Decision-Theoretic Planning of Clinical Patient Management

Besliskundig plannen van klinische patiëntenzorg

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. Dr. H.O. Voorma,
ingevolge van het besluit van het College voor Promoties
in het openbaar te verdedigen
op maandag 30 oktober 2000 des namiddags te 14:30 uur

door

Niels Bastiaan Peek

geboren op 17 april 1970, te Eindhoven

promotor: Prof. Dr. J.-J.Ch. Meyer
Faculteit Wiskunde en Informatica, Universiteit Utrecht
co-promotor: Dr. P.J.F. Lucas
Department of Computing Science, University of Aberdeen

ISBN 90–939–2547–2

SIKS Dissertation Series No. 2000–7.

The research in this thesis has been carried out under the auspices of SIKS, the Dutch Graduate School for Information and Knowledge Systems. The investigations were (partly) supported by the Netherlands Computer Science Research Foundation with financial support from the Netherlands Organisation for Scientific Research (NWO).

Contents

CHAPTER 1

Introduction

When a doctor is treating a patient, he[1] is constantly facing decisions. From the externally visible signs and symptoms he must establish a hypothesis of what might be wrong with the patient. He must decide whether the signs and symptoms provide ample evidence for this hypothesis, or whether it should be further verified by diagnostic testing. When the doctor is sufficiently certain which disease the patient has, he must decide whether treatment is necessary, and if so, which treatment is most suitable. The timing of treatment may also be critical: sometimes the condition of the patient requires immediate action. In other cases such urgency is not required and it may be beneficial to postpone risky interventions until the overall condition of the patient has improved. After the treatment, the patient's response to it must be evaluated. Has the treatment been successful, or is further remedial action required? If the treatment has not been successful, then maybe the initial hypothesis on the underlying disease should be reconsidered. If the treatment has been successful, what are the chances that the disease recurs in the future? Should the patient return for investigations at some point in the future?

All these bedside decisions are not only unavoidable but must also be made under conditions of uncertainty. The uncertainty arises from several sources, such as errors in the results of diagnostic tests, limitations in medical knowledge, and unpredictability of the future course of disease. Furthermore, clinical decisions are often highly interrelated and can therefore not be made in isolation. For instance, when choosing a diagnostic test to verify a hypothesis, it should be assessed whether such testing

[1] Anywhere we use the masculine form, the feminine form is understood as being included.

will improve the ability to choose the right treatment. If only few treatment modalities are available, then it is probably superfluous to conduct additional tests, as the choice of treatment would not be influenced by the test result. Similarly, the choice of treatment itself should take into account the possibility that the hypothesis is wrong and the patient has a different disease. It is then advantageous to choose a treatment whose consequences are amenable for correction in the future.

We will use the term *patient management* to refer to all types of decision making in situations faced by a doctor at the bedside. Decision making under uncertainty is traditionally studied in the field of *decision theory*, the mathematical theory of rational choice. As patient management involves making a large number of interrelated decisions, it can also be viewed as a form of *action planning*, traditionally a branch of *artificial intelligence* (AI). Within this view, each decision consists of a choice between multiple clinical actions, and the sequence of actions chosen has the objective of reaching a satisfactory health condition of the patient. The optimal approach to solving the problem is formulating a detailed program of action, or a *plan*. In this thesis, we study the problem of clinical patient management as action planning using decision-theoretic principles, or *decision-theoretic planning* for short.

In the following sections, we will elaborate on this conception of clinical patient management, and put it in perspective of the tradition of decision support in clinical medicine. We conclude with listing the main research questions that motivate our work, and by giving an overview of this thesis.

1.1 Supporting medical decisions

In both decision theory and AI, there exists an extensive literature on clinical-medical reasoning and problem solving; the conception of clinical decision making as a form of action planning, however, is not often adopted. In this section, we review the standard procedure for clinical patient management that is followed in today's clinical practice, and discuss the decision-theoretic and AI approaches to clinical decision support.

1.1.1 The clinical procedure

When a patient becomes aware of particular symptoms or manifestations of disease, he may decide to visit a doctor. The doctor will then roughly use the following standard procedure, schematically depicted in Figure 1.1. Firstly, he interviews the patient on his medical history and the case history of the present symptoms; this is called *history taking*. The second and third steps consist of *physical examination* of the patient and conducting *routine diagnostic tests*, respectively. The objective is to establish a *differential diagnosis*, a set of possible explanations of the symptoms along with their respective likelihoods. After this step it may be decided to choose a therapy, or to conduct additional, non-routine diagnostic tests when too much

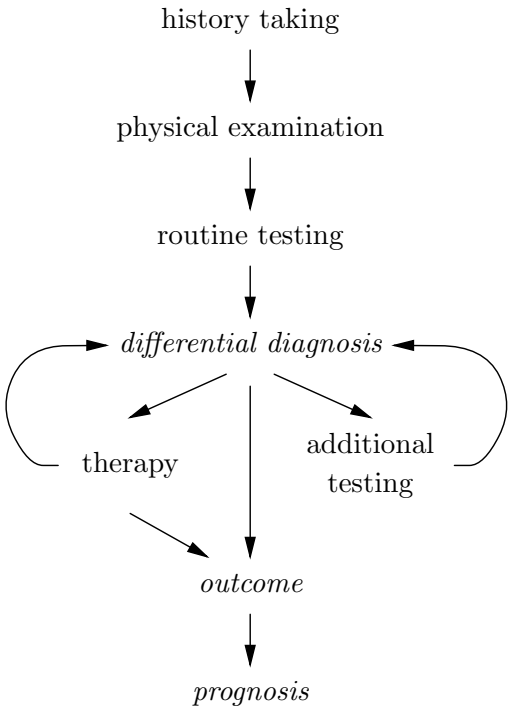


Figure 1.1: The standard clinical procedure.

uncertainty remains on the diagnosis; one may also decide from the diagnosis to forgo further clinical action. Non-routine diagnostic tests are often more intensive than routine tests in that they involve pain or risks to the patient, or are more costly or time-consuming. Therapy may be any collection of remedial actions that relieves or cures the disease. Based on the patient's response to therapy or the results of the tests, this last step may be repeated several times; at each cycle, the differential diagnosis is adjusted by incorporating the new information. The condition of the patient that results afterwards is called the *outcome* of the procedure, and his expected condition after a prolonged period of time (typically a significant number of years) is his *prognosis*. It is customary in many clinical procedures that the treatment has a *follow-up* in the form of regular tests to monitor the patient's condition. Outcome and prognosis are the major evaluation criteria for the sequence of actions that has been undertaken.

Diagnostic tests can be characterised by a number of standards. Here, we will use five test characteristics that are frequently used in all fields of clinical medicine. Firstly, is the test *invasive*, i.e. does it involve insertion of an instrument or foreign material into the body? Secondly, are there any *health risks* involved for the patient? These risks divide into *mortality* (probability of death) and *morbidity* (probability of complications). Thirdly, is the test procedure painful or otherwise burdening for the patient? Fourthly, what are the associated costs, financial or otherwise? Fifth and finally, what is the *reliability* of the test, i.e. what are the chances of obtaining a false

test outcome? Generally speaking, non-invasive, inexpensive and reliable tests that involve little risks and pain for the patients are preferred to other tests. Unfortunately, reliability often comes at the price of invasiveness, which in turn increases health risks, painfulness, and costs.

Medical therapy can be classified in various guises. The first distinction considers the object of the therapy: we distinguish *causal therapy*, which aims to fight the causes of the disease, and *symptomatic therapy*, which aims to suppress its symptoms. A second distinction is between *curative therapy*, which intends to cure the patient completely from the disease and its underlying causes, and *palliative therapy*, which intends to alleviate the patient's suffering or to prolong his duration of life. Palliative therapy is mostly symptomatic but can sometimes be classified as causal. The choice of therapeutic regimen often depends on multiple criteria, but expected outcome and prognosis are often decisive, especially for causal therapy. In comparative studies, outcome is often measured in terms of *survival* and prognosis in terms of *life expectancy*; the preferred therapeutic choice is the one that optimises both the chances of survival and the life expectancy afterwards. Of course, a tradeoff between these objectives is sometimes unavoidable. In addition, one may consider the expected *health status* of the patient: suffering from chronic post-therapeutic disabilities reduces the quality of life and is therefore considered undesirable. A final criterion may be the costs associated with the therapy.

It should be noted that the clinical procedure described so far is a correct but rather simplified description of reality. For ease of understanding we have abstracted from the roles of time and change in the description. Yet, there is often a considerable course of time during the clinical management of individual patients, which may be accompanied by many changes in the patient's situation. The significance of this fact for our conception of clinical management is threefold. First, the human body can usually not be regarded as a passive entity that patiently undergoes a doctor's acts and remains otherwise unaltered. Biological processes influence the patient's condition over time, sometimes resulting in a substantial improvement but equally often causing injurious complications. Second, it is not only the choice, but also the *timing* of clinical actions that is subject to a tradeoff: the information that is obtained from diagnostic procedures and the effectiveness of therapeutic actions may critically depend on their moment of execution. And third, it is often no longer possible to make a clear separation between the different phases of clinical management from Figure 1.1. For instance, it may be required to re-assess the patient's diagnosis multiple times in the process of management. The purpose of diagnostic actions is then not to simply refine one's knowledge of the patient's status, but also, or primarily, to update this knowledge to the current state of affairs.

To summarise the above, we can state that acknowledging the role of time and change in patient management leads to a re-conception of the clinical procedure. It departs the static view on clinical decision making where all acting is concentrated in a single abstract moment of time, and leads to a dynamic perspective where patient

management is regarded as a *process* of ongoing interaction between patient and clinician. This will be the leading perspective on clinical management in this thesis.

1.1.2 Clinical decision science

Decision theory and Artificial Intelligence (AI) emerged from research on systematic methods for problem solving and decision making that blossomed in the 1940s. These disciplines were stimulated by new possibilities for automated reasoning unleashed by the development of the computer. But although the fields had common roots, they soon grew out to become independent traditions. Decision theory favours carefully quantified problem descriptions and the application of mathematical methods to support people in delineated parts of their decision-making tasks; the field is strongly connected to probability theory and statistics. AI distinguishes itself in its concern with autonomous problem solving, its emphasis on symbolic rather than numerical information, and its interest in analogies between computer programs and human thinking.

In this section, we discuss the decision-theoretic approach to clinical decision support. Decision theory has been fruitfully applied to clinical decision-making problems, and this has led to the maturation of a discipline that is now known as *medical decision analysis*. Below, we first review the theoretical and social backgrounds of the medical decision analysis; we will then sketch the standard method of analysing clinical decisions in the field, and finally discuss some problems with the application of this method in practice.

Decision theory

The roots of decision theory are found in the *theory of games* that was formulated in the 1940s by Von Neumann and Morgenstern (1944). The theory of games is based on a mathematical characterisation of rational choice called *utility theory*. The central results of utility theory are that, given a number of assumptions on rational behaviour, decision-making objectives can be expressed as numerical quantities (called *utilities*), and optimal solutions to decision-making problems can be found by numerical maximisation. As such, utility theory provides a firm mathematical basis for decision making under uncertainty. The study of applying utility-theoretic principles to varying situations of choice has become known as decision theory (Chernoff and Moses, 1959; Raiffa and Schlaifer, 1961).[2] In this thesis, we focus on *Bayesian* decision theory, (Savage, 1972), the study of rational choice from given beliefs, as

[2]Sometimes, a distinction is made between *descriptive*, *normative* and *prescriptive* decision theory, where descriptive decision theory is part of the social sciences and studies actual decision-making behaviour, normative decision theory describes idealised decision-making behaviour as derived from the axioms of utility theory, and prescriptive decision theory is concerned with the transformation of actual to idealised decision-making behaviour. In this thesis, we are concerned with normative decision theory.

opposed to *statistical* decision theory, the study of comparative statistical inference from given data, (Barnett, 1982). Bayesian decision theory can be regarded as the synthesis of utility theory and Bayesian probability theory.

Bayesian decision theory identifies three fundamental components in each decision problem: (i) the *alternatives* to choose from, generally in the form of different courses of action or decision-making strategies, (ii) ones *beliefs* with respect to reality, expressed as probabilities on the possible outcomes after choosing an alternative, and (iii) ones *preferences* with respect to these outcomes, expressed as utilities. From a proper specification of these problem components, application of the theory provides for making a rational decision that is consistent with ones beliefs and preferences, and provably leads to the best expected outcome in the given circumstances. Importantly, decision theory distinguishes a preferable decision, i.e. a rational choice that is expected to be optimal in general, from a preferable outcome, i.e. the result of a choice that turns out to be desirable. If the effects of decisions are uncertain, then we cannot hope to always achieve preferable outcomes. Yet, we can always make preferable decisions, thus striving for the best expected outcome.

Evidence-based medicine

The methodology of applying decision theory to real-world clinical decision problems is known as *medical* or (*clinical*) *decision analysis*, (Weinstein and Fineberg, 1980; Pauker and Kassirer, 1987; Sox et al., 1988). Medical decision analysis emerged in the 1970s as a result of the rationalisation of medical practice that has taken place since the 1960s. Power has increasingly slipped from the hands of physicians to those of patients, governments, insurers, and other organisations; and simultaneously, the limitations of medical reasoning were increasingly recognised, (Dowie and Elstein, 1988). Although people are quite good at some complex tasks (e.g., pattern recognition, information synthesis and communication), they tend to fail equally often at others, such as systematic and unbiased observation, and (probabilistic) belief revision given new information. Furthermore, the number of diagnostic and therapeutic choices available to physicians has dramatically increased. The response to this recognition of medical reasoning's limitations was an attempt to formalise medical practice and to look for more rational, scientific ways to make clinical decisions. The leading paradigm that has emerged from this formalisation is called *evidence-based medicine*.

Evidence-based medicine argues that medical decisions should be based on the best available evidence from clinical research, as opposed to anecdote, habit, local practice patterns, and pathophysiological theory, (Sackett et al., 1996). The practice of evidence-based medicine means that individual clinical expertise is integrated with the best available external clinical evidence from systematic research. By best available external clinical evidence is meant clinically relevant research, often from the basic sciences of medicine, but especially from patient centred, systematic clinical research into the accuracy and precision of diagnostic tests, the power of prognostic markers, and the efficacy and safety of therapeutic regimens.

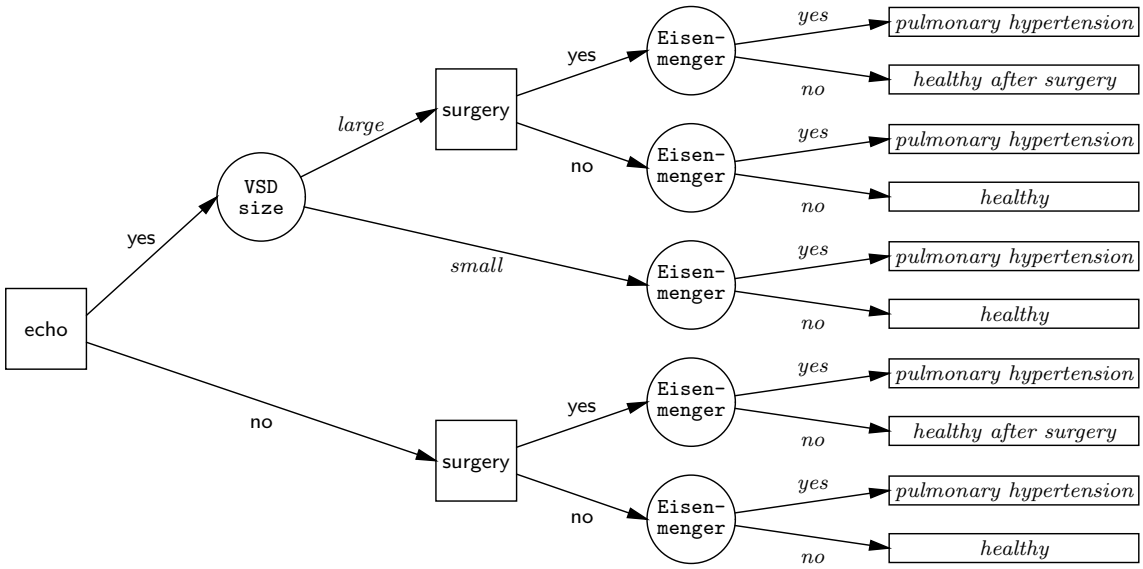


Figure 1.2: A decision tree for the management of patients with VSD.

Decision analysis

The field of medical decision analysis studies the preferred way of incorporating the best available evidence into the everyday practice of clinical patient management. As stated above, decision analysis is the methodology of applied decision theory; in a clinical context, it aims at analysing and supporting all decisions regarding diagnostic and remedial action. The traditional tool that is employed in the field is the *decision tree*, (Raiffa, 1968). Given a decision problem under uncertainty, a decision tree provides an explicit, graphical enumeration of a set of potential decision scenarios. Its form is highly intuitive, and provides for easy communication with physicians.

An example decision tree, modelling the management of patients with the congenital heart disease called *ventricular septal defect* (VSD), is depicted in Figure 1.2. We do not describe the problem domain here as Chapter 2 provides an elaborate description of this disease. For illustration purposes, the VSD management problem has been reduced to two simple yes-or-no decisions in this example: the first being whether or not to make an echocardiographic image, and the second being whether or not to submit the patient to surgery. Both decisions are represented by small squares in the tree, called *decision nodes*. Furthermore, two random variables pertaining to the clinical state of the patient are used in the decision tree: one variable, called `VSD_size`, describing the size of the VSD as seen on an echocardiographic image, and one variable, called `Eisenmenger`, describing occurrence of Eisenmenger's complex, the primary complicatory risk for VSD patients. Random variables are represented by small circles, called *chance nodes*. Both random variables can take two possible values, *small* and *large*, and *yes* and *no*, respectively. Each path in the tree ends with a description of the associated outcome in a rectangular box. In the example, we have distinguished the outcomes *healthy*, *healthy after surgery*, and *pulmonary*

hypertension.

The branches emanating from a decision node correspond to the decision alternatives available to the decision maker at that point in the decision-making process; the branches emanating from a chance node correspond to the possible values of the random variable modelled by the node. With each branch of the latter type, one now associates an estimate of the probability that the random variable in question takes that value under the given circumstances. To each possible outcome of the decision process is assigned a numerical utility value that represents the relative desirability of that outcome. A simple procedure, called *fold-back analysis*, computes the decision strategy that is optimal in decision-theoretic terms. Subsequent *sensitivity analyses* can point out the sensitivity of this strategy to potential errors in the probability estimates, and to assumptions regarding tree structure and utility assignments, (Habbema et al., 1990).

The problem of implementation

The results of a decision analysis are usually not implemented straight out in clinical practice. Instead, making an analysis has the objective of gaining more insight in the problem domain in general, and of improving the strategy for clinical management as a whole. Such improved strategies are then often formalised in *clinical guidelines* that are to be followed by all clinicians, to ensure uniform high quality of medical care. Nevertheless, the usefulness of analysing medical decisions from decision-theoretic principles has been subject to debate over the years, (e.g., see Hershey and Baron, 1987); we discuss two important arguments in this debate.

First, researchers in psychology have shown that people, including physicians, often do not follow decision-theoretic criteria when they make decisions (Kahneman et al., 1982). In fact, it has been demonstrated that people exhibit stereotypical deviations or *biases* from the criteria and their underlying assumptions. When reasoning with uncertain information, people tend not to adhere to the rules of probability theory. In clinical medicine for instance, many physicians tend to attach more importance to the fact that a patient fits the classic description of a disease than the high improbability of that disease due to its low prevalence, (Sox et al., 1988). Also, diagnostic procedures are often employed to gather additional information that is consistent but redundant; such procedures could therefore be omitted.

The fact that physicians normally do not conform to the criteria of decision theory in their reasoning creates a tension when decision-analytic methods are applied to clinical problems. Decision theory is normative in nature, and deviations from decision-theoretic criteria are therefore classified as erroneous. But most experienced physicians will be hard to convince that they are frequently making errors. Similarly, the formalisation of preferred management strategies in clinical guidelines is seldom effective: it has also been found that clinicians rarely use guidelines, no matter how well-written and how evidence-based they are, (Berg, 1997).

A second problem with the application of decision analysis in clinical practice concerns the description of decision-making objectives. It is often felt that numerical utility values fall short in expressing the difficult tradeoffs that are involved in medical care. Furthermore, it is unclear who should assess these utilities. When the doctor is to do so, he must decide on highly subjective matters for his patient. When the patient is to do it, he must rate and value conditions which he has never experienced. In response to an illness, patients have been found to commit themselves to choices that they regret in retrospect, (Hilden et al., 1992).

1.1.3 Artificial intelligence in medicine

Soon after its inception, the digital computer started to inspire computer scientists to design automated tools that could support medical workers in clinical judgement and decision making. This has to the specialist field of *artificial intelligence in medicine* as a branch of AI.

A key motivation for research in AI has always been the psychology of human problem solving. This motivation was partly inspired by a paradigmatic shift in psychology in the 1950s, where the leading behaviourist approach was traded for a *cognitive* approach that emphasises symbolic information processing and problem solving; in cognitive psychology, the primary metaphor for the human brain is the digital computer. Correspondingly, researchers in AI favoured symbolic rather than numerical problem representations; decision-theoretic and probabilistic reasoning were considered inadequate models of human problem solving, (e.g., see Simon, 1955). An additional reason for rejecting the decision-theoretic approach was the computational complexity of probabilistic inference: automated reasoning with uncertainty in a mathematically-sound manner was conceived to be impossible. Alternative reasoning strategies and heuristic methods for decision making were therefore preferred in AI.

Also in artificial intelligence in medicine, the usage of probability and decision theory has been varying. The approach was abandoned in the 1970s due to disappointing results; since the late 1980s however, decision theory has regained interest from workers in the field as a result from progress in the field of probabilistic representation and reasoning. Below, we give a brief historical overview of the intelligent decision-support systems in medicine, highlighting the main developments that relate to our work.

Diagnosis with naive Bayesian models

The earliest computer systems for clinical assistance focused on the problem of diagnostic assessment, and used a naive Bayesian model ("idiot's Bayes") to infer the most likely diagnosis from a given set of findings, (e.g. Warner et al., 1961). A naive Bayesian model is a simple probabilistic model where disease types and potential findings are represented as statistical variables, and it is assumed that the findings

are mutually independent if the patient's disease status is known. The broader problem of diagnosis as both diagnostic inference and diagnostic test selection was first considered by Gorry and Barnett (1968). In their system, diagnostic inference is also performed from a naive Bayesian model, while test selection is based on heuristic functions describing the cost of conducting diagnostic tests and the penalties of misdiagnosing. In several experimental applications, it was found that the system yielded a sharp reduction in the average number of tests performed, while establishing a diagnosis at expert level.

Noteworthy is also the work by F. de Dombal and his colleagues, (Horrocks et al., 1972; de Dombal et al., 1972; de Dombal et al., 1975), who developed a naive Bayesian model for the diagnosis of acute abdominal pain: lacking highly specific tests, clinical diagnosis of this complaint (which may be a sign of appendicitis) is a notorious problem in emergency medicine. In a series of studies in the early 1970s, De Dombal and his colleagues devoted themselves to delivering diagnostic assistance by the computer in an extremely practical format; this allowed to gain experience with the usefulness of such systems in routine clinical practice. In a controlled prospective trial, it was found that the system's diagnostic accuracy was significantly higher than that of the most senior member of the clinical team (91.8% versus 79.6%).

Notwithstanding these promising results, the probabilistic approach to diagnostic assistance was departed in the 1970s. The assumption of mutual independence in naive Bayesian models is often unsound, and may thus impede diagnostic accuracy. Yet, probabilistic reasoning in general is highly combinatorial, and was therefore held to be infeasible without the assumption of mutual independence. Furthermore, the approach did not seem to leave room for incorporating the knowledge of clinical experts. And finally, naive Bayesian models do not increase one's understanding of the domain under consideration, and are therefore soon disregarded by clinicians. For these reasons, symbolic reasoning prevailed over probabilistic reasoning for almost two decades; the predominant approach to medical decision support was within the *expert system* paradigm.

Expert systems

An *expert system* is a computer program capable of giving some sort of reasoned guidance on a small set of closely related problems. Research on expert systems arose as a subfield of AI in the 1970s from the conviction that general reasoning and problem solving by computer systems is probably an aim too high, but that satisfying results may be obtained by restricting to tightly-delineated problem domains. In these restricted settings, there was no need to mimic poorly-understood human problem-solving methods such as analogical reasoning and "common sense" inference; an expert system might be able to compete with human experts in the domain under consideration. A traditional field of application for expert systems is clinical medicine; medical specialities are thought to provide precisely the delineated problem domains where these systems excel.

One of the earliest and most influential expert systems developed was MYCIN, an advisory system on the diagnosis and treatment of the infectious diseases septicaemia and meningitis, (Shortliffe, 1976). The core of MYCIN is formed by a collection of production rules that express causal and associational relationships in the problem domain. The rules are processed by an inferential algorithm in response to queries by the user. It was shown in the MYCIN project that these techniques could be generalised for use in other domains through a separation of domain knowledge and problem-solving inference. This type of separation has become a key characteristic of expert systems: a *knowledge base* provides a declarative specification of all domain knowledge, and a generic *inference engine* responds to the user's queries by selecting relevant items from the knowledge base and employing them in systematic reasoning. Domain knowledge may be obtained from several sources, but often originates from human experts. Several other expert systems were developed by means of the EMYCIN (Essential MYCIN) system, which was obtained from MYCIN by removing domain knowledge from the system. At present, programs such as EMYCIN are known as *expert system shells*; the success of the EMYCIN system inspired many others to develop similar systems.

An important topic in expert system design is the choice of *knowledge-representation* formats. Many different representation formats have been used in expert systems or suggested for that purpose, but a substantial number of them lacked a solid mathematical basis. Therefore, only few types of knowledge representation have remained over the years to be used and re-used in many systems; apart from the most popular production-rule representation of MYCIN and EMYCIN, these are the *frame-based representation* (Minsky, 1975) and first-order predicate logic.

We describe two more influential medical expert systems, as they relate most closely to our work. CASNET (Weiss et al., 1978) is a system for assisting clinicians in the diagnosis and long-term treatment of glaucoma. It is based on a representation that models disease processes as a network of causal-associational relationships. The CASNET system incorporates many features that are relevant for clinical reasoning that were overlooked in other systems: the construction of a differential diagnosis, the representation of severity and progression of disease, and the possibility to reason with therapy and outcome.

ONCOCIN (Shortliffe et al., 1981) is a chemotherapy protocol advisor that combines rules from clinical guidelines with judgements of oncologists who have experience adjusting therapy in complex clinical situations. An evaluation showed that ONCOCIN provides advice on lymphoma treatment similar to the treatment provided in a university oncology clinic, (Hickam et al., 1985). However, the system had difficulties with anticipating problems with therapy, as this requires reasoning with uncertain and incomplete information. Therefore, a successor system, called ONYX (Langlotz et al., 1987), was designed to handle these difficult planning problems; it combines decision-theoretic and AI approaches to planning. The system's planning procedure consists of three steps: (i) the use of production rules to generate a small set of

plausible plans, (ii) the use of physiological knowledge to simulate the consequences of each plan for the patient, and (iii) the use of decision theory to rank the plans according to how well the results of each simulation meet the treatment objectives.

Probabilistic networks

The knowledge-representation formats that were developed in the field of expert systems were symbolic rather than numerical in nature, and usually aimed at forms of categorical inference. When a fully deterministic perspective on the problem domain was inadequate, attention focused on qualitative ways of handling uncertainty such as non-monotonic logic (Reiter, 1987), or on novel numerical schemes such as fuzzy logic (Zadeh, 1965), certainty factors (Shortliffe and Buchanan, 1975), or Dempster-Shafer belief functions (Shafer, 1976). Many researchers in the field held probability to be inadequate as a representation of uncertain knowledge, and probabilistic inference, i.e. the calculation of probabilities of interest, was considered computationally infeasible.

While research into the alternative representations of uncertainty continued to be pursued, probability theory made a comeback. It was perceived that exploitation of conditional independence assumptions, implicit in the qualitative structure of expert knowledge, might reduce the problem of specification and evidence propagation to a feasible level. The branch of *graphical models* (Whittaker, 1990) in statistics emphasised such qualitative structure over quantitative specification, and provided a graphical representation of conditional independence statements. The major breakthrough, halfway through 1980s, consisted of the development of methods for efficient probabilistic inference with graphical representations, and yielded the formalism of *belief networks*, (Kim and Pearl, 1983; Lauritzen and Spiegelhalter, 1988; Pearl, 1988).

Ever since, belief networks have enjoyed a rapid increase in popularity, and are now the most widely-used formalism for handling quantified uncertainty in AI. Numerous expert systems based on belief networks have been developed, with a substantial number of applications in medicine (e.g. Andreassen et al., 1987; Andreassen et al., 1998; Bellazzi et al., 1991; Middleton et al., 1991; Heckerman et al., 1992). A renewed interest in decision theory has emerged, motivated by the new possibilities to perform automated probabilistic reasoning. It was in fact found that a straightforward decision-theoretic extension to belief networks exists in the form of *influence diagrams*, a representation that had concurrently been developed in the field of decision analysis, (Howard and Matheson, 1981; Shachter, 1986). As such, influence diagrams can be regarded as a true synthesis of the symbolic and numerical reasoning methods advocated by AI and decision theory.

1.2 The dynamic perspective on patient management

It was earlier described that time and change often play a significant role in clinical situations; we concluded that patient management is therefore best regarded as a

process of ongoing interaction between patient, disease, and clinician. Within this process, the clinician faces the task of making a large number of interrelated decisions over time, where each decision consists of a choice between multiple clinical actions. The management task can thus be viewed as a form of action planning, where not only the choice, but also the timing of clinical actions is important.

Although it was already noted by Gorry (1968; 1973) that a doctor can be thought of as solving a sequence of similar decision problems, the dynamic perspective on patient management is rarely found in the literature on medical decision support. Similarly, the role of time and change in clinical reasoning has been largely neglected. Most decision-support systems are characterised by a static perspective on the clinical situation where time is left implicit, (Kahn, 1991). The crucial role of temporal reasoning for medical decision support systems is increasingly recognised, (Soper et al., 1991; Aliferis et al., 1997; Peek, 1999a; Shahar, 1999), and there are experimental indications that temporal reasoning can indeed improve the accuracy of medical decision-support systems; such was found, for instance, in the Heart Disease Program, (Long et al., 1986; Long, 1996). Yet, it is hard to assess which type of temporal representation is adequate for a given clinical domain, and even harder to design an associated inference mechanism that correctly, and accurately, implements temporal reasoning.

In this section, we elaborate on the conception of patient management as a form of action planning under uncertainty. We will first, in Subsection 1.2.1, review the field of planning in AI, where special attention is given to the increased interest in planning under uncertain conditions that has emerged in the last decade. Decision-theoretic planning is the synthesis of this type of planning with principles from decision theory. We draw up an inventory of the abstract concepts that occur in decision-theoretic planning in Subsection 1.2.2.

1.2.1 Action planning in AI

Early in the growth of the field of AI it was recognised that an important behaviour for any intelligent system was the ability to plan a set of actions to accomplish its goals. The attempt to realise programs with this ability has resulted in one of AI's main sub-disciplines – the field of *planning*. Since the first published papers on planning in the late 1950s, this field has grown and now comprises a literature that is quite diverse, as planning is related to many other subareas of AI. It is also the field that has raised the most challenging problems from both practical, theoretical, and philosophical standpoints.

The traditional approach to planning in AI employs a strictly symbolic language (e.g. predicate logic) in the formalisation of planning tasks and makes several assumptions on the type of problem that is to be solved, (McCarthy and Hayes, 1969; Fikes and Nilsson, 1971). To be specific, it is assumed that if an agent performs actions within a planning environment, the effects of these actions can be predicted with certainty,

and are moreover known to the agent. Furthermore, there exist particular states of the environment that are unambiguously preferred over others and therefore represent the goal of the planning effort: the objective is to reach such a state while minimising the number of actions. A sequence of actions that satisfies this objective is then called an *optimal plan*. A common additional assumption is that the agent knows the initial state of the environment; there is no hidden information.

Most of the research within the traditional planning approach has focused on possible ways to formalise notions such as action, time, and change, (e.g. McDermott, 1982; Allen, 1984; Lifschitz, 1987). Formal reasoning about the effects of actions over time has turned out to be notorious in the generation of philosophical problems such as the *frame problem* (McCarthy and Hayes, 1969), the *qualification problem* (McCarthy, 1980), and the *ramification problem* (Finger, 1987). But even when we disregard these problems and restrict the potential effects of actions by prior assumptions, solving planning problems is computationally intractable as inspection of an exponentially growing number of action sequences may be required to find a goal state. Nevertheless, satisfying results are sometimes obtained by heuristically searching the state space.

The last 10 to 15 years have shown a growing recognition that the assumptions of traditional planning research are unrealistic in most real-world settings, and the results of the research efforts could therefore hardly ever be put to practice. An increasing interest has emerged for *planning under uncertainty*, that is, planning in environments where the effects of actions are nondeterministic, and where part (or all) of the state of the environment is unknown to the planning agent. These conditions have important consequences for the notion of plan, and therefore also for the construction of (optimal) plans, as they require that planned behaviour is adjusted dynamically during plan execution on the basis of observations. Alternatively, a conditional planning structure must be developed that anticipates on all possible situations that may be encountered during plan execution; we speak of a *contingency plan*. Several planning systems have been described in the literature where a traditional approach is adjusted to handle various types of uncertainty, (Peot and Smith, 1992; Draper et al., 1993; Kushmerick et al., 1995).

Another traditional assumption that has been criticised is that there exist goal states, states of the environment that are unambiguously preferred over others. It was argued that a planning agent often has multiple, possibly competing planning objectives, each of which is only partially satisfiable. And even if goal states can be identified, they have to be approached differently under conditions of uncertainty as there will generally not exist plans that guarantee such states to be reached. Therefore, part of the planning task is to also make a tradeoff between multiple objectives and to formulate a planning criterion that takes uncertainties into account.

In recent years, it was recognised that action planning under uncertainty can be regarded as the task of controlling a stochastic process over time, a type of problem that has traditionally been within the realm of control theory, (Dean and Well-

man, 1991; Bertsekas, 1995). Furthermore, multi-attribute utility theory (Keeney and Raiffa, 1976), a specialisation of utility theory, was found to provide adequate ways of incorporating multiple objectives in choices under uncertainty. The synthesis of these ideas yields a form of planning that can be characterised as planning under uncertainty using decision-theoretic principles, or *decision-theoretic planning* for short.

1.2.2 Decision-theoretic planning

In this thesis, we study decision-theoretic planning as a formalisation of the dynamic perspective on clinical patient management. There are three conditions that motivate this choice of formalisation. First, this formalisation is rooted in well-established mathematical theories with a clear, formal semantics, and a large body of known formal properties. Second, the numerical basis of decision theory accords with the quantitative tradition in evidence-based medicine. The decision-theoretic approach allows us to use statistical data from clinical trials, improving the reliability of recommended decisions and increasing the chances of acceptance in clinical environments. And third, decision theory requires that knowledge encoded in a model or knowledge base be self-consistent; contradictions in the knowledge will be found at the time of building the model instead of while employing it.

We conclude this section with a conceptual framework for decision-theoretic planning, where we identify the principal ingredients that occur in this type of reasoning. This conceptual framework will serve as a starting-point for a formalisation of the task in subsequent chapters of the thesis; we describe it in terms of its five principal ingredients.

First of all, the notion of action planning assumes the existence of *actions* from which choices are to be made. A specific choice of action will be called a *decision*; instead of actions we will therefore also speak of *decision alternatives* or simply *alternatives*. A person or artificial agent charged with the responsibility of choosing between decision alternatives is called a *decision maker*. In the general case, multiple decision makers may be involved in an action planning task, but in this thesis we will take the perspective of the individual clinician who is treating a patient; we will therefore speak of *the* decision maker. We will assume that the object of decision making, i.e. the environment or part of reality to which it pertains, can be viewed as a *system*, with specific behaviour that is described in terms of *system states*. In turn, these states are characterised in terms of a number of *attributes*. In a medical context, we identify the system under consideration with a patient who receives treatment, where the states correspond to different conditions of that patient. We further assume that the decision maker has limited information about the precise state of the system. This relates to the fact that clinical settings limit the potential of knowing the patient's precise state; it is only observable in the form of symptoms, signs, and diagnostic measurements. We therefore distinguish between the actual system state and the

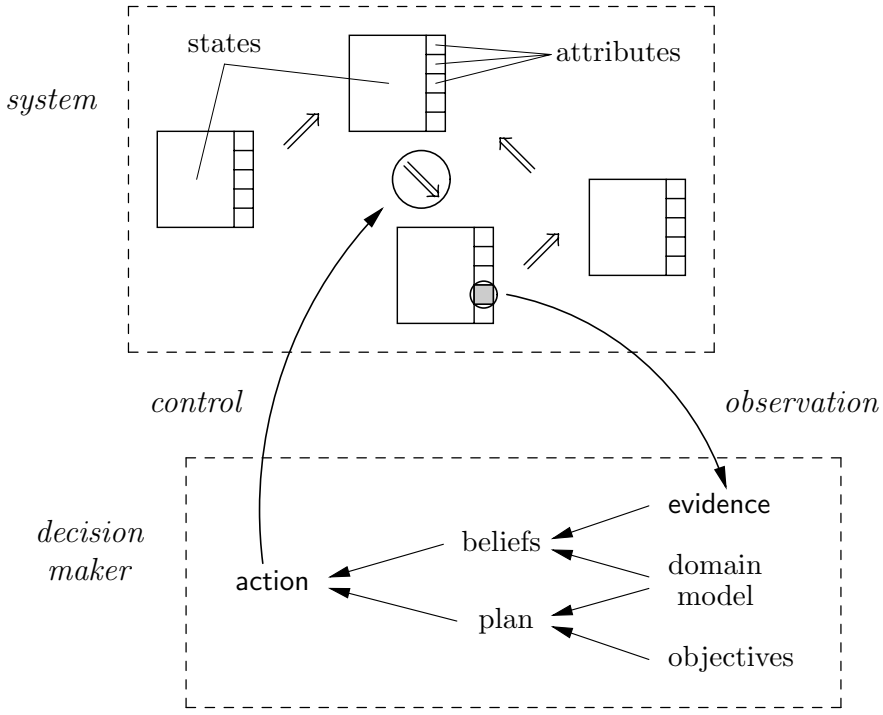


Figure 1.3: Decision maker controlling a partially-observable system by action planning.

decision maker's limited information thereof, to which we refer as his *knowledge*.

An important characteristic of the system at hand is its dynamic nature: the system's state is subject to change over time. The second ingredient that needs distinction is therefore the *time frame* for the planning task, and the potential changes in the system that may accompany temporal progression. We identify two types of change: *endogenous* and *exogenous change*. Endogenous changes are alterations to the system's state that occur independent of external intervention by the decision maker; in clinical terms, this may be regarded as the natural history of disease. Exogenous changes, in contrast, are induced by actions chosen by the decision maker; these changes are typically regarded as the effects of clinical interventions such as surgery. We assume that both types of change are, in general, nondeterministic, and can therefore not be predicted with certainty. We recall that in the framework of decision-theoretic planning, this type of action planning is regarded as the task of controlling a stochastic process over time.

The third ingredient of our framework concerns the informational relation between system and decision maker. We will assume that the decision maker has a *model* of the domain that describes the system and the expected (yet uncertain) effects of performing actions on the system state. Furthermore, actions may yield the opportunity to observe parts of the system state. We refer to such observations as *evidence*; the evidence that is collected over time constitutes the decision maker's knowledge. From the model and the evidence obtained, the decision maker develops an opinion

on unobserved parts of the system. Note that we do not take the model to be a correct description of reality: it may be inaccurate or incomplete. The decision maker's opinion is therefore subjective in nature; to emphasise the inherent subjectivity, we will rather speak of *beliefs* than of opinion. In line with our choice to use Bayesian decision theory in the formalisation, we assume that the decision maker's beliefs are expressed as probabilities.

The fourth concept that needs distinction is the *objective* of planning. Generally speaking, the motivation to select actions will be that certain system states are preferred over others, and that we hope to 'move' the system state in the preferred 'direction' by appropriate action choices. Of course, there may be multiple, competing objectives that require a tradeoff; these competing objectives are then expressed in terms of different attributes of the system. We assume all preferences and objectives to be quantifiable using the principles of utility theory.

Fifth and finally, we distinguish well-posed *problems* and *solutions*. A decision-theoretic planning problem usually consists of a decision-making situation, expressed in terms of all available information such as case-specific parameters and past observations and decisions. When the first decision is to be made, this will be limited to case-specific parameters. In addition, there may be constraints on the possible choices for the current or future decisions. As planning decisions can often not be made in isolation, a well-posed solution should consist of a contingency plan that decides on the current choice and anticipates on possible situations that may be encountered in the future, while taking into account possible constraints that have been formulated as part of the problem.

In Figure 1.3, we have illustrated our conceptual framework of planning under uncertainty. It depicts the ingredients that were identified above, as well as the mutual relationships that exist between these ingredients and their role in the planning task.

1.3 Overview of this thesis

We will now summarise the preceding sections and give an overview of the chapters to come. In this thesis we employ a dynamic perspective on clinical patient management, where doctor, patient, and disease engage in a process of continual interaction: the doctor responds to observed signs, symptoms, and results of diagnostic procedures by taking appropriate clinical action, and the patient's condition changes over time in response to the doctor's actions. Within this process, the tasks of diagnostic assessment, therapy selection, and prognostication are intertwined activities, and do not form separate phases in the management procedure. The doctor is viewed as solving a sequence of similar but mutually related decision problems over time; in abstract terms, this task is characterised as action planning under uncertainty with partial information and temporal constraints. As we employ decision theory at the fundamental level of trading off alternative choices, the overall management task is

characterised as a form of decision-theoretic planning.

The primary scientific contributions of this thesis are

1. the formalisation of the conceptual framework for decision-theoretic planning that was put forward in the previous section, and

2. the subsequent interpretation of the proposed formal structures in clinical terms.

Roughly, we first review several existing representation formalisms for decision-theoretic reasoning, and then propose a new formal framework that synthesises the ideas from these representation formalisms and from the field of symbolic planning in AI. We will then use the proposed formal framework to model a large number of aspects related to clinical patient management.

A secondary contribution is the integration of decision-theoretic and AI approaches to clinical decision support. On the one hand, we subscribe the decision-theoretic principle that a decision-support system in medicine should not uncritically reproduce a physician's heuristics and biases, or mimic the casual habits of clinical practice. Furthermore, a system should build on mathematical theories of information processing and uncertainty reasoning, as people typically tend to have difficulties at this front. Yet, we believe that important steps for applied decision theory lie at the interface between descriptive, prescriptive, and normative accounts, all of which affect each other. To address this interface, an explicit symbolic representation of the concepts that underly clinical reasoning, is required: a decision-support system should not only give advice, but also provide the user with additional insight in the problem domain. As such, our work is rooted in the tradition of symbolic representation and reasoning in AI.

Chapter 2

In Chapter 2, we describe the domain of congenital heart disease and the associated decision problems in clinical practice. We will focus on the most frequently-occurring congenital heart disease, called *ventricular septal defect* (VSD). This disorder is relatively well-understood, and carries many clinical features that are characteristic for congenital heart disease in general; it was therefore chosen as a case study in our investigation. In the VSD domain, it is difficult to predict the future course of disease, and it is therefore hard to decide if, and when, a patient should be submitted to surgery. A careful timing of diagnostic investigations can improve the quality of predictions, but the risks and costs of invasive tests always have to be evaluated against their potential benefits. We use small illustrations from the VSD domain in Chapters 3 and 4, and use it as a more detailed case study in Chapter 6.

Chapter 3

Chapter 3 discusses the formal foundations for decision making under uncertainty within the decision-theoretic paradigm. It is described how probability theory, utility theory, and their synthesis decision theory provide a framework for analysing various types of choice under uncertain conditions. This chapter is introductory in nature and may be skipped by readers that are already acquainted with these fields.

Chapter 4

Chapter 4 reviews the main existing representation formalisms that adhere to the decision-theoretic perspective. Notwithstanding their common ground, these formalisms stem from separate traditions, and therefore come in different forms. We first describe the earlier-mentioned belief networks and their decision-theoretic siblings influence diagrams. Influence diagrams provide a concise way of representing decision problems by exploiting probabilistic independencies between variables in the problem domain; there have been several applications of influence diagrams to problems of therapy planning in medicine, (Quaglini et al., 1989; Quaglini et al., 1993; Bielza et al., 1999). In their basic form, however, influence diagrams are static in nature and are therefore a poor formalisation of the concept of decision-theoretic planning.

The second representation we discuss is the *Markov decision process*. This is a mathematical model of stochastic control that incorporates many ingredients of decision-theoretic planning, (Boutilier et al., 1999). There exist two variants: fully-observable and partially-observable Markov decision processes. Partially-observable Markov decision processes are more suited to the clinical situation, as they assume that part of the problem situation remains hidden for the decision maker. These processes have recently been applied to the problem of therapy planning for adults with acquired cardiovascular diseases, (Hauskrecht, 1997b; Hauskrecht, 1998).

Although Markov decision processes have explicit notions of time and change, their representation of temporal progression is rather coarse. A more delicate representation of stochastic processes is found in the dynamic variants of belief networks and influence diagrams; this is the third and last type of representation discussed in Chapter 4. Dynamic belief networks have been used for monitoring breast-cancer patients who are being given cycles of post-operative cytotoxic chemotherapy, (Bellazzi et al., 1991). Dynamic influence diagrams can be integrated with Markov decision processes; this type of representation is broadly regarded as one of the most powerful formalisms for decision-theoretic modelling and reasoning to date.

Chapter 5

Although Markov decision processes and dynamic influence diagrams incorporate many ingredients of decision-theoretic planning, most of these ingredients are still derived, rather than primitive notions in the formalisation. In Chapter 5 we present a formal framework that integrates probability theory, decision theory, and symbolic

planning at a more fundamental level. Part of this framework is a language that allows for direct manipulation of symbolic structures that describe states, events, observations, decisions and plans. The general notion of *decision process* is defined to cover all influential relationships between system and decision maker over time; this allows to express and study their mutual interaction more directly, and provides for analysing the simplifying assumptions that are used to reduce the computational complexity of problem solving. As an example, we re-evaluate and compare partially-observable Markov decision processes and dynamic influence diagrams as implementations of this notion.

A significant part of the chapter is devoted to a theory of contingency planning, where plans are expressed as collections of decision rules. Decision rules allow for easy communication with field experts and can be directly employed, for instance, in clinical practice guidelines. Furthermore, this form of contingency plan enables us to formulate decision-making strategies at different levels of detail and with varying ranges of applicability; we investigate notions of plan consistency, plan coherence, and plan completeness. A partial order on contingency plans, based on their range of applicability, is defined; this paves the way for incremental procedures of plan construction.

Chapter 6

In Chapter 6 we return to the starting-point of our study, decision making in medical care. It is discussed how the framework from Chapter 5 is applied to decision situations in clinical patient management. We first describe how the framework can be used to create a formal description of a given clinical domain. This is essentially a modelling activity; we use elaborate examples from the VSD domain to illustrate our ideas. In the second part of the chapter, we assume a formal domain description to be available, and investigate the formalisation of clinical reasoning tasks such as diagnosis, treatment planning, and prognosis, and their mutual alternation and interaction in the dynamic perspective on patient management. We show that in each of these tasks, our framework allows for capturing the purely decision-theoretic perspective, but can also make explicit other, conceptual perspectives in clinical decision-making. A related theme is the fact that our framework leaves room for variation in formalising these reasoning tasks: as such, it allows for analysing, and comparing, existing and novel approaches to formal decision support.

Chapter 7

Chapter 7 concludes the thesis with a summary of the main results and a discussion of possible directions for further research.

CHAPTER 2

Treatment planning in paediatric cardiology

Each year, in the Netherlands approximately 1600 children are born with a heart disease. The severity of these diseases ranges from hardly noticeable to life-threatening. Fortunately, the rapid evolution in both medical knowledge and clinical technology has yielded the opportunity to manage the majority of these disorders successfully, providing the patients a normal life-expectancy with little or no disabilities.

In this chapter, we describe the domain of congenital heart disease, and the associated decision problems the paediatric cardiologist faces. We will focus on the most frequently-occurring congenital heart disease, called *ventricular septal defect* (VSD). This disorder is relatively well-understood, and carries many clinical features that are characteristic for congenital heart disease in general; it was therefore chosen as a case study in our investigation. Of course, the methods and techniques to be presented are more general and can be applied to many different decision problems. In Chapters 3 and 4, we will frequently use small illustrations from the VSD domain; in Chapter 6, we describe an elaborate model of the domain, including the formalisation process from which it originated. To facilitate the reader's understanding, Appendix A provides a quick reference of the medical terms that are introduced in this chapter.

We begin in Section 2.1 by describing the anatomy and physiology of the human heart; readers with a medical background may wish to skip this section. Section 2.2 provides a high-level introduction to the domain of congenital heart disease, and Section 2.3 gives a more detailed picture of VSD. We also describe the management of VSD patients in today's clinical practice. The chapter is concluded in Section 2.4, where

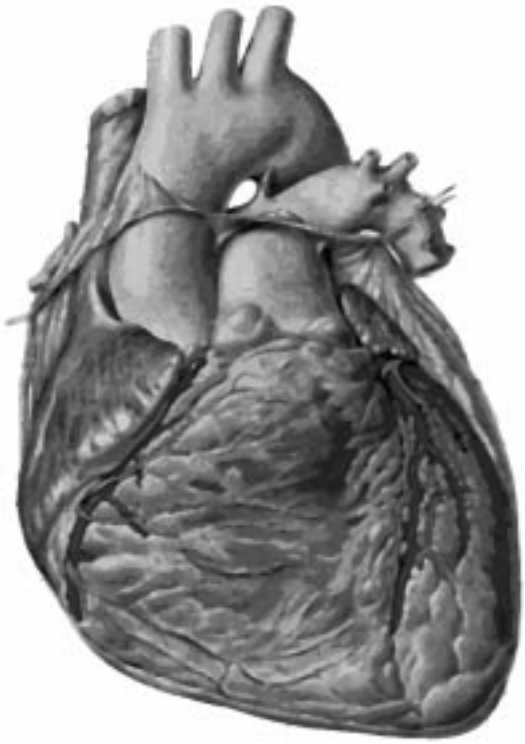


Figure 2.1: Exterior of the human heart.

we discuss this management task in more general terms and sketch the potential for formal decision support in the area.

2.1 Structure and function of the normal heart

The heart is a hollow muscular pump connected to an elaborate network of blood vessels that is spread throughout the body. Heart and vessels thus jointly form a circular transport system, usually called the *cardiovascular system*. Contractile activity of the *myocardium* (heart muscle) is regulated in detail so that the heart can pump either small or large amounts of blood as dictated by the needs of the body.

Below, we describe the anatomical structure of the heart and vessels, the circulation of blood though the body, and the associated mechanics of the cardiovascular system. We will generally confine ourselves to topics that are relevant for the discussion of congenital heart diseases and in particular VSD that is to follow; for a more elaborate treatment of the human cardiovascular system, we refer the reader to a textbook on medical physiology (e.g., Ganong, 1997; Guyton, 1986), or a specialised book on the subject (e.g. Katz, 1977).

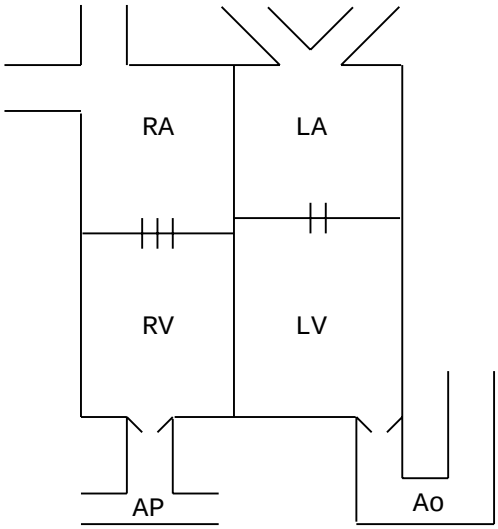


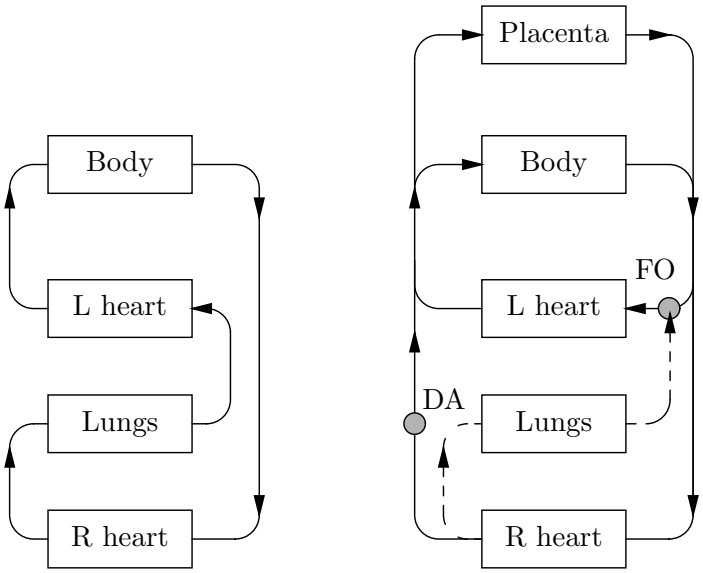
Figure 2.2: Schematic depiction of the heart. RA = right atrium, LA = left atrium, RV = right ventricle, LV = left ventricle, AP = pulmonary artery (*arteria pulmonalis*), and Ao = aorta.

Anatomy of the heart and vessels

The heart is built up from two halves, each of which is divided into two pumping chambers: a thin-walled *atrium*, and a larger and thicker *ventricle* below. The left and right atria are separated by what is called the *atrial septum*, and the left and right ventricle are separated by the greater *ventricular septum*; both are walls of muscular and fibrous tissue. Connected to the ventricles are the great arteries: the *pulmonary artery*, which extends from the right ventricle and leads to the lungs, and the *aorta*, which extends from the left ventricle and leads to the body. Bloods enters the heart through the atria and is pumped into into the great arteries by the contracting ventricles; in the schematic depiction of the heart of Figure 2.2, the blood therefore flows from top to bottom. We note that in reality, pulmonary artery and aorta extend from the upper side of the heart as in Figure 2.1: the arch-shaped aorta is seen on top, and the trunc-shaped pulmonary artery lies directly below it.

The heart contains *valves* to prevent blood from flowing in a direction other than towards the great arteries. Each valve can open only to one side and closes automatically when pressure differences give rise to a blood flow in the reverse direction; the valves are thus essential for the pumping activity of the heart. There are four heart valves: two valves between the cavities of the atria and ventricles, called *atrioventricular valves*, and two valves between the outflow tracts of the ventricles and the great arteries, called *semilunar valves*. The semilunar valve between right ventricle and pulmonary artery is called *pulmonary valve*, and the one between left ventricle and aorta is called *aortic valve*.

The vascular system consists of a extensively branching network of vessels with a large variation in size and capacity. The great arteries connected to the heart split



(a) Normal (b) Foetus

Figure 2.3: The cardiovascular system. The dashes are used to indicate that the pulmonary circulation is bypassed in the fetal circulation.

up in a number of *small arteries*, which in turn split up into a larger number of smaller *arterioles*, and finally a huge number of tiny *capillaries*. The arterioles are the major site of the resistance to blood flow in the cardiovascular system, and small changes in their calibre, which is under nerve control, can cause large changes in the total peripheral resistance. At each bifurcation in the vascular system, the cross-sectional area of the branches exceeds that of the parent vessel. As a result, the cross-sectional area in the capillaries is enormous, and the velocity of blood flow through them is very low. Slow flow provides time for the necessary exchange of substances with the surrounding tissue across the thin capillary walls. After the exchange, the blood is collected by *veins* and carried back to the heart. Here, the blood again accelerates as the cross-sectional area progressively decreases. However, the calibre of the veins exceed that of the corresponding arteries, so the velocity of venous blood only approaches and does not equal that of arterial blood.

Cardiac cycle and blood circulation

Rhythmic contraction of cardiac muscle is achieved by electrical excitation of special *pacemaker cells* in the myocardium. The conduction system made up by these cells spreads the impulses throughout the heart to let its parts beat in an orderly sequence: contraction of the atria is followed by contraction of the ventricles, and subsequently all four compartments are relaxed. All together, the events are called the *cardiac cycle*, where the phase of contraction of the cardiac muscle is referred to as *systole*, and the phase of relaxation is called *diastole*.

The cardiovascular system has two distinct parts: *pulmonary circulation* and *systemic circulation*[1]; this is schematically depicted in Figure 2.3a. Pulmonary circulation is the movement of blood from the heart to the lungs and back, and is responsible for the exchange of carbon dioxide and oxygen in the lungs. Systemic circulation is the movement of blood from the heart to body tissue and back to the heart again; it is responsible for supplying the tissue with nutrients and oxygen, and for collecting waste products. As appears from the figure, the heart can be conceived as consisting of two pumps in series with each other: the left side of the heart moves the blood through the systemic circulation, whereas the right side of the heart moves the blood through the pulmonary circulation.

The above description of the cardiovascular system pertains only to its functioning after birth. During pregnancy, the basic functions of respiration, degradation and elimination of waste products are carried out by the mother, and the cardiovascular system of the foetus is adapted for intra-uterine existence. The fetal lungs are collapsed and have no respiratory function. The flow of blood through the circulation is largely diverted around the lungs; this is accomplished as follows (see Figure 2.3b). Firstly, the small pulmonary arteries are diminished in internal diameter by a thick media, thus preventing a large blood flow by their high resistance. Secondly, there exists a special duct connecting aorta and pulmonary artery, called *ductus arteriosus*, and an oval opening between the atria, called the *foramen ovale*, which functions as a unidirectional flutter valve. Ductus arteriosus and foramen ovale act as bypasses, permitting blood from the systemic veins to enter the systemic circulation without passing through the lungs.

The critical dependence on oxygen requires rapid accommodation of the cardiovascular system to enable autonomous existence immediately after birth. In the newborn infant, the pulmonary arterioles react to the initial inflation of the lungs with oxygen by dilatation (expansion), resulting in a diminished pulmonary resistance; the pulmonary blood flow is greatly increased. The ductus arteriosus is promptly constricted so that aortic flow becomes separated from the pulmonary circuit. Furthermore, the increased pulmonary flow elevates left atrial pressure sufficiently to functionally close the foramen ovale. Anatomic obliteration of the potential aperture of the foramen ovale requires much more time but should normally be completed at the age of 3 months. During the same period, the pulmonary arteries change to thin-walled structures with increased internal diameter; these changes are accompanied by a further decrease in pulmonary resistance to blood flow.

Dynamics of the heart and circulation

The cardiovascular system functions like common hydraulic systems, and its properties can therefore be described in terms of physical quantities, such as location or dimensions, time and force. Derived haemodynamic quantities like displacement,

[1] There is in fact a third part (called the *coronary circulation*) that supplies the heart with oxygen, but this part is irrelevant for the present discussion.

velocity, acceleration, flow and pressure of blood are generally used to characterise cardiac action. The *systemic stroke volume* (Q_s) is the volume of blood ejected into the aorta during each cardiac cycle ($\pm$ 70 mL in a human adult at rest). It is thus equal to the volume of blood in the left ventricle at the moment it begins to contract (*left-ventricular end-diastolic volume*, Q_{lved}; $\pm$ 130 mL) minus the volume that remains after contraction (*left-ventricular end-systolic volume*, Q_{lves}; $\pm$ 60 mL):

$$Q_s \;=\; Q_{lved} - Q_{lves}. \tag{2.1}$$

The fraction of ejected blood (Q_s/Q_{lved}; usually around 0.6 in adults) is called the *ejection fraction*; forceful contraction of the heart can increase this fraction. The pulmonary stroke volume Q_p can be computed in a similar way the systemic stroke volume Q_s. It is easily seen, however, that under normal circumstances the volumes are equal since they merely represent different points of a closed circulatory system.

The quantity of blood pumped into the aorta by the left ventricle each minute is a volume flow rate (i.e., the displacement of a quantity of liquid per unit of time), and is known as the *systemic cardiac output*. It can be calculated as stroke volume times *heart rate* r, the number of beats per minute:

$$\dot{Q}_s \;=\; \tfrac{d}{dt}Q_s \;=\; Q_s \cdot r. \tag{2.2}$$

Cardiac output is approximately 5–6 litres/min in a healthy adult in rest; this is also roughly the complete volume of blood in the body. Again, the flow rates through all the segments of the cardiovascular system must be essentially identical; the pulmonary cardiac output $\dot{Q}_p$ is therefore equal to its systemic counterpart $\dot{Q}_s$.

The autonomic nervous system regulates cardiac output by adjusting the heart rate and, to a lesser extent, the force of myocardial contraction. Depending on the needs of the body, cardiac output is easily doubled or tripled in this way; it may ultimately grow with a factor 7. An important fact is that diastole is shortened to a much greater degree than systole when the heart rate is increased. At high rates the filling of the ventricle, which occurs in diastole, may therefore be compromised; a further increase in heart rate is then ineffective. The factors controlling cardiac output are depicted in Figure 2.4.

We note that although stroke volume and cardiac output are the same at both sides of the heart, there are considerable pressure differences. The peak systolic left ventricular pressure is about 120 mm Hg, and peak systolic right ventricular pressure is 25 mm Hg or less; there are similar differences between aortic and pulmonary pressures.

2.2 Congenital heart disease

Approximately 8 out of each 1000 newborn infants have a congenital heart disease. These diseases are generally due to some anatomical malformation of the heart, great

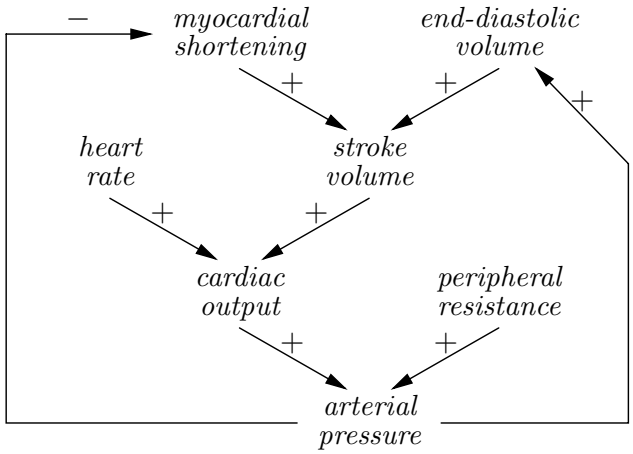


Figure 2.4: Interactions between the components that regulate cardiac output and arterial pressure (modified from Braunwald, 1974).

vessels, or both; the severity of disease and the consequences for the child vary greatly. In approximately 40% of the cases, medical intervention is not necessary as the malformation disappears by itself, or hardly influences the child's general condition and life-expectancy. The remaining 60% is more serious and can be life-threatening; these children therefore need clinical support, possibly including surgery. The majority of the children with congenital heart disease, however, can lead a normal life without surgical intervention.

Below in Section 2.2.1, we briefly discuss the three main groups of congenital disease found in the human heart. We restrict ourselves to disorders that stem from deviations in the anatomical disposition of the heart or great vessels; these can usually be traced back to errors in the development of the fetal heart. Section 2.2.2 reviews the management of patients with congenital heart disease in current medical practice. For more information on congenital disorders of the human heart, we refer to the books by Anderson et al. (1987), Garson et al. (1990), and Moss et al. (1995).

2.2.1 Types of disorders

Congenital cardiac disorders are usually classified into three groups; combinations of various disorders, however, do occur.

I. The first group is referred to as *heart defects*, and is characterised by the existence of a connection between the pulmonary and systemic circulations.

II. The second group comprises *stenoses* (narrowings), which can occur at multiple locations near or at the semilunar valves, and in the aortic arch.

III. The third group of disorders is characterised by *central cyanosis*: the systemic arterial blood is undersaturated with oxygen.

With the modern diagnostic techniques it is generally quite feasible to establish the correct diagnosis of congenitally diseased heart patient. Yet in over 90% of the cases, the preceding cause of a congenital heart disease cannot be established. It is assumed that both hereditary disposition and environmental factors are involved. Below, we elaborate on each of the three groups of congenital cardiac disorders, and on *heart failure*, a condition that accompanies most congenital heart diseases. As our interest is directed towards VSD, which belongs to the group of heart defects, we mostly confine to disorders that may occur secondary to a VSD as complication; the disease itself is described in detail in Section 2.3.

Heart defects

The existence of a connection between the two circulations causes a blood flow through this connection when the heart contracts; this is called *shunting*. Since the pressures in the systemic circulation and left side of the heart are higher than in the pulmonary circulation and right side, shunting will generally occur from systemic to pulmonary arteries, or from left to right sides of the heart; one speaks of *left-to-right* shunting. The common characteristic of disorders with left-to-right shunting is increased pulmonary blood flow. This causes an increased risk of pulmonary infections and, generally speaking, a decreased ability to exert. Examples of this type of congenital heart disease are abnormal openings in the atrial and ventricular septum, respectively called *atrial septal defect* (ASD) and, as mentioned before, *ventricular septal defect* (VSD). Another type of heart defect is a *persistent ductus arteriosus* (PDA). This pertains to the situation where the ductus arteriosus fails to close after birth, and therefore continues to connect aorta and pulmonary artery. Both ASD and PDA may occur secondary to VSD.

Stenotic disorders

The second group of congenital cardiac disorders comprises *stenoses* (localised narrowings). They are categorised by their location: *aortic stenosis* is located at or near the aortic valve, *pulmonary stenosis* is located at or near the pulmonary valve, and *coarctation* is located within the aorta. We will not elaborate on the last type of stenosis here, as the combination of VSD and coarctation is usually seen as an independent disorder. Both aortic and pulmonary stenoses may however occur secondary to a VSD. When the orifices of these valves become stenotic, this affects the pressures and flows within specific chambers of the heart and in the circulation as a whole. The direct effects of aortic stenosis are related to the impediment to blood flow across the aortic valve: it causes a pressure overload on the left ventricle, that is, left ventricular systolic pressure must increase in order to force blood through the narrowed aortic valve into the aorta. The haemodynamic effects of pulmonary stenosis can be understood in a similar fashion to aortic stenosis, now producing a pressure overload on the right ventricle.

Cyanotic disorders

The third group of congenital heart disease is characterised by undersaturation of the systemic arterial blood with oxygen. This occurs when venous and arterial blood is mixed, and is visible from a dusky blue discolouration of the tongue, lips, and conjunctivae; one speaks of *(central) cyanosis*. Similar to disorders from the group of heart defects, these cyanotic disorders comprise a connection between the circulations within or near the heart, but now the shunting of blood is directed from the right to the left side of the heart; this explains the undersaturation with oxygen of the systemic blood. Mixing of venous and arterial blood is dangerous and surgical correction is therefore often emergent. Examples of cyanotic disorders are *transposition of the great vessels*, where both the main arteries and veins are misconnected to the heart, and *Tetralogy of Fallot*, a combination of VSD and pulmonary stenosis. Furthermore, the congenital cardiac disorders from the first group (i.e., those characterised by left-to-right shunting), may in time result in cyanosis. This is due to an adverse reaction in the pulmonary circulation; the associated findings are referred to as *Eisenmenger's syndrome*. The syndrome and its underlying mechanisms is described in more detail in Section 2.3.

Heart failure

A common characteristic of most congenital heart diseases is that the cardiovascular system functions inadequate due to haemodynamic disturbances; this circumstance is referred to as *heart failure*. The heart faces an increased workload as the functioning of several vital organs critically depends on the systemic cardiac output and arterial pressure; any disturbance must therefore be compensated for. The cardiac muscle is however often so powerful during childhood that it can easily compensate for minor failure. With moderate and especially severe heart failure, however, continuous additional effort is required; this is primarily accomplished by increasing the heart rate, and to a lesser extent by narrowing the systemic arterioles to increase their resistance to blood flow. These compensatory mechanisms maintain mean aortic pressure at or near a normal level.

Heart failure is a circumstance that can be distinguished on either sides of the heart; one then speaks of *left* and *right heart failure*, respectively. Left and right failure of the heart vary considerably in their consequences. Left heart failure is generally recognised by signs of the compensatory mechanisms described above; it is therefore associated with a high heart rate and shortness of breath during rest, and a decreased ability to exert. Because fluid will congest in the lungs, left heart failure also yields an increased risk of pulmonary infections. Right heart failure is most apparent from congestion of fluid in the body; typical signs are *hepatomegaly* (enlarged liver) and *oedema* (accumulation of fluid in body tissue). Both left and right heart failure may yield an enlargement of the heart (*cardiomegaly*). We note that since both sides of the heart are part of a closed circulatory system, it is unavoidable that failure of the one side of the heart results in failure of the other, albeit after some time.

A final response to chronically increased haemodynamic demands is growth of the size of cardiac muscular cells; this is called *hypertrophy*. To a certain extent, hypertrophy is beneficial as it increases the total capacity for cardiac work, but it is also a sign of cardiovascular disease. Cardiac hypertrophy is generally distinguished by the heart chamber to which it pertains; one speaks for instance of *left-ventricular* and *right-ventricular hypertrophy*.

We conclude with a few figures on prevalence of congenital heart disease. The most common congenital heart disease is VSD, with a reported prevalence of 24–35%, (Anderson et al., 1987). Other frequently found diseases are PDA (6–12%), pulmonary stenosis (3–14%), and ASD (6–11%).

2.2.2 Clinical paediatric cardiology

In this subsection, we discuss the modalities for the paediatric cardiologist within the standard clinical procedure as described in Subsection 1.1.1. Congenital heart diseases are generally either detected from one of the externally visible signs such as cyanosis or breathing problems, or from the characteristic heart murmurs that accompany these disorders. These murmurs can be heard with cardiac auscultation, an examination that is often routinely performed with newborn infants. After detection of the disease, the infant is sent to the paediatric cardiologist for further management; below we elaborate on the clinical modalities for the paediatric cardiologist.

Physical examination

Physical examination of a child with a (suspected) congenital heart disease may consist of visual inspection of the patient, palpation, and auscultation of the heart and lungs.

Visual inspection can be used to detect central cyanosis, breathing problems, and the presence of oedema. Furthermore, the decreased systemic cardiac output with large left-to-right shunts usually causes the infant to look pale and to sweat easily.

Palpation of the chest and abdomen can be used to detect cardiomegaly and hepatomegaly, and the presence of *thrill* (abnormal vibrations of the heart).

Auscultation of the heart through a stethoscope reveals several *sounds* and possibly also several *murmurs*. Heart sounds accompany the closure of heart valves and are associated with the abrupt acceleration or deceleration of blood at these times. Cardiac murmurs are abnormal sounds, and result from vibrations set up in the bloodstream and the surrounding heart and great vessels as a result of turbulent blood flow. The major causes of murmurs are stenoses, valve insufficiencies, and heart defects.

Auscultation of the lungs, finally, can reveal so-called *pulmonary rales*, which indicate the presence of redundant fluid as a result of congestion in the pulmonary circulation.

Diagnostic tests

Routine tests in today's paediatric cardiology are electrocardiography and echocardiography; when needed, chest roentgenography, blood tests, cardiac catheterisation, and open lung biopsy are also conducted.

Electrography is a non-invasive measurement technique that records fluctuations in the action potentials of muscle cells. An *electrocardiogram* (ECG) is used to investigate disorders of the heart's electrical activation and conduction system. It can also be employed to detect left- and right-ventricular hypertrophy. The investigation is non-invasive, painless and has no risks for the patient. The reliability of the ECG is however questionable.

Echography is a measurement technique that uses the body's internal as an acoustic mirror. Application of this technique to the region of the heart is called *echocardiography*. It permits continuous recording of the position and movements of heart walls, valves and blood vessels. Additional information about intra-cardiac blood flow patterns can be obtained by employing the Doppler shift to compute blood cell velocity. The application of echocardiography is non-invasive and painless, and carries no risks to the patient; the results are reliable when the image is judged by an experienced cardiologist.

Roentgenography can be used by the paediatric cardiologist to obtain an undistorted image of the heart and vessels. It provides a means to discover cardiomegaly and increased pulmonary vascularity due to pulmonary hypertension. The procedure is non-invasive, painless, and reliable, but exposes the patient to a small amount of noxious x-rays; the number of exposures should be minimised.

Blood tests are used to measure the level of hemoglobin and oxygen in the systemic arterial blood, and to check levels of medicines and their potential side-effects. A blood test is invasive but harmless to the patient.

Cardiac catheterisation is used to measure the pressures and oxygen saturations of blood in the heart's chambers, which allows, for instance, to calculate the size of a left-to-right ventricular shunt. It is also possible to produce a moving roentgenographic image of cardiac blood flows by injecting radiopaque contrast media through the tip of the catheter. Because cardiac catheterisation is rather painful and the procedure takes quite some time, young children are often anaesthetised; hospitalisation may be required. The procedure carries a risk of stroke (due to thrombus formation) and perforation of the heart or great vessels; the usage of contrast media may cause allergic reactions, in particular

anaphylactic shock. In a small number of cases, these complications are fatal; the mortality risk is higher during the first months of life.

Open lung biopsy is a procedure for obtaining a small specimen of pulmonary tissue for microscopic analysis. It can point out whether there exists damage to the pulmonary arterioles. Although an open lung biopsy is invasive, extremely burdening to the patient, and not completely reliable, the procedure may be performed because the state of the pulmonary arterioles has major consequences for the choice of therapy.

Treatment modalities

The treatment modalities for the clinical cardiologist are medical treatment, cardiac catheterisation, and cardiac surgery.

Medical treatment can be used to relieve the symptoms of heart failure in two ways: *cardiac glycosides* enhance the strength of myocardial contraction, and *diuretics* promote the excretion of urine through their effects on kidney function, thus removing redundant fluid that is accumulating in body tissue. These treatments carry negligible risk to the patient and have little or no side-effects.

Cardiac catheterisation can not only be used for diagnostic purposes, but also to repair some cardiac lesions. Examples are stretching stenotic heart valves or blood vessels with a balloon catheter and closing unwarranted connections and defects by placing small umbrellas. The latter technique can sometimes be employed to close a small VSD. As described earlier, catheterisation carries a risk of complications and death, but it is significantly smaller than in the case of surgery.

Cardiac surgery is one of the most difficult types of operation, especially with young children. Although today it is possible to operate the heart of newborn infant within days or even hours after birth, the risks of such a procedure are still substantial. Possible postoperative complications are bleeding, arrhythmia, and pulmonary hypertension; there is also a risk of death. When there is no need for immediate intervention, as is the case with congenital heart defects such as VSD, surgery is usually postponed. At the age of one year, the mortality risks of cardiac surgery have dropped to roughly 1%. When the surgical procedure is complicated and takes much time, the chances of complications and death increase.

2.3 Ventricular septal defect

VSD (Gumbiner and Takao, 1990; Graham and Gutgesell, 1995) is a relatively well-understood congenital cardiac disorder with many clinical features that are charac-

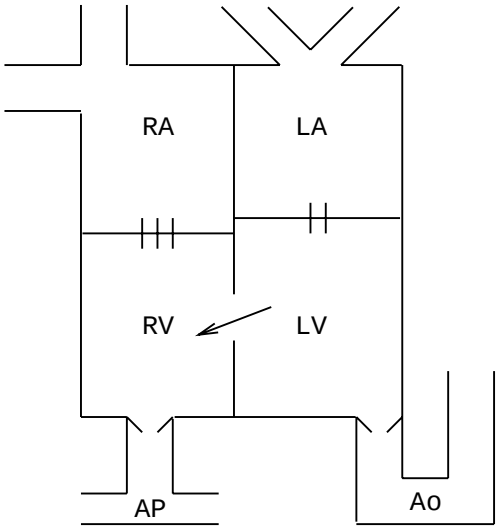


Figure 2.5: Schematic depiction of a heart with a VSD and left-to-right ventricular shunting of blood.

teristic for congenital heart disease in general. Anatomically speaking, VSD is a hole in the ventricular septum, the fibromuscular wall that separates the left and the right ventricle. As noted in Section 2.2.1, it is also the most frequently occurring congenital heart disease, with a reported prevalence of 24–35%, (Anderson et al., 1987). In approximately 80% of these cases, VSD is the sole lesion. In the Netherlands, some 400 to 500 children are born each year with the disease; the incidence is 1.5 to 3.5 out of each 1000 newborn infants. VSD occurs 2 to 3 times more often with premature births than with term births, and slightly more frequently with women (56%) than with men (44%). It is also frequently found in patient's with Down's syndrome; the majority of all VSDs (>95%) is however not associated with chromosome abnormalities.

In brief, the main pathophysiological consequences of the disease are left-to-right shunting through the defect due to ventricular pressure differences (see Figure 2.5), and as a result thereof pulmonary hypertension and heart failure; the main threat is pulmonary arteriopathy and cyanosis (Eisenmenger's syndrome). We will first elaborate on the disease's characteristics in Subsection 2.3.1, and then discuss the management of VSD patients in Subsection 2.3.2.

2.3.1 Disease characteristics

In this subsection, we describe on morphological and pathophysiological aspects of VSD, and discuss the possible complications that may accompany the disease. We also describe Tetralogy of Fallot, a congenital heart disease that is strongly related to VSD, and conclude with a summary of potential developments of VSD and its complications over time.

Morphology

A VSD is anatomically classified according to its *size* and its *type*, where type pertains to its location in the ventricular septum. Furthermore, there may be more than one defect in the septum. Below, we briefly elaborate on these classifications.

A most distinguishing feature of a VSD is its size. Defects may have different dimensions and this has important physiological and clinical consequences. Unfortunately, it is somewhat difficult to characterise the size of a given VSD as it is essentially three-dimensional in shape. In clinical practice, it is customary to use the largest diameter of the VSD to describe its size. This diameter may range up to 6 to 7 mm at birth, and extend up to 20 mm during childhood as the heart grows and the VSD expands with it; the relative size of the VSD does not increase though. Using the diameter, one usually distinguishes *small* (up to 3 mm at birth), *moderately large* (3–5 mm at birth, at most 10 mm thereafter), and *large* (more than 5 mm at birth, later 10–20 mm) VSDs.

The type of a VSD refers to its location in the ventricular septum. The septum is composed of a small membranous component that is surrounded by three muscular components. A number of VSD typologies have been proposed in the literature; we confine ourselves to the most widely used distinction proposed by Soto et al. (1980) between

- *perimembranous VSD* (85–90%; located within the membranous component, adjacent to the semilunar valves),
- *subaortic VSD* (5–7%; also located within the membranous component, located just beneath the aortic valve), and
- *muscular VSD* (5–20%; located within the muscular apical septum in the lower part of the heart, neither bordering on one of the heart valves nor on the membranous component of the septum).

Additionally, a perimembranous VSD may extend into the muscular components near the semilunar and atrioventricular valves.

There may be multiple defects in the ventricular septum: muscular defects may co-occur with perimembranous and subaortic defects, and simultaneous occurrence of a large number of small, muscular defects is also possible. This is called a *Swiss cheese septum* and sometimes regarded as separate type of VSD. We note that singular VSDs are however much more common than multiple VSDs.

It is possible, and even quite likely, that septal tissue grows on the borders of the defect such that the VSD spontaneously decreases in size; it may result in eventual closure of the defect. This is more likely to happen for small than for large VSDs; especially small muscular VSDs are amenable to spontaneous closure. Overall, about 70% of all VSDs close spontaneously, (Krovetz, 1998), the majority (54%) closing in

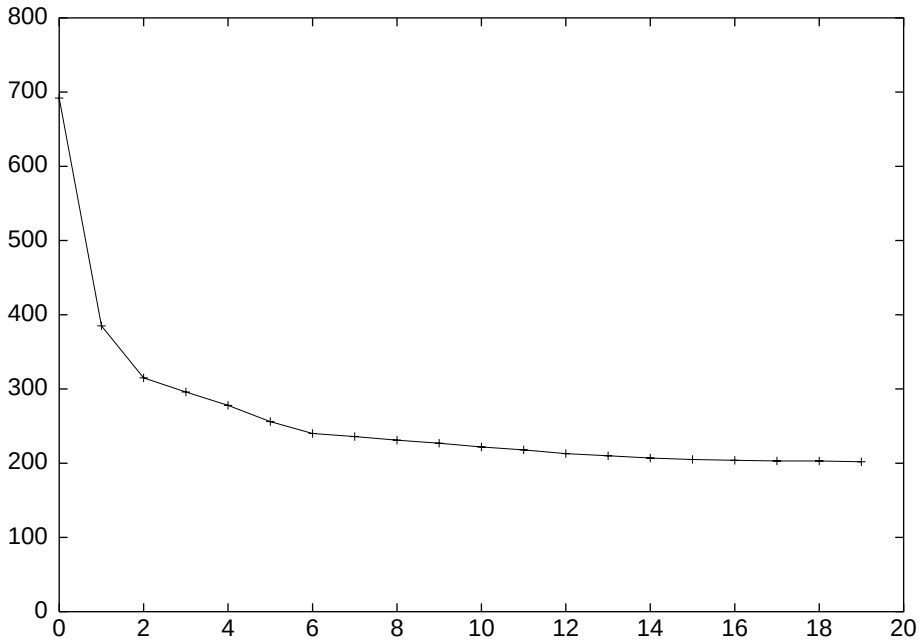


Figure 2.6: Exponential decay rate of closure. The x-axis shows age of the patients in years, the y-axis shows the number of patients with a VSD that has not closed (yet); taken from (Krovetz, 1998).

the first two years of life. Spontaneous closures have however been reported to occur up to the age of 31 years, but the rate of closure appears to follow an exponential decay rate; see the graph of Figure 2.6.

Pathophysiology

A defect in the ventricular septum yields a connection between the heart's ventricles through which blood shunts from left to right upon cardiac contraction; the disturbed circulation is depicted in Figure 2.7. Physiologically, VSDs are therefore characterised by the *shunt size*, the amount of blood flowing through the defect. Generally speaking, the shunt size depends on the size of the defect and the vascular resistances to blood flow. The type (location) of the VSD is physiologically speaking uninfluential.

The existence of a shunt invalidates many of the physiological rules that normally apply to the heart and circulation. For instance, the flow rates through different segments of the cardiovascular system are no longer identical. Left-to-right shunting causes oxygenous blood to be pumped back into the lungs again, and therefore left-ventricular output and systemic flow are smaller than right-ventricular output and pulmonary flow, respectively. The shunt size is usually described as the ratio $\dot{Q}_p : \dot{Q}_s$ of pulmonary and systemic blood flows. Under normal, healthy conditions, the amounts of blood flowing through both circulations are necessarily equal, and therefore $\dot{Q}_p : \dot{Q}_s = 1 : 1$. Left-to-right shunting will cause the ratio to go up; shunt

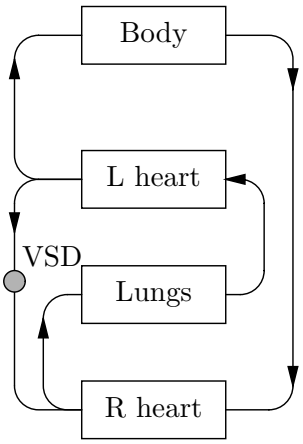


Figure 2.7: Circulation in an infant with VSD.

sizes of 2:1, 3:1, and occasionally even higher may be found with large heart defects. Conversely, right-to-left shunting will cause the ratio to decrease. It is not possible, though, that pulmonary flow is substantially smaller than systemic flow because the pressures on the right side of the heart cannot exceed those on the left side. As a general rule, the shunt size is related to pressures and resistances in the cardiovascular system as follows:

$$\dot{Q}_p : \dot{Q}_s \;=\; \frac{\Delta P_p : \Delta P_s}{R_p : R_s}, \tag{2.3}$$

where $R_p : R_s$ is the ratio of pulmonary and systemic vascular resistances (normally between 1:10 and 1:8), and $\Delta P_p : \Delta P_s$ is the ratio of pulmonary and systemic pressure lapses, that is, ΔP_p (ΔP_s) is the difference between mean arterial and venous pressures in the pulmonary (systemic) circulation. Because vascular resistances are not influenced by blood flow rates (at least not directly), it follows that shunting will induce abnormalities in pressure levels. With left-to-right shunting for instance, pulmonary arterial pressure will rise; this is called *pulmonary hypertension*. One would expect a corresponding pressure drop in the systemic arteries, but as the functioning of several vital organs critically depends on the systemic arterial pressure, this is immediately compensated for by sympathetic stimulation. In sum, both sides of the heart face a higher workload with left-to-right shunting: the right ventricle must contract more forcefully to push blood into the high-pressured pulmonary artery, while the left ventricle most work harder to maintain systemic arterial pressure.

The main consequences of left-to-right shunting are ventricular hypertrophy and heart failure. As blood flows through the VSD upon cardiac contraction, the (systolic) pressure difference between the heart's ventricles will diminish. When the VSD is small and the pressure difference remains 25 mm Hg or higher, the VSD is said to be *obstructive*. With large, typically non-obstructive VSDs, the mitral valve (between left atrium and ventricle) may be too small for the increased amount of blood that

has to flow through it, yielding signs of mitral stenosis. The valve is then said to be *functionally stenosed*; it is not truly narrowed.

VSD is accompanied by a number of pathophysiological developments during the first years of life. In the first weeks following birth, the presence of a VSD usually has little consequences. As described in Section 2.1, the muscular pulmonary arteries are then small in diameter with a thick smooth muscular wall, thus preventing massive shunting by their high resistance. During this period, the arteries change to thin-walled structures with increased internal diameter. These changes are accompanied by a decline in pulmonary vascular resistance, resulting in an increased shunt size and the associated symptoms. After 6 to 12 weeks, the pulmonary arteries have obtained their normal state, and the shunt size has reached its maximum. There are now basically two potential developments: spontaneous closure of the VSD or increasing pulmonary hypertension.

Eisenmenger's syndrome

With the passage of time, left-to-right shunting may cause severe damage to the pulmonary vascular bed, (Hopkins, 1995); this is called *pulmonary arteriopathy*. More specifically, the continuous pulmonary hypertension and overflow cause an increase of cells at the inner layer of pulmonary arterioles (*intimal hyperplasia*), a pathological condition that increases their resistance to blood flow. As a result, the size of the shunt is diminished while the pulmonary vascular pressure remains high. The initial effects of the diminished shunt are an apparent improvement of the patient's condition: the heart now faces a less extreme workload. Pulmonary arteriopathy should however be avoided as it negatively affects the respiratory function: the patient will increasingly have breathing problems.

Eventually, the pulmonary vascular resistance becomes so high as to cause reversal of the direction of shunting, and the patient becomes centrally cyanosed (similar to patients with a congenital heart disease from the third group, discussed in Section 2.2.1). The pulmonary arterioles have now become irreversibly damaged; the patient will develop severe respiratory problems and has a strongly reduced life-expectancy. The complex of findings associated with the final stage of this development is named *Eisenmenger's syndrome*, after its first describer.

Eisenmenger's syndrome represents the major threat to patients with a VSD. It is more likely to occur with large VSDs as there is then more blood shunting through the defect. The rate at which the pulmonary vascular bed is damaged differs from individual to individual, and the whole process up to the final, cyanotic stage may take many years. Spontaneous closure of a VSD with Eisenmenger's syndrome is excluded.

Complications

Several cardiac disorders may co-occur with a VSD: sometimes an additional disorder is present at birth, and it is also possible that a cardiac complication develops over time. Below, we distinguish four types of possible complication related to the heart; we restrict ourselves to cases where VSD is the primary disorder and all other disorders are secondary; this occurs with approximately 20% of all VSDs.

The first type of complication is an additional congenital heart defect, an ASD (atrium septum defect) or PDA (persistent ductus arteriosus). These complications have in common that they increase the connectivity between pulmonary and systemic circulation, and can be regarded as increasing the *functional size* of the VSD, that is, the defect size that is relevant in haemodynamic respect. As a result, these complications raise the size of the shunt, and aggravate all pathophysiological conditions associated with VSD.

The second type of complication is a stenosis of the aortic or pulmonary valve. As described in Subsection 2.2.1, such stenoses influence the pressures in the heart's ventricles as they obstruct the outflow of blood to the great arteries. In combination with a VSD, this has primarily consequences for the size of the shunt. A pulmonary stenosis reduces the shunt and is therefore beneficial from a haemodynamic point of view. It does however cause right-ventricular hypertrophy, and hampers spontaneous closure of the VSD. An aortic stenosis will increase the shunt and therefore worsen the condition of the patient; it may also cause underdevelopment of the aorta.

The third type of complication is a *malalignment of the outlet* (or *infundibular*) *septum*, the muscular component of the ventricular septum between the semilunar valves. This means that the unimpaired parts of the ventricular septum are sometimes not positioned in a single plane. In particular, the outlet septum is directed towards either the left or the right side of the heart. When it occurs, this malalignment is in fact the actual cause of the VSD; it will often also preclude spontaneous closure of the VSD. Furthermore, a malalignment may induce several further complications such as a pulmonary or aortic stenosis.

The complications described so far all constitute conditions that are present at birth. The fourth and final complication, the *aortic prolapse*, does not, in contrast, occur at birth but may develop in time; it may be caused by subaortic VSDs. Anatomically speaking, each heart valve is built up from a few leaflets, called *cusps*; one of the cusps of the aortic valve may prolapse into a subaortic VSD. The prolapsed cusp may occlude the defect, and therefore diminishes the functional size of the VSD as less blood will be able to flow through it. However, another result is that the aortic valve becomes *insufficient* (leaky): during early diastole, when the pressure gradient between the aorta and left ventricle is maximal, blood flows back from the aorta into the left ventricle. An aortic prolapse thus causes additional failure of the left heart.

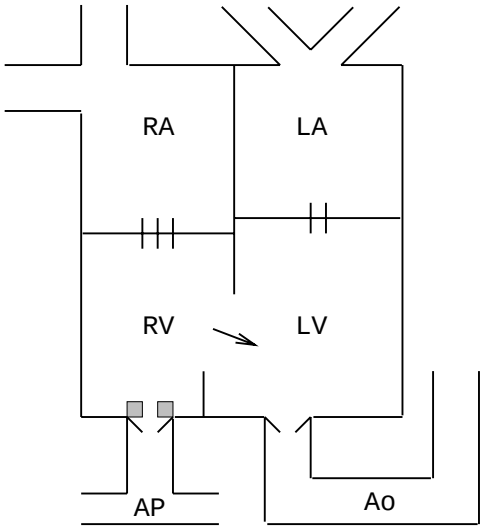


Figure 2.8: Schematic depiction of a heart with Tetralogy of Fallot. The outlet component of the ventricular septum is wrongly aligned.

Tetralogy of Fallot

The perimembranous VSD is from anatomical point of view closely related to another congenital heart disease, *Tetralogy of Fallot*. Tetralogy of Fallot comprises four lesions: a perimembranous VSD with extension to the outlet septum, an infundibular pulmonary stenosis, right-ventricular hypertrophy, and an *overriding aorta*. This means that the aorta is moved away from its normal position so that it arises partly from the right ventricle; the aorta is therefore positioned above (i.e., overrides) the ventricular septum. The severity of the disease is often expressed in terms of the amount of overriding. The actual cause of Tetralogy of Fallot is a right-malalignment of the outlet septum (see Figure 2.8). With this malalignment, VSD, pulmonary stenosis, and overriding aorta are unavoidable; right-ventricular hypertrophy is caused by the stenosed outflow in the right ventricle. Tetralogy of Fallot is a cyanotic congenital heart disease that may be present at birth, but an acyanotic form exists as large perimembranous VSD with an extension to the outlet area, accompanied by a minor malalignment of the outlet septum. The symptoms are completely similar to a normal VSD and it is usually diagnosed as such. As time progresses, the malalignment may increase, resulting in more overriding of the aorta, increased narrowing of right-ventricular outflow tract, and therefore a decrease in shunt size. Eventually, the shunting becomes bi-directional (both left-to-right and right-to-left flow). At this stage, the child is not cyanotic (yet), and one sometimes speaks of *pink Fallot*. When the overriding of the aorta exceeds 50% (i.e., the ventricular septum is positioned halfway the aorta valve), the child becomes cyanotic and the primary diagnosis is Tetralogy of Fallot; this happens with approximately 10% of all VSDs.

Spontaneous closure of a VSD that is part of a Tetralogy of Fallot is impossible due to the malalignment of the outlet septum. Typical symptoms of the disease are so-

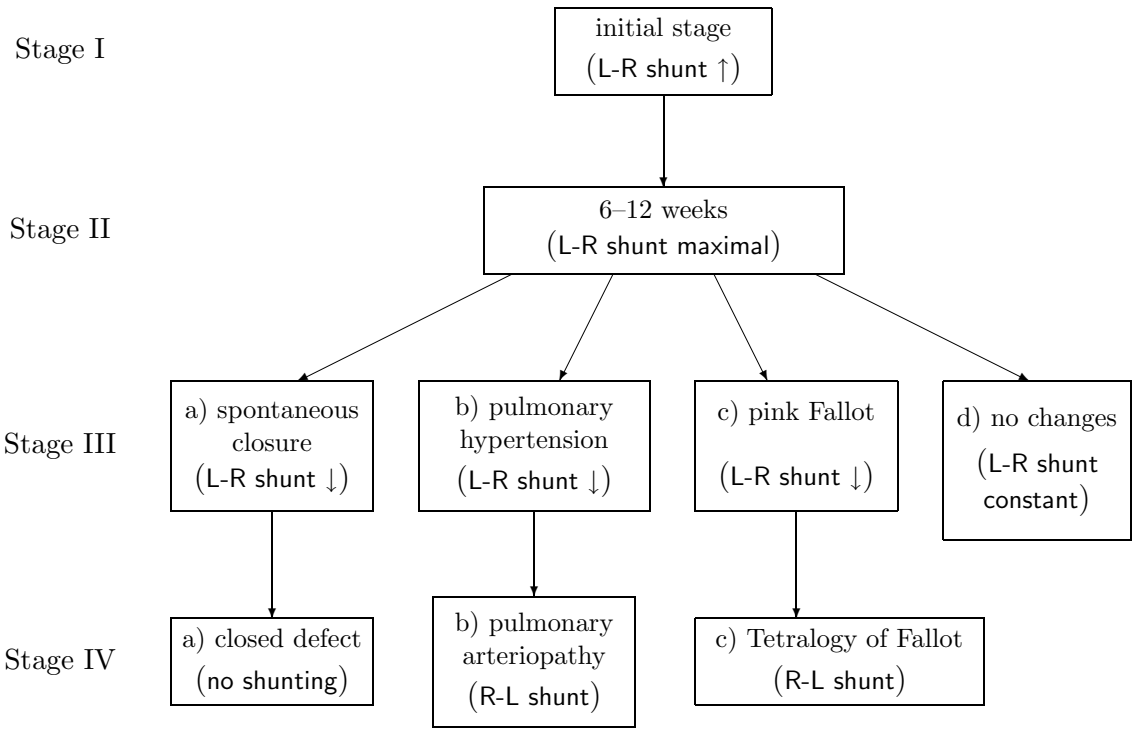


Figure 2.9: Stages in the pathophysiological development of a VSD.

called *cyanotic spells*, caused by spasms in the right-ventricular outflow tract; the stenotic pulmonary valve is then temporarily completely occluded, and the patient has an acute shortage of oxygen. Furthermore, because the pressure in the right ventricle is much higher than normal, connective tissue may grow in this chamber after some time. This will reduce the pumping ability of the heart and may cause cardiac arrhythmia.

Summary of stages

We will now briefly summarise the possible developments over time for a VSD patient; see the diagram of Figure 2.9. The first stage (Stage I in the diagram) follows birth and is characterised by the presence of a VSD without much consequences. After 6 to 12 weeks (Stage II), the pulmonary vascular resistance has dropped, resulting in increased left-to-right shunting, and possibly heart failure and the associated symptoms. It is now possible to assess the severity of the disease and its complications; the size of the shunt will not become larger than it is at this stage. Subsequently (Stage III), there are three potential developments: a) growth of septal tissue and spontaneous closure of the VSD in Stage IV, b) increasing pulmonary hypertension eventually leading to pulmonary arteriopathy in Stage IV, or c) increasing malalign-

ment of the outlet septum, resulting in Tetralogy of Fallot in Stage IV. A fourth possible development is d) that there are no significant changes. It should be noted that with different patients there may be substantial differences between the rates of these developments; Stages III-IV may cover anything between roughly 1 and 30 years.

2.3.2 Clinical management

In this subsection, we discuss the management of VSD patients in today's clinical practice. We will first follow the above division in disease stages and indicate which of the clinical modalities described in Subsection 2.2.2 are generally employed at each stage. Subsequently, we describe the management of complications that may accompany a VSD, potential pitfalls for the clinician, and issues that are yet unresolved.

Management of stages I and II: diagnosis

As described above, the first stage of VSD covers the first 6 to 12 weeks of life and is characterised by a gradual arise of left-to-right ventricular shunting. The shunt yields a systolic murmur that is sometimes detected immediately after birth from routine auscultation of the heart. Otherwise, the VSD remains unnoticed until the first symptoms of left-to-right shunting and heart failure show up after a few weeks. In either case, the patient should be sent to a paediatric cardiologist who can confirm the diagnosis from echocardiographic investigation. The presence of a VSD causes little or no harm to the patient at this stage, and the precise timing of its detection is unimportant. It is very rare that a VSD remains undetected, unless it is very small and therefore completely harmless.

The second stage is marked by the presence of typical signs and symptoms associated with VSD, and provides the first opportunity to assess the severity of the disease. With small defects ($\pm$ 35%), there will be little or no haemodynamic abnormalities and therefore hardly any symptoms. The main sign is that cardiac auscultation reveals a systolic murmur that is characteristic for VSD. Moderate-sized defects ($\pm$ 40%), induce more shunting and therefore some symptoms of heart failure such as breathing problems and difficulties with drinking; in time there may be some growth arrearage. Physical examination of the patient reveals a loud systolic murmur and a thrill, and the ECG shows left-ventricular hypertrophy. Sometimes a murmur may be audible during the diastole as well, caused by a functionally stenosed mitral valve. Large defects ($\pm$ 25%) induce a large shunt and severe symptoms of heart failure, including a poor weight gain, cardiomegaly, oedema, hepatomegaly, and pulmonary rales. Left-ventricular output is reduced, causing the patient to be sweating and to have a pale facial colour; the ECG shows hypertrophy of both ventricles. The systolic murmur will however be softer, as the pressure difference between both ventricles is now smaller.

As appears from the above, the systolic murmur will vary in duration, volume and

form depending on the shunt size; a trained physician will however recognise it as produced by a VSD. Furthermore, the combination of systolic murmur and thrill is symptomatic for VSD: no other disease causes simultaneous occurrence of these signs. It is therefore rare that a VSD is wrongly diagnosed, but it does happen that the VSD later turns out to be part of a Tetralogy of Fallot. Although the findings described above provide an indication of the size of the shunt and, therefore, of the size of the VSD, echocardiography is indispensable to assess the VSD's precise size, (Geva et al., 1988). Echocardiography can also show the location and number of defects, and the presence of additional cardiac anomalies. Furthermore, Doppler techniques can be used to establish the haemodynamic changes induced by the defect, (Helmcke et al., 1989; Danford et al., 1997). Roentgenography is usually avoided because of the noxious x-rays.

Management of stage III: the surgery decision

The principal strategy in the management of VSD patients is to avoid surgical intervention when possible; this is the case when the VSD closes spontaneously due to tissue growth, or when the defect causes little or no haemodynamic abnormalities. With mild to moderate heart failure, medical treatment (i.e., cardiac glycosides and diuretics) can be used to enhance the blood circulation, thereby supporting the child's functioning and development. We note that medical treatment cannot accelerate spontaneous closure of the VSD or preclude pulmonary arteriopathy. A secondary aim in VSD management is to avoid invasive diagnostic procedures such as cardiac catheterisation and open lung biopsy. Whereas catheterisation was routinely performed in the 1980s, it is now often obsolete because of the improved reliability of echocardiography and associated Doppler techniques, (Danford et al., 1997).

With small VSDs, the clinical course is favourable throughout infancy and childhood (Kidd et al., 1993), and these defects are likely to close spontaneously. Patients with moderate-sized defects may develop large left-to-right shunts and associated complications in infancy, but the majority of this group can also be managed medically without surgical intervention. Patients with large defects are more difficult to manage, because of the risks of mortality in the first year of life due to heart failure and associated pulmonary infections. Furthermore, pulmonary arteriopathy and cyanosis (Eisenmenger's syndrome) may develop over time as a response to continuous pulmonary overflow and hypertension. Early surgical intervention is therefore recommended for these patients: once the pulmonary arteries are damaged, surgical closure of the VSD will only worsen the condition of the patient.

For the treating clinician, the main problem is to decide if and when to submit a patient to surgery. Usually, the patient's condition is monitored without surgical intervention during the first year of life. During this period, non-invasive diagnostic tests such as cardiac auscultation, ECG, and echocardiography are conducted repeatedly, and when necessary, medical treatment is given to reduce heart failure and improve the overall condition of the patient. Sometimes, roentgenographic images of

the chest are made to inspect the size of the heart and pulmonary vascularity. After the first year of life, the risks associated with surgical intervention have dropped, and a decision whether surgery is necessary has to be made. Further tests may be performed prior to that decision to obtain more information. Cardiac catheterisation is employed when pulmonary hypertension due to pulmonary arteriopathy is suspected and the results of auscultation and echocardiography are equivocal. When the results of catheterisation are also unclear, the remaining option is use open lung biopsy to inspect the state of the pulmonary arterioles directly; it is rarely the case that this investigation is needed though.

From all diagnosed VSDs, only 1 to 2 out of 10 require surgical repair. When the patient's condition is very bad, a hospitalisation period may precede the operation for purposes of improvement. Large defects have to be closed with a patch of synthetic material; these patches sometimes hamper cardiac function. It does happen that the surgeon fails to close a VSD completely, or fails to close all VSDs when multiple defects were present; a second operation is required when the remaining defects cause haemodynamic abnormalities. The failure happens more often with muscular VSDs, because these defects are difficult to access for the surgeon and often come in multiples. Very large defects are always treated in two phases: the defect is reduced in size during the first operation, and closed completely during the second. All these circumstances increase both the mortality risks and the risks of permanent complications for the patient. The majority of patients with repair of uncomplicated VSD in infancy or early childhood have however an excellent result with no clinical signs or symptoms, and apparently normal life-expectancy, (Moller et al., 1991).

Management of stage IV: prognosis

The surgery decision at Stage III critically depends on the prognosis made by the treating physician: the expected clinical development of the patient is crucial for the preferred action. Unfortunately, even for an experienced cardiologist it is often difficult to correctly assess the prognosis for a given VSD patient. Furthermore, the three major developments (branches a,b, and c in the diagram of Figure 2.9) have similar appearances at stage III: they initially induce a gradual diminishment of most signs and symptoms as the child will suffer from less shunting and heart failure. It therefore happens that the development is misjudged: sometimes increasing pulmonary vascular resistance (Eisenmenger's syndrome) or increasing overriding of the aorta (pink Fallot), are falsely taken for spontaneous closure of the VSD. We note that in principle, it is well possible to distinguish these developments using echocardiography and cardiac catheterisation.

When the shunt size diminishes due to increased pulmonary vascular resistance, the systolic murmur will first turn into an atypically-formed murmur before disappearing completely. The patient will occasionally become cyanotic during exertion; in addition to the high systolic pulmonary arterial pressure, the diastolic pulmonary pressure is now also increasing. It can be detected from insufficiency of the pulmonary valve,

and confirmed by cardiac catheterisation. Although this finding indicates progressive pulmonary arteriopathy, the condition may be reversible at this point, and surgical closure of the VSD may still be possible. When the results of echocardiography and catheterisation are equivocal, an open lung biopsy is sometimes performed to obtain more certainty. The final stage of Eisenmenger's syndrome is marked by continual cyanosis, and is also recognisable from a loud second heart sound; the damage to the pulmonary arterioles is then definitely irreversible. Surgical closure of the VSD is discouraged as it will only worsen the patient's condition. The only alternative is lung and heart-lung transplantation, but the risks of this transplantation are extremely high, (Hopkins et al., 1996). A strongly reduced life-expectancy results for these patients: they rarely live more than 30 years.

The third development, progressive overriding of the aorta as part of a Tetralogy of Fallot also reduces shunting and heart failure. It is also sometimes falsely judged to be a spontaneous closure of the VSD, which will later cause an unpleasant surprise when the child becomes cyanotic. The mistake is however less dramatic than overseeing increasing pulmonary vascular resistance, as there are little consequences for the management strategy to be followed. Tetralogy of Fallot requires surgery in all cases, because the patient otherwise has a life-expectancy of less than 20 years; the mortality risk is approximately 5%. We do note that if the operation is performed too late, the mortality risk is higher, and connective tissue that has grown in the right ventricle may chronically hamper the pumping ability of the heart and cause arrhythmia.

Complications

In this subsection, we briefly discuss the effects of the complications ASD, PDA, aortic and pulmonary stenoses, and valvular insufficiencies on the management of VSD patients. We also describe the potential postoperative complications and their long-term consequences.

As described in Section 2.3, when one of the defects ASD or PDA is present in addition to a VSD, this will increase the shunt size and the associated symptoms. As both ASD and PDA can be repaired with little risks, it is decided more quickly to perform surgery with these complications. This is especially true for PDA, which has a significant effect on shunting.

Stenoses of the aortic and pulmonary valve also require surgery: aortic stenoses because they increase the shunt and may cause underdevelopment of the aorta, and pulmonary stenoses because they hamper spontaneous closure of the VSD. Unfortunately, these anomalies cannot be repaired along with the VSD; two cardiac operations are needed. It does happen that a subvalvular aortic stenosis (which is found relatively frequently in combination with subaortic VSDs), emerges after surgical repair of the VSD. This possibility calls for careful postoperative examination of the patient as a second operation is then required.

A possible complication of subaortic VSDs is a progressive prolapse of the aortic

valve; without surgical intervention, this would happen with 2 to 5% of all VSDs. The aortic prolapse will cause insufficiency of the aortic valve, which can be detected from a specific diastolic murmur. Echocardiography is however also required as pulmonary insufficiency yields exactly the same kind of murmur. Surgical closure of the VSD after an aortic prolapse is problematic because the valve may become more leaky; the prolapsed valve is also difficult to repair. Early intervention is therefore recommended.

Potential postoperative complications are pulmonary hypertension, infections, bleedings, stenoses, hypertrophy of the cardiac muscle, and damage to the heart's electrical conductive tissue. These complications influence the patient's life-expectancy.

Potential pitfalls

As noted above, the most dangerous error when treating a VSD patient is a failure to recognise increasing pulmonary vascular resistance in time. When pulmonary arteriopathy becomes irreversible, the patient cannot be treated anymore and has a strongly reduced life-expectancy. A similar risk is the late detection of progressive overriding of the aorta (Tetralogy of Fallot). Although this does not preclude the possibility of surgery, the long-term risks for the patient will have increased. Finally, overseeing the first signs of an aortic prolapse may complicate the possibilities of treatment.

Although cardiac catheterisation can be valuable in assessing the actual development of the disease, it is generally possible to avoid it by thorough and regular echocardiographic examinations. Unnecessary catheterisation is therefore regarded as erroneous, although the risks associated with it are small. Similar observations hold for open lung biopsies. A more serious error is surgical repair of a small, closing, or otherwise harmless VSD. Sometimes, this happens when after years of prosperous development, the patient's condition suddenly worsens due to a different cause but the worsening is misattributed to the VSD.

In many hospitals, cardiac catheterisation and chest roentgenography are still performed on a routine basis prior to surgical repair of a VSD. It was recently conjectured by Magee et al. (1998) that routine pre-operative cardiac catheterisation is probably no longer necessary because of the improved echocardiographic techniques.

2.4 Discussion

In this section, we look back on the domain of paediatric cardiology as described in the previous sections and discuss the task of managing VSD patients in more general terms. Furthermore, we give a brief historical overview of decision-support systems in the field of congenital heart disease. We conclude with establishing the criteria for a system that could presently support the paediatric cardiologist.

Review of the application domain

If we regard the domain of paediatric cardiology, the following characteristics of its current clinical practice emerge. First, little is known about what causes children to be born with a congenital heart disease, and it is therefore not to be expected in the near future that prevention is possible. The physiological mechanisms in the human cardiovascular system are however well understood and allow for extensive modelling (e.g., see John, 1995). Second, there have been rapid technological improvements over the last decades, resulting in the possibility to perform cardiac surgery at low risks and making accurate diagnostic assessments without employing invasive tests; this has led to a considerably improved life-expectancy for the patients. For the majority of VSD patients for instance, the expectations are presently very good with little or no disabilities. Third, the domain becomes rapidly more complex when multiple anomalies interact, and it is, even for uncomplicated lesions, often very difficult to establish a reliable prognosis. The planning of therapy does however require the ability to predict the interplay between the natural history of the disease and effects of clinical actions: there is always a trade-off between the benefits gained by waiting with surgical intervention in the hope that the patient's condition will improve, and the risks caused by the natural history of the disorder, (Macartney et al., 1987). The fourth and last characteristic is therefore that due to the prognostic difficulties, domain experts often differ in their opinions on management decisions, and different treatment regimes are employed at different sites. Also, it is unclear in a number of situations which diagnostic investigations should be conducted; a continual shift in opinions, caused by the rapid technological developments, appears from the specialist literature.

Past efforts on decision support

The field of paediatric cardiology has regularly inspired researchers in artificial intelligence in medicine to design decision-support systems. We give a brief historical overview of the most well-known systems.

Early work in the field was performed by Warner et al. (1961), who developed a naive Bayesian model for diagnosing congenital heart disease. In such a model, disease types and potential clinical findings are represented as statistical variables, and it is assumed that the findings are mutually independent if the disease is known. Historically, this was the first application of this type of model whose later applications are too numerous to list (e.g., de Dombal et al., 1972).

Whereas Warner et al. had concentrated on inferring the most likely diagnosis from a given set of findings, Gorry and Barnett (1968) considered the broader problem of diagnosis as both diagnostic inference and diagnostic test selection. They describe a mathematical model and associated interactive computer program to support this task. In the model, diagnostic inference is again performed from a naive Bayesian model, while test selection is based on heuristic functions describing the cost of conducting diagnostic tests and the penalties of misdiagnosing. In an application in the

field of congenital heart disease, it was found that the model yields a sharp reduction in the average number of tests performed, while establishing a diagnosis at expert level.

In the early 1980s, a diagnostic expert system called GALEN was developed by Thompson et al. (1983). GALEN's knowledge base describes 70 congenital heart diseases and disease variants, covering ± 95% of all cases found in hospital files. GALEN distinguishes possible forms of congenital heart disease in children given previous data obtained from the patient's history and current data obtained from physical examination, ECG, and roentgenography. Domain knowledge is represented as a combination of production rules and frames.

Franklin et al. (1991) developed an algorithm to help the junior doctor reaching a preliminary diagnosis for newborn babies with cyanotic congenital heart disease using information reported over the telephone (concerning findings from physical examination, blood tests, ECG and roentgenography). The algorithm was evaluated on 400 cases with encouraging results: the algorithm had a diagnostic accuracy of 76%, compared with 64% for the paediatric cardiologist and 45% for the referring doctor. Although the algorithmic formulation was attractive in its simplicity and transparency, it suffered from problems of observer variability and was unable to handle missing data. It was therefore decided to reformulate the algorithm as a probabilistic expert system as such systems are more forgiving of limitations in data, and also provide the opportunity to incorporate statistical information extracted from clinical data records. A Bayesian belief network was therefore developed to model the domain. Initially, the naive Bayesian assumption was made for this network. The results of experiments were however disappointing, and the conclusion was that a "deeper" model comprising the underlying physiological mechanisms was more appropriate, (Franklin et al., 1989). A more elaborate network, called CHILD, took such mechanisms into account. It was found to perform better than the naive Bayesian network, but still inferior to the original algorithm, (Spiegelhalter et al., 1993).

The most recent work on diagnosing congenital heart disease was performed by Reed et al. (1997). Their system, called FALLOT, allows for diagnosing multiple co-occurring congenital heart diseases such as defects, stenoses, and absent or misconnected vessels. The system uses a database of expected cues for each disease, and descriptions of how each type of cue combines when more than one disease is present. Alternative solutions are compared by the ratio of explained normal cues over total abnormal cues. Experiments showed that FALLOT approaches the level of field experts.

Present criteria for decision support

From the above review of systems that have been developed to support clinical decision making in the field, it appears that diagnosing congenital heart disease has received considerable attention, whereas therapeutic and prognostic decision making has been largely neglected. In part, this can be explained by the fact that technolog-

ical advances have only recently solved the problem of diagnosis in this field. Modern echocardiographic techniques have emerged during the 1980s and are now available in every hospital clinic; before that time however, correctly diagnosing cardiac disorders was much harder. But probably, the emphasis on diagnostic decision support has also been influenced by the greater availability of formal diagnostic tools and techniques, and the widespread fallacy that diagnosis is the major problem to solve in clinical medicine.

From our discussion of the VSD domain however, we conclude that presently, most benefits can be expected from systems that support the clinician in predicting the future course of disease. The primary focus of decision support should therefore be *prognosis*. By weighing the risks associated with predicted developments, the clinician should also be advised in the *decision and timing of intervention*. Furthermore, as prognosis basically concerns the extrapolation of past and current findings to the future, careful timing of clinical investigations can improve the quality of predictions. Additionally, the risks and costs of invasive tests have to be evaluated against their potential benefits, i.e. better management resulting from a gain in information. The final task to support is therefore the *choice and timing of diagnostic tests*. In general terms, the overall task of the envisioned support system may be characterised as *prognostic assessment and action planning under uncertainty*, where the *timing* of actions is essential.

Several additional requirements for decision support systems in the domain of congenital heart disease can be identified. It is our conviction that the paediatric cardiologist should be provided with flexible, interactive, and transparent decision support. Flexibility is required because it is customary to employ a mixture of 'hard' and 'soft' preferences in many clinical decisions. Although considerations of life-expectancy and life-quality are predominant, borderline cases may be decided by considering, for instance, the parents' reaction to the stress of raising a diseased child. A system should be highly interactive as the treating clinician will want to inspect the expected outcomes of alternative choices, and check his intuitions against the predictions of the system. Transparency, finally, is a requirement as the structure of the domain of application and the associated management problem should be apparent from a system's presentation. As noted above, the physiological mechanisms in the cardiovascular system are well understood; explicit modelling of these mechanisms will make an advisory system more liable to acceptance by clinicians in the field.

CHAPTER 3

Decision making under uncertainty

A variety of approaches exist to formalise problems of action planning under uncertainty, each of them emphasising other aspects of this type of problem. As motivated in Chapter 1, in this thesis we choose to employ (Bayesian) decision theory at the fundamental level of comparing alternative actions in situations of choice. The basics of decision theory are discussed in this chapter. Decision theory can be regarded as synthesising *(Bayesian) probability theory* and *utility theory*. Here, probability theory serves as a framework for reasoning with uncertainty, whereas utility theory provides the guidelines for rational choice under uncertainty.

This chapter is structured as follows. We first discuss probability theory in Section 3.1, and continue with utility theory in Section 3.2. The synthesis of both theories is given in Section 3.3, which also provides an overview of how different types of decision problem are analysed with this type of reasoning. We conclude with a short discussion in Section 3.4. The chapter also serves to introduce a number of formal notations that will be used in subsequent chapters. Throughout, the various notions that are introduced will be illustrated with examples from the domain of paediatric cardiology as described in the previous chapter.

3.1 Probability theory

In this section, we review the main concepts from probability theory. Our review will be concise, and is not intended to be exhaustive: it highlights the aspects of

probability theory that are crucial in decision-theoretic reasoning; for a thorough introduction to the theory, we refer to (Shiryayev, 1984; Grimmett and Stirzaker, 1992). We start introducing the language we will employ to describe elements from a domain of interest; this language, a *Boolean algebra*, consists of Boolean expressions over value assignments to a given set of variables (Birkhoff and MacLane, 1977). Let $\mathrm{dom}(w)$ denote the *domain* of a given variable w, i.e. the set of possible values that the variable may take.

Definition 3.1 (Boolean algebra) *Let W be a set of variables. The* Boolean algebra spanned by *W, denoted by $\beta(W)$, comprises all expressions built up from value assignments to elements from W, the constants $\top$* (true) *and $\bot$* (false)*, the binary operators $\wedge$* (conjunction) *and $\vee$* (disjunction)*, and the unary operator $\neg$* (negation)*.*

In the Boolean algebra $\beta(W)$, value assignments to elements from W, i.e. expressions of the form $w = v$, where $w \in W$ and $v \in \mathrm{dom}(w)$, act as Boolean variables. We use $\varphi \equiv \psi$ to indicate that the expressions $\varphi, \psi \in \beta(W)$ are equivalent under the usual axioms. Furthermore, we will write $\varphi \vdash \psi$ when $\varphi \equiv \varphi \wedge \psi$, or equally $\psi \equiv \varphi \vee \psi$. The constants $\bot$ and $\top$ now denote the universal upper and lower bounds of the distributive and complemented lattice on $\beta(W)$ induced by the relation $\vdash$.

Within our theory, a prominent part is played by conjunctions of value assignments to sets of variables; we will refer to such conjunctions as *configurations*.

Definition 3.2 (Configuration) *Let W be a set of variables. A conjunction*

$$\mathrm{c}_W \quad = \quad \bigwedge_{w \in W} w = v \tag{3.1}$$

of value assignments to the variables from W is called a configuration *of W. The set of all configurations of W is called the* universe *of W, notation Ω_W.*

Note that $\Omega_W \subseteq \beta(W)$. Furthermore, there is but one configuration of the empty set $\varnothing$, and this is the empty conjunction $\top$. We will usually write c_W to denote a configuration of W (i.e. $\mathrm{c}_W \in \Omega_W$); for a singleton set $\{w\}$, we will simply write c_w and Ω_w instead of $\mathrm{c}_{\{w\}}$ and $\Omega_{\{w\}}$, respectively. So,

$$\Omega_w \quad = \quad \{w = v \mid v \in \mathrm{dom}(w)\} \tag{3.2}$$

for each variable w.

Notation 3.3 *Let $W = \{w_1, \ldots, w_n\}$ be a set of variables. We use*

$$\mathrm{dom}(W) \quad = \quad \mathrm{dom}(w_1) \times \cdots \times \mathrm{dom}(w_n) \tag{3.3}$$

to denote the set of possible values of W, and $W = V$, where $V \in \mathrm{dom}(W)$, $V = \{v_1, \ldots, v_n\}$, as an alternative notation for the configuration

$$\mathrm{c}_W \quad \equiv \quad w_1 = v_1 \wedge \cdots \wedge w_n = v_n. \tag{3.4}$$

Variable	*Interpretation*	*Domain*
`VSD`	VSD size	*null*, *small*, *moderate*, *large*
`shunt`	shunt size	*none*, *small*, *moderate*, *large*, *reversed*
`resis`	pulmonary vascular resistance	*normal*, *increased*, *high*, *very_high*
`hfail`	heart failure	*absent*, *mild*, *moderate*, *severe*
`pmhyp`	pulmonary hypertension	*absent*, *mild*, *moderate*, *severe*
`pmart`	pulmonary arteriopathy	*false*, *true*
`closure`	spontaneous closure	*false*, *true*
`death`	death	*false*, *true*

Table 3.1: Example variables for the VSD domain.

We will regard the expression $W = V$ as an element of the Boolean algebra $\beta(W)$.

Notation 3.4 *Let W be a set of variables. The expression $\mathrm{C}_W = \bigwedge_{w \in W} w$ is called the* configuration template *of W. From the configuration template of W, any configuration $\mathrm{c}_W \in \Omega_W$ can be obtained by substituting the variables with appropriate value assignments.*

We will generally use configuration templates within formulas. The templates then provide for treating the formulas as *schemata* from which multiple instantiations can be obtained by filling in values for the variables in the templates.

Given two sets Y and Z of variables, we say that their configurations c_Y and c_Z are *compatible* when $\mathrm{c}_Y \wedge \mathrm{c}_Z \not\equiv \bot$. This holds when either Y and Z are disjoint, or c_Y and c_Z assign the same values to variables in $Y \cap Z$. Otherwise, these configurations are called *incompatible*.

Example 3.5 *Consider the set of variables listed in Table 3.1, representing concepts from the domain of paediatric cardiology; each of the variables describes a part of the clinical state of VSD patients. Example configurations from the Boolean algebra of propositions spanned by this set are*

$$\begin{array}{rcl} \varphi_1 & : & \texttt{VSD} = \mathit{small} \quad \wedge \quad \texttt{shunt} = \mathit{small}, \\ \varphi_2 & : & \texttt{shunt} = \mathit{small} \quad \wedge \quad \texttt{hfail} = \mathit{mild}, \quad \text{and} \\ \varphi_3 & : & \texttt{shunt} = \mathit{large} \quad \wedge \quad \texttt{hfail} = \mathit{severe}. \end{array}$$

We have that $\varphi_1 \in \Omega_{\{\texttt{size},\texttt{shunt}\}}$, and $\varphi_2, \varphi_3 \in \Omega_{\{\texttt{shunt},\texttt{hfail}\}}$. Furthermore, φ_1 and φ_2 are compatible, whereas neither φ_1 and φ_3, nor φ_2 and φ_3 are.

We will now consider a set X of *random variables*, and define a *joint probability distribution* on X as a function on the Boolean algebra of propositions spanned by X.

Definition 3.6 (Probability distribution) *Let X be a set of random variables, and let $P : \beta(X) \rightarrow [0, 1]$ be a function such that*

- $P(\top) = 1$,
- $P(\varphi) = 0$ *when* $\varphi \equiv \bot$, *and*
- *if* $\varphi, \psi \in \beta(X)$ *such that* $\varphi \wedge \psi \equiv \bot$, *then* $P(\varphi \vee \psi) = P(\varphi) + P(\psi)$.

Then, P is called a joint probability distribution *on X. For each $\varphi \in \beta(X)$, the function value $P(\varphi)$ is termed the* probability *of φ.*

Note that we associate probabilities with Boolean expressions instead of with sets, which is common in textbooks on probability theory. Both notions of probability are however equally expressive, (de Finetti, 1970). We say that a joint probability distribution P on X is *degenerate* when $P(\mathrm{c}_X) \in \{0, 1\}$ for all $\mathrm{c}_X \in \Omega_X$; otherwise, it is *non-degenerate*. A non-degenerate distribution P is called *strictly positive* when $P(\mathrm{c}_X) > 0$ for all $\mathrm{c}_X \in \Omega_X$; it is *uniform* when all configurations of X have equal probability, i.e. when $P(\mathrm{c}_X) = 1/|\Omega_X|$ for all $\mathrm{c}_X \in \Omega_X$.

Definition 3.7 (Conditional probability) *Let X be a set of random variables, and let P be a joint probability distribution on X. For each $\varphi, \psi \in \beta(X)$ such that $P(\psi) > 0$, the* conditional probability *of φ given ψ is defined as*

$$P(\varphi \mid \psi) \quad = \quad \frac{P(\varphi \wedge \psi)}{P(\psi)}. \tag{3.5}$$

The conditional probability $P(\varphi \mid \psi)$ expresses the amount of certainty concerning the truth of proposition φ, given that the information ψ is known with certainty. We state without proof that for a given $\psi \in \beta(X)$ with $P(\psi) > 0$, the conditional probabilities $P(\varphi \mid \psi)$ for all $\varphi \in \beta(X)$ once more constitute a joint probability distribution on X. This distribution is called the *conditional probability distribution given ψ*, and we also denote it by P^{ψ}. Conditional probability distributions are also referred to as *posterior distributions*, as they describe the respective probabilities posterior to the incorporation of a piece of information; the original (unconditional) probability distribution is then called the *prior* distribution.

The notion of *conditional independence* (Dawid, 1979) allows for qualification of the relations between variables in a joint probability distribution; it underlies the graphical representation of probability distributions which is discussed in Chapter 4.

Definition 3.8 (Conditional independence) *Let X be a set of random variables, let $Y_1, Y_2, Z \subseteq X$, and let P be a joint probability distribution on X. Then, the set*

Y_1 is said to be conditionally independent *of the set Y_2 given the set Z under the distribution P, notation $Y_1 \perp\!\!\!\perp_P Y_2 \,|\, Z$, if*

$$P(\mathrm{C}_{Y_1} \,|\, \mathrm{C}_{Y_2} \wedge \mathrm{C}_Z) \;=\; P(\mathrm{C}_{Y_1} \,|\, \mathrm{C}_Z). \tag{3.6}$$

If $Z = \varnothing$, we say that Y_1 is marginally independent*, or simply* independent*, of Y_2, and write $Y_1 \perp\!\!\!\perp_P Y_2$.*

Intuitively, $Y_1 \perp\!\!\!\perp_P Y_2 \,|\, Z$ means that if Z is known, then determining Y_2 provides no further knowledge about Y_1. Recall that C_{Y_1}, C_{Y_2} and C_Z denote the configuration templates of Y_1, Y_2, and Z, respectively, representing all possible configurations of these sets. Note that conditional independence of Y_1 from Y_2 given Z is equivalent with

$$P(\mathrm{C}_{Y_1} \wedge \mathrm{C}_{Y_2} \,|\, \mathrm{C}_Z) \;=\; P(\mathrm{C}_{Y_1} \,|\, \mathrm{C}_Z) \cdot P(\mathrm{C}_{Y_2} \,|\, \mathrm{C}_Z). \tag{3.7}$$

Therefore, conditional independence is symmetrical in Y_1 and Y_2.

We now recapitulate some well-known propositions from probability theory that allow to make inferences from a given probability distribution. We assume that X is a set of random variables, P is a strictly-positive joint probability distribution on X, and that Y and Z are arbitrary subsets of X. Note that a joint probability distribution P' on $\beta(Y)$ can be derived from the probabilities $P(\mathrm{c}_Y)$ for all configurations of Y ; P' is then called the *marginal distribution* on Y.

Proposition 3.9 (Chain rule)

$$P(\mathrm{C}_X) \;=\; P(\mathrm{C}_{X \setminus Y \cup Z} \mid \mathrm{C}_{Y \cup Z}) \cdot P(\mathrm{C}_Y \mid \mathrm{C}_Z) \cdot P(\mathrm{C}_Z) \tag{3.8}$$

Proof.

$$\begin{aligned} P(\mathrm{C}_{X \setminus Y \cup Z} \mid \mathrm{C}_{Y \cup Z}) \cdot P(\mathrm{C}_Y \mid \mathrm{C}_Z) \cdot P(\mathrm{C}_Z) \;&=\; \\ \frac{P(\mathrm{C}_{X \setminus Y \cup Z} \wedge \mathrm{C}_{Y \cup Z})}{P(\mathrm{C}_{Y \cup Z})} \cdot \frac{P(\mathrm{C}_Y \wedge \mathrm{C}_Z)}{P(\mathrm{C}_Z)} \cdot P(\mathrm{C}_Z) \;&=\; \\ \frac{P(\mathrm{C}_X)}{P(\mathrm{C}_{Y \cup Z})} \cdot \frac{P(\mathrm{C}_{Y \cup Z})}{P(\mathrm{C}_Z)} \cdot P(\mathrm{C}_Z) \;&=\; P(\mathrm{C}_X) \end{aligned} \tag{3.9}$$

□

Proposition 3.10 (Marginalisation)

$$P(\mathrm{C}_Y) \;=\; \sum_{\mathrm{c}_Z \in \Omega_Z} P(\mathrm{C}_Y \wedge \mathrm{c}_Z). \tag{3.10}$$

Proof. For each $\mathrm{c}_Y \in \Omega_Y$ we have

$$\mathrm{c}_Y \quad \equiv \quad \bigvee_{\mathrm{c}_Z \in \Omega_Z} \mathrm{c}_Y \wedge \mathrm{c}_Z. \tag{3.11}$$

As the disjuncts in this expression are mutually incompatible, i.e.

$$\mathrm{c}_Y \wedge \mathrm{c}_Z) \wedge (\mathrm{c}_Y \wedge \mathrm{c}'_Z) \quad \equiv \quad \bot \tag{3.12}$$

for all $\mathrm{c}_Z, \mathrm{c}'_Z \in \Omega_Z,\ \mathrm{c}_Z \neq \mathrm{c}'_Z$), the sum of their respective probabilities equals the marginal probability of c_Y. □

Proposition 3.11 (Conditioning)

$$P(\mathrm{C}_Y) \quad = \quad \sum_{\mathrm{c}_Z \in \Omega_Z} P(\mathrm{C}_Y \mid \mathrm{c}_Z) \cdot P(\mathrm{c}_Z). \tag{3.13}$$

Proof. Directly from the definition of conditional probability and the marginalisation property. □

Theorem 3.12 (Bayes' theorem)

$$P(\mathrm{C}_Y \mid \mathrm{C}_Z) \quad = \quad \frac{P(\mathrm{C}_Z \mid \mathrm{C}_Y) \cdot P(\mathrm{C}_Y)}{P(\mathrm{C}_Z)}. \tag{3.14}$$

Proof. Directly from the definition of conditional probability. □

We proceed by giving definitions of *expectation* and *variance*, which play a crucial role in utility theory and decision theory. The *expected* or *mean value* of a function f under distribution P is the weighted sum of its possible values where the weights are probabilities.

Definition 3.13 (Expected value) *Let $f : \Omega_X \to \mathbb{R}$ be a function over the possible configurations of X. The* expected value *of f under probability distribution P is defined as*

$$E_P(f) \quad = \quad \sum_{\mathrm{c}_X \in \Omega_X} P(\mathrm{c}_X) \cdot f(\mathrm{c}_X). \tag{3.15}$$

The variance of f is a measure of the spread around its mean; it is defined as the expected squared deviation of the values $f(\mathrm{c}_X)$, $\mathrm{c}_X \in \Omega_X$, from $E_P(f)$.

Definition 3.14 (Variance) *Let $f : \Omega_X \to \mathbb{R}$ be a function over the possible configurations of X. The* variance *of f under probability distribution P is defined as*

$$\mathrm{var}_P(f) \;=\; E_P(g), \tag{3.16}$$

where

$$g(\mathrm{c}_X) \;=\; (f(\mathrm{c}_X) - E_P(f))^2. \tag{3.17}$$

Note that when P is degenerate probability distribution with $P(\mathrm{c}_X) = 1$ for some $\mathrm{c}_X \in \Omega_X$, we have that $E_P(f) = f(\mathrm{c}_X)$ and $\mathrm{var}_P(f) = 0$.

We conclude this section by discussing the notion of *entropy*. Entropy was developed by Shannon and Weaver (1949) to characterise the uncertainty in a given probability distribution, which is regarded as inversely proportional to the information conveyed by that distribution. Uniform distributions are seen as conveying the least, and degenerate distributions as conveying the most information possible on the actual state of the variables involved.

Definition 3.15 (Entropy) *The entropy $H_P(Y)$ of the set $Y \subseteq X$ in probability distribution P on X is defined as*

$$H_P(Y) \;=\; -\sum_{\mathrm{c}_Y \in \Omega_Y} P(\mathrm{c}_Y) \log P(\mathrm{c}_Y). \tag{3.18}$$

Note that $H_P(Y) \geq 0$ for all distributions P and sets Y. The value of $H_P(Y)$ is high when there is much uncertainty regarding the set Y, i.e. when all possible configurations of Y have roughly the same marginal probability; the value of $H_P(Y)$ is low when the marginal distribution on Y is more pronounced. If there is no uncertainty, i.e. when $P(\mathrm{c}_Y) = 1$ for some $\mathrm{c}_Y \in \Omega_Y$, then $H_P(Y) = 0$.

3.2 Utility theory

Utility theory was formulated by Von Neumann and Morgenstern (1944) as an adjunct to their theory of games. Others soon recognised it in its own right as an important mathematical foundation for decision making under uncertainty. The central results of utility theory are that, given a number of assumptions on rational choice, a preference order on decision-making outcomes can be expressed as a real-valued function (called a *utility function*), and a preferred decision alternative is one that maximises the expectation of this function (Chernoff and Moses, 1959; Savage, 1972; Raiffa and Schlaifer, 1961); this is called the *Maximum Expected Utility criterion*, or *MEU criterion*.

This section is divided into three subsections. Subsection 3.2.1 reviews the fundamentals of utility theory, and presents the MEU criterion. In Subsections 3.2.2 and

3.2.3, we study further characteristics of utility functions, based on attitudes towards risk and the existence of multiple objectives. We conclude in Subsection 3.2.4 with a discussion of *quasi-utility functions*, functions that violate the assumptions of utility theory, but may nevertheless prove useful in some circumstances.

3.2.1 The MEU criterion

Formally, in utility theory a *decision* is viewed as a choice from among one or more *lotteries*. A lottery specifies a probability distribution over "prizes" (i.e. potential outcomes of the decision) by listing them along with their respective probabilities. The decision maker receives exactly one prize, drawn using the probability distribution specified by the lottery he has chosen.

Definition 3.16 (Lottery) *Let W be a set of variables, jointly describing the possible outcomes of a decision. The set $\mathcal{L}(W)$ of* lotteries *over W is defined as the smallest set satisfying:*

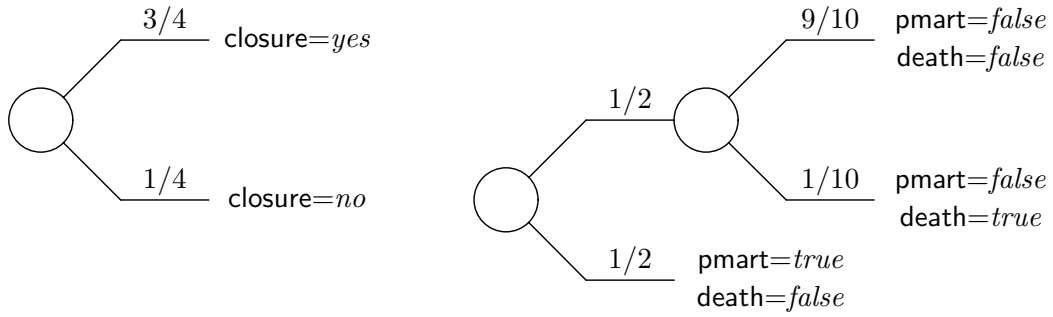
1. $\Omega_W \subset \mathcal{L}(W)$, *and*
2. *if $l_1, \ldots, l_n \in \mathcal{L}(W)$, $n \in \mathbb{N}$, $n \geq 1$, then $(p_1, l_1; \ldots; p_n, l_n) \in \mathcal{L}(W)$, where $0 \leq p_i \leq 1$, $i = 1, \ldots, n$, and $\sum_{i=1}^{n} p_i = 1$.*

A configuration $c_W \in \Omega_W$ is also called an atomic lottery *of $\mathcal{L}(W)$. A lottery $l = (p_1, l_1; \ldots; p_n, l_n)$, is called a* simple lottery *if each l_i is atomic; otherwise l is called a* compound lottery. *For a non-atomic lottery $l = (p_1, l_1; \ldots; p_n, l_n)$, a pair (p_i, l_i), $i = 1, \ldots, n$, is called a* branch *of l with* sublottery *l_i.*

We note that the prizes in a simple lottery need not be distinct, and neither do the prizes in sublotteries of a compound lottery. Furthermore, there are no formal restrictions on the number of branches, or on the recursion depth in a lottery, except that they are both finite. Without loss of generality, we do take non-atomic lotteries to be non-degenerate, i.e. to have some branches with a probability $0 < p < 1$, and therefore involve a true gamble.

Example 3.17 *Lotteries often are depicted graphically using rooted trees. Figure 3.1a shows the simple lottery* $(3/4, \texttt{closure} = \textit{true};\ 1/4, \texttt{closure} = \textit{false})$. *Figure 3.1b shows the compound lottery* $(1/2,\ (9/10, \texttt{pmart} = \textit{false} \wedge \texttt{death} = \textit{false};\ 1/10, \texttt{pmart} = \textit{false} \wedge \texttt{death} = \textit{true});\ 1/2, \texttt{pmart} = \textit{true} \wedge \texttt{death} = \textit{false})$.

A compound lottery $(p_1, l_1; \ldots; p_n, l_n)$ is taken to yield the sublottery l_i with probability p_i, $i = 1, \ldots, n$, where l_i in turn is interpreted as yielding some outcome or lottery. It is easily seen that for each lottery $l \in \mathcal{L}(W)$, there exists a unique corresponding probability distribution P_l over W that satisfies the following conditions:



(a) A simple lottery. (b) A compound lottery.

Figure 3.1: Two lotteries.

1. if $l \in \Omega_W$, then $P_l(l) = 1$ and $P_l(\mathrm{c}_W) = 0$ for all $\mathrm{c}_W \in \Omega_W \setminus \{l\}$, and
2. if $l = (p_1, l_1; \ldots; p_n, l_n)$, then $P_l(\mathrm{c}_W) = \sum_{i=1}^{n} p_i \cdot P_{l_i}(\mathrm{c}_W)$ for each $\mathrm{c}_W \in \Omega_W$.

From a formal point of view, the set $\mathcal{L}(W)$ is therefore equivalent to the set of possible probability distributions over W. The association with gambling-style lotteries, where the outcomes typically involve winning or losing sums of money, is purely metaphoric. It does help emphasise though the assumption that the decision maker is informed about the probabilities associated with branches in a lottery and about the worths of potential outcomes.

One of the results of the correspondence between lotteries and probability distributions is that compound lotteries can be reduced to simple ones using the chain rule of probability theory (Proposition 3.9), without altering their interpretation.

Example 3.18 *The compound lottery of figure 3.1b can be reduced to the simple lottery* $(9/20, \texttt{pmart} = \mathit{false} \wedge \texttt{death} = \mathit{false};\ 1/20, \texttt{pmart} = \mathit{false} \wedge \texttt{death} = \mathit{true};\ 1/2, \texttt{pmart} = \mathit{true} \wedge \texttt{death} = \mathit{false})$ *specifying the same probability distribution.*

Utility theory assumes that the lotteries over a set W can be compared to each other by the decision maker as to their desirability, meaning that there exists a (subjective) preference ordering on the set $\mathcal{L}(W)$ of all lotteries over W. For two lotteries $l, l' \in \mathcal{L}(W)$, we will write $l \prec l'$ to denote preference of l' over l by the decision maker, $l \succ l'$ for preference of l over l', and $l \simeq l'$ for indifference of the decision maker between l and l'. In the latter case, we will also say that l and l' are *equally preferred*. The symbol $\preceq$ will be used to denote preference over or indifference between lotteries.

Definition 3.19 (Preference ordering) *Let W be a set of variables and let $\mathcal{L}(W)$ be the set of lotteries over W. Then, a* preference ordering *on $\mathcal{L}(W)$ is a relation $\preceq \subseteq \mathcal{L}(W) \times \mathcal{L}(W)$ that for all $l, l', l'' \in \mathcal{L}(W)$ satisfies the following properties:*

(P1) *$\preceq$ is a linear ordering on $\mathcal{L}(W)$;* (Orderability)

(P2) *if $l \preceq l' \preceq l''$, then there exists a p with $0 \leq p \leq 1$ such that $l' \simeq (p, l; 1-p, l'')$;* (Continuity)

(P3) *for all p with $0 \leq p \leq 1$ we have $l \simeq l'$ if and only if $(p, l; 1-p, l'') \simeq (p, l'; 1-p, l'')$;* (Substitutability)

(P4) *if $l \preceq l'$, then for all p,q with $0 \leq p, q \leq 1$, we have $(p, l'; 1-p, l) \preceq (q, l'; 1-q, l)$ if and only if $p \leq q$;* (Monotonicity)

(P5) *for all p,q with $0 \leq p, q \leq 1$, we have $(q, (p, l; 1-p, l'); 1-q, l'') \simeq (qp, l; q(1-p), l'; 1-q, l'')$.* (Decomposability)

The properties P1 through P5 are generally called the *axioms of preference*. The continuity axiom (P2) states that for any three lotteries l, l', l'' with $l \preceq l' \preceq l''$, there exists a lottery composed of l and l'' that is equally preferred to l'. If l' is atomic, i.e. $l' \in \Omega_W$, then we say that l' is the *certainty equivalent* of this compound lottery. Axiom P3 asserts that lotteries that are equivalent when considered alone remain equivalent as part of a larger context, and vice versa. The fourth axiom (P4, monotonicity) asserts that a decision maker prefers the lottery that offers the greater chance of receiving the better outcome. Finally, axiom P5 (decomposability) reformulates the chain rule of probability theory in lottery terms, which accords the interpretation of lotteries as probability distributions. This axiom is sometimes called "no fun in gambling," since it prohibits one to place value on the number of steps needed to achieve an outcome.

Given a set of lotteries $L \subseteq \mathcal{L}(W)$ and a preference ordering $\preceq$ on $\mathcal{L}(W)$, we say that $l^* \in L$ is a *most preferred lottery* of L with respect to $\preceq$, when $l' \preceq l^*$ for all $l' \in L$. Note that most preferred lotteries need not be unique, but if there are multiple most preferred lotteries in a given set, then these lotteries are all equally preferred.

For a given set of lotteries $\mathcal{L}(W)$ and an associated preference ordering $\preceq$, it can be shown that the axioms of preference guarantee the existence of a real-valued function over $\mathcal{L}(W)$ that respects the decision maker's preferences with respect to these lotteries; such a function is called a *utility function*.

Theorem 3.20 (Utility function) *Let W be a set of variables, and let $\preceq$ be a preference ordering on $\mathcal{L}(W)$. Then, there exists a function $u_\preceq : \mathcal{L}(W) \to \mathbb{R}$ such that for all $l, l' \in \mathcal{L}(W)$ we have*

(U1) *$u_\preceq(l) \leq u_\preceq(l')$ if and only if $l \preceq l'$, and*

(U2) *if* $l = (p_1, l_1; \ldots; p_n, l_n)$, *then* $u_{\preceq}(l) = \sum_{i=1}^{n} p_i \cdot u_{\preceq}(l_i)$.

We say that $u_{\preceq}$ *is a* utility function *for preference ordering* $\preceq$. *The function value* $u_{\preceq}(l)$ *is called the* utility *of lottery* $l \in \mathcal{L}(W)$ *under function* $u_{\preceq}$.

Proof. See (Debreu, 1954). □

The rules U1 and U2 are often referred to as the *axioms of utility*, although they describe derived, not assumed, properties of utility functions. The first axiom states that the utility of a lottery l' is greater than the utility of a lottery l if and only if l' is preferred to l. The second axiom says that we can compute the utility of a compound lottery from the utilities of its constituents, using the rule of expectation from probability theory: if P_l is the probability distribution over W that corresponds to lottery l, then

$$u_{\preceq}(l) \quad = \quad E_{P_l}(u^m_{\preceq}), \tag{3.19}$$

where $u^m_{\preceq}$ is the function $u_{\preceq}$ restricted to Ω_W, i.e. $u^m_{\preceq} = u_{\preceq|\Omega_W}$. We refer to $u^m_{\preceq}(\mathrm{c}_W)$ as the *marginal utility* of outcome c_W.

The axioms of preference (P1–P5) guarantee not only the existence but also the uniqueness, up to positive linear transformations, of a utility function u for a given preference ordering, (Debreu, 1954). That is, if $u_{\preceq}$ is a utility function for preference ordering $\preceq$, then so is each function $u'_{\preceq}$ for which

$$u'_{\preceq}(l) \quad = \quad a \cdot u_{\preceq}(l) + b, \tag{3.20}$$

for all $l \in \mathcal{L}(W)$, where $a, b \in \mathbb{R}$, $a > 0$. The significance of this result is that the calibration commodity employed in utility functions is not the unit of outcomes' worths, but uncertainty itself. That is, we can choose an arbitrary non-empty interval $[u^{\min}, u^{\max}] \subset \mathbb{R}$ as the range of possible utilities, and calibrate the marginal utility of outcome $\mathrm{c}_W \in \Omega_W$ using

$$u^m_{\preceq}(\mathrm{c}_W) \quad = \quad p \cdot u^{\min} + (1-p) \cdot u^{\max} \tag{3.21}$$

if

$$\mathrm{c}_W \quad \simeq \quad (p, \mathrm{c}^-_W; 1-p, \mathrm{c}^+_W), \tag{3.22}$$

where c^-_W and c^+_W are the least and most preferred outcomes, respectively. Note that the continuity axiom (P3) guarantees the existence of the compound lottery in Equation 3.22, because $\mathrm{c}^-_W \preceq \mathrm{c}_W \preceq \mathrm{c}^+_W$. It also follows that $u^m_{\preceq}(\mathrm{c}^-_W) = u^{\min}$ and $u^m_{\preceq}(\mathrm{c}^+_W) = u^{\max}$. A utility function is said to be *normalised* when $u^{\min} = 0$ and $u^{\max} = 1$; such a function is unique given the corresponding preference ordering $\preceq$.

The *maximum expected utility* (MEU) criterion states that a decision maker, when confronted with a choice between multiple lotteries, should always choose the lottery with maximum (expected) utility.

Lemma 3.21 (MEU criterion) *Let W be a set of variables, let $\preceq$ be a preference ordering over $\mathcal{L}(W)$, and let $L \subseteq \mathcal{L}(W)$ be a non-empty set of lotteries over W. Then, l^* is a most preferred lottery of L with respect to $\preceq$ if and only if*

$$l^* \quad = \quad \text{argmax}_l\{u_{\preceq}(l) \mid l \in L\}, \tag{3.23}$$

where $u_{\preceq}$ is a utility function for $\preceq$.

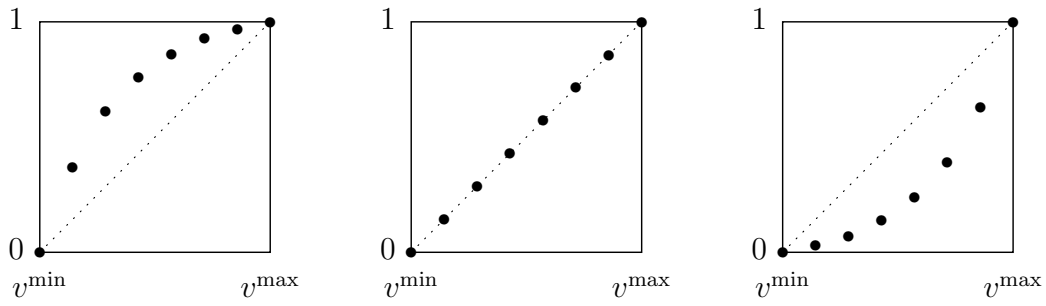
The MEU criterion follows from the axioms of preference and utility (see Heckerman, 1991, for an example of the proof). It is the main result of utility theory, as it provides a recipe for rational behaviour in situations of choice under uncertainty. It should be noted that such behaviour is guaranteed to yield preferred decisions (in the sense that they are consistent with one's preferences), but not necessarily preferred outcomes: the result of a preferred decision may still turn out to be undesirable due to the uncertainty involved.

Application of utility theory to concrete decision problems under uncertainty now boils down to three steps: (1) identification of a set of variables W that jointly cover all possible outcomes of decision making, (2) assessment of the marginal utility $u^m_{\preceq}(\mathrm{c}_W)$ for each outcome $\mathrm{c}_W \in \Omega_W$, and (3) choosing the decision alternative that maximises $E_P(u^m_{\preceq})$, where each decision alternative is taken to explicitly or implicitly yield a probability distribution P on W.

To facilitate the assessment of the marginal utility function $u^m_{\preceq}$, several qualitative characteristics of utility functions have been identified in the literature. Below, we will discuss two types of characteristics. The first type of characteristic is based on identifying a more or less objective worth, or *value*, with each lottery in $\mathcal{L}(W)$, and systematically comparing utilities with these values; this allows for assessing the *risk attitude* of the decision maker. The second type of characteristic is based on decomposition of the marginal utility function using subsets of variables from W; the individual variables are then referred to as *attributes*, and the composite function as a *multiattribute utility function*.

3.2.2 Risk attitudes

Assume that with each element c_W in the outcome space Ω_W is associated a bounded numerical *value* $v(\mathrm{c}_W) \in \mathbb{R}$. In a medical context, the value $v(\mathrm{c}_W)$ often represents life-expectancy of the patient under the circumstance described by c_W; in an economical setting, it represents the monetary gain or loss associated with the outcome c_W. The *expected value* of lottery $l \in \mathcal{L}(W)$ is now defined as $\tilde{v}(l) = E_P(v)$, and its *risk* as $r(l) = \text{var}_P(v)$, where P is the probability distribution on W that corresponds to l. Note that $\tilde{v}(\mathrm{c}_W) = v(\mathrm{c}_W)$ and $r(\mathrm{c}_W) = 0$ for all outcomes $\mathrm{c}_W \in \Omega_W$, as the probability distributions associated with outcomes are degenerate. A decision maker is now said to be *risk-neutral* if he is indifferent between all lotteries with the same expected value. He is *risk-averse* if he always prefers the lottery with the smallest risk in such



(a) Risk aversion (b) Risk neutrality (c) Risk proneness

Figure 3.2: Different risk attitudes for monotonically increasing utility. The x-axis represents value, and the y-axis represents (normalised) utility; the dots correspond to elements of Ω_W, the set of all outcomes (atomic lotteries).

cases; he is *risk-prone* if he always prefers the largest risk. This is formalised in the following definition.

Definition 3.22 (Risk attitude) *Let W be a set of variables, let $\preceq$ be a preference ordering on $\mathcal{L}(W)$, and let $v : \Omega_W \to \mathbb{R}$ be a value function for W. The decision maker is*

- risk-neutral *if $l' \simeq l$,*
- risk-averse *if $l \succ l'$, and*
- risk-prone *if $l \prec l'$,*

for all lotteries $l, l' \in \mathcal{L}(W)$ having $\tilde{v}(l) = \tilde{v}(l')$ and $r(l) < r(l')$. Risk aversion and risk proneness are more generally referred to as risk sensitivity.

Risk neutrality is generally viewed as optimal in situations where the objective is to maximise the cumulative result of a large sequence of similar decisions (e.g. gambling), whereas risk sensitivity is also applicable in single-decision situations (e.g. medical treatment).

The decision maker's attitude towards risk can also be recognised from a plot of the utilities and values in the (x, y)-plane; see Figure 3.2 for an example. We assume here *monotonically increasing utility*, i.e. $u^m_{\preceq}(\mathrm{c}_W) \geq u^m_{\preceq}(\mathrm{c}'_W)$ if $v(\mathrm{c}_W) \geq v(\mathrm{c}'_W)$. This is typically the case when v represents life-expectancy or monetary gain. Given a certain outcome $\mathrm{c}_W \in \Omega_W$, a risk-neutral decision maker is indifferent between c_W and any non-atomic lottery $l \in \mathcal{L}(W)$ having $\tilde{v}(l) = v(\mathrm{c}_W)$. That is, $u^m_{\preceq}(\mathrm{c}_W) = u_{\preceq}(l)$; we can in fact compute all marginal utilities as geometric means once the least and most

preferred outcomes have been calibrated. Under monotonically increasing utility, risk neutrality therefore implies that all outcomes $c_W \in \Omega_W$ are on the line

$$y = \frac{x - v^{\min}}{v^{\max} - v^{\min}} \tag{3.24}$$

where $v^{\min} = \min\{v(c_W) \mid c_W \in \Omega_W\}$ and $v^{\max} = \max\{v(c_W) \mid c_W \in \Omega_W\}$. A risk-averse decision maker will however prefer the certain outcome c_W to lottery l, as $r(c_W) = 0$ and $r(l) > 0$. The marginal utility $u^m_{\preceq}(c_W)$ associated with c_W will therefore be higher than one would expect by computing the geometric mean utility $u_{\preceq}(l)$ of lottery l; all certain outcomes except the least and most preferred are therefore located above the line of Equation 3.24 in this case. A similar argument can show that risk proneness implies that the certain outcomes lie below the line.

Under the (generally reasonable) assumption that the decision maker is indifferent between outcomes in Ω_W with equal values, it is possible to define marginal utility as a function of outcome value. That is, we define a function $u_v : \mathbb{R} \to \mathbb{R}$ with the restriction that

$$u_v(E_P(v)) = E_P(u_v \circ v) \tag{3.25}$$

for any joint probability distribution P on W (where '$\circ$' denotes functional composition), and take $u_{\preceq}(l) = u_v(\tilde{v}(l))$ for all $l \in \mathcal{L}(W)$. Calibration of u_v is achieved by assessing the marginal utility of elements in outcome space directly, and by establishing the decision maker's attitude towards risk. Note that monotonic utility now corresponds to monotonicity of the function u_v.

It can be shown that under monotonically increasing utility, a decision maker is risk-neutral if and only if u_v is a positive linear function, he is risk-averse if and only if his marginal utility function is concave, and he is risk-prone if and only if his marginal utility function is convex (Keeney and Raiffa, 1976); again, see Figure 3.2 for illustrations. Intuitively, these correspondences can be understood as follows. Given the prospect of obtaining some value v, $v^{\min} < v < v^{\max}$, a risk-neutral decision maker is precisely twice as eager to obtain $2v$, a risk-averse decision maker is less than twice as eager to obtain $2v$, and a risk-prone decision maker is more than twice as eager to obtain $2v$. A measure of risk aversion (proneness) is found in the *local risk aversion function*

$$q(x) = -\frac{\frac{d}{dx^2}u_v(x)}{\frac{d}{dx}u_v(x)}. \tag{3.26}$$

Concavity of u_v causes q to be positive, where higher values of q indicate stronger risk aversion; convexity of u_v causes q to be negative, where lower values of q indicate stronger risk proneness. Two utility functions are strategically equivalent (i.e. correspond to the same preference order) if and only if they have the same local risk aversion function (Keeney and Raiffa, 1976). Note that all risk-neutral utility functions are strategically equivalent, as $q(x) = 0$ for all x for these functions.

3.2.3 Multiattribute utility theory

In many decision problems under uncertainty, the assessment of a preference order on the possible outcomes of decision making is hampered by the fact that there are multiple objectives involved. It is very likely these objectives conflict with each other in that the improved achievement with one objective can only be accomplished at the expense of another. For instance, in a medical setting one may wish to maximise the patient's life-expectancy while minimising the costs of treatment. This requires, however, establishing a tradeoff between improvements in life-expectancy and the additional monetary costs.

Multiattribute utility theory (Keeney and Raiffa, 1976) is a framework for handling the value tradeoffs and uncertainties in multi-objective decision problems. In multi-attribute utility theory, the outcome space Ω_W is divided into subspaces using the variables $w_1, \ldots, w_k$ that constitute the set W; these variables are generally referred to as *attributes*. The basic approach is to identify regularities in the decision-making objectives that provide for decomposition of the utility function $u^m_\preceq$ to a form

$$u^m_\preceq(\mathrm{c}_W) \quad = \quad f(u_{w_1}(\mathrm{c}_{w_1}), \ldots, u_{w_k}(\mathrm{c}_{w_k})). \tag{3.27}$$

That is, with each attribute w_i, $i = 1, \ldots, k$, is associated a *local utility function* $u_{w_i} : \Omega_{w_i} \to \mathbb{R}$, and the utility values are subsequently combined by the function f to obtain unidimensional utility. Fundamental to the identification of regularities in the decision-maker's objectives are the concepts of *utility independence* and *additive independence*. These concepts pertain to the influence of individual attributes or sets of attributes on the decision maker's preferences: by restricting the possible preference orderings on $\mathcal{L}(W)$, they allow for simplifications in the associated utility function. Below, we assume both the global utility function $u^m_\preceq$ and all local utility functions to be normalised.

We first discuss the concept of utility independence. The underlying notion is the *conditional preference ordering*, which refers to the decision maker's preferences with respect to all lotteries over a set $Y \subseteq W$ of attributes, obtained from the preference ordering on $\mathcal{L}(W)$ by holding the other attributes fixed at a given value.

Definition 3.23 (Conditional preference ordering) *Let W be a set of variables, and let $\preceq$ be a preference ordering on $\mathcal{L}(W)$. Furthermore, let $Y \subseteq W$, and let $\mathrm{c}_{\overline{Y}} \in \Omega_{\overline{Y}}$ be a configuration of its complementary set $\overline{Y} = W \backslash Y$. The* conditional preference ordering *on $\mathcal{L}(Y)$ induced by $\mathrm{c}_{\overline{Y}}$, notation $\preceq_{\mathrm{c}_{\overline{Y}}}$, is defined as*

$$l_Y \preceq_{\mathrm{c}_{\overline{Y}}} l'_Y \quad \text{if} \quad l \preceq l' \tag{3.28}$$

for all lotteries $l_Y, l'_Y \in \mathcal{L}(Y)$, where l and l' are the lotteries over W obtained by replacing each $\mathrm{c}_Y \in \Omega_Y$ with $\mathrm{c}_Y \wedge \mathrm{c}_{\overline{Y}}$ in l_Y and l'_Y, respectively.

We now say that the set Y is *utility independent* of the remaining variables when the conditional preferences for lotteries over Y are the same, regardless of the configuration of $\overline{Y} = W \setminus Y$.

Definition 3.24 (Utility independence) *Let W be a set of variables, and let $\preceq$ be a preference ordering on $\mathcal{L}(W)$. The set $Y \subseteq W$ is* utility independent *(of its complementary set $\overline{Y}$) when each configuration $\mathrm{c}_{\overline{Y}} \in \Omega_{\overline{Y}}$ induces the same conditional preference ordering $\preceq_{\mathrm{c}_{\overline{Y}}}$ on $\mathcal{L}(Y)$. The variables from W are said to be* mutually utility independent *if every subset Y of W is utility independent of its complementary set $\overline{Y}$.*

When each configuration $\mathrm{c}_{\overline{Y}} \in \Omega_{\overline{Y}}$ induces the same conditional preference ordering $\preceq_{\mathrm{c}_{\overline{Y}}}$ on the lotteries over Y, it is reasonable to speak *the* preference ordering on $\mathcal{L}(Y)$ and similarly *the* utility function for Y, independently of the other attributes. As a result, we can speak of the least and most preferred configurations c_Y^- and c_Y^+ of Y, independently of the values of other attributes. When all attributes are utility independent, the following simplified form of utility function can be derived.

Theorem 3.25 (Multilinear utility) *If the variables of W are mutually utility independent, then there exist functions $u_w : \Omega_w \to \mathbb{R}$, $w \in W$, such that*

$$u^m_{\preceq}(\mathrm{C}_W) \;=\; \sum_{Y \subseteq W, Y \neq \varnothing} k^{|Y|-1} \prod_{w \in Y} k_w u_w(\mathrm{C}_w) \tag{3.29}$$

where $k_w = u^m_{\preceq}(\mathrm{c}^+_w \wedge \mathrm{c}^-_{W\setminus\{w\}})$ is the weight factor *for variable $w \in W$, and k is a scaling constant that is a solution to*

$$1 + k \;=\; \prod_{w \in W} (1 + k \cdot k_w). \tag{3.30}$$

Proof. See (Keeney and Raiffa, 1976). □

The result of this theorem seems impractical as the right-hand side of Equation 3.29, which is referred to as a *multilinear utility function*, is rather awkward. Fortunately, there exist simplified forms of this formula that are more intuitive.

The first form of simplified multilinear utility function is obtained from the assumption of *mutually additive independence* of the utility attributes, (Fishburn, 1970). Under mutually additive independence, we have that

$$u^m_{\preceq}(\mathrm{c}^+_W) \;=\; \sum_{w \in W} u^m_{\preceq}(\mathrm{c}^+_w \wedge \mathrm{c}^-_{W\setminus\{w\}}), \tag{3.31}$$

or equally, that $\sum_{w \in W} k_w = 1$. It then follows that $k = 0$, and the multilinear utility function reduces to the following *additive* form:

$$u^m_{\preceq}(\mathrm{C}_W) \;=\; \sum_{w \in W} k_w u_w(\mathrm{C}_w). \tag{3.32}$$

Intuitively, the attributes of utility are independent additive contributors to global utility, and in optimising an individual attribute $w \in W$, we do not have to care about

the values of other variables. The additive utility function is therefore characterised by much robustness; it is typically employed in cases where the attributes of utility represent independent factors of income.

The second form of simplified multilinear utility function is obtained from the assumption of *mutually multiplicative independence* of the utility attributes. This assumption states that for each configuration $c_w \in \Omega_w$ of an individual attribute $w \in W$, the ratio

$$\frac{u^m_{\preceq}(c_w \wedge c_{W\setminus\{w\}})}{u^m_{\preceq}(c^+_w \wedge c_{W\setminus\{w\}})} \tag{3.33}$$

is the same for all configurations $c_{W\setminus\{w\}}$ of the set $W \setminus \{w\}$ of remaining attributes. We can then use $u_w(c_w)$ to denote this ratio, and the multilinear utility function reduces to the following *multiplicative* form:

$$u^m_{\preceq}(C_W) \;=\; k^{|W|-1} \cdot \prod_{w\in W} k_w u_w(C_w). \tag{3.34}$$

It models a situation where each fluctuation in local utility has a proportional effect on global utility. A multiplicative utility function is therefore not very robust: it is sometimes compared to a chain that is only as strong as its weakest link. One typically employs such a function in the case where each attribute of utility represents an independent factor of risk.

3.2.4 Quasi-utility functions

The axioms of preference from Definition 3.19 have not been beyond dispute. Over the years, each of the axioms P1 through P5 has been criticised, for various reasons (see Bell and Raiffa, 1988, for a discussion), and many alternatives have been formulated (e.g. Fishburn, 1988). Although the five axioms still constitute the most popular formalisation of preference under uncertainty, there are sometimes reasons to depart from them. Evaluation functions that violate one or more of the axioms of preference are called *quasi-utility functions*. In this subsection, we discuss such a function and motivate the reasons for its application.

Suppose that the decision maker indicates that he wants to minimise the uncertainty with respect to the set W of variables. That is, atomic lotteries $c_W \in \Omega_W$ are preferred over (non-degenerate) non-atomic lotteries. More generally put, from a set of lotteries $L \subseteq \mathcal{L}(W)$, the decision maker prefers the lottery $l \in L$ that minimises the entropy $H_{P_l}(W)$ of W, or equivalently, maximises the negative entropy $-H_{P_l}(W)$. An appropriate utility would therefore seem

$$u(l) \;=\; -H_{P_l}(W). \tag{3.35}$$

However u is *not* a utility function, as it violates the second axiom of utility (U2; see Theorem 3.20 on p. 58): the function regards atomic lotteries as equally (and

maximally) preferred. In a given lottery

$$l \;=\; (p_1, \mathrm{c}_W^1; \ldots; p_n, \mathrm{c}_W^n) \tag{3.36}$$

we can therefore substitute each c_W^i, $i = 1, \ldots, n$ with some arbitrary $\mathrm{c}_W \in \Omega_W$ because it is equally preferred. Then, we obtain the equivalent lottery

$$l' \;=\; (p_1, \mathrm{c}_W; \ldots; p_n, \mathrm{c}_W). \tag{3.37}$$

However, l' yields the outcome c_W with certainty, and is therefore equivalent to the atomic lottery c_W itself. As a result, all lotteries from $\mathcal{L}(W)$ should be equally preferred under the preference ordering encoded by the function u. This however contradicts the fact that u will yield lower (negative) values when a lottery involves much uncertainty. The function u is therefore not a utility function but a quasi-utility function.

The reasons for the invalidity of the function u as a utility function can intuitively be understood from the following principle: it is only useful to strive for more certainty when this allows for appropriate (i.e. utility-increasing) action. In a medical setting, for instance, gathering more information on the clinical state of a patient by performing a diagnostic test is considered useful only when the resulting improved diagnosis allows for better treatment. Otherwise, the adverse effects of the test such as stress, costs, or risks prove unnecessary and the test should be avoided.

Although the above principle is generally held to be sound, there may still be reasons to use a quasi-utility function that solely aims to decrease uncertainty. We distinguish three possible reasons for doing so. Firstly, information may be of prognostic value independent of its usefulness in determining treatment, (Asch et al., 1990). For instance, many patients would be willing to pay substantial sums of money simply to know more about their disease and its probable developments. Secondly, the formalisation of the decision problem may not cover all potential treatment strategies. In particular, this may be so when the formalisation aims to support only the diagnostic process, (Gorry and Barnett, 1968). Thirdly, some (complex) decision problems suffer from a combinatorial explosion in the number of possible decision policies, and can only be solved using heuristic measures. The quasi-utility function u may then be helpful in estimating the value of diagnostic tests, (Glasziou and Hilden, 1989). We will return to complexity issues in decision making at the end of the next section.

3.3 Decision-theoretic analysis

Probability theory and utility theory provide the building blocks for *decision theory*, the discipline that formulates the principles of rational decision making under conditions of uncertainty. The art and science of analysing real-life decision problems from these principles is called *decision analysis*. In this section, we provide a sketch of the field by analysing a number of prototypical decision problems under uncertainty using

Action	*Interpretation*
echo	echocardiography
med	medical treatment
cath	cardiac catheterisation
biop	open lung biopsy
surg	perform surgery

Table 3.2: Example decision alternatives for the VSD domain.

decision-theoretic principles. Decision-theoretic reasoning is characterised by the fact that each situation of choice is ultimately reduced to a utility-theoretic tradeoff; from a conceptual point of view, however, these choice situations may be very different. We will see that a number of concepts, pertaining to the role of actions, information and time, are significant in many decision problems. While in utility theory, these concepts are left implicit, they are typically made explicit in decision-theoretic analyses. This section aims to introduce the reader to these concepts and their role decision-theoretic reasoning.

Below, we analyse three simple decision problems: (i) making a single choice under uncertainty without prior information, (ii) making such a choice with problem-specific information, and (iii) a two-stage decision problem comprising a choice whether or not to gather information and a choice that can exploit that information. In the analyses, we confine ourselves to descriptions from sets of variables with a finite set of possible values. The choices faced by a decision maker are represented by *decision variables*; the values that a decision variable may take represent a mutually exclusive and exhaustive set of actions from which a choice is made at some point in the decision problem. Such actions may include the options of waiting for something to happen, inspecting or activating a given measuring device, and gathering or investing physical resources to enable future actions. We will refer to configurations of decision variables as *decision alternatives*. The uncertain events that are relevant to the decision problem but are beyond the (direct) control of the decision maker are represented by random variables. In the terminology of Subsection 1.2.2, these random variables describe a *system* under control by the decision maker, and their possible configurations represent potential *system states*; we will henceforth refer to them as *state variables*.

Example 3.26 *Table 3.2 provides an example set of clinical actions that are available to a paediatric cardiologist for the management of VSD patients. Medical treatment may be used to control heart failure, and surgery may be used to close the defect. Information concerning the clinical state of the patient may be obtained by making echocardiographic images, cardiac catheterisation, and open lung biopsies. Throughout the remainder of this chapter, we will use examples with decision variables that range over this set of actions, and state variables that are taken from Table 3.1.*

To visualise the various types of decision problem, we will use *decision trees* (Raiffa, 1968), the traditional tool for decision-theoretic analysis. A decision tree is a rooted tree that provides an explicit, graphical enumeration of the potential scenarios involved in a given decision problem; each path from root to leaf describes such a scenario. The internal nodes of the tree represent either decisions or uncertain events, and the leaf nodes represent outcomes of the decision process. As such, a decision tree highlights the structural components of the problem, i.e. (i) the alternative actions that are available to the decision maker, (ii) the events that follow from and affect these actions, and (iii) the outcomes that are associated with each possible scenario of actions and events. We note that the structure of a decision-theoretic analysis and the corresponding decision tree describe a decision problem from the perspective of the person solving the decision problem. An analysis is limited to the details of the problem relevant for choosing the optimal decision alternative(s), and ignores all other characteristics of the problem.

Decision making in ignorance

The simplest type of decision problem under uncertainty is where there is a single moment of choice, and there is no problem-specific information available prior to that decision: the decision maker is completely ignorant of the current state of affairs. The decision maker's choice influences the system, and has the objective of reaching a satisfying system state, but the decision alternatives themselves are also subject to a tradeoff as they involve particular costs or risks. Using the results of Section 3.2, we will assume the decision maker's preferences are expressed as marginal utilities for all combinations of states and choices; the underlying preference order is left implicit from now on.

Example 3.27 *Faced with a patient that has a VSD, the primary decision that needs to be made by the clinician is whether or not to submit this patient to surgery, i.e. whether or not to select the decision alternative* `surg`*. On the one hand, operating the patient will generally result in successful closure of the defect (*`size` $=$ *null), improving the patient's condition (*`shunt` $=$ *none*, `hfail` $=$ *absent), and eliminating the risks of complications such as Eisenmenger's complex (represented by high values of the variable* `resis`*). On the other hand, a number of patients do not survive open-heart surgery (i.e.* `death` $=$ *true).*

Formally, let d be a decision variable, and let X be a set of state variables whose joint probability distribution depends on d. Let P_{c_d} be the distribution on X associated with decision alternative $\mathrm{c}_d \in \Omega_d$. For convenience, we will use $p(\mathrm{c}_X \mid \mathrm{c}_d)$ as a shorthand notation for $P_{\mathrm{c}_d}(\mathrm{c}_X)$.[1] Furthermore, a marginal utility function u is defined

[1]Note that, notwithstanding the notation, $p(\mathrm{c}_X \mid \mathrm{c}_d)$ is *not* a conditional probability as there are no marginal probabilities defined for variable d.

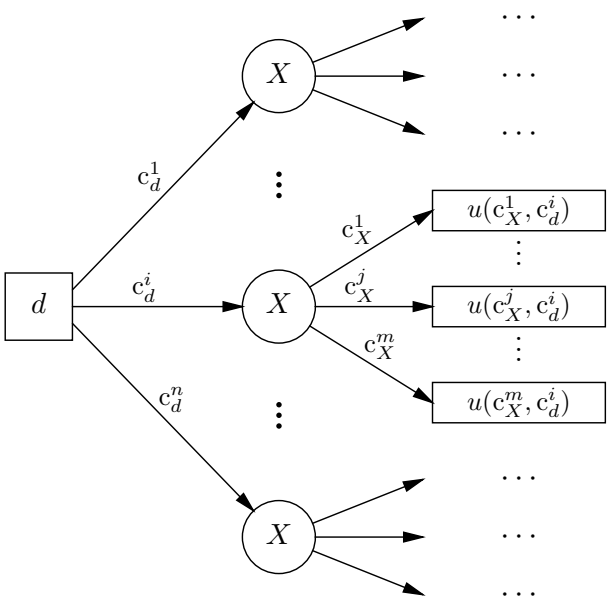


Figure 3.3: Decision tree representing singular choice under uncertainty.

over $\Omega_X \times \Omega_d$. The *expected utility* of alternative c_d, denoted by $\tilde{u}(c_d)$, now equals

$$\tilde{u}(c_d) \quad = \quad \sum_{c_X \in \Omega_X} u(c_X, c_d) \cdot p(c_X \mid c_d). \tag{3.38}$$

Following the MEU criterion, a *preferred*, or *optimal* alternative c_d^* is one that maximises expected utility, i.e.,

$$c_d^* \quad = \quad \text{argmax}_{c_d}\{\tilde{u}(c_d) \mid c_d \in \Omega_d\}. \tag{3.39}$$

Note that there need not be a unique preferred alternative. The term *regret* refers to the expected utility that is 'missed' by choosing a sub-optimal decision alternative.

Definition 3.28 (Regret) *The* regret *of decision* c_d *is defined as* $\tilde{u}(c_d) - \tilde{u}(c_d^*)$, *where* c_d^* *is an optimal alternative.*

It is easily seen that regret is non-positive, and the lower the regret, the worse the decision.

The decision problem described above is represented by the decision tree of Figure 3.3. Decisions are depicted by squares, called *decision nodes*, and labelled with the name of the corresponding decision variable; the branches emanating from a decision node correspond to the possible decision alternatives (in this case $c_d^1, \ldots, c_d^n$). Uncertain events are depicted by circles, called *chance nodes*, and labelled with the name of

the corresponding set of random variables. In general, there may be multiple nodes in a decision tree representing the same decision or uncertain event (albeit under different circumstances); in this case, there are n chance nodes, each representing the uncertain event with possible outcomes in the universe Ω_X of the set X. The branches emanating from a chance node correspond to these possible outcomes (in this case $\mathrm{c}_X^1, \ldots, \mathrm{c}_X^m$). With each of the branches is associated the probability that the event takes place in that context; for instance, the upper branch emanating from the middle chance node has associated probability $p(\mathrm{c}_X^1 \mid \mathrm{c}_d^i)$. These probabilities are not shown in the figure. Leaf nodes of the tree represent potential outcomes of the decision-making process; each leaf node is labelled with the utility that corresponds to the sequence of decisions and events along the unique path to that node starting at the root.

As is seen from the figure, the decision-tree representation of decision problems reflects the utility-theoretic conception of decisions as choices between lotteries. By *evaluating* or *solving* a decision tree is meant finding an optimal *policy* for the tree, i.e. selecting an optimal branch for each decision node in the tree. In this case, there is only one decision node, and an optimal branch is one which corresponds to an alternative that satisfies the condition in Equation 3.39.

We conclude the analysis of this decision problem by noting that it is well possible that one or more state variables are unaffected by the decision, i.e. that there exists a subset $Y \subseteq X$ for which

$$P_{\mathrm{c}_d^i}(\mathrm{C}_Y) \;=\; P_{\mathrm{c}_d^j}(\mathrm{C}_Y) \tag{3.40}$$

for all $\mathrm{c}_d^i, \mathrm{c}_d^j \in \Omega_d$. We can then regard these variables as receiving their values prior to decision making. Recall, however, that a decision analysis solely reflects the perspective of the decision maker. The circumstance described therefore has no consequences for the analysis of the decision problem, as we have assumed the decision maker to be ignorant of the values of state variables when making the decision. Below, we will lift the assumption of complete ignorance, and analyse the effect of observations.

Decision making with prior observations

It often comes about that a decision maker has some information regarding the state of the system to which his decisions pertain. In our formalisation, this information consists of configurations of one or more state variables; we will refer to these configurations as *evidence*. We will first discuss *case parameters*, the special type of evidence that is received prior to making the (first) decision, and turn later to observations that are received as a result of specific choices by the decision maker. In a medical setting, case parameters typically consist of patient-specific information such as personal and historical data, symptoms, and findings from physical examination.

Example 3.29 *In the VSD example, case parameters consist of externally visible*

signs and symptoms, typically caused by heart failure and shunting: shortness of breath, feeding and growing problems (comprised in the variable `symp` *in Table 3.1), and occurrence of central cyanosis (*`cyan` = *true).*

Formally, let as before X be a set of random variables, let d be a decision variable, and let u be a marginal utility function over $\Omega_X \times \Omega_d$. Now, assume that $Y \subseteq X$ is a set of state variables that receive their values prior to decision making, and are observed at that time. We use $p_Y(\mathrm{c}_Y)$ to denote the marginal probability of configuration c_Y, i.e. $p_Y(\mathrm{c}_Y) = P_{\mathrm{c}_d}(\mathrm{c}_Y)$ for any $\mathrm{c}_d \in \Omega_d$, and use

$$p_{X\setminus Y}(\mathrm{c}_{X\setminus Y} \mid \mathrm{c}_Y, \mathrm{c}_d) \;=\; P_{\mathrm{c}_d}(\mathrm{c}_{X\setminus Y} \mid \mathrm{c}_Y) \tag{3.41}$$

to denote the conditional probability of configuration $\mathrm{c}_{X\setminus Y}$ given c_Y when choosing alternative c_d; the derived conditional distribution $p_{X\setminus Y}$ can be seen as describing the effects of the variables from Y on the variables from $X \setminus Y$ under the various decision alternatives.

Example 3.30 *Let* $X = \{\mathtt{shunt}, \mathtt{hfail}\}$ *and consider the decision to administer cardiac glycosides to enhance the strength of myocardial contraction. This decision will not affect the shunt, but it does in general reduce heart failure. Taking* $Y = \{\mathtt{shunt}\}$*, the conditional probability* $p_{X\setminus Y}$ *describes the effects of shunting on heart failure under both decision alternatives. In general, larger shunt sizes will increase the risk of (severe) heart failure, but the risk will decrease under treatment with glycosides.*

The decision tree depicted in Figure 3.4 models the decision problem with prior observations. The root of the tree is a chance node representing the observed set Y of state variables, and is followed by decision d. The configuration of Y, which is then known to the decision maker, is used to optimise the decision. Generally speaking, all variables preceding a decision node in a decision tree are assumed to be known when making the decision; this may also include earlier decisions and observations that result from specific choices. Chance nodes not followed by a decision node, in this case representing the set $X \setminus Y$, either represent attributes of the system that remain hidden from observation, or model an uncertain event in the future. Recall that with each branch emanating from a chance node is associated the conditional probability of the uncertain event given the history of past decisions and events along the path that leads from the root to the node; here, the history consists of configurations of the set $Y \cup \{d\}$.

The form of the tree stresses the fact that with each observation from Ω_Y, we face a different subproblem, and each of these subproblems may have its own optimal solution. To formulate solutions to observation-dependent subproblems, we use *decision functions.*

Definition 3.31 (Decision function) *Let W be a set of variables, and let d be a decision variable. A* decision function *for d is a function $\delta : \Omega_W \to \Omega_d$. Configu-*

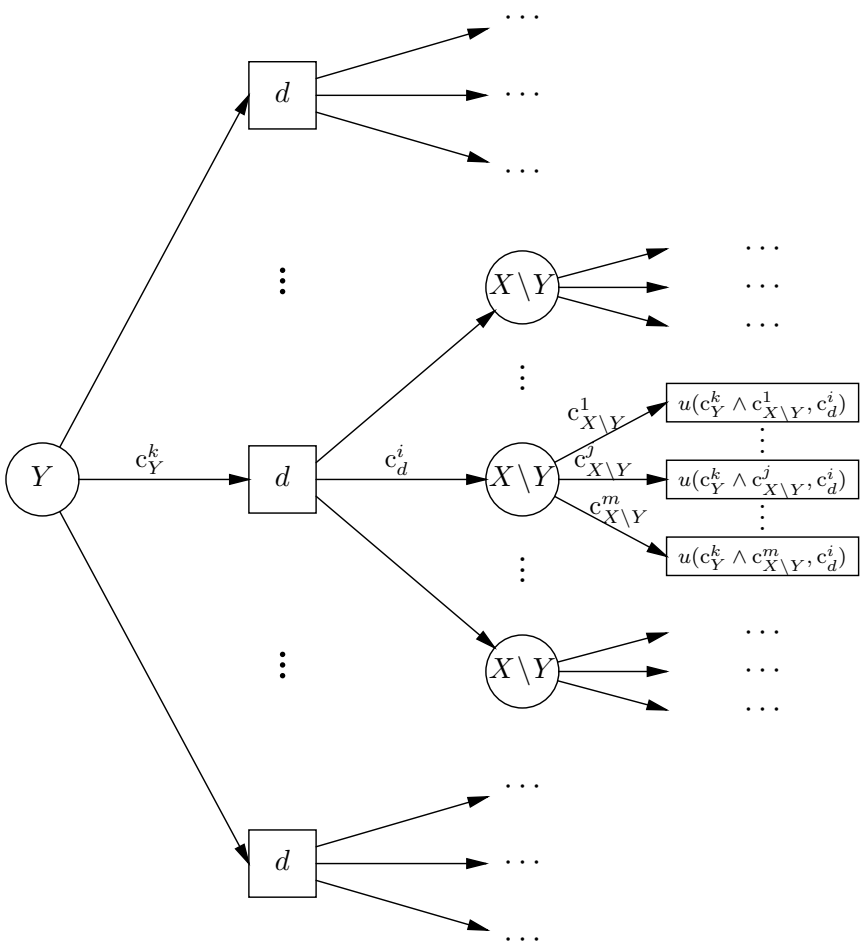


Figure 3.4: Decision tree with an observation prior to decision making.

rations of W are called inputs *to the decision function δ. The set of all decision functions for d with inputs from Ω_W is written $\Delta_{W\to d}$.*

Both configurations of decision variables and random variables can serve as inputs to a decision function. Note that when $W = \varnothing$, then the function simply picks an alternative from Ω_d; such functions can therefore be used for decisions without prior observations.

The problem faced by the decision maker is to solve the decision tree by selecting an optimal branch for each of the tree's decision nodes, or equivalently, to choose the optimal decision function for d with inputs from Ω_Y. We define the *conditional expected utility* of decision $\mathrm{c}_d \in \Omega_d$ given observation c_Y, notation $\tilde{u}(\mathrm{c}_d \mid \mathrm{c}_Y)$, as

$$\tilde{u}(\mathrm{c}_d \mid \mathrm{c}_Y) \quad = \sum_{\mathrm{c}_{X\setminus Y} \in \Omega_{X\setminus Y}} u(\mathrm{c}_Y \wedge \mathrm{c}_{X\setminus Y}, \mathrm{c}_d) \cdot p_{X\setminus Y}(\mathrm{c}_{X\setminus Y} \mid \mathrm{c}_Y, \mathrm{c}_d), \tag{3.42}$$

and the (unconditional) expected utility $\tilde{u}(\delta)$ of decision function δ as

$$\tilde{u}(\delta) \;=\; \sum_{\mathrm{c}_Y \in \Omega_Y} \tilde{u}(\delta(\mathrm{c}_Y) \mid \mathrm{c}_Y) \cdot p_Y(\mathrm{c}_Y). \tag{3.43}$$

The optimal decision function is

$$\delta^* \;=\; \mathrm{argmax}_\delta\{\tilde{u}(\delta) \mid \delta \in \Delta_{Y \to d}\}, \tag{3.44}$$

or equivalently, the function for which

$$\delta^*(\mathrm{c}_Y) \;=\; \mathrm{argmax}_{\mathrm{c}_d}\{\tilde{u}(\mathrm{c}_d \mid \mathrm{c}_Y) \mid \mathrm{c}_d \in \Omega_d\}. \tag{3.45}$$

for all $\mathrm{c}_Y \in \Omega_Y$. The latter formulation reflects the fact that there are $|\Omega_Y|$ different subproblems, for each of which an optimal solution has to be found. We remark that the notion of regret from Definition 3.28 can be defined analogously for decision functions.

The following notion is due to Howard (1966).

Definition 3.32 (Value of information) *The* expected value of information *of the set Y, notation* $\mathrm{EVI}(Y)$*, is defined as*

$$\mathrm{EVI}(Y) \;=\; \tilde{u}(\delta^*) - \tilde{u}(\mathrm{c}_d^*), \tag{3.46}$$

where $\tilde{u}(\delta^*)$ is the optimal decision function with inputs from Ω_Y, and c_d^* is the optimal decision alternative without prior information.

If Y is the largest set of state variables uninfluenced by the decision, this means that an observed configuration of Y represents all there is to know of the problem prior to decision making. We then refer to $\mathrm{EVI}(Y)$ as the *expected value of perfect information* (EVPI).

Proposition 3.33 now expresses the decision-theoretic principle that a decision maker should always use as much of the available information as possible, as long as the information is freely available.

Proposition 3.33 *The expected value of information* $\mathrm{EVI}(Y)$ *of any set Y is non-negative.*

Proof. Let c_d^* as before be the decision alternative that expectedly optimises utility when there is no prior information, and let δ^b be the constant decision function that selects c_d^* whatever evidence on the set Y is received (where b refers to the 'blindness' of this decision function). From Equation 3.43, we have that

$$\tilde{u}(\delta^b) \;=\; \textstyle\sum_{\mathrm{c}_Y \in \Omega_Y} \tilde{u}(\mathrm{c}_d^* \mid \mathrm{c}_Y) \cdot p_Y(\mathrm{c}_Y) \;=\; \tilde{u}(\mathrm{c}_d^*). \tag{3.47}$$

Now, as δ^* maximises $\tilde{u}(\delta^*(\mathrm{c}_Y) \mid \mathrm{c}_Y)$ for each $\mathrm{c}_Y \in \Omega_Y$ (Equation 3.45), we have that

$$\tilde{u}(\delta^*(\mathrm{c}_Y) \mid \mathrm{c}_Y) \quad \geq \quad \tilde{u}(\delta^b(\mathrm{c}_Y) \mid \mathrm{c}_Y) \tag{3.48}$$

for all $\mathrm{c}_Y \in \Omega_Y$, and therefore

$$\tilde{u}(\delta^*) \quad \geq \quad \tilde{u}(\delta^b), \tag{3.49}$$

so

$$\tilde{u}(\delta^*) \quad \geq \quad \tilde{u}(\mathrm{c}_d^*). \tag{3.50}$$

□

It appears from the proof that the expected value of information of Y can be viewed as the negative regret of ignoring the evidence by using the blind decision function δ^b.

We note that it is possible to re-formulate a decision problem with prior observations as a singular choice without prior observations. Let d' be a new decision variable, taking values from $\Delta_{Y \to d}$, and define, for each alternative $\delta \in \Delta_{Y \to d}$,

$$p'_{X \setminus Y}(\mathrm{c}_{X \setminus Y} \mid d' = \delta) \quad = \quad \sum_{\mathrm{c}_Y \in \Omega_Y} p_Y(\mathrm{c}_Y) \cdot p_{X \setminus Y}(\mathrm{c}_{X \setminus Y} \mid \mathrm{c}_Y \wedge \delta(\mathrm{c}_X)), \tag{3.51}$$

and

$$\tilde{u}'(d' = \delta) \quad = \quad \sum_{\mathrm{c}_{X \setminus Y} \in \Omega_{X \setminus Y}} u(\mathrm{c}_Y \wedge \mathrm{c}_{X \setminus Y}, \delta(\mathrm{c}_Y)) \cdot p'_{X \setminus Y}(\mathrm{c}_{X \setminus Y} \mid d' = \delta). \tag{3.52}$$

We then have that $\tilde{u}'(d' = \delta) = \tilde{u}(\delta)$. In this formulation, the corresponding decision tree again has the form of the tree depicted in Figure 3.3. The number of branches emanating from the root decision will however be much larger, as there exist n^z possible decision functions for d with inputs from Y if $n = |\Omega_d|$ is the number of decision alternatives and $z = |\Omega_Y|$ is the number of possible observations. This is much more than the number of n decision alternatives when there is no prior evidence. We conclude that the possibility to observe prior evidence does not introduce a fundamental difference to the type of tradeoff involved, but there is a substantial increase in complexity of the problem.

Deciding upon intermediate observations

In the previous decision problem, we assumed that evidence was freely available, and therefore provided a guaranteed increase in expected utility. In many decision problems, evidence concerning the system state is however not freely available but comes forth as a result of specific choices made during the decision process; we will refer to such choices as *test decisions*.

Definition 3.34 (Test decision) *A* test decision *is a decision with the objective to gather evidence.*

On the one hand, the evidence obtained from test decisions can be used to optimise subsequent decisions, but on the other hand the conduction of tests will generally involve certain costs or risks. In a decision-theoretic analysis, these costs and risks are discounted as a decrease in utility. The question is therefore whether this immediate decrease in utility is justified by the expected increase in utility that results from improved future decision making.

Example 3.35 *A cardiologist may choose to make an echographic image of the heart of a VSD patient (the decision alternative* `echo`*), providing information about the size of the VSD (the state variable* `size`*). In addition, cardiac catheterisation (*`cath`*) may be used to obtain evidence on shunting and vascular resistances (variables* `shunt` *and* `resis`*), and the state of the pulmonary arterioles (*`pmart`*) can be examined by open lung biopsy (the alternative* `biop`*). Each of these types of evidence is helpful in assessing the severity of disease, and therefore in deciding upon the need for surgery.*

We formalise this situation as follows. Let d_1, d_2 be decision variables, where d_1 denotes a test decision; it takes one of the values *test* and *no_test*. When $d_1 = test$, the decision maker receives evidence on the set $Y \subseteq X$ of state variables; otherwise, he has no information on the state of the system. The probability distribution P on X is not influenced by d_1 and therefore only parametrised by d_2; the utility function u depends on X, d_1, and d_2.

The decision tree that corresponds to this problem is shown in Figure 3.5. As appears from the figure, the test decision is basically a choice between a decision with prior observations (represented by the upper half of the tree) and the same decision without observations (represented by the lower half of the tree). Also note that in contrast with the trees of Figures 3.3 and 3.4, the tree of Figure 3.5 is not symmetrical: the two halves of the tree are structurally different. We will refer to this phenomenon, which is induced by test decisions, as *informational asymmetry*. Informational asymmetry abounds in all decision problems that involve test decisions, and is therefore a significant phenomenon in decision-theoretic representation and reasoning.

Due to the asymmetric nature of the problem, the analysis is also split into two parts: one part where it is decided to perform the test, and one where it is decided to skip it. First, suppose that $d_1 = test$ is selected. This means that information on the configuration of Y becomes available, and we can use this information to optimise decision d_2. That is, we will then choose a decision function $\delta \in \Delta_{Y \to d_2}$ for d_2, and the expected utility of this function is

$$\tilde{u}(\delta \mid d_1 = test) \;=\; \sum_{c_Y \in \Omega_Y} p_Y(c_Y) \cdot \tilde{u}(\delta(c_Y) \mid c_Y, d_1 = test), \tag{3.53}$$

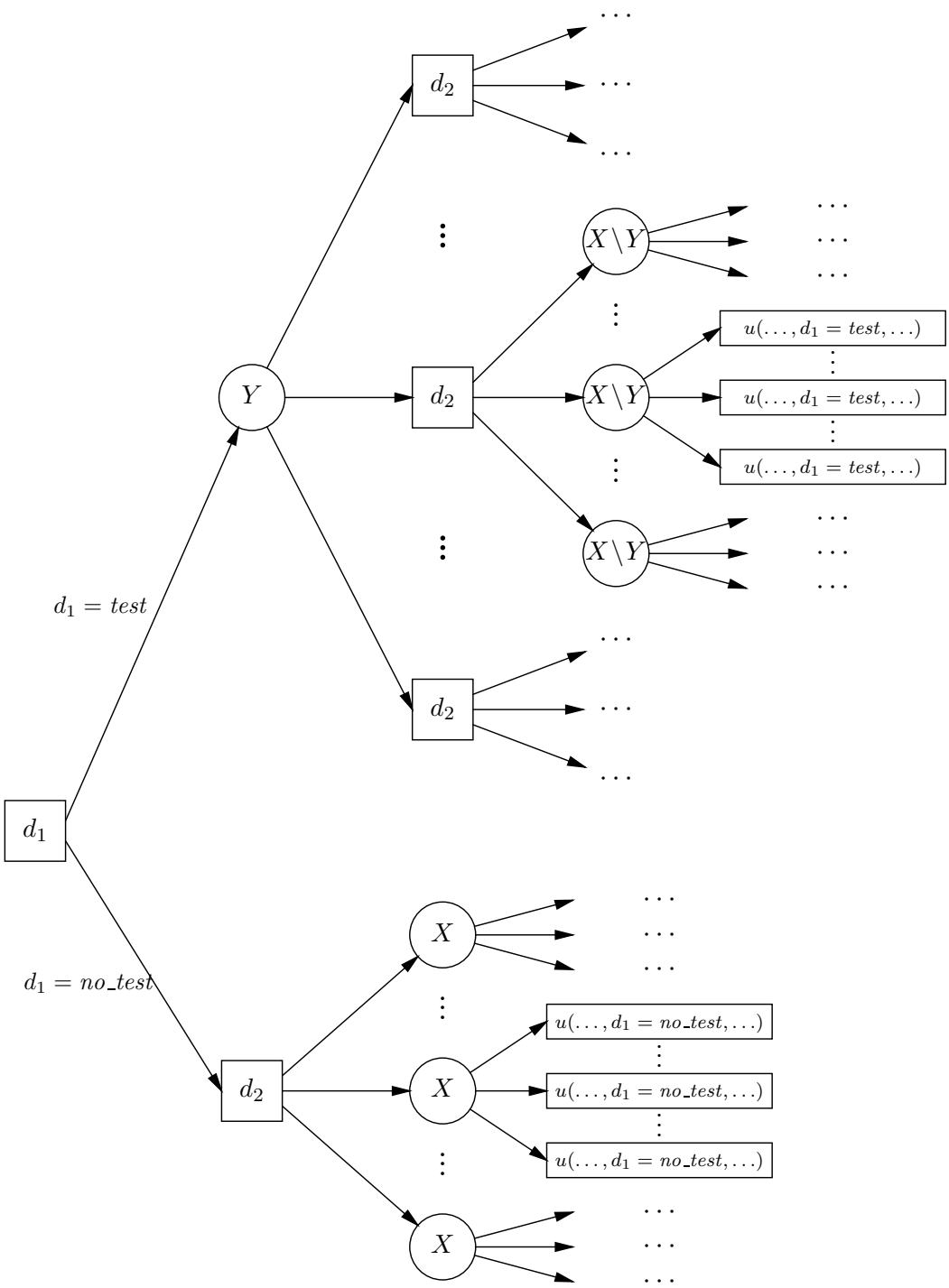


Figure 3.5: Decision tree with a test decision.

where the expected utility $\tilde{u}(\mathrm{c}_d \mid \mathrm{c}_Y, d_1 = \mathit{test})$ of decision alternative c_{d_2} given $d_1 = \mathit{test}$ and evidence c_Y is defined as

$$\tilde{u}(\mathrm{c}_{d_2} \mid \mathrm{c}_Y, d_1 = \mathit{test}) \quad = \quad \sum_{\mathrm{c}_{X\setminus Y} \in \Omega_{X\setminus Y}} u(\mathrm{c}_Y \wedge \mathrm{c}_{X\setminus Y}, d_1 = \mathit{test}, \mathrm{c}_{d_2}) \cdot p_{X\setminus Y}(\mathrm{c}_{X\setminus Y} \mid \mathrm{c}_Y \wedge \mathrm{c}_{d_2}). \tag{3.54}$$

Note that $\delta(\mathrm{c}_Y)$ in Equation 3.53 yields such a decision alternative for decision d, based on the evidence c_Y. The (maximum) expected utility of making the observation is therefore

$$\tilde{u}(d_1 = \mathit{test}) \quad = \quad \max\{\tilde{u}(\delta \mid d_1 = \mathit{test}) \mid \delta \in \Delta_{Y \to d_2}\}. \tag{3.55}$$

In contrast, when we decide to choose *no_test* for decision d_1, then no information about Y comes available and all we can do is select the alternative from Ω_{d_2} that is expected to be optimal given the prior distribution on X. The expected utility of alternative $\mathrm{c}_{d_2} \in \Omega_{d_2}$ equals

$$\tilde{u}(\mathrm{c}_{d_2} \mid d_1 = \mathit{no\_test}) \quad = \quad \sum_{\mathrm{c}_X \in \Omega_X} p_X(\mathrm{c}_X) \cdot u(\mathrm{c}_X, d_1 = \mathit{no\_test}, \mathrm{c}_{d_2}), \tag{3.56}$$

and the (maximum) expected utility of not observing is

$$\tilde{u}(d_1 = \mathit{no\_test}) \quad = \quad \max\{\tilde{u}(\mathrm{c}_{d_2} \mid d_1 = \mathit{no\_test}) \mid \mathrm{c}_{d_2} \in \Omega_{d_2}\}. \tag{3.57}$$

If $\tilde{u}(d_1 = \mathit{test}) > \tilde{u}(d_1 = \mathit{no\_test})$, then we perform the test and apply the optimal decision function for variable d_2 afterwards; otherwise, we do not make the observation and select the optimal alternative from Ω_{d_2}.

The difference $\tilde{u}(d_1 = \mathit{test}) - \tilde{u}(d_1 = \mathit{no\_test})$ in maximum expected utilities between testing and not testing is sometimes referred to as the *expected test value*. Note that when the test decision does not affect utility, i.e. when

$$u(\mathrm{C}_X, d_1 = \mathit{test}, \mathrm{C}_{d_2}) \quad = \quad u(\mathrm{C}_X, d_1 = \mathit{no\_test}, \mathrm{C}_{d_2}) \tag{3.58}$$

then the expected test value equals the expected value of information $\mathrm{EVI}(Y)$ of the set Y. If follows from Proposition 3.33 that the expected test value then be non-negative, and it is therefore recommended to perform it. In most practical settings, however, information will not be available 'for free'. That is, we usually have that

$$u(\mathrm{C}_X, d_1 = \mathit{test}, \mathrm{C}_{d_2}) \quad < \quad u(\mathrm{C}_X, d_1 = \mathit{no\_test}, \mathrm{C}_{d_2}), \tag{3.59}$$

and there is a tradeoff between loss in utility on the one hand and gain in information that may compensate that loss on the other hand.

We conclude this subsection with a few notes on the generalised problem where multiple tests are available that can be conducted serially. An example of this situation is found in the differential diagnosis task described in Subsection 2.2.2, where the

physician repeatedly selects diagnostic tests until sufficient information is available to choose appropriate therapy. Firstly, if the tests can be performed in any order, their expected values depend on the actual order in which the tests are performed, and can therefore not be assessed in isolation. The problem becomes highly combinatorial as in principle, any sequence of tests should be considered.

Example 3.36 *Suppose that a paediatric cardiologists suspects that a VSD patient suffers from pulmonary hypertension due to pulmonary arteriopathy. If his suspicion is correct, surgical closure of the VSD will worsen the patient's condition. He may now consider the tests* `cath` *(cardiac catheterisation) to inspect intra-cardiac flows and pressures and* `biop` *(open lung biopsy) to inspect the pulmonary arterioles directly. Both tests are invasive but not completely reliable. When a test is chosen and the results are equivocal, additional testing is required.*

Secondly, the existence of multiple tests will generally induce *relevantial asymmetry* in the formal analysis of the problem. This can be understood as follows. Suppose that $d_1, \ldots, d_k$ denote test decisions, where each d_i, $i = 1, \ldots, k$, takes values from an arbitrary set of tests A. Now, if $a \in A$ denotes a test that is guaranteed to provide evidence that is free of measurement errors, then selecting $d_i = a$, $i = 1, \ldots, k-1$, rules out the necessity to consider test a for any future decision d_j, $j = i+1, \ldots, k$, as conducting the test would not provide any further information: the option a has become *irrelevant*. Relevantial asymmetry is reminiscent of informational asymmetry, but pertains to decision variables instead of state variables; it generally occurs in problems where decisions (not necessarily pertaining to tests) may be repeated.

Example 3.37 *Consider once more the above example. If we assume that an open lung biopsy provides reliable information on the state of the pulmonary arterioles, then it is unnecessary to repeat this test once it has been performed.*

3.4 Discussion

In the previous sections, we have discussed the formal foundations for decision making under uncertainty within the decision-theoretic paradigm. It was described how the synthesis of probability theory and utility theory provides a framework for analysing various types of choice under uncertain conditions. Within this framework, probability theory serves to formalise the reasoning about uncertain events, whereas utility theory provides the guidelines for rational choice under uncertainty. The principal rule in any circumstance is that a person facing a decision should make the choice that is expectedly optimal with respect to his preferences.

Although decision theory is a normative theory as it prescribes the preferred behaviour of a decision maker, in part it also has a descriptive character (Kyburg, 1991).

One would expect the goal of a normative theory of decision to provide rules that would unconditionally optimise one's preferences. This is however not the goal of decision theory because we take it for granted that no theory can do this: it is part of the human condition that we cannot predict the future perfectly, and therefore cannot choose the course of action that will in fact maximise our satisfactions. So, what we take as a normative theory, depends on what we take to be possible for human agents – a matter of descriptive rather than normative character.

The implications of this 'realistic' character of decision theory are more than purely philosophical. The complexity of most practical decision problems stems from the circumstance that there is a large number of decisions involved, and each potential policy for making these decisions induces a multitude of possible scenarios. As we cannot hope to fulfil our preferences completely, we are ultimately forced to evaluate all policies by inspecting each possible scenario. In other words, practical application of decision theory is hampered by the facts that decisions can often not be made in isolation but instead have to be evaluated in the context of other decisions, and that all possible consequences of decisions have to be taken into account in these evaluations.

To illustrate the highly combinatorial nature of decision problems, consider a problem that involves k subsequent decisions.[2] If z is an upper bound on the number of different observations that may be obtained before each decision, and n is the maximum number of alternatives for each decision, then there exist $\mathcal{O}(n^{z^k})$ different decision functions for the kth decision. In a decision-tree analysis of k subsequent decisions, the tree will have $\mathcal{O}(n^k z^k)$ nodes. Evaluation of all decision-making policies, or equivalently, solving the decision tree, is therefore hopelessly intractable when k is large. It should be noted though, that the complexity of the problem does not alter the nature of the tradeoff between decision alternatives on utility-theoretic grounds: there is no fundamental difference between a simple choice between decision alternatives in a single-decision problem and a choice between policies in a problem with multiple decisions.

The decision-theoretic analyses in this chapter have been illustrated with decision trees. Decision trees are frequently used in the field of clinical decision analysis (Weinstein and Fineberg, 1980; Pauker and Kassirer, 1987) as they provide an intuitive representation of decision problems and can easily be constructed in cooperation with experienced clinicians. They are however not suited as a knowledge-representation formalism for automated reasoning systems. The reasons for this are threefold. First, decision trees describe decision problems from the viewpoint of the decision maker while leaving most of the underlying knowledge of the problem domain implicit. Second, as we discussed above, the size of a decision tree grows exponentially in the size of the problem because a tree explicitly enumerates all possible decision-making policies. Third, they can only be used to solve a single problem case, whereas

[2] We assume that there are no concurrent actions. This is sometimes referred to as the *single decision maker assumption*.

a knowledge-based system would typically cover a range of problem cases within a given domain.

While decision trees are impractical as a knowledge-representation formalism for decision-theoretic reasoning systems, there exists several other representation formalisms that adhere to the decision-theoretic perspective, and may be used as a basis for intelligent reasoning. These formalisms are discussed in the next chapter.

CHAPTER 4

Decision-theoretic representation formalisms

It was explained in the previous chapter that decision-theoretic reasoning is characterised by the fact that each situation of choice is ultimately reduced to a utility-theoretic comparison. Yet, these choice situations may be very different from a conceptual point of view, and involve rather incomparable concepts pertaining to the role of actions, information and time. In utility theory, these concepts are left implicit; the task of decision-theoretic representations is to allow for explicit reasoning with these concepts.

In this chapter, we review the main formalisms for decision-theoretic representation and reasoning in intelligent systems. Notwithstanding their common decision-theoretic ground, these formalisms stem from separate fields, and therefore come in different forms. The first formalism to be discussed, in Section 4.1, is the *influence diagram*, a graphical representation of decision problems under uncertainty. Influence diagrams provide a concise way of representing these problems by exploiting independencies between the variables involved. In their basic form however, influence diagrams are static in nature and are therefore primarily suited to evaluate and compare a pre-specified set of decision-making scenarios. When part of the problem is to determine the number and timing of decisions, the expressiveness of influence diagrams falls short.

Section 4.2 deals with decision-theoretic reasoning where problem specifications call for the generation of decision-making scenarios; this is a form of decision-theoretic planning. We will review the formalism that is now most widely adopted for this type of reasoning, the *Markov decision process*. The expressiveness of this formalism

comes at the price of a combinatorial explosion in larger problem domains, rendering it then impractical. An integration of influence diagrams and Markov decision processes exists in the form of *dynamic influence diagrams*; these are discussed in Section 4.3. The chapter is concluded in Section 4.4 by re-evaluating and comparing the representation formalisms described. As before, examples will be taken from the domain of paediatric cardiology.

4.1 Graphical representations

One of the approaches to handle the combinatorial nature of decision making under uncertainty is exploiting independencies between the variables involved. This approach is advocated by formalisms based on *graphical representations*. Graphical representation formalisms explicitly separate qualitative and quantitative information on the problem domain, where the qualitative information comprises a collection of independency statements expressed in a graph. The graphical representation of decision problems under uncertainty that is most popular today is the *influence diagram*, which is the subject of this section. As a preliminary, we will first discuss *belief networks* in Subsection 4.1.1, a formalism for representing and reasoning with multivariate probability distributions. The belief-network formalism itself does not cover decision making under uncertainty, but can be readily extended to deal with decision-theoretic concepts. Influence diagrams provide such an extension, and as such offer a compact way of representing of decision problems; they are discussed in Subsection 4.1.2.

4.1.1 Belief networks

A *(Bayesian) belief network* (Pearl, 1988) is a concise, graphical representation of a joint probability distribution on a set of random variables. In a belief network, conciseness of representation is arrived at by explicit separation of information about the probabilistic independencies holding among the variables in the distribution and the numerical quantities involved. Below, we first discuss the representation of probabilistic independence in directed graphs; we then proceed to give a formal definition of belief networks.

The graphical representation of conditional independence

Conditional independence relations can be modelled in directed and undirected graphs. Each vertex in the graph then represents one of the variables involved, and the presence of an arc (or edge, in the undirected case) between two vertices indicates that there exists a direct, influential relationship between the corresponding variables. Absence of an arc (edge) generally means that there exists some form of conditional independence. Here, we restrict ourselves to the representation of

conditional independence relations in directed graphs, formalised by the notion of *d-separation*, (Pearl, 1988; Geiger et al., 1990). This type of representation of conditional independence is used in belief networks; a prerequisite is that the graph be acyclic.

We assume the reader to be familiar with the concept of directed graph; for a useful compendium of many definitions and results from graph theory, we refer to the book by Harary (1969).

Notation 4.1 (Directed graph) *We use $G = (V(G), A(G))$ to denote a directed graph, where $V(G)$ denotes the vertices in G and $A(G)$ denotes its arcs. When $v_1 \to v_2 \in A(G)$, $v_1, v_2 \in V(G)$, we say that vertex v_1 is a* predecessor *of v_2, and conversely that v_2 is a successor of v_1. We write $\rho_G(v_2)$ for the set of all predecessors of vertex v_2.*

- *A* (directed) path *in graph G is a sequence $v_1, \ldots, v_n$ of vertices where for each $i = 1, \ldots, n-1$ we have $v_i \to v_{i+1} \in A(G)$.*
- *A* cycle *in graph G is a path $v_1, \ldots, v_n$ in G where $v_0 = v_n$.*
- *A* chain *in graph G is a sequence $v_1, \ldots, v_n$ of vertices where for each $i = 1, \ldots, n-1$ we have either $v_i \to v_{i+1} \in A(G)$ or $v_{i+1} \to v_i \in A(G)$.*
- *A* loop *in graph G is a chain $v_1, \ldots, v_n$ in G where $v_0 = v_n$.*

If $v_1, \ldots, v_n$ is a path in graph G, then each v_i, $i = 1, \ldots, n$ is called a descendant *of vertex v_1. When the graph G does not contain any cycles, it is called* acyclic*.*

Chains and loops in a directed graph G can also be viewed as paths and cycles, respectively, in the underlying, undirected graph of G.

Definition 4.2 (Chain blocking) *Let $G = (V(G), A(G))$ be an acyclic, directed graph, and let $v_1, \ldots, v_n$ be a chain in G. We say that this chain is* blocked *by a set of vertices $W \subseteq V(G)$ if it contains three consecutive vertices v_{i-1}, v_i, v_{i+1}, $1 < i < n$, for which one of the following conditions holds:*

- *arcs $v_{i-1} \leftarrow v_i$ and $v_i \to v_{i+1}$ are on the chain s, and $v_i \in W$;*
- *arcs $v_{i-1} \to v_i$ and $v_i \to v_{i+1}$ are on the chain s, and $v_i \in W$;*
- *arcs $v_{i-1} \to v_i$ and $v_i \leftarrow v_{i+1}$ are on the chain s, and neither v_i or any of its descendants is comprised in W.*

We use the above definitions to define the *d-separation criterion*, (Pearl, 1988).

Definition 4.3 (d-Separation) *Let $G = (V(G), A(G))$ be an acyclic, directed graph, and let $U, V, W \subseteq V(G)$ be (disjoint) sets of vertices in G. Then, W is said to* d-separate *U from V, notation $\langle U \mid W \mid V \rangle_G^d$, if every chain in G from a vertex in U to a vertex in V is blocked by W.*

The d-separation criterion provides for reading independency statements from a directed graph where the vertices represent random variables.

Definition 4.4 (I-map) *Let $G = (V(G), A(G))$ be an acyclic, directed graph where the vertices in G represent random variables, and let P be a joint probability distribution on $V(G)$. The graph G is called an* independency map*, or* I-map*, for P if for all $U, V, W \subseteq V(G)$ holds:*

$$\textit{if} \quad \langle U \mid W \mid V \rangle_G^d \quad \textit{then} \quad U \perp\!\!\!\perp_P V \mid W. \tag{4.1}$$

In an I-map for distribution P, any probabilistic dependency between random variables is covered by the arcs of the graph. The converse is however not true: there may be probabilistic *in*dependencies that are *not* covered by the absence of arcs: any fully-connected, acyclic graph is an I-map for the variables discerned. We therefore strive to use *minimal* I-maps, i.e. graphs where no arc could be omitted without violating condition in Equation 4.1. In the extreme case, we have that all independencies are covered by the graph topology, i.e. $U \perp\!\!\!\perp_P V \mid W$ implies $\langle U \mid W \mid V \rangle_G^d$; the graph G is then said to be a *perfect map* for P. Unfortunately, there does not exist a perfect map for every probability distribution, (Pearl et al., 1990).

The belief-network formalism

As mentioned earlier, a belief network is a concise representation of a joint probability distribution P on a set of random variables; it consists of a qualitative and a quantitative part. The qualitative part of the network is an I-map for P; the quantitative part of the network comprises the specification of a local conditional probability distribution for each vertex in the graph. Recall that $\rho_G(v)$ denotes the set of all predecessors of vertex v in graph G.

Definition 4.5 (Belief network) *Let X be a finite set of random variables. A* belief network *over X is a pair $B = (G, \Gamma)$ where*

- *$G = (V(G), A(G))$ is an acyclic, directed graph,*
- *$V(G) = X$, and*
- *$\Gamma = \{\gamma_x : \Omega_x \times \Omega_{\rho_G(x)} \to [0, 1] \mid x \in X\}$ is a set of* (conditional) probability assessment functions*, such that for each variable $x \in X$ and each configuration $\mathrm{c}_{\rho_G(x)}$ of its predecessors $\rho_G(x)$ it holds that*

$$\sum_{\mathrm{c}_x \in \Omega_x} \gamma_x(\mathrm{c}_x \mid \mathrm{c}_{\rho_G(x)}) \quad = \quad 1. \tag{4.2}$$

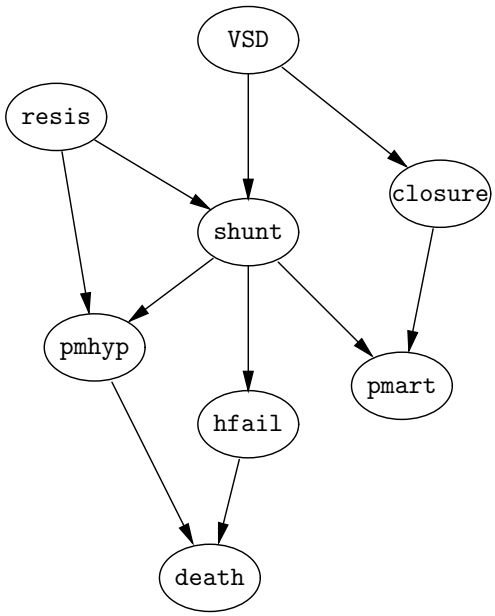


Figure 4.1: An example belief network for the VSD domain.

Example 4.6 *Figure 4.1 shows the graphical part of an example belief network for the VSD domain. The network comprises the 8 variables from Table 3.1, representing (from top to bottom) size of the VSD, pulmonary vascular resistance, spontaneous closure of the VSD, shunting, pulmonary hypertension, heart failure, pulmonary arteriopathy, and death. Note that the variables* `closure` *and* `pmart` *pertain to potential future developments. The conditional probabilities that constitute the quantitative part of the belief network are listed in Appendix B.*

The following theorem now states that the probability assessment functions taken together uniquely define a joint probability distribution on the variables discerned, that respects the independency relation portrayed by the graphical part of the network. Henceforth, we will call this distribution *the* distribution that is defined by the belief network.

Theorem 4.7 *Let X be a set of random variables, let $B = (G, \Gamma)$ be a belief network over X, and let P be a joint probability distribution on X that is defined as follows:*

$$P(\mathrm{C}_X) \;=\; \prod_{x \in X} \gamma_x(\mathrm{C}_x \,|\, \mathrm{C}_{\rho_G(x)}). \tag{4.3}$$

Then G is an I-map for P.

Proof. The acyclicity of G allows for a total topological ordering of the vertices in $V(G)$; in this ordering, each variable is conditioned only on variables preceding it when we apply the chain rule of probability. This enables to exploit the conditional

independency relations portrayed by G. By taking $P(\mathrm{c}_x \,|\, \mathrm{c}_{\rho_G(x)}) = \gamma_x(\mathrm{c}_x \,|\, \mathrm{c}_{\rho_G(x)})$ for each $x \in X$ and all configurations $\mathrm{c}_x \in \Omega_x$ and $\mathrm{c}_{\rho_G(x)} \in \Omega_{\rho_G(x)}$, the property stated in the theorem follows immediately. For further details, we refer to (Kiiveri et al., 1984). □

Since a belief network $B = (G, \Gamma)$ defines a unique joint probability distribution on the variables involved, it allows for computing any marginal or conditional probability regarding these variables. As such, the belief-network formalism offers a powerful framework for reasoning with uncertainty in a mathematically correct manner. Moreover, the graphical representation of conditional independence offers an intuitively appealing modelling language. Belief networks have therefore enjoyed a rapidly increasing popularity in AI since their introduction in the 1980s.

Several algorithms have been developed to infer probabilities from a given belief network, the most popular being the *clique-tree propagation algorithm*, (Lauritzen and Spiegelhalter, 1988; Jensen et al., 1990). In this algorithm, the belief network is 'compiled' to a computational architecture called a *junction tree* (or *clique tree*) prior to the actual inference. The worst-case computational complexity of the algorithm is exponential in the size of the graph. This is not surprising, as the problem of general probabilistic inference has been shown to be NP-hard, (Cooper, 1990). Modern computers are however capable of inferring a probability of interest within seconds for networks up to hundreds of variables, depending on the connectivity of the graph.

Problem solving with belief networks

Reasoning with a belief network generally proceeds as follows. Initially, no observations are assumed, and the marginal probability distribution of each network variable is computed. Based on these distributions, an *information-gathering procedure* now selects variables for which an observation should be obtained, or a *decision-making procedure* selects variables that are set to specific values by external intervention. After the user has obtained and entered the observations or decisions, these variables are instantiated ('freezed') to their values, and the posterior distributions are computed for the remaining variables in the network. This process is repeated until a pre-defined stopping criterion is met.

From the above description, it becomes apparent that the belief-network framework is limited to the representation of random variables and their uncertain relations and does not cover the problem-solving knowledge involved. Application of belief networks therefore requires extension of the basic framework with representations of the envisioned problem-solving task. We distinguish two types of extension that have been proposed in the literature. The first type of extension adds problem-solving knowledge at the graphical level of representation within in a belief network. An exemplary instance of this type of extension is the *influence diagram* representation, that is to be discussed in the subsection to follow. The second type of extension embeds the belief network in a larger problem-solving architecture, where a separate

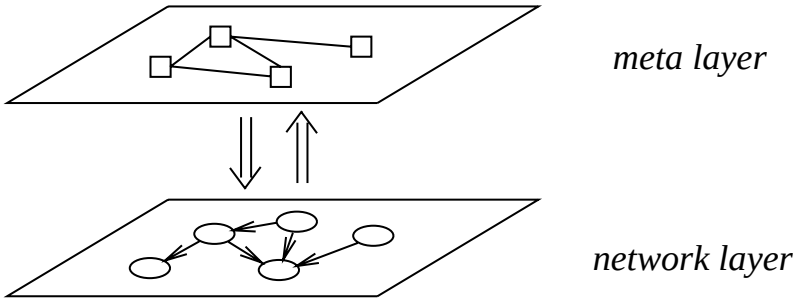


Figure 4.2: A 2-layered architecture for belief networks.

meta layer is designed to provide for control over problem-solving; see Figure 4.2. This type of architecture is implicitly present in many belief-network applications, but was first explicitly described by Van der Gaag and Wessels (1993). In their proposal, the control layer implements an information-gathering procedure for diagnostic reasoning. For further details on this architecture, we refer to the article in question.

4.1.2 Influence diagrams

An *influence diagram* (Howard and Matheson, 1981; Shachter, 1986) is a graphical representation of a decision problem under uncertainty. The representation is closely related to that of belief networks: an explicit separation is made between qualitative and quantitative information on the problem, where qualitative information is captured by an acyclic directed graph, and quantitative information is laid down in assessment functions associated with vertices in that graph. The main difference is the addition of one or more *decision nodes* and a single *value node* to the graph, where decision nodes represent decision variables and the value node accommodates a utility function. Influence diagrams are also closely related to decision trees, but they provide a far more concise representation of decision problems. We follow the type of formalisation employed in (Ndilikilikesha, 1994) and (Zhang, 1998).

Definition 4.8 (Influence diagram) *Let X be a finite set of random variables. An* influence diagram *over X is a triple $ID = (G, \Gamma, u)$, where*

- *$G = (V(G), A(G))$ is an acyclic, directed graph,*
- *$V(G) = X \cup D \cup \{v\}$, where $D = \{d_1, \ldots, d_k\}$ is a set of decision variables,*
- *the value node v has no successors in G,*
- *$\Gamma = \{\gamma_x \mid x \in X\}$ is a set of conditional probability assessment functions for the variables from X,*

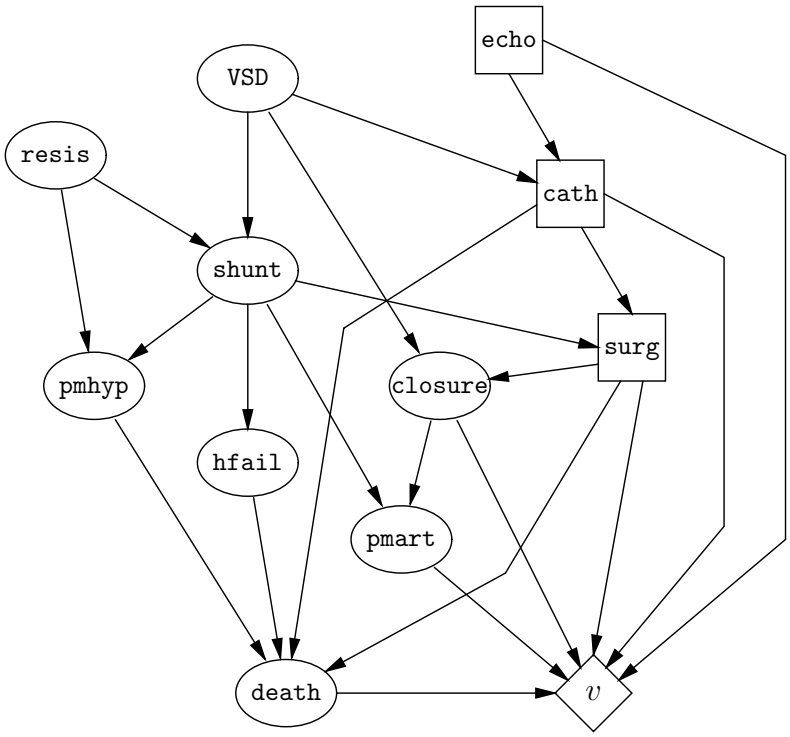


Figure 4.3: An example influence diagram for the VSD domain.

- u *is a function* $u : \Omega_{\rho_G(v)} \rightarrow \mathbb{R}$.

In an influence diagram $ID = (G, \Gamma, u)$, the interpretation of an arc in $A(G)$ depends on the node in the graph that it points to. The direct predecessors $\rho_G(x)$ of a random variable $x \in X$ represent direct influences (conditioning variables) on that variable. These arcs jointly describe conditional independence relations, as in a belief network; the definition of the set Γ of conditional probability assessment functions is the same as for belief networks. The direct predecessors $\rho_G(d)$ of a decision variable $d \in D$ are taken to represent the available information when deciding upon decision d; this type of arc therefore models *informational constraints* on the various decisions. The direct predecessors $\rho_G(v)$ of the value node v again represent direct influences, but now on the decision maker's preferences (utility).

Example 4.9 *Figure 4.3 shows an example influence diagram for the VSD domain. The diagram extends the belief network from Figure 4.1 with 3 decision nodes,* `echo` *(echocardiography),* `cath` *(cardiac catheterisation), and* `surg` *(cardiac surgery), and a value node. Each of the decision nodes can take the values 'yes' and 'no'. Echocardiography reveals the size of the VSD, catheterisation reveals the shunt, and surgery may be used to close the defect. Direct influences of these decisions on utility (as for* `echo` *and* `surg`*) relate to definite increases or decreases in utility caused by the decision (e.g., costs), whereas indirect influences on utility (through the variables*

`closure`, `pmart`, *and* `death`*) relate to contingent changes in utility (e.g. risks). The quantitative part of the influence diagram is provided in Appendix B.*

The informational constraints expressed by arcs leading to a decision node $d \in D$ have the following interpretation: it is assumed that at the time of choosing an alternative for variable d, a value is known for each of its predecessors in the graph. These values are either known by observation (for random variables) or because they were earlier chosen by the decision maker himself (for decision variables). A configuration $\mathrm{c}_{\rho_G(d)}$ of the predecessors of decision node $d \in D$ is therefore interpreted as describing an information history for decision d. We will assume a linear order on the decision variables involved in the decision problem, which determines the temporal order in which they are given a value. This is enforced by the *regularity constraint*.

Definition 4.10 (Regularity) *Let $ID = (G, \Gamma, u)$ be an influence diagram. We say that the diagram ID is* regular *when there exists a (directed) path $d_1, \ldots, d_k$ in G comprising all decision variables.*

Another common assumption is that when making decision d_i, $1 < i \leq k$, all the information that was available when making decisions $d_1, \ldots, d_{i-1}$ is taken into account. That is, node d inherits the predecessors of nodes $d_1, \ldots, d_{i-1}$ in the graph; this assumption is referred to as *no forgetting*.

Definition 4.11 (No forgetting) *Let $ID = (G, \Gamma, u)$ be an influence diagram. It is said that the* no-forgetting property *holds in diagram ID when for each pair of distinct decision nodes $d_i, d_j \in D$ we have that $\rho_G(d_j) \subseteq \rho_G(d_i)$ whenever $d_j \in \rho_G(d_i)$.*

Example 4.12 *The influence diagram of Figure 4.3 is regular but does not have the no-forgetting property. For this property to hold, the arcs* `echo` $\rightarrow$ `surg` *and* `VSD` $\rightarrow$ `surg` *should be added to the diagram.*

The direct predecessors $\rho_G(v)$ of the value node v designate which variables influence the decision maker's preferences: the value of the variable v is determined by applying the function u to v's predecessors $\rho_G(v)$ in the graph. Given a configuration $\mathrm{c}_{\rho_G(v)}$ of this set, the value of v is invariant under the possible values for all other variables. For convenience, however, we will write $u(\mathrm{c})$ to denote the utility value associated with a given configuration $\mathrm{c} \in \Omega_{X \cup D}$ of all variables in the diagram.

An influence diagram represents a decision problem, for which a solution takes the form of a set of *decision functions*, jointly called a *policy*.

Definition 4.13 (Decision function and policy) *Let $ID = (G, \Gamma, u)$ be an influence diagram. A* decision function *for decision variable $d \in D$ is a mapping*

$$\delta_d : \Omega_{\rho_G(d)} \rightarrow \Omega_d. \tag{4.4}$$

A policy *for ID is a set* $\pi = \{\delta_d \mid d \in D\}$ *of decision functions for all decision nodes in ID.*

Given a policy π, a probability distribution P_π over $X \cup D$ is induced as follows:

$$P_\pi(\mathrm{C}_{X\cup D}) \;=\; \prod_{x\in X} \gamma_x(\mathrm{C}_x \mid \mathrm{C}_{\rho_G(x)}) \cdot \prod_{d\in D} \hat{\pi}(\mathrm{C}_d \mid \mathrm{C}_{\rho_G(d)}), \tag{4.5}$$

where

$$\hat{\pi}(\mathrm{c}_d \mid \mathrm{c}_{\rho_G(d)}) \;=\; \begin{cases} 1, & \text{if } \delta_d \in \pi,\ \delta_d(\mathrm{c}_{\rho_G(d)}) = \mathrm{c}_d, \\ 0, & \text{otherwise.} \end{cases} \tag{4.6}$$

Configuration c_Y of $Y \subseteq X \cup D$ is said to be *consistent* with policy π if $P_\pi(\mathrm{c}_Y) > 0$. The *expected utility* of policy π given such a configuration c_Y equals

$$\tilde{u}(\pi \mid \mathrm{c}_Y) \;=\; \sum_{\mathrm{c}\in\Omega_{X\cup D}} u(\mathrm{c}) \cdot P_\pi(\mathrm{c} \mid \mathrm{c}_Y). \tag{4.7}$$

Expected utility is undefined for inconsistent configurations. *Evaluating* or *solving* an influence diagram is the problem of finding a policy π^* that maximises (unconditional) expected utility, i.e. $\tilde{u}(\pi^*) \geq \tilde{u}(\pi)$ for any policy π. The policy π^* and the decision functions it comprises are called *optimal*.

Example 4.14 *For the influence diagram of Figure 4.3, the unique optimal policy* π^* *consists of the following decision functions:*

$$\begin{aligned} \delta_{\texttt{echo}}(\top) &= \mathit{yes}, \\ \delta_{\texttt{cath}}(\texttt{echo} = \mathit{yes}, \texttt{VSD} = s) &= \mathit{no}, \end{aligned}$$

for all possible values $s \in \mathrm{dom}(\texttt{VSD})$*, and*

$$\delta_{\texttt{surg}}(\texttt{echo} = \mathit{yes}, \texttt{cath} = \mathit{no}, \texttt{VSD} = s) \;=\; \begin{cases} \mathit{yes}, & \text{if } s = \mathit{large}, \\ \mathit{no}, & \text{otherwise.} \end{cases}$$

In words, it is advised to conduct echocardiography, but not cardiac catheterisation, regardless of the VSD size that has been observed. Surgery is recommended with large VSDs.

Evaluating influence diagrams

A variety of methods exists to evaluate a given influence diagram; each of them basically consists of performing a sequence of alternating steps of probabilistic inference and utility maximisation. We can distinguish methods that perform the evaluation *in situ*, i.e. operating directly on the diagram given, and those that perform the evaluation after transforming the diagram into an alternative representation. The

first and still most well-known *in situ* method was described by Shachter (1986). His algorithm consists of a series of transformations on the graph that result in removal of decision nodes and chance nodes; the optimal policy is computed as a side-effect. The main drawback of the algorithm is that sometimes arcs in the graph have to reversed before a node can be removed. Arc reversal is a computationally costly operation that corresponds to application of Bayes' rule. When the influence diagram comprises nodes with a large number of direct predecessors, the reversal of arcs may result in lengthy computation times.

Another approach to solving influence diagrams is based on transforming the diagram into a belief network. This approach was first proposed by Cooper (1988) and later refined by Shachter and Peot (1992) and Zhang (1998). The transformation of influence diagram $ID = (G, \Gamma, u)$ proceeds as follows. The topology of the graph remains unaltered but now all the nodes in $V(G)$ are taken to represent random variables. Decision variables are initialised with arbitrary conditional probability assessment functions, and later adjusted to represent optimal decision functions. The value node v is provided with a binary universe $\Omega_v = \{\mathrm{c}_v^-, \mathrm{c}_v^+\}$, and

$$\gamma_v(\mathrm{c}_v^+ \mid \mathrm{c}_{\rho_G(v)}) \;=\; f(u(\mathrm{c}_{\rho_G(v)})) \tag{4.8}$$

where $f : \mathbb{R} \to [0, 1]$ is a positive linear function. Maximising the probability of c_v^+ in the belief network now corresponds to utility maximisation in the original influence diagram; the optimal policy can therefore be found using probabilistic inference on the resulting belief network. A related method transforms the influence diagram into a junction tree that is suited for probabilistic inference using the clique-tree propagation algorithm, (Shenoy, 1992; Jensen et al., 1994). Each of the solution methods mentioned here solves the diagram by considering the decision variables in reverse order: the decision function for last decision $d_k \in D$ is constructed first, and the decision function for the first decision $d_1 \in D$ is constructed last. This type of solution strategy is more generally referred to as *backward induction*. Backward induction usually performs well on problems of limited size, but becomes intractable in larger domains.

Remarks on the influence-diagram representation

An influence diagram of size $m = |V(G)|$ corresponds to a symmetrical decision tree of depth m, (Howard and Matheson, 1981; Olmsted, 1983). The correspondence is obtained by mapping the variable at rank i ($1 \leq i \leq m$) in a topological sort of the variables in the diagram to the nodes at depth i in the tree; the arcs emanating from a node in the tree correspond to its possible values. It is easily seen though, that the influence-diagram representation is characterised by far more conciseness than the representation in decision trees. Similar to belief networks, influence diagrams explicitly separate qualitative and quantitative information: qualitative information is expressed in the graph, and quantitative information is laid down in probability assessment and utility functions. As a result, the size of the graph is directly proportional to the number of variables involved in the problem.

A drawback to the influence-diagram representation is that it has problems with the handling of asymmetry in decision problems. Recall from Section 3.3 that informational asymmetry stems from test decisions and pertains to structural differences in available information for future decisions; relevantial asymmetry occurs in problems where decisions may be repeated and pertains to the irrelevance of considering particular alternatives twice. Informational asymmetry is hard to model in an influence diagram because the structure of decision functions is fixed by the graph's topology; the decision function for decision $d \in D$ always takes in inputs from $\Omega_{\rho_G(d)}$, whereas in reality the availability of observations on the set $\rho_G(d)$ may depend on previous decisions. Relevantial asymmetry is hard to model because in all situations the same set Ω_d of alternatives is considered relevant for decision variable d, and one cannot explicitly ignore alternatives in specific situations.

Example 4.15 *In the influence diagram of Figure 4.3, it is assumed that the variable* `VSD` *is known when the second decision (catheterisation) is made. Similarly, the variable* `shunt` *is assumed to be observed prior to the third decision (surgery). However, these observations are only available when the previous decisions were affirmative. In the optimal policy described in Example 4.14, this has already been taken into account.*

We will now first discuss how both problems can be dealt with without modifications to the basic representation, and then briefly review some solutions proposed in the literature that extend the influence-diagram representation.

To start with informational asymmetry, suppose that the set Ω_{d_i} of alternatives for decision variable $d_i \in D$ encompasses an action c_{d_i} that uniquely yields an observation for random variable $x \in X$. Ideally, this configuration of x is only included in situations for a subsequent decision d_j, $j > i$, when c_{d_i} was chosen for d_i. But in the influence-diagram representation, x is made a parent of decision d_j, and always needs to have a value when deciding upon d_j. So the informational asymmetry that results from observations cannot be represented graphically. Still, we can solve the problem by prohibiting that decision d_j takes advantage of information on variable x when c_{d_i} was not selected. That is, if $\rho_G(d_j) = \{d_i, x\} \cup Z$ is the set of predecessors of decision variable d_j, we require that for all configurations $\mathrm{c}'_{d_i} \in \Omega_{d_i}$ other than c_{d_i} that

$$\delta_{d_j}(\mathrm{c}'_{d_i} \wedge \mathrm{c}_x \wedge \mathrm{C}_Z) \;=\; \delta_{d_j}(\mathrm{c}'_{d_i} \wedge \mathrm{c}'_x \wedge \mathrm{C}_Z) \tag{4.9}$$

for all configurations $\mathrm{c}_x, \mathrm{c}'_x \in \Omega_x$. So the decision maker is effectively forced to disregard the value of x in deciding upon d_j when configuration c_{d_i} was not chosen for the earlier decision d_i.

We now turn to relevantial asymmetry. Suppose that we want to ignore the alternative c_d for decision variable d in situation $\mathrm{c}_{\rho_G(d)}$ because we consider c_d irrelevant in that situation. We can enforce this by choosing an appropriate utility function u. Let c'_d be an arbitrary alternative for decision d other than c_d. We take

$$u(\mathrm{c}_{\rho_G(d)} \wedge \mathrm{c}_d \wedge \mathrm{c}_Y) \;<\; u(\mathrm{c}_{\rho_G(d)} \wedge \mathrm{c}'_d \wedge \mathrm{c}_Y), \tag{4.10}$$

for each configuration c_Y of the set $Y = (X \cup D) \setminus (\{d\} \cup \rho_G(d))$ of remaining nodes in the diagram. It then follows that c_d cannot be used to maximise utility in situation $c_{\rho_G(v)}$, and will therefore be ignored. Note that this solution requires that there are arcs in the graph from node d and $\rho_G(d)$ to the value node v; it therefore induces an increase in computational complexity if one of the above solution methods is employed.

Both solutions described above have the disadvantage of encoding qualitative relationships of asymmetry at the quantitative (numerical) level of the representation. As such, these relationships remain implicit in the representation, and can only be uncovered from a given influence diagram by laborious examination of the specified quantities. This situation has inspired several authors to propose extensions to influence diagrams that allow for explicit representation of asymmetry. Jenzarli (1995) provides a treatment of relevantial asymmetry by employing *relevance arcs* to constrain the set of possible values of a decision node; the extended representations are called *information/relevance diagrams*. Smith et al. (1993) provide a treatment of both types of asymmetry in influence diagrams by adding a tree-based graphical language to specify value domains, probability assessment functions, and utility functions. Both extensions require specialised methods to compute the optimal policy.

4.2 Decision-theoretic planning

In this section, we focus on decision-theoretic representation formalisms that allow for generation of decision-making scenarios. Such representation formalisms are preferable when part of the decision problem is to assess the number and timing of actions. In these formalisms, the primitives consist of a system that evolves over time, and a set of actions from which the decision maker may choose to influence the system dynamics and to make observations. Generally speaking, the objective is to achieve a long-term goal (in terms of preferred system state), taking into account particular short-term restrictions (e.g., the costs of performing actions). As was described in Chapter 1, in AI this type of reasoning is referred to as *planning*, (Fikes and Nilsson, 1971; Allen et al., 1990). The essential part of solving a planning problem is to consider the consequences of actions before being forced to experience them; the result of these considerations is a detailed formulation of a program of action, called a *plan*.

When competing plans are compared using decision-theoretic principles, one speaks of *decision-theoretic planning*, (Dean and Wellman, 1991). We will focus on *Markov decision processes* (MDPs), the currently predominant approach to formalising decision-theoretic planning, (Boutilier et al., 1999). MDPs are models for sequential decision making based on random Markov processes; the underlying conceptualisation can be described as follows. At a specified point in time, a decision maker observes the state of a dynamic system. Based on this observation, he chooses an action. The action

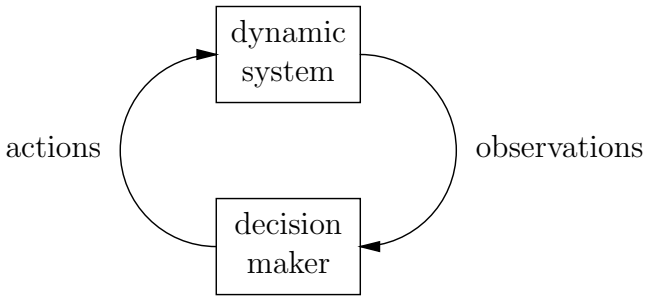


Figure 4.4: The decision process depicted.

choice produces two results: the decision maker receives an immediate reward, and the system evolves to a new state at a subsequent point in time according to an effect determined by the action choice. At this subsequent point in time, the decision maker faces a similar problem, but now the system may be in a different state. This conceptualisation is depicted in Figure 4.4. The decision maker's objective is to develop a decision-making policy that maximises the expected total reward over a predefined period of time.

The concept of decision making as described above is very general and thus covers a wide variety of decision-theoretic planning problems. Consequently, MDPs come with various characteristics, but a most important distinction is between *full* and *partial observability* of the system state. Fully-observable Markov decision processes (FOMDPs) are discussed in Section 4.2.1; partially-observable Markov decision processes (POMDPs) are discussed in Section 4.2.2.

4.2.1 Fully-observable Markov decision processes

Research on MDP theory and algorithms was initiated in the 1950s and 1960s by R. Bellman and R. Howard. Here, we follow the model formulation of Blackwell (1965) with slight adjustments, and analyse some of the central properties of fully-observable MDPs. For an in-depth analysis of FOMDPs, we refer the reader to the monograph by Puterman (1994).

Definition 4.16 (MDP model) *Let X be a finite set of random variables. A* Markov decision process *over X is described by a 4-tuple $\mathcal{M} = (T, A, \Theta, R)$, where*

- *T is a set of linearly ordered* decision moments*,*
- *A is a set of available* actions*,*
- *$\Theta = \{\theta_t : \mathrm{dom}(X) \times A \times \mathrm{dom}(X) \to [0,1] \mid t \in T\}$ is a set of* transition probability functions*, and*
- *$R = \{r_t : \mathrm{dom}(X) \times A \to \mathbb{R} \mid t \in T\}$ is a set of* reward functions*.*

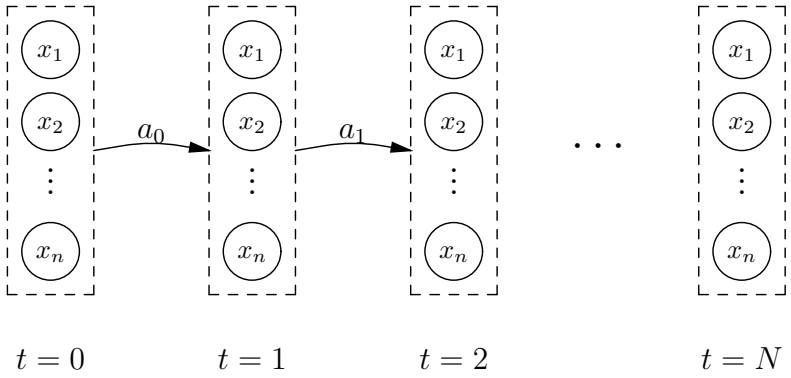


Figure 4.5: Process dynamics in MDPs.

In principle, few restrictions exist on the above elements of $\mathcal{M}$ in order to jointly qualify as an MDP. The primary characteristic of an MDP is the *Markov property*: the effects of actions are described by stochastic transitions on the system state that depend on the last state and action choice only; the sequence of subsequent system states within a given evolution of the decision process is therefore a Markov chain. We restrict the discussion of MDPs to cases where the set T is discrete, and the sets X and A are finite. As in the previous sections, by finiteness of X we mean that X is a finite set of discrete variables with a finite domain, and therefore the set $\mathrm{dom}(X)$ of system states is also finite. Below, we will first describe the process dynamics in MDPs. We then turn to criteria to evaluate and compare decision processes, and to the formulation of decision-making policies for FOMDPs. We conclude the section with a brief discussion of solution methods, and some remarks on the FOMDP representation.

Process dynamics

In an MDP model $\mathcal{M} = (T, A, \Theta, R)$, the set T explicitly denotes the times at which the decision maker is expected to choose an action; the explicitness in the representation of these decision moments contrasts with the earlier discussed decision-theoretic representation formalisms where the notion of time was left implicit. As we take T to be discrete, we can assume without loss of generality that $T = \{0, 1, 2, \ldots\} \subseteq \mathbb{N}$, where the ordering $<$ on the natural numbers represents temporal precedence; the time point $t = 0$ is called the *initial moment* of the decision process. When there exists a finite maximum element $N \in \mathbb{N}$ in T, the model is said to have a *finite horizon* of *length* N; otherwise, it is said to have an *infinite horizon*. In the present discussion, we will focus on finite-horizon models and make a few remarks on the generalisation to infinite horizons. Note that in the finite case, the action choice at the final decision moment $t = N$ is meaningless with respect to evolution of the system state.

The dynamic system under (partial) control by the decision maker is described by a set X of random variables, where each joint value $S \in \mathrm{dom}(X)$ represents a possible

state of that system.[1] Similar to the representation of time, state dynamics are made explicit in MDPs: all the variables in X obtain a new value at each decision moment (although it is possible that some values have not changed as compared to their previous values). An expression of the form $X = S$ states that the system occupies state $S \in \mathrm{dom}(X)$; we regard the elements of the set X as *attributes*, each describing a different aspect of the dynamic system. The number of possible system states equals $|\mathrm{dom}(X)| = |\Omega_X|$. The set A represents the actions, or equally, decision alternatives, that are available to the decision maker at each decision moment. Note that there are no restrictions on the action-selection procedure: actions may be chosen multiple times, and it is even possible to repeat a single action all the time. Figure 4.5 schematically depicts the described process dynamics.

The effects of actions on system dynamics are described by the set Θ of time- and action-dependent transition probability functions, where $\theta_t(S, a, S')$, $\theta_t \in \Theta$, denotes the probability that state $S' \in \mathrm{dom}(X)$ results after performing action $a \in A$ in state $S \in \mathrm{dom}(X)$ at decision moment $t \in T$. In infinite-horizon MDPs, action effects are usually assumed to be independent of time, i.e. $\theta_t = \theta_{t'}$ for all time points $t, t' \in T$; the transition probabilities are then said to be *stationary*. A special case exists when the action effects are *deterministic*, i.e. $\theta_t(S, a, S') \in \{0, 1\}$ for all $t \in T$, $a \in A$, and $S, S' \in \mathrm{dom}(X)$. Then, a given initial system state and sequence of action choices fixes the evolution of the system over time. Generally speaking, however, the effects are stochastic and a multitude of evolutions is possible.

We will now introduce some notations to guide the remaining discussion. Let $\tau \in T$ be a decision moment. A sequence

$$\sigma \quad = \quad S_0, \ldots, S_\tau \tag{4.11}$$

of subsequent system states (i.e. $S_t \in \mathrm{dom}(X)$, $t = 0, \ldots, \tau$) represents a potential evolution of the system and is called a *state sequence* up to time point τ. If $m = |\Omega_X|$ is the cardinality of the state space, there exist $m^{\tau+1}$ different state sequences up to that time point. A sequence

$$\alpha \quad = \quad a_0, \ldots, a_\tau \tag{4.12}$$

of subsequent action choices (i.e. $a_t \in A$, $t = 0, \ldots, \tau$) represents concrete decision-making behaviour and is called an *action sequence* up to time point τ. If there are $k = |A|$ different actions to choose from, there exist $k^{\tau+1}$ different action sequences up to that time point. The pair $h = (\sigma, \alpha)$ represents a potential realisation of the decision process and is called a *decision-making history*. We will use H_τ to denote the set of all possible histories up to time point τ. The set of full-length histories H_N now represents all potential *outcomes* of the decision process; the cardinality of H_N is $(mk)^{N+1}$.

Given an action sequence $\alpha = a_0, a_1, \ldots, a_\tau$ and an initial state $S \in \mathrm{dom}(X)$, a (conditional) probability distribution $P_\tau^{(S,\alpha)}$ on state sequences up to time point $\tau \in T$ is

[1] Recall from Section 3.1 that $\mathrm{dom}(X) = \mathrm{dom}(x_1) \times \cdots \times \mathrm{dom}(x_n)$ if $X = \{x_1, \ldots, x_n\}$.

induced as follows:

$$P_\tau^{(S,\alpha)}(\sigma) \;=\; \prod_{t=0}^{\tau-1} \theta_t(S_t, a_t, S_{t+1}) \tag{4.13}$$

for all state sequences $\sigma = S_0, \ldots, S_\tau$ having $S_0 = S$, and $P_\tau^{(S,\alpha)}(\sigma) = 0$ for all other state sequences. In this probability distribution, the system state at time point t is independent of the decision-making history given the action choice and system state at time point $t-1$; the sequence of subsequent system states is therefore a Markov chain. Furthermore, the action a_τ at time point τ does not appear in the equation and is uninfluential. In the overall decision problem, the action choice at the last decision moment $t = N$, is therefore irrelevant to the system's evolution.

Example 4.17 *An example MDP model for the VSD domain can be devised as follows. The set X represents the clinical state of the patient and is composed of the attributes* `VSD`*,* `resis`*,* `shunt`*,* `pmhyp`*,* `pmart`*,* `hfail`*, and* `death`*. There are 6 decision moments, at respectively 3 months, 6 months, 12 months, 24 months, 4 years, and 8 years after birth. The available actions, finally, are the modalities available to the cardiologist to manage a VSD patient:* $A = \{\texttt{echo}, \texttt{med}, \texttt{cath}, \texttt{surg}, \texttt{biop}\}$*. Note that in this MDP model, spontaneous closure of the VSD is represented implicitly by diminishing values for the attribute* `VSD` *at subsequent decision moments. Similarly, the Eisenmenger syndrome is represented by increasing values for* `pmhyp` *(pulmonary hypertension) due to pulmonary arteriopathy (*`pmart`*=true).*

Evaluation criteria

In an MDP model $\mathcal{M} = (T, A, P, R)$, the set R comprises reward functions r_t, $t \in T$, that describe time-dependent preferences of the decision maker with respect to states and actions: $r_t(S, a)$, $r_t \in R$, denotes the (numerical) reward received when the decision maker chooses action $a \in A$ at time point $t \in T$ and the current state is $S \in \mathrm{dom}(X)$. It is important to note that this reward value reflects relative (un)desirability of that state and action at time point t only; states and actions at other time points are disregarded within the reward functions. Furthermore, reward values may be positive as well as negative; in the latter case one often speaks of *costs*. Similar to transition probability functions, we speak of *stationary reward functions* when they are independent of time, i.e. when $r_t = r_{t'}$ for all $t, t' \in T$; this is customary for infinite-horizon models. We note that although the action choice at the final decision moment $t = N$ will not influence the system's evolution, it does affect the reward received at that time point.

To rank the potential outcomes $h \in H_N$ of a decision process, the rewards received at subsequent time points have to be combined using an *evaluation metric*. Examples of such metrics are *total reward*, *average reward*, and variations thereof. We focus

here on the *total discounted reward* metric, which is defined as

$$u(h) \;=\; \sum_{t=0}^{N} \lambda^t r_t(S_t, a_t), \tag{4.14}$$

where $h = (S_0, \ldots, S_N, a_0, \ldots, a_N)$, and $0 < \lambda \leq 1$ is a real-valued *discount factor*. The value $u(h)$ is the total discounted reward associated with history h, and is also referred to as its *utility* under this metric; when $\lambda = 1$, we simply speak of *total reward*. From a utility-theoretic point of view, the function u provides a preference ordering on the set H_N of outcomes, where the states and actions at subsequent time points are taken to be additive-independent attributes of utility. The discount factor is generally justified economically (as a representation of interest, when the rewards represent monetary gains), mechanically (as a representation of physical decay), or psychologically (people tend to care more about near than about distant future). The discount factor is also a prerequisite to infinite-horizon MDPs, as there would otherwise be no upper bound on the function u.

Given an action sequence $\alpha = a_0, \ldots, a_N$, the expected utility $\tilde{u}_\alpha(S)$ of initial system state S is now defined as

$$\tilde{u}_\alpha(S) \;=\; \sum_{h \in H_N,\, h=(\sigma,\alpha)} u(h) \cdot P_N^{(S,\alpha)}(\sigma). \tag{4.15}$$

The decision maker's objective is to maximise expected utility by choosing an appropriate sequence of actions.

Formulation of solutions

As the evolution of the system cannot be predicted with certainty, the decision maker will have to respond to observations in due course when choosing his actions; otherwise, expected-utility maximisation is not guaranteed. Under the assumption of full observability, solving the decision problem formulated by an MDP model therefore amounts to finding a *policy* $\pi = \{\delta_t \mid t \in T\}$, where

$$\delta_t : \mathrm{dom}(X) \to A \tag{4.16}$$

is a decision function prescribing the action choice at time point $t \in T$ given the actual system state. Potentially, there are k^m different decision functions to choose from at each time point, where again $k = |A|$ and $m = |\Omega_X|$. The number of different policies is therefore $k^{m^{(N+1)}}$. If the functions are identical for all time points, i.e. when $\delta_t = \delta_{t'}$ for all $t, t' \in T$, then the policy is said to be *stationary*; the number of different policies than reduces to k^m. Stationarity of the policy can be assumed when both the transition probability and the reward functions are stationary; in that case, investigating non-stationary policies would not help to increase expected utility. Such policies are therefore the standard type of solution to infinite-horizon FOMDPs.

Given a policy $\pi = \{\delta_t \mid t \in T\}$ and an initial system state $S \in \mathrm{dom}(X)$, a (conditional) probability distribution $P_\tau^{(S,\pi)}$ on histories up to time point τ is induced as follows:

$$P_\tau^{(S,\pi)}(h) = \prod_{t=0}^{\tau-1} \theta_t(S_t, a_t, S_{t+1}) \tag{4.17}$$

for all $h = (S_0, \ldots, S_\tau, a_0, \ldots, a_\tau)$ having $S_0 = S$ and $\delta_t(S_t) = a_t$; for all other $h \in H_\tau$, we have $P_\tau^{(S,\pi)}(h) = 0$. The expected utility $\tilde{u}_\pi(S)$ of initial system state S under policy π now equals

$$\begin{aligned} \tilde{u}_\pi(S) &= E_{P_N^{(S,\pi)}}(u) \\ &= \sum_{h \in H_N} u(h) \cdot P_N^{(S,\pi)}(h). \end{aligned} \tag{4.18}$$

The task of computing $\tilde{u}_\pi(S)$ for a given policy π and initial state S is called *policy evaluation*. We say that a policy is *optimal* when it maximises $\tilde{u}_\pi(S)$ for all $S \in \mathrm{dom}(X)$; the task of finding such a policy is called *solving* the FOMDP. Note that the utility function must have an upper bound in order to compare policies; this condition is satisfied when all the reward functions are bounded, and, in the case of infinite-horizon models, the discount rate λ is smaller than 1.

Solution methods

The standard approach to solving FOMDPs is based on decomposing the decision process using the Markov property. Define the *maximum expected partial utility* $\tilde{\nu}_t^*(S_t)$ of state $S_t \in \mathrm{dom}(X)$ at time point $t \in T$ as follows:

$$\tilde{\nu}_t^*(S_t) = \max_{a_t \in A} \left\{ r_t(S_t, a_t) + \lambda \cdot \sum_{S_{t+1} \in \mathrm{dom}(X)} \theta_t(S_t, a_t, S_{t+1}) \cdot \tilde{\nu}_{t+1}^*(S_{t+1}) \right\}, \tag{4.19}$$

if $t < N$, and $\tilde{\nu}_N^*(S_N) = \max_{a_N \in A}\{r_N(S_N, a_N)\}$, otherwise. The value $\tilde{\nu}_t^*(S_t)$ is the maximum expected total reward that is to be received during the future steps of the decision process. If π^* is an optimal policy, we have that

$$\tilde{u}_{\pi^*}(S) = \tilde{\nu}_0^*(S) \tag{4.20}$$

for all initial system states $S \in \mathrm{dom}(X)$. The recursion described by Equation 4.19 is usually named a *Bellman equation*, after Richard Bellman, the researcher who introduced this method. It reflects the fact in the FOMDP representation, a multi-stage decision problem can be reduced to a series of inductively-defined single-stage decision problems. Computational methods based on Bellman equations are generally referred to as *stochastic dynamic programming*; as they solve decision problems in reverse order are considered to be an efficient form of backward induction. An example

method is the *value iteration algorithm* (Bellman, 1957), which finds the maximum expected utility for a given FOMDP by iteratively computing $\tilde{\nu}_N^*, \tilde{\nu}_{N-1}^*, \ldots, \tilde{\nu}_0^*$; the optimal policy is inductively constructed as a side-effect. For finite-horizon FOMDPs, the algorithm has a computational complexity that is polynomial in the size of the state space $|\Omega_X|$, the number of actions $|A|$, and horizon length N. Value iteration can also be used to compute approximately optimal stationary policies for infinite-horizon FOMDPs, given a predefined bound on the loss in expected utility that the decision maker is willing to accept. It is possible though that the number of required iterations grows exponentially in the discount rate λ, (Puterman, 1994).

Remarks on the FOMDP representation

Markov decision processes provide an intuitive and expressive representation for decision problems that involve temporal progression and continuous interaction of a decision maker with a dynamic system. The underlying theory of stochastic Markov processes has been studied extensively in control theory. Although the Markov property on temporal evolution of the system may seem restrictive, it is not truly prohibitive: any non-Markovian system of finite order (i.e. whose dynamics depend on at most n previous states for some finite $n \in \mathbb{N}$) can be converted to an equivalent, though larger Markovian model, (Luenberger, 1979). So, the Markov property essentially applies to the model and not to the system itself.

We have discussed FOMDPs with finite and infinite planning horizons. Conceptually speaking, there exists a third possibility where it is not possible to determine the number of action choices needed to solve a particular decision problem in advance; the horizon is then said to be *indefinite*. Such problems can be cast in FOMDPs with a finite horizon, where a number of decision moments remain unused; the horizon length N serves as an upper bound on the number of decisions. The set A of actions then has to be supplemented with a special action ε, which has the interpretation of *skipping* the current decision moment without taking action. With this action, it is also possible to model decision-making scenarios where only a small number of moments are selected for intervention whereas the other moments pass uncontrolled. In other words, it allows us to model the *timing* of actions. We do note, however, that from a formal point of view there exists no difference between ε and other actions in A.

There are two significant drawbacks to FOMDPs that appear in their application to real-world problems. The first drawback to FOMDPs is their coarse representational granularity: the transition probabilities and rewards are described by functions whose complexity grows polynomially in the size of the state space and number of actions. In many problem domains, the number of probabilities and rewards that have to be specified is therefore enormous.

Example 4.18 *In the MDP model for VSD of Example 4.17, there are* $|\Omega_X| = 4 \cdot 4 \cdot 4 \cdot 4 \cdot 2 \cdot 4 \cdot 2 = 4096$ *possible clinical states, and there are 5 actions to choose from. Therefore,* $4096 \cdot 5 \cdot 4096 \approx 8.4 \cdot 10^6$ *transition probabilities need to be specified for*

each decision moment (assuming non-stationarity). Furthermore, reward values have to be assessed for $4096 \cdot 5 = 20480$ *state-action pairs.*

Furthermore, when the system dynamics and reward mechanisms consist of several interacting factors, little insight is to be gained from a system whose reasoning is based on a representation where such factors remain completely implicit. The second drawback to FOMDPs is the assumption of full observability of the system state. In many decision problems under uncertainty, this assumption is highly unrealistic. In medical domains, for example, it is often the underlying, unobservable physiological state of the patient that determines state progression over time, and a model in which this mechanism is ignored will rarely be adequate.

Example 4.19 *In the VSD domain, the attributes* `hfail` *(heart failure) and* `death` *can be easily observed in all circumstances. This is however not the case for the other attributes of the model in Example 4.17: size of the VSD, pulmonary arteriopathy, pulmonary vascular resistance, shunting, and pulmonary hypertension are attributes whose precise values often remain hidden for the treating clinician.*

Both drawbacks can be alleviated by extending the basic FOMDP formalism. In the next subsection, we discuss partially-observable Markov decision processes. In these processes, the assumption of full observability is lifted. In Section 4.3 we discuss the alternative representation of transition probability functions in dynamic probabilistic networks, which allows for a representation of system dynamics at the level of individual state attributes.

4.2.2 Partially-observable Markov decision processes

Partially observable Markov decision processes, or *POMDPs* for short, are a generalisation of FOMDPs that allow for uncertainty regarding the system state from the perspective of the decision maker: instead of observing the entire state of the system prior to each decision, the decision maker now observes a subset of state variables after the decision, where the action that is chosen determines the observability of state variables. Therefore the tradeoff between actions now does not only concern their immediate and long-term effects, but also their information-gathering properties. This introduces severe complications to the decision problem, especially when the observation opportunities are limited; in those cases, complex decision strategies may be needed to guarantee optimal control of the system. POMDPs were first described in the literature by A. Drake in 1962, and later (independently) by K. Aström in 1965.

Definition 4.20 (POMDP model) *Let* X *be a set of random variables. A* partially-observable Markov decision process *over* X *is a 5-tuple* $\mathcal{M}^{\text{PO}} = (T, A, \Theta, R, O)$*, where*

- (T, A, Θ, R) *describes a Markov decision process over* X*, and*
- $O = \{o_t : A \to \wp(X) \mid t \in T\}$ *is a set of* observation functions.

When the decision maker chooses action $a \in A$ at time point $t \in T$, he receives evidence on the subset of state variables $Y = o_t(a)$ immediately thereafter, i.e. the actual configuration of Y is observed; the state transition that results from the action choice takes place afterwards. When $o_t(a) = X$ for all time points $t \in T$ and all actions $a \in A$ and the initial system state is known to the decision maker, the model reduces to a FOMDP. In the typical case though, we have that $o_t(a) \subsetneq X$, and at each time point, part of the system state remains hidden for the decision maker. The initial state is equally unknown; instead, a probability distribution P_0 on the initial state is given. This distribution may implicitly contain evidence on a set of case parameters $Y \subseteq X$, i.e. when $P_0(Y = S_Y) = 1$ for some $S_y \in \mathrm{dom}(Y)$. Process dynamics and evaluation criteria are the same as in the fully-observable case; we will therefore not elaborate on them. The formulation of solutions and the algorithmic methods to solve POMDPs are however significantly different; we will focus on these issues in the discussion below.

Example 4.21 *The FOMDP model for the VSD domain from Example 4.17 can be extended to a POMDP with a set* O *of observations functions, where for each* $o_t \in O$ *we have*

$$\begin{aligned} o_t(\texttt{echo}) &= \{\texttt{VSD}, \texttt{hfail}, \texttt{death}\}, \\ o_t(\texttt{med}) &= \{\texttt{hfail}, \texttt{death}\}, \\ o_t(\texttt{cath}) &= \{\texttt{shunt}, \texttt{hfail}, \texttt{death}\}, \\ o_t(\texttt{surg}) &= \{\texttt{VSD}, \texttt{hfail}, \texttt{death}\}, \text{ and} \\ o_t(\texttt{biop}) &= \{\texttt{pmart}, \texttt{hfail}, \texttt{death}\}. \end{aligned}$$

The variables `hfail` *and* `death` *are always observable and therefore included in each of the observation sets.*

Formulation of solutions

The principal difference between FOMDPs and POMDPs is that under partial observability, the stochastic process is non-Markovian from the viewpoint of the decision maker. This is due to the uncertainty regarding the system states: any of the past actions and observations in a concrete evolution of the decision process may influence the decision maker's beliefs concerning the current system state. Decision-making policies therefore have to be based on all what is known from the past if we want to maximise expected utility; simply considering the most recent observation is insufficient. The result of this fact is that the problem of solving an infinite-horizon POMDP is undecidable, (Madani et al., 1999).

Formally, the decision maker's knowledge ψ of the past will consist of a sequence of alternating action choices and observations:

$$\psi = a_0, S_{Y_0}, a_1, S_{Y_1}, \ldots, a_\tau, S_{Y_\tau} \tag{4.21}$$

where $S_{Y_t} \in \mathrm{dom}(Y_t)$ and $Y_t = o_t(a_t)$ for all $t = 0, \ldots, \tau$; we refer to ψ as a *decision-making context*. Note that the observation at time point t follows the action choice at that time point. Let Ψ_τ be the set of all possible decision-making contexts up to time point $\tau \in T$. In POMDPs, a decision function δ_t^{PO} for time point $t \in T$, $t > 0$, in policy π is a partial function

$$\delta_t^{\mathrm{PO}} : \Psi_{t-1} \to A, \tag{4.22}$$

and a decision function δ_0^{PO} for the initial time point $t = 0$ is a constant from the set of actions (i.e. $\delta_0^{\mathrm{PO}} \in A$) as the decision-making context is then empty. It will often be an action that is expected to provide much evidence and to have little impact on the system state. The functions for non-initial time points only need to cover decision-making contexts that are compatible with earlier decision functions and can therefore be defined as partial functions.

Now, let H_τ as in the previous subsection denote the set of all possible decision-making histories up to decision moment $\tau \in T$, i.e. each $h \in H_\tau$ consists of sequences of entire system states and action choices up to that moment. We define the mapping $\overline{o}_{\tau'} : H_\tau \to \Psi_{\tau'}$, $\tau' \leq \tau$, from decision-making histories to contexts as follows:

$$\overline{o}_{\tau'}(h) \;=\; a_0, S_{Y_0}, \ldots, a_{\tau'}, S_{Y_{\tau'}} \tag{4.23}$$

if $h = (S_0, \ldots, S_\tau, a_0, \ldots, a_\tau)$, where for all $t = 0, \ldots, \tau'$ we have $o_t(a_t) = Y_t$ and S_{Y_t} is the subvalue associated with Y_t in S_t, i.e. $X = S_t \vdash Y_t = S_{Y_t}$. In words, $\overline{o}_{\tau'}(h)$ is the context that represents precisely the decision maker's knowledge of history h up to time point τ'. Given a policy $\pi = \{\delta_t^{\mathrm{PO}} \mid t \in T\}$ and an initial system state $S \in \Omega_X$, a (conditional) probability distribution $P_\tau^{(S,\pi)}$ on decision-making histories up to time point τ is again induced using

$$P_\tau^{(S,\pi)}(h) \;=\; \prod_{t=0}^{\tau-1} \theta_t(S_t, a_t, S_{t+1}) \tag{4.24}$$

for all $h = (S_0, \ldots, S_\tau, a_0, \ldots, a_\tau)$ having $S_0 = S$ and $\delta_t^{\mathrm{PO}}(\overline{o}_t(h)) = a_t$. As with FOMDPs, we have $P_\tau^{(S,\pi)}(h) = 0$ for all other histories $h \in H_\tau$. Given a probability distribution P_0 on initial states, the expected utility $\tilde{u}_\pi$ of policy π equals

$$\begin{aligned} \tilde{u}_\pi \;&=\; \sum_{S \in \Omega_X} P_0(S) \cdot E_{P_N^{(S,\pi)}}(u) \\ &=\; \sum_{S \in \Omega_X} P_0(S) \cdot \sum_{h \in H_N} u(h) \cdot P_N^{(S,\pi)}(h). \end{aligned} \tag{4.25}$$

In principle, this provides the criterion to compare decision-making policies and therefore to computationally solve POMDPs. Unfortunately, the space of possible policies is immense, and in the worst case each possible policy needs to be considered when solving a POMDP. When $k = |A|$ is the number of available actions and

$$m \;=\; \max\{|\Omega_Y| \mid Y = o_t(a), t \in T, a \in A\} \tag{4.26}$$

is an upper bound on the number of possible observations, then the size of decision-making policies is bounded by m^N and the number of distinct policies is $\mathcal{O}(k^{m^N})$. As a consequence, the problem of solving a POMDP is PSPACE-complete, (Papadimitriou and Tsitsiklis, 1987).

Throughout the years, several simpler types of POMDP policy have been investigated. The simplest type of POMDP policy basically ignores the unreliability caused by partial observability, and bases action choices on the most recent observation only; such a policy is called *memoryless*. Memoryless policies often have a very poor performance, but notwithstanding their simple formulation, the problem of finding the optimal memoryless policy is NP-hard, (Littman, 1994; Littman, 1996). Better performance can be achieved by taking into account a number of recent actions and observations; we then speak of a *history-window policy*, (White and Scherer, 1994; Platzman, 1977). More generally, we can think of *finite-memory policies*, which base decisions on a finite amount of information about the past, but unlike in history-window policies, the information can be obtained arbitrarily long ago. Because they are more expressive, general finite-memory policies can be defined that perform better that any history-window policy. However, there are also POMDP problems for which no finite memory is sufficient to guarantee optimality, and we have to resort to general context-dependent policies.

Solution methods

The fact that system states are partially unknown to the decision maker in a POMDP also means that we cannot decompose the decision problem using the Markov property (as in Equation 4.19 on page 99). Therefore, stochastic dynamic programming techniques do not apply, and we seem to be forced to explicitly search through the space of policies, and evaluate each of them by iterating over all possible decision-making histories. This approach is not feasible because of the highly combinatorial nature of these structures. Fortunately, there exists an alternative way to solve POMDPs which is more efficient: we can transform the POMDP into an equivalent model where stochastic dynamic programming is applicable. This is the predominant approach to solving POMDPs, and we will briefly describe it here.

The key insight underlying the approach is that from the probability distribution P_0 on initial states and a context $\psi \in \Psi_t$ of actions and observations, we can compute a probability distribution P_{t+1} on system states at time point $t+1$ that adequately summarises all the available information. Such a probability distribution is called a *belief state* at time point $t+1$, as it can be regarded as representing the beliefs of the decision maker concerning the system state at that time point. As belief states provide an adequate summary of the decision-making context, the Markov property is regained. From a given belief state P_t at time point $t \in T$, and an action choice $a_t \in A$ and observation S_{Y_t} that follow, we can compute the next belief state P_{t+1}:

$$P_{t+1}(S_{t+1}) \quad = \quad \sum_{S_t \in \Omega_X} \theta_t(S_t, a_t, S_{t+1}) \cdot P_t(S_t \mid S_{Y_t}) \tag{4.27}$$

for all $S_{t+1} \in \Omega_X$. Therefore, the P_{t+1} is independent of $P_0, \ldots, P_{t-1}$ given P_t, action a_t, and observation S_{Y_t}: the sequence of belief states is a Markov chain. We can therefore transform any POMDP to a fully-observable MDP over belief states; we refer to a transformed POMDP as a *belief MDP*. The belief MDP has the property that an optimal Markovian policy for it will give rise to optimal behaviour for the original POMDP, (Aström, 1965; Smallwood and Sondik, 1973).

The belief MDP is Markovian, and can therefore be solved using value iteration, (Sondik, 1971). However, the state space of the belief MDP is the set of all probability distributions over X. When the state space in the original POMDP contains m states, i.e. $|\Omega_X| = m$, then the corresponding belief space has the complexity of the m-dimensional unit simplex $[0, 1]^m$. Value functions and policies have to be defined as functions over this highly complex, continuous space. This is the main obstacle that now needs to be overcome and most algorithms for solving POMDPs are actually different approaches to handle it. The details of these approaches fall outside the scope of this thesis; we refer the interested reader to the surveys by Monahan (1982) and Lovejoy (1991), and to the papers by Littman et al. (1995) and Cassandra et al. (1997) for recent advances.

Remarks on the POMDP representation

POMDPs provide a powerful framework for decision-theoretic planning where both uncertainty in action effects and imperfect observability are essential. They allow for the expression of many decision-making scenarios, including reasoning with incomplete information, and planning of both information-gathering and intervening actions. POMDP models have been developed in such diverse domains as machine maintenance, medical diagnosis and treatment planning, and acquisition of cognitive skills.

In contrast with influence diagrams, POMDPs have no problems with the representation of informational asymmetry. Observability of state variables is associated with actions instead of decision moments; the asymmetric nature of decision problems with imperfect observability is therefore easily handled by the formalism, a point which is also illustrated by Example 4.21 on page 102. Dealing with relevantial asymmetry is more troublesome. It is possible to extend the basic formalism with a specification of explicitly excluded actions in various decision-making contexts, but the standard method of solving POMDPs by stochastic dynamic programming on the belief MDP is unable to take such exclusions into account, as it solves the decision problem in reverse order.

As was noted above, the problem of solving POMDPs without restrictions is PSPACE-complete, and therefore the applicability of exact solution methods is necessarily limited to (very) small POMDPs. To make things worse, also the problem of computing approximately-optimal policies is extremely demanding from a computational point of view, and the same holds even for the seemingly uncomplicated problem of evaluating a given policy, (Goldsmith and Mundhenk, 1998). Any attempt to apply

POMDPs of considerable size without restrictions is deemed to run aground in intractability. Recent years have shown a growing interest in POMDPs from within the AI community, and several new algorithmic results have been obtained, (Cassandra et al., 1997; Zhang and Lee, 1998). It remains an open issue whether feasible solution methods can be developed for practical applications, and if so, which problem characteristics enable feasibility.

To date, little attention has been given to special forms of POMDP, where one or more domain-dependent restrictions are assumed on the components of a POMDP model. For instance, in certain domains one might assume that several actions have similar effects on state evolution or observability. Another, often reasonable assumption is that a partial ordering on system states exists, where state transitions can only lead to 'higher' states in the ordering. Probably, such assumptions can effectively be exploited by solution methods and the improved efficiency might enable practical application of POMDPs to certain domains. We will return to this issue in subsequent chapters.

Although POMDPs alleviate one of the main drawbacks to FOMDPs, the assumption of full observability, this does not hold for the other drawback, their coarse representational granularity. In the next section, we discuss the alternative, fine-grained representation of transition probability functions in dynamic probabilistic networks.

4.3 Dynamic networks

In this section we will discuss extensions to graphical representation formalisms that allows for explicit temporal modelling and reasoning. As we saw in the examples of Section 4.1, it often occurs that the variables in belief network and influence diagrams refer to events at different time points. This is in fact unavoidable in influence diagrams, as in that case the decision variables represent a temporally ordered sequence of decisions. However, the notion of time is left implicit in the representation, which is unsatisfactorily when it plays an important role in the problem domain. Both belief networks and influence diagrams have been extended to handle temporal modelling and reasoning, and the latter extension can moreover be used as a graphical representation of Markov decision processes.

Dynamic belief networks

Throughout the years, several temporal extensions to belief networks have been proposed in the literature (e.g. Berzuini, 1990; Dagum et al., 1992; Hanks et al., 1995; Aliferis and Cooper, 1996). Although these extensions differ in a number of respects, the unifying idea is to replicate a set of random variables X over a predefined set T of points in time, and define a probabilistic network over the resulting set of variable replications. We refer to this set as the *temporal extension* of X; as in the previous section, we assume that $T = \{0, \ldots, N\} \subset \mathbb{N}$.

Definition 4.22 (Temporal extension) *Let X be a set of random variables, and let T be a set of time points. The* temporal extension *of X over T is defined as the set $X(T) = \{x(t) \mid x \in X, t \in T\}$ of temporally indexed random variables, where each $x(t) \in X(T)$ can take the same values as x.*

Within the temporal extension $X(T)$ of X over T, the subset $X(t) = \{x(t) \mid x \in X\}$, $t \in T$, is called the *time slice* at time point t; it represents the state of the variables of X at that time point.

A *dynamic belief network* (DBN) now is a belief network over a temporally extended set of random variables, with the restriction that arcs in the graph may not point to variables in the past.

Definition 4.23 (Dynamic belief network) *Let X be a finite set of random variables, and let $T = \{0, \ldots, N\}$ be a set of time points. A* dynamic belief network *over X and T is a belief network $B = (G, \Gamma)$ over the temporal extension $X(T)$ (of X over T), where $x_1(t_1) \to x_2(t_2) \in A(G)$ only if $t_1 \leq t_2$.*

An arc between variables within a single time slice is called *synchronous* and is seen to represent a causal or influential mechanism whose realisation takes negligible time as compared to the time steps in T. An arc between variables of different time slices is called *temporal* and represents a mechanism that is more time-consuming (a process or evolution). The topological restriction is that temporal arcs may not point "backwards" in time; the obvious underpinning is that otherwise the causal interpretation of arcs is lost. The restriction can also be formulated as the requirement that a variable's predecessors are not located in future time slices, i.e.

$$\rho_G(x(t)) \cap X(\{t+1, \ldots, N\}) \;=\; \varnothing \tag{4.28}$$

for all $x \in X$ and $t \in T$, $t < N$. A further restriction may be that each variable's predecessors are located in the same or previous time slice; we then say that the network is *Markovian*.

Definition 4.24 (Markovian DBN) *Let $B = (G, \Gamma)$ be a dynamic belief network over X and T. We say that B is* Markovian *when for all $x \in X$ and $t \in T$, $t > 0$, we have*

$$\rho_G(x(t)) \subset X(t-1) \cup X(t). \tag{4.29}$$

Proposition 4.25 *A Markovian DBN represents a Markov process over X.*

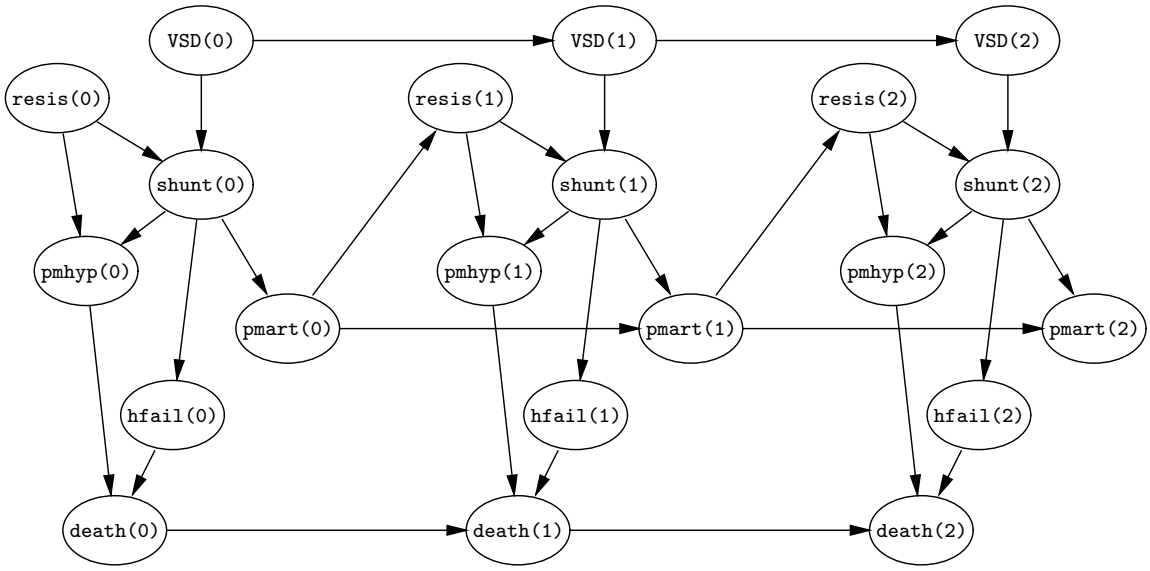


Figure 4.6: A dynamic belief network for the VSD domain.

Proof. Let $B = (G, \Gamma)$ be a Markovian DBN over X and T. Now, let $t \in T$, $t > 1$, and consider an arbitrary chain in G from a variable in one of the time slices $X(0), \ldots, X(t-1)$ to a variable in time slice $X(t)$. From the topological restrictions on Markovian DBNs it follows that the chain is blocked by time slice $X(t-1)$. As a result, we have that

$$\langle X(t) \mid X(t-1) \mid X(\{0, \ldots, t-1\}) \rangle_G^d, \tag{4.30}$$

and therefore

$$X(t) \perp\!\!\!\perp_P X(\{0, \ldots, t-1\}) \mid X(t-1) \tag{4.31}$$

for all probability distributions P on $X(T)$ for which G is an I-map. □

Example 4.26 *A dynamic belief network for the VSD domain is shown in Figure 4.6. The network is Markovian, and contains 7 of the 8 variables that were used in the static belief network of Example 4.6: the variable* `closure` *is omitted as spontaneous closure of the VSD is modelled implicitly by diminishing values over time for the variable* `VSD`.

DBNs allow for explicit temporal reasoning over the variables involved, and are therefore preferred over static belief networks when time and change are essential ingredients of the problem domain. The improved expressiveness of DBNs comes however at the price of a significantly increased computational cost of probabilistic inference. Recall from Subsection 4.1.1 that probabilistic inference in general is NP-hard: the

number of computations may be exponential in the number of variables involved. As DBNs are defined over the temporal extension of the set of domain variables, inference may quickly become intractable in these networks when the set T is large. Under certain conditions, tractable inference is possible though. An example is found in the Kjærulff's (1992) work, which extends the clique-tree propagation method for belief-network inference to Markovian DBNs.

Dynamic influence diagrams

Dynamic influence diagrams (DIDs; Dean and Kanazawa, 1989; Tatman and Shachter, 1990; Provan and Clarke, 1993) are defined in a similar way as DBNs. As with their static counterparts, the main difference is that influence diagrams include a set of decision variables and a value node. In a DID, each decision node is associated with one of the time points involved.

Definition 4.27 (Dynamic influence diagram) *Let X be a set of random variables, and let $T = \{0, \ldots, N\}$ be a set of time points. A* dynamic influence diagram *over X and T is an influence diagram $DID = (G, \Gamma, u)$ over $X(T)$, where*

- $V(G) = X(T) \cup D(T) \cup \{v\}$,
- $D(T) = \{d(t) \mid t \in T\}$ *is a set of temporally indexed decision variables,*
- $d(t) \rightarrow d(t+1) \in A(G)$ *for all* $t = 0, \ldots, N-1$,
- *for all variables* $y(t_1), z(t_2) \in X(T) \cup D(T)$ *we have that* $y(t_1) \rightarrow z(t_2) \in A(G)$ *only if* $t_1 \leq t_2$.

In the graph of a DID there exists a directed path comprising all decision variables; any DID is therefore a regular influence diagram. The interpretation of the various types of arcs in a DID is completely similar to their interpretation in influence diagrams; we therefore do not elaborate on it here. Furthermore, the value node v is again a sink node in G and u is a utility function over the universe $\Omega_{\rho_G(v)}$ of its predecessors.

Example 4.28 *Figure 4.7 shows a DID for the VSD domain that extends the earlier described DBN with three decision nodes $d(0)$, $d(1)$, and $d(2)$ and a value node. It is assumed that each of decision variables can take one of the values* `echo`, `cath`, `biop`, `surg`, *and ε; the latter value again has the interpretation of skipping the current decision moment without taking action. The utility function associated with the value node is based on the three decisions and the values of three random variables from the last time slice, representing final size of the VSD, state of the pulmonary arterioles, and mortality.*

The Markov property can also be formulated for dynamic influence diagrams.

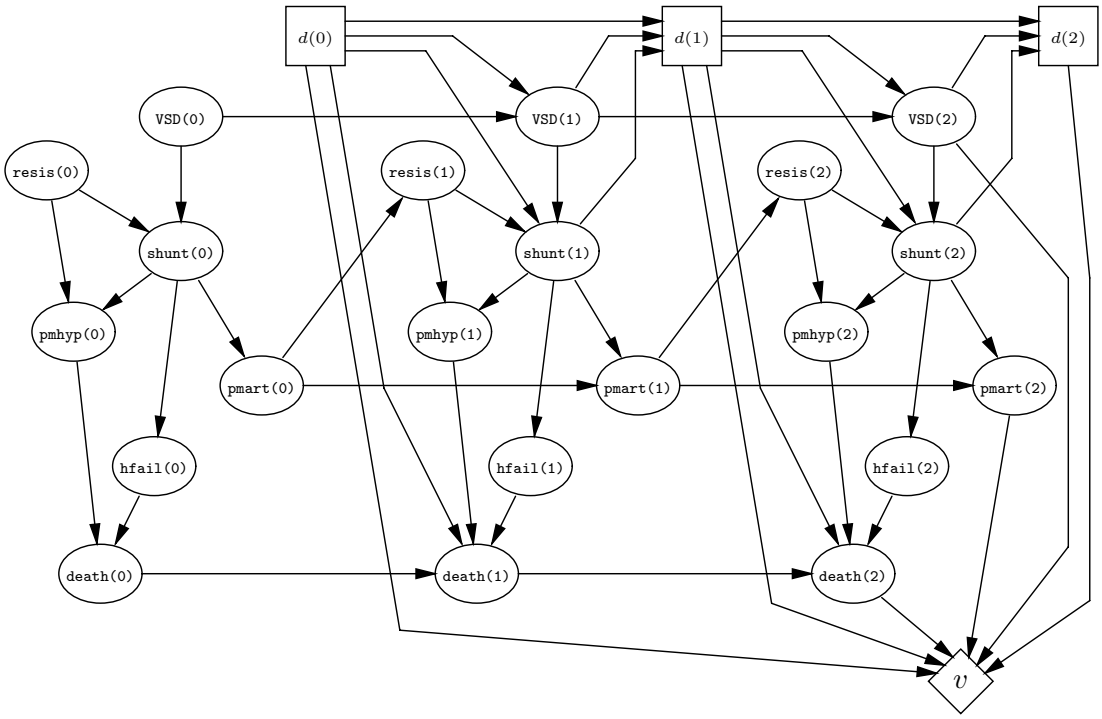


Figure 4.7: A dynamic influence diagram for the VSD domain.

Definition 4.29 (Markovian DID) *Let $DID = (G, \Gamma, u)$ be a dynamic influence diagram over X and T. We say that the diagram is* Markovian *when for all $x \in X$ and $t \in T$, $t > 0$, we have*

$$\rho_G(x(t)) \subset X(t-1) \cup X(t) \cup \{d(t)\}. \tag{4.32}$$

Note that the above topological restriction only pertains to arcs leading to random variables in the diagram, and is therefore essentially a restriction on the conditional independency structure of the underlying probability distribution. There are no specific restrictions on the arcs leading to decision variables (modelling informational constraints) or to the value node (modelling influences on utility) other than those for DIDs in general. We remark that the DID of Figure 4.7 is Markovian as all arcs leading to random variables depart from the same or the previous time slice.

Graphical representation of POMDPs

Both POMDPs and Markovian DIDs are representations of sequential decision problems under uncertainty that include notions of time, state dynamics, and observability, and both representations incorporate the Markov assumption. It is not surprising then, that we can sometimes regard a Markovian DID as a graphical representation of an underlying POMDP. Unfortunately, this not always true: there exists POMDPs than cannot be expressed as a Markovian DID, and conversely, there exist Markovian

DIDs that cannot be taken to represent a POMDP. The main reasons that the correspondence does not always hold is that POMDPs make certain assumptions on the structure of the utility function where DIDs do not, and that the notions of observability differ in both formalisms. When the correspondence does hold, the graphical representation of POMDPs in DIDs integrates all concepts that were described in this chapter; it is broadly regarded as one of the most powerful formalisms for decision-theoretic modelling and reasoning to date. Note that a DID is much more economical in its space requirement than a POMDP and is therefore the preferred representation in practical settings; we will shortly illustrate this with an example.

A Markovian dynamic influence diagram $DID = (G, \Gamma, u)$ can be viewed as representing a POMDP under the following three conditions. First, there should be no observations prior to the first decision (as POMDPs do not know case parameters). That is, decision variable $d(0)$ may not have predecessors in the graph. Second, there may not be observational time lags, i.e. a given variable $x(t_1)$ becomes observable immediately at $t_1 + 1$, or not at all. Formally, if $x(t_1) \in \rho_G(d(t_2))$ for some $t_2 > t_1 + 1$, then also $x(t_1) \in \rho_G(d(t_1 + 1))$. Third, the utility function can be written as the (discounted) sum of temporal rewards. That is, there exist a set of reward functions $R = \{r_t : \Omega_{X(t)} \times \Omega_{d(t)} \rightarrow \mathbb{R} \mid t \in T\}$ and a discount factor $\lambda \in \mathbb{R}$ such that

$$u(\mathrm{C}_{\rho_G(v)}) \;=\; \sum_{t=0}^{N} \lambda^t r_t(\mathrm{C}_{X(t)}, \mathrm{C}_{d(t)}). \tag{4.33}$$

So, the time slices of the diagram are mutually additive independent sets in the preference ordering induced by the utility function; this is called *time separability* (Luenberger, 1973), and can be exploited to perform stochastic dynamic programming to solve the diagram, (Tatman and Shachter, 1990). Note that if $\rho_G(d(t)) = X(t-1)$ for all time points $t \in T$, $t > 1$, in a dynamic influence diagram, then the underlying Markov decision process is *de facto* fully observable; otherwise, it is partially observable.

We can also follow the opposite direction and ask ourselves under which conditions a given POMDP $\mathcal{M}^{\mathrm{PO}} = (T, A, \Theta, R, O)$ can be represented by a Markovian DID. This is possible when the observation functions in O are symmetrical; recall from Subsection 4.1.2 that influence diagrams cannot handle informational asymmetry. That is, a DID representation for the POMDP exists when $o_t(a_1) = o_t(a_2)$ for all decision moments $t \in T$ and actions $a_1, a_2 \in A$; no further conditions on the POMDP are required.

Example 4.30 *The POMDP for the VSD domain of Example 4.21 (page 102) cannot be represented by a DID as it embodies informational asymmetry. If we take* $o_t(a) = \{\texttt{VSD}, \texttt{shunt}\}$ *for all actions* $a \in \{\texttt{echo}, \texttt{med}, \texttt{cath}, \texttt{surg}, \texttt{biop}\}$ *however, then the Markovian DID of Figure 4.7 represents the POMDP, provided that a time-separable utility function is employed.*

It should be noted that in the DID representation, conditional independence relations are exploited to arrive at a model of much finer representational granularity than in the MDP representation: probabilistic relations are quantified at the level of individual variables instead of at the level of complete time slices. The main advantages are (i) a significant reduction in the number of probabilities that have to be specified, and (ii) a model that is easier to comprehend, and therefore easier constructed and maintained. Similar observations hold for the specification of rewards: the fact that often many variables are uninfluential on utility is readily exploited in the representation.

Example 4.31 *Recall from Example 4.18 that the FOMDP and POMDP models for the VSD domain described in the previous section required the specification of* $8,4 \cdot 10^6$ *transition probabilities and 20480 reward values. The DID representation of Figure 4.7, requires "only" 2150 conditional probabilities and 2000 utility values.*

We do not study the correspondence between Markovian DIDs and POMDPs in more detail here; this is done in Section 5.5 of the next chapter. In that chapter, we develop a formal framework for decision-theoretic planning that is more expressive than both POMDPs and DIDs and therefore allows for a smooth comparison.

4.4 Discussion

In this section, we summarise this chapter's review of representation formalisms and evaluate these formalisms with respect to their expressiveness, flexibility, compactness, and representational granularity. The evaluation will provide the motivation for a new, more elaborate framework for decision-theoretic planning which is presented in Chapter 5.

The first decision-theoretic representation formalism discussed in this chapter was the influence diagram. Influence diagrams are graphical representations that can be viewed as extending the belief-network representation with (i) decision variables to model the relevant moments of choice, and (ii) a value node to encode a utility function over the network variables that influence the decision maker's preferences. Influence diagrams can also be regarded as a compact representation of symmetric decision trees: whereas the size of a decision tree grows exponentially in the number of variables involved, the size of an influence diagram will be proportional to that number. It should be noted though that the complexity of a variable's conditional probability assessment function grows exponentially in the number of the variable's parents in graph. The complexity of these functions may therefore be considerable in dense graphs.

There are several drawbacks to the influence-diagram representation. The meaning of arcs in the graph depends on the type of node to which they lead, which is often felt as confusing. Influence diagrams are therefore generally judged as being less

comprehensive than their purely probabilistic siblings, belief networks, especially in large domains. Another drawback is that influence diagrams have difficulties with handling asymmetry in decision problems. Also, influence diagrams are designed to solve a limited number of decision problems for which a set of decision policies is implicitly pre-specified in the representation; they lack the generative character of planning formalisms. These drawbacks can be largely traced back to the fact that influence diagrams strongly interweave different types of domain knowledge into a single representation. The resulting compactness comes at the price of limited flexibility and expressiveness.

A possible way to overcome this problem is to use an underlying, generative model of the problem domain and construct suitable influence diagrams from it for specific problem types; this is called *knowledge-based model construction*, (Breese et al., 1994). Examples are found in the work of Agosta (1996), who uses classical (i.e. non-probabilistic) planning methods to construct the decisional part of an influence diagram, and Egar et al. (1992), who describe a method for the automatic generation of influence diagrams using a graph-grammar and a set of pre-defined network chunks.

The underlying domain model for an influence diagram can also be cast as a Markov decision process, the second decision-theoretic representation discussed in this chapter. In contrast with influence diagrams, MDPs possess the characteristic properties of a planning formalism: MDPs allow for the generation of a large number of decision-making scenarios from generic definitions of the interactions between domain elements. MDPs also explicitly cover the notions of time and change, which are indispensable for a true planning system. A drawback to the original MDP representation is that the effects of actions are expressed at the level of system states rather than at the level of individual variables. If we choose to express state dynamics in dynamic influence diagrams however, then we can alleviate this drawback and take advantage of probabilistic independencies between domain variables in representation and reasoning.

Dynamic influence diagrams were the third and last representation formalism discussed in this chapter, and can be regarded as the most powerful formalism for decision-theoretic planning to date. Although it captures important qualitative and quantitative features of a decision process, it considers all state variables and all decisions as similar: it does not identify the roles of different variables within that process. If we revert to the type of decision-making task we are seeking to support, the planning of clinical diagnostic and therapeutic actions, then it appears that we can identify very different roles of variables and decisions, and that these roles are crucial in problem solving. For instance, diagnostic decisions are aimed at gathering more information about a patient, whereas treatment decisions aim to cure the patient on the basis of that information. It is immediately apparent that one would not consider a decision policy that performs diagnosis after treatment. Similarly, some state variables pertain to the internal physiological state of a patient, whereas others may represent the outcomes of investigative procedures.

In the next chapter, we present a formal framework for decision-theoretic planning that builds on the ideas that underlie the representation formalisms discussed here. The approach we take is however more fundamental as we define concepts such as *choice*, *history*, *decision process*, *prediction*, and *plan* directly from the primitive elements of a planning domain, instead of indirectly from a given representation. Among the framework's components are symbolic and probabilistic languages to allow for explicit reasoning with these concepts, and to analyse and describe their mutual relationships.

CHAPTER 5

A framework for decision-theoretic planning

Making a sequence of decisions under conditions of uncertainty can be regarded as a form of *planning* – this was one of the introductory remarks to Markov decision processes in the previous chapter. However true this may be, the standard formulation of these processes hardly reminds of the type of symbolic planning that is traditionally studied in AI. In particular, it does not allow for reasoning about the relations between decisions and states at the level of individual variables, or for problem solving through explicit manipulation of symbolic plans.

In this chapter, we present a formal framework for planning under uncertainty that obviates these deficiencies by integrating probabilistic and symbolic reasoning. The objective is to provide a language for decision-theoretic planning domains which allows for direct manipulation of symbolic structures that describe decisions and plans. The essence of planning under uncertainty is the continual existence of contingencies that should be taken into account when action choices are made; a significant part of this chapter is therefore devoted to a theory of *contingency planning*. The framework links up with the decision-theoretic representation formalisms of the previous chapter, but takes a more fundamental approach in the formalisation of planning and decision processes. The main difference is that our framework is not intended as a computational architecture, but serves to study the nature of decision-theoretic planning problems.

This chapter is organised as follows. In Section 5.1, we develop a general, formal language to describe probabilistic planning processes. Section 5.2 discusses the notions of control and observability for this language. Then, in Section 5.3, we develop the

theory of contingency planning within our framework. In Section 5.4 we describe the formulation of planning objectives in a decision-theoretic fashion. This completes our framework and provides the opportunity to re-evaluate and compare the representation formalisms of the previous chapter; this is the subject of Section 5.5. The chapter is concluded with a discussion in Section 5.6, where we evaluate the framework and discuss related work by other authors.

5.1 Formal foundations

In this section we develop the formal structures to describe planning processes under uncertainty in our framework. Conceptually speaking, our framework is reminiscent of the formalism of partially-observable Markov decision processes: the planning task is envisioned as the problem of controlling a dynamic system over time, where the effects of action choices are uncertain and the state of the system is partially hidden from observation. We will no longer refer to a person that decides upon the action choices as a decision maker but use the term *planning agent* instead. It emphasises the fact that making decisions is just one of the tasks that this person faces, among other tasks such as evaluating and repairing given plans and investigating various properties of plans such as their robustness, the associated costs and risks, and expected effects on the system.

We start this section by developing a symbolic language over the elements of a planning domain (Subsection 5.1.1); this is followed by the formal characterisation of decision processes (Subsection 5.1.2).

5.1.1 The planning language

The starting point of our framework for decision-theoretic planning is a specification of the *state space* of a dynamic system, an *action set* constituting the alternatives that are available to the planning agent, and a *time frame* that is considered relevant for the planning task. A joint specification of these elements will be called a *decision basis*.

Definition 5.1 (Decision basis) *A* decision basis *is a tuple* $\mathcal{B} = (X, A, T)$*, where* X *is a set of system attributes,* A *is a set of actions, and* T *is a time frame.*

In a decision basis $\mathcal{B} = (X, A, T)$, the set X consists of variables that jointly describe the state space of a dynamic system. Conceptually, each variable $x \in X$ describes some attribute of this system, and we therefore use the term *system attributes*. As in the previous chapters, the notion of *configuration* is used as a shorthand notation for the joint assignment of values to a set of variables. With this notation, the set Ω_X of configurations of X covers all possible states that the system may occupy. The

set A of actions models the decision alternatives that are potentially available to the planning agent; it represents the available means to control and observe the dynamic system. The set T is a linearly ordered set of points in time where a snapshot of the current system state is considered and the decision maker can choose an action.

A decision basis describes the principal elements of a planning domain but not their various types of interaction. To specify such interactions, a decision basis is supplemented with models of observation and control. *Control models*, defined in Subsection 5.2.1, describe the effects of action choices on the dynamic system. *Observation models*, which are introduced in Section 5.2.2, describe the observability of that system under different action choices.

As in the previous chapters, we assume state spaces and action sets to be finite unless stated otherwise. The reason for this assumption is that the finite case links up with current medical practice: clinical decision problems and solutions are generally described in finitely discretised variables. Furthermore, infinite and especially continuous sets radically change the nature of the mathematics involved; they are therefore not covered in the present exposition. We note that finiteness of a state space implies that it is described by a finite number of attributes, and each attribute has a finite value domain.

Time and change

The *time frame* T is a set of points in time, or *decision moments*, relevant to the planning task. A linear ordering, representing temporal precedence, is assumed to exist on this set. A most general option would be to assume T being a closed interval of the set of the real numbers, and use the ordering $<$ on $\mathbb{R}$ for temporal precedence. As we want also to restrict ourselves to finiteness for time frames, we take $T = \{0, 1, \ldots, N\} \subset \mathbb{N}$. The elements $t = 0$ and $t = N$ are called the *initial* and *final moments* in T, respectively.

Definition 5.2 (Time segment) *Let T be a finite, linearly ordered set of time points. A subset $T' \subseteq T$ of subsequent points in T is called a* time segment *in T, notation $T' \sqsubseteq T$. We say that* $\inf T'$ *and* $\sup T'$ *are the* initial *and* final *moments of T', and $|T'|$ is called its* length*. If, in addition,* $\inf T' = \inf T$*, then T' is called an* initial segment *of T; if* $\sup T' = \sup T$*, then T' is called a* final segment *of T.*

As the segment relation is reflexive, the time frame T itself is also a segment, with the unique property that it is both initial and final.

Notation 5.3 *For $t_1, t_2 \in T$, we will use $[t_1, t_2]$ as a shorthand notation for the time segment $\{t \in T \mid t_1 \leq t \leq t_2\}$.*

As temporal progression is of central concern in our framework, an important role is played by the dynamics of the system's state and changes in the planning agent's

behaviour over time. A natural way to formalise such changes is accomplished by regarding state and behaviour as functions of time; we refer to such functions are *state* and *action sequences*, respectively.

Definition 5.4 (Sequence) *Let T' be a time segment in T. A* state sequence *over T' is a mapping $\sigma : T' \to \mathrm{dom}(X)$, and an* action sequence *over T' is a mapping $\alpha : T' \to A$. The sets of all possible state and action sequences over T' are denoted by $\mathcal{S}(T')$ and $\mathcal{A}(T')$, respectively.*

A state sequence σ over time segment T' describes states of the dynamic system at subsequent time points in T'. There exist $|\Omega_X|^n$ possible state sequences over T' when $n = |T'|$; we use σ to denote the unique 'empty' state sequence over the empty time segment $T' = \varnothing$. Conceptually, $\check{\sigma}$ describes the system's dynamics prior to the realisation of a concrete state. Similarly, an action sequence α over T' describes all subsequent decisions during the time segment. There exist $|A|^n$ possible action sequences over T'; we use $\check{\alpha}$ to denote the unique 'empty' action sequence over the empty time segment. Conceptually, this action sequence represents behaviour of the planning agent prior to any moment of choice.

A *planning history* is a pair of state and action sequences, and thus provides a complete description of the planning process over the time segment at hand.

Definition 5.5 (History) *Let T' be a time segment. A* planning history *over T' is a pair $h = (\sigma, \alpha)$, where σ and α are a state sequence and an action sequence over T', respectively. The set of all possible histories over T' is denoted by $\mathcal{H}(T')$.*

We assume that at each time point t during the realisation of a history $h = (\sigma, \alpha)$, the dynamic system has occupied state $\sigma(t)$ before action $\alpha(t)$ is chosen, and that information on that state $\sigma(t)$ may be used by the planning agent in his choice. There exist $|\Omega_X|^n \cdot |A|^n$ possible histories over time segment T' when $n = |T'|$. To avoid abundance of parentheses, we will write $\mathcal{S}[t_1, t_2]$, $\mathcal{A}[t_1, t_2]$, and $\mathcal{H}[t_1, t_2]$ instead of $\mathcal{S}([t_1, t_2])$, $\mathcal{A}([t_1, t_2])$, and $\mathcal{H}([t_1, t_2])$, respectively.

Planning expressions

We will now define a symbolic planning language from the elements of a decision basis. The planning language builds on the definition of Boolean algebra of logical propositions over value assignments from Section 3.1. Whereas in the previous chapters such algebras were primarily used to capture joint value assignments to sets of variables (configurations), we will employ their full expressive power in the current framework.

Definition 5.6 (Planning expression) *Let $\mathcal{B} = (X, A, T)$ be a decision basis. The set $\Phi(T)$ of* planning expressions *over T is the Boolean algebra $\beta(W)$ spanned by the set $W = X(T) \cup D(T)$, where*

- $X(T) = \{x(t) \mid x \in X, t \in T\}$ *is the temporal extension of* X *over* T *where each* $x(t) \in X(T)$ *takes values from* $\mathrm{dom}(x)$*, and*
- $D(T) = \{d(t) \mid t \in T\}$ *is a set of decision variables for* T*, where each* $d(t) \in D(T)$ *takes values from the action set* A*.*

The elements of $X(T)$ will be called *state variables*; the variable $x(t) \in X(T)$ describes the state of system attribute $x \in X$ at time point $t \in T$. The planning language associated with a given decision basis $\mathcal{B} = (X, A, T)$ allows to describe various types of statement concerning the evolution of a decision process over time frame T by combining value assignments to state and decision variables. We say that a given expression $\varphi \in \Phi(T)$ is *consistent* when $\varphi \not\equiv \bot$; otherwise, it is said to be *inconsistent*. For a given time segment $T' = [t_1, t_2]$, we will also use $X[t_1, t_2]$ and $D[t_1, t_2]$ instead of $X([t_1, t_2])$ and $D([t_1, t_2])$ as a shorthand notation for sets of state and decision variables over T'. As was described above, we will assume that the timing of decisions and events follows the ordering of T, where at each time point $t \in T$, the uncertain event described by the set $X(t)$ precedes decision $d(t)$.

We will now assume that $\mathcal{B} = (X, A, T)$ is a decision basis and distinguish a number of special types of planning expression from $\Phi(T)$. The first type of planning expression we distinguish is a simple fact regarding the system at an isolated time point; it formalises the notion of *system state*.

Definition 5.7 (System state) *A configuration* $\mathrm{c}_{Y(t)} \in \Phi(T)$ *of a set of state variables* $Y(t) \subseteq X(t)$ *is called a* system state *at time point* $t \in T$*. When* $Y(t)$ *is strictly smaller than* $X(t)$*, the system state is said to be* partial*; otherwise, it is said to be* complete*.*

Note that the system states $\mathrm{c}_{Y(t_1)}$ and $\mathrm{c}_{Y(t_2)}$ cannot be conflicting when $t_1 \neq t_2$, as $Y(t_1)$ and $Y(t_2)$ are then disjoint sets of variables.

The second type of expression denotes the planning agent's behaviour at an isolated time point and is called an *action choice*.

Definition 5.8 (Action choice) *An assignment* $d(t) = a$ *of action* $a \in A$ *to decision variable* $d(t)$ *is called a* decision *or* action choice *at time point* $t \in T$*.*

As the expression $d(t) = a$ is a configuration of variable $d(t)$, we will also write $\mathrm{c}_{d(t)}$ to denote a decision at time point t, thus leaving implicit which action is chosen. The partial versus complete distinction does not exist for action choices. Also action choices for different time points cannot be conflicting.

A third type of expression is obtained when we collect a set of complete system states over a given time segment $T' \sqsubseteq T$. This yields a configuration $\mathrm{c}_{X(T')}$ of $X(T')$, the

temporal extension of X over T'. It is easily seen that there exists for such a configuration a unique corresponding state sequence $\sigma \in \mathcal{S}(T')$ whose value assignments are compatible with the assignments in $\mathrm{c}_{X(T')}$ at all time points in T', i.e.

$$\text{if } \sigma(t) = S \text{ then } \mathrm{c}_{X(T')} \vdash X(t) = S \tag{5.1}$$

for all $t \in T'$ and $S \in \mathrm{dom}(X)$. In cases where there is no type confusion, we will use σ to denote $\mathrm{c}_{X(T')}$, and consider it to be an element of the planning language $\Phi(T)$.

Similarly, a collection of action choices over T' yields a configuration $\mathrm{c}_{D(T')}$ of the set $D(T')$, which uniquely corresponds to action sequence $\alpha \in \mathcal{A}(T')$ if $\mathrm{c}_{D(T')}$ and α choose the same actions at each time point $t \in T'$, or formally,

$$\text{if } \alpha(t) = a \text{ then } \mathrm{c}_{D(T')} \vdash d(t) = a \tag{5.2}$$

for all $t \in T'$ and $a \in A$. We will also write α instead of $\mathrm{c}_{D(T')}$ when there is no ambiguity. Note that with these correspondences, we have that $\check{\sigma} \equiv \top$ and $\check{\alpha} \equiv \top$ as these empty sequences represent empty conjunctions.

The fifth type of expression we distinguish is a conjunction $\sigma \wedge \alpha$ of state and action sequences σ and α over a given time segment T'. This conjunction trivially corresponds to the planning history $h = (\sigma, \alpha)$ over T', and also in this case we will use these notions interchangeably, i.e. we generally take $h \in \Phi(T)$, where h is a configuration of the set $X(T') \cup D(T')$. Using this notation, we can write each planning expression as a disjunction of planning histories.

Proposition 5.9 *Let T' be a non-empty time segment, and let $\varphi \in \Phi(T')$ be an arbitrary planning expression over T'. Then,*

$$\varphi \quad \equiv \quad h_1 \vee \cdots \vee h_k, \tag{5.3}$$

where $h_1, \ldots, h_k \in \Omega_{X(T') \cup D(T')}$.

Proof. We construct the set $H_\varphi = \{h_1, \cdots, h_k\}$ by induction on φ.

- If $\varphi \equiv \top$ then $H_\varphi = \mathcal{H}(T')$ and if $\varphi \equiv \bot$ then $H_\varphi = \varnothing$;
- if $\varphi \equiv x(t) = s$, where $x \in X$ and $t \in T'$,
  then $H_\varphi = \{(\sigma, \alpha) \in \mathcal{H}(T') \mid \sigma(t) = S,\ X = S \vdash x = s\}$;
- if $\varphi \equiv d(t) = a$, where $a \in A$ and $t \in T'$,
  then $H_\varphi = \{(\sigma, \alpha) \in \mathcal{H}(T') \mid \alpha(t) = a\}$;
- if $\varphi \equiv \psi \wedge \psi'$ then $H_\varphi = H_\psi \cap H_{\psi'}$;
- if $\varphi \equiv \psi \vee \psi'$ then $H_\varphi = H_\psi \cup H_{\psi'}$; and

- if $\varphi \equiv \neg\psi$ then $H_\varphi = \mathcal{H}(T') \setminus H_\psi$.

It is easily verified that $\varphi \equiv \bigvee_{h \in H_\varphi} h$ for each $\varphi \in \Phi(T')$. □

We will refer to the right-hand side of Equation 5.3 as the *normal form* of expression φ. If φ represents the planning agent's knowledge, then we can think of the histories $h_1, \ldots, h_k$ as the actual range of possibilities over the given time segment. Each of these histories comprises statements of which the agent is uncertain; the number k provides an indication of his uncertainty. For instance, $k = |\mathcal{H}(T')|$ if $\varphi \equiv \top$, $k = 1$ if $\varphi \equiv h$ for some $h \in \mathcal{H}(T')$, and $k = 0$ if $\varphi \equiv \bot$.

Decision rules

We now turn to expressions that explicitly guide the action choices of the planning agent; such expressions will be called *decision rules*. We first introduce the notion of *choice context*, which refers to expressions that denote possible information that may be used in choosing an action at a given time point. Recall that each decision may be based on all decisions and system states in the past and the contemporaneous system state.

Definition 5.10 (Choice context) *Let $t \in T$ be a decision moment. Any configuration of a subset of $X[0,t] \cup D[0,t-1]$ is called a* choice context *for moment t.*

When $\varphi \vdash \mathrm{c}_{x(t_1)}$ (i.e. $\mathrm{c}_{x(t_1)}$ is one of the conjuncts in φ), we say that state variable $x(t_1)$ is *covered* by φ, and similarly if $\varphi \vdash \mathrm{c}_{d(t_2)}$ we say that φ *covers* a decision at time point t_2. Note that by definition, $t_1 \leq t$ and $t_2 < t$ if φ is a choice context for moment t. It is possible that a choice context does not cover any state variables and it is also possible that a choice context does not cover any decisions.

A choice context for time point t does not commit to a decision for that time point; such commitments are expressed in *decision rules*. In the notation of these rules, we write $\varphi \rightarrow \psi$ instead of $\neg\varphi \vee \psi$.

Definition 5.11 (Decision rule) *A* decision rule *for time point $t \in T$ is an expression of the form*

$$\varphi \;\rightarrow\; d(t) = a, \tag{5.4}$$

where φ is a choice context for t. We refer to φ as the antecedent *of the rule and to $d(t) = a$ as its* consequent.

The above decision rule prescribes to choose action a at time point t given the information conveyed by choice context φ. It is said to be *applicable* in all choice contexts ψ for time point t satisfying $\psi \vdash \varphi$. The number of variables that is referred to in φ is the *complexity* of the rule.

Two decision rules $\varphi_1 \rightarrow d(t)=a_1$ and $\varphi_2 \rightarrow d(t)=a_2$ are said to be *conflicting* when their choice contexts are compatible (i.e. $\varphi_1 \wedge \varphi_2 \not\equiv \bot$), but they do prescribe different action choices (i.e. $a_1 \neq a_2$). Conflicting rules induce a contradiction once we start reasoning with evidence that renders both rules applicable: if ψ is a choice context satisfying $\psi \vdash \varphi_1 \wedge \varphi_2$, then

$$\psi \wedge (\varphi_1 \rightarrow d(t)=a_1) \wedge (\varphi_2 \rightarrow d(t)=a_2) \quad \vdash \quad \bot. \tag{5.5}$$

In Section 5.2, we will discuss observability of system states and the consequences thereof for the possibilities to apply decision rules. If the choice context φ in decision rule $\varphi \rightarrow d(t) = a$ is observable, then it is straightforward for the planning agent to determine the applicability of the rule. Otherwise, the planning agent will be uncertain about the truth of φ, and straightforward application of the rule is not possible. A solution may be found in considering a more general rule that abstracts from unobservable variables; such a rule is called a *generalisation*.

Definition 5.12 (Generalisation) *The decision rule $\varphi_1 \rightarrow d(t) = a$ is called an* generalisation *of the rule $\varphi_2 \rightarrow d(t)=a$ when $\varphi_2 \vdash \varphi_1$.*

Conversely, we say that $\varphi_2 \rightarrow d(t)=a$ is a *specialisation* of $\varphi_1 \rightarrow d(t)=a$ in that case. Because a generalisation is less restrictive in its antecedent, a wider range of cases will render it applicable; the rule is also less complex. When $\varphi_1 \not\equiv \varphi_2$, we will speak of *proper generalisation* and *proper specialisation*, respectively.

5.1.2 Decision processes

Using the language of planning expressions, we now formally define the notion of *decision process*.

Definition 5.13 (Decision process) *Let $\mathcal{B} = (X, A, T)$ be decision basis. A* decision process *for $\mathcal{B}$ is a joint probability distribution P on $X(T) \cup D(T)$.*

A decision process P describes the intertwined reaction over time of the dynamic system to the behaviour of the planning agent and vice versa: it covers the state changes induced by control properties of the actions chosen and the agent's responses to his perceptions of those changes. As such, a decision process P comprises both a description of the planning domain and a decision-making strategy: it implements the meta-level perspective from an external observer. Because we are concerned with action planning under uncertainty, we have chosen to model a decision process as a joint probability distribution on the variables of the planning domain. We will make extensive use of the fact that all probabilistic expressions pertaining to decision processes take arguments from the language $\Phi(T)$.

Definition 5.14 (Possibility) *Let $\mathcal{B} = (X, A, T)$ be decision basis and let P be a decision process for $\mathcal{B}$. We say that the proposition $\varphi \in \Phi(T)$ is* possible *in P when $P(\varphi) > 0$; otherwise, φ is* impossible. *Proposition φ is* inevitable *in P when $P(\varphi) = 1$.*

Impossibility and inevitability of propositions are extreme and usually rare conditions, given the probabilistic nature of the planning problems that we investigate. Note that inevitability of φ renders $\neg\varphi$ impossible and *vice versa*.

When solving a particular problem, we will generally consider a family of related decision processes and attempt to select the process that is expected to provide a most satisfying fulfilment of our objectives. As each decision process implicitly defines a decision-making strategy for the domain in question, this can also be regarded as a procedure for plan selection. We will return to this topic in Section 5.3 on contingency planning. In this section, we first discuss probabilistic expressions that describe predictions in decision processes, and then investigate the relation between symbolic decision rules and decision processes. We will throughout assume that $\mathcal{B} = (X, A, T)$ is a decision basis.

State predictions and system dynamics

We distinguish two types of probabilistic expression that are centrally important in the characterisation of decision processes: *choice* and *state predictions*. A state prediction is a probabilistic expression with respect to the expected system at some time point, given (partial) information on the past history of states and action choices.

Definition 5.15 (State prediction) *Let P be a decision process. An expression of the form $P(\mathrm{c}_{Y(t)} \mid \varphi)$, where $\varphi \in \Phi[0, t-1]$ and $Y(t) \subseteq X(t)$, is called a* state prediction *for time point $t \in T$. If $\varphi \equiv h$ for some history $h \in \mathcal{H}[0, t-1]$, then the expression is a* minimal state prediction. *We say that φ is* sufficient knowledge *for state predictions for time point $t \in T$ if φ is possible in P and*

$$\psi \vdash \varphi \quad \Rightarrow \quad P(\mathrm{C}_{X(t)} \mid \psi) = P(\mathrm{C}_{X(t)} \mid \varphi) \tag{5.6}$$

for all $\psi \in \Phi[0, t-1]$ possible in P. When there is no other $\varphi' \in \Phi[0, t-1]$, $\varphi \vdash \varphi'$, such that φ' is also sufficient knowledge for state predictions for that time point, then φ is called minimally sufficient knowledge.

The term *minimal state prediction* for expressions of the form $P(\mathrm{c}_{Y(t)} \mid h)$ refers to the fact that the planning history h provides a maximum of information on the past and therefore this type of prediction requires less predictive 'effort' than all other types of state predictions. Minimal state predictions have the special property that they provide 'pure' descriptions of the dynamic system under consideration, while non-minimal state predictions also involve the effects of choices that are made by the planning agent.

The intuition behind Equation 5.6 is as follows. If φ is sufficient knowledge for predictions at time point t, then knowing φ is sufficient to calculate the associated probabilities, and no additional information of the history over $[0, t-1]$ would change that probability. If φ is, in addition, minimally sufficient knowledge, then that property does no longer hold once we remove information from φ. Note that each history $h \in \mathcal{H}[0, t-1]$ that is possible in P is trivially sufficient knowledge for state predictions at time point t, but often not minimally sufficient.

The system dynamics are *deterministic* in decision process P when each minimal state prediction with respect to the complete system state is certain (i.e. either 0 or 1).

Definition 5.16 (Deterministic system dynamics) *The system dynamics in decision process P are said to be* deterministic *if for each time point $t \in T$ we have*

$$P(\mathrm{C}_{X(t)} \mid \mathrm{C}_{X[0,t-1] \cup D[0,t-1]}) \quad \in \quad \{0, 1\}. \tag{5.7}$$

Otherwise, the system dynamics are stochastic.

Note that the configuration template $\mathrm{C}_{X[0,t-1] \cup D[0,t-1]}$ in this equation is used as a shorthand for all possible planning histories over time segment $[0, t-1]$. The assumption of deterministic system dynamics is common in the symbolic planning formalisms that have traditionally been studied in AI. Here, we will assume that the behaviour is generally stochastic, although it may be possible to foretell some events with certainty.

In many planning problems it will be final state predictions that are of particular interest. For instance, given an initial configuration $S_1 \in \mathrm{dom}(X)$ of the system, one may be interested in the probability of reaching a particular *goal configuration* $S_2 \in \mathrm{dom}(X)$ at the final time point $t = N$ (i.e. a configuration that is preferred over all others), which is expressed by the state prediction $P(X(N) = S_2 \mid X(0) = S_1)$. Alternatively, one may ask what that probability is if we assume that action sequence $\alpha \in \mathcal{A}(T)$ is followed; the corresponding prediction is $P(X(N) = S_2 \mid X(0) = S_1, \alpha)$. Generalisations to multiple (equally preferred) goal configurations, or planning with less prior information, are straightforward.

We now recapitulate some notions from the theory of random processes within our framework. The first notion, *accessibility*, concerns the possibility to travel between state values.

Definition 5.17 (Accessibility) *Let P be a decision process, let $x \in X$ be a system attribute, and let $s_1, s_2 \in \mathrm{dom}(x)$ be possible values of x. If*

$$P(x(t_2) = s_2 \mid x(t_1) = s_1) \quad > \quad 0 \tag{5.8}$$

for all time points $t_1, t_2 \in T$, $t_2 > t_1$, where $x(t_1) = s_1$ is possible, then we say that value s_2 is accessible *from value s_1 in P.*

An attribute is said to be *progressive* when the accessibility relation among its values is limited to following a total ordering.

Definition 5.18 (Progression) *System attribute $x \in X$ is* progressive *in decision process P when there exist a total ordering $\preceq_x$ on its possible values such that value $s_2 \in \mathrm{dom}(x)$ is accessible from value $s_1 \in \mathrm{dom}(x)$ in P only if $s_1 \preceq_x s_2$.*

Values from which no further evolution is possible are called *absorbing*.

Definition 5.19 (Absorption) *The value $s \in \mathrm{dom}(x)$ is said to be* absorbing *in decision process P if neither other value of attribute x is accessible from it in P.*

Note that this implies that $P(x(t) = s \mid x(t-1) = s) = 1$ whenever $x(t-1) = s$ is possible, for all non-initial time points $t \in T$. We exclude the trivial case of absorption where $x(t) = s$ is inevitable for all $t > 0$.

Proposition 5.20 *A progressive attribute has absorbing values.*

Proof. Let $x \in X$ be a progressive attribute in decision process P, where $\preceq_x$ is the ordering on its possible values. The maximal element with respect to $\preceq_x$ is only accessible from itself. □

Definition 5.21 (Static attribute) *Attribute $x \in X$ is said to be* static *in decision process P when all its values are absorbing in P.*

A static attribute can be regarded as obtaining its value at the initial time point, from which it never departs thereafter. Again, we exclude the trivial case where $x(t) = s$ is inevitable for some $s \in \mathrm{dom}(x)$ and all $t \in T$. In a medical context, the patient's gender, chromosome abnormalities, and chronic diseases would typically be represented by static attributes.

It should be noted that each of the notions accessibility, progression, absorption, and static are defined here in terms of non-minimal state predictions, and therefore partially derive from the planning agent's behaviour. As an example, consider the decision process P where action $a \in A$ is chosen unconditionally and with certainty at all times (i.e. $d(t) = a$ is inevitable for all $t \in T$), and suppose that

$$P(x(t) = \mathit{false} \mid x(t-1) = \mathit{true}, d(t-1) = a) \;=\; 0 \tag{5.9}$$

for some Boolean attribute $x \in X$ and all time points $t \in T$, $t > 0$. Then, the value *false* is not accessible from the value *true*. Moreover, the attribute x is progressive and has absorbing value *true*. Yet, all these properties might not have hold were another decision-making strategy employed in process P. It would therefore be premature to

think of these properties as pertaining to the dynamic system only: they also stem, in part, from the planning agent's behaviour. We will return to this topic in Section 5.2, where the notions are generalised to properties of dynamic systems.

We conclude the discussion of state predictions and system dynamics by remarking that our formalisation of decision-theoretic planning also allows for the reverse type of reasoning. That is, instead of predicting a system state in the future, we can also derive what has probably occurred in the past. For instance, consider once again the goal state $S_2 \in \mathrm{dom}(X)$ in decision process P. By inspecting probabilities of the form $P(X(0) = S_1 \mid X(N) = S_2)$, we can get a feeling for which initial states $S_1 \in \mathrm{dom}(X)$ are likely to lead to the goal S_2, given the decision-making strategy implemented by P.

Choice predictions and planning behaviour

A choice prediction is a probabilistic expression with respect to the expected decision at some time point, given (partial) information on the past history of states and action choices and the contemporaneous system state.

Definition 5.22 (Choice prediction) *Let P be a decision process. An expression of the form $P(\mathrm{c}_{d(t)} \mid \varphi)$, where φ is a choice context for time point $t \in T$, is called a* choice prediction *for that time point. If $\varphi \equiv \mathrm{c}$ for some configuration* c *of the set $X[0,t] \cup D[0,t-1]$, then the expression is a* minimal choice prediction*. We say that φ is* sufficient knowledge *for choice predictions for time point $t \in T$ if φ is possible in P and*

$$\psi \vdash \varphi \;\;\Rightarrow\;\; P(\mathrm{C}_{d(t)} \mid \psi) = P(\mathrm{C}_{d(t)} \mid \varphi) \tag{5.10}$$

for all choice predictions ψ for time point t that are possible in P. When no other choice context φ' for t, $\varphi \vdash \varphi'$, is also sufficient knowledge for choice predictions for that time point, then φ is called minimally sufficient knowledge*.*

Choice predictions essentially encode the behaviour of the planning agent in a decision process. This behaviour is *deterministic* when each minimal choice prediction is certain; otherwise his behaviour is (at least partially) *randomised*.

Definition 5.23 (Deterministic planning behaviour) *A decision process P is said to comprise* deterministic planning behaviour *if for each time point $t \in T$ we have*

$$P(\mathrm{C}_{d(t)} \mid \mathrm{C}_{X[0,t] \cup D[0,t-1]}) \in \{0, 1\}. \tag{5.11}$$

Otherwise, the process comprises randomised planning behaviour*.*

Note that this condition implies that all choice predictions based on sufficient knowledge are certain.

We will generally not consider randomised behaviours as satisfying solutions to planning problems, but similar to incomplete contingency plans, there exists a theoretical interest for them as intermediate representations during problem solving. If both the dynamics of the system and the planning agent's behaviour are deterministic, and the marginal distribution over initial system states is degenerate, then P is also degenerate and it is possible to make all state and choice predictions with certainty. These conditions therefore obviate the usage of probability distributions (or any other representation of uncertainty) in the formalisation. In any other case, there exist uncertain state and choice predictions.

Proposition 5.24 *Let P be a decision process with deterministic planning behaviour. Then, for all time points $t \in T$ and state sequences $\sigma \in \mathcal{S}[0,t]$ that are possible in P there exists a unique action sequence $\alpha \in \mathcal{A}[0,t]$ having $P(\alpha \mid \sigma) = 1$.*

Proof. We will prove this property by induction on T. First, let $t = 0$. Then $P(\alpha \mid \sigma)$ is a minimal choice prediction for each $\alpha \in \mathcal{A}[0,t]$, and there can only be one state sequence $\alpha \in \mathcal{A}[0,t]$ having $P(\alpha \mid \sigma) = 1$, whereas it is zero for all others. Now, suppose that the property holds for time points $0, \ldots, t$, $t < N$. Let $\sigma \in \mathcal{S}[0,t]$ be a state sequence with $P(\sigma) > 0$ and let $\alpha \in \mathcal{A}[0,t]$ be the unique action sequence such that $P(\alpha \mid \sigma) = 1$. Consider a state sequence $\sigma \wedge \mathrm{c}_{X(t+1)} \in \mathcal{S}[0,t+1]$ for which $P(\sigma \wedge \mathrm{c}_{X(t+1)}) > 0$. Again we have that $P(\mathrm{c}_{d(t+1)} \mid \alpha \wedge \sigma \wedge \mathrm{c}_{X(t+1)})$ is a minimal choice prediction for each $\mathrm{c}_{d(t+1)} \in \Omega_{d(t+1)}$, and is therefore deterministic; let $\mathrm{c}_{d(t+1)}$ be the unique action choice for which the prediction is certain. It now follows that

$$\begin{aligned} P(\alpha \wedge \mathrm{c}_{d(t+1)} \mid \sigma \wedge \mathrm{c}_{X(t+1)}) &= P(\mathrm{c}_{d(t+1)} \mid \alpha \wedge \sigma \wedge \mathrm{c}_{X(t+1)}) \cdot P(\alpha \mid \sigma \wedge \mathrm{c}_{X(t+1)}) \\ &= P(\mathrm{c}_{d(t+1)} \mid \alpha \wedge \sigma \wedge \mathrm{c}_{X(t+1)}) \cdot P(\alpha \mid \sigma) \\ &= 1, \end{aligned} \tag{5.12}$$

and $P(\alpha' \wedge \mathrm{c}'_{d(t+1)} \mid \sigma \wedge \mathrm{c}_{X(t+1)}) = 0$ for all other $\alpha' \wedge \mathrm{c}'_{d(t+1)} \in \mathcal{A}[0,t+1]$. □

We will now relate the planning behaviour that is explicitly described by decision rules to the behaviour that is implicitly present in decision processes.

Definition 5.25 (Implementation) *Decision process P is said to* implement *the decision rule $\varphi \to \mathrm{c}_{d(t)}$ when $\varphi \to \mathrm{c}_{d(t)}$ is inevitable in P; the implementation is* strict *if in addition φ is possible.*

Implementation of decision rule $\varphi \to \mathrm{c}_{d(t)}$ by decision process P refers to the situation where the logical implication of action choice $\mathrm{c}_{d(t)}$ in choice context φ is acknowledged by the fact that $P(\varphi \to \mathrm{c}_{d(t)}) = 1$. The next proposition serves to sharpen the intuitions for this concept a bit further.

Proposition 5.26 *Let P be a decision process. The following statements are equivalent:*

1. *P implements the decision rule $\varphi \rightarrow \mathrm{c}_{d(t)}$;*
2. *$P(\varphi \wedge \mathrm{c}_{d(t)}) = P(\varphi)$; and*
3. *$P(\varphi \wedge \neg \mathrm{c}_{d(t)}) = 0$.*

Proof. Using the marginalisation property $P(\mathrm{c}_{d(t)}) = P(\varphi \wedge \mathrm{c}_{d(t)}) + P(\neg\varphi \wedge \mathrm{c}_{d(t)})$, we find that the first two statements are equivalent as

$$\begin{aligned} P(\varphi \rightarrow \mathrm{c}_{d(t)}) &= P(\neg\varphi \vee \mathrm{c}_{d(t)}) \\ &= P(\neg\varphi) + P(\mathrm{c}_{d(t)}) - P(\neg\varphi \wedge \mathrm{c}_{d(t)}) \\ &= P(\neg\varphi) + P(\varphi \wedge \mathrm{c}_{d(t)}) \\ &= 1 - P(\varphi) + P(\varphi \wedge \mathrm{c}_{d(t)}), \end{aligned} \tag{5.13}$$

and therefore

$$\begin{aligned} P(\varphi \rightarrow \mathrm{c}_{d(t)}) = 1 &\Leftrightarrow P(\varphi \wedge \mathrm{c}_{d(t)}) - P(\varphi) = 0 \\ &\Leftrightarrow P(\varphi \wedge \mathrm{c}_{d(t)}) = P(\varphi). \end{aligned} \tag{5.14}$$

Equivalence of the first and third statements follows from the fact that

$$\begin{aligned} \neg(\varphi \rightarrow \mathrm{c}_{d(t)}) &\equiv \neg(\neg\varphi \vee \mathrm{c}_{d(t)}) \\ &\equiv \varphi \wedge \neg \mathrm{c}_{d(t)}, \end{aligned}$$

and so

$$P(\varphi \wedge \neg \mathrm{c}_{d(t)}) = 1 - P(\varphi \rightarrow \mathrm{c}_{d(t)}).$$

□

As appears from the proof, decision rule implementation is a trivial matter when φ is impossible in P, because then $P(\varphi \wedge \psi) = 0$ for any proposition $\psi \in \Phi(T)$, and any decision rule having φ as its antecedent is thus implemented by P. For this reason, it is often useful to require strict implementation, where $P(\varphi) > 0$, and so Equation 5.14 can be rewritten to

$$P(\mathrm{c}_{d(t)} \mid \varphi) = 1. \tag{5.15}$$

Proposition 5.27 *If decision process P implements decision rule $\varphi \rightarrow d(t) = a$, then it also implements each specialisation of that rule.*

Proof. Let $\psi \rightarrow d(t) = a$ be a specialisation of $\varphi \rightarrow d(t) = a$, i.e. ψ is a choice context for time point t such that $\psi \vdash \varphi$. If φ is impossible in P, then so is ψ,

and $P(\psi \wedge d(t){=}a) = 0$; hence, $\psi \rightarrow d(t){=}a$ is trivially implemented by P. If φ is possible in P, then

$$\begin{aligned}
P(\psi) &= P(\psi \wedge \varphi) && \text{(as } \psi \vdash \varphi\text{)} \\
&= \sum_{\mathrm{c}_{d(t)} \in \Omega_{d(t)}} P(\psi \wedge \varphi \wedge \mathrm{c}_{d(t)}) && \text{(marginalisation)} \\
&= \sum_{\mathrm{c}_{d(t)} \in \Omega_{d(t)}} P(\psi \mid \varphi \wedge \mathrm{c}_{d(t)}) \cdot P(\mathrm{c}_{d(t)} \mid \varphi) \cdot P(\varphi) && \text{(chain rule)} \\
&= P(\psi \mid \varphi \wedge d(t){=}a) \cdot P(\varphi) && \text{(Equation 5.15)} \\
&= \frac{P(\psi \wedge \varphi \wedge d(t){=}a)}{P(\varphi \wedge d(t){=}a)} \cdot P(\varphi) && \text{(conditional prob.)} \\
&= P(\psi \wedge \varphi \wedge d(t){=}a) && \text{(Proposition 5.26)} \\
&= P(\psi \wedge d(t){=}a), && \text{(as } \psi \equiv \psi \wedge \varphi\text{)}
\end{aligned}$$

so the rule $\psi \rightarrow d(t){=}a$ is then also implemented by P. □

Note that P may strictly implement the rule $\varphi \rightarrow d(t){=}a$, but that the implementation of some (or all) of its specialisations may be non-strict.

Corollary 5.28 *If decision rule $\varphi \rightarrow d(t){=}a$ is strictly implemented by decision process P, then its antecedent φ is sufficient knowledge for choice predictions at time point t in P.*

Proof. This follows immediately from the fact that all specialisations of the rule are also implemented by P. □

It is also easily seen that if decision process P implements the decision rule $\varphi \rightarrow d(t){=}a$ and φ is minimally sufficient knowledge for choice predictions at time point t (implying that the implementation is strict), then there is no proper generalisation of the rule that is also implemented by P.

5.2 Control and observation

We now turn to models that describe the effects of action choices on the evolution of the dynamic system under consideration, and to models that describe their effects on the planning agent's knowledge of that system. The former type of model is called a *control model*; the latter type is called an *observation model*. Throughout we assume $\mathcal{B} = (X, A, T)$ to be a decision basis.

5.2.1 Models of control

A *model of control* describes the stochastic behaviour of the dynamic system over time under different action regimes.

Definition 5.29 (Control model) *A* control model *for decision basis $\mathcal{B}$ is a set $\Xi = \{\xi_\alpha \mid \alpha \in \mathcal{A}(T)\}$ of probability distributions on $X(T)$, where for all time points $t \in T$ and all actions sequences $\alpha_1, \alpha_2 \in \mathcal{A}(T)$ we have*

$$\xi_{\alpha_1}(\mathrm{C}_{X[0,t]}) \;=\; \xi_{\alpha_2}(\mathrm{C}_{X[0,t]}) \tag{5.16}$$

whenever α_1 and α_2 are identical over time segment $[0, t-1]$.

The effects of a given action sequence $\alpha \in \mathcal{A}(T)$ on the dynamic system is expressed by the probability distribution $\xi_\alpha \in \Xi$ on $X(T)$; we refer to ξ_α as the *control distribution* for sequence α. As control distributions describe how the dynamic system will react to choosing action sequences, we can infer how the planning agent exerts control over that system by comparing different control distributions. Because we have assumed that

$$X(0), d(0), X(1), d(1), \ldots, X(N), d(N) \tag{5.17}$$

is the alternating sequence events and choices in a decision process, it is not possible for an action choice at time point $t \in T$ or thereafter to influence the uncertain events at time points $0, \ldots, t$: these events have then already been realised as the past sequence of system states. The resulting restriction on models of control, expressed in Equation 5.16, is therefore that the marginal distributions on $X[0,t]$ induced by action sequences α_1 and α_2 are the same when $\alpha_1(i) = \alpha_2(i)$ for all $i = 0, \ldots, t-1$. Note that the action choice at time point $t = N$ will not influence any system state.

A result from this property, that actions in the present and future cannot change the system in the past, is that each partial action sequence over some segment $[0, t-1]$, $t \in T$, induces a unique probability distribution over $X[0,t]$: no knowledge of further actions would help to predict the system's states so far. As it is often convenient to be able to refer to such distributions directly, we introduce a separate notation for them.

Notation 5.30 *Let Ξ be a control model for decision basis $\mathcal{B}$, and let $t \in T$ be a point in time. Given an action sequence $\alpha \in \mathcal{A}[0, t-1]$, we use ξ^t_α to denote the unique probability distribution on $X[0,t]$ induced by Ξ.*

That is,

$$\xi^t_\alpha(\mathrm{C}_{X[0,t]}) \;=\; \xi_{\alpha'}(\mathrm{C}_{X[0,t]}) \tag{5.18}$$

for all $\alpha' \in \mathcal{A}(T)$ that extend α, i.e. $\alpha' \vdash \alpha$. A special case is found at the initial time point $t = 0$; here, the only possible action sequence is the empty action sequence $\breve{\alpha}$. As such, $\xi^0_{\breve{\alpha}}$ specifies the unique distribution on initial system states, which withdraws from control by the planning agent.

It should be noted that the restriction from Equation 5.16 does not hold for decision processes. Decision processes implement the perspective of an external observant that regards system and agent involved in their interaction. For the observant, seeing that the planning agent decides to choose $d(t) = a$, $a \in A$, provides indirect information on earlier system states if he is aware of the planning agent's decision-making strategy. For instance, if the planning agent is known to choose $d(t) = a$ only when $x(t-1) = s$, $x \in X$, $s \in \mathrm{dom}(x)$, then observing decision $d(t) = a$ will make the external observant infer that $x(t-1) = s$ holds. This will even apply if all action choices over time segment $[0, t-1]$ were already known. So, while action choices cannot change the past, they can change an external observant's *beliefs* with respect to the past. The effects of the agent's choices can furthermore be perfectly in line with the specifications of a control model; we will shortly return to this issue.

We will assume the general form of Definition 5.29 for control models, although it is not practical for concrete applications of the framework. Its complexity stems from the fact that it provides an extensional description of the effects of all possible actions on all possible state variables over time, and that no attempt is made to avoid such descriptions for state variables that essentially remain unaffected. A more manageable control specification would typically use smaller descriptions, building on one or more *frame axioms*, assumptions of unaffectedness for variables not mentioned in the description (McCarthy and Hayes, 1969). We return to this subject in due course.

A control model Ξ is said to be *strictly positive* when

$$\xi_\alpha(\mathrm{C}_{X(T)}) \;>\; 0 \tag{5.19}$$

for all control distributions $\xi_\alpha \in \Xi$. In words, this means that "everything is always possible" with respect to the dynamic system. In a mathematical sense, this is a convenient condition, but it will often be regarded as unnatural in real-world settings; we will only assume it when explicitly stated so.

Control and decision processes

Given a control model Ξ, we say that decision process P is *compliant with* (or *complies with*) that model when its probabilities are fully consistent with the ones provided by the model.

Definition 5.31 (Control compliance) *Decision process P is* compliant with *control model Ξ when*

$$P(\mathrm{C}_{X(T)} \mid \alpha) \;=\; \xi_\alpha(\mathrm{C}_{X(T)}) \tag{5.20}$$

for each action sequence $\alpha \in \mathcal{A}[0, t-1]$ that is possible in P.

The control model can then be regarded as providing one 'half' of the decision process, the behaviour of the dynamic system.

Proposition 5.32 *A control model determines all minimal state predictions in a decision process.*

Proof. Let Ξ be a control model for decision basis $\mathcal{B} = (X, A, T)$, and let P be a compliant decision processes. Furthermore, let $t \in T$, $h \in \mathcal{H}[0, t-1]$, and $\mathrm{c}_{X(t)} \in \Omega_{X(t)}$. If $t = 0$, then h is empty and

$$P(\mathrm{c}_{X(t)} \mid h) \;=\; \xi^0_{\hat{\alpha}}(\mathrm{c}_{X(t)}). \tag{5.21}$$

If $t > 0$, we write $h \equiv \sigma \wedge \alpha$, and

$$\begin{aligned} P(\mathrm{c}_{X(t)} \mid h) &= \frac{P(\mathrm{c}_{X(t)} \wedge \sigma \wedge \alpha)}{P(\sigma \wedge \alpha)} \\ &= \frac{P(\mathrm{c}_{X(t)} \wedge \sigma \mid \alpha) \cdot P(\alpha)}{P(\sigma \mid \alpha) \cdot P(\alpha)} \\ &= \frac{P(\mathrm{c}_{X(t)} \wedge \sigma \mid \alpha)}{P(\sigma \mid \alpha)} \\ &= \frac{\xi^t_\alpha(\sigma \wedge \mathrm{c}_{X(t)})}{\xi^{t-1}_{\alpha'}(\sigma)} \end{aligned} \tag{5.22}$$

where $\alpha' \in \mathcal{A}[0, t-2]$ is the action sequence obtained by removing $\alpha(t-1)$ from α. The equality $P(\sigma \mid \alpha) = \xi^{t-1}_{\alpha'}(\sigma)$ follows from the requirement in Definition 5.29 that the action choice at time point $t-1$ cannot influence the states at preceding time points. □

When two decision processes both comply with the same control model, then they will have the same minimal state predictions and we say they are *control-equivalent*. Note that not necessarily *all* state predictions are the same in these decision processes: non-minimal predictions will generally differ as they also depend on the implemented behaviour of the planning agent.

Notation 5.33 *Given a control model Ξ, we will use*

- *$\mathbb{P}_\Xi$ to denote the set of all decision processes that comply with it, and*
- *$\mathbb{P}_\Xi^{\mathrm{det}}$ to denote its proper subset of decision processes with deterministic planning behaviour.*

Note that the set $\mathbb{P}_\Xi$ is uncountable as it comprises all decision process with randomised planning behaviour; the set $\mathbb{P}_\Xi^{\mathrm{det}}$ is finite.

System dynamics revisited

Given the relation between models of control and decision processes, we can now reconsider the notions of accessibility and progression, absorption, and static attribute (Definitions 5.17, 5.18, 5.19, and 5.21, respectively). These notions have been defined above for decision processes, but allow for straightforward generalisations to models of control.

Given a model of control Ξ and a system attribute $x \in X$, we will say that value $s_2 \in \mathrm{dom}(x)$ is *accessible* from value $s_1 \in \mathrm{dom}(x)$ under Ξ when there exists an action sequence $\alpha \in \mathcal{A}(T)$ such that

$$\xi_\alpha(x(t_2) = s_2 \mid x(t_1) = s_1) \;>\; 0 \tag{5.23}$$

for all time points $t_1, t_2 \in T$, $t_2 > t_1$, where $\xi_\alpha(x(t_1) = s_1) > 0$. This is equivalent to the statement that there exists a compliant decision process P in which s_2 is accessible from s_1, as is expressed by the next proposition.

Proposition 5.34 *Let Ξ be a control model, let $x \in X$ be a system attribute, and let $s_1, s_2 \in \mathrm{dom}(x)$ be possible value of x. Then, s_2 is accessible from value s_1 under Ξ if and only if there exists a decision process $P \in \mathbb{P}_\Xi$ in which s_2 is accessible from value s_1.*

Proof. ($\Rightarrow$) Let $\alpha \in \mathcal{A}(T)$ be an action sequence for which

$$\xi_\alpha(x(t_2) = s_2 \mid x(t_1) = s_1) \;>\; 0 \tag{5.24}$$

for all time points $t_1, t_2 \in T$, $t_2 > t_1$, where $\xi_\alpha(x(t_1) = s_1) > 0$, and let P be a decision process that complies with Ξ and in which α is inevitable, i.e. $P(\alpha) = 1$. Then, $P \in \mathbb{P}_\Xi$ and s_2 is accessible from s_1 in P.
($\Leftarrow$) Let $P \in \mathbb{P}_\Xi$ be a decision process such that

$$P(x(t_2) = s_2 \mid x(t_1) = s_1) \;>\; 0 \tag{5.25}$$

for all time points $t_1, t_2 \in T$, $t_2 > t_1$, where $x(t_1) = s_1$ is possible. Then, there must exist an action sequence $\alpha \in \mathcal{A}(T)$ that is possible in P and

$$P(x(t_2) = s_2 \mid x(t_1) = s_1, \alpha) \;>\; 0 \tag{5.26}$$

for t_1, t_2 as above, and as P complies with Ξ, we then also find that

$$\xi_\alpha(x(t_2) = s_2 \mid x(t_1) = s_1) \;>\; 0. \tag{5.27}$$

Hence, s_2 is accessible from s_1 under Ξ. □

The notion of *progression* is now generalised to models of control by taking it to be applicable to attribute x when there exists a total ordering $\preceq_x$ on its possible values such that value s_2 is accessible from value s_1 under model Ξ only if $s_1 \preceq_x s_2$.

Corollary 5.35 *The attribute $x \in X$ is progressive under control model Ξ if and only if it is progressive in each decision process $P \in \mathbb{P}_\Xi$.*

Proof. Directly from Proposition 5.34 and the generalised notion of progression. □

Note that attribute x may be progressive in some compliant decision processes although it is generally not progressive under control model Ξ; this is then due to the behaviour of the planning agent in these processes. It does imply that the model Ξ allows for this type of control, and that the system dynamics are to some extent deterministic.

We say that the value $s \in \mathrm{dom}(x)$ *absorbing* under control model Ξ if neither other value of attribute x is accessible from it under Ξ. As with decision processes, we exclude the trivial case of absorption where $\xi_\alpha(x(t) = s) = 1$ for all $\xi_\alpha \in \Xi$ and all $t \in T$.

Corollary 5.36 *The value $s \in \mathrm{dom}(x)$ is absorbing under control model Ξ if and only if it is absorbing in each decision process $P \in \mathbb{P}_\Xi$.*

Proof. Directly from Proposition 5.34 and the generalisation of absorption. □

Again there may be compliant decision processes in which s is absorbing while this does not hold in general under the given control model. In such decision processes, the planning agent is apparently holding x in value s once this value has been reached. Such a decision-making strategy may be employed when s is strongly preferred over the other values of x.

The notion of *static attribute* finally, is generalised by taking it to be applicable to x under control model Ξ when all its values are absorbing under Ξ. Here we again exclude the trivial case where $\xi_\alpha(x(t) = s) = 1$ for some $s \in \mathrm{dom}(X)$ and all $\xi_\alpha \in \Xi$, $t \in T$.

Corollary 5.37 *Attribute $x \in X$ is static under control model Ξ if and only if it is static in each decision process $P \in \mathbb{P}_\Xi$.*

Proof. Directly from Proposition 5.34 and the generalised notion of static. □

Unresponsiveness and influence

We now turn to control relations at the level of individual actions and state variables. These relations concern the question whether action choices can influence a given state variable, and if so, the extend to which this is possible. The basis for investigating these relations is provided by the notion of *unresponsiveness*, which is due to (Heckerman and Shachter, 1995).

Definition 5.38 (Unresponsiveness) *Let Ξ be a model of control, let $t \in T$ be a time point, and let $Y, Z \subseteq X[0,t]$. We say that the set Y is* conditionally unresponsive *to a choice between action sequences $\alpha_1, \alpha_2 \in \mathcal{A}[0, t-1]$ given Z when*

$$\xi^t_{\alpha_1}(\mathrm{C}_Y \mid \mathrm{C}_Z) \;=\; \xi^t_{\alpha_2}(\mathrm{C}_Y \mid \mathrm{C}_Z). \tag{5.28}$$

When $Z = \varnothing$, we say that Y unconditionally unresponsive *to that choice.*

In words, this means that knowing which of these action sequences was chosen does not influence our beliefs with respect the set Y once we know the configuration of Z. The condition can also be formulated as

$$P(\mathrm{C}_Y \mid \mathrm{C}_Z,\, \alpha_1) \;=\; P(\mathrm{C}_Y \mid \mathrm{C}_Z,\, \alpha_2) \tag{5.29}$$

for all decision processes $P \in \mathbb{P}_\Xi$ where both α_1 and α_2 are possible. It is trivially true when $Y \subseteq Z$; in all other cases, it provides us with information regarding the influence of action choices on state variables. For instance, if the set Y is unconditionally unresponsive to choosing between any pair action sequences, then Y is apparently completely beyond the decision maker's control. In the special case where $Y = x(T)$ for some system attribute $x \in X$, we will say that this attribute is *uncontrollable*.

Definition 5.39 (Controllability) *We say that system attribute $x \in X$ is* uncontrollable *when $x(T)$ is unconditionally unresponsive to a choice between any pair of action sequences from $\mathcal{A}(T)$.*

Unresponsiveness is a rather course-grained notion as it considers the relation between complete action sequences and sets of state variables. One will often be interested in control relations between individual action choices and state variables. A first step towards describing these relations is made by the notion of *influence*, where the decisional side of the relation is limited to a single action choice.

Definition 5.40 (Influence) *Let Ξ be a model of control, let $t_1, t_2 \in T$ be time points where $t_1 < t_2$, and let $Y, Z \subseteq X[0,t_2]$. We say that the set Y is* conditionally uninfluenced *by the action choice at time point t_1 given Z, when Y is conditionally unresponsive given Z to a choice between any pair $\alpha_1, \alpha_2 \in \mathcal{A}[0, t_2 - 1]$ of action sequences satisfying $\alpha_1(t) = \alpha_2(t)$ for all $t = 0, \ldots, t_1 - 1, t_1 + 1, \ldots, t_2 - 1$.*

In words, this means that knowing the action choice at time point t_1 does not influence our beliefs with respect the set Y once we know the configuration of Z. The condition can also be formulated as

$$P(\mathrm{C}_Y \mid \mathrm{C}_Z,\, d(t_1) = a_1) \;=\; P(\mathrm{C}_Y \mid \mathrm{C}_Z,\, d(t_1) = a_2) \tag{5.30}$$

for all decision processes $P \in \mathbb{P}_\Xi$ and all actions $a_1, a_2 \in A$. Again, it is trivially true when $Y \subseteq Z$, but highly informative in other cases. Statements of the form of

Equation 5.30 are sometimes regarded as *frame axioms*, as they identify aspects of the system that are not changed by certain decisions under given conditions, (Darwiche and Pearl, 1994). Examples are the Markov assumption in Markov decision processes and the conditional independency statements depicted by the graph of an influence diagram; we will return to this topic in Section 5.5.

Given a particular decision $d(t_1) \in D(T)$, we now identify three types of state variables: those that are never influenced by the decision, those that may be uninfluenced if we have enough information, and those that remain to be influenced even if we know all other variables.

Definition 5.41 (Affectedness) *Let Ξ be a model of control, and let $t_1, t_2 \in T$ be time points.*

- *State variable $x(t_2)$ is* unaffected *by decision $d(t_1)$ when $\{x(t_2)\}$ is unconditionally uninfluenced by that choice.*
- *State variable $x(t_2)$ is* indirectly affected *by decision $d(t_1)$ when $x(t_2)$ is not unaffected but $\{x(t_2)\}$ is conditionally uninfluenced by that choice given the set $X[0, t_2] \setminus \{x(t_2)\}$.*
- *Otherwise, $x(t_2)$ is* directly affected *by the choice.*

Note that $x(t_2)$ is definitely unaffected by the action choice at time point t_1 when $t_1 \geq t_2$, as we have required in the definition of control models that the past cannot be changed. The same holds when attribute x is uncontrollable, as is stated by the next proposition.

Proposition 5.42 *If system attribute $x \in X$ is uncontrollable, then for each time point $t \in T$, state variable $x(t)$ is unaffected by all decisions.*

Proof. If there were some $x(t) \in X(T)$ that was not unaffected by all decisions, then $\{x(t)\}$ would be responsive to a choice of action sequence, contradicting the definition of uncontrollability. □

If state variable $x(t_2)$ is affected by the action choice at time point t_1, this may proceed in a direct and in an indirect fashion. In the latter case, the effect of the decision on $x(t_2)$ can be regarded as a *ramification* of the effect on one or more directly affected state variables, (Dean and Wellman, 1991). The variable $x(t_2)$ must therefore be related to one or more of these directly affected state variables.

Proposition 5.43 *Let Ξ be a model of control, and let $t_1, t_2 \in T$ be time points. If variable $x(t_2)$ is indirectly affected by the decision at time point t_1, then there exists*

a state variable $x'(t_3) \in X[0, t_2] \setminus \{x(t_2)\}$ *that is directly affected by the decision, and a decision process* $P \in \mathbb{P}_\Xi$ *where*

$$\{x(t_2)\} \not\perp\!\!\!\perp_P \{x'(t_3)\} \mid \{d(t_1)\}. \tag{5.31}$$

Proof. Suppose that

$$\{x(t_2)\} \perp\!\!\!\perp_P \{x'(t_3)\} \mid \{d(t_1)\} \tag{5.32}$$

for all decision processes $P \in \mathbb{P}_\Xi$ and all state variables $x'(t_3) \in X[0, t_2] \setminus \{x(t_2)\}$. Then also

$$\{x(t_2)\} \perp\!\!\!\perp_P X[0, t_2] \setminus \{x(t_2)\} \mid \{d(t_1)\} \tag{5.33}$$

for all decision processes $P \in \mathbb{P}_\Xi$. Yet, as $x(t_2)$ is indirectly affected by the decision, we have

$$\{x(t_2)\} \perp\!\!\!\perp_P \{d(t_1)\} \mid X[0, t_2] \setminus \{x(t_2)\} \tag{5.34}$$

for all decision processes $P \in \mathbb{P}_\Xi$. Therefore, variable $x(t_2)$ must be unconditionally independent of decision $d(t_1)$. This contradicts the assumption that $x(t_2)$ is indirectly affected, and therefore not unaffected, by the action choice at time point t_1. □

We conclude this section with noting that direct and indirect affectedness are properties of control models that cannot necessarily be retraced in every compliant decision process. For instance, in some decision processes all system attributes may occupy an absorbing state at the very beginning, preventing any change and therefore any influence from the planning agent. Furthermore, knowing that such relations hold in a given model of control is, in general, insufficient for drawing conclusions with respect to state dynamics: the underlying attributes of affected and unaffected variables may exhibit all types of behaviour over time, ranging from complete inertia to permanent change. Another point to note (again) is that decision processes encode responses of the system to the decision-making behaviour of the planning agent and the agent's responses to observed system states. As a result, knowing that the planning agent chose action $a \in A$ at time point $t \in T$ will not only influence our beliefs with respect to present and future state variables that are known to be affected by that decision, but also with respect to *past* state variables that are known to have influenced the planning agent's decision.

5.2.2 Models of observation

We will now formulate the notion of *observability* within our framework. As in the POMDP formalism, it is assumed that action choices may yield the opportunity to inspect part of the system state. Whereas in POMDPs such opportunities immediately follow the action choices, our definition is more loose and allows for more types of observability.

Definition 5.44 (Observation model) *Let $\mathcal{B} = (X, A, T)$ be a decision basis. An* observation model *for $\mathcal{B}$ is a set $O = \{o_t \mid t \in T\}$ of functions*

$$o_t : \mathcal{A}[0, t-1] \to \wp(X[0, t]). \tag{5.35}$$

We refer to $o_t \in O$ as the *observation function* for time point $t \in T$. The set $o_t(\alpha) \subseteq X[0, t]$ of state variables can be observed by the planning agent at that time point, given that he performed the action choices in sequence $\alpha \in \mathcal{A}[0, t-1]$. A special case is the observation function for the initial time point $t = 0$, which is essentially typed

$$o_0 : \{\breve{\alpha}\} \to \wp(X(0)). \tag{5.36}$$

This function provides the subset of observable variables from $X(0)$ before any decision has been made; following the terminology of Chapter 3, we will refer to $o_0(\breve{\alpha})$ as the set of *case parameters*. Furthermore, note that no observation function takes the last decision (at time point $t = N$) into account; this decision is therefore not only uninfluential with respect to the system's evolution, but also with respect to observability.

Assumption 5.45 (Monotonicity of observation) *We take each observation model O to be* monotonic*, i.e.*

$$o_t(\alpha) \quad \subseteq \quad o_{t+1}(\alpha \wedge d(t) = a) \tag{5.37}$$

for all time points $t \in T$, $t < N$, action sequences $\alpha \in \mathcal{A}[0, t-1]$, and each action $a \in A$.

Monotonicity of observation means that observability persists through time: $o_t(\alpha)$ remains observable at future time points under any additional action choices. This is comparable to the 'no-forgetting' property (Definition 4.11 on page 89) in influence diagrams.

The typical kind of planning behaviour in our framework can now be described as follows. At time point $t = 0$, the planning agent observes the state of the set of case parameters $Y_0 = o_0(\breve{\alpha})$. Let $\psi_0 \in \Omega_{Y_0}$ denote the observed configuration of Y_0; we refer to ψ_0 as *evidence*. We can also regard ψ_0 as a choice context for time point $t = 0$; it is indeed the choice context that is used by the planning agent to select his first action, say $a_0 \in A$. The action choice extends his *knowledge* of the decision process to $\psi_0 \wedge d(0) = a_0$. In addition, he is now able to observe the state of the variables in $Y_1 = o_1(d(0) = a_0) \setminus Y_0$ (the set Y_0 remains observable because of the assumption of monotonicity, but it does not add new information). Let ψ_1 be the observed configuration of Y_1. The joint evidence now available is $\psi_0 \wedge \psi_1$; the planning agent's knowledge of the decision process equals $\psi_0 \wedge d(0) = a_0 \wedge \psi_1$. It is this knowledge, a choice context for time point $t = 1$, that is used to make the next

decision. This procedure continues to the final decision moment, each time extending the planning agent's information on past and present.

We do note that it is well possible that the planning agent does not always use all his knowledge in choosing his actions. He may be convinced that some evidence, or some earlier decisions, have become irrelevant for the planning of future actions in the light of other knowledge. Of course, there may also be practical reasons for such ignorance in the planning process: as time proceeds, the wealth of available information may exceed the computational capacity of the agent. In either case, the complexity of decision rules used by the planning agent is smaller than what is potentially possible, and sufficient knowledge for predictions regarding the agent's action choices is easier obtained than one might theoretically expect.

Manifest and hidden variables

From the coarse grained definition of observability above we now derive a more delicate relation between individual decisions and state variables. We say that a state variable is *manifest* (at a given time point) under some decision when that decision leads to observability of that state variable.

Definition 5.46 (Manifestness) *Let $O = \{o_t \mid t \in T\}$ be an observation model for decision basis $\mathcal{B} = (X, A, T)$, let $t_1, t_2, t_3 \in T$ be arbitrary time points, where $t_1 < t_3$.*

- *State variable $x(t_2) \in X(t_2)$ is called* manifest *at time point t_3 under decision $d(t_1) = a$ when $x(t_2) \in o_{t_3}(\alpha)$ for each action sequence $\alpha \in \mathcal{A}[0, t_3 - 1]$ having $\alpha(t_1) = a$.*
- *State variable $x(t_2)$ is called* potentially manifest *at time point t_3 when there exists an action sequence $\alpha \in \mathcal{A}[0, t_3 - 1]$ such that $x(t_2) \in o_{t_3}(\alpha)$.*
- *State variable $x(t_2)$ is called* absolutely manifest *at t_3 when $x(t_2) \in o_{t_3}(\alpha)$ for all action sequences $\alpha \in \mathcal{A}[0, t_3 - 1]$.*

Note that any form of manifestness implicitly requires that $t_2 \leq t_3$ ("we cannot glance into the future"). Potential manifestness of state variable $x(t_2)$ means that it is possible to choose an action sequence such that the variable is observable at time point t_3. State variables that lack this property withdraw from observation in all scenarios and are said to be *hidden* at time point t_3. Absolute manifestness of state variable $x(t_2)$ means that one can always observe its value at time point t_3, irrespective of the actions chosen. This holds, for instance, always for the set of case parameters.

Proposition 5.47 *Let O be an observation model. All variables from the set $o_0(\check{\alpha})$, the case parameters under O, are absolutely manifest at all time points.*

Proof. Directly from the assumption of observational monotonicity. □

Potential and absolute manifestness actually coincide at the initial time point, as observability withdraws from the planning agent's influence there.

Example 5.48 *In a clinical setting, potentially manifest variables typically describe signs and symptoms of disease and outcomes of diagnostic tests. It depends on the formalisation of the planning problem that is faced by the treating clinician whether such variables are also absolutely manifest. If each decision moment denotes a point in time where the patient is seen by the clinician, then all symptoms are absolutely manifest at all times. In contrast, when part of the planning problem is to decide when the patient should visit the clinic, then these variables are manifest only upon such visits.*

There are some further subtleties to be noted when state variable $x(t_2)$ is manifest at time point t_3 under decision $d(t_1) = a$; these subtleties are related to time. For clarity, we assume that decision $d(t_1) = a$ is not only a sufficient but also a *necessary* condition for the manifestness of $x(t_2)$ at time point t_3 just now, i.e. $x(t_2) \in o_{t_3}(\alpha)$ only if $\alpha(t_1) = a$. When $t_2 < t_1$, this means that the decision allows to inspect $x(t_2)$ with hindsight, where it was previously unobservable; we can think of seeing the light of a star that has actually ceased to exist already. When $t_1 = t_2$, observing $x(t_2)$ coincides with the moment of action; this is what we find in POMDPs. When $t_2 > t_1$ finally, there is a time lag between the decision and the actual moment of observation. In a medical setting, this occurs for instance with the decision to apply cell cultures for inspecting body tissue.

Below, we will first investigate the role of observations in action planning, and then the relation between observability and decision processes.

Verifiable choice contexts and operational rules

A most important aspect of observability relates to contingency planning. An observation model specifies which information on the dynamic system potentially belongs to the planning agent's knowledge: such information can therefore be used in making decisions.

Definition 5.49 (Verifiability) *Let O be an observation model and let φ be a choice context for time point $t_3 \in T$. We say that φ is* verifiable *under O when each state variable $x(t_2) \in X[0, t_3]$ covered by φ is manifest at t_3 under some decision $\mathrm{c}_{d(t_1)}$ where $\varphi \vdash \mathrm{c}_{d(t_1)}$.*

Note that again we may have either $t_1 < t_2$, $t_1 = t_2$ or $t_1 > t_2$. The empty choice context $\top$ is trivially verifiable under each observation model. From all other verifiable choice contexts we can obtain (strictly) smaller choice contexts that are also verifiable

by removing evidence, or, when the choice context does not contain evidence, decisions. The converse is not true: we cannot always extend a verifiable choice context to obtain a great choice context that is also verifiable. In those cases, the context in question represents maximal information for the next decision from the viewpoint of the planning agent. These choice context have the form $\alpha \wedge \mathrm{c}_Z$, where $\alpha \in \mathcal{A}[0, t-1]$ and $\mathrm{c}_Z \in \Omega_Z$, $Z = o_t(\alpha)$.

If choice context φ in decision rule $\varphi \rightarrow d(t) = a$ is verifiable under a given observation model O, then it is straightforward for the planning agent to determine the applicability of the rule; we say that the rule is *operational* under O.

Definition 5.50 (Operationality) *Let O be an observation model. A decision rule δ is* operational *under O when that model renders the antecedent of δ verifiable; otherwise, the rule is* non-operational *under O.*

From the above observation that each verifiable choice context subsumes smaller choice contexts that are also verifiable, we can conclude that each operational decision rule has operational generalisations.

Observability and decision processes

We have seen in Subsection 5.1.2 that decision processes implicitly define the planning agent's behaviour. We will now examine the relation between observability and this implicit behaviour; more specifically, we will examine the extent to which this behaviour is truly based on the planning agent's knowledge under a given observation model.

Recall that the notion of *sufficient knowledge* from Definition 5.22 provides a means to discover the information that is used in planning decisions in a given decision process P. If choice context φ is sufficient knowledge for choice predictions at time point $t \in T$, then adding more information to φ will not help to improve those predictions. For instance, if the planning behaviour is deterministic, then knowing φ is sufficient to determine the action choice at time point t with certainty, but this does not hold for choice contexts with less information. If, in addition, φ is minimally sufficient knowledge, then we cannot leave out any of the information in φ without introducing more uncertainty in the predictions. Therefore, if φ is minimally sufficient knowledge but not verifiable under a given observation model O, this behaviour could not be implemented by the planning agent.

This brings us to the following definition.

Definition 5.51 (Observational compliance) *We say that decision process P* is compliant with *observation model O, if a choice context φ is verifiable under O whenever it is minimally sufficient knowledge for choice predictions at some time point.*

Notation 5.52 *Let Ξ and O be control and observation models, respectively. We use*

- $\mathbb{P}_O$ *to denote that the set of all decision processes that are compliant with O,*
- $\mathbb{P}_O^{\det}$ *to denote that the set of all decision processes with deterministic planning behaviour that are compliant with O, and*
- $\mathbb{P}_{\Xi,O}^{\det} = \mathbb{P}_{\Xi}^{\det} \cap \mathbb{P}_O^{\det}$ *to denote the set of all decision processes with deterministic planning behaviour that are compliant with both Ξ and O.*

Proposition 5.53 *Let O be an observation model and let $P \in \mathbb{P}_O$. If P strictly implements the decision rule δ and none of its proper generalisations, then δ is operational under O.*

Proof. Assume that $\delta \equiv \varphi \rightarrow \mathrm{c}_{d(t)}$. As the implementation of δ is strict, we know that $P(\varphi) > 0$. From Corollary 5.28 on page 129, we also know that φ is sufficient knowledge for choice predictions at time point t. Now suppose that φ is not verifiable under O and hence the rule is not operational. Apparently, φ is not minimally sufficient knowledge for choice predictions at time point t. This contradicts the assumption that no proper generalisation of δ is implemented by P. □

Note that if decision process P implements a rule δ non-strictly, then that rule is not guaranteed to be operational under model O. However, this is not problematic as its antecedent is impossible in that process, and therefore the rule is not applied.

5.3 Contingency planning

In this section we investigate the notion of *contingency plan*. Such a plan prescribes decision-making behaviour for a large number of situations in a planning problem under uncertainty. Recall from Section 5.1 that decision rules express action choices contingent upon preceding events and decisions. We define a set of decision rules for mutually exclusive situations to be a *contingency plan*, or *plan* for short.

Definition 5.54 (Plan) *A* contingency plan *is a set π of decision rules, where for each pair $\varphi_1 \rightarrow \mathrm{c}_{d(t_1)}, \varphi_2 \rightarrow \mathrm{c}_{d(t_2)} \in \pi$ of rules we have either $t_1 \neq t_2$ or $\varphi_1 \wedge \varphi_2 \equiv \bot$.*

A contingency plan π prescribes action choices for a collection of choice contexts in a given planning domain. We require that all rules for a given time point have mutually incompatible antecedents; a plan can therefore not contain conflicting rules or generalisations of its own rules. We note that the size of a plan π may (and often will) exceed the length of time frame T, as many decision rules for a single time point can co-exist without being conflicting: these rules then prescribe action choices for

different contexts. Furthermore, informational asymmetry in the decision problem is easily handled by varying the observed state variables in rules' antecedents with the actions chosen.

The motivation to describe planning behaviour by sets of decision rules is twofold. First, this notion of plan is very flexible as it allows for describing planning behaviour at varying levels of complexity. Within a single plan, some decision rules may consider a large number of past actions and observations in their antecedent, whereas others consider only a few. This is a significant difference with the representation formalisms of Chapter 4, where plans consist of sets of *functions*, and each of these functions has a fixed complexity. Second, as a plan is a set of rules, we can construct a plan by iteratively adding decision rules for arbitrary time points and choice contexts. This paves the way for new algorithms for decision-theoretic planning that are based on manipulation of symbolic structures.

Below, we first define a partial ordering on contingency plans and a normal form of plans (Subsection 5.3.1), and subsequently discuss the relation between contingency plans and decision processes (Subsection 5.3.2), the notion of plan completeness (Subsection 5.3.3), and finally the operationality of plans under a given model of observation (Subsection 5.3.4).

5.3.1 Plan ordering and normal form

In this subsection, we define an ordering relation on the set of all contingency plans for a given decision basis. In addition, a notion of plan equivalence, and a normal form of plans are introduced.

We will start with defining the notion of *coverage*, which provides the basis for comparing contingency plans. The underlying idea is that a given choice context ψ for time point $t \in T$ essentially represents a multitude of more specific contexts – that is, unless ψ is a most specific choice context for that time point. Recall that ψ is most specific when it is a configuration of the set $X[0,t] \cup D[0,t-1]$; otherwise, more specific contexts are obtained by adding information to ψ.

We now say that context ψ is *covered* by a given plan π, if one of the decision rules in π applies in context ψ, or if we can always find an applicable rule in the plan by adding information to ψ. In the former case, we have that $\varphi \rightarrow \mathrm{c}_{d(t)} \in \pi$ for some $\mathrm{c}_{d(t)} \in \Omega_{d(t)}$, where $\psi \vdash \varphi$. In the latter case, there exist a collection of mutually incompatible choice contexts $\varphi_1, \ldots, \varphi_m$, each covered by π and more informative than ψ, and such that

$$\psi \quad \equiv \quad \varphi_1 \vee \cdots \vee \varphi_m. \tag{5.38}$$

This is not a trivial condition, as choice contexts are defined to be conjunctions of positive literals (value assignments). Therefore, the above equivalence implies that there exist a set $Y \subseteq X[0,t] \cup D[0,t-1]$ of variables, such that each φ_i, $i = 1, \ldots, m$,

extends ψ with a different configuration c_Y^i of Y, and that the enumeration of these configurations by $\varphi_1, \ldots, \varphi_m$ is exhaustive. That is, $\varphi_i \equiv \psi \wedge \mathrm{c}_Y^i$ for all $i = 1, \ldots, m$, and

$$\Omega_Y \quad = \quad \{\mathrm{c}_Y^1, \ldots, \mathrm{c}_Y^m\}. \tag{5.39}$$

Therefore, we can regard context ψ as covered by the plan: irrespective of the actual configuration of Y, we can always find an applicable rule. It should be noted that the rules associated with $\varphi_1, \ldots, \varphi_m$ may prescribe different actions choices; context ψ is thus not specific enough provide an unambiguous choice, although it is covered by the plan.

The notion of coverage is now formally defined as follows.

Definition 5.55 (Coverage) *The* coverage *of contingency plan π at time point $t \in T$, written* $\mathrm{cover}(\pi, t)$*, is the smallest set of choice contexts for time point t such that $\psi \in \mathrm{cover}(\pi, t)$ if*

- $\psi \rightarrow \mathrm{c}_{d(t)} \in \pi$ *for some* $\mathrm{c}_{d(t)} \in \Omega_{d(t)}$*,*
- $\varphi \in \mathrm{cover}(\pi, t)$ *and* $\psi \vdash \varphi$*, or*
- $\varphi_1, \ldots, \varphi_m \in \mathrm{cover}(\pi, t)$ *and* $\psi \equiv \varphi_1 \vee \cdots \vee \varphi_m$*.*

When $\psi \in \mathrm{cover}(\pi, t)$, we use $\pi_t(\psi) \subseteq A$ to denote the set of actions that may be prescribed by plan π in that context, i.e., $a \in \pi_t(\psi)$ if and only if there exists a rule $\varphi \rightarrow d(t) = a \in \pi$ such that $\psi \wedge \varphi \not\equiv \bot$.

The first and second clause jointly describe the former case identified above, while the third clause describes the latter case.

Given that $\psi \in \mathrm{cover}(\pi, t)$, the plan π is said to be *equivocal* for context ψ when $|\pi_t(\psi)| > 1$; otherwise, the plan is called *unequivocal* for that context. As a result of our definition of contingency plan, equivocality is avoided for context ψ when ψ is covered by π due to the former case described above. Then, there exists a rule $\varphi \rightarrow d(t) = a \in \pi$ such that $\psi \vdash \varphi$; as there can be no other rule in π that applies in context ψ, we have $\pi_t(\psi) = \{a\}$. The latter case described above may induce equivocality, although this is not necessarily so. When context ψ is covered by π at time point t, but ψ is not specific enough to permit a single decision rule to be applicable, then each of the more specific contexts φ_i, $i = 1, \ldots, m$, may lead to a different decision $d(t) = a_i$. As a result, $\pi_t(\psi) = \{a_1, \ldots, a_m\}$, and π is equivocal for context ψ. Notwithstanding the ambivalence, if $a \notin \pi_t(\psi)$ then π will certainly not prescribe the action choice $d(t) = a$ in context ψ or any of its specialisations; we say this action choice is not an *option* in context ψ at time point t. Also note that we may have that $a_i = a_j$ for all $i, j = 1, \ldots, m$, rendering π unequivocal for context ψ.

A plan is guaranteed to be unequivocal when ψ is a most specific choice context for time point t and $\psi \in \mathrm{cover}(\pi, t)$.

Proposition 5.56 *Let π be a contingency plan, and let $\psi \in \text{cover}(\pi, t)$. If $\psi \equiv \text{c}$ for some configuration* c *of the set $X[0,t] \cup D[0,t-1]$, then π is unequivocal for ψ at time point t.*

Proof. Suppose that $|\pi_t(\text{c})| > 1$. This means that there exists different rules $\varphi_1 \rightarrow d(t) = a_1, \varphi_2 \rightarrow d(t) = a_2 \in \pi$, such that both $\text{c} \wedge \varphi_1 \not\equiv \bot$ and $\text{c} \wedge \varphi_1 \not\equiv \bot$. As c is a most specific choice context for this time point, it follows that $\text{c} \vdash \varphi_1 \wedge \varphi_2$, and therefore $\varphi_1 \wedge \varphi_2 \not\equiv \bot$. By Definition 5.54, this implies that π is not a proper plan. □

For choice contexts covered by a given plan, we can determine the set of options by inspecting all choice contexts that are more informative.

Proposition 5.57 *Let π be a contingency plan. If $\psi \in \text{cover}(\pi, t)$, then*

$$\pi_t(\psi) \;=\; \bigcup_{\varphi \in \text{cover}(\pi,t),\, \varphi \vdash \psi} \pi_t(\varphi). \tag{5.40}$$

Proof. We will first proof that if $a \in \pi_t(\psi)$, then $a \in \pi_t(\varphi)$ for some $\varphi \in \text{cover}(\pi, t)$, $\varphi \vdash \psi$, and then that the converse statement also holds.

- Suppose that $a \in \pi_t(\psi)$. By Definition 5.55, we know that there exists a rule $\varphi' \rightarrow d(t) = a \in \pi$ such that $\psi \wedge \varphi' \not\equiv \bot$. From the second clause in the definition we have that $\psi \wedge \varphi' \in \text{cover}(\pi, t)$ as $\psi \wedge \varphi' \vdash \varphi'$. We find that $a \in \pi_t(\psi \wedge \varphi')$, which completes this part of the proof as $\psi \wedge \varphi' \vdash \psi$.
- Now suppose that $a \in \pi_t(\varphi)$ for some $\varphi \in \text{cover}(\pi, t)$ that is more informative than ψ, i.e. $\varphi \vdash \psi$. Then $\varphi \wedge \psi \not\equiv \bot$, and again by Definition 5.55, it follows that $a \in \pi_t(\psi)$.

□

Given the notion of coverage, we can now define the *subplan-superplan* relation.

Definition 5.58 (Subplan) *Let π, π' be contingency plans. The plan π' is said to be a* subplan *of π when for each time point $t \in T$, $\psi \in \text{cover}(\pi', t)$ implies that $\psi \in \text{cover}(\pi, t)$ and $\pi'_t(\psi) \subseteq \pi_t(\psi)$.*

Conversely, we say that π is a *superplan* of π' in that case. The next proposition states that the subplan-superplan relation is completely determined by those choice contexts for which the subplan is unequivocal.

Proposition 5.59 *Let π, π' be contingency plans. The following two statements are equivalent:*

1. *π' is a subplan of π; and*
2. *for all $t \in T$, if $\varphi \in \text{cover}(\pi', t)$ and $\pi'_t(\varphi) = \{a\}$, then $\varphi \in \text{cover}(\pi, t)$ and $\pi_t(\varphi) = \{a\}$.*

Proof. We only prove the non-trivial case that the second statement implies the first, and in particular that a correct result is obtained for choice contexts for which plan π' is unequivocal. Let $t \in T$, and let $\varphi \in \text{cover}(\pi', t)$, where $|\pi'_t(\varphi)| > 1$. We have to prove that also $\varphi \in \text{cover}(\pi, t)$, and $\pi'_t(\psi) \subseteq \pi_t(\psi)$. Now, by the definition of coverage, we know that for each action $a \in \pi'_t(\varphi)$ there exists a more informative choice context $\psi \in \text{cover}(\pi', t)$ such that $\pi'_t(\psi) = \{a\}$. The second statement above now says that also $\psi \in \text{cover}(\pi, t)$ and $\pi_t(\psi) = \{a\}$; as a consequence, we have for the less informative choice context φ that $\varphi \in \text{cover}(\pi, t)$ and $a \in \pi_t(\varphi)$. □

An instance of the subplan-superplan relation is found in cases where for each decision rule $\delta \in \pi'$ there exists an generalisation in π; this will extend the coverage of the plan without modifying the action prescriptions. As each rule is a trivial generalisation of itself, π' is also a subplan of π when $\pi' \subseteq \pi$.

The notion of subplan induces a partial ordering on the set of all possible contingency plans for a given decision basis. Here, the empty plan $\pi = \varnothing$ is the universal lower bound (as it is a subplan of all plans including itself). There is no universal upper bound: any plan that does not allow extension without introducing a rule conflict is maximal with respect to the ordering. When two plans share their place in the ordering, we say that they are *equivalent*.

Definition 5.60 (Plan equivalence) *We say that contingency plans π and π' are* equivalent *when π is a subplan of π' and vice versa.*

We will henceforth refer to a non-equivalent subplan (superplan) as a *proper* subplan (superplan). Equivalent plans represent the same planning behaviour but it should be noted that they may be syntactically different. We will regard them as different representations of the same planning strategy, and prefer the smallest representation. This representation is called the *normal form*.

Definition 5.61 (Normal form) *We say that contingency plan π is in* normal form *when for all time points $t \in T$ and choice contexts ψ, if $|\pi_t(\psi)| = 1$, then there exists a decision rule $\delta \in \pi$ that is applicable in context ψ.*

Proposition 5.62 *If contingency plan π is in normal form, then there does not exist an equivalent plan with less decision rules.*

Proof. Suppose that π is in normal form, but there exists an equivalent plan π' with less rules. The latter circumstance implies that there exists a time point $t \in T$ and

a choice context ψ for t such that $|\pi_t(\psi)| = 1$ (and therefore $|\pi'_t(\psi)| = 1$) but plan π uses more rules than plan π' to prescribe this behaviour. That is, the coverage of ψ at time point t in plan π must be realised by a set of rules

$$\{\varphi_1 \to d(t)=a, \ldots, \varphi_m \to d(t)=a\} \subseteq \pi, \tag{5.41}$$

$m > 1$, where each antecedent φ_i, $i = 1, \ldots, m$, is strictly more specific than ψ (i.e. $\varphi_i \vdash \psi$ and $\varphi_i \not\equiv \psi$), but they jointly cover the context ψ (i.e. $\varphi_1 \vee \cdots \vee \varphi_m \equiv \psi$), and therefore $\pi_t(\psi) = a$. However, this contradicts our earlier assumption that ψ is in normal form, and we conclude that there does not exist such a plan π'. □

To summarise, a plan that is not in normal form contains a subset of related decision rules that can jointly be described by a single, more general rule. In the normal form of plans, such savings are not possible: a plan in normal form uses as less rules as possible, and all rules have the highest level of generalisation that is possible without sacrificing precision. To inspect the 'true' complexity of a contingency plan, one should therefore inspect the complexity of its rules in normal form.

5.3.2 Plans and decision process

We now turn to the *implementation* of contingency plans by decision processes. A plan is said to be implemented by a given process when that process implements all its decision rules. Recall that a decision rule is strictly implemented when its antecedent has nonzero probability.

Definition 5.63 (Implementation of plans) *Let P be a decision process, and let π be a contingency plan. The process P is said to* implement *the plan π when it implements all decision rules $\delta \in \pi$, and to* strictly implement *that plan when it strictly implements all its rules.*

It was earlier shown in the discussion of subplan-superplan relations that a contingency plan is fully determined by the choice contexts for which it is unequivocal in its prescribed action choice. Here, we will prove a similar property with respect to plan implementation.

Proposition 5.64 *Let P be a decision process, and let π be a contingency plan. Then, the following two statements are equivalent.*

1. *P strictly implements π; and*
2. *for all $t \in T$, $\psi \in \mathrm{cover}(\pi, t)$ with $\pi_t(\psi) = \{a\}$ it holds that $P(d(t) = a \mid \psi) = 1$.*

Proof. ($1 \Rightarrow 2$) Suppose that P implements π, and let $t \in T$ be a time point and $\psi \in \text{cover}(\pi, t)$ be a choice context such that $\pi_t(\psi) = \{a\}$. From these conditions we know that π has a subset

$$\{\varphi_1 \rightarrow d(t){=}a, \ldots, \varphi_m \rightarrow d(t){=}a\}, \tag{5.42}$$

$m \geq 1$, of decision rules such that $\varphi_1 \vee \cdots \vee \varphi_m \equiv \psi$. Each of these rules is strictly implemented by P, so

$$P(d(t){=}a \mid \varphi_1 \vee \cdots \vee \varphi_m) \;=\; 1, \tag{5.43}$$

and therefore

$$P(d(t){=}a \mid \psi) \;=\; 1. \tag{5.44}$$

($2 \Rightarrow 1$) Consider an arbitrary decision rule $\varphi \rightarrow d(t){=}a$ in plan π. We have that $\varphi \in \text{cover}(\pi, t)$, and $\pi_t(\varphi) = \{a\}$. So $P(d(t) = a \mid \varphi) = 1$, and P strictly implements the rule, and therefore also strictly implements the plan π. □

We note that as the above proposition holds for strict implementation, it also holds for implementation in general (strict and non-strict).

The next two corollaries follow directly from Propositions 5.59 and 5.64.

Corollary 5.65 *If a decision process strictly implements contingency plan π, then it also strictly implements all its subplans.*

Corollary 5.66 *If a decision process strictly implements contingency plan π, then it also strictly implements all its equivalent plans.*

Again, these statements also hold for the weaker concept of implementation in general.

The next proposition says that any action choice that is not an option in a given choice context is also excluded by a decision process that implements the plan.

Proposition 5.67 *Let P be a decision process that implements contingency plan π. If $\psi \in \text{cover}(\pi, t)$ and $a \notin \pi_t(\psi)$, then $P(\psi \wedge d(t){=}a) = 0$.*

Proof. Let $\varphi_1, \ldots, \varphi_m$ be the smallest set of most specific choice contexts for time point t such that $\psi \equiv \varphi_1 \vee \cdots \vee \varphi_m$. As $\psi \in \text{cover}(\pi, t)$ and each φ_i, $i = 1, \ldots, m$, is more specific than ψ, we also have $\varphi_i \in \text{cover}(\pi, t)$. Furthermore, $|\pi_t(\varphi_i)| = 1$ as it is most specific (Proposition 5.56). However, $a \notin \pi_t(\varphi_i)$; otherwise we would also have that $a \in \pi_t(\psi)$. By Proposition 5.64, we know that $P(\varphi_i \wedge d(t){=}a) = 0$, and therefore

$$P(\psi \wedge d(t){=}a) \;=\; P((\varphi_1 \wedge d(t){=}a) \vee \cdots \vee (\varphi_m \wedge d(t){=}a)) \;=\; 0.$$

□

We have earlier seen that decision-rule implementation is a trivial matter when the antecedent of the rule has zero probability in the given decision process; the notion of strict implementation is therefore much stronger and more interesting. Unfortunately, not all contingency plans permit strict implementation. This is due to the fact that the flexible nature of contingency plans allows for several forms of internal contradiction in plans. In the remainder of this subsection, we will further investigate this topic.

Definition 5.68 (Presupposition) *The expression $\varphi_1 \wedge \mathrm{c}_{d(t_1)}$, where φ_1 is a choice context for time point t_1, is called a* presupposition *of decision rule $\varphi_2 \rightarrow \mathrm{c}_{d(t_2)}$ when $t_1 < t_2$ and $\varphi_2 \vdash \varphi_1 \wedge \mathrm{c}_{d(t_1)}$.*

Intuitively, presuppositions of a given decision rule describe planning behaviour that must have been present for the rule to be applicable. When the rule's antecedent does not contain action choices, then the rule does not make any presuppositions. Note, however, that a presupposition of the form $\varphi_1 \wedge \mathrm{c}_{d(t_1)}$ does not express that decision $\mathrm{c}_{d(t_1)}$ is *necessarily* made in context φ_1. It says that decision $\mathrm{c}_{d(t_1)}$ has been made in that context, and the rule making the presupposition implicitly asserts that such is *possible*.

Definition 5.69 (Exclusion) *Let π be a contingency plan. We say that planning history $h \in \mathcal{H}(T)$ is* excluded *by π when there exists a decision rule $\delta \in \pi$ such that $h \wedge \delta \equiv \bot$. The set of all planning histories excluded by plan π is denoted by H_π^-.*

Exclusion of planning history h stems from the fact that it matches with the antecedent φ of some decision rule $\varphi \rightarrow d(t) = a \in \pi$, while contradicting the action choice $d(t) = a$. Then, $h \wedge \delta \equiv \bot$, and it is easily seen that we could never obtain history h by following the plan.

Proposition 5.70 *Let π be a contingency plan and let P be a decision process that implements π. Then, $P(h) = 0$ for all excluded histories $h \in H_\pi^-$.*

Proof. Let $h \in H_\pi^-$ be an excluded planning history and let $\varphi \rightarrow \mathrm{c}_{d(t)} \in \pi$ be the decision rule that is responsible for its exclusion, i.e. $h \vdash \varphi \wedge \neg \mathrm{c}_{d(t)}$ (there may be more rules with this property but that is irrelevant for the argument). From the implementation of π by P and Proposition 5.26, we have that $P(\varphi \wedge \neg \mathrm{c}_{d(t)}) = 0$, and so $P(h) = 0$. □

More generally, we also say that proposition $\varphi \in \Phi(T)$ is *excluded* by plan π when each $h \in \mathcal{H}(T)$, $h \vdash \varphi$, is excluded by π. It follows that also $P(\varphi) = 0$ when decision process P implements π.

The forms of internal plan contradiction mentioned above can be traced back to 'wrong' presuppositions. We can identify two cases where this happens. The first case is where a presupposition is directly contradicted by other rules in the plan through exclusions; we then say that the plan is *inconsistent*.

Definition 5.71 (Plan consistency) *A contingency plan π is said to be* inconsistent *if one of its rules makes a presupposition that is excluded by the plan itself. Otherwise, the plan is* consistent.

Proposition 5.72 *An inconsistent plan does not permit strict implementation.*

Proof. Suppose that decision process P implements contingency plan π, and assume that the plan is inconsistent. So there exists a decision rule $\delta \in \pi$ that presupposes $\varphi \wedge d(t)=a$, while this situation is excluded by the plan itself. It follows that $P(\varphi \wedge d(t)=a) = 0$ and therefore the antecedent of rule δ also has zero probability in process P. We conclude that the rule is not strictly implemented by P. □

Note that we cannot restore consistency by adding rules to an inconsistent plan: any superplan of an inconsistent plan is also inconsistent.

Inconsistent plans do not permit strict implementation, but unfortunately, there exist also consistent plans that do not permit that. This is then caused by the second case of internal plan contradiction, where multiple presuppositions are contradictory. In this case, the plan is said to be *incoherent*.

Definition 5.73 (Plan coherence) *Let π be a contingency plan. We say that the plan is* incoherent *if there exist decision rules $\delta_1, \delta_2 \in \pi$ that presuppose $\sigma \wedge d(t)=a_1$ and $\sigma \wedge d(t)=a_2$, respectively, where $\sigma \in \mathcal{S}[0,t]$ and $a_1 \neq a_2$. Otherwise, the plan is said to be* coherent.

Proposition 5.74 *An incoherent plan does not permit strict implementation by a decision process with deterministic planning behaviour.*

Proof. Suppose that P is a decision process with deterministic planning behaviour that strictly implements the incoherent plan π. The incoherence of π implies that there exist decision rules $\delta_1, \delta_2 \in \pi$ which presuppose $\sigma \wedge d(t)=a_1$ and $\sigma \wedge d(t)=a_2$, respectively, where $\sigma \in \mathcal{S}[0,t]$ and $a_1 \neq a_2$. Furthermore, from the strictness of the implementation we must have $P(\sigma) > 0$. Now, recall from Proposition 5.24 on page 127 that there exists a unique action sequence $\alpha \in \mathcal{A}[0,t]$ for which $P(\alpha \mid \sigma) = 1$, and $P(\alpha' \mid \sigma) = 0$ for all other $\alpha' \in \mathcal{A}[0,t]$. It follows that $P(d(t) = \alpha(t) \mid \sigma) = 1$, and $P(d(t) = a \mid \sigma) = 0$ for all actions $a \in A$, $a \neq \alpha(t)$. Therefore, we cannot have both $P(\sigma \wedge d(t)=a_1) > 0$ and $P(\sigma \wedge d(t)=a_2) > 0$ as $a_1 \neq a_2$. So, the implementation by P of either rule δ_1 or rule δ_2 is non-strict. □

Also the superplans of an incoherent plan are incoherent.

5.3.3 Plan completeness

We now turn to the notion of *plan completeness*. We say that a contingency plan is complete when it covers all possible initial states at time point $t = 0$, and covers all choice contexts for future time points that are compatible with prescribed action choices.

Definition 5.75 (Plan completeness) *A contingency plan π is* complete *when*

1. *if* c *is a configuration of* $X(0)$*, then* $\mathrm{c} \in \mathrm{cover}(\pi, 0)$*, and*
2. *for all* $t < N$*, if* c *is a configuration of* $X[0,t] \cup D[0,t-1]$ *and* $\mathrm{c} \in \mathrm{cover}(\pi, t)$*, then* $\mathrm{c} \wedge d(t) = a \in \mathrm{cover}(\pi, t+1)$ *where* $\pi_t(\mathrm{c}) = \{a\}$*.*

Otherwise, the plan is said to be incomplete*.*

Intuitively, the property of plan completeness ensures that we always know which action to choose if we have followed the plan at the preceding time points, whatever states the system occupies. Note that the requirement that $\mathrm{c} \wedge d(t) = a \in \mathrm{cover}(\pi, t+1)$ does not imply the existence of a rule with that antecedent in π: there will typically be rules that take into account some observation at time point $t+1$ while being less specific with respect to preceding time points. These rules should however jointly cover the case $\mathrm{c} \wedge d(t) = a$.

We can identify two types of plan that are trivially complete. The first type of plan does not take any state information into account and has the form

$$\pi \;=\; \{\top \rightarrow d(t) = \alpha(t) \mid t \in T\} \tag{5.45}$$

where $\alpha \in \mathcal{A}(T)$ is an action sequence over T. The second type of plan takes all possible state information into account and comprises a rule of the form

$$\mathrm{c}_{X(0)} \rightarrow \mathrm{c}_{d(0)} \tag{5.46}$$

for each $\mathrm{c}_{X(0)} \in \Omega_{X(0)}$, and a rule of the form

$$\mathrm{c}_{X(0)} \wedge \mathrm{c}_{d(0)} \wedge \mathrm{c}_{X(1)} \wedge \mathrm{c}_{d(1)} \wedge \cdots \mathrm{c}_{X(t-1)} \rightarrow \mathrm{c}_{d(t)} \tag{5.47}$$

for all time points $t \in T$, $t > 0$, and all configurations of $X(0), \ldots, X(t)$, where $\mathrm{c}_{d(0)}, \ldots, \mathrm{c}_{d(t-1)}$ are the unique decisions for these configurations prescribed by other rules of the plan.

It is easily seen that both of these plans prescribe a unique sequence of decisions when all possible state information is provided. This is, in fact, a general property of complete plans.

Proposition 5.76 *A complete contingency plan prescribes a unique action sequence for each state sequence over T.*

Proof. Let π be a complete plan, and let $\sigma \in \mathcal{S}(T)$ be a state sequence over $T = \{0, \ldots, N\}$, where $\sigma \equiv \mathrm{c}_{X(0)} \wedge \cdots \wedge \mathrm{c}_{X(N)}$. We will prove the proposition by induction on T. For the base case, the completeness of π ensures that $\mathrm{c}_{X(0)} \in \mathrm{cover}(\pi, 0)$. Furthermore, $|\pi_t(\mathrm{c}_{X(0)})| = 1$ as $\mathrm{c}_{X(0)}$ is a most specific context for time point $t = 0$ (Proposition 5.56). So, the plan prescribes a unique initial action. For the induction step, suppose that given σ, the plan prescribes a sequence $\mathrm{c}_{d(0)}, \ldots, \mathrm{c}_{d(t)}$ of decisions up to time point $t < N$. Now, let

$$\psi \;=\; \mathrm{c}_{X(0)} \wedge \mathrm{c}_{d(0)} \wedge \mathrm{c}_{X(1)} \wedge \cdots \wedge \mathrm{c}_{d(t-1)} \wedge \mathrm{c}_{X(t)}. \tag{5.48}$$

Apparently, $\psi \in \mathrm{cover}(\pi, t)$ and $\pi_t(\psi) = \{a\}$, where $\mathrm{c}_{d(t)} \equiv d(t) = a$. The completeness of plan π now ensures that $\psi \wedge \mathrm{c}_{d(t)} \in \mathrm{cover}(\pi, t+1)$, and so also the more specific $\psi \wedge \mathrm{c}_{d(t)} \wedge \mathrm{c}_{X(t+1)} \in \mathrm{cover}(\pi, t+1)$. Again using Proposition 5.56, we conclude that π prescribes a unique decision at time point $t+1$ for this context. □

The trivially complete plan of Equation 5.45 does of course always prescribe the *same* action sequence α, regardless of the states of the dynamic system.

The result of Proposition 5.76 proves to be quite useful in the establishment of additional properties of complete plans, as it allows us to conclude that planning history $h = (\sigma, \alpha)$ is excluded by complete plan π when α is *not* the unique action sequence prescribed by π given state sequence σ. The next proposition shows that complete plans are the uppermost plans in the partial ordering that are of practical use, as greater plans are always inconsistent. Note that any superplan of a complete plan is also complete.

Proposition 5.77 *A complete plan does not have proper superplans that are consistent.*

Proof. Suppose that contingency plan π is complete, and that π' is a consistent, proper superplan of π. That is, for all $t \in T$ and choice contexts ψ, if $\pi_t(\psi) = \{a\}$ then also $\pi'_t(\psi) = \{a\}$ (Proposition 5.59), but there exists a time point $t \in T$ and a choice context $\psi \in \mathrm{cover}(\pi', t)$ for which $\pi'_t(\psi) = \{a\}$, but $\psi \notin \mathrm{cover}(\pi, t)$. Assuming that both π and π' are written in normal form, we know that there exists a decision rule $\varphi \rightarrow d(t) = a \in \pi'$ that applies in context ψ (i.e. $\psi \vdash \varphi$). This rule is however not contained in π. Now, let $h \in \mathcal{H}(T)$, $h = (\sigma, \alpha)$, be a planning history such that $h \vdash \varphi$; from Proposition 5.76 we know that α is the unique action sequence prescribed by π' for state sequence σ. We distinguish two cases:

1. If α is also prescribed by plan π, there must exist a rule $\varphi' \rightarrow d(t) = a \in \pi$ with $h \vdash \varphi'$. But then π' cannot be a proper plan as also $\varphi' \rightarrow d(t) = a \in \pi'$ and $\varphi \wedge \varphi' \not\equiv \bot$.

2. If α is not prescribed by plan π, then h must be excluded by π: there exists a decision rule $\delta \in \pi$ for which $h \wedge \delta \equiv \bot$. As π' is a superplan of π, this rule is also contained in π', and history h is similarly excluded by π'. We can repeat this argument for each history h for which $h \vdash \varphi$; therefore, φ is itself excluded by π': the plan π' is inconsistent.

We conclude that such a consistent, proper superplan of π does not exist. □

Corollary 5.78 *A consistent and complete plan does not have proper subplans that are complete.*

Proof. By contrapositioning Proposition 5.77 □

Definition 5.79 (Completion) *Let π be a plan. A consistent and complete superplan of π is called a* completion *of π.*

We can think of an incomplete plan as a concise representation of all its completions. By adding decision rules to the plan, the range of possible completions is narrowed until a single consistent and complete plan remains. A incoherent plan does not have any completions as its incoherence unavoidably leads to inconsistencies when completeness is approached.

Proposition 5.80 *A consistent and complete plan is coherent.*

Proof. Suppose that plan π is consistent and complete, and moreover incoherent. The latter assumption implies that there exist decision rules $\delta_1, \delta_2 \in \pi$ that presuppose some $\sigma \wedge d(t) = a_1$ and $\sigma \wedge d(t) = a_2$, respectively, where $\sigma \in \mathcal{S}[0,t]$ and $a_1 \neq a_2$. Now, let $\alpha \in \mathcal{A}[0,t]$ be the unique action sequence up to time point t that is prescribed by π given state sequence σ. We cannot have both $\alpha(t) = a_1$ and $\alpha(t) = a_2$. This means that either one of the presuppositions is excluded by π, contradicting our earlier assumption that π is consistent. We conclude that π must be coherent. □

Corollary 5.81 *An incoherent plan does not have any completions.*

Proof. Directly from Proposition 5.80 and the fact that any superplan of an incoherent plan is also incoherent. □

Below, we assume all plans to be consistent, but we will tolerate incomplete yet coherent plans as intermediate steps in a solution process. These incomplete plans are then taken to represent a set of possible complete plans.

As we saw in Proposition 5.76, a principal property of complete plans is that they prescribe a unique sequence of actions, given a complete sequence of system states over the time frame T. This reminds of the similar finding for decision processes with deterministic planning behaviour of Proposition 5.24 in Subsection 5.1.2. There does indeed exist an intimate relationship between complete plans and such processes; we conclude this subsection with a further investigation of this relationship.

We first establish that there is only one decision process that implements a given complete plan under an assumed model of control.

Theorem 5.82 *Let Ξ be a control model, and let π be a complete contingency plan for decision basis $\mathcal{B}$. Then, the decision process $P \in \mathbb{P}_\Xi$ that implements π is unique.*

Proof. Suppose that $P_1, P_2 \in \mathbb{P}_\Xi$ are different decision processes that both implement π. If P_1 and P_2 are different, then there must exist a proposition $\varphi \in \Phi(T)$ for which $P_1(\varphi) \neq P_2(\varphi)$. As we can write φ as a disjunction of mutually incompatible decision-making histories over T (as in Equation 5.3), this would imply that there exists a history $h \in \mathcal{H}(T)$ for which $P_1(h) \neq P_2(h)$. We will now prove by induction on T that such a history does not exist under the given conditions.
The base case consists of proving that P_1 and P_2 induce the same marginal probability distribution on $X(0) \cup \{d(0)\}$. This proceeds as follows. First, as P_1 and P_2 both comply with model Ξ, we have that

$$P_1(\mathrm{C}_{X(0)}) \;=\; P_2(\mathrm{C}_{X(0)}). \tag{5.49}$$

Second, as P_1 and P_2 both implement the complete plan π, we also have that

$$P_1(\mathrm{C}_{d(0)} \mid \mathrm{c}_{X(0)}) \;=\; P_2(\mathrm{C}_{d(0)} \mid \mathrm{c}_{X(0)}) \tag{5.50}$$

for all $\mathrm{c}_{X(0)} \in \Omega_{X(0)}$ for which $P_1(\mathrm{c}_{X(0)}), P_2(\mathrm{c}_{X(0)}) > 0$, and therefore

$$P_1(\mathrm{C}_{X(0)\cup\{d(0)\}}) \;=\; P_2(\mathrm{C}_{X(0)\cup\{d(0)\}}). \tag{5.51}$$

The induction hypothesis is that

$$P_1(\mathrm{C}_{X[0,t]\cup D[0,t]}) \;=\; P_2(\mathrm{C}_{X[0,t]\cup D[0,t]}) \tag{5.52}$$

for $t < N$, from which we have to show that the same equality holds for the next point in time. Again we use the joint compliance of P_1 and P_2 with control model Ξ, now yielding

$$P_1(\mathrm{C}_{X(t+1)} \mid \mathrm{C}_{X[0,t]\cup D[0,t]}) \;=\; P_2(\mathrm{C}_{X(t+1)} \mid \mathrm{C}_{X[0,t]\cup D[0,t]}), \tag{5.53}$$

and therefore

$$P_1(\mathrm{C}_{X[0,t+1]\cup D[0,t]}) \;=\; P_2(\mathrm{C}_{X[0,t+1]\cup D[0,t]}). \tag{5.54}$$

As P_1 and P_2 both implement the same complete plan, we have

$$P_1(\mathrm{C}_{d(t+1)} \mid \mathrm{C}_{X[0,t+1]\cup D[0,t]}) \;=\; P_2(\mathrm{C}_{d(t+1)} \mid \mathrm{C}_{X[0,t+1]\cup D[0,t]}), \tag{5.55}$$

and so

$$P_1(\mathrm{C}_{X[0,t+1]\cup D[0,t+1]}) \;=\; P_2(\mathrm{C}_{X[0,t+1]\cup D[0,t+1]}). \tag{5.56}$$

By induction on T we now find that

$$P_1(h) \;=\; P_2(h) \tag{5.57}$$

for all decision-making histories h over T. □

It is easily seen from the proof that the decision process $P \in \mathbb{P}_\Xi$ implementing a given plan π is not only unique, but also has deterministic planning behaviour. Conversely, a given decision process with such behaviour can be shown to implement a consistent and complete plan.

Theorem 5.83 *Let Ξ be a control model. Each decision process $P \in \mathbb{P}_\Xi^{\mathrm{det}}$ implements a consistent and complete contingency plan for decision basis $\mathcal{B}$.*

Proof. We construct the plan π that is implemented by P as follows. Initialise $\pi = \varnothing$. For each $\mathrm{c}_{X(0)} \in \Omega_{X(0)}$ we find a configuration $\mathrm{c}_{d(0)}$ of decision variable $d(0)$ such that

$$P(\mathrm{c}_{X(0)} \wedge \mathrm{c}_{d(0)}) \;=\; P(\mathrm{c}_{X(0)}). \tag{5.58}$$

and add the decision rule $\mathrm{c}_{X(0)} \to \mathrm{c}_{d(0)}$ to π. Note that such a configuration of decision variable $d(0)$ exists for each $\mathrm{c}_{X(0)} \in \Omega_{X(0)}$ as the planning behaviour is deterministic in process P. Then, for each $\mathrm{c}_{X[0,1]} \in \Omega_{X[0,1]}$ we find configurations $\mathrm{c}_{d(0)}$ and $\mathrm{c}_{d(1)}$ of decision variables $d(0)$ and $d(1)$ such that

$$P(\mathrm{c}_{X[0,1]} \wedge \mathrm{c}_{d(0)}) \;=\; P(\mathrm{c}_{X[0,1]}) \tag{5.59}$$

and

$$P(\mathrm{c}_{X[0,1]} \wedge \mathrm{c}_{d(0)} \wedge \mathrm{c}_{d(1)}) \;=\; P(\mathrm{c}_{X[0,1]} \wedge \mathrm{c}_{d(0)}). \tag{5.60}$$

This time we add the rule $\mathrm{c}_{X[0,1]} \wedge \mathrm{c}_{d(0)} \to \mathrm{c}_{d(1)}$ to π. Again we are sure to find configurations satisfying these requirements given the deterministic planning behaviour in P. We can now repeat the second step for time points $t = 2, \ldots, N$; it is easily verified that the plan thus obtained is consistent and complete, and implemented by P. □

One might be inclined to think that there must exist a bijection between $\mathbb{P}_\Xi^{\mathrm{det}}$ and the set of all consistent and complete plans for a given decision basis. This is not the case as our representation of decision strategies is not unique: multiple, syntactically

different plans may occupy the same place in the plan lattice and therefore describe the same strategy; such plans have previously been defined to be equivalent. There is however only one plan at each place in the lattice that has the normal form. We will refer to the consistent and complete plan π in normal form as *the* complete plan that is implemented by decision process $P \in \mathbb{P}_{\Xi}^{\mathrm{det}}$.

It should be noted that the above theorems deal with implementation in general, not strict implementation. The possibility to strictly implement complete plans critically depends on the characteristics of the control model Ξ: only when the model is strictly positive, we are sure to find a decision process $P \in \mathbb{P}_{\Xi}^{\mathrm{det}}$ that strictly implements a complete plan π. Therefore, Theorems 5.82 and 5.83 only apply to strict implementation when the control model in question is strictly positive.

5.3.4 Plan operationality

The final topic of contingency plans to be discussed is their operationality. Recall that we have previously defined a decision rule to be operational under a given observation model when its antecedent is verifiable under that model. The definition of plan operationality is a straightforward generalisation.

Definition 5.84 (Plan operationality) *Let O be an observation model. A contingency plan π is said to be* operational *under O when each of its decision rules is operational under that model.*

The following result is of crucial importance for incremental procedures of plan construction.

Proposition 5.85 *Let O be an observation model. Any consistent and coherent plan that is operational under O has an operational completion.*

Proof. Let π be a plan that is operational under O. We will add operational rules to the plan until it is complete, thus obtaining an operational completion. As the given plan is coherent, we are sure to be able to do this without introducing inconsistencies. Suppose that π is complete for time points $0, \ldots, t-1$ but not thereafter: t is the first point in time for which the plan does not always prescribe a unique action choice when a state sequence over $[0, t]$ is given. As π is complete for the preceding time points, there exists a unique action sequence $\alpha \in \mathcal{A}[0, t-1]$ that is prescribed by π for $\sigma(0), \ldots, \sigma(t-1)$ (Proposition 5.76). Now, let c_Y be the unique configuration of $Y = o_t(\alpha)$ smaller than σ, i.e. $\sigma \vdash \mathrm{c}_Y$. We add an operational rule

$$\alpha \wedge \mathrm{c}_Y \rightarrow d(t) = a, \tag{5.61}$$

$a \in A$, to the plan π, yielding an operational superplan. If the plan is now complete for time point t, we proceed with the next time point; otherwise, the procedure is

repeated for other state sequences. The plan that is finally obtained is a complete superplan of the original plan π that is operational under O. □

Without proof we remark that each operational plan has an operational normal form.

We have also defined a decision process P to comply with a given observation model O if minimally sufficient knowledge for choice predictions is always verifiable under the model: only then the planning behaviour that is implicitly specified by the process is reproducible under the limitations in observability described by O. It followed that if P strictly implements the rule δ and neither of its proper generalisations, that rule must be operational under O (Proposition 5.53). The next proposition is an immediate result hereof.

Proposition 5.86 *Let O be an observation model and let $P \in \mathbb{P}_O$. Any complete plan that is strictly implemented by P is operational under O.*

Proof. As P strictly implements a complete plan, it necessarily comprises deterministic planning behaviour (i.e. $P \in \mathbb{P}_O^{\mathrm{det}}$), and the complete and consistent plan implemented by it is unique up to equivalence. Let π denote this plan in normal form. We have to prove that P does not implement any proper generalisation of a decision rule in $\delta \in \pi$; the operationality of π under O then follows from Proposition 5.53. Now, suppose the $\varphi \rightarrow d(t) = a \in \pi$, so $P(d(t) = a \mid \varphi) = 1$, but there is also a proper generalisation $\psi \rightarrow d(t) = a$ of that rule (i.e. $\varphi \vdash \psi$, $\varphi \not\equiv \psi$) for which $P(d(t) = a \mid \psi) = 1$. This generalisation must be comprised in a complete and consistent plan π' that is implemented by P and equivalent to π. But then there must be a subset $\{\varphi_1 \rightarrow d(t) = a, \ldots, \varphi_m \rightarrow d(t) = a\} \subseteq \pi$, $m > 1$, of rules in π such that $\psi \equiv \varphi_1 \vee \cdots \vee \varphi_m$, $\varphi = \varphi_i$ for some $1 \leq i \leq m$. This contradicts our earlier assumption that the plan π was in normal form, and we conclude that P does not implement any proper generalisation of a decision rule $\delta \in \pi$. Hence, each of its rules is operational under O. □

5.4 Planning objectives

In this section we discuss the guidance of the planning agent's behaviour through formalisation of his planning objectives; this will complete our framework for decision-theoretic planning. To fulfil the principles of decision theory, the objectives of planning for a given domain should be formulated using a *utility function*. Given a utility function, we can start reasoning about the value and validity of choosing decision alternatives, observing state information, applying decision rules, and choosing contingency plans.

Utility and preference

Recall from Sections 3.2 and 3.3 of Chapter 3 that we can express a preference order on probability distributions by a numerical utility function, and that utilities associated with nondeterminate, uncertain scenarios (i.e. probability distributions over outcomes) can be derived from the utilities associated with outcomes using the rule of expectation from probability theory. It is therefore sufficient to specify a utility value for each possible outcome of the planning process to lay down a complete specification of the planning objectives; such utility values are called *marginal utilities*, and the values associated with uncertain scenarios are called *expected utilities*. Within our framework, the set $\mathcal{H}(T)$ of full-length planning histories represents all possible outcomes of decision processes for decision basis $\mathcal{B} = (X, A, T)$. We therefore take the objectives of planning to be formulated using a marginal utility function

$$u : \mathcal{H}(T) \to \mathbb{R} \tag{5.62}$$

on this set; the value $u(h)$ is the marginal utility associated with planning history $h \in \mathcal{H}(T)$. We now define the notions of *conditional utility* and *conditional preference*.

Definition 5.87 (Conditional utility and preference) *Let u be a marginal utility function for decision basis $\mathcal{B} = (X, A, T)$, and let $\varphi \in \Phi(T)$ be a planning expression. The* (expected) conditional utility $\tilde{u}_\varphi(P)$ *of decision process P given φ is defined as*

$$\tilde{u}_\varphi(P) \;=\; \sum_{h \in \mathcal{H}(T)} P(h \mid \varphi) \cdot u(h), \tag{5.63}$$

provided that φ is possible in P; otherwise, $\tilde{u}_\varphi(P)$ is undefined. We say that P is conditionally preferred to *decision process P' given φ when $\tilde{u}_\varphi(P) \geq \tilde{u}_\varphi(P')$.*

Sometimes, we refer to

$$\tilde{u}_\top(P) \;=\; \sum_{h \in \mathcal{H}(T)} P(h) \cdot u(h) \;=\; E_P(u) \tag{5.64}$$

as *the (expected) utility* of decision process P, and omit the subscript, i.e. $\tilde{u}(P) = \tilde{u}_\top(P)$. The notion of conditional preference induces a preference ordering on decision processes, which allows us to formalise the concept of *optimality*.

Definition 5.88 (Optimality) *Let u be a utility function for decision basis $\mathcal{B} = (X, A, T)$, and let $\varphi \in \Phi(T)$ be a planning expression. Given a set of $\mathbb{P}$ of decision processes, we refer to the process*

$$P^*_\varphi \;=\; \arg\max_P \{\tilde{u}_\varphi(P) \mid P \in \mathbb{P}\} \tag{5.65}$$

as optimal *given u and φ. A contingency plan π is said to be* optimal *given u, φ, and $\mathbb{P}$, when there exists an optimal decision process P^*_φ that strictly implements π; each of the decision rules in π is then also called* optimal *given u, φ, $\mathbb{P}$.*

Note that this definition is simply a reformulation of the MEU criterion (Lemma 3.21 on page 60) in terms of our framework. As with any application of utility-theoretic principles, optimality does not guarantee uniqueness: that there may be multiple optimal decision processes in $\mathbb{P}$ for a given utility function and planning expression. Another point to note is that decision rule δ is optimal under the above definition whenever there exists an optimal decision process that strictly implements δ. However, a contingency plan that is assembled by collecting such rules need *not* be optimal: each of the rules may be part of a different optimal solution.

If we employ our framework purely for decision-theoretic action planning, the construction of optimal plans can be regarded as searching the set of $\mathbb{P}_{\Xi,O}$ of decision processes that comply given models Ξ and O and control and observation to find the unconditionally optimal process $P^* \in \mathbb{P}_{\Xi,O}$ under a given utility function u. From the normative decision-theoretic perspective, the planning behaviour implemented by process P^* is then preferred over any other type of behaviour under the circumstances described by Ξ and O. It is generally reasonable to restrict the search to the subset $\mathbb{P}^{\,\mathrm{det}}_{\Xi,O}$ of decision processes that implement fully deterministic planning behaviour. Then, each process corresponds to a complete and consistent contingency plan π that is operational under O. Using the theory of contingency planning from the previous section, plan construction can then be formulated as a procedure of adding optimal operational decision rules to an optimal, consistent and coherent plan until that plan is complete.

The contributors to utility

It will often occur that there is a particular set of variables that designate utility values, while the other variables are, strictly speaking, unimportant. Variables from this first set are then called *direct contributors* to utility.

Definition 5.89 (Contribution) *Let u be a utility function for decision basis $\mathcal{B} = (X, A, T)$. The set of* direct contributors *to u is the smallest set $Y \subseteq X(T) \cup D(T)$ such that*

$$u(\mathrm{C}_Y \wedge \mathrm{c}_Z) \;=\; u(\mathrm{C}_Y \wedge \mathrm{c}'_Z) \tag{5.66}$$

for all configurations $\mathrm{c}_Z, \mathrm{c}'_Z$ of the complementary set $Z = (X(T) \cup D(T)) \setminus Y$.

If, in a given decision process P, the values of all direct contributors to utility are known, then the values of other variables are irrelevant to determine the utility value of that process. It is easily seen that this paves the way for a simplified form of marginal utility function.

Proposition 5.90 *Let u be a utility function for decision basis $\mathcal{B} = (X, A, T)$, and let Y be the set of all direct contributors to u. Then, there exists a function $u' : \Omega_Y \to \mathbb{R}$ such that for each history $h \in \mathcal{H}(T)$ we have $u(h) = u'(\mathrm{c}_Y)$, where $h \vdash \mathrm{c}_Y$.*

Furthermore, the expected conditional utility of decision process P given planning expression $\varphi \in \Phi(T)$ equals

$$\tilde{u}_\varphi(P) \;=\; \sum_{\mathrm{c}_Z \in \Omega_Z} P(\mathrm{c}_Y \mid \varphi) \cdot u'(\mathrm{c}_Y), \tag{5.67}$$

if φ is possible in P.

Proof. The result follows immediately if we take $u'(\mathrm{c}_Y) = u(\mathrm{c}_Y \wedge \mathrm{c}_Z)$, where c_Z is an arbitrary configuration of the set $Z = (X(T) \cup D(T)) \setminus Y$. □

Note, that if we do *not* know the values of all direct contributors to utility, then all information may influence the expected utility of a given decision process; this is not limited to information on direct contributors. Instead of referring to individual state variables, we will often say that system attribute $x \in X$ is an direct contributor to utility, to mean that this holds for all variables $x(t)$, $x(t) \in x(T)$.

Time-separable utility

In Subsection 3.2.3 of Section 3.2 we discussed multiattribute utility theory and the notion of *utility independence*, which builds on the concept of conditional preference ordering. In the current formulation, a given set $Y \subseteq X(T) \cup D(T)$ of variables is utility independent (of the other variables) when each configuration of Y induces the same conditional preference ordering on decision processes. We now define the notion of *time-separability*, which is derived from this concept. A utility function is time-separable when all planning histories over a limited time segment induce the same conditional preference ordering on decision processes.

Definition 5.91 (Time-separability) *A utility function u is* time-separable *when for each time segment $T' \sqsubseteq T$ and each pair of decision process P_1, P_2 we have*

$$\tilde{u}_{h_1}(P_1) \leq \tilde{u}_{h_1}(P_2) \;\Rightarrow\; \tilde{u}_{h_2}(P_1) \leq \tilde{u}_{h_2}(P_2) \tag{5.68}$$

for all histories $h_1, h_2 \in \mathcal{H}(T')$.

It follows from Theorem 3.25 on page 64 that under the assumption of time-separability, marginal utilities can be written as a multilinear function of independent utility values for subsequent time points:

$$u(h) \;=\; \sum_{T' \sqsubseteq T, T' \neq \varnothing} k^{|T'|-1} \prod_{t \in T'} u_t(\sigma(t), \alpha(t)), \tag{5.69}$$

where $h = (\sigma, \alpha)$, $u_t : \mathrm{dom}(X) \times A \to \mathbb{R}$ is the *subutility function* associated with time point $t \in T$, and k is a scaling constant; we do not use weight factors here as normalisation of the utility functions is not assumed. A special case of time-separable

utility occurs when $u_{t_1} = u_{t_2}$ for all time points $t_1, t_2 \in T$. The procedure for judging state-action pairs is then essentially the same at all time points; we speak of *stationary subutility*.

It was also discussed in Subsection 3.2.3 that under various additional assumptions, more convenient forms of multiattribute utility function are derivable; we described the *additive* and the *multiplicative* utility functions. We will now focus on both types of multiattribute function as specialisations of time-separable utility functions.

In the time-separable additive utility function, marginal utility values are equal to the sum of subutility values for subsequent time points:

$$u(h) \quad = \quad \sum_{t \in T} u_t(\sigma(t), \alpha(t)) \tag{5.70}$$

where $h = (\sigma, \alpha)$. As described in the previous chapter, time-separable additive utility models are employed in Markov decision processes; we will return to this topic in the next section where partially-observable Markov decision processes are reconsidered in the light of our framework. Here, we give two examples of time-separable additive utility, one where subutilities are solely dependent on action choices (and represent costs), and one where subutilities are solely dependent on system states (and represent planning goals).

Example 5.92 (Action costs) *An example of time-separable additive utility is found in models where each action has a particular* cost*, and the objective is to minimise the total costs. Let* $cost(a)$ *denote the cost associated with action* $a \in A$*; it is interpreted as the negative stationary subutility associated with choosing that action, i.e.*

$$u_t(S, a) \quad = \quad -cost(a) \tag{5.71}$$

for all $S \in \mathrm{dom}(X)$*. The marginal utility of decision-making history* $h \in \mathcal{H}(T)$*,* $h = (\sigma, \alpha)$*, equals*

$$u(h) \quad = \quad \sum_{t \in T} u_t(\sigma(t), \alpha(t)) \quad = \quad -\sum_{t \in T} cost(\alpha(t)). \tag{5.72}$$

Note that in this example, the decision variables are direct contributors to utility.

Example 5.93 (Goal-directed planning) *We can also employ time-separable additive utility functions to induce goal-directed behaviour in planning systems. The underlying conception of a planning problem is then a disparity between the current state and an envisioned state; the subutilities therefore depend on the states occupied by the dynamic system. Let* $Y \subseteq X$ *be a set of system attributes, let* $S_Y \in \mathrm{dom}(Y)$ *be the* goal configuration *of* Y*, and let* u *be a time-separable additive utility function.*

- *To express that S_Y is always preferred over other values of Y, we use stationary subutilities, where*

$$u_t(S_X, a) = \begin{cases} 1, & \text{if } X = S_X \vdash Y = S_Y, \\ 0, & \text{otherwise,} \end{cases} \tag{5.73}$$

for all $t \in T$, $S_X \in \mathrm{dom}(X)$ and $a \in A$. That is, a unit subutility value is obtained each time that S_Y is the substate for Y in the overall system state S_X. The expected utility $\tilde{u}(P)$ of decision process P equals the expected number of occurrences of value S_Y during time frame T.

- *To express that S_Y is a* final-state goal *(i.e. S_Y should be achieved at the final time point $t = N$), we use*

$$u_t(S_X, a) = \begin{cases} 1, & \text{if } X = S_X \vdash Y = S_Y \text{ and } t = N, \\ 0, & \text{otherwise,} \end{cases} \tag{5.74}$$

for all $t \in T$, $S_X \in \mathrm{dom}(X)$ and $a \in A$. The unit subutility value is obtained only when S_Y is the substate for Y at time point $t = N$. The expected utility $\tilde{u}(P)$ of decision process P now equals the probability that $Y(N) = S_Y$ is achieved: $\tilde{u}_1(P) = P(Y(N) = S_Y)$.

This type of utility function can also be extended to cover deadline goals, goals of prevention, and goals of maintenance; see (Haddawy and Hanks, 1998) for a discussion.

In the above example, neither decision variables nor system attributes from the set $X \setminus Y$ contribute directly to utility. When S_Y is a final-state goal, only state variables from the set $Y(N)$ are direct contributors.

The second type of time-separable utility function we discuss is called *multiplicative*, and takes the form

$$u(h) = \prod_{t \in T} u_t(\sigma(t), \alpha(t)) \tag{5.75}$$

where $h = (\sigma, \alpha)$. We recall that this type of function is characterised by much less robustness than the additive type of function: it can be compared to a chain that is only as strong as its weakest link. A typical application of such functions is found in cases where subutilities represent complemented risks.

Example 5.94 (Survival probability) *Let $risk_t(S, a)$ be the risk of contracting some particular damage or disorder at time point $t \in T$ in system state $X(t) = S$ under action choice $d(t) = a$. In a medical setting, we can think of the risk of particular complications or death. When the objective is to minimise the overall risks during the planning process, this is accomplished by maximising the multiplicative time-separable utility function of Equation 5.75, where*

$$u_t(S, a) = 1 - risk_t(S, a) \tag{5.76}$$

for all $t \in T$, $S \in \mathrm{dom}(X)$ and $a \in A$. The marginal utility $u(h)$ now expresses the probability that the risk is avoided during time frame T under history $h \in \mathcal{H}(T)$, and is consequently called the survival probability *of outcome h. The expected utility $\tilde{u}(P)$ of decision process $P \in \mathbb{P}^{\mathrm{det}}_{\Xi,O}$ expresses the probability of avoiding the risk when executing the contingency plan associated with P.*

5.5 POMDPs and DIDs revisited

We now return to the decision-theoretic representation formalisms that were discussed in Chapter 4, to re-evaluate and compare them in the light of our framework. In that chapter, we concluded that partially-observable Markov decision processes (POMDPs) and dynamic influence diagrams (DIDs) provide the most powerful representations of decision-making problems to date: these representation formalisms cover notions as time, action, uncertainty, state dynamics, contingent choice, and utility. Below, we will describe both types of representation in our framework (Subsection 5.5.1 and 5.5.2, respectively), highlighting the characteristics of these representations, and discussing the aspects where they are restrictive or make implicit assumptions. We conclude with a comparison in Subsection 5.5.3.

5.5.1 POMDPs

Recall from Section 4.2 that a Markov decision process is a stochastic Markov process where the transition probabilities are chosen by a planning agent[1]: each action has an associated transition probability function. In the partially-observable case, which we consider here, these action choices also influence the planning agent's possibilities to inspect the state of the system. Furthermore, at each point in time, a reward is received that depends on the current system state and action choice; the objective is to maximise the expected sum of all rewards.

Formally, a POMDP over a given set X of system attributes is described by a 5-tuple $\mathcal{M}^{\mathrm{PO}} = (T, A, \Theta, R, O)$, where T is a set of decision moments, A is a set of possible actions, Θ is a set of transition probability functions, R is a set of reward functions, and O is a set of observation functions. In addition, a probability distribution P_0 on initial system states is given.

We can reformulate this specification in our framework, where $\mathcal{B} = (X, A, T)$ is the induced decision basis; no modifications to the sets X, A, and T are required. Below, we describe how associated models of control and observation and a utility function are induced from Θ, O, and R, respectively; we also describe the type of contingency plan that serves as a solution to the given POMDP.

[1] We use the term *planning agent* instead of *decision maker* in this chapter.

Control

In the POMDP model $\mathcal{M}^{\mathrm{PO}}$, the planning agent's control over the dynamic system is described by the set

$$\Theta = \{\theta_t : \mathrm{dom}(X) \times A \times \mathrm{dom}(X) \to [0,1] \mid t \in T\} \tag{5.77}$$

of time- and action-dependent transition probability functions: if $S_1 \in \mathrm{dom}(X)$ is the system state at time point $t \in T$, $t < N$, and the planning agent chooses action $a \in A$, the system occupies state $S_2 \in \mathrm{dom}(X)$ at time point $t+1$ with probability $\theta_t(S_1, a, S_2)$. So, we can regard control model Ξ for decision basis $\mathcal{B}$ as a translation of Θ if and only if

$$P(X(t+1) = S_2 \mid X(t) = S_1, d(t) = a) \;=\; \theta_t(S_1, a, S_2) \tag{5.78}$$

for all decision processes $P \in \mathbb{P}_{\Xi}$, time points $t \in T$, $t < N$, actions $a \in A$, and states $S_1, S_2 \in \mathrm{dom}(X)$. It is assumed in these decision processes that the Markov condition holds, i.e. no states or decisions prior to time point t influence this transition probability:

$$X(t+1) \perp\!\!\!\perp_P X[0,t], D[0,t] \mid X(t), d(t). \tag{5.79}$$

Consequently, the probability that action sequence $\alpha \in \mathcal{A}(T)$ yields state sequence $\sigma \in \mathcal{S}(T)$ is obtained by simply multiplying the associated transition probabilities:

$$\xi_\alpha(\sigma) \;=\; P_0(X = \sigma(0)) \cdot \prod_{t=0}^{N-1} \theta_t(\sigma(t), \alpha(t), \sigma(t+1)), \tag{5.80}$$

where P_0 is the probability distribution on initial states. In this way, we can derive the control distribution ξ_α for each action sequence $\alpha \in \mathcal{A}(T)$, thus obtaining the model of control Ξ induced by Θ.

The Markov assumption is impractical when (part of) the system described by the set X may be subject to changes that are best understood as accumulative effects of certain conditions over time. We recall however from Section 4.2 that from a theoretical perspective, the assumption is not truly restrictive as a given non-Markovian specification can always be converted to a Markovian specification by extending the set X with 'memory' attributes, (Luenberger, 1979): the Markov property essentially applies to the specification and not to the system itself. Yet, the smaller, non-Markovian representation may be preferred for practical reasons in some domains.

In terms of influence, we can say that at each time point $t \in T$, $t < N$, the set $X[t+1, N]$ is conditionally uninfluenced by decisions prior to time point t given $X(t)$. As the POMDP formalism does however not specify control relations below the level of complete system states, we cannot make statements about the effect of decisions on individual state variables without inspecting large numbers of transition probabilities. For instance, we cannot tell whether a state variable $x(t) \in X(t)$ is

unaffected, directly affected, or indirectly affected by the decision at time point $t-1$. The same holds for accessibility relations, properties of progression and absorption, and the existence of static attributes: these notions may well apply to parts of the system under consideration, but such will not show from the representation.

Observability

Observability in the POMDP $\mathcal{M}^{\mathrm{PO}}$ is specified by the set O of observation functions. Each of these functions has the form

$$o_t : A \to \wp(X), \tag{5.81}$$

where $o_t(a) \subseteq X$ is the set of system attributes that is observable at time point $t \in T$ when choosing action $a \in A$. So, a state variable $x(t)$ is manifest at future time points only when the decision at time point t caused it to be so; neither future nor past decisions can influence its manifestness. In addition, there may be system attributes whose marginal probabilities in the initial state distribution P_0 are degenerate; we consider these variables as case parameters, i.e. as variables that are observed prior to the first decision.

We now translate the above observation functions to an observation model O' for decision basis $\mathcal{B}$ as follows. Let $Y \subseteq X$ be the set of all case parameters, i.e. Y the largest subset of system attributes for which $P_0(\mathrm{c}_Y) = 1$ for some configuration $\mathrm{c}_Y \in \Omega_Y$. For the initial time point $t = 0$, we take

$$o_0(\breve{\alpha}) \;=\; Y(0) \tag{5.82}$$

and for each other time point $t \in T$, $t > 0$, we take

$$o'_t(\alpha) \;=\; Y(0) \cup o_0(\alpha(0)) \cup \cdots \cup o_{t-1}(\alpha(t-1)) \tag{5.83}$$

for all action sequences $\alpha \in \mathcal{A}[0, t-1]$.

We have already discussed that a principal property of case parameters is that they are absolutely manifest at all time points; here, this holds therefore for elements of the set $Y(0)$. Any other state variable $x(t) \in X(T) \setminus Y(0)$ is absolutely manifest at future time points in observation model O' if and only if $x \in o_t(a)$ for all actions $a \in A$. In general, state variable $x(t) \in X(t) \setminus Y(0)$ is manifest at time points greater than t only under decision $d(t)=a$ when $x \in o_t(a)$. Potential manifestness of the variable is determined by checking if there exists an action $a \in A$ such that $x \in o_t(a)$; as there are no time lags between action and observing in the POMDP formalism, no other time points need to be inspected. We deduce that a given choice context φ for time point $t \in T$ is verifiable under observation model O', if for each state variable $x(t') \in X[0, t]$ covered by φ we have $x(t') \in Y(0)$ or $x \in o_t(a)$, where $d(t')=a$ is a decision in φ.

Utility function

In the Markov decision processes described by $\mathcal{M}^{\mathrm{PO}}$, a reward $r_t(S, a)$ received at time point $t \in T$ when the decision maker chooses action $a \in A$ and the system occupies state $S \in \mathrm{dom}(X)$. The objective is to maximise the discounted sum of expected rewards of the entire time frame T; the utility function employed in these processes can therefore be characterised as time-separable additive:

$$u(\sigma, \alpha) \;=\; \sum_{t \in T} \lambda^t \cdot r_t(\sigma(t), \alpha(t)) \tag{5.84}$$

for all histories $(\sigma, \alpha) \in \mathcal{H}(T)$, where λ is the discount factor employed. We recall that the subsequent rewards associated with a given planning history can be regarded as independent factors of income, where the discount factor represents a form of interest (when the rewards represent monetary gains), physical decay, or psychological preference of near over distant future. There are no possibilities in the POMDP representation to use utility functions that are not time-separable additive, or to make an explicit distinction between direct and indirect contributors to utility: each reward function takes the action choice and the complete system state into account.

Solution form

Recall that a decision-making policy for a given POMDP consists of a set $\pi = \{\delta_t^{\mathrm{PO}} \mid t \in T\}$ of decision functions of the form

$$\delta_t^{\mathrm{PO}} : \Psi_t \rightarrow A, \tag{5.85}$$

where Ψ_t is a set of choice contexts for time point t. In Section 4.2, we assumed that each Ψ_t, $t \in T$, comprised all maximally-informative choice contexts for time point t; here we will make the weaker assumption that the elements of each Ψ_t are mutually exclusive, i.e. $\varphi_1 \wedge \varphi_2 \equiv \bot$ for all $\varphi_1, \varphi_2 \in \Psi_t$, $\varphi_1 \neq \varphi_2$. Such a policy is translated to a contingency plan π' for decision basis $\mathcal{B}$ as follows:

$$\pi' \;=\; \{\varphi \rightarrow d(t){=}a \mid t \in T,\, \varphi \in \Psi_t,\, a = \delta_t^{\mathrm{PO}}(\varphi)\}. \tag{5.86}$$

The set of decision rules thus obtained is a proper contingency plan in the sense of Definition 5.54, as the elements of each Ψ_t, $t \in T$, are mutually exclusive and the functional formulation prohibits multiple action choices to be assigned to a single choice context. The plan π' is not necessarily complete, consistent or coherent; for these conditions to hold, the policy π must have additional properties that are derived from Definitions 5.75, 5.71, and 5.73. As these derivations are relatively straightforward, we do not elaborate on these properties here.

Recall that a given choice context φ for time point $t \in T$ is verifiable, when each state variable $x(t') \in X[0, t]$ covered by φ is a case parameter (and $t' = 0$), or is manifest under some decision $d(t'){=}a$ in φ. Therefore, if all choice contexts in $\Psi_0, \ldots, \Psi_N$ obey this restriction, then the plan π' is operational under observation model O'.

One might for instance require, as we did in the previous chapter, that each $\varphi \in \Psi_t$ has the form

$$\varphi \equiv \alpha \wedge \mathrm{c}_Z \tag{5.87}$$

where $\alpha \in \mathcal{A}[0, t-1]$ denotes all previous decisions, and c_Z is a configuration of the set $Z = o'_t(\alpha)$ of all state variables that are observable under α. The decision rules are then based on maximal information for the planning agent as each context $\varphi \in \Psi_t$, $t \in T$ is verifiable under O', but no greater context is also verifiable. As a result, any optimal contingency plan can be expressed in this form. The plan formulation under this requirement is however also the least efficient: there may be smaller, equivalent plans when particular information is not really used in making the decisions. It is possible that this holds for the optimal decision-making policy; unfortunately, such cannot, in general, be determined in advance.

5.5.2 DIDs

Dynamic influence diagrams are graphical representations of sequential decision-making problems under uncertainty. Their attractiveness as a representation formalism stems to a great extent from the ability to depict influential and observational relations between state variables, decision variables, and utility. Formally, a dynamic influence diagram over a set of system attributes X and a time frame T is described by a triple $DID = (G, \Gamma, u)$, where G is a directed graph over $X(T)$, $D(T)$, and a value node v, Γ is a set of probability assessment functions for the variables from $X(T)$, and u is a utility function.

We derive a decision basis $\mathcal{B} = (X, A, T)$ from the diagram by taking $A = \mathrm{dom}(d(0)) \cup \cdots \cup \mathrm{dom}(d(N))$ to be the set available actions; the necessary restriction at each time point $t \in T$ is that action choice $d(t) = a$ is excluded when $a \notin \mathrm{dom}(d(t))$. We will now formulate the models and control and observation and the utility function for decision basis $\mathcal{B} = (X, A, T)$ that correspond to the influence diagram DID.

Control

The general, non-Markovian type of dynamic influence diagram places little restrictions on the possibilities for the planning agent to influence the evolution of the system state; the primary characteristic of the influence-diagram representation is that it takes advantage of probabilistic independencies and has a graphical way of expression influential relations among decision and state variables. From the definition of dynamic influence diagram, it follows that the probabilistic relations expressed by the set of probability assessment functions $\Gamma = \{\gamma_{x(t)} \mid x \in X, t \in T\}$ are translated to a control model $\Xi = \{\xi_\alpha \mid \alpha \in \mathcal{A}(T)\}$ as

$$\xi_\alpha(\sigma) = \prod_{x(t) \in X(T)} \gamma_{x(t)}(\mathrm{c}_{x(t)} \mid \mathrm{c}_{\rho_G(x(t))}) \tag{5.88}$$

for each action sequence $\alpha \in \mathcal{A}(T)$ and state sequence $\sigma \in \mathcal{S}(T)$, where for all $x(t) \in X(T)$, $\mathrm{c}_{x(t)}$ and $\mathrm{c}_{\rho_G(x(t))}$ are the configurations of $x(t)$ and $\rho_G(x(t))$ in planning history $h = (\sigma, \alpha)$, i.e. $h \vdash \mathrm{c}_{x(t)} \wedge \mathrm{c}_{\rho_G(x(t))}$. Note that it is possible that there is no configuration $\mathrm{c}_{\rho_G(x(t))}$ of the set of the parents $\rho_G(x(t))$ of variable $x(t)$ that is compatible with action sequence α; this occurs when there is a decision variable $d(t') \in \rho_G(x(t))$ and $\alpha(t') \notin \mathrm{dom}(d(t'))$. It is then immaterial which probability is assigned to $\xi_\alpha(\sigma)$ as this action choice is excluded when solving the planning problem.

The very strength of the influence-diagram formalism is that we can read off influential relations between decisions and state variables directly from the graph G: a state variable $x(t_1) \in X(T)$ is affected by decision $d(t_2) \in D(T)$ if and only if $d(t_2)$ is an ascendant of $x(t_1)$ in G. If, in addition, $d(t_2)$ is a direct ascendant (i.e. predecessor) of $x(t_1)$ in G, then $x(t_1)$ is directly affected by $d(t_2)$. If for a given system attribute $x \in X$, $x(t)$ does not have decisional ascendants for any time point $t \in T$, then x is uncontrollable. The explicit, graphical representation of these relations provides the possibility to exploit them during problem solving without inspecting large numbers of control probabilities that give rise to them. It is not possible to make accessibility relations, properties of progression and absorption, or the existence of static attributes explicit in influence diagrams.

An additional Markov assumption can be employed in dynamic influence diagrams; the assumption then appears from the graph as

$$\rho_G(x(t)) \subset X(t-1) \cup X(t) \cup \{d(t)\} \tag{5.89}$$

for all $x \in X$ and $t \in T$, $t > 0$. As with POMDPs, it implies that at each time point $t \in T$, $t < N$, the set $X[t+1, N]$ is conditionally uninfluenced by decisions prior to time point t given $X(t)$.

Observability

Observability is a more constrained notion in influence diagrams. Recall that upon making decision $d(t)$, $t \in T$, all predecessors of variable $d(t)$ in graph G are assumed to be manifest. So, the observation function for time point t is induced as

$$o_t(\alpha) \;=\; X[0,t] \cap \rho_G(d(t)) \tag{5.90}$$

for all action sequences $\alpha \in \mathcal{A}[0, t-1]$; variable $d(t)$ cannot have predecessors in $X[t+1, N]$. When $t = 0$, we have $o_0(\check{\alpha}) = \rho_G(d(0))$; the predecessors of decision $d(0)$, necessarily a subset of $X(0)$, therefore act as case parameters. The influence-diagram formalism requires that set of predecessors of any future decision variable $d(t) \in D(T)$ should include all earlier decisions and their predecessors (the 'no-forgetting' property), to assure that all information that was earlier available can also be used for decision $d(t)$. Formally, we have

$$\rho_G(d(t)) \;\supset\; \rho_G(d(t-1)) \tag{5.91}$$

for all $t > 0$, and therefore

$$o_t(\alpha) \supseteq o_{t-1}(\alpha') \tag{5.92}$$

for all action sequences $\alpha \in \mathcal{A}[0, t-1]$ and $\alpha' \in \mathcal{A}[0, t-2]$, where $\alpha \vdash \alpha'$. So, this requirement corresponds to the condition of monotonicity of observation in our framework (Assumption 5.45 on page 138).

Verifiability of a given choice context φ for time point t under the observation model thus obtained is easily determined: the context is verifiable whenever each state variable $x(t')$ covered by φ is an element of $\rho_G(d(t))$. It is not necessary to inspect the action choices in φ; there is no distinction between potential and absolute manifestness of state variables in influence diagrams. This illustrates an important characteristic of the influence-diagram representation of observability that needs mentioning here once more. The manifestness of state variables simply progresses with time and does not depend on the actual actions chosen. We have already discussed this point in Chapter 4, where we have termed this phenomenon *informational symmetry*; it prohibits the representation of information-seeking actions and is therefore a serious restriction.

Utility function

The utility function u employed in influence diagram *DID* is a function of the predecessors of the value node v, i.e.

$$u : \Omega_{\rho_G(v)} \to \mathbb{R}. \tag{5.93}$$

In terms of our framework, the predecessors of node v are direct contributors to utility, whereas all other variables are indirect contributors. Generalisation to marginal and expected utilities proceeds as in Proposition 5.90: the marginal utility $u^m(h)$ of planning history $h \in \mathcal{H}(T)$ equals $u(\mathrm{c}_{\rho_G(v)})$, where $\mathrm{c}_{\rho_G(v)}$ is configuration of the set $\rho_G(v)$ in h, i.e. $h \vdash \mathrm{c}_{\rho_G(v)}$. The expected conditional utility of decision process P given planning expression $\varphi \in \Phi(T)$ equals

$$\tilde{u}_\varphi(P) = \sum_{\mathrm{c}_{\rho_G(v)} \in \Omega_{\rho_G(v)}} P(\mathrm{c}_{\rho_G(v)} \mid \varphi) \cdot u(\mathrm{c}_{\rho_G(v)}), \tag{5.94}$$

if φ is possible in P.

It is well possible that the induced marginal utility function u^m is time-separable, but this cannot be made explicit in the influence-diagram representation of utility; we can only detect it by inspecting large numbers of utility values. In the literature, extensions to the representation have been proposed where time-separability is made explicit through the usage of multiple value nodes; we refer to the paper by Tatman and Shachter (1990) for further details of this extension.

Solution form

A decision-making policy for the dynamic influence diagram takes the form of a set

$$\pi \quad = \quad \{\delta_{d(t)} \mid d(t) \in D(T)\} \tag{5.95}$$

where each $\delta_{d(t)} \in \pi$ is a function

$$\delta_{d(t)} : \Omega_{\rho_G(d(t))} \rightarrow \mathrm{dom}(d(t)) \tag{5.96}$$

that selects an action $a \in \mathrm{dom}(d(t))$ for each possible configuration $\mathrm{c}_{\rho_G(d(t))}$ of the predecessors of $d(t)$ in the graph. To translate this policy to a contingency plan π' for decision basis $\mathcal{B}$, each decision function $\delta_{d(t)} \in \pi$ is converted to a set of decision rules of the form

$$\mathrm{c}_{\rho_G(d(t))} \quad \rightarrow \quad d(t) = a, \tag{5.97}$$

where $\mathrm{c}_{\rho_G(d(t))} \in \Omega_{\rho_G(d(t))}$ and $a = \delta_{d(t)}(\mathrm{c}_{\rho_G(d(t))})$. It is easily seen that the antecedents of these rules are verifiable as all non-decisional predecessors of $d(t)$ are assumed to be manifest at time point t. We conclude that the plan thus obtained consists of operational decision rules and is therefore also operational itself. In fact, we could only obtain non-operational plans if we included non-predecessors in the domain of a decision variable's function.

The contingency plan π' that is obtained from translating a decision-making policy for the influence diagram is complete as it provides, at each time point, a decision rule for each possible combination of past decisions and observations. For the same reason, the plan will also be more than complete as it specifies more decision-making information that is strictly necessary. As a result, the plan is inconsistent and has many decision rules that will never be applied; strict implementation of the plan is impossible. To avoid this redundancy, one could decide to use partial functions in the policy π. To obtain a complete plan, it is sufficient that

$$\mathrm{c}_{\rho_G(d(t-1))} \wedge d(t-1) = a \quad \in \quad \mathrm{cover}(\pi', t), \tag{5.98}$$

when $\mathrm{c}_{\rho_G(d(t-1))} \in \mathrm{cover}(\pi', t)$ and $\pi'_{t-1}(\mathrm{c}_{\rho_G(d(t-1))}) = \{a\}$, for all non-initial time points $t \in T$. So, $\delta_{d(t)}$ need not be defined for all configurations of $\rho_G(d(t))$, but only for those where

$$\mathrm{c}_{\rho_G(d(t+1))} \quad \vdash \quad \mathrm{c}_{\rho_G(d(t-1))} \wedge d(t-1) = a, \tag{5.99}$$

and $\delta_{d(t-1)}(\mathrm{c}_{\rho_G(d(t-1))}) = a$. Omitting any of these choice context renders the plan π' incomplete, while including more choice contexts unavoidably yields inconsistencies. The initial time point is an exception to this rule; decision function $\delta_{d(0)}$ needs to be defined for all configurations of $\rho_G(d(t))$, so that

$$\Omega_{X(0)} \quad \subseteq \quad \mathrm{cover}(\pi', 0) \tag{5.100}$$

and the plan π' is therefore complete.

It is worthwhile to note that the standard form of policy representation in influence diagrams is rather rigid. Decision functions are always expressed in terms of maximal information; when it is possible to make optimal decisions with less information, the formalism does not allow one to exploit this circumstance to arrive at a smaller representation; the true complexity of a given policy can only be determined after extensive analysis.

5.5.3 Comparison

We conclude this section with a brief summary and comparison of our findings with respect to the POMDPs and DID representations. We have shown that our framework is more general than both POMDPs and DIDs in the sense that it allows to reformulate domain specifications in either type of representation. Note this also holds for fully-observable Markov decision processes (FOMDPs), as this type of representation is essentially a restricted form of POMDP; we have not elaborated on FOMDPs as we regard partial observability as an essential ingredient to action planning under uncertainty. Another point to note is that both POMDPs and DIDs also serve as computational architectures for problem solving, and several characteristics of these formalisms have computational motivations. Yet, as we currently do not consider computational aspects, we have excluded them from the discussion and have restricted ourselves to considering representational issues.

POMDPs incorporate a strong notion of temporal locality in the representation of control, observation, and utility: the Markov assumption exists in their notion of control, state observation proceeds instantaneous without time lags, and the utility function is time-separable and additive. Yet, POMDPs also have a rather coarse representation of control and utility as they do not specify control relations or reward functions below the level of complete system states. There is a significant potential for more economic representations at these points.

DIDs make less global assumptions than POMDPs and arrive at a compact representation by making the various relations among domain variables explicit in a directed graph. As such, the representation of control and utility is more fine-grained and provides for reading off relations of direct and indirect affectedness, and of direct and indirect contribution to utility. A drawback to the graphical representation is that it is unable to handle informational asymmetry, and therefore has problems to model test decisions. Furthermore, the standard form of decision-making policies for DIDs does not allow for representational savings when decisions can be made on the basis of less than maximal information.

Given the above translations of POMDPs and DIDs to our framework, we can identify the problem types for which either type of representation can be used. When a decision-theoretic planning problem has Markovian control and observation, and

a time-separable additive utility function, then it can be expressed as a POMDP. When a problem has a symmetrical observation model, it can be expressed as a DID. When all these conditions apply, the problem can be cast both as a POMDP and a (Markovian) DID.

5.6 Discussion

In this section we evaluate our framework for decision-theoretic planning and relate it to the work of Poole (1993; 1996) and Haddawy (1996).

We have presented a formal framework that integrates notions from Bayesian probability theory, decision theory, and symbolic planning. The building blocks of our framework are the elements of a decision basis: a specification of a dynamic system in terms of its attributes and possible values, the possible actions to choose from, and the time frame for a planning task under consideration. From these elements we derived a Boolean planning language for the given specification, and defined the notion of decision process as a probability distribution over planning expressions. We believe that the strength of our framework is that given these definitions, we can express most concepts that occur in decision-theoretic planning. We have described how to formalise a large number of concepts related to stochastic control, partial observability, contingency planning, and decision-theoretic planning objectives. As was shown in Section 5.5, the framework allows to describe and analyse existing representation formalisms such as POMDPs and DIDs.

In our framework, decision processes provide a meta-level description of the intertwined dynamics of planning agent and stochastic system. Using the rules of conditional probability, we can switch from this description to object levels that focus on the agent's behaviour (using choice predictions) or on the system's reactions to that behaviour (using state predictions). The approach we have taken can be regarded as strictly Bayesian: the collection of marginal probabilities on initial states serves as a prior distribution, and observations and the (expected) effects of action choices are used to repeatedly update this distribution. It is mainly a matter of convenience that decision processes integrate all these distributions over time: this type of formalisation supports a smooth analysis of the relations between subsequent distributions.

Although the chapter has focused on an integrated approach to decision-theoretic planning, most ingredients of our framework could well be used in isolation, or in combination with other formalisms. For instance, the elaborate theory of contingency planning that was presented in Section 5.3 is valuable in its own right and has a potential for application in combinations with other formalisms for reasoning with uncertainty and incomplete information.

We note that it is not necessary to employ our framework in its full generality: it is

also suited to analyse and compare planning tasks in more restricted circumstances. For instance, one may consider planning with deterministic system dynamics but an unknown initial state, or planning to reach a goal state in a stochastic environment without any observations. Another possibility is to assume that the planning agent does not have a correct conception of reality and reasons from erroneous models of control or observation, or, more radically, that he does not have such models at all. When a (correct) model of control is lacking, the agent's assignment can then be regarded as a *reinforcement learning* task (Kaelbling et al., 1996; Sutton and Barto, 1998). When a (correct) model of observation is lacking, prospective planning by formulating decision rules is no longer useful as the operationality of these rules is uncertain; the only resort for the planning agent is performing real-time planning, where the choice of actions and their execution coincide.

The framework of D. Poole

D. Poole (1993; 1996) has developed a framework that combines notions from probability theory, decision theory and symbolic planning. The framework aims at symbolic reasoning over contingency plans in uncertain environments, where the objective is to maximise expected utility. The main differences with our work are that Poole (i) uses acyclic Horn clauses to specify the effects of actions on states, (ii) borrows the notion of *situation* from the situation calculus (McCarthy and Hayes, 1969) to specify action sequences, and (iii) introduces uncertainty by setting the truth of certain combinations of logical atoms externally through a stochastic mechanism. The result of such uncertainty is that situations (i.e. action sequences) do not uniquely determine the resulting state (as is traditionally the case in the situation calculus), but yields a probability distribution on states instead.

The usage of Horn clauses to specify the effects of actions builds on the conception that action effects are in principle deterministic and known to the planning agent, but there may be conditions in the effect description that are subject to uncertainty and whose truths are moreover unknown to the planning agent. Using a frame axiom, persistence of all facts not mentioned in the effect description is assumed. The Markov assumption is implicitly present in the semantics of effect descriptions. The specification of utility values is also structured into Horn clauses, and allows for various types of decomposition (e.g. additivity and multiplicativity). A contingency plan has the form of an algorithmic procedure with sequencing and conditioning operators. Such a plan selects a situation (i.e. sequence of actions), depending on the observations in due course. A more general form of plan that includes loops, nondeterministic choice, and procedural abstraction is possible.

The framework of P. Haddawy

P. Haddawy (1996) has developed a first-order logical language called $\mathcal{L}_{\text{tca}}$ that integrates earlier work on temporal logics, probability logics, and action modelling. This language is not intended as a computational architecture (it is, in fact, not

finitely axiomatizable), but serves as a formal foundation and analytical framework for planning systems.

Haddawy's language is characterised by a high level of expressiveness, which is already reflected at the most basic level. In the description of a planning domain, an explicit distinction is made between facts and events, and various types of interaction between these entities are possible. The underlying notion of time is continuous, and temporal reasoning is not based on time points but on time intervals. Furthermore, the language allows for concurrent actions.

The notions of probability and possibility (and its pendant inevitability) are formalised by modal operators with an indication of time. Possibility is defined independently of probability, for the reason that in an uncountably infinite outcome space, events can have zero probability but are still possible. The choice of formalisation of these notions has important semantic consequences. First, probabilistic expressions can be nested and combined with possibility and inevitability. The nesting of probabilistic expressions requires higher-order probabilistic semantics. This is resolved by making the assumption that probability equals the expected value of a distribution over probability distributions. Second, Haddawy's language incorporates a dynamic perspective on time as each of these modal operators incorporates a reference to a time point. Thus, the probability and possibility of facts and events to occur are relative to the point in time that represents 'now'.

The language comes equipped with an elaborate model-theoretic semantics; a number of intuitively appealing properties are obtained by placing constraints on the models. For instance, it is ensured that the past is inevitable and cannot be influenced by actions or events. Furthermore, when actions are chosen, this results in attempts, which may or may not succeed. A successful attempt yields one of the potential effects of the action; each potential effect is described indirectly in the language by a secondary event that induces the effect upon successful action completion. The reasoning about effects of actions is therefore highly qualitative in nature.

Conclusion

To conclude we provide a brief comparison of the three frameworks. From a conceptual point of view, our framework is most closely related to the framework of D. Poole, as they are both founded on the conception of a planning agent that is controlling and observing a dynamic system over time. This is the same conception that underlies the POMDP representation formalism. Both frameworks are further characterised by a static perspective on time, and a considerable attention to the notion of contingency plan. Poole's framework is however, in terms of expressiveness, also comparable to POMDPs, whereas our work explicitly seeks more expressiveness: we believe that the assumptions underlying POMDPs can be too restrictive, and should be not be incorporated in a general framework. Additionally, the framework of Poole employs a large variety in formal structures that jointly provide for succinct symbolic representation of decision-theoretic planning notions. But the disadvantage

of this approach is that it obscures rather than illuminates a theoretical analysis: it allows for switching between various types of representation, but leaves the underlying, integrating perspective unaddressed. Our framework in contrast highlights this perspective through a unity in formalisation.

The similarity between our framework and P. Haddawy's logical language $\mathcal{L}_{\text{tca}}$ is that both are characterised by a high level of uniformity in the formalisation with the objective of supporting a clear analysis and insight to the type of problem that is being studied. An important difference is that Haddawy's work is rooted in the traditional, logical-analytical approach to formalising planning problems. An asset of the logical language $\mathcal{L}_{\text{tca}}$ is that it is highly expressive: it incorporates uncountably infinite outcome spaces, a distinction between facts and events, higher-order probabilities, a dynamic perspective on time, and a notion of possibility that is independent of probability. However, there is no concept of information or observation in the language. Therefore, Haddawy's notion of plan is rather naive: it simply consists of a collection of action attempts, with the objective of achieving a propositional goal; the concept of planning for contingencies is not included. Similarly, the concept of planning objective is weak: it is assumed that a planning agent aims at reaching a propositional goal, while utility-theoretic notions are lacking. We believe that these notions are indispensable in domains where uncertainty truly abounds; they are therefore prominent in our framework. It is questionable whether the language $\mathcal{L}_{\text{tca}}$ is truly suited for analysing and founding decision-theoretic planning systems: higher-order probabilities and uncountably infinite outcome spaces are superfluous for this task, while the notions of observation, contingency plan, and expected utility are most important.

CHAPTER 6

Clinical modelling and reasoning

In this chapter we return to the starting-point of our study, decision making in medical care. It is discussed how the framework for decision-theoretic planning that was developed in the previous chapter is applied to model medical-clinical domains, and to implement patient management tasks.

The chapter is organised as follows. Section 6.1 describes how the components of our framework can be used to develop a model of given clinical domain. To illustrate our ideas, we will use examples from the application domain that was described in Chapter 2, the domain of paediatric cardiology. In previous chapters we have used examples of the disorder VSD to illustrate various parts of the theory, but these examples have often been small and greatly simplified representations of reality; here, we will use more comprehensive and realistic examples. We recall that for the reader's convenience, a glossary of specialist terms from the VSD domain is found in Appendix A.

Section 6.2 focuses on the medical reasoning tasks of diagnosis, therapy planning, and prognosis. We investigate how these tasks can be approached within our framework, and how they relate to each other in the dynamic perspective on patient management that was sketched in Chapter 1. A number of concepts that occur in everyday clinical reasoning, such as diagnostic hypothesis, clinical indication, and future scenario, are formalised in the framework and their role in clinical reasoning is analysed. We conclude the chapter in Section 6.3 with a discussion of our findings.

6.1 Building a domain model

In general terms, the projection of a given clinical domain to our framework's components is envisioned as follows. In a decision basis $\mathcal{B} = (X, A, T)$, the system attributes of the set X describe the condition of a patient that requires or undergoes clinical management. The set A of actions comprises the diagnostic and therapeutic modalities that are available to the treating clinician. The time frame T finally represents the space of time for the task at hand, and the potential moments of clinical action. Describing these sets is typically the first phase in the formalisation of a given domain; we discuss it thoroughly in Subsections 6.1.1 and 6.1.2.

The second phase will often consist of specifying the domain's control and observation models. A model of control Ξ has basically two functions: it specifies intra-state relations between attributes (such as the probabilities of making particular observations) and the developments of the patient's condition over time under different therapeutic regimens (such as the likelihood of disease progression and complications). A model of observation O specifies which attributes of the patient's condition can be inspected by performing diagnostic tests, and the time scale at which these tests yield the evidence longed for. The specification of control and observation is discussed in Subsections 6.1.3 and 6.1.4.

The third and final phase concerns modelling the objectives of patient management with a utility function; optimal decisions in varying clinical circumstances can then be derived from this function. In Subsection 6.1.5, we describe the possibilities one has in formalising the objectives of management with a utility function.

6.1.1 Describing the patient's condition

As stated above the attributes from the set X jointly describe the condition of a patient that requires or undergoes clinical management. The set Ω_X of all possible configurations of X should therefore cover the range of all possible patient conditions that are considered relevant in this respect. Instead of *system state* we will rather speak of *clinical state* in this chapter, to emphasise the role of state variables. We take the clinical state description provided by the set X to include a description of disease in terms of both its underlying, internal pathophysiology (i.e. disease and complications) and its external manifestations (i.e. signs, symptoms, and test results). Therefore, the set X is assumed to be composed of a set I of *internal* and a set E of *external attributes*, i.e.

$$X \;=\; I \cup E, \tag{6.1}$$

where the sets I and E are disjoint. The set I represents the internal, pathophysiological state of a patient: these attributes determine the patient's health in present and future and are therefore of principal interest to the treating physician. The set

E represents manifestations of the internal state and consists of attributes that are (at least potentially) observable.

Example 6.1 *In order to find a suitable description of the VSD domain in terms of a set of clinical attributes, the following procedure was employed. First, the field expert was interviewed to elicit the main terms and concepts of the problem domain. Using this informal domain description, we discerned seven groups of attributes:*

1. VSD pathology*, describing the primary pathological condition,*
2. cardiac complications*, i.e. additional abnormalities of the heart and main arteries,*
3. physiology*, pertaining to the haemodynamics of the circulatory system, such as blood pressures, shunting, and vascular resistances,*
4. pulmonary complications *resulting from haemodynamic abnormalities,*
5. clinical complications *resulting from diagnostic or therapeutic procedures,*
6. signs and symptoms *obtained from history taking and physical examination of the patient, and finally*
7. clinical measurements *obtained from diagnostic investigations.*

Here, groups 1–5 jointly constitute the set of I of internal attributes, and groups 6 and 7 jointly constitute the set E of external attributes. The second step consisted of searching through the literature on congenital heart disease and VSD, which brought a number of refinements to the pathophysiological part of the attribute set, i.e. groups 1–4. The third and final step was to study medical records of VSD patients that were treated during the years 1981–1996 at the Leiden University Medical Center; these records were then also collected in a database. This provided a thorough inventory of the types of information on VSD patients that is available to the clinician, and yielded refinements to groups 6 and 7.

Table 6.1 lists all internal attributes from the VSD domain. The attributes in group 6 (signs and symptoms) are observable through physical examination of the patient; they are shown in Table 6.4 on page 192. The attributes of group 7 (clinical measurements) are observable through a variety of diagnostic procedures; they are listed in Table 6.5 on page 194.

When all clinical attributes have been identified, the next task is to establish their value domains. As our framework assumes all value domains to be finite, this requires (finite) discretization of continuous-valued attributes. In cases where this is done, it is recommendable to keep track of the fact that the underlying concept is non-discrete, and furthermore that the values of such an attribute, both before and after

Group	Name	Description
1 *VSD pathology*	`outlet_pos`	malalignment of the outlet septum
	`VSD_type`	type (location) of the VSD
	`VSD_ext`	extension of the VSD
	`VSD_size`	size of the VSD
2 *cardiac complications*	`ASD`	atrium septum defect
	`PDA`	persistent ductus arteriosus
	`aort_sten`	aortic stenosis
	`aort_prolapse`	aortic prolapse
	`pulm_sten`	pulmonary stenosis
3 *physiology*	`resis`	pulmonary vascular resistance
	`shunt`	direction and size of shunting
	`ventr_press`	interventricular pressure ratio
	`art_press`	interarterial pressure ratio
	`aort_regurg`	aortic valve regurgitation
	`pulm_regurg`	pulmonary valve regurgitation
	`pulm_flow_hyp`	pulmonary flow-hypertension
	`cardiomegaly`	cardiomegaly
	`LVH`	left-ventricular hypertrophy
	`RVH`	right-ventricular hypertrophy
	`ox_sat`	oxygen saturation of the blood
4 *pulmonary complications*	`pulm_inf`	pulmonary infections
	`pulm_art`	pulmonary arteriopathy
5 *clinical complications*	`stroke`	stroke
	`bleeding`	bleeding
	`arrhythmia`	arrhythmia
	`perforation`	perforation of the heart or great vessels
	`death`	death

Table 6.1: Internal attributes for the VSD domain.

Name: `VSD_type`

Definition: Type (location) of the VSD

Group: 1 (VSD pathology)

Domain:

perimembranous	(adjacent to triscuspid and mitral valve, with possible extensions into the inlet and/or outlet area)
subaortic	(located in the outlet area just beneath the aortic valve, non-perimembranous)
muscular	(surrounded by muscular tissue, non-perimembranous, non-subaortic)

Type: unordered categorical

Comments: We exclude the rare subpulmonary and doubly-committed (= both subaortic and subpulmonary) VSDs. Note that when `VSD_size` = *null*, apparently there is no VSD present (anymore), and this variable is meaningless.

Figure 6.1: Model description of the attribute `VSD_type`.

discretisation, permit a linear ordering. Furthermore, one should preferably postpone the discretisation step as long as possible, and once it is carried out, it should be well-documented and possibly verified after the model is completed.

Example 6.2 *For each attribute in the VSD domain, we established a value domain in cooperation with the field expert. An example description, for the attribute* `VSD_type`*, is given in Figure 6.1. Several attributes (e.g.* `VSD_size`*,* `shunt`*, and* `ox_sat`*) are continuous-valued in their natural form. However, we found that there existed commonly-used discretizations for such attributes in clinical practice. Where possible we used these discretisations, thus staying as close as possible to the clinician's vocabulary. An example is provided in Figure 6.2, which describes the attribute* `VSD_size`*; this attribute has an age-dependent discretization.*

6.1.2 Clinical modalities

We now turn to the modalities for the clinician in managing patients. These are classified along two dimensions. First, the actions that the clinician may choose from must be distinguished; this yields the set A of the decision basis. Second, the moments in time where such actions may be performed have to be established, providing the time frame T.

Diagnostic and therapeutic actions

We now turn the second element of a decision basis, the action set A. This set should comprise all diagnostic and therapeutic modalities that are available to the

Name: `VSD_size`

Definition: Largest diameter of the VSD

Group: 1 (VSD pathology)

Domain:

null	(no defect)
small	(1–3 mm at 0–6 months, 1–5 mm thereafter)
moderate	(3–5 mm at 0–6 months, 5–10 mm thereafter)
large	($>$ 5 mm at 0–6 months, $>$ 10 mm thereafter)

Type: ordered discretized

Comments: A perimembranous VSD without extension(s) is by definition small; otherwise, it is moderately large or large.

Figure 6.2: Model description of the attribute `VSD_size`.

treating clinician, including the most basic modality in all circumstances, the *absence* of taking action. We take the set A therefore to be composed of three disjoint sets, comprising *test actions*, *treatment actions*, and a single *empty action*:

$$A \;=\; A_{\text{test}} \cup A_{\text{treat}} \cup \{\varepsilon\}. \tag{6.2}$$

Test actions are characterised by the fact that they yield the opportunity to observe one or more external state variables. Treatment actions are distinguished by their potential to produce changes in the clinical state of the patient. Theoretically speaking, the distinction may become blurred as test actions might also affect the patient's state and treatment actions sometimes provide observations. It is nevertheless useful to maintain the distinction between these action types as they are conceptually very different, and this allows for reasoning with the *purpose* of action choices. Furthermore, one will rarely have difficulties with distinguishing test actions from treatment actions in a given clinical domain.

The empty action ε denotes the absence of taking action; we use this action to skip decision moments when future times seem more appropriate for activity, or to 'fill up' the remaining points in a time frame when our objectives are already met at an early stage. The empty action does neither yield observations nor induce state changes. We note that a result of our assumptions is that no action exists that yields observations on internal variables.[1]

Example 6.3 *For the VSD model, we have identified 7 test actions and 2 treatment actions; they are listed in Table 6.2. The first test action, physical examination of the patient provides for observing all clinical attributes from group 6 (signs and symptoms). The other test actions represent diagnostic procedures that are more involved;*

[1]This is not really a restriction as one could always add an external variable that represents the outcome of a perfect test on one of the internal variables.

Type	Action	Description
test	*exam*	physical examination
	ECG	electrocardiography
	echo	echocardiography
	Xray	chest roentgenography
	bltest	blood test
	cath	cardiac catheterisation
	biop	open lung biopsy
treat	*med*	medical treatment
	surg	cardiac surgery

Table 6.2: Clinical actions for the VSD domain.

each of them yields observations for attributes of group 7 (clinical measurements). The available treatment modalities are medical treatment (reducing the symptoms of heart failure with cardiac glycosides and diuretics), and surgical repair of the cardiac lesion.

When test actions change the patient's state this is generally due to adverse (and unintended) effects: some diagnostic procedures, typically those that are invasive, may cause complications or even death. There are basically two options for incorporating such effects in a decision-theoretic model. The first option is to model these effects explicitly by having clinical attributes that describe potential complications. These attributes are then directly (though nondeterministically) affected by decisions to perform such actions, while the attributes themselves are direct (and negative) contributors to utility. As a result, the decisions are indirect contributors to utility. The second option leaves the adverse effects of test actions implicit and has the decisions as direct contributors to utility. This option will generally yield a simpler model (as it has less clinical attributes), but this model is also less expressive and less transparent. A mixture of both options was employed in the VSD model.

Example 6.4 *In the current formalisation of the VSD domain, the only test action that affects the patient's state is cardiac catheterisation: it may cause clinical complications such as stroke, perforation of the heart or great vessels, and even death. The other test actions do not affect the patient's clinical state in the model. This is not fully realistic for chest roentgenography and open lung biopsies: roentgenography exposes the patient to a small amount noxious x-rays and an open lung biopsy removes a small specimen of tissue from the lungs. There is widespread consensus that these investigations should be avoided if possible, and otherwise be minimised in the number of times they are employed. In both cases we have therefore chosen to let these actions contribute directly, and negatively, to utility.*

Time frame and temporal map

To complete the specification of a decision basis, one needs to assess the space of time for the task at hand, and the potential moments of clinical action. This information is then translated to a time frame T; as before we will assume time frames to be finite, and to have the form $T = \{0, 1, \ldots, N\} \subset \mathbb{N}$. The number N, which represents the maximum number of actions chosen, should be chosen sufficiently large to provide for accomplishing the medical task at hand.[2] Furthermore, one should be aware of the envisioned range of operational time that planning task takes, and the actual periods of time between subsequent decision moments. We formalise this by associating a *temporal map* with a time frame T of decision moments.

Definition 6.5 (Temporal map) *A* temporal map *for time frame T is a monotonically non-decreasing function $\tau : T \to \mathbb{R}$.*

A temporal map τ allows for projecting the decision moments in T onto arbitrarily distributed points on the real line, where the real numbers are in turn interpreted on a linear time scale (e.g., minutes, days, years). We assume that the function τ is non-decreasing, i.e. if $t_1, t_2 \in T$ and $t_1 < t_2$, then also $\tau(t_1) \leq \tau(t_2)$. The difference between $\tau(t)$ and $\tau(t+1)$, $t \in T$, $t < N$, now represents the actual period of time between time points t and $t+1$; we refer to these time periods as *episodes*. We allow that episodes differ considerably in size, thus lifting the implicit assumption that decision moments are uniformly spread across time.

We will often use $\tau(t)$ to represent the age of a patient at decision moment $t \in T$; this has also been our modelling choice in the VSD domain.

Example 6.6 *A paediatric cardiologist typically wants to see a VSD patient a number of times during childhood to monitor the patient's development. After the first visit where the VSD is diagnosed (usually in the first 3 months of life), this is often at 6 months, 12 months, and 24 months after birth, and, if necessary, at the ages of 4 and 8 years. In a naive formalisation of this domain we would therefore use a time frame with 6 decision moments (i.e. $N = 6$), and use $\tau(t)$, $t \in T$, to represent the number of months since the patient's birth, i.e. $\tau(0) = 3$, $\tau(1) = 6$, $\tau(2) = 12$, $\tau(3) = 24$, $\tau(4) = 48$, $\tau(5) = 96$. This formalisation has however the disadvantage that we can only choose one action at each consultation. In reality, multiple diagnostic actions and possibly a treatment action are performed each time.*

To overcome the problem in the above example, we can use a special construction where $\tau(t) = \tau(t+1)$ for one or more decision moments $t \in T$. This means that the passage of time between decision moments t and $t+1$ is considered negligible in terms

[2] Recall that the final decision moment in a time frame is virtually meaningless as we cannot model its future consequences. This moment serves to evaluate the post-therapeutic state of the patient.

of the model and is therefore abstracted from. Yet, the precedence relation between these moments, as provided by the ordering on time frame T, remains: the decision at time point t is made before the decision at time point $t+1$.

Example 6.7 *A more realistic formalisation of the VSD domain has 8 decision moments at each of the ages where the cardiologist may see the patient. That is,* $N = 48$*, where*

$$\begin{array}{rcccccl} \tau(0) & = & \ldots & = & \tau(7) & = & 3, \\ \tau(8) & = & \ldots & = & \tau(15) & = & 6, \\ \vdots & & & & & & \\ \tau(40) & = & \ldots & = & \tau(47) & = & 96. \end{array}$$

This allows to perform the maximum of 7 diagnostic actions and a single treatment action at each age; the possibility to perform multiple treatment actions is not required as the treatments are mutually exclusive. The outcome of the decision process is evaluated also at the age of 8 years, i.e. $\tau(48) = 96$*.*

Given a decision basis $\mathcal{B} = (X, A, T)$, elements of the planning language $\Phi(T)$ for that basis describe situations, events, and scenarios that may occur in the clinical management of patients. Decision processes provides an integrated perspective on the management process, by describing the relationships between them. It describes how the internal and external attributes of a patient's clinical state are related, how the treating physician responds to observed external manifestations, and how in turn the patient's internal state develops over time in response to this treatment. We recall from the previous chapter that decision processes are thought as being derived from specifications of the temporal, informational, and probabilistic relations between system (patient) and planning agent (physician). In the next subsection, we consider the models of control and observation, that specify effects of the physician's actions on the patient's development and on his own informational situation.

6.1.3 Prognostic models of disease

After the basic ingredients of a problem domain have been formalised in a decision basis, one needs to construct models of control and observation. These models are crucial for virtually all reasoning tasks one might wish to implement, and their roles in clinical problem solving are also, to some extent, intertwined. The latter circumstance is due to the fact that we can identify three aspects of control models in clinical settings. The first aspect concerns the *target population* of the envisioned application: the model of control characterises this population. The second aspect concerns the development of disease and complications over time; here, temporal relations between internal clinical states are predominant. The third aspect, finally, concerns diagnostic

aspects of the problem domain, as cast in the probabilistic relations between internal and external variables.

The first and second aspects of control models are basically prognostic in nature; these aspects are discussed in this subsection. In the third aspect, a model of control interacts with the employed notion of observability; this aspect is discussed in Subsection 6.1.4, in combination with the role of observation models in clinical specifications. Note that from all formal standpoint, all three aspects concern *probabilistic* properties of the domain at hand; this is the reason that in our framework, a single model incorporates these conceptually rather dissimilar aspects.

Recall from Section 5.2 that formally, a model of control for a given decision basis $\mathcal{B} = (X, A, T)$ is a set $\Xi = \{\xi_\alpha \mid \alpha \in \mathcal{A}(T)\}$ of probability distributions on $X(T)$, with the restriction that

$$\xi_{\alpha_1}(\mathrm{C}_{X[0,t]}) \;=\; \xi_{\alpha_2}(\mathrm{C}_{X[0,t]}) \tag{6.3}$$

for all time points $t \in T$ and all action sequences $\alpha_1, \alpha_2 \in \mathcal{A}(T)$ that are identical over time segment $[0, t-1]$. In words: action choices cannot change the past. As in the previous chapter, we use ξ^t_α as a shorthand notation for the unique marginal probability distribution on $X[0,t]$ under action sequence $\alpha \in \mathcal{A}[0,t]$, given model Ξ. When a temporal map τ is employed to model the operational time for the management task at hand, a model of control should preferably be expressed in terms of the image of time frame T under τ instead of T itself; we do not elaborate on this issue here.

Target population

Given a model of control Ξ, the probability distribution on initial clinical states is described by $\xi^0_{\breve{\alpha}}$, where $\breve{\alpha} \in \mathcal{A}(\varnothing)$ is the unique action sequence over the empty time segment. So, $\xi^0_{\breve{\alpha}}(I(0) = S)$ is the probability that a patient has internal state $S \in \mathrm{dom}(I)$ at the start of the decision process. The distribution $\xi^0_{\breve{\alpha}}$ thus describes the characteristics of the *target population* of patients we wish to manage: it incorporates assumptions with respect to delimitation of the problem domain, restrictions stemming from choices in the formalisation, and domain knowledge in form of figures on the prevalence of diseases and complications, prior to clinical management.

Delimitations of the problem domain are normally categorical and will therefore induce extreme probabilities in the distribution $\xi^0_{\breve{\alpha}}$.

Example 6.8 *It was noted in Chapter 2 that the improved echocardiographic imaging of recent years techniques have virtually solved the problem of correctly diagnosing a VSD. A global assumption in our model is therefore that VSD is the primary and correct diagnosis at the start of the decision process. That is,*

$$\xi^0_{\breve{\alpha}}(\texttt{VSD\_size}(0) = \mathit{null}) \;=\; 0,$$

or equally, each of the individuals in the target population has a VSD at time point $t = 0$*. Furthermore,*

$$\begin{aligned} \xi^0_{\breve{\alpha}}(\texttt{outlet\_pos}(0) = \textit{severe\_right}) &= 0, \\ \xi^0_{\breve{\alpha}}(\texttt{outlet\_pos}(0) = \textit{very\_severe\_right}) &= 0, \end{aligned}$$

as severe malalignments of the infundibular septum indicate the existence of Tetralogy of Fallot. This would invalidate VSD being the primary diagnosis.

Also restrictions stemming from choices in the formalisation often result in extreme probabilities.

Example 6.9 *With a perimembranous VSD, the small membranous component of the ventricular septum is completely lacking. In addition, muscular parts of the septum may fail to exist when that VSD is moderate or large in size. Depending on the location of the missing parts, the VSD is then said to* extend *in the direction of the inlet or outlet areas, or both. In the VSD model, we use the internal attribute* `VSD_ext` *to describe these extensions. They can only occur with perimembranous VSDs, not subaortic or muscular ones:*

$$\begin{aligned} \xi^0_{\breve{\alpha}}(\texttt{VSD\_ext}(0) = \textit{none} \mid \texttt{VSD\_type}(0) = \textit{subaortic}) &= 1, \\ \xi^0_{\breve{\alpha}}(\texttt{VSD\_ext}(0) = \textit{none} \mid \texttt{VSD\_type}(0) = \textit{muscular}) &= 1. \end{aligned}$$

Furthermore, a perimembranous VSD without extension(s) is by definition small, and with extension(s), it is moderately large or large:

$$\begin{aligned} \xi^0_{\breve{\alpha}}(\texttt{VSD\_size}(0) = \textit{small} \mid \texttt{VSD\_ext}(0) = \textit{none}, & \\ \texttt{VSD\_type}(0) = \textit{perimembranous}) &= 1, \\ \xi^0_{\breve{\alpha}}(\texttt{VSD\_size}(0) = \textit{small} \mid \texttt{VSD\_ext}(0) \neq \textit{none}, & \\ \texttt{VSD\_type}(0) = \textit{perimembranous}) &= 0. \end{aligned}$$

A final example illustrates the role of figures on the prevalence on diseases and complications, prior to clinical management, in the probability distribution on initial states.

Example 6.10 *The literature on VSD reports the following prevalence of different VSD types:*

$$\begin{aligned} \xi^0_{\breve{\alpha}}(\texttt{VSD\_type}(0) = \textit{perimembranous}) &= 0.85, \\ \xi^0_{\breve{\alpha}}(\texttt{VSD\_type}(0) = \textit{subaortic}) &= 0.05, \quad \text{and} \\ \xi^0_{\breve{\alpha}}(\texttt{VSD\_type}(0) = \textit{muscular}) &= 0.10. \end{aligned}$$

We also know that 80% of all VSDs is not accompanied by cardiac complications:

$$\begin{aligned} \xi^0_{\breve{\alpha}}(\texttt{ASD}(0) = \textit{none}, \texttt{PDA}(0) = \textit{none}, \texttt{aort\_sten}(0) = \textit{none}, & \\ \texttt{pulm\_sten}(0) = \textit{none}) &= 0.80. \end{aligned}$$

Characteristics of the target population, as described by ξ^0_α, can be regarded as static prognostic features of the domain: they allow us to make predictions about individuals in the population without having any further information about them. We now turn to dynamic prognostic features of the problem domain, as described by the temporal relations between internal states.

Natural history

The developments in patients' internal states over time are described by probability distributions on the set $I(T)$; these distributions are also induced by the control model Ξ. For instance, given an initial internal state $S_1 \in \mathrm{dom}(I)$ and a specific choice of action sequence $\alpha \in \mathcal{A}(T)$, the probability of reaching internal state $S_2 \in \mathrm{dom}(I)$ at the final time point $t = N$ equals $\xi_\alpha(I(N) = S_2 \mid I(0) = S_1)$. We first discuss the special case where action sequence α incorporates the empty action choice at all time points, i.e. $\alpha(t) = \varepsilon$ for all $t \in T$. This action sequence denotes the absence of all action, and the associated probability distribution can therefore be taken to model the *natural history* of disease. The effects of other action choices, in particular (treatment) actions aimed at altering the patient's development, are discussed thereafter.

Example 6.11 *The natural history of VSD patients is often mild, as the majority ($\pm$ 70%) of defects closes spontaneously due to tissue growth:*

$$\xi_\alpha(\texttt{VSD\_size}(N) = \mathit{null} \mid \texttt{VSD\_size}(0) \neq \mathit{null}) \;=\; 0.70$$

if α denotes the absence of all action. The majority (54%) of defects even closes in the first two years of life:

$$\xi_\alpha(\texttt{VSD\_size}(t) = \mathit{null} \mid \texttt{VSD\_size}(0) \neq \mathit{null}) \;=\; 0.54$$

if α denotes the absence of all action and $\tau(t) = 24$ (months).

In general, if changes may occur to internal attributes in the absence of action, we say that the patient's clinical state is subject to *endogenous changes*.

Definition 6.12 (Endogenous change) *Let $\alpha \in \mathcal{A}(T)$ be the action sequence that denotes absence of all action, i.e. $\alpha(t) = \varepsilon$ for all $t \in T$. The clinical state described by the set X is subject to* endogenous changes *when*

$$\xi_\alpha(I(t) = S \mid I(0) = S) \;\neq\; 1 \tag{6.4}$$

for some time point $t \in T$ and internal state $S \in \mathrm{dom}(I)$.

If the clinical state is not subject to endogenous changes, the planning agent holds the initiative for state transitions, as all internal attributes remain static in the absence of action. This circumstance will facilitate the planning task, but unfortunately, it

Action	Precondition attributes		Postcondition attributes	
	decisive	*additional*	*target*	*complicatory*
med	shunt resis	–	LV_failure	–
surg	outlet_pos VSD_type VSD_size resis	LV_failure RV_failure pulm_inf	outlet_pos VSD_size ASD PDA aort_prolapse	bleeding arrhythmia death

Table 6.3: VSD treatments with pre- and postcondition attributes.

is rather rare in medical domains. The human body can usually not be regarded as a passive entity that patiently undergoes a doctor's acts and remains otherwise unaltered; the spontaneous closure of VSDs as described in Example 6.11 provides an illustration.

The natural history in medical domains is often characterised by a mixture of static and non-static internal attributes, where state dynamics are further restricted by progression and absorption of one or more non-static attributes.

Example 6.13 *The type of a VSD is fixed as the location of a given defect cannot change; the attribute* VSD_type *is therefore static. As described above, however, the size of a VSD will often decrease, and an increase in size is impossible; so, the attribute* VSD_size *is progressive. The presence of a subaortic VSD may cause an aortic valve prolapse, and this complication will then become increasingly more serious over time; the attribute* aort_prolapse *is progressive. Similarly, a large left-to-right shunt may cause pulmonary arteriopathy over time (Eisenmenger's reaction); the damage to the pulmonary arterioles is initially reversible, but eventually becomes irreversible; the value irreversible_arteriopathy of the attribute* pulm_art *is absorbing.*

In the above example we have assumed absence of medical interference. Some, but not all, of the clinical attributes mentioned have different properties under treatment; the effects of performing treatment actions are discussed next.

Effects of treatment

Generally speaking, the goal of remedial clinical action is to 'move' the patient's state into a preferred 'direction' with the ultimate objective of reaching a healthy condition. It depends on the state of the patient whether this is possible: the effectiveness of clinical therapies varies with the condition to which they are applied. We can therefore think of the current clinical state of the patient as a *precondition* that influences the effectiveness of treatment actions. In this precondition, there will often be a small number of *decisive* attributes for successful treatment, while the risk of complications depends on *additional* attributes, related to the general condition of the patient.

Example 6.14 *There are two treatment modalities in the VSD domain: medical treatment with cardiac glycosides (to reduce the effects of left-to-right shunting) and cardiac surgery (to close the VSD and repair possible other malformations of the heart). Medical treatment will be effective when the heart is potentially capable of compensating the haemodynamic disturbance caused by shunting; the shunt size and vascular resistances are primary factors in this respect. Surgical closure of VSDs is usually successful, but may fail with severe malalignments of the outlet septum, with VSDs that are difficult the access (e.g. muscular defects), and with VSDs that are very large in size. Furthermore, when a VSD is closed in the presence of high pulmonary vascular resistance, this will cause acute right-ventricular failure and death. There is an increased risk of other complications when the general condition of the patient is bad because he suffers from severe (chronic) heart failure and pulmonary infections. Both treatment modalities and their respective precondition attributes are shown in Table 6.3.*

Upon successful application of treatment, one or more attributes of the patient's condition will change to a more desirable state. When the treatment fails, such a state is not reached, or it is reached at the price of complications. In sum, we can say that the patient's state after treatment is a *postcondition* that results from application of the treatment under the precondition that was present in the patient's state beforehand. In terms of Subsection 5.2.1, postcondition attributes are the attributes in state variables that are directly affected by the action. If a state variable is indirectly affected, we do not consider its attribute a part of the action's postcondition, because changes to that variable can be understood in terms of changes to directly affected variables and the normal mechanisms that relate directly and indirectly affected variables. We illustrate this with an example.

Example 6.15 *The delivery of cardiac glycosides directly affects the severity of heart failure from which a VSD patient suffers; indirect results of the treatment are a reduction in heart size (when it was increased), and a diminished risk of pulmonary infections. These indirect results are however normal consequences of a reduction in heart failure; the only postcondition attribute of the action 'med' is therefore* `LV_failure`*. Similarly, surgical closure of a VSD eliminates the haemodynamic disturbances caused by the defect, and therefore indirectly affects attributes such as* `shunt`*,* `art_press`*,* `ventr_press`*, and many other related physiological, complicatory, and symptomatic attributes. The relevant postcondition attribute of the action 'surg' in this respect is* `VSD_size`*; once the situation of* `VSD_size` *= null is reached, the haemodynamic situation becomes the same as in a healthy patient.*

Again, two types of postcondition attributes are identified: *target* attributes, i.e. attributes whose state we wish to change by application of the action, and *complicatory* attributes, i.e. attributes that come into play involuntarily. Complicatory attributes describe the possible side effects of drugs and the potential complications of clinical

interventions. We note that this distinction is, in general, *context-sensitive*: what is considered a complication in some domains may be seen as a target in other. From a probabilistic point of view, there is neither a difference between decisive and additional precondition attributes, nor between target and complicatory postcondition attributes; we could never make such distinctions from a given model of control. It is possible though, to employ additional structures to formalise these distinctions; we do not elaborate on this issue.

Now, let Y_1 and Y_2 be the sets of pre- and postcondition attributes of treatment action $a \in A_{\text{treat}}$. If we assume that the effects of actions are both *immediate*, i.e. do not require additional time, and *short-lived*, i.e. do not extend over future time points, then specifying the effects of action a boils down to establishing, for each non-final time point $t \in T$ and each of pair of configurations $(c_{Y_1(t)}, c_{Y_2(t+1)}) \in \Omega_{Y_1(t)} \times \Omega_{Y_2(t+1)}$, the probability that $c_{Y_2(t+1)}$ results when $d(t)=a$ is chosen in the context of $c_{Y_1(t)}$. This approach to modelling treatment effects described above can be regarded as probabilistic version of the STRIPS method of modelling actions effects, (Fikes and Nilsson, 1971). This method was developed for classical, logic-based planning systems; recently, a similar probabilistic version was used by Boutilier et al. (1997) in an algorithm for solving (fully-observable) Markov decision processes.

From a planning perspective, it is important to recognise the extent to which treatment actions can alter the normal course of disease. More particularly, it should be assessed which types of endogenous change can be speeded up, slowed down, stopped, or undone. One will generally find that static, progressive, or absorbing attributes that occur in the postcondition of some treatment action cease to have this property under effective application of the action; attributes that do not occur in any postcondition may withdraw from this type of control.

Example 6.16 *Recall from Example 6.13 that the attribute* `VSD_size` *is progressive, with smaller defect sizes being more likely over time; surgical closure of the VSD can be seen as speeding up this process. Furthermore, it was described that an aortic valve prolapse may complicate a subaortic VSD, and that this complication will then become increasingly more serious over time. Surgical repair of the VSD will however stop this progression; moreover, repair of the prolapsed valve is often possible. Both circumstances appear implicitly from the fact that* `VSD_size` *and* `aort_prolapse` *are postcondition attributes of the action 'surg', as listed in Table 6.3. Finally, the attribute* `pulm_art` *was described to have absorbing state irreversible_arteriopathy. In contrast with the previous conditions, there is no clinical means to undo this complication:* `pulm_art` *is not a postcondition attribute of any treatment action.*

6.1.4 Clinical findings

We now turn to the specification of clinical findings of disease in our framework; these include symptoms reported by the patient, signs appearing from physical examina-

Name	Description
`LV_failure`	left-ventricular failure
`RV_failure`	right-ventricular failure
`paleness`	pale facial colour
`sweating`	sweating
`sys_mur`	systolic murmurs
`dia_mur`	diastolic murmurs
`heart_sounds`	heart sounds
`thrill`	thrill
`vous_card`	*voussure cardiaque*
`cyanosis`	central cyanosis

Table 6.4: Attributes representing signs and symptoms (group 6).

tion and measurements obtained from diagnostic procedures. There are two basic relations that are important here. First, the model of observation O describes which external attributes are observable under different types of action choice. Second, in the model of control Ξ we find described the reliability of observations with respect to the internal condition of the patient. In the fields of medical decision analysis and biostatistics, it is customary to express this reliability in term of *sensitivity* and *specificity*. Below, we will first discuss the possibilities to observe external variables, and then describe how sensitivity and specificity of observations are cast in control functions.

External manifestation

Recall that a model of observation for decision basis $\mathcal{B}$ takes the form of a set $O = \{o_t \mid t \in T\}$ of functions $o_t : \mathcal{A}[0, t-1] \rightarrow \wp(X[0,t])$, with the restriction of monotonicity, i.e. that

$$o_t(\alpha) \quad \subseteq \quad o_{t+1}(\alpha \wedge d(t) = a) \tag{6.5}$$

for all time points $t \in T$, $t < N$, action sequences $\alpha \in \mathcal{A}[0, t-1]$, and each action $a \in A$.

In Subsection 5.2.2, we derived the notion of *manifestness* from a given model of observation O; a distinction was made between *potentially manifest*, *absolutely manifest*, and *hidden* variables. In a clinical setting, potentially manifest variables typically describe signs and symptoms of disease and outcomes of diagnostic tests. So from our earlier assumptions, it follows that variables from the set $E(T)$ are potentially manifest. It depends on the formalisation of the clinical problem whether such potentially manifest variables are also absolutely manifest. If each decision moment denotes a point in time where the patient is seen by the clinician, then all variables that represent clinical symptoms are absolutely manifest at all times. In contrast, when part of the planning problem is to decide if and when the patient should visit the clinic, these variables are manifest only upon such visits. Variables from the set $I(T)$ are

never observable and therefore hidden under all circumstances.

Example 6.17 *One of the aspects in the management of VSD patients is the planning and timing of clinical visits. When a patient visits the clinic at decision moment $t \in T$, his signs and symptoms at that moment are assessed by the paediatric cardiologist; the attributes representing signs and symptoms in the VSD domain are listed in Table 6.4. Depending on his findings, the cardiologist can now decide to postpone further action to a next visit, to conduct additional diagnostic tests, or to submit the patient surgery; the latter decision will however not be taken without further investigations. The possibilities for diagnostic investigation are ECG, echocardiography, blood tests, cardiac catheterisation, and open lung biopsies. The attributes that represent measurements obtained from these tests are shown in Table 6.5.*

The reliability of observations

Generally speaking, observing an external attribute has the objective of getting a better picture of one or more internal attributes. As the relationship between internal and external attributes is usually stochastic in nature, one could say that these internal attributes are observed indirectly, through a noisy channel; we say that the external attribute provides a *measurement* on the internal attributes. Frequently, there exist multiple measurements on a given internal attribute, each differing in reliability, or, in other words, in the amount of 'noise' on the channel.

Example 6.18 *It was described in Chapter 2 that continual left-to-right shunting through a VSD may cause severe damage to the pulmonary arterioles (Eisenmenger's reaction). The damaged arterioles have an increased resistance to blood flow, causing a reduced shunt. Eventually, the resistance becomes so high as to cause reversal of the shunt direction; the venous and arterial blood are now mixed in the heart, yielding an undersaturation of the systemic arterial blood with oxygen. To assess whether this condition applies to a given patient, one can measure the oxygen saturation with a blood test; in our formalisation, the external attribute* `oxsat_test` *represents a measurement of the internal attribute* `oxsat`*. In addition, undersaturation yields a dusky blue discolouration of the tongue, lips, and conjunctivae. These clinical signs, jointly referred to as* cyanosis*, are easily detected from visual inspection of the patient. In our formalisation, the external attribute* `cyanosis` *therefore represents a second, but less reliable measurement on the internal attribute* `oxsat`*.*

To establish the reliability of indirect observations, we must assess the properties of the channel through which they are made. Let internal variable $x(t) \in I(T)$ be indirectly observable by measuring the value of external variable $y(t) \in E(T)$,[3] and suppose we are interested in knowing whether $x(t)$ has the value $s \in \mathrm{dom}(x)$, because

[3]For brevity, we assume that there is no time lag between these internal and external events. One could however also assume that attribute y is observed at some later time point.

Name	Description	Test
LVH_ECG	LVH measurement	*ECG*
RVH_ECG	RVH measurement	*ECG*
CT-ratio	cor-thoracic ratio	*roentgenography*
pulm_vasc	pulmonary vascularity	*roentgenography*
outlet_pos_echo	observed malalignment	*echocardiography*
VSD_type_echo	observed VSD type	*echocardiography*
VSD_ext_echo	observed VSD extension	*echocardiography*
VSD_size_echo	observed VSD size	*echocardiography*
PDA_echo	observed PDA	*echocardiography*
aort_sten_echo	observed aortic stenosis	*echocardiography*
aort_prolapse_echo	observed aortic prolapse	*echocardiography*
pulm_sten_echo	observed pulmonary stenosis	*echocardiography*
aort_regurg_echo	aortic valve regurgitation	*echocardiography*
pulm_regurg_echo	pulmonary valve regurgitation	*echocardiography*
ox_sat_test	oxygen saturation of the blood	*blood test*
shunt_cath	observed shunt	*catheterisation*
LV_press_cath	observed left-ventricular pressure	*catheterisation*
RV_press_cath	observed right-ventricular pressure	*catheterisation*
aort_press_cath	observed aortic pressure	*catheterisation*
pulm_press_cath	observed pulmonary wedge pressure	*catheterisation*
pulm_art_biop	open-lung biopsy result	*open lung biopsy*

Table 6.5: Attributes representing clinical measurements (group 7).

that value is crucial for the patient's prognosis. We briefly describe the decision-theoretic approach for using observations on $y(t)$ to estimate whether this is the case; for details, we refer to (Weinstein and Fineberg, 1980). First, we choose a subset M_s^+ of values from $\mathrm{dom}(y)$ that we associate with value s. If there exists a linear ordering $\preceq_y$ on $\mathrm{dom}(y)$, this will usually amount to selecting two *cut-off values* $m_s^-, m_s^+ \in \mathrm{dom}(y)$ such that

$$M_s^+ \;=\; \{m \in \mathrm{dom}(y) \mid m_s^- \preceq_y m \preceq_y m_s^+\}$$

We say that $y(t)$ and M_s^+ provide a *test* on $x(t) = s$.

The *sensitivity*, or *true-positives rate*, of this test is now defined as the proportion of patients where $x(t) = s$ is true, and a 'positive' test result, i.e. a value from M_s^+, is found; it is equal to

$$\sum_{m \in M_s^+} \xi_\alpha^t(y(t) = m \mid x(t) = s), \tag{6.6}$$

where $\alpha \in \mathcal{A}[0, t-1]$ is the past sequence of actions. The *specificity*, or *true-negatives rate*, of the test is defined as the proportion of patients where $x(t) = s$ is false, and a 'negative' test result, i.e. a value outside M_s^+, is found; it is equal to

$$\sum_{m \in M_s^-} \xi_\alpha^t(y(t) = m \mid x(t) \neq s) \tag{6.7}$$

where $M_s^- = \mathrm{dom}(y) \setminus M_s^+$.

Example 6.19 *Cyanosis is not a fully reliable indication of the oxygen saturation of the systemic arterial blood. When the oxygen saturation has the normal level of 93% or higher, cyanosis is still seen in 2% of the cases. When the oxygen saturation is below 93% but above 86%, cyanosis is seen in 60% of the cases. Otherwise, the cyanotic features are definitely present. So, if we are interested in the value '$> 93\%$' of attribute* `oxsat`*, the sensitivity of a test with attribute* `cyanosis` *and value 'no' has a sensitivity 0.98. If we assume that oxygen saturations below 86% occur with 5% of the patients and saturations between 86% and 93% with 15% of the patients, a specificity of 0.7 is obtained.*

6.1.5 Specification of objectives

In the previous section, we have explained how a decision basis and models of control and observation can be employed to develop a formal description of a given clinical domain. We now arrive at the two remaining components of our framework: utility functions and contingency plans. In the traditional decision-theoretic approach, the components are purely prescriptive in character. One then uses a utility function u to specify the objectives of clinical decision making, and the associated computational problem is to derive a contingency plan π that is operational under the given model of observation, and is optimal with respect to the function u. The plan π then describes the ideal, rational decision-making behaviour that is consistent with the objectives described by the utility function.

In this subsection we concentrate on tradeoffs in the design of a utility function that is to be employed within this traditional decision-theoretic approach. This is not to say, however, that other approaches to decision making and planning are excluded. In Subsection 6.2.2, for instance, we will use a set of pre-specified heuristic decision rules to guide decision making.

We will investigate two possible methods to formalising the objectives of patient management with a utility function. The first method stems from the classical planning systems in AI and identifies *goal states*. The second method is usually employed in clinical decision analysis and is based on metrics for *duration and quality of life*. It is assumed that a decision basis $\mathcal{B} = (X, A, T)$ and models of control and observation Ξ and O are given, where as before, a distinction is made between internal and external attributes of clinical states.

Planning for clinical goals

A most basic way of formalising planning objectives with goal states exists in the assumption of *final-state goals*. As described in Section 5.4, a final-state goal is a condition that is unambiguously preferred at the end of the decision process. Here, we assume that such goals are formulated in terms of a set $Y \subseteq I$ of internal attributes

that are considered most relevant for the well-being of the patient. We now distinguish a subset $\mathcal{G} \subset \mathrm{dom}(Y)$ of states of the set Y that we would like to achieve at the final time point $t = N$; the corresponding utility function u_1 defined as

$$u_1(h) \;=\; \begin{cases} 1, & \text{if } h \vdash Y(N) = S_Y \text{ and } S_Y \in \mathcal{G}, \\ 0, & \text{otherwise.} \end{cases} \tag{6.8}$$

In clinical settings, final-state goals will usually be expressed in terms of the internal attributes that are crucial for the patient's post-therapeutic condition.

Example 6.20 *The management of a VSD patient is considered successful when, at the end of the management process, the VSD is closed without further complications. In particular, pulmonary arteriopathy must have been avoided. Furthermore, when surgery was necessary to close the defect, one prefers the situation where this has not yielded any clinical complications. A final-state goal for the VSD domain is therefore*

$$\begin{array}{l} \texttt{VSD\_size}(N) = \mathit{null} \wedge \texttt{ASD}(N) = \mathit{none} \wedge \texttt{PDA}(N) = \mathit{none} \\ \wedge\, \texttt{aort\_sten}(N) = \mathit{none} \wedge \texttt{aort\_prolapse}(N) = \mathit{none} \\ \wedge\, \texttt{pulm\_sten}(N) = \mathit{none} \wedge \texttt{pulm\_art}(N) = \mathit{normal} \\ \wedge\, \texttt{arrhythmia}(N) = \mathit{false} \wedge \texttt{perforation}(N) = \mathit{false} \\ \wedge\, \texttt{stroke}(N) = \mathit{false} \wedge \texttt{bleeding}(N) = \mathit{false} \wedge \texttt{death}(N) = \mathit{false}. \end{array} \tag{6.9}$$

One of the drawbacks to the assumption of final-state goals is that temporal aspects of management objectives are neglected. The clinical condition of Example 6.20 is typically preferred at all times; one would want to achieve it as soon as possible, but this aspect of the objective is not formalised by function u_1. A more sophisticated approach is therefore to use the time-separable additive function u_2, defined as

$$u_2(h) \;=\; \sum_{t \in T} k_t \cdot r(\sigma(t)), \tag{6.10}$$

for $h = (\sigma, \alpha)$, where

$$r(S_X) \;=\; \begin{cases} 1, & \text{if } X(t) = S_X \vdash Y(t) = S_Y \text{ and } S_Y \in \mathcal{G}, \\ 0, & \text{otherwise,} \end{cases} \tag{6.11}$$

and $k_0, \ldots, k_N$ are scaling constants. Function u_2 counts the number of decision moments during time frame T where a goal state is occupied. In principle, a unit reward is received for each of these moments. The scaling factors can be used to incorporate the fact that decision moments may not be distributed uniformly over time under a given temporal map τ, by taking

$$k_t \;=\; \tau(t+1) - \tau(t), \tag{6.12}$$

for all $t < N$. Occupying a goal state at non-final decision moment $t \in T$ is now rewarded directly proportional to the period of real time between t and its successor

moment. The factor k_N is typically much greater, as reaching a goal state at the final time point $t = N$ is crucial for the patient's post-therapeutic prognosis.

A second drawback, to both functions u_1 and u_2, is that they neglect the actions that are undertaken during the management process. In practice, we will want to minimise the number of actions taken, especially those that involve high costs or a lot of discomfort for the patient. To incorporate this secondary objective into function u_2, we associate a numerical penalty $cost(a)$ with each action $a \in A$, and try to minimise the total costs during the management process:

$$\begin{aligned} u_3(h) &= u_2(h) - m \cdot \sum_{t \in T} cost(\alpha(t)) \\ &= \sum_{t \in T} k_t \cdot r(\sigma(t)) - m \cdot cost(\alpha(t)), \end{aligned} \tag{6.13}$$

for $h = (\sigma, \alpha)$, where m is a factor that weighs the action costs with the rewards that are received from achieving goal states. If we assume that the penalty function *cost* yields monetary values, then the weight factor m is assessed by determining the average amount of money one is willing to pay for reaching a goal state. Note that the benefits of reaching a goal state highly depend on one's current condition, and that it may therefore be difficult to assess the factor m.

Example 6.21 *In the VSD domain, one would typically associate high costs with diagnostic actions such as cardiac catheterisation and open lung biopsy, and very high costs with cardiac surgery. The latter costs cause one to postpone a decision to operate a VSD patient as long as the chances of spontaneous closure are high, and the costs of surgery are not outweighed by the reward of reaching the goal condition (of a closed VSD) earlier.*

A final drawback to functions u_1, u_2, and u_3 is that they do not formulate a tradeoff between reaching and not reaching a goal state: goal states are basically rated infinitely more valuable than other states. This is problematic when goal states cannot be reached with certainty; striving for a less ideal condition might be better than pursuing the unreachable. A possible way to obviate this drawback is introducing a value function

$$v : \mathrm{dom}(Y) \rightarrow [0, 1], \tag{6.14}$$

where $v(S_Y)/v(S'_Y) = j$, $S_Y, S'_Y \in \mathrm{dom}(Y)$ expresses that state S_Y is j times more desirable than state S'_Y. Typically, goal states will be considered more desirable than non-goal states, i.e. $v(S_Y) > v(S'_Y)$ if $S_Y \in \mathcal{G}$ and $S'_Y \notin \mathcal{G}$, but the precise definition of the function v depends on subjective preferences. This leads to utility function u_4, in which the binary function r is replaced by the value function v:

$$u_4(h) = \sum_{t \in T} q_t(\sigma(t)) - m \cdot cost(\alpha(t)), \tag{6.15}$$

for $h = (\sigma, \alpha)$, where

$$q_t(S_X) \;=\; \begin{cases} (\tau(t+1) - \tau(t)) \cdot v(S_Y), & \text{if } t < N, \\ k_N \cdot v(S_Y), & \text{if } t = N, \end{cases} \tag{6.16}$$

and $S_Y \in \mathrm{dom}(Y)$ is the subvalue associated with Y in S_X. This is roughly the point where the symbolic, qualitative type of modelling from AI is traded for an approach that is more numerical in character.

Maximising life-expectancy

We now turn to the second approach to formalising clinical objectives, stemming from the field of clinical decision analysis. In this approach, a quantitative performance metric is designed, usually based on duration and quality of life. We first focus on duration, i.e. life-expectancy, and incorporate life-quality thereafter. As above, we assume $Y \subseteq I$ to be a set of internal attributes that is crucial for the patient's condition in present and future. Furthermore, this set is taken to comprise a Boolean state variable `death`, whose state *true* is absorbing, i.e.

$$\xi_\alpha(\texttt{death}(t+1) = \mathit{true} \mid \texttt{death}(t) = \mathit{true}) \;=\; 1 \tag{6.17}$$

for all control functions $\xi_\alpha \in \Xi$ and time points $t \in T$, $t < N$, where

$$\xi_\alpha(\texttt{death}(t) = \mathit{true}) \;>\; 0. \tag{6.18}$$

We assume that when `death`(t)=*false*, the mortality risk at time point $t+1$ is determined by the state and action at that time point; these risks are encoded in the model of control Ξ. The probability of survival up to time point $t \in T$ in decision process P equals $P(\texttt{death}(t) = \mathit{false})$.

A simple way of formalising life-expectancy is to estimate it from the final state of the set Y. That is, we use a prognostic function

$$\mathrm{LE} : \mathrm{dom}(Y) \to \mathbb{R} \tag{6.19}$$

that specifies the (estimated) post-therapeutic life-expectancy of the patient, given the final state of the set Y, and on the scale of the temporal map τ. Naturally, $\mathrm{LE}(S_Y) = 0$ for all states $S_Y \in \mathrm{dom}(Y)$ that comprise the value *true* for the variable `death`. For other states of Y, we must assess some life-expectancy estimate from clinical studies or field experts; note that this may be as difficult as obtaining probability assessments for a model of control.

Example 6.22 *The attributes of the final-state goal in Example 6.20 are the ones that determine a patient's life-expectancy. Patients for which the goal state is achieved have a normal life-expectancy; when a minor disorder such as a small VSD or a cardiac arrhythmia remains, this life-expectancy is somewhat reduced. More serious are stenoses of the aortic and pulmonary valve, and an aortic valve prolapse. Patients with pulmonary arteriopathy have a strongly-reduced life-expectancy: they usually do not reach middle age.*

The utility associated with outcome $h \in \mathcal{H}(T)$ is now defined as

$$u_5(h) \;=\; \mathrm{LE}(S_Y), \tag{6.20}$$

where $S_Y \in \mathrm{dom}(Y)$ is the (unique) state of Y for which $h \vdash Y(N) = S_Y$.

It is interesting to note that we can actually regard utility function u_1 of Equation 6.8, the function that implements the simplest form of planning for final-state goals, as a qualitative abstraction of function u_5, where all states that have an associated life-expectancy above some predefined level l are treated as goals:

$$\mathcal{G} \;=\; \{S_Y \in \mathrm{dom}(Y) \mid \mathrm{LE}(S_Y) > l\}. \tag{6.21}$$

Utility function u_5 also suffers from a similar shortcoming as function u_1: it neglects the temporal aspects of clinical objectives. Function u_5 falls short when the management process itself takes considerable time and carries non-negligible mortality risks; these circumstances are found, for instance, in the management of chronic diseases (e.g. acquired cardiovascular disorders) and incurable diseases (e.g. cancer and AIDS). We can solve this by adding the life time *during* the decision process to function u_5:

$$u_6(h) \;=\; u_5(h) + \max\{\tau(t) - \tau(0) \mid h \vdash \texttt{death}(t) = \mathit{false}\}. \tag{6.22}$$

The utility value $u_6(h)$ is equal to the difference $\tau(t) - \tau(0)$ when the patient dies at time point $t \in T$ in history h. In words, this is the amount of real life time (on the scale of the temporal map τ) the patient has had during the management process. Otherwise, $u_6(h)$ equals $\mathrm{LE}(S_Y) + \tau(N) - \tau(0)$, where $h \vdash Y(N) = S_Y$: the post-therapeutic life-expectancy plus the life time during clinical management.

Functions u_5 and u_6 now both suffer from a similar shortcoming as functions u_1 and u_2: they neglect the actions that are undertaken during the management process. Here, the solution we employ is completely the same as earlier: we take penalties associated with actions into account to avoid abundant clinical activity. This yields the refined function u_7, defined as

$$u_7(h) \;=\; u_6(h) - m' \cdot \sum_{t \in T} \mathit{cost}(\alpha(t)), \tag{6.23}$$

for $h = (\sigma, \alpha)$, where m' is again a weight factor. If we assume that the penalty function *cost* yields monetary values, then the weight factor m' is assessed by determining the amount of money one is willing to pay for a given period of life (e.g. a year) in a healthy condition. This type of tradeoff is characteristic for the decision-theoretic approach to clinical decision support; how difficult it may be to determine such an amount of money, it does assign a clear, context-independent meaning to the factor m'. Recall that the interpretation of the comparable weight factor m in Equation 6.13 was much more problematic.

It is customary in clinical decision analysis to also include a measure of life quality in utility functions: a shorter life time in good health is often preferred over a longer life time with disease or disabilities. To this end, we introduce an additional function

$$\mathrm{Q} : \mathrm{dom}(Y) \rightarrow [0, 1], \tag{6.24}$$

where $\mathrm{Q}(S_Y)$ is a subjective, relative measure of life quality for state $Y = S_Y$. We have $\mathrm{Q}(S_Y) = 1$ when S_Y is a state of good health, and have lower values for states with disease or disability. These lower values are calibrated as follows: when $\mathrm{Q}(S_Y) = q$, $0 < q < 1$, this means that the patient is indifferent between 1 year in a healthy condition and $1/q$ years in state $Y = S_Y$. Finally, if $Y = S_Y \vdash \texttt{death} = \mathit{true}$, then $\mathrm{Q}(S_Y) = 0$.

We now arrive at the following definition for a utility function that is based on quality-adjusted life-expectancy, discounted by action costs:

$$u_8(h) \;=\; \sum_{t \in T} q'_t(\sigma(t)) - m' \cdot \mathit{cost}(\alpha(t)), \tag{6.25}$$

for $h = (\sigma, \alpha)$, where

$$q'_t(S_X) \;=\; \begin{cases} \mathrm{Q}(S_Y) \cdot (\tau(t+1) - \tau(t)), & \text{if } t < N, \\ \mathrm{Q}(S_Y) \cdot \mathrm{LE}(S_Y), & \text{if } t = N, \end{cases} \tag{6.26}$$

and $S_Y \in \mathrm{dom}(Y)$ such that $X = S_X \vdash Y = S_Y$.

If we compare Equations 6.15 and 6.16 with Equations 6.25 and 6.26, it appears immediately that we have now reached a point where the two approaches to formalising clinical objectives coincide. If the weight factors m and m' and the functions v and Q are equal, and moreover $\mathrm{LE}(S_Y) = k_N$ for all $S_Y \in \mathrm{dom}(Y)$, then there is no difference between functions u_4 and u_8. Where the first two assumptions seem perfectly reasonable, the third assumption, that LE is a constant function, curtails the expressiveness of utility function u_8. We conclude that the goal-based approach to formalising clinical objectives approximates the decision-analytic approach based on duration and quality of life. The latter approach is slightly more expressive, and has, in addition, some semantical advantages; it does require though the assessment of a prognostic function for life-expectancy.

6.2 Supporting management tasks

In the preceding section, the emphasis has been on modelling and specification of clinical domains; the remainder of this chapter is devoted to the clinical tasks of diagnosing, treating, and prognosticating a patient. Starting from our framework for decision-theoretic planning, we describe various choices one has in formalising these tasks, and the consequences of these choices from conceptual and formal perspectives.

Throughout we assume that a given clinical domain is described by a decision basis $\mathcal{B} = (X, A, T)$, and that Ξ and O are associated models of control and observation. As before, we take the set X to be composed of sets I and E of internal and external variables, and the set A to comprise the empty action ε. Furthermore, a utility function u is taken to describe the clinical objectives of managing patients in the domain under consideration; it is assumed that u is normalised, i.e. $0 \leq u(h) \leq 1$ for all histories $h \in \mathcal{H}(T)$.[4]

In each of the subsections to come, we will start from the following setting. A physician faces the problem of choosing an action at time point $t \in T$ during the management process. At the time points preceding t, various events and actions have taken place; this past clinical history of the patient is described by history $h \in \mathcal{H}[0, t-1]$, where $h = (\sigma, \alpha)$. Furthermore, the present clinical state of the patient is described by the configuration $\mathrm{c}_{X(t)} \in \Omega_{X(t)}$. Of course, this description of the state of affairs is based on an omniscient perspective; the treating physician is typically equipped with much less information. The sequence α consists of clinical actions that have previously been undertaken, and is therefore known to the physician, but the patient's past and present clinical conditions will be partially (or largely) unknown to him. To some extent, these conditions have been observed in the form of external manifestations of disease; we use $\chi \in \Phi(T)$ to represent the collected findings. We assume that χ is a configuration of the observable set $o_t(\alpha)$, where $o_t \in O$ is the observation function for time point t. We will refer to χ as the *evidence* at time point t, and to $\alpha \wedge \chi$ as the clinician's *knowledge*.

Conceptually, the physician's problem can now be summarised as follows. From the available evidence χ, he must establish a diagnosis for the patient. Using the diagnosis, an action choice for the current time point is to be made; the tradeoff between possible actions should take into account the associated prospects for the patient in the future, as far the knowledge $\alpha \wedge \chi$ allows for making such predictions. Furthermore, the physician must be aware of his own management strategy in the future as this may have consequences for the present choice. Overall, we can therefore say that the decision involves diagnostic, therapeutic, and prognostic aspects that are highly intertwined. This holds in fact for all decisions during the management process; the main difference is that they are based on different information and with varying prospects for the patient. In the subsections below, we will now successively consider the diagnostic, therapeutic, and prognostic aspects of such decisions.

6.2.1 Diagnosis

Diagnosis is the task of determining the likely cause(s) of observed findings. Often medical diagnostic systems simply generate a list of diseases that might account for the given findings; we find this for instance in the classical diagnostic system

[4] This is purely a matter of convenience and does not affect generality: recall that for any utility function u there exist an equivalent utility function u' that is normalised.

INTERNIST-1, (Miller et al., 1982). It depends on the domain under consideration and the purpose of the diagnosis whether such is adequate. In complex domains with multiple interacting mechanisms leading to the observations, diagnosis often requires the reconstruction of the likely scenario that produced them, as further management of the patient is otherwise too difficult. Furthermore, in a domain with a significant degree of uncertainty, one can often not reliably determine the operational (i.e. 'true') cause of pathological findings, and a set of possible explanations, called the *differential diagnosis*, has to be generated.

Before any type of diagnostic reasoning can take place, one must establish the universe of potential *diagnostic hypotheses* in a given domain. There are, already at this preliminary point, various choices involved; we will first draw up an inventory of the possibilities here. Thereafter, we will formalise the notion of *diagnostic explanation*, and consider some possibilities to formalise the *diagnostic relevance* of potential explanations. We finally discuss possibilities to make diagnoses more concise, and briefly touch upon diagnostic test selection.

The diagnostic hypothesis

A first step in any formalisation of diagnostic problem solving consists of identifying the potential diagnostic hypotheses in the domain under consideration. In general, we can say that diagnostic hypotheses at time point t aim to classify disorders in the unobserved, internal part of past and present clinical conditions, described by $\sigma \wedge \mathrm{c}_{X(t)}$. We will therefore assume that diagnostic hypotheses are planning expressions from $\Phi(T)$ that refer to variables from $I[0,t]$ only. We will use Ψ_t to denote all possible hypotheses that are considered in diagnostic reasoning when deciding upon the action choice at time point t, where each hypothesis $\psi \in \Psi_t$ is a conjunction of value assignments. Note that from an omniscient perspective, we can state that hypothesis ψ is correct in when $\sigma \wedge \mathrm{c}_{X(t)} \vdash \psi$.

We now shift our focus of attention from the concept of 'hypothesis' to the concept of 'disorder'. Hypotheses from the set Ψ_t are taken to describe possible disorders of the patient, but surely, not every expression that refers to variables from $I[0,t]$ would be considered to do so. The questions arise as to which type of expression would be considered to describe a disorder, and which set of expressions would cover the complete range of possible disorders? These questions are in fact quite challenging; the concept of 'disease' has not without reason been called a "many-headed monster" (Habbema and Hilden, 1981). We describe four distinctions, relating to *temporal scope*, *pathological interpretation*, *exclusiveness*, and *exhaustiveness* of diagnostic hypotheses.

Temporal scope The first distinction concerns the *temporal scope* of reference of disorders. When diagnostic hypotheses only aim to classify the internal state of the patient at time point t, we speak of *static disorder descriptions*. The elements from Ψ_t then only refer to variables from $I(t)$. A more general approach is when hypotheses may also extend over the past, and elements of Ψ_t may thus include a description

of the disease history. We speak of *dynamic disorder descriptions*; an example of this type of disorder description is used in the system CASNET, (Weiss et al., 1978). The use of dynamic disorder descriptions is motivated by the fact that classifying the present disease status of a patient is often insufficient for purposes of management; the degree of progression and severity of disease need also to be taken into account. We find this in the domain of paediatric cardiology.

Example 6.23 *Consider the diagnostic hypothesis "an originally moderately-sized VSD that is now closed". This is expressed by the conjunction*

$$\texttt{VSD\_size}(t) = \mathit{null} \ \wedge \ \texttt{VSD\_size}(0) = \mathit{moderate}.$$

The hypothesis "a large perimembranous VSD that has progressed into a Tetralogy of Fallot" is formalised as

$$\begin{aligned}&\texttt{VSD\_size}(t) = \mathit{large} \ \wedge \ \texttt{outlet\_pos}(t) = \mathit{severe\_right} \ \wedge \\ &\texttt{VSD\_size}(0) = \mathit{large} \ \wedge \ \texttt{VSD\_type}(0) = \mathit{perimembranous},\end{aligned}$$

and "a muscular VSD with recently acquired, reversible pulmonary arteriopathy" as

$$\begin{aligned}\texttt{pulm\_art}(t) = \mathit{reversible\_arteriopathy} \ \wedge \ \texttt{pulm\_art}(t-1) = \mathit{normal} \ \wedge \\ \texttt{VSD\_type}(0) = \mathit{muscular}.\end{aligned}$$

The increased expressiveness of the dynamic disorder descriptions comes at the price of an exponentially growing universe of possible disorders: if one distinguishes k disorders in static descriptions, there are potentially k^{t+1} different disorder histories to be considered in dynamic descriptions at time point t. Of course, often many of these histories would *a priori* be ruled out because of their high implausibility. We note that there exists a relationship between the restriction to static disorder descriptions and the Markov assumption in decision processes: in both cases one considers the past to be irrelevant with respect to the future, given the present state of affairs.

Pathological interpretation A second distinction concerns the pathological interpretation of diagnostic hypotheses. One may choose to let hypotheses represent *normal* states of affairs or *abnormal* states of affairs, and a mixture of types of also possible. When all hypotheses represent a normal, healthy condition, diagnostic reasoning proceeds from the apparent mismatch between given findings and those that would be expected; the objective is to find out which hypothesis cannot possibly be true, which thus localises the disorder. When all hypotheses represent an abnormal, pathological condition, diagnostic reasoning consists of finding the hypothesis that best predicts the given findings, again localising the disorder. The distinction between models of normal and abnormal conditions is often used in *model-based diagnosis*, (Reiter, 1987; de Kleer et al., 1992; Lucas, 1998), where reasoning proceeds from a behavioural or causal model that describes a system in terms of functional

relations. Model-based diagnosis is traditionally formalised using classical and non-monotonic logic, and implements a form of purely symbolic reasoning.

The distinction between normal and abnormal situations is seldomly encountered in the literature on diagnosis in probabilistic and decision-theoretic settings, because probabilistic reasoning does not, in principle, require this distinction to be made. From a semantical point of view, the distinction is of course equally valid; the hidden assumption is generally that both normal and abnormal conditions occur in diagnostic universes.

Exclusiveness The third distinction concerns the *exclusion* of hypotheses by others. If the elements of Ψ_t are mutually exclusive (i.e. $\psi_1 \wedge \psi_2 \equiv \bot$ for all $\psi_1, \psi_2 \in \Psi_t$, $\psi_1 \neq \psi_2$), then Ψ_t provides a set of competing hypotheses, and only one hypothesis can describe the true state of affairs in a given situation. This is often a convenient property in formal diagnostic reasoning; it is frequently employed in probabilistic and decision-theoretic settings. Note that if x_1 and x_2 are internal attributes that are considered relevant for the patient's diagnosis, then configurations of $x_1(t)$ and $x_2(t)$ can be regarded as partial (single-disorder) hypotheses about the patient's state, while configurations of the set $\{x_1(t), x_2(t)\}$ are compound (multiple-disorder) hypotheses. The assumption of mutual exclusion however thwarts this possibility, as configurations of $x_1(t)$ and $x_2(t)$ do not exclude each other and could therefore not be both elements of the universe of hypotheses.

Exhaustiveness The fourth and final distinction concerns the range of disorders identified: the universe Ψ_t of disorders may or may not be *exhaustive*. We say that such is the case when for each possible past state sequence $\sigma \in \mathcal{S}[0, t-1]$ and each possible present state $\mathrm{c}_{X(t)} \in \Omega_{X(t)}$ there exists a correct hypothesis $\psi \in \Psi_t$ (i.e. $\sigma \wedge \mathrm{c}_{X(t)} \vdash \psi$). If this is not the case, then we may find ourselves in the situation that we cannot find an explanation in Ψ_t for a given set of findings χ, and the diagnostic procedure stalls at its very first step because of a failure to generate hypotheses. It may simply indicate that the patient is healthy and no clinical action is required, but also that we have crossed the boundaries of our formalisation's application domain. Exhaustiveness therefore interacts with the pathological interpretation of hypotheses in the diagnostic universe: a non-exhaustive universe may indicate that only healthy or only pathological conditions are covered. In contrast, we would usually expect an exhaustive universe to contain at least one element that represents a healthy condition.

Example 6.24 *Let Y be the group of internal clinical attributes that describe VSD pathology, i.e. $Y = \{$*`VSD_size`, `VSD_type`, `VSD_ext`, `outlet_pos`*$\}$. The set*

$$\Psi_t \;=\; \{\psi \mid \psi \in \Omega_{Y(t)}\}$$

is a universe of static and mutually-exclusive disorder descriptions at time point $t \in T$ that is exhaustive. Note that some hypotheses in Ψ_t are redundant: the attribute `VSD_ext` *for instance, can only have the value 'none' when* `VSD_type` $\neq$

perimembranous (See Example 6.9 on page 187). All hypotheses in Ψ_t that are consistent with

$$\texttt{VSD\_size}(t) = \mathit{null} \wedge \texttt{outlet\_pos}(t) = \mathit{normal}$$

represent healthy conditions; all others represent pathological conditions where (at least) a VSD exists.

The set

$$\Psi'_t \;=\; \{\psi \mid \psi \in \Omega_{Y[0,t]}\}$$

is a universe of dynamic disorder descriptions with the same properties. In Ψ'_t, even much more hypotheses are redundant. For instance, the attribute `VSD_type` *is static, but Ψ'_t includes all scenarios where the attributes changes over time. Now, let $Y = \{\texttt{VSD\_size}, \texttt{outlet\_pos}\}$ and $Z = \{\texttt{VSD\_type}, \texttt{VSD\_ext}\}$. The set*

$$\Psi''_t \;\subseteq\; \Psi_t \cup \{\psi_1 \wedge \psi_2 \mid \psi_1 \in \Omega_{Y[0,t]}, \psi_2 \in \Omega_{Z(t)}\}$$

is a universe of (partially) dynamic disorder descriptions; the elements of this universe may no longer be mutually exclusive. The pathological interpretation of hypotheses in Ψ'_t and Ψ''_t is similar to the interpretation of hypotheses in Ψ_t. The set Ψ''_t is exhaustive if $\Psi''_t = \Omega_{Y[0,t]\cup Z(t)}$.

Diagnostic explanations

Once it is decided which universes of diagnostic hypotheses are considered, we can focus on diagnostic reasoning. There are basically two conceptions of diagnostic reasoning, a narrow and a broad conception. In the narrow conception, diagnostic reasoning is restricted to determining the likely cause(s) of observed findings. This is purely a matter of inference and does not involve decision making. The broad conception of diagnosis also considers the selection of test actions to gather additional observations on the patient's state. Diagnosis is then regarded as a sequential decision problem where diagnostic inference and test selection alternate, (Gorry and Barnett, 1968).

Here, we focus on the narrow conception of diagnostic reasoning. We will introduce a number of abstract concepts related to this task; throughout we discuss the possibilities one has to make these concepts concrete when implementing a particular form of diagnosis. We remark that the analyses here start with a given model of control Ξ and the assumption that this model is known to the diagnosing clinician; he is only limited in his information about the internal state of the patient. Furthermore, we assume that Ψ_t is a universe of mutually exclusive and jointly exhaustive diagnostic hypotheses. From the exhaustiveness of Ψ_t, it follows that its elements describe both healthy and pathological conditions; we do not make assumptions regarding the temporal scope of hypotheses in Ψ_t.

We say that hypothesis $\psi \in \Psi_t$ is a *potential explanation* of the evidence χ when it might have produced that evidence under action sequence α.

Definition 6.25 (Explanation) *Let $\alpha \wedge \chi$ be the available knowledge at time point $t \in T$. Diagnostic hypothesis $\psi \in \Psi_t$* potentially explains *the evidence χ under action sequence α when*

$$\xi_\alpha^t(\psi \wedge \chi) \;>\; 0, \tag{6.27}$$

where ξ_α^t is the control distribution for time point t under action sequence α. The set of all potential explanations of χ is denoted by $expl_t(\alpha, \chi)$.

The set $expl_t(\alpha, \chi)$ of potential explanations may be empty for two reasons. The first reason is that $\xi_\alpha^t(\chi) = 0$, and therefore there are no potential explanations of evidence χ: the model does not cover the possibility that χ occurs. The patient under consideration is therefore someone outside the target population of our formalisation.

Example 6.26 *Thrill (abnormal vibrations of the heart) and a typical systolic murmur heard by auscultation of the heart are characteristic signs of a VSD. Any patient having a large VSD with moderate heart failure in the first year of life will show these signs. A patient that seems to have heart failure but misses thrill and murmur will not have a VSD, and is therefore not part of the target population of our formalisation of the VSD domain:*

$$\xi_\alpha^t(\texttt{thrill}(t) = no \wedge \texttt{sys\_mur}(t) = no \wedge \texttt{LV\_failure}(t) = moderate) \;=\; 0$$

for all action sequences $\alpha \in \mathcal{A}[0, t-1]$ and all time points $t \in T$ where $\tau(t) \leq 12$ (months).

The second reason is that although the patient is part of our target population, i.e. $\xi_\alpha^t(\chi) > 0$, there is no diagnostic hypothesis in the universe Ψ_t that is consistent with it, i.e. $\xi_\alpha^t(\psi \wedge \chi) = 0$. This situation can only occur when the universe Ψ_t is not exhaustive. Depending on the pathological interpretation of the hypotheses in Ψ_t, it may indicate that the patient is healthy, or rather diseased, but further information is not available at this point.

In what follows, we will assume that $\xi_\alpha^t(\chi) > 0$, and that the set $expl_t(\alpha, \chi)$ of potential explanations is non-empty. We note that $expl_t(\alpha, \chi)$ is often much greater than the set $\Psi_t \setminus expl_t(\alpha, \chi)$ of diagnostic hypotheses that are rejected under evidence ψ. Furthermore, the latter set also bears interesting information, as these hypotheses identify what definitely *not* applies to the patient.

The concept of relevance

When uncertainty abounds in the domain of application, and this often occurs in clinical medicine, the set $expl_t(\alpha, \chi)$ will be large and therefore impractical. The solution that is generally employed in medicine is to order the set $expl_t(\alpha, \chi)$ using a concept of *relevance*. We thus obtain a list of diagnoses that are ranked from most to least relevant given the available knowledge with respect to the patient's state; such a list is called a *differential diagnosis*.

Definition 6.27 (Differential diagnosis) *Let $\alpha \wedge \chi$ be the available knowledge at time point $t \in T$. A* differential diagnosis *for α and χ is a set*

$$dd_t(\alpha, \chi) \quad = \quad \{(\psi, r) \mid \psi \in expl_t(\alpha, \chi), r \in [0, 1]\} \tag{6.28}$$

of pairs (ψ, r), where $r = r_{\alpha,\chi}(\psi)$ is the diagnostic relevance of hypothesis ψ in the light of α and χ.

There are several ways to implement the concept of relevance. Perhaps the most straightforward implementation is based on *likelihood*, or equally, posterior probability. The relevance of hypothesis $\psi \in expl_t(\alpha, \chi)$ is then defined as

$$r_{\alpha,\chi}(\psi) \quad = \quad \xi^t_\alpha(\psi \mid \chi). \tag{6.29}$$

The most relevant hypothesis is therefore the one that is most likely given the available knowledge. Note that this implementation of relevance implies that when $r_{\alpha,\chi}(\psi) = 1$ for some $\psi \in \Psi_t$, then all other explanations of χ are excluded from the differential diagnosis as they have zero probability. Conversely, when a differential diagnosis has only one element (ψ, r), this implies that the hypothesis ψ has unit relevance.

The differential diagnosis will usually be a starting point for further clinical action; this may consist of diagnostic testing to gather more information on the patient's condition, or may consist of therapy aimed to improve that condition. In both cases, the objective is to reach a better prognosis for the patient: indirectly through better opportunities to treat the patient with the information gathered, or directly through reaching a hopefully better health status. The above implementation of relevance does however not consider the patient's prognosis in ranking the explanations, and may therefore fall short in providing directions for further management.

The second implementation of relevance we describe is based on the impact of diagnostic information on the patient's prognosis, as expressed by expected utility, and in particular its impact on future decisions; we refer to it as *prognostic relevance*. Let

$$\tilde{u}^*(\varphi) \quad = \quad \max\{\tilde{u}_P(\varphi) \mid P \in \mathbb{P}^{\mathrm{det}}_{\Xi,O}, P(\varphi) > 0\} \tag{6.30}$$

be the maximum expected utility of the situation described by φ. That is, $\tilde{u}^*(\varphi)$ is a measure for the patient's prognosis under the best possible management for situation φ. With this notation, $\tilde{u}^*(\alpha \wedge \chi)$ is the maximum expected utility in the given decision-making context described by α and χ, and $\tilde{u}^*(\alpha \wedge \chi \wedge \psi)$ is the maximum expected utility when it is known that $\psi \in expl_t(\alpha, \chi)$ is the operational cause of the evidence χ. Note that when action $a \in A$ is an optimal choice for time point t in that situation, then

$$\tilde{u}^*(\alpha \wedge \chi \wedge \psi) \quad = \quad \tilde{u}^*(\alpha \wedge \chi \wedge \psi \wedge d(t){=}a) \tag{6.31}$$

as the function u^* implicitly assumes an optimal contingency plan to be followed, and $\alpha \wedge \psi$ provide sufficient knowledge for predicting the choice at time point t.

The prognostic relevance of hypothesis $\psi \in expl_t(\alpha, \chi)$ is now defined as

$$r'_{\alpha,\chi}(\psi) \;=\; \frac{\tilde{u}^*(\alpha \wedge \chi \wedge \psi) - \tilde{u}^-(\alpha \wedge \chi \wedge \psi)}{1 - \tilde{u}^-(\alpha \wedge \chi \wedge \psi)} \cdot \xi^t_\alpha(\psi \mid \chi), \tag{6.32}$$

where

$$\tilde{u}^-(\alpha \wedge \chi \wedge \psi) \;=\; \min\{\tilde{u}^*(\alpha \wedge \chi \wedge \psi \wedge d(t){=}a) \mid a \in A\} \tag{6.33}$$

is the maximum expected utility of the *worst* action choice at time point t when hypothesis ψ is true. That is, a diagnostic hypothesis $\psi \in expl_t(\alpha, \chi)$ is irrelevant (i.e. $r'_{\alpha,\chi}(\psi) = 0$), when the present action choice is uninfluential with respect to the patient's prognosis, i.e. $\tilde{u}^*(\alpha \wedge \chi \wedge \psi) = \tilde{u}^-(\alpha \wedge \chi \wedge \psi)$, or ψ does not explain the evidence, i.e. $\xi^t_\alpha(\psi \mid \chi) = 0$. The relevance of hypothesis ψ increases with the influence of the action choice to follow when ψ is the true cause, and the probability $\xi^t_\alpha(\psi \mid \chi)$ that ψ is indeed the case, given α and χ.

Two remarks are in place here. First, our notion of prognostic relevance at time point t is strongly connected to the action choice at that very moment, and disregards the effects of all other decisions. Theoretically speaking, a hypothesis that is considered virtually irrelevant now may be of the greatest importance at the next time point, and *vice versa*. Although this scenario is not to be expected in most domains, it does illustrate the temporal character of diagnostic concepts employed here. Second, the prognostic implementation of relevance may cause the correct hypothesis ψ to be excluded from the differential diagnosis, regardless of its likelihood: the fact that ψ is correct does not preclude that $\tilde{u}^*(\alpha \wedge \chi \wedge \psi) = \tilde{u}^-(\alpha \wedge \chi \wedge \psi)$. The likelihood-based implementation of relevance will in contrast always include the correct hypothesis.

Prognostic relevance reduces to likelihood-based relevance when

$$\tilde{u}^*(\alpha \wedge \chi \wedge \psi) \;=\; 1, \tag{6.34}$$

for all explanations $\psi \in expl_t(\alpha, \chi)$, meaning that knowing the cause of disease would always, and unconditionally, yield the opportunity of giving the patient the best possible prognosis; indeed, a highly unlikely situation. On the other hand, it must be acknowledged that the prognostic perspective may consider a highly probably (and maybe true) hypothesis as most irrelevant. This seems undesirable from a conceptual point of view; one could therefore choose to present both likelihood-based and prognostic relevances to the user of a decision-support system.

Conciseness in diagnosis

It is often complained that diagnostic programs produce moderately long lists of diagnoses, containing many diagnoses that a knowledgeable physician would regard as completely irrelevant, (Berner et al., 1994). When the evidence χ is little symptomatic (i.e. does not exclude many explanations from the differential diagnosis), our approach carries the risk of meeting a similar complaint.

To anticipate this situation, we now lift the requirement that all explanations of given evidence be contained in a differential diagnosis. That is, we define a differential diagnosis to be a subset

$$dd_t(\alpha, \chi) \quad \subseteq \quad \{(\psi, r) \mid \psi \in expl_t(\alpha, \chi), r > 0\} \tag{6.35}$$

of possible explanations and their associated relevances; one typically omits the explanations that are found to be little relevant. Of course, this carries the risk that the correct hypothesis is no longer included.

The most basic strategy is now to define a *relevance threshold* p, and to limit the explanations in the differential diagnosis to those whose relevance exceeds this threshold:

$$dd_t(\alpha, \chi) \quad = \quad \{(\psi, r) \mid \psi \in expl_t(\alpha, \chi), r > p\}. \tag{6.36}$$

The probability q that the correct hypothesis is still contained the differential diagnosis equals

$$q \quad = \quad \xi_\alpha^t(\bigvee_{(\psi,r)\in dd_t(\alpha,\chi)} \psi \mid \chi), \tag{6.37}$$

which can be rewritten to

$$q \quad = \quad \sum_{(\psi,r)\in dd_t(\alpha,\chi)} \xi_\alpha^t(\psi \mid \chi), \tag{6.38}$$

in the case that all hypotheses are mutually exclusive, as we have assumed here.

Now, if the implementation of relevance is based on likelihood, it is not very probable that we omit the correct hypothesis from the differential diagnosis if the threshold p is small. Yet, the prognostic implications of the unlikely hypotheses that are being omitted may be considerable. If, in contrast, the prognostic implementation of relevance is employed, we may be omitting all highly plausible hypotheses, possible including the correct one, simply because their impact on the future is modest. Again, this seems undesirable if we want the user of a decision support system to understand the system's recommendations.

We therefore propose to assemble differential diagnoses using a mixture of both approaches: while prognostic value is of principal importance for decision making, highly likely explanations should also be presented to the user. In this mixed approach, a differential diagnosis for α and χ is defined as

$$dd_t(\alpha, \chi) \quad \subseteq \quad \{(\psi, r) \mid \psi \in expl_t(\alpha, \chi)\},$$

where for each pair $(\psi, r) \in dd_t(\alpha, \chi)$ we have

$$\xi_\alpha^t(\psi \mid \chi) \geq p_1 \quad \text{or} \quad \tilde{u}^*(\alpha \wedge \chi \wedge \psi) - \tilde{u}^-(\alpha \wedge \chi \wedge \psi) \geq p_2,$$

where $0 < p_1 \ll 1$ is a likelihood threshold, and $0 < p_2 \ll 1$ is threshold for prognostic value.

Test selection

The established differential diagnosis $dd_t(\alpha, \chi)$ now provides a starting-point for choosing the action at time point t. Generally speaking, there are three possibilities for this choice. The first possibility is to decide that no action is required at the current time point (and possibly thereafter), and to select the empty action ε. This may be reasonable when the diagnosis indicates that the patient is in healthy condition, or when the prognostic relevance of all potential explanations of the evidence χ is low: this would indicate that acting is presently little helpful.

The second possibility is that the diagnosis indicates that therapy is required for the patient, because a disorder has been identified and there exists a remedial action $a \in A_{\text{treat}}$ for it. We will describe this possibility in more detail in the next subsection. It usually requires that one is sufficiently certain about the disorder; preferably, a single explanation in the differential diagnosis stands out with high relevance, or multiple explanations are equally relevant but require the same therapy. The third possibility is that one is not sufficiently certain about the disorder to start with therapy, and therefore wants to refine the differential diagnosis first. This requires a test action $a \in A_{\text{test}}$ to be selected at the present time point. We conclude this subsection on diagnosis with a brief and informal discussion of this possibility.

In general, we can say that the usefulness of a diagnostic test $a \in A_{\text{test}}$ is assessed by considering its expected gain in information compared to its associated costs and risks. The tradeoff in test selection therefore requires some *information measure* on differential diagnoses to be formulated. As with the concept of diagnostic relevance, information measures can be defined in purely probabilistic terms, but it is also possible to take a decision-theoretic approach where the (impact of diagnostic information on) the patient's prognosis is important. For a review of information measures in diagnostic test selection, we refer to (Glasziou and Hilden, 1989).

Now, if $\alpha \wedge \chi$ represents the available knowledge at time point t, and action a is chosen, then

$$Y \;=\; o_{t+1}(\alpha \wedge d(t){=}a) \setminus o_t(\alpha) \tag{6.39}$$

is the set of additionally observed variables at time point $t+1$. Depending on the configuration of Y that is actually observed, this may help to increase the diagnostic information. Of course, it is not known in advance which configuration will be observed; the probability of observing configuration $\mathrm{c}_Y \in \Omega_Y$ is

$$\xi^{t+1}_{\alpha \wedge d(t){=}a}(\mathrm{c}_Y \mid \chi), \tag{6.40}$$

and this leads to new, updated differential diagnosis $dd_{t+1}(\alpha \wedge d(t){=}a, \chi \wedge \mathrm{c}_Y)$ at time point $t+1$. Given a measure of information, one can derive the expected gain in information from action a by considering these possibilities.

We do not pursue the matter of test selection here, and conclude this subsection with noting that the dynamic perspective employed here involves that diagnostic

assessment is performed whenever new information comes available. In principle, this happens at each time point during the decision process; the diagnostic process is therefore never regarded as 'finished'.

6.2.2 Therapy selection

From a strictly decision-theoretic standpoint, there can only be one criterion for action selection in decision processes, which is the maximisation of expected utility. This holds for all actions, include those that pertain to therapy. Conceptually speaking, however, most clinical decisions are not guided by the objective of expected-utility maximisation, but are typically based on a mixture of considerations regarding the patient's condition, pathophysiological knowledge, the clinical modalities at hand, and experience with similar patients. Of course, such considerations are also implicitly present in any utility-theoretic tradeoff.

In this subsection, we describe a therapy selection method that starts from a symbolic description of pathophysiological knowledge. This knowledge is taken to be expressed as a contingency plan π with non-operational decision rules, that select action choices on the basis of diagnostic hypotheses. The underlying idea is as follows. In many clinical domains, most physicians would be able to indicate the best therapy for a given patient if the disease were known with certainty. Such indications are then simply derived from pathophysiological knowledge of the domain that provides for predicting the future course of disease with high confidence.

Example 6.28 *In the VSD domain, a clinician would normally submit a patient to surgery if he is sure that the patient's VSD is accompanied with a malalignment of the outlet septum or an aortic prolapse. It is known that in both cases, recovery of the patient is excluded without clinical intervention. Malalignments of the outlet septum prohibit spontaneous closure of the VSD; while this is not the case with prolapsed aortic valves, a prolapse will progressively cause aortic insufficiency and heart failure, even if the VSD closes spontaneously.*

Formally, we assume that the plan π comprises decision rules of the form

$$\psi \rightarrow d(t) = a \tag{6.41}$$

where $\psi \in \Psi_t$ describes a possible disorder at time point t, and $a \in A_{\text{treat}}$ is a treatment action. These rules are non-operational as the expressions in the set Ψ_t pertain to internal, hidden state variables: such as expressions are usually not part of the clinician's knowledge, thus prohibiting straightforward application of the plan in practice. There are no specific requirements with respect to the plan π in terms of completeness, consistency, or coherence: we only assume that it is a proper plan in the sense that its rules are non-contradictory.

We assume that some notion of diagnosis is employed which has yielded a differential diagnosis $dd_t(\alpha, \chi)$ from the knowledge $\alpha \wedge \chi$ that is available. Our aim now is to arrive, from this differential diagnosis, at a *differential indication* for time point t.

Definition 6.29 (Differential indication) *Let $\alpha \wedge \chi$ be the available knowledge at time point $t \in T$. A* differential indication *for time point t given α and χ is a function*

$$di_{\alpha,\chi} : A_{\text{treat}} \to [0, 1], \tag{6.42}$$

such that $\sum_{a \in A} di_{\alpha,\chi}(a) = 1$.

A differential indication $di_{\alpha,\chi}$ assigns a real number $di_{\alpha,\chi}(a)$ to each possible action $a \in A$, given α and χ: this number is interpreted as a measure of confidence that $d(t) = a$ is the best action choice in the given circumstances. The indication can thus be regarded as a stochastic decision rule for time point t with antecedent $\alpha \wedge \chi$. When $di_{\alpha,\chi}(a) = 1$ for some action $a \in A_{\text{treat}}$, the differential indication is equivalent to the (deterministic) decision rule $\alpha \wedge \chi \to d(t) = a$; we say that the indication is *unambiguous*. When there is less clarity, i.e. when $di_{\alpha,\chi}(a) > 0$ for multiple actions $a \in A_{\text{treat}}$, the differential indication does not provide an unambiguously preferred action choice, and a more elaborate tradeoff is necessary. In such cases, one has to resort to an explicit prognostic analysis.

We now propose the following therapy selection procedure. Let $\alpha \wedge \chi$ be the available knowledge at time point t, and let $dd_t(\alpha, \chi)$ be the associated differential diagnosis. It is immaterial which notion of diagnosis has lead to the realisation of $dd_t(\alpha, \chi)$; but we do assume that $dd_t(\alpha, \chi)$ is non-empty. The associated differential indication $di_{\alpha,\chi}$ is constructed as follows:

$$di_{\alpha,\chi}(a) \;=\; m \cdot \sum_{(\psi,r) \in dd_t(\alpha,\chi)} r \cdot g_\pi(\psi, a) \tag{6.43}$$

for all actions $a \in A_{\text{treat}}$, where

$$g_\pi(\psi, a) \;=\; \begin{cases} 1, & \text{if } \psi \to d(t) = a \in \pi, \\ 0, & \text{otherwise,} \end{cases} \tag{6.44}$$

and $m \in \mathbb{R}$ is a normalisation factor to ensure that $\sum_{a \in A_{\text{treat}}} di_{\alpha,\chi}(a) = 1$.

This procedure can be understood as follows. We consider all explanations in the differential diagnosis to derive an indication for treatment, and no other explanations. If, for some explanation ψ in $dd_t(\alpha, \chi)$, a decision rule $\psi \to d(t) = a$ exists in π, then $g_\pi(\psi, a) = 1$, and therefore the relevance r associated with ψ is added to the confidence we have in action a (Equation 6.43). The simplest case is therefore when

$$dd_t(\alpha, \chi) \;=\; \{(\psi, 1)\} \tag{6.45}$$

and

$$\psi \rightarrow d(t) = a \quad \in \quad \pi, \tag{6.46}$$

and thus the differential indication $di_{\alpha,\chi}$ selects action a unambiguously, $di_{\alpha,\chi}(a) = 1$, an intuitive result. The same indication would however be arrived at when

$$dd_t(\alpha, \chi) \quad = \quad \{(\psi, r)\} \tag{6.47}$$

for some $r > 0$, because the differential indication is being normalised after all explanations have been considered. When multiple explanations occur in the differential diagnosis and each of them has an associated decision rule in π, the differential indication will only be unambiguous when all these rules point at the same action choice. Otherwise, it will reflect the relevances of the corresponding explanations in the diagnosis.

The procedure described above may fail if for some reason (for instance a lack of pathophysiological knowledge), one is unable to formulate decision rules for particular disorders. The procedure will simply neglect these omissions in the plan π, as $g_\pi(\psi, a) = 0$ for all $a \in A_{\text{treat}}$ when $\psi \notin \text{cover}_t(\pi)$. Still, ψ may be a relevant explanation of the evidence χ, and call for remedial action. Below, we discuss two possible solutions to this problem.

The first solution consists of assuming a default level of confidence $0 < \epsilon \ll 1$ for each action $a \in A_{\text{treat}}$ in cases where an explanation in the differential diagnosis is not covered by the plan π. It requires the following modification to the function g_π:

$$g'_\pi(\psi, a) \quad = \quad \begin{cases} 1, & \text{if } \psi \rightarrow d(t) = a \in \pi, \\ \epsilon, & \text{if } \psi \notin \text{cover}_t(\pi), \\ 0, & \text{otherwise.} \end{cases} \tag{6.48}$$

Note that the third clause applies when there exists a rule $\psi \rightarrow d(t) = a' \in \pi$, $a' \neq a$, and therefore $\psi \in \text{cover}_t(\pi)$. The above modification to the function g has the effect of 'flattening' the differential indication in the presence of an explanation that is not covered by the plan π; extreme measures of confidence (0 and 1) are then excluded. The extent of flattening depends on the parameter ϵ and on the relevance of explanations for which no decision rule exists: the term $\epsilon \cdot r$ is added for each action $a \in A_{\text{treat}}$. Of course, a normalisation step follows thereafter. In case that $\pi = \varnothing$, we obtain $di_{\alpha,\chi}(a) = 1/k$ for all $a \in A_{\text{treat}}$, where $k = |A_{\text{treat}}|$.

The second solution we describe is based on the observation that in some cases, inadequate pathophysiological knowledge does not prohibit one to formulate *negative* decision rules, or, in the clinical jargon, *contra-indications*.

Example 6.30 *Surgical closure of a VSD is strongly discouraged in the presence of irreversible pulmonary arteriopathy. The increased resistance to blood flow in the pulmonary system will then cause a high pressure overload on the right ventricle, possibly resulting in acute cardiac death.*

We will now allow that the plan π also comprises decision rules of the form

$$\psi \rightarrow d(t) \neq a, \tag{6.49}$$

where again $\psi \in \Psi_t$ describes a possible disorder at time point t, and $a \in A_{\text{treat}}$ is a treatment action. This form of rule says that action choice $d(t) = a$ is dissuaded when ψ is the cause of disease. We allow that multiple contra-indications are associated with a single disorder ψ, but not that all possible action choices are dissuaded for ψ. When a 'normal', positive decision rule $\psi \rightarrow d(t) = a$ occurs in π, we do not allow any other rule, positive or negative, to occur in π having ψ as its antecedent.

Example 6.31 *Formally, the contra-indication described in Example 6.30 has the form*

$$\texttt{pulm\_art}(t_1) = \mathit{irreversible\_arteriopathy} \;\; \rightarrow \;\; d(t_2) \neq \mathit{surg}$$

for all time points $t_1, t_2 \in T$, $t_1 \leq t_2$.

The therapy selection procedure is modified as follows. First, a *penalty* is given to actions for which a contra-indication exists in π:

$$g''_\pi(\psi, a) \;\; = \;\; \begin{cases} 1, & \text{if } \psi \rightarrow d(t) = a \in \pi, \\ -1, & \text{if } \psi \rightarrow d(t) \neq a \in \pi, \\ 0, & \text{otherwise} \end{cases} \tag{6.50}$$

This means that we are basically weighing the arguments against and in favour of particular treatments, where the weights' sizes are obtained from the relevances of explanations in the differential diagnosis, and their signs are provided by decision rules. The modification does imply that a total negative weight may be assigned to some action. As this would conflict with our definition of differential diagnosis, normalisation now requires that

$$di''_{\alpha,\chi}(a) \;\; = \;\; m_1 \cdot \Big(\sum_{(\psi,r) \in dd_t(\alpha,\chi)} r \cdot g''_\pi(\psi, a) \;+\; m_2 \Big), \tag{6.51}$$

for all $a \in A_{\text{treat}}$, where $m_1, m_2 \in \mathbb{R}$ are normalisation factors.

We note that it is well possible to combine both solutions to the problem of inadequate knowledge; we do not pursue that possibility here. To conclude this subsection we remark that the procedure outlined above selects treatment actions directly on the basis of a differential diagnosis and a collection of heuristic rules that formalise a clinician's knowledge and experience. Consequently, the reliability of this procedure crucially depends on the quality of both, and on the quality of their synergy. If the relevance measure used in assembling the differential diagnosis is based on likelihood, then the decision rules should embody long-term prognostic knowledge; if the relevance measure itself is based on long-term prognosis, then the rules may be aimed at a short-term result.

Finally, even if the quality of both sources and their synergy is high, the procedure may fail to provide a pronounced indication for therapeutic action. Such happens, for instance, when there are multiple explanations in the diagnosis that are more or less equally relevant, but different treatment actions are associated with them. A more elaborate prognostic analysis of the possible choices is then required; this is discussed in the next subsection.

6.2.3 Prognostic assessment

In this subsection we consider the final reasoning task in clinical medicine: *prognostication*, assessing a patient's prognosis. The notion of prognosis has already occurred in the context of diagnosis, where it was used as a measure for the relevance of diagnostic hypotheses. It was however given a mostly numerical characterisation: the maximum expected utility was taken to characterise the patient's prognosis. In this section, we investigate a more symbolic approach to the notion of prognosis.

We develop a notion of prognosis that is patient-centred and builds on the concepts of *future scenario* and *differential prognosis*. Here, a future scenario is an explicit symbolic description of future events, and a differential prognosis is a set of such scenarios and their associated importances in the given clinical situation. The differential prognosis aims to provide a picture of the patient's prospects, where the associated expected utility provides a summarising aggregate of these prospects. We will see that this approach yields a number of similarities between the notions of diagnosis and prognosis.

Future scenarios

In general, the set $\mathcal{H}[t+1, N]$ of planning histories over time segment $[t+1, N]$ denotes the full range of prospective possibilities at time point t; of course, it will often occur that many of these possibilities are highly implausible or even excluded in a given situation. We can nevertheless think of $\mathcal{H}[t+1, N]$ to represent all potential future scenarios at time point t. However, the level of detail within these scenarios is too high. When the length of time segment $[t+1, N]$ is considerable, each history $h \in \mathcal{H}[t+1, N]$ contains a torrent of information, and the number of possible histories is enormous. It is then simply not feasible to form a picture of the patient's prospects by inspecting all scenarios that seem reasonable in a given situation.

We therefore seek to design a universe Ω_t of future scenarios at time point t at a higher level of abstraction than $\mathcal{H}[t+1, N]$. The elements of Ω_t are planning expressions that describe a number of events during time segment $[t+1, N]$, but much less than complete planning histories over this time segment: information that is irrelevant for prognostic purposes is omitted. We assume that the elements of Ω_t are mutually exclusive, i.e. $\omega_1 \wedge \omega_2 \equiv \bot$ for all scenarios $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$, and jointly exhaustive, i.e. for all histories $h \in \mathcal{H}[t+1, N]$ there exists a scenario $\omega \in \Omega_t$ such that $h \vdash \omega$.

It must now be decided which information is irrelevant for prognostic purposes. There exist multiple ways of doing so, and the tradeoff involved is ultimately subjective. Yet, we can establish guidelines for separating relevant from irrelevant prognostic information. We describe an informal and a formal guideline here.

The informal guideline is based on the general objective of therapy and the characteristic prospects for patients in the domain under consideration. If the therapy is curative and there is good hope of obtaining a healthy condition at the end of the management process, one will focus on the internal clinical attributes, as these attributes represent the primary factors of the patient's condition. Furthermore, the final outcome will be considered more relevant than the trajectory leading to it: both clinical states prior to the final time point $t = N$ and clinical actions required to reach the healthy condition may be largely neglected. So, in this case, the universe Ω_t of future scenarios would contain expressions that mostly refer to variables from $I(N)$. If, on the other hand, the prospects are worse and the current management strategy is palliative rather than curative in character, external attributes, especially those pertaining to the symptoms of disease, may considered as more relevant. In addition, events at near rather distant time points will be emphasised, and there will be more attention to clinical actions if they involve a lot of discomfort; the patient may wish to forego such actions unless they really improve his prospects. In this case, the universe Ω_t would contain expressions that refer to a number of variables from $E[t+1, N]$, most decision variables from $D[t+1, N]$, and probably also a small number of variables from $I[t+1, N]$.

The formal guideline for separating relevant from irrelevant prognostic information is based on the notion of *direct contribution* to utility from Definition 5.89 on page 159. Recall that the elements of Z are direct contributors to utility when $Z \subseteq X(T) \cup D(T)$ is the smallest set such that

$$u(\mathrm{C}_Z \wedge \mathrm{c}_Y) \;=\; u(\mathrm{C}_Z \wedge \mathrm{c}'_Y) \tag{6.52}$$

for all configurations $\mathrm{c}_Y, \mathrm{c}'_Y$ of the complementary set $Y = (X(T) \cup D(T)) \setminus Z$, consisting of indirect contributors to utility. It depends on the type of utility function employed which variables contribute directly to utility; for example, a utility function based solely on final-state goals (as described in Section 6.1.5), would only have variables from $I(N)$ contribute directly. In general, variables from the set

$$Z_t \;=\; Z \cap (X[t+1, N] \cup D[t+1, N]) \tag{6.53}$$

are prognostic variables at time point t that contribute directly to utility. Given a future configuration $\mathrm{c}_{Z_t} \in \Omega_{Z_t}$ of this set, the other prognostic variables are irrelevant for determining the associated utility. From a formal standpoint, the set Ω_{Z_t} of all configurations of Z_t is therefore a most suitable universe of future scenarios.

We note that ideally, both guidelines lead to the same set of variables that is considered relevant for prognostic purposes: a utility function should reflect one's attitude towards future developments.

Differential prognosis

We now turn to the notion of *differential prognosis*, which is defined analogously to the notion of differential diagnosis; the roles of diagnostic explanations are filled by future scenarios. As before, we assume that the physician faces a choice at time point $t \in T$, while having performed action sequence $\alpha \in \mathcal{A}[0, t-1]$ and obtained evidence $\chi \in \Phi(T)$. The choice at time point t is taken to be restricted to some candidate set $A_t \subseteq A$, $|A_t| \geq 1$, of actions that have previously been selected, for instance using the procedure described in the previous subsection. In addition, it is assumed that some future management strategy for time points $t+1, \ldots, N$, is available; only the action choice at the current time point t, or equivalently, the decision rule for time point t with antecedent $\alpha \wedge \chi$, has yet to be decided upon. It is irrelevant what the strategy for time point $t+1, \ldots, N$, looks like: it may range from complete absence of action (i.e. $d(t') = \varepsilon$ for all $t' \in [t+1, N]$) to a detailed plan of action that anticipates many contingencies. In either case, we take a partial contingency plan π, consisting of decision rules for time points $t+1, \ldots, N$, to describe the strategy; we think of plan π as the available knowledge of the future.

Now, note that at time point t, it is immaterial which considerations have led to choosing the actions $\alpha(0), \ldots, \alpha(t-1)$ in sequence α. These choices will probably have resulted from following some decision-making strategy, but with hindsight, the only relevant fact is that this has led to choosing the actions mentioned. A similar observation holds for the action choice at time point t. In a prognostic evaluation of decision $d(t) = a$, we do not care about the reasons for this choice, but only about its consequences. We can therefore think of the physician as choosing from the set

$$\mathbb{P} \;=\; \{P_a \in \mathbb{P}^{\mathrm{det}}_{\Xi,O} \mid a \in A_t\} \tag{6.54}$$

of decision processes, where each $P_a \in \mathbb{P}$ implements action sequence α, decision $d(t) = a$, and plan π. With each action choice $d(t) = a$, $a \in A_t$, is now also associated a differential prognosis that is based on the universe of future scenarios Ω_t and whose concrete form derives from P_a; the decision among these choices is made by comparing prognoses. The notion of differential prognosis is defined as follows.

Definition 6.32 (Differential prognosis) *Let Ω_t be the universe of future scenarios at time point $t \in T$, $t < N$. A* differential prognosis *at time point t is a set*

$$dp_t \;=\; \{(\omega, q) \mid \omega \in \Omega_t,\, q \in [0,1]\}, \tag{6.55}$$

of pairs (ω, q), where q is called the prognostic importance *of scenario ω in differential prognosis dp_t.*

As with differential diagnoses, where only explanations of nonzero relevance were included, we assume that $q > 0$ for each element (ω, q) of a differential prognosis dp_t. Another correspondence between the two notions is that there exist different ways

for obtaining the numbers associated with the elements of differential sets. We have earlier discussed several implementations of the concept of diagnostic relevance; we will now discuss three possibilities to implement the notion of prognostic importance.

The first, and most obvious, implementation is based on the likelihood of future scenarios. From the above observation is follows that, given a particular choice of action $a \in A_t$, the likelihood q of future scenario $\omega \in \Omega_t$ equals

$$q \;=\; P_a(\omega \mid \chi), \tag{6.56}$$

where $P_a \in \mathbb{P}$. This implementation of importance emphasises the probability of future scenarios: the most important scenario ω is the one that is most likely given the available knowledge about past and future. When scenario ω has unit likelihood ($q = 1$), this implies that ω is also the only future scenario in the differential prognosis dp_t^a associated with decision $d(t)=a$, i.e. $dp_t^a = \{(\omega, 1)\}$. Such may occur when we are approaching the final point in the time frame, or when many internal variables have reached an absorbing state. Otherwise, there will be multiple scenarios with nonzero likelihood; one may choose to omit highly unlikely scenarios to obtain a succinct picture of the patient's prospects.

The likelihood-based implementation of prognostic importance does not reflect the preferences for different future scenarios: it neglects utility. The second implementation of importance we discuss in contrast neglects likelihood, and focuses completely on utility: it regards future scenarios with little utility as most important, as these scenarios are to be avoided. That is, the importance q associated future scenario ω now equals

$$q \;=\; \begin{cases} 1 - \tilde{u}_{\chi\wedge\omega}(P_a), & \text{if } P_a(\omega \mid \chi) > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{6.57}$$

where

$$\tilde{u}_{\chi\wedge\omega}(P_a) \;=\; \sum_{h\in\mathcal{H}(T)} P_a(h \mid \chi, \omega) \cdot u(h) \tag{6.58}$$

is the expected conditional utility of decision process P_a given χ and ω, as defined in Definition 5.87 on page 158. This implementation of importance provides for making *worst-case* and *best-case* analyses, by considering the smallest and largest importances found in a differential prognosis. If the difference between these two extremes is small, then the patient faces a stable future in terms of his preferences; note that this also occur when there is much uncertainty regarding which *concrete* scenario is to be expected. Furthermore, the importances in a differential prognosis will not sum to unity, and at least theoretically, all importances may be 0 or 1.

A disadvantage of the utility-based implementation of importance is the fact that scenarios marked by very high or low expected utility may have a very low likelihood; the insight gained from worst-case and best-case analyses is then rather limited. The

third implementation of importance therefore synthesises the two implementations described above by weighing utility with likelihood. That is, the importance q associated future scenario ω is defined as

$$q = \begin{cases} P_a(\omega \mid \chi) \cdot (1 - \tilde{u}_{\chi \wedge \omega}(P_a)), & \text{if } P_a(\omega \mid \chi) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{6.59}$$

With this implementation, a future scenario obtains high importance only if it is likely but undesirable; other scenarios are considered less important. An noticeable property is now that the importances in the differential prognosis dp_t^a associated with decision $d(t)=a$ will sum to the complement of $\tilde{u}_\chi(P_a)$, the expected conditional utility of decision process P_a given χ:

$$\begin{aligned} \sum_{(\omega,q) \in dp_t^a} q &= \sum_{\omega \in \Omega_t} P_a(\omega \mid \chi) \cdot (1 - \tilde{u}_{\chi \wedge \omega}(P_a)) \\ &= 1 - \sum_{\omega \in \Omega_t} P_a(\omega \mid \chi) \cdot \tilde{u}_{\chi \wedge \omega}(P_a) \\ &= 1 - \tilde{u}_\chi(P_a). \end{aligned} \tag{6.60}$$

As such, this implementation of importance reflects the decision-theoretic viewpoint; an optimal action choice $a^* \in A_t$ is one that minimises this sum, and therefore maximises expected utility:

$$\begin{aligned} a^* &= \text{argmin}_{a \in A_t} \{ \sum_{(\omega,q) \in dp_t^a} q \} \\ &= \text{argmin}_{a \in A_t} \{1 - \tilde{u}_\chi(P_a)\} \\ &= \text{argmax}_{a \in A_t} \{\tilde{u}_\chi(P_a)\}. \end{aligned} \tag{6.61}$$

We note that it depends on the purpose of prognostication which implementation of prognostic importance is most suitable. When the purpose is to select an action that is optimal in decision-theoretic terms, then the third implementation, that is based on weighed utilities of future scenarios, is the best choice. Prognostic information can however also have a value in itself, without contributing to optimal decision making, (Asch et al., 1990). When the purpose of prognostication is merely to inform the patient on his prospects, the likelihood-based and utility-based implementation of prognostic importance may be preferable.

Concluding remarks

We conclude this subsection with some further remarks on the approach to prognostication that was described above.

A first remark concerns the fact that with the likelihood-based implementation of importance in differential prognoses, we have basically returned to the formal foundation of decision theory: we can then interpret a differential prognosis dp_t as a

simple lottery over the universe Ω_t of future scenarios. Formally, the corresponding lottery is

$$l \quad = \quad (q_1, \omega_1; \ldots; q_n, \omega_n), \tag{6.62}$$

if $\Omega_t = \{\omega_1, \ldots, \omega_n\}$, and q_i, $i = 1, \ldots, n$, is the importance (i.e. likelihood) associated with scenario ω_i. As we constructed a differential prognosis for each possible action choice $d(t) = a$ at time point t, this means that we have reduced this decision to a choice between lotteries, which is also the starting-point of utility theory.

A second remark concerns the fact that multiple situations may give rise to equal expected utilities for two actions $a_1, a_2 \in A$, i.e. to

$$\tilde{u}_\chi(P_{a_1}) \quad = \quad \tilde{u}_\chi(P_{a_2}). \tag{6.63}$$

The first situation is where all future scenarios are expected with the same likelihoods for both actions, and the action choice itself does not affect utility. One can then truly say that there is absolutely no difference between these actions in the given situation. The second situation is where differences in likelihood only pertain to variables that do not influence utility directly. This may happen, for instance, if the utility function focuses on a small number of attributes that have already reached an absorbing state at time point t. The fact that other attributes may show considerable differences in their probable developments under actions a_1 and a_2 will then be masked by the selective nature of the utility function. The third situation, finally, is where also the probable developments for direct contributors to utility differ, but the weighed sum of utilities associated with future scenarios is still the same. In utility-theoretic terms, this means that one is indifferent between the lotteries represented by prognoses $dp_t^{a_1}$ and $dp_t^{a_2}$.

Finally, we have seen that the considerations that surround the notion of differential prognosis are roughly similar as those surrounding the notion of differential diagnosis, discussed in Subsection 6.2.1. Notwithstanding these similarities, however, there is an importance difference between diagnostic and prognostic reasoning. The diagnostic task concerns inferring the probable or relevant causes of given evidence; this is often called *evidential reasoning*, as reasoning proceeds from the evidence. Prognosis, in contrast, concerns the prediction of effects from given causes; this is therefore often called *causal reasoning*. In a probabilistic formalisation, as we employ here, both types of reasoning ultimately reduce to probabilistic inference. Yet, evidential reasoning seeks to uncover what is truly the case, and can strictly speaking come up with only one correct answer. Causal reasoning, in contrast, is unavoidably at error when it comes up with only one answer in a domain with inherent uncertainty: the essence of the future is contingency.

6.3 Discussion

In this chapter we have described how the framework of Chapter 5 can be employed to implement the dynamic perspective on patient management that was sketched in Chapter 1. Many notions related to clinical modelling and reasoning have passed in review in the preceding sections. It has, however, by no means been our intention to be in some sense 'complete'; we have shown how many clinical decision-making notions can be formalised within our framework, but equally many notions have been left out of consideration. Instead, we have aimed to stress that there is often room for substantial variation when formalising clinical notions within the framework; our purpose has been to illustrate the choices one has during the formalisation process, and in particular, how the effects of such choices are analysed.

The first part of the chapter concentrated on translating the ingredients of a given clinical domain to formal structures of our framework. In Section 6.1 we have successively considered the formal description of a patient's clinical condition, modalities for the treating physician, prognostic aspects of the domain at hand, and findings from diagnostic investigations. In addition, we have compared two approaches to formalising the objectives of patient management: one approach, from the field of symbolic planning, based on the identification of goal states, and one approach based on duration and quality of life, from clinical decision analysis. It was concluded that, notwithstanding the conceptual differences, these two approaches may ultimately coincide. Overall, we have tried to give the reader a feeling for the role of the framework's components in clinical decision-support settings.

The second part of the chapter focused on the clinical tasks of diagnosis, therapy selection, and prognosis; in particular, we have formally investigated the roles of these tasks within the dynamic perspective on patient management that is advocated in this thesis. From a strictly decision-theoretic standpoint, there can only be one criterion for action selection in decision processes, which is the maximisation of expected utility. Conceptually speaking, however, clinical patient management is not guided by the objective of expected-utility maximisation, but typically based on considerations regarding the explanatory power of competing diagnostic classifications, pathophysiological knowledge of the domain, experience with the application of therapy, and future scenarios for the patient's development. We have shown how these considerations can be made explicit in the reasoning process, without losing the underlying decision-theoretic tradeoff.

We now conclude the chapter by covering a subject that has been largely neglected: the difficult practice of clinical modelling. Furthermore, we re-evaluate our framework in the light of the experience gained, and compare our work with related investigations by other authors.

Name: `LV_failure`

Definition: left-ventricular failure

Group: 5 (signs and symptoms)

Domain:

none	
mild	(tiredness, increased heart rate and breathing frequency)
moderate	(also feeding problems, failure to thrive)
severe	(also pulmonary rales)

Type: ordered categorical

Comments: Tiredness, feeding problems and "failure to thrive" (growing problems, unhappiness) are subjective measures. Furthermore, heart rate is subject to large variations.

Figure 6.3: Model description of the attribute `LV_failure`.

The practice of clinical modelling

Constructing a model from medical reality is often hard and laborious; it is sometimes considered to be the major obstacle building decision-support systems. We have hardly touched upon this issue in the preceding sections, as it is beyond the scope of this thesis. Yet, to give an impression of the type of difficulties one may encounter, we will here discuss three problems we had during the modelling of the VSD domain.

The first problem concerned the proper description of clinical states. We believe that it is generally recommendable to mimic the clinical vocabulary as much as possible when a domain is formalised. Unfortunately, one may then encounter problems with finding precise, unambiguous definitions for clinical attributes and their value domains. This is due to the fact that there are often concepts that are used in everyday clinical practice whose definition is rather imprecise. In the VSD domain, we encountered this problem with the 'fuzzy' concept *heart failure*.

As described in Chapter 2, the heart is said to fail when it cannot fulfil its primary function, circulating the blood through the body. Recall that this condition may be the result of left-to-right shunting of blood through a VSD. However, the very moment that there are signs of heart failure, the body will respond by preventing a decrease in arterial blood flow and pressure: the heart rate will increase and the systemic vascular resistance is increased by narrowing many small blood vessels in the body. So the effects of shunting on the circulation are immediately compensated for, and it is really the effects of the compensatory mechanisms (e.g., shortness of breath) that are observable to the clinician. We therefore decided to define the notion of heart failure purely on the basis of these symptomatic effects; see the full model description for the attribute in Figure 6.3 for details.

The second and third problems concerned the construction of models of control. For the VSD domain, we have constructed a dynamic belief network (as described in Section 4.3) to model the probable developments in clinical state over time. In building

a belief network, two closely related tasks can be discerned: the construction of the graphical part of the network, and its subsequent quantification. Constructing the graphical part of a network consists of designing a directed acyclic graph that adequately models the conditional independence relations holding between the domain variables; the quantification of a probabilistic network consists of assessing the probabilistic parameters that are needed to calibrate the probability distribution modelled by the network.

Both tasks turned out to be, to some extent, problematic. The construction of the graph was hampered by the circumstance that the human cardiovascular system contains several homeostatic feedback mechanisms. It is often not possible to identify conditional independence relations between the variables involved, and therefore all these variables (representing flows, pressures, and resistances in different parts of circulation) will have to be directly connected to each other in the graph. This causes another problem, as the number of probabilistic parameters associated with network variables grows exponentially in their numbers of predecessor variables. Furthermore, it is generally unclear what the direction of the arcs between such variables should be, as their relationship is best characterised as one of continual mutual interaction. For a more elaborate discussion of the problem related to modelling the cardiovascular system with belief networks, we refer to the papers by Peek and Ottenkamp (1997) and Long et al. (1997).

Even more difficult it was to quantify the network thus obtained. This was, in fact, little surprising and certainly not unique for the VSD application: belief-network quantification requires the assessment of numerous probabilities, and is therefore a notorious bottleneck in building belief-network applications, (Druzdzel et al., 1995). Probability assessment is a topic of active methodological research in the field of uncertainty reasoning, (e.g., see Spiegelhalter et al., 1990; Druzdzel and Van der Gaag, 1995; Renooij and Witteman, 1999). In medical domains, probabilities are normally obtained from frequencies reported in the literature, statistics from clinical databases, or subjective estimation by clinical experts; it is also possible to employ a combination of sources.

During the construction of the dynamic belief network for VSD, a recently proposed method for efficient belief-network quantification was experimentally verified with a (static) part the network. This method, developed by Coupé et al. (1998), is based on performing *sensitivity analyses* to identify the most influential probabilistic parameters. The main idea is to quantify a belief network by iteratively refining highly-influential probabilities, while the network is initially supplied with rough estimates. Instead of putting an effort in obtaining accurate probability estimates for all network parameters, this approach thus focuses on those parameters that are most influential on the network's performance. The procedure was experimentally verified on a belief network for the VSD domain that was supplied with subjective probabilities. The results suggest that the approach is indeed successful: a satisfying network performance is obtained when a limited number of highly-influential probabilities is

elicited from a "high-quality source" (in this case, a field expert) while the others are obtained from less informed sources. However, the effects of the refinements on network performance were found to be non-monotonic; in practice it is therefore difficult to establish the number of refinements that is required. For further details of this investigation, we refer to (Coupé et al., 1999) and (Peek et al., 1999).

Re-evaluation of the framework

From the experience gained in this chapter with applying our framework for decision-theoretic planning in clinical settings, we can now make a re-evaluation. We list some strengths and weaknesses of the framework that have shown up in the preceding sections.

A first and foremost strength of the framework is definitely the fact that time is a primitive, not a derived, concept, and that all compound concepts involve references to time. This allows to describe clinical patient management in a highly dynamic fashion; notions of time and change are not restricted to isolated parts of the formalisation, as occurs in most decision-analytic tools (e.g. Hazen, 1992; Sonnenberg and Beck, 1993). The second strength is the expressive symbolic planning language $\Phi(T)$ within the framework. This enabled us, in Section 6.2, to formalise the notions of diagnosis and prognosis as collections of past and future scenarios respectively, thus emphasising their dynamic characters. A final advantage of the framework is that it leaves room for approaching the concepts that occur in clinical reasoning from multiple perspectives within a single formalisation. In the formalisms discussed in Chapter 4 for instance, this is only possible by introducing additional representations of the same problem.

Of course, we have also seen that the current framework has several limitations. First, all attributes from the set X occur as clinical variables at all time points. This is sometimes unnatural from a conceptual point of view, and somewhat awkward in formal respects. In some situations, particular attributes are redundant and do not have a clear interpretation. For instance, when there is no VSD, it seems strange to speak of its size and type. Yet, the attributes `VSD_size` and `VSD_type` necessarily occur at all time points in the VSD model. Second, the point-based formalisation of time in our framework is limited in its expressiveness and flexibility; this required, for instance, the introduction of a temporal map τ to model the "true" passage of time between decision moments. Third and finally, the probabilistic foundation for modelling the effects of action choices is semantically rather weak. As we saw in Subsection 6.1.3, one would preferable distinguish between decisive and additional preconditions, and target and complicatory postconditions of clinical actions; from a probabilistic standpoint, however, this distinction is meaningless.

Related work

We conclude with a brief comparison of the work by Magni and Bellazzi (1997; 1998), Leong (1998a; 1998b), and Hauskrecht (1997a; 1997c), with our investigation.

P. Magni and R. Bellazzi (1997; 1998) propose to employ fully-observable Markov decision processes for therapy planning in medicine. They explore a novel graphical formalism, called *influence view*, for describing the transition probability functions in an MDP model. Roughly, an influence view is a dynamic belief network with two time slices that models the transition probability function associated with a given action. In addition to the variables describing the patient's clinical states, influence views contain *context variables* for population-specific parametrisation of the transition probabilities, and *transition variables* that facilitate the modelling process. The technique has been applied on the problem of therapy planning for patients with *hereditary spherocytosis*. With moderate and severe forms of this disease, surgical removal of the spleen is recommended, but the mild variant is more difficult to decide upon. Traditional decision-analytic approaches based on decision trees typically model static therapies where surgery is considered only once at some given age of the patient. The authors show the MDP approach allows to reconsider surgery each year: a dynamic therapy.

T.-Y. Leong (1998a; 1998b) has developed a framework planning under uncertainty called DYNAMOL (for DYNamic decision MOdeling Language). Similar to our framework, it integrates ideas from AI, decision theory, and control theory, and is aimed at analysing and supporting decision making in clinical medicine. In DYNAMOL, planning problems are cast as semi-Markov decision processes, a generalisation of fully-observable Markov decision processes where the state transitions induced by action choices take a variable, and stochastic, amount of time; the underlying notion of time is however discrete. Considerable attention is spent in the DYNAMOL framework to the graphical representation of transition probability functions.

An important drawback to the work of both Magni and Bellazzi, and Leong, is the assumption of full observability at all times. This is awkward in clinical medicine, and therefore yields a rather simplified description of reality: informational relations among patient, disease, and physician cannot be analysed. The approach precludes, for instance, to analyse the notion of diagnosis in clinical reasoning, in the manner of Subsection 6.2.1.

M. Hauskrecht (1997a; 1997c) employs partially-observable Markov decision processes (POMDPs) to model clinical management problems. As he correctly states, the issue of imperfect information in clinical management is often significant, and can therefore not be neglected. Hauskrecht also employs a graphical representation of transition probability functions. This representation has a unique hierarchical feature: the state description is considered at multiple levels of complexity, where state attributes are only included when they are relevant for the clinical condition at hand. Such an approach solves the above described problem that all clinical attributes occur as variables at all time points in our framework. The work of Hauskrecht focuses on methods for approximately solving POMDP problems where the additional hierarchical structure is exploited; it leaves most issues regarding clinical modelling and reasoning unaddressed.

A drawback to all the work described above is that it hardly extends beyond the classical description of Markov decision processes. The primary additions are graphical representations, while an expressive language for decision-theoretic planning concepts is lacking. In this chapter, we have shown that our framework is more expressive, and thus allows a thorough analysis of these concepts in clinical modelling and reasoning.

CHAPTER 7

Conclusions

This chapter finishes the thesis with a review of the topics that have been in addressed; throughout, we summarise the main conclusions of our work. As each of the Chapters 2 to 6 has already provided a detailed discussion of results obtained, a general perspective is adopted here. We conclude with some directions for future research.

The dynamic perspective on patient management

The starting point of our investigation has been a dynamic perspective on clinical patient management, where doctor, patient, and disease engage in a process of continual interaction: the doctor responds to observed signs, symptoms, and results of diagnostic procedures by taking appropriate clinical action, and the patient's condition changes over time in response to the doctor's actions. Within this process, the tasks of diagnostic assessment, therapy selection, and prognostication are intertwined activities, and do not form separate phases in the management procedure. The doctor is viewed as solving a sequence of similar and mutually related decision problems over time; this task was characterised as action planning under uncertainty with partial information and temporal constraints.

Decision-theoretic representation and reasoning

We have employed Bayesian decision theory at the fundamental level of trading off alternative choices in decision making. Decision theory builds on probability theory for reasoning with uncertainty, and utility theory for rational choice under uncertainty.

The traditional tool for decision-theoretic analysis is the decision tree; decision trees were used to illustrate the decision-theoretic analyses in Chapter 3. Decision trees provide an intuitive representation of decision problems and can easily be constructed in cooperation with field experts. They are however not suited as a knowledge-representation formalism for automated reasoning systems, because they restrict to a single problem case and adhere to a single viewpoint on that problem. Furthermore, decision trees explicitly enumerate all potential scenarios in a decision problem, yielding a representation that grows exponentially in size with the size of the problem. These observations motivated a review of other decision-theoretic representation formalisms in Chapter 4.

Three representation formalisms were reviewed: *influence diagrams*, *Markov decision processes*, and *dynamic influence diagrams*. Influence diagrams provide a concise way of representing decision-theoretic problems by exploiting probabilistic independencies between variables in the problem domain. They do have the drawback however of mixing different types of knowledge in a single representation, which hampers their expressiveness; in addition, they do not include notions of time or change. Markov decision processes are mathematical models of stochastic control that incorporate the decision-theoretic perspective on choice tradeoffs; the partially-observable variant (POMDP) adheres to the clinical situation where part of a problem's information is hidden for the decision maker. Markov decision processes have explicit notions of time and change, but their representation of temporal progression is rather coarse. A more delicate representation is found in dynamic influence diagrams. These diagrams extend traditional influence diagrams with a notion of time, which makes them suitable for representing decision-theoretic planning problems. Dynamic influence diagrams can also be integrated with Markov decision processes, thus combining both approaches.

Integrating planning and decision theory

The core of our own work was presented in Chapters 5 and 6. In Chapter 5, we presented a formal framework for decision-theoretic planning that was subsequently used to model and analyse problems of clinical patient management in Chapter 6. The framework integrates notions from decision theory, uncertainty reasoning, and symbolic planning, and provides a broader perspective on decision-theoretic planning that the representation formalisms of Chapter 4. In brief, its building blocks consist of a *decision basis*, models of *control* and *observation*, and a *utility function*. The three elements X, A, and T of a decision basis describe the fundamental components of a problem domain: the dynamic system under partial control by the planning agent, the possible actions to choose from, and the time frame for the planning task. These elements give rise to a propositional algebra that serves as a *planning language* for the domain, and allows for direct manipulation of symbolic structures that describe the logical relations between states, events, observations, decisions and plans.

A second fundamental notion in our framework is the *decision process*. A decision

process provides a meta-level description of the intertwined dynamics of planning agent and stochastic system. Using the rules of probability theory, we can switch from this meta-level description to object levels that focus on the agent's behaviour (using choice predictions) or on the system's reactions to that behaviour (using state predictions). Furthermore, by analysing conditional independence relations in a decision process, we can identify the influential and informational relationships between system states and decisions: this provides for characterising the extent to which action choices can change the system's development, and the extent to which observed state information is used in making these choices.

Contingency planning

A significant part of Chapter 5 was devoted to a theory of *contingency planning*. In our framework, plans are expressed as collections of symbolic decision rules. These rules allow for easy communication with field experts and can be directly employed, for instance, in clinical guidelines. Furthermore, this form of contingency plan enables us to formulate decision-making strategies at different levels of detail and with varying ranges of applicability. We investigated properties of contingency plans related to their completeness, consistency, coherence, and operationality under a given model of observation. In addition, it was analysed how decision processes can implement contingency plans, and shown how the properties mentioned above then appear as properties from these decision processes. We conclude that a successful synthesis of symbolic and numerical approaches to planning has been obtained in our framework, and that this synthesis provides for thoroughly analysing the concepts that come into play when time and change are essential in decision making under uncertainty.

Explicit prognostic models

In Chapter 6 we discussed how the framework for decision-theoretic planning is applied to model medical-clinical domains, and to implement patient management tasks. The first part of the chapter concentrated on *modelling*: translating the ingredients of a given clinical domain to formal structures. A notable property of our framework is that it allows to develop an explicit prognostic model that describes the probable developments in a patient's clinical condition over time. Prognostic models found in the literature usually concentrate on a single aspect of the patient's condition (most often, mortality), and supplement a decision-making scheme without being part of it, (Wyatt and Altman, 1995; Lucas and Abu-Hanna, 1999). In our approach, prognosis concerns all of the patient's condition, and is fully integrated with the model for decision making. While this approach is more involved and requires a more elaborate domain specification to be assessed, it is also more realistic from a clinical point of view: prognosis concerns both near and distant future, and multiple aspects of the patient's condition should be taken into account in prognostication.

Explicit clinical reasoning

A leading theme in our work has been the explicit representation of concepts that are involved in decision making and planning under uncertainty. Decision-theoretic reasoning is characterised by the fact that each situation of choice is ultimately reduced to a utility-theoretic tradeoff; from a conceptual point of view, however, these choice situations may be very different. A number of concepts, pertaining to the role of actions, information and time, are significant in many planning problems; our aim has been to allow for explicit reasoning with these concepts, while adhering to the decision-theoretic perspective.

The second part of Chapter 6 focused on the medical reasoning tasks of diagnosis, therapy planning, and prognosis. We investigated how these tasks can be described within our framework, and how they relate to each other in the dynamic perspective on patient management. The clinical management of patients is often based on considerations regarding the explanatory power of competing diagnostic classifications, pathophysiological knowledge of the domain, experience with the application of therapy, and future scenarios for the patient's development. It was shown how these considerations can be made explicit in the reasoning process, without losing the underlying decision-theoretic tradeoff. It is our firm belief that such will facilitate the user's understanding, and hence acceptance, of the recommendations of a decision-support system.

Future research

A number of topics in our work have received only modest attention and require further investigation; there are also several starting-points for continuation. We describe the main directions for future research.

Computational methods A most glaring omission in our work has been the analysis of computational methods for decision making. It was already noted in Chapter 3 that practical application of decision-theoretic methods is hampered by the fact that decision-theoretic reasoning is often highly combinatorial, because all contingencies, with respect to both future events and future decisions, have to be taken into account when making a choice. As discussed in Chapter 4, this circumstance has been found to impede the application of partially-observable Markov decision processes in practice, because the associated methods for constructing optimal decision-making policies are intractable. We have decided to drop the matter of computation in our investigation, and focus on representation and reasoning. For any practical application of the results, however, further inquiries regarding computational methods are indispensable.

We believe that a promising approach to enhancing the feasibility of computational methods is the exploitation of domain-dependent constraints, specialised medical knowledge, and clinical experience. For instance, the theory of contingency planning from Section 5.3 paves the way for incremental procedures of plan construction, as it

shows how a collection of decision rules may serve as a partial solution to the planning problem, and indicates which extensions lead to a complete solution. As such, one may start with a given collection of rules that are employed by field experts, and extend this collection to obtain a complete contingency plan; this obviates the need to exhaustively search the space of all plans. A related approach starts from rules that require some modification, for instance because they are yet non-operational. An example of this type of approach was already given in Subsection 6.2.2, where a collection of such non-operational decision rules was used for decision-making by relating them to a given differential diagnosis.

Another possibility is to place formal constraints on the type of contingency plan allowed as a solution to planning problems. In the VSD domain for instance, it is reasonable to assume that surgery is required at most once for a given patient, and that surgery is never considered before routine clinical investigations have been performed. This restricts the space of possible plans to inspect when devising a management strategy; we have earlier investigated this approach in the context of POMDPs, (Peek, 1999b). Placing such constraints can be regarded as employing an informed heuristic method for decision making; it may be able to compete well with exact reasoning methods. As was shown in Chapter 6, our framework allows to incorporate and analyse such heuristics.

Representation and modelling A second omission in Chapters 5 and 6 has been the concrete representation of the framework's components. Before any practical application can be realised, data structures must be designed that can describe these components; this holds in particular for models of control. Note that under restricted conditions (as described in Section 5.5) it may be possible to use a POMDP or dynamic influence diagram for this purpose; in general, however, this is not possible as our framework is less restrictive than these representation formalisms. For the VSD domain, we have employed a dynamic belief network to represent the model of control, where a number of probability assessment functions are parametrised to specify the influence of actions on state variables; unfortunately, the belief network was not yet completed as of writing this thesis. The approach seems however reasonable in general, but must be further investigated.

A related issue is the fact that structured methods for knowledge acquisition and modelling are required to make practical applications of the framework within reach; we have already touched upon this issue in the concluding discussion of the preceding chapter. Again, this concerns in particular models of control, as these models require the assessment of many probabilistic parameters. Furthermore, the framework presently does not incorporate a mechanism to update its probabilistic parameters in the wake of new evidence. Such a mechanism is however needed allow for handling inter-patient variability. For instance, where some VSD patients show many symptoms of little left-to-right shunting, others are in better condition and show little symptoms of much shunting. The treating physician will often discover this, and adjust his expectations with respect to the patient in question.

Extending the framework An interesting topic for future research is extending the current framework to increase its expressiveness and range of applicability. It was already noted that the representation of time in our framework is not very flexible and may hamper temporal reasoning in practice. This drawback might be obviated by employing an interval-based temporal representation, where the underling time axis is continuous. In addition, the planning language may be extended to allow for more expressive temporal reasoning. Possibilities are the inclusion of existential and universal temporal quantifiers, and the addition of modal operators to capture a dynamic perspective on time.

It is also possible to investigate a multi-agent approach to decision making and planning in our framework. In many domains, including clinical medicine, planning tasks that are accomplished in collaboration or competition with others. This requires an extension of the framework with multiple, parallel decisions at each time point. Furthermore, one will need to model how information is exchanged between the agents making these decisions, and what their respective planning objectives are.

Other applications in health care Finally, the current framework has other potential applications in clinical health care that have not been addressed here, but may provide fruitful lines of future research. In particular, the framework provides starting-points to analyse existing clinical practice or clinical guidelines. This is envisioned as follows. To analyse existing clinical practice, a registration of decisions in a real-world clinical setting, over some period of time, is required. This registration can be summarised as a decision process in our framework, and then allows for analysing the decision-making behaviour. One may investigate the consistency of the recorded behaviour, its implicit clinical objectives, and the information that is used in making decisions. An existing set of clinical guidelines is analysed by translating them to a contingency plan in our framework; they are then amenable to formal study. The plan may be characterised in terms of completeness, consistency, and coherence, and again one can investigate the implicit objectives of the guidelines, and which information is used in making decisions.

APPENDIX A

Glossary of medical terms

aorta (Ao)	largest artery of the systemic circulation
aortic insufficiency	abnormal regurgitant flow of blood through an incompetent aortic valve
aortic stenosis (AoS)	narrowing at or near the aortic valve orifice
aortic prolapse	downward displacement of aortic cuspal material
aortic valve	valve between left ventricle and aorta
arteriole	small artery
atrioventricular valves	heart valves between atria en ventricles
atrial septum	the wall that separates left and right atrium
atrial septal defect (ASD)	abnormal opening in the atrial septum
atrium	upper heart chamber
capillary	minute blood vessel
cardiac cycle	the sequence of contraction and relaxation of the cardiac muscle
cardiac output	the volume of blood ejected by the heart per minute

cardiomegaly	pathological enlargement of the heart
cardiovascular system	the circulatory system of heart and vessels
coarctation	narrowing of the aortic arch
cyanosis	bluish discolouration of skin and mucous membranes
diastole	the phase of relaxation in the cardiac cycle
ductus arteriosus	short vessel connecting PA and Ao before birth
Eisenmenger's syndrome	the complex of findings associated with a VSD, severe pulmonary hypertension, and cyanosis
foramen ovale	communication between the two atria of the foetal heart
heart failure	ineffective pumping of the heart
hepatomegaly	enlarged liver
hypertrophy	overgrowth of cardiac muscle
mitral valve	valve between left atrium and ventricle
myocardium	heart muscle
oedema	accumulation of fluid in body tissue
outlet septum	the part of the ventricular septum between both semilunar valves
overriding aorta	misplacement of the aorta above the ventricular septum
persistent ductus arteriosus (PDA)	failure of closure of the ductus arteriosus at birth
pulmonary arteriopathy	a pathological condition of arterioles
pulmonary artery (PA)	short wide vessel arising from the right ventricle and conveying deoxygenated blood to the lungs
pulmonary circulation	the circulation of blood through the lungs
pulmonary hypertension	abnormally elevated blood pressure within the pulmonary circulation
pulmonary insufficiency	abnormal regurgitant flow of blood through an incompetent pulmonary valve

pulmonary stenosis	narrowing at or near the pulmonary valve orifice
pulmonary valve	valve between right ventricle and pulmonary artery
semilunar valves	valves between ventricles and great arteries
shunt	a passage or anastomosis between two natural channels or chambers
stenosis	localised narrowing or stricture of a duct or canal
stroke volume	the volume of blood ejected by the heart at one stroke
systemic circulation	the circulation of blood through the body
systole	the phase of contraction of the heart
Tetralogy of Fallot	a congenital disorder which consists of pulmonary stenosis, VSD, overriding of the aorta and right-ventricular hypertrophy
thrill	abnormal vibrations of the heart
tricuspid valve	valve between right atrium and ventricle
ventricle	lower heart chamber
ventricular septum	the wall that separates left and right ventricle
ventricular septal defect (VSD)	abnormal opening in the ventricular septum

APPENDIX B

Network quantifications

In this appendix we list the quantifications of the belief network from Figure 4.1 and the influence diagram from Figure 4.3. These estimates were provided by J. Ottenkamp, a senior paediatric cardiologist of the Leiden University Medical Center in the Netherlands. For each conditional probability $P(\mathrm{c}_x \mid \mathrm{c}_{\rho_G(x)})$, we was asked to estimate the number of patients satisfying the statement c_x, out of a given hunderd patient conforming to the statements in configuration $\mathrm{c}_{\rho_G(x)}$. Consequently, we do not list probabilities here but numbers in the range $0 \ldots 100$. The utility function for the influence diagram is based on interviews with the cardiologist on the VSD domain; we will motivate it at the definition.

B.1 Belief network

`VSD` (existence and size of a VSD)

prior distribution of `VSD` values			
none	*small*	*moderate*	*large*
0	35	40	25

`resis` (pulmonary vascular resistance)

prior distribution of `resis` values			
normal	*increased*	*high*	*very_high*
82	10	7	1

`closure` (spontaneous closure of the VSD)

given	conditional distribution of `closure` values	
`VSD`	*true*	*false*
none	100	0
small	95	5
moderate	70	30
large	25	75

`shunt` (size and direction of shunting)

given		conditional distribution of `shunt` values				
`VSD`	`resis`	*none*	*small*	*moderate*	*large*	*reversed*
none	anything	100	0	0	0	0
small	*normal*	0	98	2	0	0
	increased	0	99	1	0	0
	high	0	100	0	0	0
	very_high	2	89	0	0	9
moderate	*normal*	0	40	40	20	0
	increased	0	45	40	15	0
	high	0	60	30	10	0
	very_high	3	64	18	5	10
large	*normal*	0	10	30	60	0
	increased	0	15	35	50	0
	high	0	30	40	30	0
	very_high	6	59	14	9	12

`pmhyp` (pulmonary hypertension)

given		conditional distribution of `pmhyp` values			
`resis`	`shunt`	*absent*	*mild*	*moderate*	*severe*
normal	*none*	99	1	0	0
	small	94	5	$\frac{1}{2}$	$\frac{1}{2}$
	moderate	98	2	0	0
	large	95	3	2	0
increased	*none*	3	95	2	0
	small	30	60	$9\frac{1}{2}$	$\frac{1}{2}$
	moderate	40	55	3	2
	large	60	39	1	0
high	*none*	1	2	95	2
	small	20	45	30	5
	moderate	5	30	65	0
	large	30	65	5	0
very_high	*none*	0	0	2	98
	small	10	20	45	25
	moderate	2	10	40	48
	large	8	30	52	10
anything	*reversed*	0	0	0	100

`hfail` (heart failure)

given	conditional distribution of `hfail` values			
`shunt`	*absent*	*mild*	*moderate*	*severe*
none	98	1	$\frac{1}{2}$	$\frac{1}{2}$
small	92	5	2	1
moderate	15	70	10	5
large	5	30	50	15
reversed	1	14	35	50

`pmart` (pulmonary arteriopathy)

given		conditional distribution of `pmart` values	
`closure`	`shunt`	*true*	*false*
true	*none*	0	100
	small	0	100
	moderate	1	99
	large	5	95
	reversed	80	20
false	*none*	0	100
	small	5	95
	moderate	15	85
	large	$57\frac{1}{2}$	$42\frac{1}{2}$
	reversed	100	0

`death` (death)

We assume that the predecessors `pmhyp` and `hfail` of `death` are independent contributors to mortality risks and use a multiplicative model to combine their influence on `death`.

given	conditional distribution of `death` values	
`pmhyp`	*true*	*false*
absent	0	100
mild	$\frac{1}{2}$	$99\frac{1}{2}$
moderate	2	98
severe	5	95

given	conditional distribution of `death` values	
`hfail`	*true*	*false*
absent	0	100
mild	0	100
moderate	1	99
severe	4	96

For example: the mortality risk associated with mild pulmonary hypertension and moderate heart failure is $1 - (0.995 \cdot 0.99) \approx 0.015$, and the risk associated with moderate pulmonary hypertension and severe heart failure is $1 - (0.98 \cdot 0.96) \approx 0.06$.

B.2 Influence diagram

The influence diagram of Figure 4.3 extends the belief network with three decision nodes and a value node. We provide the extended tables for the variables `closure` and `death` (to which a decisional predecessor has been added), and the utility function.

`closure` (spontaneous closure of the VSD)

given		conditional distribution of `closure` values	
`VSD`	`surg`	*true*	*false*
none	*yes*	100	0
small		99	1
moderate		95	5
large		85	15
none	*no*	100	0
small		95	5
moderate		70	30
large		25	75

`death` (death)
We extend the multiplicative model for the variable `death` with providing the risk factors associated with cardiac catheterisation and surgery.

given	conditional distribution of `death` values	
`cath`	*true*	*false*
no	0	100
yes	$\frac{1}{2}$	$99\frac{1}{2}$

given	conditional distribution of `death` values	
`surg`	*true*	*false*
no	0	100
yes	5	95

utility function

The predecessors of the value node are the random variables `closure`, `pmart`, and `death`, and the decision variables `echo`, `cath`, and `surg`. The utility function is calibrated on a scale of 0 to 100, where 0 is associated with the least, and 100 is associated with the most preferred outcome. A utility of 0 is associated with all outcomes where `death` = *true*.

given						
`death`	`pmart`	`closure`	`surg`	`cath`	`echo`	utility
false	*false*	*true*	*no*	*no*	*no*	100
					yes	99
				yes	*no*	96
					yes	95
			yes	*no*	*no*	90
					yes	89
				yes	*no*	86
					yes	85
false	*false*	*false*	*no*	*no*	*no*	90
					yes	89
				yes	*no*	86
					yes	85
			yes	*no*	*no*	80
					yes	79
				yes	*no*	76
					yes	75
false	*true*	*false*	*no*	*no*	any	30
				yes	any	29
			yes	any	any	25
false	*true*	*true*	*no*	any	any	20
			yes	any	any	15
true	any	any	any	any	any	0

Bibliography

Agosta, J.M. (1996). Constraining influence diagram structure by generative planning: an application to the optimization of oil spill response, *Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence (UAI–96)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 11–19.

Aliferis, C.F. and Cooper, G.F. (1996). A structurally and temporally extended Bayesian belief network model: definitions, properties, and modeling techniques, *Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence (UAI–96)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 28–39.

Aliferis, C.F., Cooper, G.F., Pollack, M.E., Buchanan, B.G. and Wagner, M.M. (1997). Representing and developing temporally abstracted knowledge as a means towards facilitating time modeling in medical decision-support systems, *Computers in Biology and Medicine* **27**(5): 411–434.

Allen, J. (1984). Towards a general theory of action and time, *Artificial Intelligence* **23**: 123–154.

Allen, J., Hendler, J. and Tate, A. (eds) (1990). *Readings in Planning*, Morgan Kaufmann Publishers, Palo Alto, California.

Anderson, R.H., Macartney, F.J., Shinebourne, E.A. and Tynan, M. (eds) (1987). *Paediatric Cardiology*, Vol. I, Churchill Livingstone, Edinburgh.

Andreassen, S., Leibovici, L., Schonheyder, H.C., Kristensen, B., Riekehr, C., Kjær, A.G. and Olesen, K.G. (1998). A decision theoretic approach to empirical treat-

ment of bacteraemia originating from the urinary tract, *Proceedings of the Workshop on Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP–98)*, pp. 48–53.

Andreassen, S., Woldbye, M., Falck, B. and Andersen, S.K. (1987). MUNIN - a causal probabilistic network for interpretation of electromyographic findings, *Proceedings of the Tenth International Joint Conference on Artificial Intelligence (IJCAI–87)*, pp. 366–372.

Asch, D.A., Patton, J.P. and Hershey, J.C. (1990). Knowing for the sake of the knowing: the value of prognostic information, *Medical Decision Making* **10**: 47–57.

Aström, K.J. (1965). Optimal control of Markov processes with incomplete state information, *Journal of Mathematical Analysis and Applications* **10**: 174–205.

Barnett, V. (1982). *Comparative Statistical Inference*, John Wiley & Sons, New York.

Bell, D.E. and Raiffa, H. (1988). Risky choice revisited, *in* Bell et al. (1988), chapter 5, pp. 99–112.

Bell, D.E., Raiffa, H. and Tversky, A. (eds) (1988). *Decision Making. Descriptive, Normative and Prescriptive Interactions*, Cambridge University Press, Cambridge.

Bellazzi, R., Berzuini, C., Quaglini, S., Spiegelhalter, D.J. and Leaning, M. (1991). Cytotoxic chemotherapy monitoring using stochastic simulation on graphical models, *in* M. Stefanelli, A. Hasman, M. Fieschi and J. Talmon (eds), *Proceedings of the Third Conference on Artificial Intelligence in Medicine*, Vol. 44 of *Lecture Notes in Medical Informatics*, Springer Verlag, Berlin, pp. 227–238.

Bellman, R.E. (1957). *Dynamic Programming*, Princeton University Press, Princeton, New Jersey.

Berg, M. (1997). *Rationalizing Medical Work: Decision-Support Techniques and Medical Practices*, MIT Press, Cambridge, Massachusetts.

Berner, E.S., Webster, G.D., Shugerman, A.A., Jackson, J.R., Algina, J., Baker, A.L., Ball, E.V., Cobbs, C.G., Dennis, V.M., Frenkel, E.P., Hudson, L.D., Mancall, E.L., Rackley, C.E. and Taunton, O.D. (1994). Performance of four computer-based diagnostic systems, *New England Journal of Medicine* **330**(25): 1792–1796.

Bertsekas, D.P. (1995). *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, Massachusetts.

Berzuini, C. (1990). Representing time in causal probabilistic networks, *in* Henrion et al. (1990), pp. 15–28.

Bielza, C., Ríos Insua, S. and Gómez, M. (1999). Influence diagrams for neonatal jaundice management, *in* Horn et al. (1999), pp. 138–142.

Birkhoff, G. and MacLane, S. (1977). *A Survey of Modern Algebra*, MacMillan, New York.

Blackwell, D. (1965). Discounted dynamic programming, *Annals of Mathematical Statistics* **36**: 226–235.

Boutilier, C., Brafman, R.I. and Geib, C. (1997). Prioritized goal decomposition of Markov decision processes: toward a synthesis of classical and decision theoretic planning, *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI–97)*, pp. 1156–1162.

Boutilier, C., Dean, T. and Hanks, S. (1999). Planning under uncertainty: structural assumptions and computational leverage, *Journal of Artificial Intelligence Research* **11**: 1–94.

Braunwald, E. (1974). Regulation of the circulation, *New England Journal of Medicine* **290**: 1124.

Breese, J.S., Goldman, R.P. and Wellman, M.P. (1994). Introduction to the special section on knowledge-based construction of probabilistic and decision models, *IEEE Transactions on Systems, Man, and Cybernetics* **24**(11): 1577–1579.

Cassandra, A.R., Littman, M.L. and Zhang, N.L. (1997). Incremental pruning: a simple, fast, exact method for partially observable Markov decision processes, *Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–97)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 54–61.

Chernoff, H. and Moses, L.E. (1959). *Elementary Decision Theory*, John Wiley & Sons, New York.

Cooper, G.F. (1988). A method for using belief networks as influence diagrams, *Proceedings of the Workshop on Uncertainty in Artificial Intelligence*, pp. 55–63.

Cooper, G.F. (1990). The computational complexity of probabilistic inference using Bayesian belief networks, *Artificial Intelligence* **42**: 393–405.

Coupé, V.M.H., Peek, N.B., Ottenkamp, J. and Habbema, J.D.F. (1999). Using sensitivity analysis for efficient quantification of a belief network, *Artificial Intelligence in Medicine* **17**: 223–247.

Coupé, V.M.H., Van der Gaag, L.C. and Habbema, J.D.F. (1998). Sensitity analysis: an aid for belief-network quantification, *Technical Report UU–CS–98–10*, Department of Computer Science, Utrecht University.

Dagum, P., Galper, A. and Horvitz, E. (1992). Dynamic network models for forecasting, *Proceedings of the Eighth Annual Conference on Uncertainty in Artificial Intelligence (UAI–92)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 41–48.

Danford, D.A., Martin, A.B., Fletcher, S.E., Gumbiner, C.H., Cheatham, J.P., Hofschire, P.J. and Kugler, J.D. (1997). Children with heart murmurs: can ventricular septal defect be diagnosed reliably without an echocardiogram?, *Journal of the American College of Cardiology* **30**(1): 243–246.

Darwiche, A. and Pearl, J. (1994). Symbolic causal networks for reasoning about actions and plans, *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI–94)*, pp. 238–244.

Dawid, A.P. (1979). Conditional independence in statistical theory (with discussion), *Journal of the Royal Statistical Society (Series B)* **41**: 1–31.

de Dombal, F.T., Horrocks, J.C., Walmsley, G. and Wilson, P.D. (1975). Computer-aided diagnosis and decision-making in the acute abdomen, *Journal of Royal College of Physicians of London* **9**(3): 211–218.

de Dombal, F.T., Leaper, D.J., Staniland, J.R., McCann, A.P. and Horrocks, J.C. (1972). Computer-aided diagnosis of acute abdominal pain, *British Medical Journal* **2**: 9–13.

de Finetti, B. (1970). *Theory of Probability*, John Wiley & Sons, New York.

de Kleer, J., Mackworth, A.K. and Reiter, R. (1992). Characterizing diagnoses and systems, *Artificial Intelligence* **52**: 197–222.

Dean, T. and Kanazawa, K. (1989). A model for reasoning about persistence and causation, *Computational Intelligence* **5**: 142–150.

Dean, T. and Wellman, M. (1991). *Planning and Control*, Morgan Kaufmann, San Mateo, California.

Debreu, G. (1954). Representation of a preference ordering by a numerical function, *in* R.M. Thrall, C.H. Coombs and R.L. Davis (eds), *Decision Processes*, John Wiley & Sons, New York.

Dowie, J. and Elstein, A. (eds) (1988). *Professional Judgment: A Reader to Clinical Decision Making*, Cambridge University Press, Cambridge.

Drake, A.W. (1962). *Observation of a Markov Process through a Noisy Channel*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Draper, D., Hanks, S. and Weld, D. (1993). Probabilistic planning with information gathering and contingent execution, *Technical report 93-12-04*, Department of Computer Science and Engineering, University of Washington.

Druzdzel, M.J. and Van der Gaag, L.C. (1995). Elicitation of probabilities for belief networks: combining qualitative and quantitative information, *Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence (UAI–95)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 141–148.

Druzdzel, M.J., Van der Gaag, L.C., Henrion, M. and Jensen, F.V. (eds) (1995). *Working Notes of the Workshop on Building Probabilistic Networks: Where do the Numbers come from?*, Workshop held in conjunction with IJCAI–95, Montreal, Quebec, Canada.

Egar, J.W., Puerta, A.R. and Musen, M.A. (1992). Graph-grammar assistance for modeling of decisions, *Proceedings of Knowledge Acquisition Workshop (KAW–92)*, SRDG Publications, Calgary, pp. 7.1–7.19.

Fikes, R. and Nilsson, N.J. (1971). STRIPS: a new approach to the application of theorem proving to problem solving, *Artificial Intelligence* **2**: 189–208.

Finger, J.J. (1987). *Exploiting Constraints in Design Synthesis*, Ph.D. thesis, Stanford University.

Fishburn, P.C. (1970). *Utility Theory for Decision Making*, John Wiley & Sons, New York.

Fishburn, P.C. (1988). Normative theories of decision making under uncertainty, *in* Bell et al. (1988), chapter 4, pp. 78–98.

Franklin, R.C.G., Spiegelhalter, D.J., Macartney, F. and Bull, K. (1989). Combining clinical judgements and statistical data in expert systems: over the telephone management for critical congenital heart disease in the first month of life, *International Journal of Clinical Monitoring and Computing* **6**: 157–166.

Franklin, R.C.G., Spiegelhalter, D.J., Macartney, F. and Bull, K. (1991). Evaluation of a diagnostic algorithm for heart disease in neonates, *British Medical Journal* **302**: 935–939.

Ganong, W.F. (1997). *Review of Medical Physiology*, Appleton & Lange, Stamford, Connecticut.

Garson, A., Bricker, J.T. and McNamara, D.G. (eds) (1990). *The Science and Practice of Paediatric Cardiology*, Vol. II, Lea & Febiger, Philadelphia.

Geiger, D., Verma, T. and Pearl, J. (1990). d-Separation: from theorems to algorithms, *in* Henrion et al. (1990), pp. 139–148.

Geva, T., Hegesh, J. and Frand, M. (1988). Reappraisal of the approach to the child with heart murmurs: is echocardiography mandatory?, *International Journal of Cardiology* **19**: 107–114.

Glasziou, P. and Hilden, J. (1989). Test selection measures, *Medical Decision Making* **9**: 133–141.

Goldsmith, J. and Mundhenk, M. (1998). Complexity issues in Markov decision processes, *Proceedings of the IEEE Conference on Computational Complexity*.

Gorry, G.A. (1973). Computer-assisted clinical decision making, *Methods of Information in Medicine* **12**: 45–51.

Gorry, G.A. and Barnett, G.O. (1968). Experience with a model of sequential diagnosis, *Computers and Biomedical Research* **1**: 490–507.

Graham, T.P. and Gutgesell, H.P. (1995). Ventricular septal defects, *in* Moss et al. (1995), pp. 724–746.

Grimmett, G.R. and Stirzaker, D.R. (1992). *Probability and Random Processes*, Oxford University Press, Oxford.

Gumbiner, C.H. and Takao, A. (1990). Ventricular septal defect, *in* Garson et al. (1990), pp. 1002–1022.

Guyton, A.C. (1986). *Textbook of Medical Physiology*, Saunders, Philadelphia.

Habbema, J.D.F., Bossuyt, P.M.M., Dippel, D.W.J., Marshall, S. and Hilden, J. (1990). Analysing clinical decision analyses, *Statistics in Medicine* **9**: 1229–1242.

Habbema, J.D.F. and Hilden, J. (1981). The measurement of performance in probabilistic diagnosis. IV. Utility considerations in therapeutics and prognosis, *Methods of Information in Medicine* **20**: 80–96.

Haddawy, P. (1996). A logic for time, chance, and action for representing plans, *Artificial Intelligence* **80**: 243–308.

Haddawy, P. and Hanks, S. (1998). Utility models for goal-directed decision-theoretic planners, *Computational Intelligence* **14**(3).

Hanks, S., Madigan, D. and Gavrin, J. (1995). Probabilistic temporal reasoning with endogenous change, *Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence (UAI–95)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 245–253.

Harary, F. (1969). *Graph Theory*, Addison Wesley, Reading, Massachusetts.

Hauskrecht, M. (1997a). Dynamic decision making in stochastic partially observable medical domains: ischemic heart disease example, *in* Keravnou et al. (1997), pp. 296–299.

Hauskrecht, M. (1997b). Incremental methods for computing bounds in partially observable Markov decision processes, *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI–97)*, pp. 734–739.

Hauskrecht, M. (1997c). *Planning and Control in Stochastic Domains with Imperfect Information*, Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hauskrecht, M. (1998). Hierarchical solution of Markov decision processes using macro-actions, *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–98)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 220–229.

Hazen, G.B. (1992). Stochastic trees: a new technique for temporal decision modeling, *Medical Decision Making* **12**: 163–178.

Heckerman, D.E. (1991). *Probabilistic Similarity Networks*, MIT Press, Cambridge, Massachusetts.

Heckerman, D.E., Horvitz, E.J. and Nathwani, B.N. (1992). Toward normative expert systems: part I, The Pathfinder project, *Methods of Information in Medicine* **31**: 90–105.

Heckerman, D.E. and Shachter, R.D. (1995). Decision-theoretic foundations for causal reasoning, *Journal of Artificial Intelligence Research* **3**: 405–430.

Helmcke, F., deSouza, A., Nanda, N.C., Villacosta, I., Gatewood Jr., R., Colvin, E. and Soto, B. (1989). Two-dimensional and color Doppler assessment of ventricular septal defect of congenital origin, *American Journal of Cardiology* **63**: 1112–1116.

Henrion, M., Shachter, R.D., Kanal, L.N. and Lemmer, J.F. (eds) (1990). *Uncertainty in Artificial Intelligence*, Vol. 5, North-Holland, Amsterdam.

Hershey, J.C. and Baron, J. (1987). Clinical reasoning and cognitive processes, *Medical Decision Making* **7**: 203–211.

Hickam, D.H., Shortliffe, E.H., Bischoff, M.B., Scott, A.C. and Jacobs, C.D. (1985). The treatment advice of a computer-based cancer chemotherapy protocol advisor, *Annals of Internal Medicine* **103**: 928–936.

Hilden, J., Glasziou, P.P. and Habbema, J.D.F. (1992). A pitfall in utility assessment – patients' undisclosed investment decisions, *Medical Decision Making* **12**(1): 39–43.

Hopkins, W.E. (1995). Severe primary pulmonary hypertension in congenital heart disease, *Current Opinion in Cardiology* **10**: 517–523.

Hopkins, W.E., Ochoa, L.L., Richardson, G.W. and Trulock, E.P. (1996). Comparison of the hemodynamics and survival of adults with severe primary pulmonary hypertension or Eisenmenger syndrome, *Journal of Heart and Lung Transplantation* **15**: 100–105.

Horn, W., Shahar, Y., Lindberg, G., Andreassen, S. and Wyatt, J. (eds) (1999). *AIMDM '99: Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, Springer Verlag, Berlin.

Horrocks, J.C., McCann, A.P., Staniland, J.R. and de Dombal, F.T. (1972). Computer-aided diagnosis: description of an adaptable system, and operational experience with 2,034 cases, *British Medical Journal* **2**: 5–9.

Howard, R.A. (1966). Information value theory, *IEEE Transactions on Systems Science and Cybernetics* **2**(1): 22–26.

Howard, R.A. and Matheson, J.E. (1981). Influence diagrams, *in* R.A. Howard and J.E. Matheson (eds), *The Principles and Applications of Decision Analysis*, Vol. II, Strategic Decisions Group, Menlo Park, California, pp. 719–762.

Jensen, F., Jensen, F.V. and Dittmer, S.L. (1994). From influence diagrams to junction trees, *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–94)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 367–373.

Jensen, F.V., Lauritzen, S.L. and Olesen, K.G. (1990). Bayesian updating in causal probabilistic networks by local computations, *Computational Statistics Quarterly* **4**: 269–282.

Jenzarli, A. (1995). Information/relevance influence diagrams, *Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence (UAI–95)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 329–337.

John, L.R. (1995). An inverse transmission line model of the lower limb arterial system, *in* H. Power and R.T. Hart (eds), *Computer Simulations in Biomedicine*, Computational Mechanics Publications, Boston, Massachusetts.

Kaelbling, L., Littman, M.L. and Moore, A.W. (1996). Reinforcement learning: a survey, *Journal of Artificial Intelligence Research* **4**: 237–285.

Kahn, M.G. (1991). Modeling time in medical decision-support programs, *Medical Decision Making* **11**: 249–264.

Kahneman, D., Slovic, P. and Tversky, A. (eds) (1982). *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York.

Katz, A.M. (1977). *Physiology of the Heart*, Raven Press, New York.

Keeney, R.L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley & Sons, New York.

Keravnou, E., Garbay, C., Baud, R. and Wyatt, J. (eds) (1997). *AIME '97: Proceedings of the Sixth Conference on Artificial Intelligence in Medicine Europe*, Springer Verlag, Berlin.

Kidd, L., Driscoll, D. and Gersony, W. (1993). Second natural history study of congenital heart defects: results of treatment of patients with ventricular septal defects, *Circulation* **87**: 138–151.

Kiiveri, H., Speed, T.P. and Carlin, J.B. (1984). Recursive causal models, *Journal of the Australian Mathematical Society (Series A)* **36**: 30–52.

Kim, J. and Pearl, J. (1983). Bayesian updating in causal probabilistic networks by local computations, *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI–83)*, pp. 190–193.

Kjærulff, U. (1992). A computational scheme for reasoning in dynamic probabilistic networks, *Proceedings of the Eighth Annual Conference on Uncertainty in Artificial Intelligence (UAI–92)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 121–129.

Krovetz, L.J. (1998). Spontaneous closure of ventricular septal defect, *American Journal of Cardiology* **81**: 100–101.

Kushmerick, N., Hanks, S. and Weld, D. (1995). An algorithm for probabilistic planning, *Artificial Intelligence* **76**(1–2): 239–286.

Kyburg, H. (1991). Normative and descriptive ideals, *in* R. Cummins and J. Pollock (eds), *Philosophy and AI: essays at the interface*, MIT Press, Cambridge, Massachusetts, pp. 129–139.

Langlotz, C.P., Fagan, L.M., Tu, S.W., Sikic, B.I. and Shortliffe, E.H. (1987). A therapy planning architecture that combines decision theory and artificial intelligence techniques, *Computers and Biomedical Research* **20**: 279–303.

Lauritzen, S.L. and Spiegelhalter, D.J. (1988). Local computations with probabilities on graphical structures and their applications in expert systems, *Journal of the Royal Statistical Society (Series B)* **50**(2): 157–224.

Leong, T.-Y. (1998a). Modeling medical decisions in DynaMol: a new general framework of dynamic decision analysis, *Proceedings of the Ninth World Congress on Medical Informatics (MEDINFO-98)*.

Leong, T.-Y. (1998b). Multiple perspective dynamic decision making, *Artificial Intelligence* **105**: 209–261.

Lifschitz, V. (1987). Formal theories of action, *Proceedings of the Tenth International Joint Conference on Artificial Intelligence (IJCAI–87)*, pp. 966–972.

Littman, M.L. (1994). Memoryless policies: theoretical limitations and practical results, *in* D. Cliff, P. Husbands, J.-A. Meyer and S. Wilson (eds), *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*, MIT Press, Cambridge, Massachusetts.

Littman, M.L. (1996). *Algorithms for Sequential Decision Making*, Ph.D. thesis, Department of Computer Science, Brown University.

Littman, M.L., Cassandra, A.R. and Kaelbling, L.P. (1995). Efficient dynamic-programming updates in partially observable Markov decision processes, *Technical Report CS-95-19*, Department of Computer Science, Brown University.

Long, W.J. (1996). Temporal reasoning for diagnosis in a causal probabilistic knowledge base, *Artificial Intelligence in Medicine* **8**: 193–215.

Long, W.J., Fraser, H. and Naimi, S. (1997). Reasoning requirements for diagnosis of heart disease, *Artificial Intelligence in Medicine* **10**: 5–24.

Long, W.J., Naimi, S., Criscitiello, M.G. and Jayes, R. (1986). Using a physiological model for prediction of therapy effects in heart disease, *Proceedings of the Conference on Computers in Cardiology*, pp. 15–20.

Lovejoy, W.S. (1991). A survey of algorithmic methods for partially observed Markov decision processes, *Annals of Operations Research* **28**: 47–66.

Lucas, P.J.F. (1998). Analysis of notions of diagnosis, *Artificial Intelligence* **105**: 295–343.

Lucas, P.J.F. and Abu-Hanna, A. (1999). Prognostic methods in medicine (editorial), *Artificial Intelligence in Medicine* **15**: 105–119.

Luenberger, D.G. (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Menlo Park, California.

Luenberger, D.G. (1979). *Introduction to Dynamic Systems: Theory, Models and Applications*, John Wiley & Sons, New York.

Macartney, F.J., Spiegelhalter, D.J. and Rigby, M.L. (1987). Medical management, *in* Anderson et al. (1987), pp. 421–442.

Madani, O., Condon, A. and Hanks, S. (1999). On the undecidability of probabilistic planning and infinite-horizon partially observable Markov decision problems, *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI–99)*, pp. 541–548.

Magee, A.G., Boutin, C., McBrindle, B.W. and Smallhorn, J.F. (1998). Echocardiography and cardiac catheterization in the preoperative assessment of ventricular septal defect in infancy, *American Heart Journal* **135**: 907–913.

Magni, P. and Bellazzi, R. (1997). DT-planner: an environment for managing dynamic decision problems, *Computer Methods and Programs in Biomedicine* **54**: 183–200.

Magni, P. and Bellazzi, R. (1998). The optimal dynamic therapy: a decision-theoretic approach, *in* P. Borne, M. Ksouri and A. El Kamel (eds), *Proceedings of the IMACS Multiconference on Computational Engineering in Systems Applications (CESA–98)*.

McCarthy, J. (1980). Circumscription – a form of nonmonotonic reasoning, *Artificial Intelligence* **13**: 27–39.

McCarthy, J. and Hayes, P.J. (1969). Some philosophical problems from the standpoint of artificial intelligence, *Machine Intelligence* **4**: 463–502.

McDermott, D.V. (1982). A temporal logic for reasoning about processes and plans, *Cognitive Science* **6**: 101–155.

Middleton, B., Shwe, M.A., Heckerman, D.E., Henrion, M., Horvitz, E.J., Lehmann, H.P. and Cooper, G.F. (1991). Probabilistic diagnosis using a reformulation of the Internist-1/QMR knowledge base. II. Evaluation of diagnostic performance, *Methods of Information in Medicine* **30**: 256–267.

Miller, R.A., Pople, H.E. and Myers, J.D. (1982). INTERNIST-1, an experimental computer-based diagnostic consultant for general internal medicine, *New England Journal of Medicine* **307**: 468–476.

Minsky, M. (1975). A framework for representing knowledge, *Memo 306*, MIT–AI Laboratory, Cambridge, Massachusetts.

Moller, J.H., Patton, C., Varco, R.L. and Lillehei, C.W. (1991). Late results (30 to 35 years) after operative closure of isolated ventricular septal defect from 1954 to 1960, *American Journal of Cardiology* **68**: 1491–1497.

Monahan, G.E. (1982). A survey of partially observed Markov decision processes: theory, models, and algorithms, *Management Science* **28**(1): 1–16.

Moss, A.J., Adams, F.H., Forrest, H. and Emmanouilides, G.C. (eds) (1995). *Heart Disease in Infants, Children, and Adolescents*, Williams & Wilkins, Baltimore.

Ndilikilikesha, P.C. (1994). Potential influence diagrams, *International Journal of Approximate Reasoning* **10**: 251–285.

Olmsted, S.M. (1983). *On Representing and Solving Decision Problems*, Ph.D. thesis, Department of Engineering-Economic Systems, Stanford University.

Papadimitriou, C.H. and Tsitsiklis, J.N. (1987). The complexity of Markov decision processes, *Mathematics of Operations Research* **12**(3): 441–450.

Pauker, S.G. and Kassirer, J.P. (1987). Decision analysis, *New England Journal of Medicine* **316**: 1630–1641.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, Palo Alto, California.

Pearl, J., Geiger, D. and Verma, T. (1990). The logic of influence diagrams, *in* R.M. Oliver and J.Q. Smith (eds), *Influence Diagrams, Belief Nets, and Decision Analysis*, John Wiley & Sons, New York, pp. 67–87.

Peek, N.B. (1999a). Explicit temporal models for decision-theoretic planning of clinical management, *Artificial Intelligence in Medicine* **15**(2): 135–154.

Peek, N.B. (1999b). A specialised POMDP form and algorithm for clinical patient management, *in* A. Abu-Hanna and P. Lucas (eds), *Working Notes of the AIMDM'99 Workshop on Prognostic Models in Medicine: Artificial Intelligence and Decision Analytic Approaches*, pp. 39–43.

Peek, N.B., Coupé, V. and Ottenkamp, J. (1999). Focused quantification of a belief network using sensitivity analysis, *Proceedings of the Eleventh Belgium-Netherlands Conference on Artificial Intelligence*, pp. 123–130.

Peek, N.B. and Ottenkamp, J. (1997). Developing a decision-theoretic network for a congenital heart disease, *in* Keravnou et al. (1997), pp. 157–168.

Peot, M. and Smith, D. (1992). Conditional nonlinear planning, *Proceedings of the First International Conference on Artificial Intelligence Planning Systems*, pp. 189–197.

Platzman, L.K. (1977). *Finite-memory Estimation and Control of Finite Probabilistic Systems*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Poole, D. (1993). Probabilistic Horn abduction and Bayesian networks, *Artificial Intelligence* **64**: 81–129.

Poole, D. (1996). A framework for decision-theoretic planning I: combining the situation calculus, conditional plans, probability and utility, *Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence (UAI–96)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 436–445.

Provan, G.M. and Clarke, J.R. (1993). Dynamic network construction and updating techniques for the diagnosis of acute abdominal pain, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(3): 299–307.

Puterman, M.L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, New York.

Quaglini, S., Bellazzi, R., Stefanelli, M. and Locatelli, F. (1993). Sharing and reusing therapeutic knowledge for managing leukemic children, *in* S. Andreassen, R. Engelbrecht and J. Wyatt (eds), *AIME '93: Proceedings of the Fourth Conference*

on Artificial Intelligence in Medicine Europe, IOS Press, Amsterdam, pp. 319–330.

Quaglini, S., Berzuini, C., Bellazzi, R., Stefanelli, M. and Barosi, G. (1989). Therapy planning by combining AI and decision theoretic techniques, *in* D. Hunter (ed.), *Proceedings of the Second Conference on Artificial Intelligence in Medicine*, pp. 125–134.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choice under Uncertainty*, Addison-Wesley, Reading, Massachusetts.

Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory*, Addison-Wesley, Reading, Massachusetts.

Reed, N.E., Gini, M., Johnson, P.E. and Moller, J.H. (1997). Diagnosing congenital heart defects using the Fallot computational model, *Artificial Intelligence in Medicine* **10**: 25–40.

Reiter, R. (1987). A theory of diagnosis from first principles, *Artificial Intelligence* **32**: 57–95.

Renooij, S. and Witteman, C. (1999). Talking probabilities: communicating probabilistic information with words and numbers, *International Journal of Approximate Reasoning* **22**: 169–194.

Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B. and Richardson, W.S. (1996). Evidence-based medicine: What it is and what it isn't, *British Medical Journal* **312**: 71–72.

Savage, L.J. (1972). *The Foundations of Statistics*, Dover, New York.

Shachter, R.D. (1986). Evaluating influence diagrams, *Operations Research* **34**(6): 79–90.

Shachter, R.D. and Peot, M.A. (1992). Decision making using probabilistic inference methods, *Proceedings of the Eighth Annual Conference on Uncertainty in Artificial Intelligence (UAI–92)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 276–283.

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

Shahar, Y. (1999). Timing is everything: temporal reasoning and temporal data maintenance in medicine, *in* Horn et al. (1999), pp. 30–46.

Shannon, C.E. and Weaver, W. (1949). *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Illinois.

Shenoy, P.P. (1992). Valuation-based systems for Bayesian decision analysis, *Operations Research* **40**(3): 463–484.

Shiryayev, A.N. (1984). *Probability Theory*, Springer, Berlin.

Shortliffe, E.H. (1976). *Computer-based Medical Consultations: MYCIN*, Elsevier, New York.

Shortliffe, E.H. and Buchanan, B. (1975). A model of inexact reasoning in medicine, *Mathematical Biosciences* **23**: 351–379.

Shortliffe, E.H., Scott, A.C., Bischoff, M.B., Campbell, A.B., van Melle, W. and Jacobs, C.D. (1981). ONCOCIN: An expert system for oncology protocol management, *Proceedings of the Seventh International Joint Conference on Artificial Intelligence (IJCAI–81)*, pp. 876–881.

Simon, H.A. (1955). A behavioral model of rational choice, *Quarterly Journal of Economics* **69**: 99–118.

Smallwood, R.D. and Sondik, E.J. (1973). The optimal control of partially observable Markov processes over a finite horizon, *Operations Research* **21**: 1071–1088.

Smith, J.Q., Holtzman, S. and Matheson, J. (1993). Structuring conditional relationships in influence diagrams, *Operations Research* **41**(2): 280–297.

Sondik, E.J. (1971). *The Optimal Control of Partially Observable Markov Processes*, Ph.D. thesis, Department of of Electrical Engineering, Stanford University.

Sonnenberg, F.A. and Beck, J.R. (1993). Markov models in medical decision making: a practical guide, *Medical Decision Making* **13**: 322–338.

Soper, P., Ranaboldo, C. and Abeysinghe, G. (1991). A temporal model for clinical and resource management in vascular surgery, *in* D. Karagiannis (ed.), *Database and Expert Systems Applications*, Springer-Verlag, Berlin, pp. 549–552.

Soto, B., Becker, A.E., Moulaert, A.J., Lie, J.T. and Anderson, R.H. (1980). Classification of ventricular septal defect, *British Heart Journal* **43**: 332–343.

Sox, H.C., Blatt, M.A., Higgins, M.C. and Marton, K.I. (1988). *Medical Decision Making*, Butterworths, Boston, Massachusetts.

Spiegelhalter, D.J., Dawid, A.P., Lauritzen, S.L. and Cowell, R.G. (1993). Bayesian analysis in expert systems, *Statistical Science* **8**(3): 219–283.

Spiegelhalter, D.J., Franklin, R.C.G. and Bull, K. (1990). Assessment, criticism and improvement of imprecise subjective probabilities for a medical expert system, *in* Henrion et al. (1990), pp. 285–294.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, Massachusetts.

Tatman, J.A. and Shachter, R.D. (1990). Dynamic programming and influence diagrams, *IEEE Transactions on Systems, Man, and Cybernetics* **20**(2): 365–379.

Thompson, W.B., Johnson, P.E. and Moen, J.B. (1983). Recognition-based diagnostic reasoning, *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI–83)*, pp. 236–238.

Van der Gaag, L.C. and Wessels, M.L. (1993). Selective evidence gathering for diagnostic belief networks, *AISB Quarterly* (86): 23–34.

Von Neumann, J. and Morgenstern, O. (1944). *The Theory of Games and Economic Behavior*, John Wiley & Sons, New York.

Warner, H.R., Toronto, A.F., Veasy, L.G. and Stephenson, R. (1961). A mathematical approach to medical diagnosis: application to congenital heart disease, *Journal of the American Medical Association* **177**: 177–183.

Weinstein, M.C. and Fineberg, H.V. (1980). *Clinical Decision Analysis*, Saunders, Philadelphia, Pennsylvania.

Weiss, S.M., Kulikowski, C.A., Amarel, S. and Safir, A. (1978). A model-based method for computer-aided medical decision making, *Artificial Intelligence* **11**: 145–172.

White, C.C. and Scherer, W.T. (1994). Finite-memory suboptimal design for partially observed Markov decision processes, *Operations Research* **42**(3): 439–455.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*, John Wiley & Sons, New York.

Wyatt, J.C. and Altman, D.G. (1995). Commentary: Prognostic models: clinically useful of quickly forgotten?, *British Medical Journal* **311**: 1539–1541.

Zadeh, L.A. (1965). Fuzzy sets, *Information and Control* **8**: 338–353.

Zhang, N.L. (1998). Probabilistic inference in influence diagrams, *Computational Intelligence* **14**(4): 476–497.

Zhang, N.L. and Lee, S.S. (1998). Planning with partially observable Markov decision processes: advances in exact solution method, *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–98)*, Morgan Kaufmann Publishers, Palo Alto, California, pp. 523–530.

Samenvatting

In de geneeskundige zorg voor individuele patiënten moet een arts voortdurend beslissingen nemen. Op basis van klachten en symptomen moet hij een hypothese vormen over de mogelijke kwaal van de patiënt, hij moet vaststellen of nader diagnostisch onderzoek nodig is om deze hypothese te bevestigen, hij moet besluiten of therapeutisch handelen vereist is, en beslissen waaruit het zorgtraject na een eventuele ingreep moet bestaan. Al deze beslissingen zijn aan elkaar gerelateerd, en het geheel kan daarom opgevat worden als een vorm van *plannen*. In de artificiële intelligentie (AI) wordt plannen van oudsher bestudeerd voor situaties die een grote mate van voorspelbaarheid kennen. Een belangrijke karakteristiek van medische beslissingen is evenwel de rol van onzekerheid: de kwaal van de patiënt is niet altijd met zekerheid vast te stellen, en ook is de patiënts toekomstige ontwikkeling vaak niet goed te voorspellen. Het nemen van beslissingen onder onzekerheid is het terrein van de besliskunde; in dit proefschrift onderzoeken we het medisch beslisproces als een vorm van *plannen* met besliskundige principes, oftewel *besliskundig plannen*.

Het is gebruikelijk om een onderscheid te maken tussen drie basisproblemen in de kliniek: *diagnose* (wat is er mis?), *therapieselectie* (wat kan er aan gedaan worden?), en *prognose* (wat zal er gebeuren?). In de literatuur worden deze drie problemen vaak apart bestudeerd, en ook beslissingsondersteunende systemen concentreren zich doorgaans op een van deze problemen. De achterliggende gedachte is dat het zorgproces een relatief statische procedure is waarin deze drie problemen na elkaar worden opgelost. Deze problemen kennen echter een zodanige onderlinge samenhang dat zij niet afzonderlijk kunnen worden opgelost; zij spelen allen voortdurend een rol in het zorgproces. In ons onderzoek gaan we daarom uit van een *dynamisch perspectief* op medische zorg, waarbij op gezette tijden interactie plaatsvindt tussen arts en patiënt, en de verandering in de klinische toestand van de patiënt over de tijd centraal staat.

Deze interactiemomenten strekken zich, afhankelijk van aard van de kwaal, uit over een kortere of langere tijdspanne, en kennen in principe allen diagnostische, therapeutische, en prognostische aspecten. Op alle interactiemomenten moeten door de arts een of meer beslissingen genomen worden, en vaak is ook de *timing* van interactiemomenten van belang.

De theorie in het proefschrift wordt geïllustreerd met voorbeelden uit het domein van aangeboren hartafwijkingen. In het bijzonder wordt ingegaan op de behandeling van patiënten met een *ventrikelseptumdefekt* (VSD), de meest voorkomende hartafwijking bij pasgeborenen. Hoofdstuk 2 geeft een inleiding tot dit domein, en bespreekt de klinische en pathofysiologische karakteristieken van het VSD. De aandoening is niet acuut en kent bovendien in veel gevallen een voorspoedig natuurlijk verloop; het is echter ook mogelijk dat ernstige pulmonale complicaties optreden. De behandeling van een VSD patiënt is een proces dat zich doorgaans uitstrekt over meerdere jaren. De centrale beslissing is of, en zo ja, wanneer, een chirurgische ingreep nodig is. Daarnaast moet worden vastgesteld met welke frequentie de patiënt gezien wordt door de kindercardioloog, en in hoeverre invasief diagnostisch onderzoek noodzakelijk is.

In Hoofdstuk 3 worden de formele beginselen van de besliskunde besproken. De besliskunde is gebaseerd op de kansrekening voor het redeneren met onzekerheid, en op de utiliteitstheorie voor het maken van rationele keuzes bij beslissingen onder onzekerheid. In het hoofdstuk wordt onder andere getoond hoe deze eenvoudige uitgangspunten de mogelijkheid bieden om allerlei verschillende soorten beslisproblemen te analyseren.

De besliskunde ligt ook ten grondslag aan diverse *representatieformalismen* voor beslissingsondersteunende systemen. In Hoofdstuk 4 beschrijven en analyseren we drie van dergelijke representatieformalismen: *influence diagrams*, *Markov beslisprocessen*, en *dynamische influence diagrams*. Het *influence diagram* (ID) is een grafische representatie die nauw verwant is aan het *belief-network* formalisme voor probabilistisch redeneren. *Influence diagrams* worden een gekenmerkt door een grote mate van compactheid, maar ondersteunen alleen beslisproblemen met een statisch karakter; zij vormen daarom een zwakke implementatie van het concept besliskundig plannen. In Markov beslisprocessen spelen tijd en toestandsdynamiek wel een grote rol: hierin wordt het nemen van beslissingen geformaliseerd als het besturen van een stochastisch Markovproces. Met name de partieel-observeerbare variant (*partially-observable Markov decision process*, POMDP), waarin keuzes genomen moeten worden op basis van beperkte informatie, sluit goed aan bij de klinische situatie. Een nadeel van POMDPs is evenwel dat zij een aantal sterke assumpties met betrekking tot het toepassingsdomein bevatten, en dat de representatie zeer snel groeit in de grootte van het probleem. Het *dynamische influence diagram* is een uitbreiding van het *influence diagram* met een expliciete notie van tijd en toestandsverandering. Het is mogelijk om dit formalisme te combineren met POMDPs; deze combinatie van formalismen vormt op dit moment een van de krachtigste besliskundige representaties.

De belangrijkste innovatieve bijdragen van het proefschrift worden geleverd in Hoofdstukken 5 en 6. In Hoofdstuk 5 verlaten we concrete representatievormen, en beschouwen het begrip 'besliskundig plannen' op een fundamenteler niveau. We ontwikkelen een uitgebreid theoretisch raamwerk voor deze vorm van plannen met een minimum aan restrictieve assumpties; dit stelt ons onder andere in staat om de eerder besproken representatieformalismen beter te analyseren. Het uitgangspunt van het raamwerk is, net als bij POMDPs, het perspectief van een *planning agent* die probeert een dynamisch systeem te 'besturen' op basis van partiële informatie over de toestand van dat systeem. De kern van het raamwerk wordt gevormd door een Boolse algebra waarin de noties van systeemtoestand, gebeurtenis, observatie, beslissing, en plan worden geformaliseerd. Hieruit wordt vervolgens het begrip *beslisproces* geconstrueerd; een beslisproces beschrijft alle probabilistische en informationele relaties tussen de planning agent en het dynamische systeem over de tijd.

Het hoofdstuk bevat een uitgebreide theorie van *voorwaardelijk plannen*, waarbij een plan wordt uitgedrukt als een verzameling (conditionele) beslisregels. Dergelijke regels zijn eenvoudig te communiceren met specialisten op het toepassingsgebied, en kunnen bijvoorbeeld worden gebruikt in klinische protocollen. Omgekeerd biedt de theorie de mogelijkheid om een verzameling gegeven beslisregels (bijvoorbeeld geformuleerd door een specialist of afkomstig uit een richtlijn of protocol), te controleren op interne consistentie en samenhang, volledigheid, en *performance* met betrekking tot een zekere klinische doelstelling. Daarnaast is het mogelijk om beslisregels te formuleren op verschillende detailniveaus en met een variërend toepassingsbereik.

Hoofdstuk 6 bestaat uit twee delen. In het eerste deel bespreken we hoe een gegeven klinisch toepassingsdomein geformaliseerd kan worden in het raamwerk, oftewel, hoe we de verschillende aspecten van zo'n domein kunnen modelleren. Met behulp van uitgebreide illustraties uit het VSD-domein, gaan we in op het modelleren van een patiënts pathofysiologische conditie, van de klinische modaliteiten voor de behandelend arts, en van bevindingen zoals resultaten van diagnostiek. Daarnaast wordt beschreven hoe we in het raamwerk het verwachte verloop van een patiënts conditie over de tijd, en de invloed daarop van de diverse modaliteiten, kunnen formaliseren. Ook bespreken we de specificatie van klinische doelstellingen; we vergelijken op *doeltoestand* gebaseerde specificaties (zoals vaak wordt gebruikt in AI-benaderingen van plannen) met op *levensverwachting* gebaseerde specificaties (zoals vaak wordt gebruikt in de klinische besliskunde).

Ten slotte keren we in het tweede deel van Hoofdstuk 6 terug naar het uitgangspunt van ons onderzoek, het dynamisch perspectief op klinische zorg. Vanuit het raamwerk voor besliskundig plannen bestuderen we de diagnostische, therapeutische, en prognostische aspecten van afzonderlijke beslissingen, en onderzoeken we ook de relaties tussen deze aspecten. Het raamwerk biedt steeds een zekere mate vrijheid bij het formaliseren van deze onderdelen van het klinisch redeneren; we laten zien hoe een strikt besliskundige benadering verruild kan worden voor een meer heuristische benadering, en wat het effect van zo'n keuze is.

Dankwoord

Graag wil ik op deze plaats de mensen bedanken die, direct of indirect, een bijdrage hebben geleverd aan de totstandkoming van dit proefschrift.

Gedurende het onderzoek heb ik kunnen profiteren van de ruime kennis en ervaring van mijn begeleider Peter Lucas op het vakgebied. Veel van de hier gepresenteerde ideeën heb ik te danken aan onze discussies over het onderwerp, en daarvoor ben ik hem zeer erkentelijk. Tevens ben ik de afgelopen jaren door Peter gevormd als wetenschappelijk onderzoeker. Ik heb van hem geleerd om in het onderzoek met rigueur te werk te gaan, maar ook om tegendraadse opvattingen niet schuwen, en een relativerende kijk op het academisch bedrijf te behouden. Daarmee zal ik het zeker nog even vol kunnen houden in de wetenschap.

Ook wil ik de overige leden van het Tetrade-project bedanken, in het bijzonder Jaap Ottenkamp en Veerle Coupé. Jaap Ottenkamp heeft mij met veel geduld en toewijding zijn kennis bijgebracht van de kindercardiologie en de behandeling van VSD patiënten; dit heeft mij veel inzicht in de medische besluitvorming gegeven. Met Veerle Coupé heb ik gewerkt aan een onderzoek dat uiteindelijk buiten het bestek van dit proefschrift is gevallen, maar onze plezierige samenwerking was voor mij een positieve stimulans; wij zijn bovendien in de loop der tijd goed bevriend geraakt.

In de laatste fase van het onderzoek was een belangrijke rol weggelegd voor mijn promotor John-Jules Meyer, die nauwgezet het gehele manuscript heeft becommentarieerd, en mij met zijn enthousiasme vertrouwen heeft gegeven in het eindresultaat.

Mijn speciale dank gaat uit naar Silja Renooij, die, op geheel vrijwillige basis, grote delen van het manuscript heeft gelezen en minutieus van commentaar heeft voorzien. Ook Linda van der Gaag, Veerle Coupé, Marc du Chatinier, Eveline Helsper, Marco

Wiering, en Wouter Kusters wil ik graag bedanken voor hun commentaar op delen van het manuscript.

Mijn jaren op de Uithof zouden niet dezelfde zijn geweest zonder mijn onvervangbare kamergenoot Tanja Vos; in het grijs van beton en computers heeft zij veel kleur aangebracht. Er werd wel eens gevloekt, en ook wel eens gescholden in rookhol A117, maar we hebben vooral een hoop gedeeld en veel gelachen.

De Afdeling Klinische Informatiekunde van het AMC wil ik graag bedanken voor de tijd en de gelegenheid die zij mij hebben gegeven om het proefschrift af te ronden.

Zoals iedere promovendus heb ik het vervullen van mijn opdracht niet altijd als een feest ervaren; er zijn momenten geweest dat ik graag “de hele rommel in de sloot had gegooid”. Dat zou ook zeker zijn gebeurd als ik niet de onvoorwaardelijke steun had gekregen van de mensen die mij zeer na staan. Daarom wil ik hier ten slotte mijn moeder, en nogmaals Wouter, Marc, en mijn allerliefste bedanken, omdat zij er steeds voor mij zijn geweest.

Curriculum Vitae

Niels Bastiaan Peek werd geboren op 17 april 1970 te Eindhoven. Hij groeide op in het nabijgelegen Waalre, en bezocht het Hertog Jan College te Valkenswaard. Na het behalen van zijn V.W.O. Atheneum B diploma op 17 juni 1988 toog hij naar Utrecht om daar Informatica en Cognitieve Kunstmatige Intelligentie (CKI) te gaan studeren. Tijdens de doctoraalfase koos Niels voor een specialisatie in de richting van *kennissystemen* en *probabilistisch redeneren*. Op 24 juni 1994 studeerde hij af op een scriptie getiteld '*A filter for belief networks*'. Het doctoraaldiploma Informatica werd *cum laude* behaald; de scriptie werd bovendien bekroond met de Scriptieprijs CKI '93/'94.

Na zijn afstuderen trad Niels in dienst van de Rijksuniversiteit Leiden en verrichtte onderzoek naar juridische kennissystemen. In november 1995 keerde hij terug naar Utrecht om daar bij het Informatica Instituut te gaan werken als onderzoeker in opleiding. Het was werkzaam in het Tetrade-project, een door NWO gefinancierd samenwerkingsverband tussen onderzoeksgroepen uit Utrecht, Rotterdam, en Amsterdam. In het project werd onderzocht hoe medische beslissingsondersteuning geboden kan worden met een synthese van technieken uit de artificiële intelligentie en de klinische besliskunde; Niels' bijdrage mondde uit in deze dissertatie.

Sinds 1 april 2000 werkt Niels Peek als postdoc onderzoeker bij de Afdeling Klinische Informatiekunde van het Academisch Medisch Centrum te Amsterdam.

In zijn vrije tijd bezoekt Niels graag de concertzaal en het bioscooptheater. Zijn vakanties brengt hij vaak wandelend door in bergachtige streken. Ook leest hij wel eens een boek.

Titles in the SIKS Dissertation Series

98–1 Johan van den Akker, *DEGAS - An Active, Temporal Database of Autonomous Objects.*

98–2 Floris Wiesman, *Information Retrieval by Graphically Browsing Meta-Information.*

98–3 Ans Steuten, *A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective.*

98–4 Dennis Breuker, *Memory versus Search in Games.*

98–5 Eduard Oskamp, *Computerondersteuning bij Straftoemeting.*

99–1 Mark Sloof, *Physiology of Quality Change Modelling. Automated Modelling of Quality Change of Agricultural Products.*

99–2 Rob Potharst, *Classification using Decision Trees and Neural Nets.*

99–3 Don Beal, *The Nature of Minimax Search.*

99–4 Jacques Penders, *The practical Art of Moving Physical Objects.*

99–5 Aldo de Moor, *Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems.*

99–6 Niek Wijngaards, *Re-design of Compositional Systems.*

99–7 David Spelt, *Verification Support for Object Database Design.*

99–8 Jacques Lenting, *Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation.*

2000–1 Frank Niessink, *Perspectives on Improving Software Maintenance.*

2000–2 Koen Holtman, *Prototyping of CMS Storage Management.*

2000–3 Carolien Metselaar, *Sociaal-organisatorische Gevolgen van Kennistechnologie. Een Procesbenadering en Actorperspectief.*

2000–4 Geert de Haan, *ETAG, A Formal Model of Competence Knowledge for User Interface Design.*

2000–5 Ruud van der Pol, *Knowledge-based Query Formulation in Information Retrieval.*

2000–6 Rogier van Eijk, *Programming Languages for Agent Communication.*

2000–7 Niels Peek, *Decision-Theoretic Planning of Clinical Patient Management.*

2000–8 Veerle Coupé, *Sensitivity Analyis of Decision-Theoretic Networks.*