expectation (with respect to beliefs over preferences) of the various possible utility functions. However, as we mentioned in Section 3.3, Boutilier [21] demonstrates that such a decision rule is not invariant to the linear scaling of the possible utility functions. Furthermore, in defining an adaptive utility and in the arguments of Section 4.2, we had to assume that the various classical utilities had been suitably scaled to ensure commensurability. Thus there exists a problem in determining the appropriate scaling of each function to ensure commensurability, *i.e.*, to ensure that different utility values from differing utility functions can be meaningfully compared.

To illustrate this problem consider two possible utility functions $u(d_i|\theta_1) = 2i$ and $u(d_i|\theta_2) = I_{\{i=1\}}$ for $i \in \{1, 2\}$ (here $I_{\{i=1\}}$ is used to represent the indicator function that returns value 1 if $i = 1$ and value 0 otherwise). Assume that prior beliefs are such that $P(\theta = \theta_1) = P(\theta = \theta_2) = 0.5$, so the adaptive utility maximiser should select decision d_2. However, the function $\tilde{u}(d_i|\theta_2) = 3I_{\{i=1\}}$ is a positive linear transformation of $u(d_i|\theta_2) = I_{\{i=1\}}$, and thus represents exactly the same preferences under θ_2. Yet when $\tilde{u}(\cdot|\theta_2)$ is used, the adaptive utility maximiser should select decision d_1. The issue then is, knowing that both $u(\cdot|\theta_2)$ and $\tilde{u}(\cdot|\theta_2)$ (and also the infinite number of alternative positive linear transformations of $u(\cdot|\theta_2)$) represent exactly the same preference ordering over the set of decisions, which is appropriate for performing our adaptive utility calculations? Clearly we have to be careful as each of $u(\cdot|\theta_2)$ and $\tilde{u}(\cdot|\theta_2)$ leads to the selection of a different decision.

Boutilier [21] resolved this problem by making an assumption of extremum equivalence. This assumption requires that, under each possible utility function, there exists the same most favourable reward r^* and the same least favourable reward r_*. Furthermore, it is also required that each possible utility function is normalised so as to give the same utility value for r^* and the same utility value for r_*.

When extremum equivalence does hold Boutilier shows that, by viewing the decision problem as the selection of a compound gamble in which the first stage is a gamble over the utility function and the second a standard gamble over rewards (similar to our repeated application of Anscombe & Aumann's theory in Section 4.2), there exists a method for scaling classical utilities so that commensurability is possible.

This assumption of extremum equivalence, however, is unreasonable for many decision problems. For example, we may wish for a particular state of mind to represent a preference reversal. This could indeed be the case in Example 3.1.1, where one state of mind may represent a preference for apples, whilst an alternative leads to a preference for bananas. Furthermore, even in situations where it is reasonable to assume that there exists a reward that is certainly considered best and another that is certainly considered worst, we may wish for the strength of preference, with respect to at least one more option, to vary.

As was discussed in the previous section, an adaptive utility function is an actual utility function, ranking decisions by their expected utilities (under the assumption of commensurability) with respect to beliefs over the correct preference ranking as determined by the state of mind θ. For this reason we can construct an adaptive utility function in the same way as classical utilities are traditionally constructed. To do this, first assume that (r^*, θ^*) is the best reward and state of mind pair (*i.e.*, receiving reward r^* if θ^* were true would be at least as preferable to receiving any other reward under any other state of mind), and that (r_*, θ_*) is the worst reward and state of mind pair (*i.e.*, receiving reward r_* if θ_* were true is at least as less bad as receiving any other reward under any other state of mind). Under this assumption the adaptive utility value of any other reward and state of mind pair (r, θ) is the number $p \in [0, 1]$ such that the DM is indifferent between receiving r under state of mind θ for certain, or playing the gamble paying r^* under state of mind θ^* with probability p and r_* under state of mind θ_* otherwise.

Before continuing it is important to note that this construction method, and the method suggested by Boutilier, requires the DM to consider preferences over reward and state of mind pairs. Indeed, in doing this we are now viewing the state of mind as an attribute of decision outcome, rather than as simply a part of the state space. This task is far from trivial, requiring the DM to make hypothetical choices that can not be played in reality (the DM is not permitted to choose the true value of θ). Although introducing a state of mind makes the task of utility elicitation harder, the assumption that hypothetical choices can be made is also included in the Completeness Axiom A1 of classical utility theory (see Section 2.2). Nevertheless, in an adaptive utility setting the DM must agree to comparability between all distributions over possible reward and state of mind combinations.

The requirement to consider distributions over reward and state of mind pairs is more than what is required when utility is not uncertain, and we note that the implication of repeated use of Axiom A1 in the construction of an adaptive utility function requires not only that there is a system of classical orderings (one for each value of θ), but that these be grouped and combined together when decisions are considered under differing values of θ. The alternative is not to use Axiom A1, leading to interval utility theories as discussed in Chapter 2 (the cost being that the DM may be presented with a situation without a full preference ordering over the set of decisions).

It is assumed that in applications of adaptive utility theory, a DM will select various possible classical utility forms through relevant properties that the state of mind utility parameter will introduce into the problem. Formally, however, a DM should be aware of the influence this parameter has on reported preferences, and perhaps this task may be made easier by providing a description of what the state of mind actually represents.

For example, in our apple and banana example we could say that one state of mind represents the situation where "bananas are just as nice as my favourite food", whilst another could represent the situation where "bananas are as bad as my least favourite food".

For more interesting problems, in particular with many or even an infinite[3] amount of possible values for θ, this task becomes increasingly difficult or even impossible. In this case it would appear that there is no other possibility than to return to selecting possible classical utility functions because of appropriate properties they have.

The construction method that assigns adaptive utility values by considering preferences over lotteries that have as their return a reward and state of mind pair, bounds all adaptive utility values to the interval $[0,1]$. However, a suitable transformation can lead to the function being restricted to any finite interval $[a,b]$ (as an adaptive utility function, just like a classical utility function, is unique up to a positive linear transformation). Yet, if the DM really were to use this construction method, then a major motivation for the use of adaptive utilities is lost, *i.e.*, preferences over rewards are again assumed to be known outright (even worse there is a greater number of rewards to consider due to the additional attribute of the state of mind). Instead, we demonstrate below that in applications where the DM is able to state her collection of possible classical utility functions, and providing these are scaled so as to be commensurable, she is be able to create her adaptive utility function directly from them.

To achieve commensurability of classical utility functions $u(r|\theta_1)$, $u(r|\theta_2)$, ..., $u(r|\theta_n)$ we constrain them all to fall within the interval $[0,1]$ and normalise the adaptive utility function to cover this interval by specifying $u(r^*|\theta^*) = 1$ and $u(r_*|\theta_*) = 0$ (here we implicitly reject the trivial situation where all reward and state of mind pairs are viewed as equally preferable). Note that in determining r^* and r_* it is sufficient to check only those results that maximise or minimise some classical utility

[3]A derivation of adaptive utility theory for continuous beliefs over θ has been avoided in this work, but to enable (in Chapter 6) examination of the effects adaptive utility has on utility diagnostics *etc.*, we assume that there is no difficulty in allowing continuous distributions over θ.

function. We then scale each classical utility function $u(r|\theta_k)$ as follows:

With the notation $(p_1 a_1, \ldots, p_n a_n)$ representing the lottery paying a_i with probability p_i, and (r_k, θ_k) representing receiving r_k when θ_k is the true state of mind:

- If $u(r|\theta_k)$ is such that $u(r_1|\theta_k) = u(r_2|\theta_k)$ for all $r_1, r_2 \in \mathcal{R}$, then we set $u(r|\theta_k) = p$ with $p \in [0,1]$ such that $\big(1(r,\theta_k)\big) \sim \big(p(r^*,\theta^*),(1-p)(r_*,\theta_*)\big)$.
- Otherwise there exists $\overline{r}_{\theta_k}, \underline{r}_{\theta_k} \in \mathcal{R}$ such that, for any other $r \in \mathcal{R}$, we have $u(\overline{r}_{\theta_k}|\theta_k) \geq u(r|\theta_k) \geq u(\underline{r}_{\theta_k}|\theta_k)$ and strict relation $u(\overline{r}_{\theta_k}|\theta_k) > u(\underline{r}_{\theta_k}|\theta_k)$. We scale such a utility function by using the two constraints $u(\overline{r}_{\theta_k}|\theta_k) = \overline{q}_{\theta_k}$ and $u(\underline{r}_{\theta_k}|\theta_k) = \underline{q}_{\theta_k}$, where $\overline{q}_{\theta_k}, \underline{q}_{\theta_k} \in [0,1]$ are respectively determined by considering the lottery making $\big(1(\overline{r}_{\theta_k},\theta_k)\big) \sim \big(\overline{q}_{\theta_k}(r^*,\theta^*),(1-\overline{q}_{\theta_k})(r_*,\theta_*)\big)$ and $\big(1(\underline{r}_{\theta_k},\theta_k)\big) \sim \big(\underline{q}_{\theta_k}(r^*,\theta^*),(1-\underline{q}_{\theta_k})(r_*,\theta_*)\big)$ true.

Under this scaling of classical utilities, commensurability is ensured without the assumption that extremum equivalence holds. The following example illustrates how knowledge of the form of each of a collection of classical utility functions simplifies formulation of the adaptive utility function.

Example 4.3.1

A DM has been given a diagnosis for which there are several incompatible treatments. The result r in this scenario is a two dimensional vector (s,t), with s measuring severity of the side-effect and t measuring the relative time until complete recovery. Both s and t run between 0 and 1, with a higher score representing a better situation. The DM has no experience of requiring such medical treatment and is thus uncertain over her preferences for different possible values of multi-attributed r. Hence we assume that her two possible states of mind are:

- θ_1: It is most important the DM recovers as quickly as possible.
- θ_2: The DM views recovery time as only being fairly important.

We assume that these possibilities are represented by a positive linear transformation of the classical utility functions $u(r|\theta_1) = t$ and $u(r|\theta_2) = 0.5t + 0.5s$, respectively, and will seek to scale these functions in order to ensure they are commensurable.

Rather than trying to construct an adaptive utility function based on comparing all combinations of possible results r and possible states of mind θ, we aim to find a suitable scaling of the above classical functions so that ${}_au(r) = E_\theta[u(r|\theta)]$. First the DM must provide the best and worst combination of r and θ. Suppose that these are $(r^*, \theta^*) = \big((1,1), \theta_1\big)$ and $(r_*, \theta_*) = \big((0,0), \theta_1\big)$, respectively. As both of these relate to the case where $\theta = \theta_1$, we scale the function $u(r|\theta_1)$ to cover the interval $[0,1]$ which, as it currently does so, requires no alteration.

Next the DM considers the best and worst results under θ_2. These are again the pairs $(1,1)$ and $(0,0)$ respectively. We now determine probability values $\overline{q}_{\theta_2}$ and $\underline{q}_{\theta_2}$ such that both the relations $\big(1\big((1,1),\theta_2\big)\big) \sim \big(\overline{q}_{\theta_2}\big((1,1),\theta_1\big),(1-\overline{q}_{\theta_2})\big((0,0),\theta_1\big)\big)$ and $\big(1\big((0,0),\theta_2\big)\big) \sim \big(\underline{q}_{\theta_2}\big((1,1),\theta_1\big),(1-\underline{q}_{\theta_2})\big((0,0),\theta_1\big)\big)$ hold. Assuming that, upon consideration, the DM assigns $\overline{q}_{\theta_2} = 0.7$ and $\underline{q}_{\theta_2} = 0.2$, we are able to determine the values of constants a and b in the generic form for a utility function representing the same preferences as $u(r|\theta_2)$, *i.e.*, $u(r|\theta_2) = a(0.5s + 0.5t) + b$.

The scaling is then fixed by finding constants a and b which simultaneously solve the equations $a\big(0.5(1) + 0.5(1)\big) + b = 0.7$ and $a\big(0.5(0) + 0.5(0)\big) + b = 0.2$. This leads to $b = 0.3$ and $a = 0.4$, and provides us with our scaling for $u(r|\theta_2)$, *i.e.*, $u(r|\theta_2) = 0.4(0.5s + 0.5t) + 0.3$. To determine ${}_au(r)$ only the DM's subjective beliefs $P(\theta = \theta_1) = 1 - P(\theta = \theta_2)$ are now required. $\diamond$

Example 4.3.2 below also illustrates this system for construction of an adaptive utility function, and is specifically concerned with uncertainty over the diagnostic of risk aversion (a concept later discussed in Chapter 6).

Example 4.3.2

We assume that $\mathcal{R} = [0, 100]$, $\theta \in \{10, 20, 30\}$, and that classical utilities are such that $u(r|\theta) = a_\theta \log(r + \theta) + b_\theta$ with $a_\theta > 0$ (here θ influences local risk aversion and a_θ and b_θ set the scale). To place these utility functions on a suitable scale to be used in constructing the adaptive utility function, we make the assumption that, under any value of θ, $u(100, \theta) = 1$ and $u(0, \theta) = 0$ (this is extremum equivalence).

This provides 6 independent linear equations with 6 unknowns:

$$\begin{aligned}
a_{10}\log(110) + b_{10} &= 1, \qquad & a_{10}\log(10) + b_{10} &= 0\\
a_{20}\log(120) + b_{20} &= 1, \qquad & a_{20}\log(20) + b_{20} &= 0\\
a_{30}\log(130) + b_{30} &= 1, \qquad & a_{30}\log(30) + b_{30} &= 0
\end{aligned}$$

This results in (where numeric answers are given to 2 decimal places):

$$\begin{aligned}
a_{10} &= \frac{1}{\log(110) - \log(10)} = 0.96, \qquad & b_{10} &= -\frac{\log(10)}{\log(110) - \log(10)} = -0.96\\
a_{20} &= \frac{1}{\log(120) - \log(20)} = 1.29, \qquad & b_{20} &= -\frac{\log(20)}{\log(120) - \log(20)} = -1.67\\
a_{30} &= \frac{1}{\log(130) - \log(30)} = 1.57, \qquad & b_{30} &= -\frac{\log(30)}{\log(130) - \log(30)} = -2.32
\end{aligned}$$

Appropriately substituting these into the functions $u(r|\theta)$ results in ensuring commensurability.

Chapter 5

Applications

This chapter focuses on the use of the adaptive utility concept in sequential decision making. We consider implications for decision selection strategies and possibilities for the DM to learn about her utilities as she progresses through her decision sequence. The chapter concludes with discussion and examples of the use of adaptive utility in the specific area of reliability theory.

5.1 Sequential Decision Problems

In a one-off solitary decision problem, the use of adaptive utility and the permittance of uncertainty over preferences does not lead to any difference in selection strategy from that arising from the assumption that the DM's true utility function equals the expectation of possible utility functions. Adaptive utility is thus a generalisation of classical expected utility theory that is of little benefit in this setting. Nevertheless, there are many situations in life where a DM is required to make a sequence of decisions, and where the outcome and selection of previous decisions may well be relevant for changing beliefs and options for future decisions. In such situations connections over the whole sequence of choices are relevant when determining an optimal decision selection strategy. The solution to such sequential decision problems is often the focus of Bayesian statistical decision theory (see, for example, [15, 34, 46]).

The type of problem under consideration in this chapter can be described in the following way. We assume that the DM has to make a sequence of n decisions $d^1, d^2, \ldots, d^n$, and that following the selection of each decision d^i the DM receives a return r^i. Furthermore, when considering the selection of decision d^i, the DM is aware of past decisions $d^1, d^2, \ldots, d^{i-1}$ she had previously selected and the returns $r^1, r^2, \ldots, r^{i-1}$ they respectively led to. The DM's objective is to maximise the utility of the entire sequence of decisions given by a function $u(d^1, d^2, \ldots, d^n)$. An influence diagram of a two stage problem is given in Figure 5.1 below, with the arc connecting node r^1 to r^2 representing the setting that beliefs over outcome r^2 of decision d^2 can depend on the outcome r^1 that was observed following selection of decision d^1. Also note that, in order to keep the diagram reasonably simple, Figure 5.1 now omits the state of nature ω; however, in a full graphical representation ω would be included as an additional node with arcs entering both r^1 and r^2.

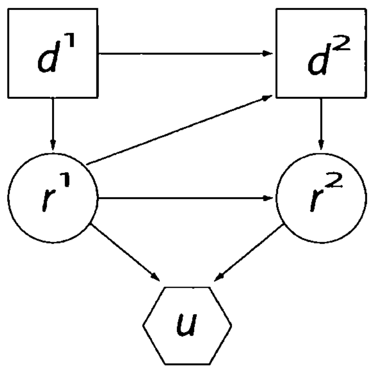

Figure 5.1: Influence diagram for classic 2-period sequential problem.

Interesting problems in this area arise when there is initial uncertainty over what the actual outcome of any decision selection will be, but where observations of past outcomes are relevant in determining likely outcomes of future decisions. For this reason it can often be beneficial, in a decision sequence of suitable length, to initially select decisions in which there is large uncertainty as to the likely outcome. The reason for this is that the observation of the initial return provides information that can be used in future decision selection.

If a decision which has large prior uncertainty over its outcome is selected and is noted to lead to a beneficial outcome, then it would appear that the selection of that decision will lead to a beneficial outcome in general, and so it should be selected again. However, if the outcome of such a decision was seen to be bad, then the DM can learn about this and avoid that decision in the future, hence only suffering the bad outcome once.

Determining an optimal selection strategy for a finite length decision sequence is a dynamic programming problem, and is solved through the use of backward induction and Bellman's Equation [12, 13]. This solution requires the DM to consider all possible histories she may have observed by the time she comes to selecting the n-th decision. She then determines what would be the optimal decision d^n for each possible decision history, and this will form her decision strategy for that time point. Formally, given a decision history H^n that details all her past decisions $d^1, \dots, d^{n-1}$ and the rewards $r^1, \dots, r^{n-1}$ that they respectively led to, the DM determines a function π^n which takes as its argument H^n and returns a feasible entry from $\mathcal{D}^n(H^n)$, the set of decisions available at the n-th selection point given history H^n.

Following dynamic programming, with knowledge of π^n the DM seeks to determine π^{n-1}, a function that takes history H^{n-1} as its argument and returns an element of $\mathcal{D}^{n-1}(H^{n-1})$. This is done by considering the DM's beliefs about the outcome of various decisions in $\mathcal{D}^{n-1}(H^{n-1})$ and the likelihood of moving to any of the histories in H^n (and what that entails for the decision d^n that will be selected according to optimal strategy π^n). Continuing in this manner of considering current beliefs and noting future optimal strategies permits the DM to keep determining optimal strategies for earlier decisions until eventually she has found π^1.

Formally, with $U^i = u(d^1, \dots, d^i, \pi^{i+1}(H^{i+1}), \dots, \pi^n(H^n))$ and $U^n = u(d^1, \dots, d^n)$, the DM should select decision strategy π^i, for $i = 1, \ldots, n-1$, in the following manner:

$$\pi^i(H^i) = \arg \max_{d^i \in \mathcal{D}^i(H^i)} E_{H^{i+1}|H^i, d^i}[\cdots E_{H^n|H^{n-1}, \pi^{n-1}(H^{n-1})}[U^i]] \tag{5.1}$$

The DM should select π^n such that:

$$\pi^n(H^n) = \arg \max_{d^n \in \mathcal{D}^n(H^n)} U^n \tag{5.2}$$

Equation (5.1) contains a nested sequence of expectations, each of which requires the DM to consider a distribution of the form $P_{H^j|H^{j-1}, \pi^{j-1}(H^{j-1})}$. This conditioning argument implies that given H^{j-1} and $\pi^{j-1}(H^{j-1})$, only the reward r^{j-1} following decision $\pi^{j-1}(H^{j-1})$ remains uncertain. Also note that both Equations (5.1) and (5.2) use the more formal version of utility as a function with argument the decision that was selected. As mentioned in Subsection 1.1.1, however, one usually considers a utility function as a function with the return from the selected decision as its argument, and in Equation (5.2) it is the history H^n that is used to determine updated beliefs for the return r^n following decision selection d^n.

In an adaptive utility setting, we no longer assume that the DM's true preferences are known, but rather that her actual utility function $u(d^1, \ldots, d^n)$ for decision stream $d^1, \ldots, d^n$ is uncertain, with such uncertainty being represented by uncertainty over the state of mind θ. In such a situation, Chapter 4 demonstrated that the DM should make decision selection by seeking to maximise their adaptive utility function ${}_au(d^1, \ldots, d^n) = E_\theta[u(d^1, \ldots, d^n|\theta)]$.

However, if the DM is not able to learn about θ as she is moving through her decision sequence, then once again we return to the classical situation, with the only difference that $E_\theta[u(d^1, \ldots, d^n|\theta)]$ replaces $u(d^1, \ldots, d^n)$ in Equations (5.1) and

(5.2). For this reason we will assume in this chapter that the DM is able to learn about θ as she progresses through her decision sequence. Section 5.2 will look into the specific details of how this could be the case.

Assuming the DM is able to learn about θ as she moves through her decision sequence, by the time she considers decision d^j, prior distribution P_θ will no longer be relevant in representing beliefs over her state of mind. Instead posterior distribution $P_{\theta|G^j}$ should be used, with G^j the additional information about θ received by the time d^j is to be selected. This is different to the classical theory, as now the function that the DM is seeking to maximise, ${}_a u(d^1, \ldots, d^n)$, will no longer be assumed constant over the duration of the decision sequence, hence providing motivation for referring to this utility function as an *adaptive* utility function. Nevertheless, formally we should state that an adaptive utility function also uses the DM's beliefs about state of mind θ as an additional argument, in which case (viewing an adaptive utility as a function of both decision selection and beliefs over θ) the function to be maximised does not change. However, as opposed to decisions, the DM is not free to select her beliefs over θ and as such we have chosen to drop its inclusion as a formal argument of the function.

In order to learn about θ we assume the DM observes utility information z^i following selection of decision d^i, and an influence diagram representing this situation for a 2-period sequential decision problem is given in Figure 5.2. Note that, as with Figure 5.1, a full representation should include a node for ω with arcs entering both r^1 and r^2, but this has again been removed in order to simplify the diagram. Also note that Figure 5.2 represents a situation in which the information learnt about the DM's utility function, as represented by the z^1 node, does not depend on the reward r^1 obtained following selection of decision d^1. However, if the utility information did depend on the actual reward obtained, as may be appropriate in some situations, then an additional arc would be required going from node r^1 to node z^1.

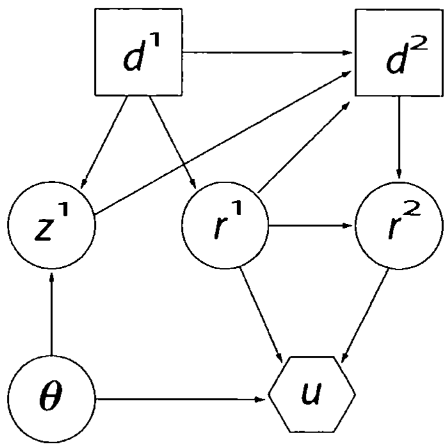

Figure 5.2: Influence diagram for an adaptive utility 2-period sequential problem.

The utility information z^i can be anything that the DM deems relevant for learning about her true preferences as characterised by her state of mind θ. Depending on the particulars of the problem z^i may or may not depend on the reward r^i obtained following the selection of decision d^i, but as will be discussed below, interesting adaptive utility problems only arise if z^i does in some way depend on the actual selection of decision d^i. Section 5.2 will further mention how such utility information may arise and will also discuss how the learning of utility takes place in some literature examples that consider a similar setting.

As discussed above, Figure 5.2 explicitly includes an arc from d^1 to z^1, meaning that the information received about utility parameter θ will depend on the decision d^1 that is chosen. However, in [29, pp.133-135] Cyert & DeGroot make the following statement (where they use notation such that θ is the uncertain utility parameter and δ_i is the i-th period decision):

> "Indeed, even in a two-period problem in which ... the information obtained about θ in the first period does not depend on which decision δ_1 is chosen or which consequence r_1 occurs, the decision maker must take possible changes in utility into account when choosing δ_1 ... The impor-

> tant feature of this result from our present point of view is that when the decision maker chooses δ_1[1], he or she must consider the full range of possible new expected utility functions ... that might be obtained from the observation ... It should be noted that both δ_1 and δ_2 will be different from the decisions that would be optimal in traditional theory ... It should be emphasised, moreover, that even though the decision maker has no control over the information about θ that will be generated in the first period, the optimal δ_1 in our approach will also be different from the one in traditional theory."

We believe Cyert & DeGroot are mistaken in this statement. We interpret it as indicating that a different d^1 could be selected if utility information were expected after its selection, even if such information is independent of d^1. However, this would be similar to suggesting that prior beliefs over what will be the DM's posterior distribution over a parameter of interest, following observation of relevant data, can be different to the current prior distribution over that parameter.

In situations where the selection of d^1 has no influence on the observation z^1 that is informative about θ, the DM should select the same decision that they would have selected if it were assumed that no information about the state of mind would be made available. Only if the DM is able to influence z^1 by selection of d^1 would there be an additional value to selecting a particular d^1 due to the information gained over θ. This would appear the more interesting case due to its difference to classical theory, and hence the inclusion of the arc between d^1 and z^1 in Figure 5.2.

If utility information is independent of decision selection, the sequential adaptive utility setting is similar to the Assumed Permanence approach of Witsenhausen [108] that was discussed in Subsection 2.3.3. In that theory it was accepted that

[1] In the original text this term is given as θ_1, yet we believe this to be a typing error. Also, omissions in the quotation, represented by ellipses, refer to material irrelevant to this discussion.

future preferences may not be the same as current preferences, but that decisions should be made under the assumption that they will be the same. Once further information is known, new preferences are identified and new decisions are made under the assumption that such preferences will again remain constant, and this form of decision making continues indefinitely.

In the adaptive utility setting we do not require true preferences to change, instead it is only assumed that they are uncertain. Yet, when utility information is independent of decision selection, decisions should be made under the assumption that true preferences are equal to current expectations, even though it is accepted that in the future there may be reason for the DM to change her beliefs over her true preferences. When such information is gained, expectations of true preferences may change and future decisions will again be selected under the assumption that true preferences are equal to the new expectations, and so on. This idea of equating expected future preferences with current expected preferences, but permitting the possibility that in the future the DM's expectations may be found to be incorrect, is analogous to Goldstein's Temporal Coherence concept for comparing current and future beliefs (see, *e.g.*, [50]).

In [29], Cyert & DeGroot restrict attention to offering a solution to the two-period sequential decision problem for uncertain utility. However, following a similar argument as outlined above for the classical sequential problem, we can provide the solution to a generic n-period problem. In this case, assuming independence between the state of nature and the state of mind, we use ${}_aH^j$ to represent the relevant decision making history prior to selection of decision d^j, *i.e.*, ${}_aH^j$ lists the past decision sequence $d^1, \ldots, d^{j-1}$, the outcomes $r^1, \ldots, r^{j-1}$, and utility information $z^1, \ldots, z^{j-1}$ that these respectively led to. Furthermore, under the notational convention that ${}_aU^i = {}_au(d^1, \ldots, d^i, \pi^{i+1}({}_aH^{i+1}), \ldots, \pi^n({}_aH^n))$ and ${}_aU^n = {}_au(d^1, \ldots, d^n)$ (where the expectation over θ in the adaptive utility is performed with beliefs conditioned on history ${}_aH^i$), the DM should select decision strategy π^i, for $i = 1, \ldots, n-1$, in the following manner:

$$\pi^i({}_aH^i) = \arg \max_{d^i \in \mathcal{D}^i({}_aH^i)} E_{{}_aH^{i+1}|{}_aH^i,d^i}[\cdots E_{{}_aH^n|{}_aH^{n-1},\pi^{n-1}({}_aH^{n-1})}[{}_aU^i]] \tag{5.3}$$

The DM should select π^n by:

$$\pi^n({}_aH^n) = \arg \max_{d^n \in \mathcal{D}^n({}_aH^n)} {}_aU^n \tag{5.4}$$

Note that Equation (5.3) contains a nested sequence of expectations, each of which requires the DM to consider a distribution of the form $P_{{}_aH^j|{}_aH^{j-1},\pi^{j-1}({}_aH^{j-1})}$. This conditioning argument implies that given ${}_aH^{j-1}$ and $\pi^{j-1}({}_aH^{j-1})$, only the reward r^{j-1} and utility information z^{j-1} following decision $\pi^{j-1}({}_aH^{j-1})$ remain uncertain. Also note that in both Equations (5.3) and (5.4) we have again used the more formal version of an adaptive utility function as a function that has as its argument the decision that was selected. Again, however, it is presumed that it would be easier for elicitation purposes to consider the adaptive utility function as a function with argument the return from the selected decision, with ${}_aH^n$ used to determine updated beliefs for the return r^n in Equation 5.4.

Equation (5.3) details how the DM should select decisions given the history she has so far observed. The DM starts by considering Equation (5.4) and determines the final decision she would make for each conceivable history ${}_aH^n$ that could have been observed by that time. Once the DM has determined decision strategy π^n, she considers how she would make decision d^{n-1}. Now Equation (5.3) is used and the DM determines the optimal decision d^{n-1} for all possible histories ${}_aH^{n-1}$ that she would have observed by that time. This is done by considering updated beliefs over the likely reward outcome for each decision, but also now by considering updated beliefs over θ. Determining the probability of moving to any history ${}_aH^n$ for each decision d^{n-1}, the DM seeks to pick decisions to maximise expected adaptive utility. This process continues backward through decisions until an optimal first period decision is found. The important difference in the adaptive utility setting is that, provided decision selection influences the utility information that is to be received, the DM can find decisions which, whilst under current beliefs over preferences would appear to be sub-optimal, are in fact optimal due to the greater source of information over preferences they are likely to reveal. That is to say, adaptive utility allows the DM to be 'forward looking' in both learning of decision outcomes and of preferences. We illustrate this using the Apple or Banana example (see Example 3.1.1).

Example 5.1.1

The DM is seeking to determine whether to select an apple or banana, but is uncertain of her preference relation between the two. We assume in an n-period problem that $\mathcal{D}^i = \{d_A, d_B\}$, with $d^i = d_A$ representing selection of an apple in period i, and d_B the selection of a banana. The DM's adaptive utility function is ${}_a u(d^1, \ldots, d^n) = E_\theta[\sum_{i=1}^n I_{\{d^i=d_A\}} + \theta \sum_{i=1}^n I_{\{d^i=d_B\}}] = E_\theta[k + (n-k)\theta]$, with k the number of times $d^i = d_A$ (note that this function does not distinguish between ordering in the decision sequence, only the number of times decisions are selected). Here θ represents the additional increase in utility from selecting $d^i = d_B$, and we assume commensurable utility (see Section 4.3) such that $\theta \in \{\theta_0 = 0, \theta_1 = 2\}$.

We assume that selection of the banana informs the DM of her preferences (*i.e.*, she becomes certain of the correct value for θ), and that never selecting the banana means that she never observes any additional information. In this case utility information z^i will correspond to the value of θ if the decision is made to select a banana, whilst if an apple is selected no utility information is available. Also assume prior beliefs $P(\theta = \theta_0) = p$. In this situation we can consider, for a given length n of the problem, a value for p such that the DM will make initial decision selection $d^1 = d_B$. Certainly this will always be the case if $p \leq 0.5$, as this represents a situation where prior beliefs are such that the DM expects that she will enjoy the banana more than the apple. However, even when $p > 0.5$, meaning the DM expects that she will not prefer the banana over the apple, there exists a smallest value n_B such that, for all $n \geq n_B$, the DM will find that it is still optimal to select the banana in the first period.

We only consider possible decision sequences in which, if decision d_B is ever chosen, then it must be chosen in the first period. We will demonstrate the irrationality of selecting $d^1 = d_A$ and then, for some $i > 1$, selecting $d^i = d_B$ when we discuss value of information in Chapter 6 (see Example 6.1.3). Hence, selecting $d^1 = d_A$ means that $d^i = d_A$ for all $i = 1, \dots, n$ and so, for an n-period problem, the adaptive utility return will be ${}_a u(d^1 = d_A, \dots, d^n = d_A) = n$.

However, if $d^1 = B$, then with prior probability p the DM discovers that her true preferences are for apples (the case that $z^1 = 0$), otherwise she discovers that her true preferences are for bananas (the case that $z^1 = 2$). In the former case the DM will select $d^i = d_A$ for all $i > 1$, and in the latter the DM will select $d^i = d_B$ for all $i > 1$. This results in an expected adaptive utility value of ${}_a u(d^1 = d_B, d^2, \dots, d^n) = p(n-1) + 2(1-p)n$. Hence, for $n \geq 2$ the DM should select $d^1 = d_B$ when p is such that $p < \frac{n}{n+1}$. Clearly $\frac{n}{n+1}$ is a monotonically increasing function of n with limit equal to 1 and hence, provided the DM accepts the possibility that bananas could be preferred to apples ($p \neq 1$), there will be some smallest value $n_B = \frac{p}{1-p}$ for the length of the decision sequence from which on it would be optimal to select $d^1 = d_B$.
$\diamond$

The use of backward induction for solving dynamic programming problems suffers from the so-called 'curse of dimensionality' (see Bellman [13, p.XII]), the effect of which means that a small increase in the number of variables in the sequential problem leads to a drastic increase in the number of calculations required for its solution. It is computationally very expensive to solve dynamic optimisation problems in this manner when the dimensions of state variables are large. This is the case for classical problems, where the DM has to calculate a nested sequence of expectations before finding the optimal decision strategy. In the case of generalised distributions (where a mixture of probability densities and discrete probabilities are used to represent beliefs) it may be that, given the current state in the solution of integrals, this is an impossible task and a solution must instead be approximated. The additional requirement in adaptive utility of considering a possibly multi-dimensional state of mind enlarges the dimension of the problem's variables, thus exasperating the situation. Hence in using adaptive utilities the DM should be aware that the benefit of permitting uncertain preferences comes with an increased cost in computational complexity.

The computational complexity of solving an adaptive utility problem is currently the greatest hindrance in its use for solving real-world problems. Nevertheless, it may well be possible to identify forms of utility functions that are not only reasonable for modelling possible preferences, but which also greatly reduce the computational requirements of solution algorithms. Just as the theory of conjugate probability distributions simplifies computation of posterior distributions by simply keeping track of a few summary statistics (see, *e.g.*, Raiffa & Schlaifer [84]), it may be possible to identify forms of utility functions that allow a DM to quickly determine the value of a nested sequence of expectations. A similar problem was considered by Lindley [70], who sought to find a conjugate family of utility functions that would be suitably 'matched' to distributions of the exponential family, and which would at the same time also be suitable for modelling realistic preferences. Unfortunately, the ideas which Lindley employs for easily determining the expected utility of a solitary decision do not appear to generalise readily to the sequential case, and there appears to have been no further development in the literature for identification of a conjugate utility family for this situation.

To demonstrate what might be possible, consider the polynomial utility function that takes the following form:

$$u(r^1, \ldots, r^n|\theta) = \sum_{k_0=0}^{m_0} \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} a_{k_0,k_1,\ldots,k_n} (r^1)^{k_1} \cdots (r^n)^{k_n} \theta^{k_0} \tag{5.5}$$

This function appears as a polynomial in all the arguments of $r^1, \ldots, r^n$, and θ. However, as demonstrated in Appendix C, it has the property that, when beliefs take the form of a series of Normal distributions and updating occurs through Normal-Normal conjugacy, it allows a closed and tractable solution for when we sequentially take expectations with respect to beliefs over all the variables $r^1, \ldots, r^n$, and θ.

Although this class of utility functions is not suitable for modelling all realistic preference relations, there are situations where it would be reasonable. For example, the utility function $u(r^1, \ldots, r^n|\theta) = \sum_{i=1}^n \theta^i r^i$ is included as a special case and is suitable for representing the situation where the DM is uncertain of the appropriate discounting rate θ to be used within the Exponential Discounting Model of sequential decision making (see Subsection 2.3.1). Another possibility would be if $n = 2$ and $u(r^1, r^2|\theta) = \theta u_1(r^1) + (1 - \theta)u_2(r^2)$, with u_i a known polynomial function of r^i. Here θ may represent an unknown trade-off weight. Unfortunately, however, the extension to $n \geq 3$ is not contained within the polynomial class discussed above, but may be included if we allow θ to be multi-dimensional and consider polynomials of its components. In such a case it may well be suitable to use a multi-variate Normal distribution with the hope that this still allows a similar result to that discussed in Appendix C.

A further extension for a multi-attribute utility function that does not make such strong independence assumptions between preferences over return levels in differing periods is that of the multiplicative utility function:

$$u(r^1, \ldots, r^n|\theta) = \left[\Pi_{i=1}^n [\theta^* \theta_i u_i(r^i) + 1] - 1\right] / \theta^* \tag{5.6}$$

Here the parameters $\theta = (\theta_1, \ldots, \theta_n)$ (with $\theta_i \in (0, 1)$) and $\theta^* > -1$ represent non-zero scaling constants (see, *e.g.*, Keeney [63]). Again, provided u_i is a polynomial function of r^i and that only two components of θ are uncertain, this function is a member of the polynomial class discussed above. If more than two of the θ_i are uncertain, then again there may be a possibility of using a multi-variate Normal distribution for θ.

Despite those situations mentioned above, there are many preference relations that can not be represented by a polynomial utility function, *e.g.*, exponential or logarithmic utilities, the latter of which is a common assumption for the form of a DM's preferences over monetary returns (see Bernoulli [18]). Nevertheless, we can approximate such functions through Taylor Polynomials, which have been previously considered for approximating utility functions by Diamond & Gelles [36] and Hlawitschka [55].

For an infinitely differentiable function f of a single variable x, the Taylor Series is defined on an open interval around a as $T(x) = \sum_{n=0}^{\infty} \frac{d^n f(x)}{dx^n}|_{x=a}(x-a)^n/n!$. The function f can then be approximated to a specified degree of accuracy by taking a partial sum of this series, and each such partial sum will be of the form of a polynomial in x. This result can also be generalised for approximating multi-variate functions, hence allowing a greater class of utility functions to be approximated by the polynomial class discussed above.

5.2 Utility Information

In the previous section it was argued that, when using adaptive utility, interesting problems are of a sequential nature, and are such that the DM is able to influence the information she receives about her state of mind θ through appropriate decision selection. We now discuss a few forms that such information may take.

In the Apple or Banana problem, last encountered in Example 5.1.1, it was assumed that θ was fully revealed once the DM made the selection of a banana. Such a situation may well model a sequential decision problem in which there are one or more reward types that have never been experienced by the DM, and in this setting the selection of decisions leading to such novel rewards provides the DM with valuable information about her preferences. Indeed, in this case, and depending on the duration of the sequential problem, it will be of benefit for the DM to begin her decision sequence by selecting decisions that are likely to lead to such novel rewards so as she may learn of her preferences for them. Only once all reward types have been experienced should the DM begin to select decisions that have known higher utility.

It may be, however, that there is residual error and only a noisy observation of the true utility value is possible. For example, this could be the case if θ represented the true utility value of a perfectly ripe banana, and the DM is unsure if she had consumed an under-ripe or over-ripe banana. In this case the state of the banana being either under-ripe, over-ripe, or perfect could be represented by an unobservable state of nature, and a likelihood function then identified for determining the probability that the observation z^i equaled the true value of θ. If information about θ is indeed based upon such noisy observations, then the DM will never be certain of her true preferences. Moreover, the usefulness of such information (as measured by its value, a concept to be introduced in Chapter 6) will be less than for observations which inform the DM of the true value of θ, thus the duration of the sequential problem may have to be increased to warrant initial selection of a decision that has uncertain utility.

In [29], Cyert & DeGroot suggest that a DM may be able to determine whether the actual utility value was above or below that which had been previously expected, and that this is then a source of information relevant in updating beliefs over state of mind θ. If true preferences are additive, $u(d^1, \ldots, d^n) = \sum_{i=1}^n \tilde{u}(d^i)$, with $\tilde{u}$ considered a one-period utility function that the DM is uncertain of. In this case the suggestion of Cyert & DeGroot can be incorporated through categorical data z^i, arising following selection of d^i, such that $z^i = 0$ if ${}_a\tilde{u}(d^i) < \tilde{u}(d^i)$ and $z^i = 1$ if ${}_a\tilde{u}(d^i) \geq \tilde{u}(d^i)$. The true value of $\tilde{u}(d^i)$ does not need to be known, indeed, that case was discussed above where the DM observed true utility value by appropriate decision selection. Instead, all that is required is to know whether or not ${}_a\tilde{u}(d^i)$ was greater or less than $\tilde{u}(d^i)$, or in other words, whether the adaptive utility was based on beliefs that turned out to be optimistic or pessimistic (this form of utility information will be employed in Example 5.3.2).

We refer to the case $z^i = 1$ as a negative surprise, for it represents a situation where the DM anticipated to achieve a greater utility amount than was actually experienced. Similarly, the case $z^i = 0$ will be referred to as positive surprise, for the DM anticipated a lower utility level than that which was actually noticed.

There are many ways in which information about utility can be gained following decision selection. It will of course depend on the particular problem under consideration, and examples from the literature on problems similar to the adaptive utility setting developed here are given below. However, once a form of information has been identified (and we appreciate that this can be a difficult task), a likelihood function can be devised detailing the probabilistic connections between d^i, z^i, r^i, ω and θ. Bayesian updating can then be performed following the observation of z^i from selection of d^i and outcome r^i, and relevant posterior probabilities concerning θ determined in the usual way (see Section 5.1).

Crawford & Shum [26] and Erdem & Keane [39] consider sequential decision problems where attribute values of possible returns can be uncertain. In the case of [26] the decision is which anti-ulcer drug prescription should be given, and the attributes of the rewards are the curative effects and the symptomatic effects of the drug chosen. In this case learning occurs through direct prescription experiences, leading to noisy signals of true values.

The work of Erdem & Keane was discussed in Section 3.3 and uses known data to model the decision making strategy of US customers. The decision relates to selection of a liquid detergent brand, and there is uncertainty over various attributes of those brands. In addition to direct experience of using a specific brand, learning also occurs through external advertisements the DM had been subjected to. Both these sources of information are assumed to provide noisy signals of correct values, and Erdem & Keane derive a functional form for a Bayesian learning framework.

A note should be made here that both [26] and [39] claim that utility is uncertain only because of uncertainty in some specific attribute levels. Indeed, [39, p.4] contains the quote "the utility to be derived from a product is not known with certainty at the instance of purchase due to consumer uncertainty about brand attributes". This is different to the type of uncertainty in utility being discussed in this thesis, where even if attributes of rewards are known, we permit the DM to remain uncertain of her preferences.

Finally, and as described in Section 3.3, Chajewska *et al.* [22] also consider a similar problem to the adaptive utility setting discussed here. However, whilst we consider the situation of a DM being uncertain of her own preferences, Chajewska *et al.* consider the problem an analyst faces when seeking to determine the utility function of a medical patient. The patient is considering some form of prenatal test for diagnosing the presence of a chromosomal abnormality, and her utility function is assumed to have a multi-attribute argument, with attributes consisting of fetus status, possibility for loss of fetus, knowledge of fetus status, and possibility for future successful pregnancy.

The analyst faces a sequential decision problem in which decisions relate to questions the analyst can put forward for determining the patient's preferences. Such questions are of the form of whether the patient would agree or disagree to taking part in some specified standard gamble (Section 1.1.1 discusses how utility values can be elicited in this manner), and a prior distribution about the patient's utility function is formed from information held about a population of similar patients. The feedback from such questions places constraints on the patient's true utility function and provides information regarding its form. The analyst decides which question to ask by seeking to maximise expected value of information (see Section 6.1) from the feedback, continuing in this process until no question would provide expected value above some specified threshold.

5.3 Adaptive Utility in Reliability Problems

The possible use of adaptive utility for decision making within the area of reliability theory was illustrated by Houlding & Coolen [56], and the material within this section draws heavily upon that work. The use of Bayesian statistical decision theory for solving system reliability problems has been applied, for example, in [23, 75, 78], and further references are available therein. However, to permit uncertainty over preferences through the use of an adaptive utility function is a novel approach which prevents unnecessary restriction to an assumed measure of preference.

Classical utility theory assumes known preferences over all possible outcomes. However, when considering vague or novel outcomes this is often an unreasonable assumption. In reliability theory, the cost of system failure, especially for a newly designed system, is a possible outcome that a DM may wish to avoid specifying a fixed constant for. For this reason we concentrate in this section on two examples in reliability theory where adaptive utility could be considered to be of use. As mentioned in Section 5.1, the solution algorithm for sequential adaptive utility problems is intractable, currently making the theory's use in interesting problems extremely computationally expensive. Hence we restrict attention to considering problems consisting of sequences of 3 decisions, with this duration being selected as it allows demonstration of the differences between classical and adaptive utility sequential decision making, but also only requires relatively easy computations.

Example 5.3.1 considers a system which has known failure modes, but where system failure leads to unknown damage for the manufacturer, *e.g.*, loss of customers, decrease in reputation, or financial cost for warranty *etc.* It is assumed that the manufacturer (DM) has opportunity to fix these failure modes for a known financial cost, and the decision must then be made whether or not to do so. The decision problem is thus to determine the optimal strategy of correcting system failures when damage through non-action is uncertain and with optimality defined by maximising expected adaptive utility. In this case the adaptive utility setting permits the DM to be initially uncertain of the trade-off between the loss in reputation and/or customers following system failure and the financial cost of fixing failure modes, with opportunity to learn about this trade-off parameter following trial runs or pilot schemes.

Example 5.3.1

Consider a newly designed system that has associated with it two known failure modes, types A and B, which occur independently of each other. First let the time between decision epochs be split into three periods of equal duration (*e.g.*, three days or three weeks). Assume that in any given period there is the possibility that one failure of type A and one failure of type B can occur. Hence, it is assumed that at most two failures can occur in each period (*i.e.*, at most one for each type of failure) and the number of failures per type within a decision epoch will be modelled through a Binomial distribution.

The true state of nature is $\omega = (\omega_A, \omega_B) \in \Omega = [0,1] \times [0,1]$, with ω_A and ω_B representing the unknown probabilities of a failure of type A or B, respectively, in any given period. Prior beliefs are elicited and are such that $\omega_A, \omega_B \in \{0.005, 0.1\}$ with $P(\omega_A = 0.005) = 0.7$ and $P(\omega_B = 0.005) = 0.2$.

Failure of the system incurs a cost or damage. However, the DM is assumed to be uncertain over how she will feel about the effects of such a cost, and she does not currently know whether it will be viewed as Severe (S) or Mild (M). In both cases the actual cost for a given failure type is known and fixed, it is only the DM's perceived attitude towards this cost that is uncertain. For example, each failure of the system may lead to the loss of a certain number of customers, or may lead to a certain financial penalty, but the effects that these situations will have are unknown.

Failures of type A or B are both modelled by an independent Bernoulli distribution. The DM's state of mind will be denoted as $\theta = (\theta_A, \theta_B) \in \Theta = [0,1] \times [0,1]$, with θ_A and θ_B respectively representing the true, but unknown, trade-off parameter concerning damage from system failure of type A or B and the known financial cost of permanently fixing them. Prior beliefs over θ stipulate that $\theta_A, \theta_B \in \{0.3, 0.7\}$, with distributions specified by $P(\theta_A = 0.7) = 0.8$ and $P(\theta_B = 0.7) = 0.3$.

The first decision occurs immediately and subsequent decisions occur after three periods each. At each decision epoch the DM can permanently fix one failure mode (assuming they have not already done so), and hence, depending on the decision history, the set of all available decisions in period i is given by $\mathcal{D}^i \subseteq \{d_A, d_B, d_N\}$. Here decisions d_A and d_B represent the permanent fix of failure modes A or B respectively, whilst decision d_N represents the null decision in which no permanent fix is undertaken. The set of feasible decision sequences consists of those decision streams d^1, d^2, d^3 from the set $\mathcal{D}^1 \times \mathcal{D}^2 \times \mathcal{D}^3$ where decisions d_A and d_B appear at most once (we assume the DM is not permitted to fix a failure type that has already been permanently fixed).

We do not seek to claim the form that a DM in this situation should take for her utility function, but to illustrate the theory, and for the purpose of this example, we will assume that the DM's utility, once categorised by θ, is the following function of feasible decision stream d^1, d^2, d^3 and beliefs over state of nature ω:

$$\begin{aligned} u(d^1, d^2, d^3|\theta) \; = \; & E_\omega\Big[-1000(\theta_A \omega_A \prod_{i=1}^{3}[1 - I_{\{d^i = d_A\}}] + \theta_B \omega_B \prod_{i=1}^{3}[1 - I_{\{d^i = d_B\}}]) \\ & -20(g_A + g_B)\Big] \end{aligned} \tag{5.7}$$

Here the terms g_A and g_B are used to represent:

$$g_A = \begin{cases} 0 & d^1, d^2, d^3 \neq d_A \\ i & \text{if } d_A \text{ selected at epoch } i \end{cases}, \quad g_B = \begin{cases} 0 & d^1, d^2, d^3 \neq d_B \\ i & \text{if } d_B \text{ selected at epoch } i \end{cases} \tag{5.8}$$

This utility function reflects a situation in which the DM prefers low failure rates and low probability of any failures being kind S. Also note that observed failures over the 3 epochs do not lead to any direct loss of utility.

Such a situation could arise when conducting laboratory tests before market launch, where failures in testing are assumed to be negligible, but where failures following launch lead to loss (either financial or otherwise) for the company. The utility function also penalizes the DM for making decisions to fix failures late in the decision stream. However, by later periods it is expected that more information over uncertain parameters will be available for the DM and so she will be better placed to make decisions that reflect her true preferences. The state of mind $\theta = (\theta_A, \theta_B)$ plays the role of either increasing or decreasing the importance of probability of system failure in contrast to the known cost of permanently fixing such failures. The values of constants involved make this problem non-trivial, *i.e.*, it is not immediately apparent that a specific strategy dominates another. Indeed, once the first decision has been chosen, all feasible decision streams are optimal for some collection of observations.

Information received between epochs consists of the number of failures that occurred over the three periods in that epoch, whether they were of type A or B, and also whether these failures were perceived to be of kind S or M. Such information is clearly relevant for updating beliefs over ω and θ, and the information received in epoch i will be represented by $z^i = \{z^i_{AS}, z^i_{AM}, z^i_{BS}, z^i_{BM}\}$. Here z^i_{jk} represents the number of failures observed in epoch i that were induced by failure type j and were perceived to be of severity level k.

This problem can be solved by the adaptive utility algorithm given in Section 5.1, and the relevant results are summarised below. Table 5.1 allows the DM to determine the optimal final epoch decision strategy π^3 given the relevant history up to that point. Unfortunately some of the expressions are too long to include and are instead briefly discussed in Appendix B. In Table 5.1, expected adaptive utility equations are represented by E's, and these are functions of the observations z^1 and z^2. L's represent lists of observed histories. Table 5.2 gives the optimal second

Ordered Decision History	d^3	Expected Utility	Max
d_A, d_B	d_N	-60	Always
d_A, d_N	d_B	-80	Otherwise
d_A, d_N	d_N	$E1$	If history in $L1$
d_B, d_A	d_N	-60	Always
d_B, d_N	d_A	-80	Otherwise
d_B, d_N	d_N	$E2$	If history in $L2$
d_N, d_A	d_B	-100	Otherwise
d_N, d_A	d_N	$E3$	If history in $L3$
d_N, d_B	d_A	-100	Otherwise
d_N, d_B	d_N	$E4$	If history in $L4$
d_N, d_N	d_A	$E5$	If history in $L5$
d_N, d_N	d_B	$E6$	If history in $L6$
d_N, d_N	d_N	$E7$	If history in $L7$

Table 5.1: Summary of π^3 for Example 5.3.1.

epoch strategy π^2 given knowledge of decision strategy π^3, and Table 5.3 gives the optimal initial decision given knowledge of future strategies π^2 and π^3.

The conclusion is that the optimal first epoch decision is $d^1 = d_B$, and further decisions should be selected as indicated by Tables 5.1 and 5.2. Decision d_B is the optimal first period decision primarily because, based upon prior beliefs, the DM perceives she will only solve one system failure type, so it is unlikely that both decisions d_A and d_B will be included within the decision stream, yet also her prior beliefs are such that it would be more beneficial if failure type B were fixed as $E_{\omega_A}[E_{\theta_A}[\omega_A\theta_A]] < E_{\omega_B}[E_{\theta_B}[\omega_B\theta_B]]$. It is unlikely that decisions d_A and d_B will both be selected in the sequence because, before making her first decision, the DM can determine which collection of observations would lead to her making the choice to fix both failure types. However, under the prior beliefs specified in this example, such decision histories have small associated probability of occurring.

Decision History	d^2	Max Expected Utility	Max
d_A	d_B	-60	Otherwise
d_A	d_N	$E8$	If history in $L8$
d_B	d_A	-60	Otherwise
d_B	d_N	$E9$	If history in $L9$
d_N	d_A	$E10$	If history in $L10$
d_N	d_B	$E11$	If history in $L11$
d_N	d_N	$E12$	If history in $L12$

Table 5.2: Summary of π^2 for Example 5.3.1.

d^1	Max Expected Utility	Max
d_A	-53.1 (1dp)	
d_B	-39.3 (1dp)	Yes
d_N	-52.6 (1dp)	

Table 5.3: Summary of π^1 for Example 5.3.1.

As an example of reading Tables 5.1-5.3, consider the history where no failures are observed in any period. Because the optimal first decision does not depend on observed histories, the DM selects $d^1 = d_B$, so it is only possible to observe histories in which a failure of type A can occur. In Table 5.2 we only need to consider rows with decision d_B as the decision history, and we check if the history with no failures of type A is included in list $L9$ or not. The list $L9$ contains just two possibilities, one where no failure of type A occurs (which has prior probability approximately 0.91) and one where 3 mild failures occur (which has prior probability approximately 2.7×10^{-5}). As the considered history is included in $L9$, the DM should select decision d_N in the second epoch.

For the final epoch the DM considers rows in Table 5.1 where the ordered history is d_B, d_N, and thus the relevant question is whether the observed failures of type A (between all three decision epochs) falls into the list $L2$. At this stage there are 100 possible combinations for recorded failures of type A, and 45 of these fall into list $L2$. However, only 10 of the possible 100 histories obtain no observed failures of type A between the first and second decision epochs, with 6 of these falling in $L2$. Given that no failures were observed before the second epoch, the updated prior probability (as viewed at the time of the second epoch) that no failures will be observed after the second epoch is approximately 0.92. Furthermore, the history with no failures of type A observed after the first or the second decision epochs does indeed fall in $L2$ (see Appendix B). Thus for our considered decision history we find that the DM should select decision d_N again.

Tables 5.1-5.3 are created without knowledge of earlier optimal decisions, and so include decision histories that become redundant when optimal strategies have been found. For example, the history in $L1$ which consists of no failures of type B between decision epochs is redundant once we know that the DM should select $d^1 = d_B$, hence failures of type B will not occur. If, for whatever reason, the DM did not make the optimal decision d_B in the first period, then knowledge of the optimal decision given such an alternative decision history could once again prove to be useful.

An interesting feature of this example is that, in certain situations, the observation of system failure can lead to an increase in expected adaptive utility (in the classical situation, such an observation can only lead to a decrease in expected utility due to an increase in the expected value of ω). This result arises because the utility function does not include a specific cost for observed failures, but rather only takes into account their associated probability of occurring and the probability that such failures are of type S. Although the observation of a system failure will tend to increase the expected value for either ω_A or ω_B, if such a failure were of type M, then it is likely that it would lead to a decrease of the expected value for one of θ_A or θ_B, and hence also possibly to a decrease of one of the products $\omega_A\theta_A$ or $\omega_B\theta_B$.

An intriguing note to make on this feature is that it can result in situations, such as the list $L9$ described above, whereby a failure type is only not rectified provided the DM observes no failures or observes a certain number of failures of type M (depending on the number of failures of type S also to be observed). Indeed, the suggestion is certainly true in this example for all the decision histories considered at the end of the first decision epoch, *i.e.*, lists $L8$, $L9$ and $L12$. This shows an interesting form of a monotonic relation that would appear in line with intuition, as it would appear reasonable to not exert cost to fix system failures if they were seen not to be a serious impediment upon the usefulness of the system, even if they were seen to occur frequently. $\diamond$

Example 5.3.1 sought to illustrate the potential for adaptive utility within reliability problems. In particular, we envision that it demonstrates potential for decision problems such as those a company could face when it has opportunity to test run a new system before commencement of full market launch. It could be that new software is tested by releasing Beta versions, and that feedback can be gained over how potential customers felt if an aspect of the software were to fail, *i.e.*, whether the software were still deemed useful with or without this aspect, and the implication this has on determining preferences over reliability and financial cost for fixing the failure. Another example could be a new machine that is useful to hospitals, with trial runs made in several hospitals to determine user feedback before launch across the entire health service. The use of adaptive utility in such settings allows the DM to remain uncertain and not to commit to a presumed trade-off parameter between finance and reputation.

Example 5.3.2 below also illustrates the use of adaptive utility within reliability decision problems, and considers a situation that is motivated by the hypothesis that individuals may tend to overweight the subjective cost of system failure compared to that which is the true effect.

Example 5.3.2

A DM has use of a machine that is in working order $100(1-\omega)\%$ of the time (we do not assume deterioration of the machine). The proportion of time that the machine is in working order thus represents the true state of nature, hence $\omega \in \Omega = [0,1]$. At each decision epoch the DM has the opportunity to permanently replace the possibly unreliable machine with one that is known to work perfectly (provided she has not already done so). The cost of making this exchange at decision epoch $i = 1, 2, 3$ is assumed to be $c - i/2$, *i.e.*, the cost of the new machine decreases over time. Depending on the decision history, the decisions available in epoch i are represented by $\mathcal{D}^i \subseteq \{d_S, d_R, d_D\}$, where d_S means stick with the possibly unreliable machine, d_R is to replace it, and d_D is a dummy decision used to construct feasible decision sequences. We thus assume that the set of feasible decision sequences are those in $\mathcal{D}^1 \times \mathcal{D}^2 \times \mathcal{D}^3$ where the decision d_R is included at most once and where decision d_D is selected if and only if d_R had been previously selected. The DM's utility function for this problem is assumed to be:

$$u(d^1, d^2, d^3|\theta) = \begin{cases} E_\omega[-3\omega^\theta] & \text{if } d^1, d^2, d^3 \neq d_R \\ E_\omega[-c + i/2 - (i-1)\omega^\theta] & \text{if } d_R \text{ selected first in epoch } i \end{cases} \tag{5.9}$$

The state of mind θ is interpreted as a parameter that alters how the DM views specific reliability levels ω. In this case we assume that $\theta \in \Theta = [1/2, 2]$ and that prior beliefs are represented by the triangular distribution with probability density function over this range $f(\tilde{\theta}) = 8(2 - \tilde{\theta})/9$, *i.e.*, under prior beliefs the DM assigns higher probability density to regions that overweight the subjective cost of reliability level ω than would be the case if true preferences were known to coincide with the linear rule with $\theta = 1$. This utility function is an additive function where the DM loses ω^{θ} in each period she uses the older machine. However, if in epoch i the DM switches to the new machine, she no longer looses any more utility but has to pay a one-off cost of $-c + i/2$. If $\omega = 0$ it does not matter what value θ takes and the DM should never switch (both machines would be just as good as each other).

To prevent the example becoming too complicated, it is assumed that ω is known and is such that $\omega = 1/2$. We also assume $c = 2$. Hence the only remaining uncertainty is with respect to beliefs over the correct state of mind θ. It is assumed that the DM can learn about this parameter through utility information z^i that is received in epoch i. Noting that, for feasible decision streams, the parameterised utility function can be decomposed into the equivalent expression $\sum_{i=1}^{3} \tilde{u}(d^i|\theta)$, with $\tilde{u}(d^i|\theta)$ given below, the DM is able to evaluate an expected value for $\tilde{u}(d^i|\theta)$ with respect to beliefs she holds over θ at the beginning of epoch i:

$$\tilde{u}(d^i|\theta) = \begin{cases} E_{\omega}[-\omega^{\theta}] & \text{if } d^i = d_S \\ -c + i/2 & \text{if } d^i = d_R \\ 0 & \text{if } d^i = d_D \end{cases} \tag{5.10}$$

To give the form of information z^i, it is assumed that, following selection of decision d^i, the DM can compare her prior expected value for $\tilde{u}(d^i|\theta)$ with the value it was actually noted to have. However, it is not assumed that the precise difference can be stated, only that the DM can determine whether or not her expectations were too pessimistic or too optimistic, and this will form categorical data where $z^i = 0$

in the first case and $z^i = 1$ in the latter (Section 5.2 discussed this form of utility information).

Noting that $\tilde{u}(d^i|\theta)$ is known for certain when either of decisions d_R or d_D are selected, utility information z^i is only informative over θ if decision d_S can still be selected in a future epoch. The probability of observing $z^i = 0$, given θ, in this case is equal to:

$$\begin{aligned} P\Big(z^i = 0\Big|\theta = \tilde{\theta}, d_S\Big) &= P\Big(E_\theta[\tilde{u}(d_S|\theta)] < \tilde{u}(d_S|\theta)\Big|\theta = \tilde{\theta}\Big) \\ &= P\Big(E_\theta[-(1/2)^\theta] < -(1/2)^\theta\Big|\theta = \tilde{\theta}\Big) \\ &= P\Big(\theta > \frac{\log(E_\theta[(1/2)^\theta])}{\log(1/2)}\Big|\theta = \tilde{\theta}\Big) \\ &= \begin{cases} 1 & \text{if } \tilde{\theta} > \frac{\log(E_\theta[(1/2)^\theta])}{\log(1/2)} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{5.11}$$

This problem is again solved via the sequential adaptive utility algorithm given in Section 5.1, and summary results are now discussed. Table 5.4 determines π^3 given the decision history ${}_aH^2$ by the time of the final decision. Note that the observation z^2 does not have any effect on this decision, because if $d^1 = d_S$ then z^1 provides enough information about the utility function for the DM to know which decision she should select, and if $d^1 = d_R$ then all future decisions must be the dummy decision d_D. For the second epoch, knowing decision strategy π^3, optimal strategy π^2 can be found through Table 5.5. Finally, for the first epoch, knowing decision strategies π^2 and π^3, the optimal first period decision π^1 is determined through Table 5.6.

The optimal strategy is thus to stick with the unreliable machine before possibly replacing it depending on whether expected utility for sticking was found to be too optimistic or too pessimistic. Note that in the classical situation, where it is assumed that preferences are characterised by the expectation of θ under prior beliefs, and where such a preference structure is enforced throughout the problem (*i.e.*, no utility information is available), the DM would do best to replace the machine immediately.

This is because under expected prior beliefs $E_\theta[\tilde{u}(d_S|\theta)] < -0.5$. So to never replace, or to replace at the second or third opportunity, leads to expected utility return less than -1.5. However, to change immediately leads to utility return of -1.5 exactly. Thus, due to the extra information that such a decision offers, incorporating the possibility that the DM's current expectations over preferences may be wrong leads to the more conservative decision strategy (in the sense that an irrevocable decision is not immediately selected) of selecting decision d_S in the first epoch. ◇

Ordered Decision History	d^3	Max Expected Utility	Max
d_S, d_S	d_S	$E_{\theta\|_a H^2}[u(d_S, d_S, d_S\|\theta)]$	If $z^1 = 0$
d_S, d_S	d_R	$E_{\theta\|_a H^2}[u(d_S, d_S, d_R\|\theta)]$	If $z^1 = 1$
d_S, d_R	d_D	$E_{\theta\|_a H^2}[u(d_S, d_R, d_D\|\theta)]$	Yes
d_R, d_D	d_D	$E_{\theta\|_a H^2}[u(d_R, d_D, d_D\|\theta)]$	Yes

Table 5.4: Summary of π^3 for Example 5.3.2.

Ordered Decision History	d^2	Max Expected Utility	Max
d_S	d_S	$E_{_aH^2\|_aH^1}[E_{\theta\|_aH^1}[u(d_S, d_S, \pi^3(_aH^2)\|\theta)]]$	If $z^1 = 0$
d_S	d_R	$E_{_aH^2\|_aH^1}[E_{\theta\|_aH^1}[u(d_S, d_R, d_D\|\theta)]]$	If $z^1 = 1$
d_R	d_D	$E_{_aH^2\|_aH^1}[E_{\theta\|_aH^1}[u(d_R, d_D, d_D\|\theta)]]$	Yes

Table 5.5: Summary of π^2 for Example 5.3.2.

d^1	Max Expected Utility	Max
d_S	-1.42	Yes
d_R	-1.5	

Table 5.6: Summary of π^1 for Example 5.3.2.

The sequential problem discussed in Example 5.3.2 is similar to the maintenance problems that are discussed by Baker [9]. Risk aversion (a concept to be discussed in Section 6.2) relates to a DM's dislike of entering into actuarially fair gambles, *i.e.*, gambles with expected return 0. The greater the level of a DM's risk aversion, the greater the amount of utility she would be prepared to forgo in order to avoid entering such gambles. Baker considers how a general trend of risk aversion in DMs and organisations can lead to overmaintenance of systems, and considers implications in principle-agent theory, where the principle relates to an organisation (not wanting overmaintenance), and where the agent relates to a maintenance engineer (who overmaintains due to excessive risk aversion). Baker finds that incentives based on the total cost of maintenance and failures can reduce over maintenance, and that it may be optimal for management to pay such an incentive.

If the utility function of Example 5.3.2 had been expressed as a function with the reliability measure ω as argument, then we could represent one-period preferences through $\tilde{u}(\omega|\theta) = -\omega^{\theta}$. In this case θ affects the level of risk aversion, with a larger value for θ representing a reduced level of risk aversion[2]. Thus the implication of using adaptive utility in this setting, is that the DM can now remain uncertain over whether or not she is undermaintaining or overmaintaining the system (through either replacing the unreliable machine too late or too early, respectively). Beliefs over state of mind θ can in this sense be seen as beliefs over the DM's risk aversion, and through the connection identified by Baker, beliefs over supposed overmaintenance. Additionally, in a general principle-agent theory problem, adaptive utility could be of use for permitting a situation in which the principle is initially uncertain of the utility function of the agent and *vice versa*.

[2] For this utility function the Arrow-Pratt measure of absolute risk aversion, which is introduced in Chapter 6, is given by $\frac{1-\theta}{\omega}$.

Chapter 6

Adaptive Utility Diagnostics

This chapter considers the implication of uncertain preferences for two classical utility diagnostics. First we consider the relationship between value of sample information and utility uncertainty. Following this we consider implications of uncertain preferences upon classical risk aversion, and we introduce the concept of trial aversion.

6.1 Value of Information

The classical concept of value of sample information, that arises as a utility diagnostic in the type of decision problem under consideration here, is discussed by DeGroot [35]. Further discussion can be found in [16, 33, 84].

Let observable random quantity X represent a currently unknown piece of information. In discussing the expected amount of information in X, or the expected value of X, we are referring to its fair utility value, *i.e.*, the maximum amount of utility the DM would forgo in order to know X. In a decision problem, the expected value of X is the expected difference (with respect to beliefs over X) between the maximum expected utility obtainable through decision selection with knowledge of X, and the maximum expected utility obtainable without knowledge of X[1].

[1] See [48, p.529] for common alternative meanings not considered here.

The set of possible information statements for X will be denoted by $\mathcal{X}$, and a particular statement by x. Under this notation, problems of interest in the classical treatment of value of sample information occur when X can inform the DM about the likely outcome of an available decision. That is to say, there exists a possible combination of return r, decision d, and information statement x such that $P(r|d, X = x) \neq P(r|d)$. The value of such information x is denoted by $I_\omega(x)$, and the value of unknown information X is denoted by $I_\omega(X)$ (the subscript ω is used to show what the information relates to). Note that as X is a random quantity, so is $I_\omega(X)$.

In [35, 84], the expected value of information X, $E_X[I_\omega(X)]$, is defined by:

$$E_X[I_\omega(X)] = E_X[\max_{d \in \mathcal{D}}\{E_{\omega|X}[u(d)]\}] - \max_{d \in \mathcal{D}}\{E_\omega[u(d)]\} \tag{6.1}$$

A particular consequence of Definition (6.1) is that $E_X[I_\omega(X)] \geq 0$.

Although in some situations a DM may wish not to observe $X = x$ (*e.g.*, discovering $X = x$ caused other knowledge to be forgotten or had a specific utility attachment itself due to emotional effects), DeGroot [35] argues that, for standard statistical decision problems, $I_\omega(x) \geq 0$ for all $x \in \mathcal{X}$. As DeGroot notes, once $X = x$ is known, the use of determining the expected utility of a decision with respect to beliefs over ω as represented by distribution P_ω is only relevant in so far as it permits calculation of what the DM believed was the optimal decision before knowledge of x. Actual beliefs are now represented by $P_{\omega|x}$, and all expected utilities of decisions should be evaluated through use of this distribution. Hence both Raiffa & Schlaifer [84] and DeGroot [35] define $I_\omega(x)$ as follows:

$$I_\omega(x) = \max_{d \in \mathcal{D}}\{E_{\omega|x}[u(d)]\} - E_{\omega|x}[u(d')], \text{ with } d' = \arg\max_{d \in \mathcal{D}}\{E_\omega[u(d)]\} \tag{6.2}$$

.

This definition implies that $I_\omega(x) \geq 0$ and leads directly to Equation (6.1). However, to ensure $I_\omega(x)$ is well-defined, we make the additional condition that d' is the decision the DM would have selected before she knew $X = x$.

As a heuristic justification for this, note that, whilst $E_X[I_\omega(X)]$ is an *ex ante* statistic, $I_\omega(x)$ is an *ex post* diagnostic, the calculation of which requires the DM to have considered her decision problem without the knowledge that $X = x$. Thus the DM should have been able to determine decision d' she would have selected. It is important to do this because, if $d' = \arg\max_{d \in \mathcal{D}}\{E_\omega[u(d)]\}$ is not unique, then whilst alternatives do not affect expected utility before $X = x$ is known, they can affect the value of $I_\omega(x)$ after $X = x$ is known. However, before $X = x$ is known, the only relevant quantity is $E_X[I_\omega(X)]$, which does not depend on selection of such d'.

Before moving to an adaptive utility setting, we note that $E_X[I_\omega(X)]$ and $I_\omega(x)$ were referred to by Raiffa & Schlaifer [84, Ch.4] as the Expected Value of Sample Information (EVSI) and the Conditional Value of Sample Information (CVSI), respectively. We also note that EVSI depends only on the following:

- The set of feasible decisions $\mathcal{D}$.
- The utility function u representing preferences.
- The distribution P_ω representing prior uncertainty over ω.
- The likelihood function $P_{X|\omega}$ that is combined with P_ω to produce posterior $P_{\omega|X}$ or predictive P_X.

The introduction of uncertain preferences, however, leads to interesting questions regarding the classical treatment of value of information. Equations (6.1) and (6.2) must be generalised to deal with adaptive utilities, but also, value of information concerning ω may be affected by the level of uncertainty the DM has about her preferences. Furthermore, there is a value for information z relating to the DM's true preferences, and this should also be quantifiable.

To tackle such problems, we first assume that θ and ω are independent, and we begin with the case of X being only relevant for determining ω. Under these assumptions we propose that the EVSI and the CVSI are as defined in Equations (6.3) and (6.4), respectively (the notation I_ω is replaced by ${}_aI_\omega$ to highlight that we are considering value of information when preferences are uncertain):

$$E_X[{}_aI_\omega(X)] = E_X[\max_{d\in\mathcal{D}}\{E_{\omega|X}[{}_au(d)]\}] - \max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\} \tag{6.3}$$

$${}_aI_\omega(x) = \max_{d\in\mathcal{D}}\{E_{\omega|x}[{}_au(d)]\} - E_{\omega|x}[{}_au(d')], \text{ with } d' = \arg\max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\} \tag{6.4}$$

Equations (6.3) and (6.4) are analogous to Equations (6.1) and (6.2), respectively, the only difference is that $u(d)$ is replaced by the adaptive utility function ${}_au(d)$. They can be justified in precisely the same manner as Equations (6.1) and (6.2), and again the EVSI, as given by Equation (6.3), follows from the formula for the CVSI, as given by Equation (6.4). Equations (6.3) and (6.4) generalise Equations (6.1) and (6.2), returning to them in the case that preferences are known with certainty. Furthermore, both ${}_aI_\omega(x)$ and $E_X[{}_aI_\omega(X)]$ are non-negative. However, the important point is that now, in addition to those components listed above, the EVSI and CVSI, for information relating to ω, also depend on beliefs over state of mind θ (through the effect this has in determining the adaptive utility function ${}_au$).

In the adaptive utility setting, the EVSI keeps the property of additivity, whereby the expected value of receiving two pieces of information X_1 and X_2 simultaneously equals the expected value of receiving them one after the other (provided that no decision was made in the interim). That is to say, the following property is true (where, when $X_1 = x_1$ is known and beliefs have been updated based solely on this, the expected additional value of X_2 is denoted by $E_{X_2|x_1}[{}_aI_\omega(X_2|x_1)]$):

$$E_{X_1,X_2}[{}_aI_\omega(X_1,X_2)] = E_{X_1}\left[E_{X_2|X_1}[{}_aI_\omega(X_2|X_1)] + {}_aI_\omega(X_1)\right] \tag{6.5}$$
$$= E_{X_1,X_2}[{}_aI_\omega(X_2|X_1)] + E_{X_1}[{}_aI_\omega(X_1)]$$

This result is derived in the same way as is done in the classical situation:

$$\begin{aligned} E_{X_1,X_2}[{}_aI_\omega(X_1,X_2)] &= E_{X_1,X_2}[\max_{d\in\mathcal{D}}\{E_{\omega|X_1,X_2}[{}_au(d)]\}] - \max_{d\in\mathcal{D}} E_\omega[{}_au(d)] \qquad (6.6)\\ &= E_{X_1,X_2}[\max_{d\in\mathcal{D}}\{E_{\omega|X_1,X_2}[{}_au(d)]\}] - E_{X_1}[\max_{d\in\mathcal{D}}\{E_{\omega|X_1}[{}_au(d)]\}]\\ &\quad + E_{X_1}[\max_{d\in\mathcal{D}}\{E_{\omega|X_1}[{}_au(d)]\}] - \max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\}\\ &= E_{X_1}\left[E_{X_2|X_1}[\max_{d\in\mathcal{D}}\{E_{\omega|X_1,X_2}[{}_au(d)]\}] - \max_{d\in\mathcal{D}}\{E_{\omega|X_1}[{}_au(d)]\}\right]\\ &\quad + E_{X_1}[{}_aI_\omega(X_1)]\\ &= E_{X_1}[E_{X_2|X_1}[{}_aI_\omega(X_2|X_1)]] + E_{X_1}[{}_aI_\omega(X_1)]\\ &= E_{X_1,X_2}[{}_aI_\omega(X_2|X_1)] + E_{X_1}[{}_aI_\omega(X_1)] \end{aligned}$$

Example 6.1.1 below demonstrates how the EVSI can depend on the DM's beliefs over her state of mind θ. In particular, this example leads to the possible suggestion that a decrease in uncertainty over θ necessarily leads to an increase in EVSI. This suggestion will be further discussed following the example.

Example 6.1.1

Consider a one-period problem with $\mathcal{D} = \{d_A, d_B\}$ and pay off matrix as below:

	ω_1	ω_2
d_A	r_1	r_2
d_B	r_2	r_1

Prior beliefs are such that $P(\omega_1) = P(\omega_2) = 0.5$, and the DM's likelihood function is such that information X will fully inform her of ω, *i.e.*, $P_{X|\omega}(x_j|\omega_i) = \delta_{ij}$ for $j = 1,2$ (with δ_{ij} representing the usual Kronecker delta). Thus the predictive distribution is $P(x_1) = P(x_2) = 0.5$.

Beliefs over preferences are assumed to be such that $P(\theta_1) = p = 1 - P(\theta_2)$, with $p \in [0, 1]$. Commensurable utilities are assumed to be $u(r_1|\theta_1) = u(r_2|\theta_2) = 1$ and $u(r_2|\theta_1) = u(r_1|\theta_2) = 0$ (in what follows $u(r|\theta)$ and $u(d|\theta)$ are interchanged as discussed in Subsection 1.1.1). In this case the EVSI is maximised when $p \in \{0, 1\}$, *i.e.*, when the DM knows her preferences with certainty, and is minimised when $p = 0.5$, *i.e.*, when the DM believes either state of mind is equally likely. To see this note that $\max_{d\in\mathcal{D}}\{E_\omega[_a u(d)]\} = 0.5$ regardless of p:

$$\begin{aligned} \max_{d\in\mathcal{D}}\{E_\omega[_a u(d)]\} &= \max_{d\in\mathcal{D}}\{pE_\omega[u(d|\theta_1)] + (1-p)E_\omega[u(d|\theta_2)]\} \qquad (6.7) \\ &= \max\{0.5p + 0.5(1-p), 0.5p + 0.5(1-p)\} \\ &= 0.5 \end{aligned}$$

The EVSI is thus maximised when p is such that $E_X[\max_{d\in\mathcal{D}}\{E_{\omega|X}[_a u(d)]\}]$ is maximal:

$$\begin{aligned} E_X[\max_{d\in\mathcal{D}}\{E_{\omega|X}[_a u(d)]\}] &= 0.5[\max_{d\in\mathcal{D}}\{E_{\omega|x_1}[_a u(d)]\} + \max_{d\in\mathcal{D}}\{E_{\omega|x_2}[_a u(d)]\}] \qquad (6.8) \\ &= 0.5[\max\{p, 1-p\} + \max\{1-p, p\}] \\ &= \max\{p, 1-p\} \end{aligned}$$

Hence $E_X[\max_{d\in\mathcal{D}}\{E_{\omega|X}[_a u(d)]\}]$ is maximised when $p \in \{0, 1\}$, and is minimised when $p = 0.5$.

Representing θ_1 and θ_2 via numerical values permits calculation of the variance of θ, $V[\theta]$, as a function of p. If we assume that $\theta_1 = 1$ and $\theta_2 = 0$, then Figure 6.1 presents a standard plot (left hand side) and a parametric plot (right hand side) of $E_X[_a I_\omega(X)]$ and $V[\theta]$ over the range $p \in [0, 1]$. Both plots demonstrate that $E_X[_a I_\omega(X)]$ decreases as $V[\theta]$ increases. ◇

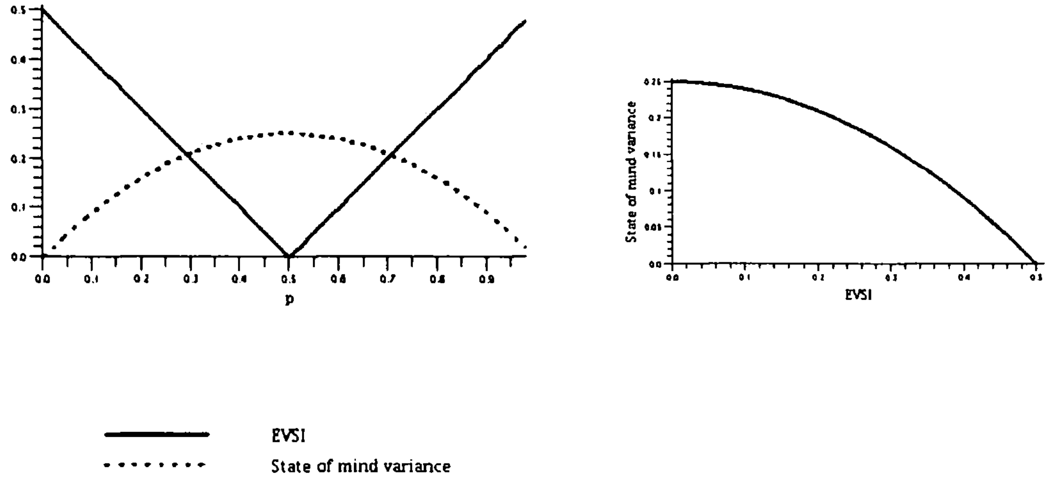

Figure 6.1: EVSI and state of mind variance $V[\theta]$ in Example 6.1.1 for $p \in [0, 1]$

As mentioned previously, a possible suggestion arising from Example 6.1.1, is that a decrease in uncertainty over θ necessarily leads to an increase in EVSI for information relating to the outcome of a decision. The argument could be that, if the DM is uncertain about what she prefers, then she will not know which type of reward to aim for in decision selection, and thus information concerning the likely return of any decision is of little use. Nevertheless, although true in Example 6.1.1, it is not generally true that EVSI concerning the outcome of a decision can only increase if uncertainty concerning preferences decreases. This will be demonstrated in Example 6.1.2.

There are several potential methods for measuring uncertainty. In the classical problem with known utility, Gould [51] considers measuring uncertainty by, for example, variance, the Shannon measure of entropy [96], and the Rothschild & Stiglitz measure of spread in distributions with equal mean [88] (under which Y_1 is deemed more uncertain than Y_2 if $Y_1 = Y_2 + \epsilon$, where, conditional upon Y_2, ϵ is uncorrelated random noise with mean 0 and positive variance). Gould demonstrates that in the classical situation, with finite Ω, EVSI does not necessarily increase when:

- The number of elements in Ω having non-zero probability increases.
- Less probability is concentrated on any single element of Ω.
- The variance of ω increases.
- Uncertainty over ω in the Rothschild & Stiglitz sense increases.

Seeking a similar result to that found by Gould, Example 6.1.2 below provides a counter example to the possible suggestion that, all else constant, a decrease in uncertainty over state of mind θ necessarily leads to an increase in EVSI.

Example 6.1.2

Consider again the setting of Example 6.1.1, but where commensurable utilities are now such that $u(r_1|\theta_1) = u(r_2|\theta_1) = 1$, $u(r_1|\theta_2) = 0$, and $u(r_2|\theta_2) = 2$.

In this setting, $E_X[{}_aI_\omega(X)]$ monotonically increases as p decreases. To see this, note that $\max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\}$ is again independent of p:

$$\begin{aligned}\max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\} &= \max_{d\in\mathcal{D}}\{pE_\omega[u(d|\theta_1)] + (1-p)E_\omega[u(d|\theta_2)]\} \qquad (6.9)\\ &= \max\{p+(1-p), p+(1-p)\}\\ &= 1\end{aligned}$$

Hence p influences EVSI only through the term $E_X[\max_{d\in\mathcal{D}}\{E_{\omega|X}[{}_au(d)]\}$:

$$\begin{aligned}E_X[\max_{d\in\mathcal{D}}\{E_{\omega|X}[{}_au(d)]\}] &= 0.5\max_{d\in\mathcal{D}}\{E_{\omega|x_1}[{}_au(d)]\} + 0.5\max_{d\in\mathcal{D}}\{E_{\omega|x_2}[{}_au(d)]\} \quad (6.10)\\ &= 0.5\max\{p, p+2(1-p)\} + 0.5\max\{p+2(1-p), p\}\\ &= \max\{p, p+2(1-p)\}\\ &= 2-p\end{aligned}$$

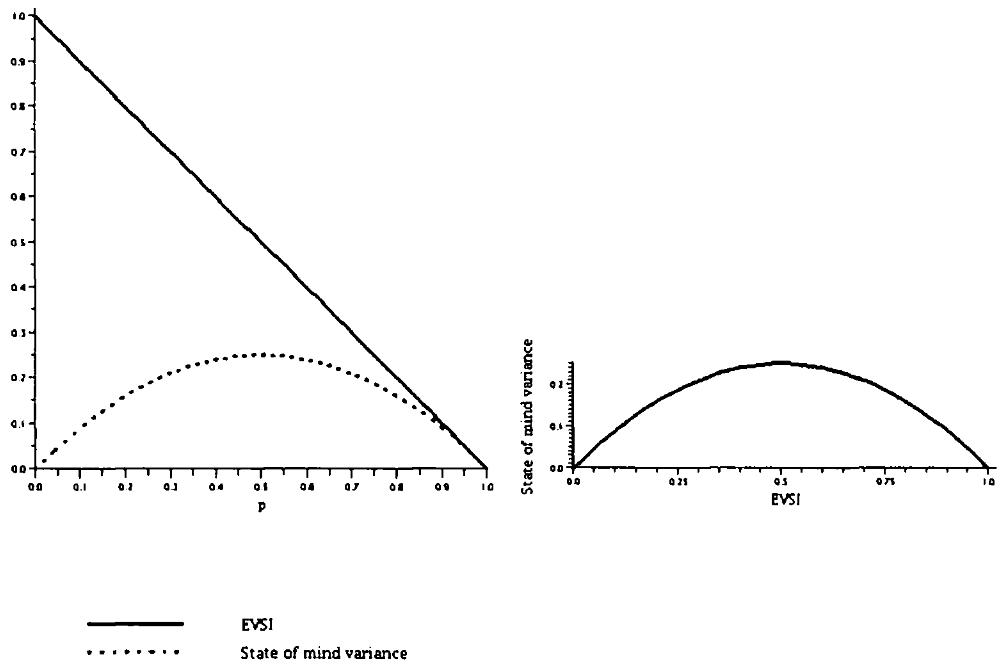

Figure 6.2: EVSI and state of mind variance $V[\theta]$ in Example 6.1.2 for $p \in [0, 1]$

Assigning $\theta_1 = 1$ and $\theta_2 = 0$ again allows for the relationship between EVSI and $V[\theta]$ to be plotted, and Figure 6.2 contains a standard plot (left hand side) and a parametric plot (right hand side) of $E_X[{}_aI_\omega(X)]$ and $V[\theta]$ for $p \in [0, 1]$. ◇

Example 6.1.2 thus demonstrates that:

- EVSI does not necessarily decrease as more elements of Θ have non-zero probability.
- EVSI does not necessarily decrease as greater probability is assigned to any single element of Θ.
- EVSI is not necessarily minimal when elements of Θ are equally probable.

Example 6.1.2 also demonstrates that EVSI is not necessarily increased when $V[\theta]$ is decreased. Furthermore, this result is invariant to the numerical representation assigned to θ_1 and θ_2.

To see this, note that the numerical assignment only affects $V[\theta]$, and not the EVSI. Assigning the generic representation of $\theta_1 = v_1$ and $\theta_2 = v_2$, where $v_1, v_2 \in \mathbb{R}$ and $v_1 \neq v_2$, the variance of θ is $V[\theta] = p(1-p)(v_1 - v_2)^2$. Hence, as a function of p, $V[\theta]$ is maximised when $p = 0.5$ and monotonically decreases as p moves away from 0.5 towards either 0 or 1. The EVSI, however, is monotonically decreasing in p. Thus, regardless of the assignment v_1 and v_2, as p increases from 0.5 towards 1, both EVSI and $V[\theta]$ decrease.

In the classical setting, Gould [51] claims that no simple relation exists between EVSI and uncertainty over ω. He notes that, with all other factors of the problem constant, a change in uncertainty affects both the maximum expected utility considering X, $E_X[\max_{d \in \mathcal{D}}\{E_{\omega|X}[u(d)]\}]$, and the maximum expected utility not considering X, $\max_{d \in \mathcal{D}}\{E_\omega[u(d)]\}$, the difference of which forms the EVSI, as given in Equation (6.1). This argument, suitably generalised, also explains why an increase in uncertainty over θ does not necessarily lead to a decrease in EVSI. A change in beliefs over θ can affect both the value of the maximum expected adaptive utility considering X, $E_X[\max_{d \in \mathcal{D}}\{E_{\omega|X}[{}_a u(d)]\}]$, and the maximum expected adaptive utility not considering X, $\max_{d \in \mathcal{D}}\{E_\omega[{}_a u(d)]\}$, the difference of which gives the EVSI for the adaptivity utility case as specified in Equation (6.3).

For an increase in uncertainty over θ to lead to a decrease in EVSI, it must be that the former is decreased (increased) by more (less) than the latter. In Examples 6.1.1 and 6.1.2, $\max_{d \in \mathcal{D}}\{E_\omega[{}_a u(d)]\}$ did not depend on the value of p. However, whilst in Example 6.1.1 $E_X[\max_{d \in \mathcal{D}}\{E_{\omega|X}[{}_a u(d)]\}]$ decreased as $V[\theta]$ increased, the opposite happened in Example 6.1.2 for $p \in [0.5, 1]$. Essentially, this alternative result occurred because increasing p away from 0.5 in Example 6.1.2 meant that it was more likely that both outcomes would be equally desirable. However, the problem description implies that both outcomes are equally likely under either decision, and hence increasing p towards 1 leads to a reduction in value for information as the decision that is actually selected becomes increasingly irrelevant.

Information X need not only be informative of state of nature ω, and alternatives include the case that X is only informative of θ, or the case that X is a vector of independent values (X_1, X_2), with X_1 being informative of ω only and X_2 being informative of θ only. In these situations an analogous argument to that given previously will demonstrate that the EVSI is (where notation is such that if X is relevant to θ, then it is included as a conditioning argument in the adaptive utility function) $E_X[{}_aI_\theta] = E_X[\max_{d\in\mathcal{D}}\{E_\omega[{}_au(d|X)]\}] - \max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\}$, and $E_X[{}_aI_{\omega,\theta}(X_1, X_2)] = E_{X_1,X_2}[\max_{d\in\mathcal{D}}\{E_{\omega|X_1}[{}_au(d|X_2)]\}] - \max_{d\in\mathcal{D}}\{E_\omega[{}_au(d)]\}$, respectively.

As mentioned in Chapter 5, interesting problems for adaptive utility are necessarily of a sequential nature, hence we now consider an n-period sequential problem under the assumption that prior beliefs over θ and ω are independent. In this situation we demonstrate that the value of observed information X, relevant for determining ω, decreases if it is observed later in the sequence of decisions. Using notation ${}_aI_\omega(x; j)$, with j denoting the period $X = x$ is to be observed, the CVSI is as indicated below:

$$\begin{aligned} {}_aI_\omega(x; j-1) &= \max_{d_j,\dots,d_n\in\mathcal{D}}\{E_{\omega|x}[{}_au(d'_1,\dots,d'_{j-1},d_j,\dots,d_n)]\} \\ &\quad - E_{\omega|x}[{}_au(d'_1,\dots,d'_n)] \text{ (for } j \geq 2) \end{aligned} \tag{6.11}$$

$${}_aI_\omega(x; 0) = \max_{d_1,\dots,d_n\in\mathcal{D}}\{E_{\omega|x}[{}_au(d_1,\dots,d_n)]\} - E_{\omega|x}[{}_au(d'_1,\dots,d'_n)]$$

$$\text{with } d'_1,\dots,d'_n = \arg\max_{d_1,\dots,d_n\in\mathcal{D}}\{E_\omega[{}_au(d_1,\dots,d_n)]\}$$

Because increasing j in Equation (6.11) means that the necessary maximisation must be performed under an extended number of constraints (decisions $d_1,\dots,d_{j-1}$ will have been fixed before $X = x$ is known), we find that ${}_aI_\omega(x; j)$ can only decrease as j increases.

This implies that the value of specific information can only diminish if it is observed later in the decision sequence. Furthermore, because of its connection to the CVSI, the EVSI for unknown information X can also only decrease if X is to become known later in the decision sequence. Such a result is indeed in line with intuition, where we expect information to be of greater use, and hence of larger value, if it is available for an increased number of decisions.

In the setting of Chapter 5, information was made available via two sources. First, the DM was able to learn about the true state of nature ω by observing the particular return r^i that decision d^i led to. Second, the DM was able to learn about her true state of mind θ by observing utility information z^i, also following decision d^i. By the time the DM selects decision d^i, beliefs will thus be represented by distributions $P_{\omega|_a H^i}$ and $P_{\theta|_a H^i}$. In this case the EVSI for $X^i = (r^i, z^i)$ is as given below (where ${}_aU^i = {}_au(d^1, \ldots, d^{i-1}, \pi^i({}_aH^i), \ldots, \pi^n({}_aH^n)|_aH^i))$:

$$E_{X^i}[{}_aI_{\omega,\theta}(X^i; i|_aH^i)] = E_{X^i}E_{\omega,\theta|_aH^i}[G] - E_{\omega,\theta|_aH^i}[G'] \tag{6.12}$$

$$\begin{aligned} G &= E_{{}_aH_{i+1}|X^i,\pi^i({}_aH^i),\omega,\theta} \cdots E_{{}_aH^n|\pi^{n-1}({}_aH^{n-1}),\omega,\theta}[{}_aU^i] \\ G' &= G\text{, but with } X^i \text{ omitted and with } r^i \text{ and } z^i \text{ removed from all histories.} \end{aligned}$$

Equation (6.12) shows that the EVSI of X^i will depend on the decision d^i that is chosen in period i. The previous decisions $d^1, \ldots, d^{i-1}$ will also influence the EVSI.

Example 6.1.3

To demonstrate the use of Equation (6.12), we return to the apple or banana example that was last discussed in Example 5.1.1. We assume a sequence length of $n = 3$, and that prior beliefs are $P(\theta = 1.5) = 0.4 = 1 - P(\theta = 0.5)$, so $E[\theta] = 0.9$. We further assume that selection of a banana leads to the correct value of θ being observed with probability 0.7, whilst nothing is learnt if the selection of an apple is made.

To determine the EVSI for the information concerning θ that is gained if the DM's first decision is selection of a banana, we must consider the possible future optimal decision streams that would arise depending on the actual information received. If an apple is initially selected no information will be observed, which clearly has value 0. If the banana is selected in the first period, then either utility information z^i is such that $z^1 = 1.5$ or $z^1 = 0.5$, with prior predictive probabilities determined via $P(z^1 = y) = P(z^1 = y|\theta = 1.5)P(\theta = 1.5) + P(z^1 = y|\theta = 0.5)P(\theta = 0.5)$. Hence we find that $P(z^1 = 1.5) = 0.46 = 1 - P(z^1 = 0.5)$. The maximum expected adaptive utility for the three cases of $z^1 = 1.5$, $z^1 = 0.5$, and for not receiving information z^1 (though a banana is still selected), are 3.37, 2.72, and 2.92, respectively. Putting these values into Equation (6.12) results in an EVSI of 0.1005 (to 4 s.f.). $\diamond$

A remark about Example 6.1.3 is that the expected adaptive utility for the decision to choose a banana in the first period is equal to EVSI+$E[\theta]$ = 1.0005. Indeed, when it is meaningful to talk about expected adaptive utility of a single decision within a sequence of decisions (*i.e.*, when adaptive utility is of an additive form), we can decompose the 'full' expected adaptive utility of any decision into the following two components:

- The 'pure' expected adaptive utility arising from receiving the return itself.
- The expected adaptive utility of the information that is gained regarding true preferences.

This can be seen via Equation (6.12), where $E_{X^i}[{}_aI_{\omega,\theta}(X^i; i|_aH^i)]$ represents the expected adaptive utility of the information that is gained regarding true preferences and likely outcomes of decisions, whilst $E_{\omega,\theta|_aH^i}[G']$ represents the pure expected adaptive utility arising from just the return itself (and no information is recorded). The final term in Equation (6.12), $E_{X^i}E_{\omega,\theta|_aH^i}[G]$, represents the full adaptive utility of selecting a particular decision. Suitable rearrangement then shows:

$$E_{X^i} E_{\omega,\theta|_a H^i}[G] = E_{X^i}[{}_a I_{\omega,\theta}(X^i; i|_a H^i)] + E_{\omega,\theta|_a H^i}[G'] \tag{6.13}$$

Also note that the EVSI is a natural method for measuring the usefulness of information. If we had assumed in Example 6.1.3 that selecting a banana would certainly inform the DM of her true state of mind, which is clearly a more useful source of information, then we would have found that the EVSI was equal to 0.3, a greater value than the 0.1005 reported above.

We conclude this section by considering and illustrating implications of EVSI for uncertain preferences in a reliability example. For this reason we return to the setting of Example 5.3.2.

Example 6.1.4

Example 5.3.2 considered a 3-period sequential decision problem where, at each decision epoch, the DM had to decide whether she wished to remain with an unreliable machine (assuming it had not previously been replaced), decision d_S, or whether she wished to change to a new and fully reliable machine, decision d_R. A dummy decision d_D was used to create feasible decision streams and was selected if and only if the DM had previously selected decision d_R.

In that example it was assumed that the DM's true preferences for decision streams were represented by the additive utility function $u(d^1, d^2, d^3|\theta) = \sum_{i=1}^{3} \tilde{u}(d^i|\theta)$, with:

$$\tilde{u}(d^i|\theta) = \begin{cases} -(1/2)^\theta & \text{if } d^i = d_S \\ i/2 - 2 & \text{if } d^i = d_R \\ 0 & \text{if } d^i = d_D \end{cases} \tag{6.14}$$

Prior beliefs about θ were assumed to be given by probability density function $f(\tilde{\theta}) = 8(2 - \tilde{\theta})/9$, for $\tilde{\theta} \in [1/2, 2]$.

Information z^i observed after selection of decision $d^i = d_S$ informed the DM whether, prior to its selection, her expected adaptive utility for that decision had been too great, $z^i = 1$ (negative surprise), or too small, $z^i = 0$ (positive surprise). No information about preferences could be observed if decisions d_R or d_D were selected. In this setting it was optimal to select $d^1 = d_S$, and then to either continue selecting $d^i = d_S$ if $z^1 = 0$, or replace in the second period if $z^1 = 1$, leading to maximum expected adaptive utility of -1.42.

As no information is available to the DM if she selects either of decisions d_R or d_D, both of these have an associated EVSI equal to 0. Hence all adaptive utility for them is associated with the actual reward outcome they lead to. Furthermore, Equation (6.14) leads to ${}_a\tilde{u}(d^i = d_R) = i/2 - 2$. However, if the DM selects decision d_S, then not only does she spend a period working with the unreliable machine, but she also gains information about her preferences, and both of these events have an associated expected adaptive utility value. In the first period, under prior beliefs, ${}_a\tilde{u}(d^1 = d_S) = -0.51$, whilst the EVSI for z^1 is approximately 0.06 (the maximum expected adaptive utility without noting z^1 is approximately -1.49, whilst this is approximately -1.23 and -1.61 for $z^1 = 0$ and $z^1 = 1$, respectively). Hence the full contribution of $d^1 = d_S$ is the sum of these two values (-0.51 and 0.06) and is thus approximately -0.45.

By the second period, without taking into account possible knowledge of z_1 (we have incorporated its EVSI into the full expected adaptive utility of decision $d_1 = d^S$), the pure (and also full) expected adaptive utility of decision $d_2 = d^R$ is equal to -1 (and because it is known that in period 3 there will be no utility loss, this is equivalent to -0.5 in each of the next two periods). The decision $d_2 = d^S$, not taking into account z_1, again has pure expected adaptive utility value approximately -0.51, but also has EVSI approximately 0.04 (reduced from the EVSI for $d_1 = d^S$ because the observation is being made later in the decision sequence), making its full expected adaptive utility value approximately -0.47.

Finally, in period 3, forgetting previous information z^1 and z^2, the decision d_R has expected adaptive utility value exactly -0.5, whilst alternative d_S again has pure expected adaptive utility value approximately -0.51. Furthermore, as this is the final decision and there will not be further opportunity to use any utility information, the decision $d^3 = d_S$ has associated EVSI equal to 0 and so -0.51 is the full value for decision $d^3 = d_S$.

Due to the additive nature of the utility function in this example, and because in such a situation it is meaningful to consider the one period utility value of a decision, we can decompose the maximum expected adaptive utility of -1.42 into the contributions from the full maximum expected adaptive utilities of the one period decisions. In this case we find that -1.42 equals the full expected adaptive utilities of decisions $d^1 = d_S$ (-0.45), $d^2 = d_S$ (-0.47), and $d^3 = d_R$ (-0.5).

Note that this decomposition, whilst allowing the DM to determine maximum expected adaptive utility, does not inform the DM of the optimal decision selection strategy, which is instead found through application of Equations (5.3) and (5.4). We further note that, if z^1 had been recorded following selection of decision $d^1 = d_S$, then z^2 has EVSI equal to 0. This can be verified either directly from Equation (6.12), or by noting that according to Table 5.4, the final period decision does not depend upon the value of z^2. $\diamond$

6.2 Risk and Trial Aversion

The following introductory review on the concept of classical risk aversion is based upon the theory developed independently by Arrow [6] and Pratt [83]. In classical utility theory, risk aversion is a diagnostic relating to a DM's preference for avoiding actuarially fair gambles. A DM who seeks to partake in such gambles is referred to as risk seeking, whilst a DM who is indifferent to such gambles is referred to as risk neutral.

If the reward set $\mathcal{R}$ is sufficiently rich that it can be identified with a finite interval of $\mathbb{R}$, *i.e.*, $\mathcal{R} = [a, b] \subset \mathbb{R}$, then for any decision d within the convex set of feasible decisions, the expected utility of d will equal $u(r)$ for some suitable reward $r \in \mathcal{R}$. Such r is referred to as the certainty equivalence for decision d, and will be denoted here by c_d. Thus the certainty equivalence of a decision is that reward $r = c_d$ making the DM indifferent between receiving c_d for sure, or selecting the possibly uncertain (with respect to its outcome) decision d. Furthermore, we are able to determine the expected value of the reward arising from any given decision, and this will be denoted by e_d. The risk premium associated with decision d, denoted ρ_d, is then defined to be the difference of c_d from e_d, *i.e.*, $\rho_d = e_d - c_d$.

Over a subset $[\alpha, \beta] \subset \mathbb{R}$, the DM is said to be risk averse if, for any decision d with all possible returns falling in $[\alpha, \beta]$, the risk premium $\rho_d \geq 0$. Similarly, the DM is said to be risk seeking if $\rho_d \leq 0$, and is said to be risk neutral if $\rho_d = 0$. Thus we see that, under such a definition, risk aversion is a concept related to the DM's aversion or willingness for selecting a decision that has greater uncertainty over its likely outcome in comparison to another with equal expected return, but which has reduced outcome uncertainty. An additional result is that, presuming that the function $u(r)$ is non-decreasing and twice-differentiable over region $[\alpha, \beta]$, the DM is risk averse over region $[\alpha, \beta]$ if and only if her utility function for rewards in that region is strictly concave. Similarly, a strictly convex utility function relates to the DM being risk prone, and a linear utility function corresponds to risk neutrality.

A closely related method of measuring a DM's level of risk aversion is the Arrow-Pratt measure of absolute risk aversion, see *e.g.*, Arrow [6] and Pratt [83] (Pratt refers to this as risk aversion in the small, *i.e.*, over a region $[\alpha, \beta]$, rather than global risk aversion). If $u(r)$ is a twice continuously differentiable function with positive first derivative, then absolute local risk aversion $l(r)$ is given by:

$$l(r) = -\frac{u''(r)}{u'(r)} \tag{6.15}$$

Because $u'(r) > 0$, the sign of $l(r)$ depends on, and is opposite to, the sign of $u''(r)$. This in turn determines whether $u(r)$ is convex or concave. However, the absolute local risk aversion $l(r)$ also has additional useful features. For example, for two utility functions u and $\tilde{u}$, $l(r) = \tilde{l}(r)$ for all possible r if and only if u is a positive linear transformation of $\tilde{u}$, implying that they represent the same preference relations. Moreover, if $l(r) > \tilde{l}(r)$ for all r in an interval $[\alpha, \beta]$, then for any decision d with all possible returns in interval $[\alpha, \beta]$, $\rho_d > \tilde{\rho}_d$. Hence, $l(r)$ also gives an indication of the strength of a DM's risk aversion.

Furthermore, given only the absolute local risk aversion $l(r)$, an analyst, knowing the general solution to a homogeneous second-order ODE with variable coefficients, can derive the DM's utility function:

$$\begin{aligned} l(r) = -\frac{u''(r)}{u'(r)} \quad &\Rightarrow \quad u''(r) + l(r)u'(r) = 0 \\ &\Rightarrow \quad e^{\int l(r)dr}[u''(r) + l(r)u'(r)] = 0 \\ &\Rightarrow \quad \frac{d}{dr}\left\{e^{\int l(r)dr}u'(r)\right\} = 0 \\ &\Rightarrow \quad e^{\int l(r)dr}u'(r) = k_1 \\ &\Rightarrow \quad u(r) = k_1 \int e^{-\int l(r)dr}dr + k_2 \end{aligned} \tag{6.16}$$

As a utility function is only unique up to a positive linear transformation, the values of constants k_1 and k_2 are irrelevant. Only the sign of k_1 is important, and this can be determined through the constraint that $u'(r) > 0$. Equation (6.16) also has an important implication for adaptive utility. As was discussed in Chapter 4, a state of mind θ characterises the DM's preferences and hence characterises the DM's true utility function. Thus, Equation (6.16) demonstrates that a state of mind will also characterise the DM's true absolute local risk aversion.

The connection between the DM's absolute local risk aversion and her utility function means that, rather than only considering possibilities for the true utility function $u(r|\theta)$, a DM can derive her adaptive utility function from a list of possible candidates for her absolute local risk aversion, $l(r|\theta)$, and their associated subjective probabilities of being true. This is demonstrated in the following example.

Example 6.2.1

A DM has opportunity to invest in a new venture, with V meaning investment and $\overline{V}$ non-investment. Event S ($\overline{S}$) indicates that the venture was (not) successful, and prior beliefs are such that $P(S) = 0.4 = 1 - P(\overline{S})$. If the DM invests and event S occurs, then she will receive £200000, but if she invests and event $\overline{S}$ occurs, then she will receive £0. If she does not invest, then she will keep the £500 that was necessary for the investment.

The DM is uncertain of her utility function for monetary returns over the interval $[0, 200000]$, but is comfortable with considering possibilities for her own level of risk aversion. Indeed, with probability 0.8 she believes that she has a positive risk aversion, but one which diminishes as her wealth increases. Alternatively, she considers she has a positive risk aversion, but constant over all monetary values considered in this problem.

The DM's true absolute local risk aversion over this region will be denoted by $l(r|\theta)$, and from the above we assume that with probability 0.2, $l(r|\theta) = l(r|\theta_1) = 0.001$, and with probability 0.8, $l(r|\theta) = l(r|\theta_2) = \frac{1}{r+1}$. Application of Equation (6.16) thus results in (where $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$):

$$l(r|\theta) = l(r|\theta_1) = 0.001 \quad \Rightarrow \quad u(r|\theta_1) = -\lambda_1 e^{-0.001r} + \lambda_2, \text{ (with } \lambda_1 > 0) \quad (6.17)$$

$$l(r|\theta) = l(r|\theta_2) = \frac{1}{r+1} \quad \Rightarrow \quad u(r|\theta_2) = \mu_1 \ln(1+r) + \mu_2, \text{ (with } \mu_1 > 0)$$

As described in Section 4.3, such possible classical utility functions require scaling to ensure that they are commensurable, *i.e.*, so they can be meaningfully compared. Hence the DM must now consider her preferences over rewards that are conditioned on different possibilities for the state of mind θ. Assuming that the correct value for the local risk aversion will not influence the DM's preferences for receiving the particular amounts[2] of £200000 or £0 (*e.g.*, the DM is just as happy to receive £200000 if $\theta = \theta_1$ as she is if $\theta = \theta_2$), we can scale each of the two possible classical utility functions by suitably choosing values for constants λ_1, λ_2, μ_1 and μ_2 to ensure that $u(\pounds 200000|\theta_1) = u(\pounds 200000|\theta_2) = 1$ and $u(\pounds 0|\theta_1) = u(\pounds 0|\theta_2) = 0$.

We find $\lambda_1 = 1$, $\lambda_2 = 1$, $\mu_1 = 0.08$ and $\mu_2 = 0$. The adaptive utility is thus:

$$ {}_a u(r) = E_\theta[u(r|\theta)] = 0.2(1 - e^{-0.001r}) + 0.8(0.08\ln(1 + r)) \tag{6.18} $$

Using this adaptive utility function in combination with beliefs over the state of nature allows the DM to determine the adaptive utility of the two decisions V and $\overline{V}$. In this case we find ${}_a u(V) = 0.49$, whilst ${}_a u(\overline{V}) = 0.4$. Thus the DM should invest in the venture. Note, however, that if it were assumed that $\theta = \theta_2$ with certainty, then the DM should not invest. ◇

As discussed above, risk aversion is a concept relating to the curvature of utility as a function of r. However, in an adaptive utility setting, there is more than one possible utility function, and the utility function $u(r|\theta)$ can be considered a function of both r and θ. We consider the implication of this by first returning to the apple or banana example.

[2] This statement can only be true for two amounts. To include a third would lead to no solutions for the scaling of the two possible classical utilities. Once two values have been selected, the true absolute local risk aversion will affect preferences over every other particular reward.

Example 6.2.2

Consider again the apple or banana example, last seen in Example 6.1.3, but now with three possible states of mind, $\theta \in \{1, 2, 3\}$. The case $\theta = 1$ corresponds to a preference for apples, the case $\theta = 2$ corresponds to indifference, and the case $\theta = 3$ corresponds to a preference for bananas.

We assume prior beliefs over θ are such that $P(\theta = i) = 1/3$ for $i = 1, 2, 3$. Thus $E[\theta] = 2$, corresponding to the situation that the DM is indifferent between apples and bananas. However, consider possible classical utility functions $u_1(b|\theta) = \theta - 1$, $u_2(b|\theta) = e^\theta - e^2 + 1$, $u_3(b|\theta) = \ln(\theta/2) + 1$, and $u_i(a|\theta) = 1$ for $i = 1, 2, 3$, with commensurablility assumed within each u_i over the various possible values for the state of mind. Each of the utilities u_i agrees with the above meanings for the possible values of θ, however, all three lead to different decision selection in a one-period adaptive utility problem:

$$u_1(b|\theta) = \theta - 1 \quad \Rightarrow \quad {}_a u(b) = 1 \tag{6.19}$$

$$u_2(b|\theta) = e^\theta - e^2 + 1 \quad \Rightarrow \quad {}_a u(b) = 3.7$$

$$u_3(b|\theta) = \ln(\theta/2) + 1 \quad \Rightarrow \quad {}_a u(b) = 0.9$$

We see that, if possible classical utilities are as expressed by u_2, the DM should select the banana, whilst if they are expressed by u_3, she should select the apple. This is despite the DM's prior beliefs being such that, in expectation, bananas are just as good as apples. $\diamond$

In Example 6.2.2 there is no risk in the result of decision selection, as selection of either fruit leads to consumption of that fruit with certainty. Indeed, the only form of risk in the problem arises through uncertainty over true preferences.

All three utility possibilities have the same prior beliefs about the state of mind θ. The only difference is the effect that various values for θ have on the possibilities for the DM's utility function. Thus, just as a DM may be averse to actuarially fair gambles with uncertain outcomes, certain forms of adaptive utility functions demonstrate an aversion to decisions whose utility values depend on an uncertain state of mind.

Such a form of adaptive utility would be appropriate if, despite prior beliefs making preferences indifferent in expectation, the DM would prefer the apple because she feels safe in the knowledge of what to expect from it. She may be averse to trying the banana because, even though there is the potential for even greater pleasure, she does not wish to take the chance of consuming something she dislikes. To distinguish this form of aversion from classical risk aversion, as reviewed at the beginning of this section, we will refer to it as *trial aversion*. In Example 6.2.2, we would refer to u_1 as a *trial neutral* function, u_2 as a *trial seeking* function, and u_3 as a *trial averse* function.

To formally define trial aversion, we require that Θ be a continuous space[3], so as to ensure that $E[\theta]$ always has a meaning as a possible state of mind itself, *i.e.*, that $E[\theta] \in \Theta$. In this is indeed the case, then we claim trial aversion to be a geometrical feature of $u(r|\theta)$ when viewed as a function of θ only (*i.e.*, r is assumed constant). Given a reward r with uncertain utility value, we define a DM as being *trial averse* with respect to r if $u(r|E[\theta]) > {}_a u(r)$. In addition, the DM is said to be *trial seeking* if $u(r|E[\theta]) < {}_a u(r)$, and *trial neutral* if $u(r|E[\theta]) = {}_a u(r)$. Note that this definition implies that a DM will be trial neutral for any reward she knows the true utility value of.

[3]We assume this is possible, yet accept that Chapter 4 assumed discrete Θ when presenting arguments for selecting decisions so as to maximise expected adaptive utility. Further discussion on the development of adaptive utility theory for continuous Θ is presented in Chapter 7.

If $u(r|\theta)$ is a non-decreasing and twice differentiable function of θ, then, for a given reward r, a DM is trial averse if $u(r|\theta)$ is strictly concave as a function of θ. Similarly, a DM is trial seeking if $u(r|\theta)$ is strictly convex in θ, and is trial neutral if $u(r|\theta)$ is linear in θ.

We use the term 'trial aversion', not only to distinguish it from the classical meaning of risk aversion, but also because in an n-period sequential problem the greater the trial aversion the less likely the DM is to select the uncertain reward in order to learn about it. For example, if there exists an alternative reward r' with known utility value such that $u(r|E[\theta]) > u(r') > {}_a u(r)$ (*i.e.*, the true utility of r' does not depend on the value of θ), then, in a one-period problem, the DM should select r' over r. The reason she should not select r if θ were uncertain is because her trial aversion implies that the potential cost of discovering that θ is worse than expected, outweighs the potential benefit of discovering that θ is better than expected.

When $u(r|\theta)$ is considered over the region $\mathcal{R} \times \Theta$, risk aversion is a geometrical feature concerning the curvature of the function along the r-axis for a fixed value of θ. Trial aversion is instead an orthogonal concept relating to the curvature of the function along the θ axis for a fixed value of r. Another difference between the two is that, as discussed above, it is meaningful to discuss local risk aversion, where all possible outcomes of decisions fall in some subset of $\mathcal{R}$. However, trial aversion is necessarily a global concept, as the DM must consider all possible states of mind when determining $E[\theta]$ or ${}_a u(r)$.

Just as a risk aversion, for a given state of mind θ, can be measured through the Arrow-Pratt measure of absolute risk aversion, a DM's degree of trial aversion for a given reward level r can be measured through an analogous measure of absolute trial aversion. Provided $u(r|\theta)$ is a twice continuously differentiable function with positive first derivative, we denote this measure by $t_r(\theta)$ and define it by:

$$t_r(\theta) = -\frac{\partial^2 u(r|\theta)}{\partial \theta^2} / \frac{\partial u(r|\theta)}{\partial \theta} \tag{6.20}$$

The function $t_r(\theta)$ has analogous properties to the Arrow-Pratt measure of absolute risk aversion $l(r)$. Indeed, it applies the same calculation as the Arrow-Pratt measure, but simply upon a different variable. In particular, the DM is trial averse if $t_r(\theta) > 0$, trial seeking if $t_r(\theta) < 0$, and trial neutral if $t_r(\theta) = 0$. This is because we required $\partial u(r|\theta)/\partial\theta > 0$, and so the sign of $t_r(\theta)$ is opposite to that of $\partial^2 u(r|\theta)/\partial\theta^2$ and hence determines whether $u(r|\theta)$ is a convex or concave function of θ.

Applying this measure to the three utility functions considered in Example 6.2.2, we find that for u_1, $t_b(\theta) = 0$, for u_2, $t_b(\theta) = -1$, and for u_3, $t_b(\theta) = 1/\theta$. All of these are in line with the description of u_1 as being trial neutral, u_2 as being trial seeking, and u_3 as being trial averse.

Since the expected utility hypothesis of Bernoulli [18], it has been assumed that most DMs act in a risk averse manner, especially when decisions concern monetary returns (an exception being the act of gambling, but this can be explained through a utility for the exhiliration that such an activity provides). As Arrow [6] notes, this hypothesis explains many economic activities such as insurance, or the aversion for entering high risk investments. Indeed, the assumption that a utility function is bounded implies that eventually the DM must be risk averse beyond a certain reward level.

However, it would appear that DMs are not necessarily trial averse in their attitudes towards decision selection. Trial aversion relates to an unwillingness to experiment, or to select the uncertain. Nevertheless, the opposite to this form of behaviour can be observed frequently in everyday life. Indeed, every experience that a DM encounters must have been novel to her at some point. DMs often order meals they have never tasted before over others they are more familiar with, or a DM with a severe medical problem may readily select a remedy which offers only the faintest possibility of providing a lifestyle which she has never before experienced.

That is not to say DMs are generally trial seeking, and not always do DMs wish to try novel rewards. For example, a DM may be averse to trying a new pastime such as attending a football match, or on holidaying in a new and different location. Although we do not wish to over-generalize, it may be that there is a connection between level of trial aversion and the age of a DM, with a potential hypothesis being that at a younger age DMs demonstrate a level of trial seeking, with trial aversion increasing as experience increases.

Perhaps the most obvious example of trial aversion within society, however, is the preference of some DMs to stick with products that are produced by a familiar brand. Indeed, Erdem & Keane [39] state "estimates indicate that consumers are risk-averse[4] with respect to variation in brand attributes, which discourages them from buying unfamiliar brands."

Just as risk aversion can explain the existence of insurance companies, in this situation trial aversion can be used to explain the existence of marketing companies. If a DM was not trial averse with respect to trying a new brand, then all that would be needed for her to select a new product would be to know of its existence. No longer would there be any need for marketing companies to attempt to 'sell' the new product over existing possibilities, *e.g.*, by offering trial samples. Indeed, it was for real world observations such as this that Cyert & DeGroot [27] first considered a mathematical model for decision making with uncertain preferences.

[4]Under the arguments presented in this section, we would refer to this behaviour as trial aversion.

Chapter 7

Conclusions and Future Directions

This chapter provides a short summary of the main results of this thesis, followed by a discussion of potential directions for further research.

7.1 Conclusions

This thesis has extended the adaptive utility concept that was first introduced by Cyert & DeGroot [27]. In Chapter 2 we reviewed various suggestions in the decision theory literature for solving single and sequential decision problems. Chapter 3 provided motivation for permitting uncertain utility and offered a formal definition of an adaptive utility function. An interpretation of a utility parameter, our so-called state of mind, was proposed in Chapter 4, and we discussed its role within decision analysis. Chapter 4 also focused on important foundational issues of adaptive utility that had been previously overlooked in the literature. The main result of the chapter is that, under the assumption that the DM agrees with the classical system of expected utility axioms and has scaled possible utility functions to ensure commensurability, the logical strategy is to select decisions so as to maximise expected adaptive utility.

Applications within sequential decision problems were illustrated in Chapter 5, where an algorithm was developed for solving n-period sequential problems and a discussion was given on how decision selection had to influence likely utility information for the theory to deviate from classical, known utility, solutions. Chapter 5 also discussed some possible methods for receiving utility information and a couple of small hypothetical examples were presented to illustrate the potential of adaptive utility for sequential problems within reliability theory.

Chapter 6 considered the implication of uncertain preferences for the classical utility diagnostics of value of information and risk aversion, an area that has not been discussed in any previous work on adaptive utility or related uncertain utility theories. Whilst it was shown that the DM's beliefs over her state of mind could influence the expected value of sample information, it was demonstrated that no simple monotone relation exists. Moreover, the hypothesis that a decrease in uncertainty over the state of mind would necessarily lead to an increase in expected value of sample information was shown to be false. Section 6.1 also considered the value of information relating to the DM's true preferences, and it was shown that, when it is meaningful to talk about the utility value of a single decision within a sequence, the expected adaptive utility of that decision could be decomposed into a part arising from expected decision outcome, and a part arising from expected value of information related to the DM's preferences.

Section 6.2 considered effects on risk aversion, and it was shown that the state of mind not only characterised the DM's true complete preference ordering over decisions, but also has a one-to-one relationship with the Pratt-Arrow measure of absolute risk aversion. For this reason we demonstrated that a DM could derive her adaptive utility function from only considering possibilities for her absolute risk aversion. Finally, we introduced trial aversion, a concept relating to a DM's aversion or willingness to try rewards or select decisions that are novel, or are such that the DM is uncertain about her preferences for them in comparison to alternative options.

7.2 Future Directions

Following the formal approach to the development of adaptive utility presented in this thesis, there are many areas for further development of the theory. In particular, we accept that, though of philosophical interest, the current state of the theory means that the benefit of allowing the DM to remain uncertain over her preferences is overshadowed by the complexity of implementation in non-trivial decision problems. For this reason we conclude the thesis with a number of options for future development that we have identified and believe to be necessary before the theory can be applied to a wide variety of interesting problems.

As was mentioned at the end of Section 5.1, the employment of dynamic programming and backward induction for solving sequential decision problems suffers from the so-called curse of dimensionality, and the inclusion of a possibly multi-dimensional state of mind for representing uncertainty over preferences only exasperates the situation. A DM must then consider whether or not the relative benefit of remaining uncommitted to a specific and assumed correct utility function outweighs the increased cost of computational complexity that is required in solving an adaptive utility problem. Such computational complexity is currently the greatest hindrance in solving an adaptive utility problem, but as is discussed in Section 5.1, it may well be possible to identify forms of utility functions that are not only reasonable for modelling possible preferences, but which also greatly reduce the computational requirements of solution algorithms.

In addition to seeking combinations of probability distributions and forms of utility functions that reduce the computational complexity of a sequential adaptive utility problem, there are further issues that can be considered for enhancing the use of the theory in solving real-world decision problems. For example, issues involving the elicitation of the DM's beliefs should be addressed, and additional problem types in which it would be regarded as highly beneficial to permit the DM to remain uncertain over preferences should be identified.

Work on elicitation is required as now a collection of possible utility functions must be considered and a probability distribution over them specified. This is a non-trivial task, especially if the DM is not trained to think about uncertainties. Bedford *et al.* [10] consider the problem of probability elicitation[1], whilst Chajewska *et al.* [22] consider the problem of eliciting utilities (see the discussion of this work in Section 5.2). However, there appears to be no discussion in the literature on how a DM who is uncertain of her preferences may elicit her own possible utility functions and the probabilities of each representing true preferences. The identification of a source of utility information following decision selection and determination of an appropriate likelihood function will also be difficult, yet necessary, hence any method of easing this task will be of much use.

In addition to developing the theory so as to make it more suitable for real-world application, there are a number of possibilities for development on a theoretical level. For example, the adaptive utility setting considered here requires the DM to have a known and precise prior probability distribution over her true preferences. However, it may be that this is too strong a requirement for some DMs to commit to, and instead imprecise or nonparametric extensions may be more reasonable. Sections 2.1.2 and 2.2.3 discussed recent research into imprecise probability and utility theories, but the combination of imprecise utility information with learning following decision implementation may be an interesting area for further development.

Augustin & Coolen [8] and Coolen [25] discuss nonparametric predictive inference as an alternative to a precise parametric probability distribution for quantifying uncertainty, and it may be that such a theory could also be useful for quantifying uncertainty over utilities. For example, suppose a DM has known utility values for ten types of fruit, with fruits numbered from 1 to 10 such that $u(i) \leq u(i+1)$, with $i = 1, \ldots, 9$, and where utilities are bounded in the interval $[0, 1]$. Further suppose that the DM is afforded opportunity to select a new fruit x with unknown

[1] Bedford *et al.* consider elicitation problems within reliable system design, discussing the need to support decision making with suitable subjective assessments about likely future system reliability.

utility. The nonparametric setting of Coolen [25] would assume that beliefs are $P(u(x) \in [0, u(1)]) = P(u(x) \in [u(i), u(i+1)]) = P(u(x) \in [u(10), 1]) = 1/11$. The advantage of such a setting is that it allows the DM to remain uncommitted to specifying a precise probability distribution over the value of $u(x)$. Furthermore, the values of known utilities $u(1), \ldots, u(10)$ will influence the DM's decision as to whether or not she should try fruit x, for if many fruits were seen to have low utility, with only one or two having larger values, then with high probability the new fruit x will also have low utility. The number of fruits the DM has knowledge of will also affect the DM's decision to try the new fruit, for if only 1 or 2 fruits had known utility, then there would be increased uncertainty over $u(x)$.

Unfortunately, as with the theories that imprecisely quantify uncertainty (see Subsection 2.2.3), such a nonparametric setting leads to a class $\mathcal{P}$ of possible probability distributions for the value of $u(x)$. This means there will be a set of possible expected utility values, and so the theory can not guarantee the identification of a uniquely admissible decision.

The argument in Section 4.2 for the existence and uniqueness of an adaptive utility function was only applicable in the context that decisions are lotteries with finite support, *i.e.*, the set of possible rewards $\mathcal{R}$ was assumed finite. This argument also assumed a finite set Θ of possible states of mind. Nevertheless, we later assumed that an adaptive utility function could be defined for when either $\mathcal{R}$ or Θ are continuous, *e.g.*, the discussion of risk aversion in Section 6.2. To be meaningful and ensure that $E[\theta] \in \Theta$, the concept of trial aversion that was introduced in Section 6.2 specifically required a continuous set of possible states of mind Θ. For this reason development of adaptive utility theory for continuous $\mathcal{R}$ or Θ is required. However, as with our repeated application of Anscombe & Aumann's subjective utility [5], we expect this simply requires repeated application of traditional expected utility results for permitting continuous $\mathcal{R}$ (for example, Herstein & Milnor [54] provide the expected utility result in the case that decisions are discrete distributions over a continuous reward space).

There are also a number of possibilities in extending the work of Chapter 6 concerning utility diagnostics. For example, it was demonstrated in Section 6.1 that a decrease in uncertainty over the state of mind does not necessarily lead to an increase in expected value of information. However, it may be that such a relation does exist for certain utility forms or decision settings, and identification of these cases, and understanding why they display such a property, will be of interest. The relationship between expected value of information and trial aversion also requires investigation. In comparison to a trial seeking DM, we would expect that the length of a sequential problem must be extended before a trial averse DM would select an unfamiliar decision or reward (assuming that such decision selection leads to information concerning preferences). Whether this is because a trial averse DM places less expected value on utility information than does a trial seeking DM, or whether it is because a trial averse DM places less 'pure' value on decisions, is also of interest.

Other types of sequential decision problems, as opposed to that considered in this thesis, may also be considered in an adaptive utility setting. For example, the sequential problem of sampling termination is considered by DeGroot [34], and it may be of interest to consider generalising classical results in that setting, *e.g.*, the Sequential Probability Ratio Test.

Finally, a suggested hypothesis in Section 6.2 was that DMs become increasingly trial averse with age. Though we have sought to provide a normative theory for the problem of sequential decision making, the descriptive validity of adaptive utility and uncertain preferences, along with such hypotheses concerning trial aversion, may well be of interest to psychologists and economists.

Appendix A

Glossary

c_d	Certainty equivalence of decision d.
$d, \mathcal{D}$	$d \in \mathcal{D}$ is a decision, with $\mathcal{D}$ the set of possible decisions.
e_d	Expected return of decision d.
$H^i, {}_aH^i$	H^i lists history of decisions and returns prior to i-th decision, ${}_aH^i$ also lists observed information about the state of mind.
$I_y, {}_aI_y$	I_y is a function for determining value of information about uncertain quantity y, with ${}_aI_y$ used when utility is uncertain.
l	Function detailing absolute local risk aversion
$P, \mathcal{P}$	$P \in \mathcal{P}$ is a probability distribution, with $\mathcal{P}$ the set of probability distributions.
$r, \mathcal{R}$	$r \in \mathcal{R}$ is a return following decision selection, with $\mathcal{R}$ the set of possible returns.
t_r	Function detailing absolute trial aversion at reward level r.
$u, {}_au$	u is a traditional utility function, ${}_au$ is an adaptive utility function.
$z, \mathcal{Z}$	$z \in \mathcal{Z}$ is information about the state of mind, with $\mathcal{Z}$ the set of possible information statements.
θ, Θ	$\theta \in \Theta$ is the true state of mind, with Θ the set of possible states of mind.
π^i	Decision strategy for period i.
ρ_d	Risk premium for decision d.
ω, Ω	$\omega \in \Omega$ is the state of nature, with Ω the set of possible states of nature.
$\succeq, \succ, \sim$	Binary preference relations.

Appendix B

Extension to Example 5.3.1

The calculation of expected adaptive utilities and the maximisation over various decision histories that were necessary in Example 5.3.1 (denoted by the E's and L's in that example) were performed using the software package Maple 10. It is not feasible to include them in full detail within this thesis. For example, lists $L5$, $L6$ and $L7$ must between them contain all 8 dimensional vectors representing possible observed failure histories by the final period, of which there are 10,000. The expected adaptive utility functions are also rather messy calculations of up to 8 variables. Nevertheless, given the problem description, the adaptive utility function and prior beliefs, these calculations can easily be verified using the algorithm given in Section 5.1.

Example 5.3.1 is based on a hypothetical scenario and it is hoped that the omission of these calculations do not prevent the reader from appreciating the implications of, and possibilities for, adaptive utility theory. However, as an example of how these calculations are performed, and as a demonstration of the general procedure, the description of how list $L2$ was determined is included below. Examination of the row in Table 5.1 containing $E2$ and $L2$ informs us that these were relevant if decision d_B had previously been made. Thus the current decision is whether or not to permanently fix failure types A, and as beliefs over the number and severity of future failures of type A will not depend on previous observed failures of type B, attention is restricted to the timing and perceived severity of failures of type A.

Because Table 5.1 represents the decision to be made in the final epoch, entries in $L2$ will be of the form of a 4 dimensional vector $(z^1_{AS}, z^1_{AM}, z^2_{AS}, z^2_{AM})$, with z^i_{Aj} being the observed number of failures of type A that occurred with perceived severity j in epoch i. Furthermore, as there can be at most 3 type A failures of any perceived severity within an epoch, there are 100 possible decision histories that must be considered.

To evaluate the expected adaptive utility of the decision $d^3 = d_N$, the updated beliefs over θ_A and ω_A for all possible histories must be determined (examination of the adaptive utility function assumed for this problem demonstrates that beliefs over θ_B and ω_B are no longer relevant). Those histories that lead to updated beliefs over these parameters being such that the expected adaptive utility of $d^3 = d_N$ is greater than -80 are then placed in list $L2$.

Application of Bayes' Theorem and use of independencies within the problem leads to the result that for given history $H = (z^1_{AS}, z^1_{AM}, z^2_{AS}, z^2_{AM})$, updated beliefs over θ_A are such that $P(\theta_A = 0.7|H) = 1 - P(\theta_A = 0.3|H)$, with $P(\theta_A = 0.7|H)$ described following simplification by:

$$P(\theta_A = 0.7|H) = \frac{4 \times 7^{z^1_{AS}+z^2_{AS}} \times 3^{z^1_{AM}+z^2_{AM}}}{4 \times 7^{z^1_{AS}+z^2_{AS}} \times 3^{z^1_{AM}+z^2_{AM}} + 3^{z^1_{AS}+z^2_{AS}} \times 7^{z^1_{AM}+z^2_{AM}}} \quad \text{(B.0.1)}$$

Similarly, updated beliefs over parameter ω_A given history H can also be determined and are such that $P(\omega_A = 0.005|H) = 1 - P(\omega_A = 0.1|H)$, with $P(\omega_A = 0.005|H)$ described following simplification by:

$$P(\omega_A = 0.005|H) = \frac{7 \times 199^{6-k}}{7 \times 199^k + 2^{12} \times 5^6 \times 3^{13-2k}} \quad \text{(B.0.2)}$$

with $k = z^1_{AS} + z^1_{AM} + z^2_{AS} + z^2_{AM}$

Determining the updated expected values of parameters θ_A and ω_A from the above distributions, and putting them into the example's adaptive utility function, then provides expected adaptive utility equation $E2$. Finally, list $L2$ is generated by considering each possible history and determining whether or not for that history equation $E2$ has a greater value than -80 (the expected adaptive utility of decision $d^3 = d_N$). The full list of histories in $L2$ is given below (note that these are expressed in the form $(z^1_{AS}, z^1_{AM}, z^2_{AS}, z^2_{AM})$):

(0,0,0,0) (0,0,0,1) (0,0,0,2) (0,0,0,3) (0,0,1,0) (0,0,1,2) (0,1,0,0) (0,1,0,1) (0,1,0,2)
(0,1,0,3) (0,1,1,1) (0,1,1,2) (0,2,0,0) (0,2,0,1) (0,2,0,2) (0,2,0,3) (0,2,1,0) (0,2,1,1)
(0,2,1,2) (0,2,2,1) (0,3,0,0) (0,3,0,1) (0,3,0,2) (0,3,0,3) (0,3,1,0) (0,3,1,1) (0,3,1,2)
(0,3,2,0) (0,3,2,1) (1,0,0,0) (1,0,0,2) (1,0,0,3) (1,1,0,1) (1,1,0,2) (1,1,0,3) (1,1,1,2)
(1,2,0,0) (1,2,0,1) (1,2,0,2) (1,2,0,3) (1,2,1,1) (1,2,1,2) (2,0,0,3) (2,1,0,2) (2,1,0,3)

Appendix C

A Conjugate Utility Class

This appendix demonstrates how the use of a utility function from the polynomial class of functions, as defined in Equation (5.5), leads to a closed and tractable solution to the nested sequence of expectations in Equations (5.3) and (5.4) when beliefs over the problem's variables are represented by Normal distributions.

First we assume prior beliefs over the state of nature ω are such that it follows a Normal distribution with known mean μ and known variance σ^2. We also assume that the distribution of reward r following decision d is Normal, with unknown mean $\mu_d(\omega) = \alpha_d\omega + \beta_d$ (with α_d and β_d known constants) and known variance σ_d^2.

Due to the conjugacy property of members of the exponential family of distributions (see, for example, Bernardo & Smith [17] or DeGroot [34]), posterior beliefs for ω can be easily found following the observation of rewards $r^1, \ldots, r^j$ when decisions $d^1, \ldots, d^j$ were made, respectively:

$$f_{d^1,\ldots,d^j}(r^1, \ldots, r^j|\omega) = \prod_{i=1}^{j} f_{d^i}(r^i|\omega) \propto \exp\left\{ - \sum_{i=1}^{j} \frac{(r^i - \mu_{d^i}(\omega))^2}{2\sigma_{d^i}^2} \right\} \quad \text{(C.0.1)}$$

$$f(\omega) \propto \exp\left\{ - \frac{(\omega - \mu)^2}{2\sigma_{d^i}^2} \right\}$$

$$\Rightarrow f(\omega|\{r^i, d^i\}_{i=1}^{j}) \propto \exp\left\{ - \frac{(\omega - \mu)^2}{2\sigma_{d^i}^2} - \sum_{i=1}^{j} \frac{(r^i - \mu_{d^i}(\omega))^2}{2\sigma_{d^i}^2} \right\}$$

Using that $\mu_d(\omega) = \alpha_d\omega + \beta_d$ and making suitable rearrangement then leads to:

$$f(\omega|\{r^i, d^i\}_{i=1}^j) \propto \exp\left\{ - \frac{1}{2\sigma^2/(1+\sigma^2\sum_{i=1}^j \frac{\alpha_{d^i}^2}{\sigma_{d^i}^2})}\left(\omega - \frac{\mu + \sigma^2\sum_{i=1}^j \frac{\alpha_{d^i}(r^i-\beta_{d^i})}{\sigma_{d^i}^2}}{1+\sigma^2\sum_{i=1}^j \frac{\alpha_{d^i}^2}{\sigma_{d^i}^2}}\right)^2\right\}$$

Hence posterior beliefs for ω follow a Normal distribution with mean η_1^j and variance η_2^j, where:

$$\eta_1^j = \frac{\mu + \sigma^2\sum_{i=1}^j \frac{\alpha_{d^i}(r^i-\beta_{d^i})}{\sigma_{d^i}^2}}{1+\sigma^2\sum_{i=1}^j \frac{\alpha_{d^i}^2}{\sigma_{d^i}^2}} \qquad \eta_2^j = \frac{\sigma^2}{1+\sigma^2\sum_{i=1}^j \frac{\alpha_{d^i}^2}{\sigma_{d^i}^2}} \tag{C.0.2}$$

Similarly, assuming that prior beliefs about θ follow a Normal distribution with known mean ν and known variance τ^2, and assuming that the distribution of utility information z^i following selection of decision d^i is also Normal with unknown mean $\nu_d(\theta) = \phi_d\theta + \psi_d$ (with ϕ_d and ψ_d known constants) and known variance τ_d^2, then after the observation of $z^1, \ldots, z^j$ posterior beliefs about θ follow a Normal distribution with mean λ_1^j and variance λ_2^j, where:

$$\lambda_1^j = \frac{\nu + \tau^2\sum_{i=1}^j \frac{\phi_{d^i}(z^i-\psi_{d^i})}{\tau_{d^i}^2}}{1+\tau^2\sum_{i=1}^j \frac{\phi_{d^i}^2}{\tau_{d^i}^2}} \qquad \lambda_2^j = \frac{\tau^2}{1+\tau^2\sum_{i=1}^j \frac{\phi_{d^i}^2}{\tau_{d^i}^2}} \tag{C.0.3}$$

In this setting it can also be shown that, knowing that previous decisions $d^1, \ldots, d^{j-1}$ led to rewards $r^1, \ldots, r^{j-1}$, respectively, the posterior predictive distribution of reward r^j following selection of decision d^j is also Normal with mean $\gamma_1^j = \alpha_{d^j}\eta_1^{j-1} + \beta_{d^j}$ and variance $\gamma_s^j = \alpha_{d^j}^2\eta_2^{j-1} + \sigma_{d^j}^2$.

Equations (5.3) and (5.4) detailed the solution algorithm for an n-period sequential adaptive utility decision problem. However, as was discussed in Chapter 5, such an algorithm requires that the DM compute a nested sequence of expectations, which in most cases leads to an intractable solution.

Nevertheless, if the DM's beliefs over problem variables are as detailed within this appendix, then one possibility for ensuring a closed form solution to such a sequence of expectations is to consider the use of a utility function that is a member of the the polynomial class that was given in Equation (5.5).

In determining optimal decision d^j when decision history ${}_aH^j$ is known, the DM is required to compute two types of expectations. The first of these is of the form $E_{\theta|_aH^j}[u(r^1,\ldots,r^n|\theta)]$, whilst the second is of the form $E_{{}_aH^{k+1}|_aH^k,d^k}[{}_au(r^1,\ldots,r^n)]$ for $k=j,\ldots,n-1$.

Presuming $u(r^1,\ldots,r^n|\theta)$ is a member of the polynomial class of Equation (5.5), we may solve $E_{\theta|_aH^j}[u(r^1,\ldots,r^n|\theta)]$ by the following:

$$\begin{aligned} E_{\theta|_aH^j}[u(r^1,\ldots,r^n|\theta)] &= \int_\Theta u(r^1,\ldots,r^n|\theta)f(\theta|_aH^j) \qquad \text{(C.0.4)} \\ &= \int_\Theta \sum_{k_0=0}^{m_0}\sum_{k_1=0}^{m_1}\cdots\sum_{k_n=0}^{m_n} a_{k_0,k_1,\ldots,k_n}(r^1)^{k_1}\cdots(r^n)^{k_n}\theta^{k_0}f(\theta|_aH^j) \\ &= \sum_{k_0=0}^{m_0}\sum_{k_1=0}^{m_1}\cdots\sum_{k_n=0}^{m_n} a_{k_0,k_1,\ldots,k_n}(r^1)^{k_1}\cdots(r^n)^{k_n}\int_\Theta \theta^{k_0}f(\theta|_aH^j) \\ &= \sum_{k_0=0}^{m_0}\sum_{k_1=0}^{m_1}\cdots\sum_{k_n=0}^{m_n} a_{k_0,k_1,\ldots,k_n}(r^1)^{k_1}\cdots(r^n)^{k_n}E_{\theta|_aH^j}[\theta^{k_0}] \end{aligned}$$

Given ${}_aH^j$, posterior beliefs over θ follow a Normal distribution with known mean and variance. Hence determination of $E_{\theta|_aH^j}[\theta^{k_0}]$ is the same as the determination of the raw moments of a Normal distribution, and these can be expressed as a polynomial function of the mean and variance of the distribution (see Papoulis [82]).

The same technique can also be applied to determine $E_{{}_aH^{k+1}|_aH^k,d^k}[{}_au(r^1,\ldots,r^n)]$, and although the following can be equally well applied when determining predictive beliefs over both r^k and z^k, we will assume utility information is not included in decision histories.

Noting that ${}_a u(r^1, \ldots, r^n) = E_{\theta|_a H^j}[u(r^1, \ldots, r^n|\theta)]$ is a polynomial function of $r^1, \ldots, r^n$, we use the result that, given ${}_a H^k$, the predictive distribution of r^k when decision d^k is selected is Normal with known mean and variance. Hence again the solution is equivalent to determining the raw moments of a Normal distribution, again leading to another polynomial of the returns $r^1, \ldots, r^{k-1}, r^{k+1}, \ldots, r^n$. Sequentially solving expectations of the form $E_{{}_a H^{k+1}|_a H^k, d^k}[{}_a u(r^1, \ldots, r^n)]$ by taking k backwards from n to 1 then permits calculation of necessary expected adaptive utilities.

Bibliography

[1] M. Abdellaoui, C. Barrios, and P.P. Wakker. Reconciling Introspective Utility with Revealed Preference: Experimental Arguments Based on Prospect Theory. *Journal of Econometrics*, 138:356–378, 2007.

[2] M. Ahlbrecht and M. Weber. Hyperbolic Discounting Models in Prescriptive Theory of Intertemporal Choice. *Zeitschrift für Wirtschafts - und Sozialwissenschaften*, 115:535–568, 1995.

[3] M. Allais. The Foundations of a Positive Theory of Choice involving Risk and a Criticism of the Postulates and Axioms of the American School (*English Translation*). In M. Allais and O. Hagen, editors, *Expected Utility Hypothesis and the Allais Paradox*, pages 27–145. D. Reidel Publishing Company, 1979.

[4] Y. Amihud. Critical Examination of the New Foundation of Utility. In M. Allais and O. Hagan, editors, *Expected Utility Hypotheses and the Allais Paradox*, pages 149–160. D. Reidel Publishing Company, 1979.

[5] F.J. Anscombe and R.J. Aumann. A Definition of Subjective Probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

[6] K.J. Arrow. The Theory of Risk Aversion. In *Essays in the Theory of Risk-Bearing*, chapter 3. Markham, Chicago, 1971.

[7] T. Augustin. Expected Utility within a Generalized Concept of Probability - A Comprehensive Framework for Decision Making under Ambiguity. *Statistical Papers*, 43:5–22, 2005.

[8] T. Augustin and F.P.A. Coolen. NONPARAMETRIC PREDICTIVE INFERENCE AND INTERVAL PROBABILITY. *Journal of Statistical Planning and Inference*, 124:251–272, 2004.

[9] R.D. Baker. RISK AVERSION IN MAINTENANCE: OVERMAINTENANCE AND THE PRINCIPLE-AGENT PROBLEM. *IMA Journal of Management Mathematics*, 17:99–113, 2006.

[10] T. Bedford, J. Quigley, and L. Walls. EXPERT ELICITATION FOR RELIABLE SYSTEM DESIGN. *Statistical Science*, 21(4):428–450, 2006.

[11] D.E. Bell. A UTILITY FUNCTION FOR TIME STREAMS HAVING INTER-PERIOD DEPENDENCIES. *Operations Research*, 25(3):448–458, 1977.

[12] R. Bellman. ON THE THEORY OF DYNAMIC PROGRAMMING. *Proceedings of the National Academy of Sciences*, 1952.

[13] R. Bellman. DYNAMIC PROGRAMMING. Princeton University Press, 1957.

[14] Y. Ben-Haim. INFORMATION GAP DECISION THEORY: DECISIONS UNDER SEVERE UNCERTAINTY. Second edition. Academic Press, 2006.

[15] J.O. Berger. STATISTICAL DECISION THEORY AND BAYESIAN ANALYSIS. Springer, 1993.

[16] J.M. Bernardo. EXPECTED INFORMATION AS EXPECTED UTILITY. *The Annals of Statistics*, 7(3):686–690, 1979.

[17] J.M. Bernardo and A.F.M. Smith. BAYESIAN THEORY. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, 1994.

[18] D. Bernoulli. EXPOSITION OF A NEW THEORY ON THE MEASUREMENT OF RISK (*translation*). *Econometrica*, 22:22–36, 1954 (original in *Comentarii Academiae Scientiarum Imperialis Petropolitanae* - 1738).

[19] S.E. Bodily and C.C. White. OPTIMAL CONSUMPTION AND PORTFOLIO STRATEGIES IN A DISCRETE-TIME MODEL WITH SUMMARY-DEPENDENT PREFERENCES. *The Journal of Financial and Quantitative Analysis*, 17(1):1–14, 1982.

[20] G. Boole. An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities. Walton and Maberly, 1854.

[21] C. Boutilier. On the Foundations of Expected Expected Utility. *18th International Joint Conference on AI*, pages 712–717, 2003.

[22] U. Chajewska, D. Koller, and R. Parr. Making Rational Decisions using Adaptive Utility Elicitation. *Proceedings 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence*, pages 363–369, 2000.

[23] T. Chavez. A Decision-Analytic Stopping Rule for Validation of Commercial Software Systems. *IEEE Transactions on Software Engineering*, 26(9):907–918, 2000.

[24] R.T. Clemen. Making Hard Decisions: An Introduction to Decision Analysis. Duxbury, 1996.

[25] F.P.A. Coolen. On Nonparametric Predictive Inference and Objective Bayesianism. *Journal of Logic, Language and Information*, 15:21–47, 2006.

[26] G.S. Crawford and M. Shum. Uncertainty and Learning in Pharmaceutical Demand. *Econometrica*, 73:1137–1173, 2005.

[27] R.M. Cyert and M.H. DeGroot. Adaptive Utility. In R.H. Day and T. Groves, editors, Adaptive Economic Models, pages 223–246. Academic Press, New York, 1975.

[28] R.M. Cyert and M.H. DeGroot. Learning Applied to Utility Functions. In A. Zellner, editor, *Bayesian Analysis in Econometrics and Statistics*, chapter 13, pages 159–168. North-Holland Publishing Company, 1980.

[29] R.M. Cyert and M.H. DeGroot. Bayesian Analysis and Uncertainty in Economic Theory, chapter 9: Adaptive Utility, pages 127–143. Chapman and Hall, 1987.

[30] Marquis de Condorcet. Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix. L'Imprimerie Royale, 1785.

[31] B. de Finetti. THEORY OF PROBABILITY, volume 1. John Wiley & Sons, 1974.

[32] B. de Finetti. A SHORT CONFIRMATION OF MY STANDPOINT. In M. Allais and O. Hagen, editors, *Expected Utility Hypotheses And The Allais Paradox*, page 162. D. Reidel Publishing Company, 1979.

[33] M.H. DeGroot. UNCERTAINTY, INFORMATION AND SEQUENTIAL EXPERIMENTS. *Annals of Mathematical Statistics*, 33:404–419, 1962.

[34] M.H. DeGroot. OPTIMAL STATISTICAL DECISIONS. McGraw-Hill Book Company, 1970.

[35] M.H. DeGroot. CHANGES IN UTILITY AS INFORMATION. *Theory and Decision*, 17:287–303, 1984.

[36] H. Diamond and G. Gelles. GAUSSIAN APPROXIMATION OF EXPECTED UTILITY. *Economic Letters*, 64:301–307, 1999.

[37] J. Drèze and A. Rustichini. STATE-DEPENDENT UTILITY AND DECISION THEORY. In S. Barbera, P.J. Hammond, and C. Seild, editors, HANDBOOK OF UTILITY THEORY, 2: EXTENSIONS, pages 839–892. Kluwer Academic Publishers, Dordrecht, 2004.

[38] D. Ellsberg. RISK, AMBIGUITY, AND THE SAVAGE AXIOMS. *Quarterly Journal of Economics*, 75(4):643–669, 1961.

[39] T. Erdem and M.P. Keane. DECISION-MAKING UNDER UNCERTAINTY: CAPTURING DYNAMIC BRAND CHOICE PROCESSES IN TURBULENT CONSUMER GOODS MARKETS. *Marketing Science*, 15:1–20, 1996.

[40] M. Farrow and M. Goldstein. TRADE-OFF SENSITIVE EXPERIMENTAL DESIGN: A MULTICRITERION, DECISION THEORETIC, BAYES LINEAR APPROACH. *Journal of Statistical Planning and Inference*, 136(2):498–526, 2006.

[41] W. Fellner. DISTORTION OF SUBJECTIVE PROBABILITIES AS A REACTION TO UNCERTAINTY. *Quaterly Journal of Economics*, 75(4):670–689, 1961.

[42] P.C. Fishburn. THE FOUNDATIONS OF EXPECTED UTILITY. D. Riedel, Dordrecht., 1982.

[43] P.C. Fishburn and A. Rubinstein. TIME PREFERENCE. *International Economic Review*, 23(3):677–694, 1982.

[44] I. Fisher. THE THEORY OF INTEREST *(Reprint)*. Augustus M. Kelly, 1967.

[45] S. French. UTILITY: PROBABILITY'S YOUNGER TWIN? In P.R. Freeman and A.F.M. Smith, editors, *Aspects of Uncertainty, A Tribute to D.V. Lindley*, pages 171–179. John Wiley & Sons, 1994.

[46] S. French and D. Rios Insua. STATISTICAL DECISION THEORY. Kendal's Library of Statistics. Arnold, 2000.

[47] I. Gilboa and D. Schmeidler. MAXMIN EXPECTED UTILITY WITH NON-UNIQUE PRIOR. *Journal of Mathematical Economics*, 18:141–153, 1989.

[48] P.K. Goel and J. Ginebra. WHEN IS ONE EXPERIMENT 'ALWAYS BETTER THAN' ANOTHER? *The Statistician*, 52(4):515–537, 2003.

[49] M. Goldstein. THE PREVISION OF A PREVISION. *Journal of the American Statistical Association*, 78(384):817–819, 1983.

[50] M. Goldstein. TEMPORAL COHERENCE. In J.M. Bernardo, D.V. Lindley, and A.F.M. Smith, editors, *Bayesian Statistics 2*, pages 231–248. North-Holland, 1985.

[51] J.P. Gould. RISK, STOCHASTIC PREFERENCE, AND THE VALUE OF INFORMATION. *Journal of Economic Theory*, 8:64–84, 1974.

[52] I. Hacking. THE EMERGENCE OF PROBABILITY. Cambridge University Press, 1975.

[53] C.M. Harvey. VALUE FUNCTIONS FOR INFINITE-PERIOD PLANNING. *Management Science*, 32(9):1123–1139, 1986.

[54] I.N. Herstein and J. Milnor. AN AXIOMATIC APPROACH TO MEASURABLE UTILITY. *Econometrica*, 21:291–297, 1953.

[55] W. Hlawitschka. THE EMPIRICAL NATURE OF TAYLOR-SERIES APPROXIMATIONS TO EXPECTED UTILITY. *The American Economic Review*, 84(3):713–719, 1994.

[56] B. Houlding and F.P.A. Coolen. SEQUENTIAL ADAPTIVE UTILITY DECISION MAKING FOR SYSTEM FAILURE CORRECTION. *Journal of Risk and Reliability. Proceedings of the IMech E PART O*, 221(4):285–295, 2007.

[57] N.E. Jensen. AN INTRODUCTION TO BERNOULLIAN UTILITY THEORY: I UTILITY FUNCTIONS. *Swedish Journal of Economics*, 69(3):163–183, 1967.

[58] J.B. Kadane, M.J. Schervish, and T. Seidenfeld. RETHINKING THE FOUNDATIONS OF STATISTICS. Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge University Press, 1999.

[59] J.B. Kadane, M.J. Schervish, and T. Seidenfeld. IS IGNORANCE BLISS? Submitted to *JASA Theory and Decision*, 2007.

[60] D. Kahneman and A. Tversky. PROSPECT THEORY: AN ANALYSIS OF DECISIONS UNDER RISK. *Econometrica*, 47(2):263–292, 1979.

[61] M. Kaplan. DECISION THEORY AS PHILOSOPHY. Cambridge University Press, 1996.

[62] E. Karni. A DEFINITION OF SUBJECTIVE PROBABILITIES WITH STATE-DEPENDENT PREFERENCES. *Econometrica*, 61:187–198, 1993.

[63] R.L. Keeney. MULTIPLICATIVE UTILITY FUNCTIONS. *Operations Research*, 22:22–34, 1972.

[64] R.L. Keeney and H. Raiffa. DECISIONS WITH MULTIPLE OBJECTIVES: PREFERENCES AND VALUE TRADEOFFS. Wiley, 1976.

[65] M. Kilka and M. Weber. WHAT DETERMINES THE SHAPE OF THE PROBABILITY WEIGHTING FUNCTION UNDER UNCERTAINTY? *Management Science*, 47(12):1712–1726, 2001.

[66] A.N. Kolmogorov. FOUNDATIONS OF THE THEORY OF PROBABILITY *(English Translation)*. Chelsea Publishing Company, 1950.

[67] T.C. Koopmans. STATIONARY ORDINAL UTILITY AND IMPATIENCE. *Econometrica*, 28(2):287–309, 1960.

[68] I. Levi. ON INDETERMINATE PROBABILITIES. *Journal of Philosophy*, 71(13):391–418, 1974.

[69] I. Levi. RATIONALITY UNBOUND. In W. Sieg, editor, *Acting and Reflecting*, pages 211–221. Kluwer Academic Press, 1990.

[70] D.V. Lindley. A Class of Utility Functions. *The Annals of Statistics*, 4(1):1–10, 1976.

[71] D.V. Lindley. Making Decisions. London: Wiley, second edition, 1985.

[72] R.D. Luce and H. Raiffa. Games and Decisions: Introduction and Critical Survey. Dover, 1957.

[73] R.D. Luce and D. von Winterfeldt. What Common Ground Exists for Descriptive, Prescriptive, and Normative Utility Theories? *Management Science*, 40(2):263–279, 1994.

[74] H.M. Markowitz. Portfolio Selection. John Wiley and Sons, New York, 1959.

[75] K. McDaid and S.P. Wilson. Deciding How Long to Test Software. *The Statistician*, 50:117–134, 2001.

[76] R.F. Meyer. Preferences over Time. In R.L. Keeney and H. Raiffa, editors, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, chapter 9, pages 473–514. John Wiley & Sons, 1976.

[77] R.F. Meyer. State-Dependent Time Preference. In D.E. Bell, R.L. Keeney, and H. Raiffa, editors, *Conflicting Objectives in Decisions*, chapter 11, pages 232–244. John Wiley & Sons, 1977.

[78] N. Morali and R. Soyer. Optimal Stopping in Software Testing. *Naval Research Logistics*, 50(1):88–104, 2003.

[79] D.G. Morrison. On the Consistency of Preference in Allais' Paradox. *Behavioural Science*, pages 373–383, 1967.

[80] H. Moskowitz, P.V. Preckel, and A. Yang. Decision Analysis with Incomplete Utility and Probability Information. *Operations Research*, 41:864–879, 1993.

[81] R. Nau. The Shape of Incomplete Preferences. In *Proceedings in Informatics 18*, pages 408–422. ISIPTA, Carleton Scientific, 2003.

[82] A. Papoulis. Probability, Random Variables and Stochastic Processes. McGraw-Hill, 1991.

[83] J.W. Pratt. RISK AVERSION IN THE SMALL AND IN THE LARGE. *Econometrica*, 32:122–136, 1964.

[84] H. Raiffa and R. Schlaifer. APPLIED STATISTICAL DECISION THEORY. Division of Research, Graduate School of Business Administration, Harvard University, Boston, 1961.

[85] F. Ramsey. A MATHEMATICAL THEORY OF SAVINGS. *Economic Journal*, 38:543–59, 1928.

[86] F.P. Ramsey. THE FOUNDATIONS OF MATHEMATICS AND OTHER ESSAYS. Routledge and Kegan Paul, London., 1931.

[87] J. Rawls. A THEORY OF JUSTICE. *Clarendon Press, Oxford*, 1972.

[88] M. Rothschild and J. Stiglitz. INCREASING RISK: I. A DEFINITION. *Journal of Economic Theory*, 2:225–243, 1970.

[89] L.J. Savage. THE FOUNDATIONS OF STATISTICS. John Wiley & Sons, 1954.

[90] M.J. Schervish, T. Seidenfeld, and J.B. Kadane. STATE-DEPENDENT UTILITIES. *Journal of the American Statistical Association*, 85:840–847, 1990.

[91] M.J. Schervish, T. Seidenfeld, J.B. Kadane, and I. Levi. EXTENSIONS OF EXPECTED UTILITY AND SOME LIMITATIONS OF PAIRWISE COMPARISONS. In *Proceedings in Informatics 18*, pages 496–510. ISIPTA, Carleton Scientific, 2003.

[92] D. Schmeidler. SUBJECTIVE PROBABILITY AND EXPECTED UTILITY WITHOUT ADDITIVITY. *Econometrica*, 57(3):571–587, 1989.

[93] T. Seidenfeld, M.J. Schervish, and J.B. Kadane. A REPRESENTATION OF PARTIALLY ORDERED PREFERENCES. *Annals of Statistics*, 23(6):2168–2217, 1995.

[94] A. Sen. SOCIAL CHOICE THEORY: A RE-EXAMINATION. *Econometrica*, 45(1):53–89, 1977.

[95] R.D. Shachter. PROBABILISTIC INFERENCE AND INFLUENCE DIAGRAMS. *Operations Research*, 36(4):589–604, 1988.

[96] C.E. Shannon and W. Weaver. THE MATHEMATICAL THEORY OF COMMUNICATIONS. Urbana, Il: University of Illinois Press, 1949.

[97] H.A. Simon. A BEHAVIORAL MODEL OF RATIONAL CHOICE. *Quarterly Journal of Economics*, 69:99–118, 1955.

[98] R.H. Strotz. MYOPIA AND INCONSISTENCY IN DYNAMIC UTILITY MAXIMIZATION. *The Review of Economic Studies*, 23(3):165–180, 1955 - 1956.

[99] A. Tversky and D. Kahneman. ADVANCES IN PROSPECT THEORY - CUMULATIVE REPRESENTATION OF UNCERTAINTY. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.

[100] E. von Böhm-Bawerk. CAPITAL AND INTEREST VOLUME 2: POSITIVE THEORY OF CAPITAL *(English Translation)*. Libertarian Press, 1959.

[101] J. von Neumann and O. Morgenstern. THEORY OF GAMES AND ECONOMIC BEHAVIOUR *(Second Edition)*. Princeton University Press, 1947.

[102] P. Wakker and A. Tversky. AN AXIOMATIZATION OF CUMULATIVE PROSPECT THEORY. *Journal of Risk and Uncertainty*, 7(2):147–175, 1993.

[103] A. Wald. STATISTICAL DECISION FUNCTIONS. John Wiley & Sons, 1950.

[104] P. Walley. STATISTICAL REASONING WITH IMPRECISE PROBABILITIES. Chapman and Hall, 1991.

[105] P. Weller. CONSISTENT INTERTEMPORAL DECISION-MAKING UNDER UNCERTAINTY. *Warwick Economic Research Papers*, 96, 1976.

[106] C.C. White. SEQUENTIAL DECISION MAKING UNDER UNCERTAIN FUTURE PREFERENCE. *Operations Research*, 32(1):148–168, 1984.

[107] P.M. Williams. NOTES ON CONDITIONAL PREVISIONS. *International Journal of Approximate Reasoning*, 44(3):366–383, 2007.

[108] H.S. Witsenhausen. ON THE UNCERTAINTY OF FUTURE PREFERENCES. *Annals of Economic and Social Measurement*, 3(1):91–94, 1974.

[109] M.E. Yaari. UNCERTAIN LIFETIME, LIFE INSURANCE, AND THE THEORY OF THE CONSUMER. *The Review of Economic Studies*, 32(2):137–150, 1965.