Chapter 1
Introduction

L. Enrique Sucar
National Institute for Astrophysics, Optics and Electronics, Mexico

Eduardo Morales
National Institute for Astrophysics, Optics and Electronics, Mexico

Jesse Hoey
University of Waterloo, Canada

ABSTRACT

This chapter gives a general introduction to decision-theoretic models in artificial intelligence and an overview of the rest of the book. It starts by motivating the use of decision-theoretic models in artificial intelligence and discussing the challenges that arise as these techniques are applied to develop intelligent systems for complex domains. Then it introduces decision theory, including its axiomatic bases and the principle of maximum expected utility; a brief introduction to decision trees is also presented. Finally, an overview of the three main parts of the book –fundamentals, concepts, and solutions– is presented.

ARTIFICIAL INTELLIGENCE AND DECISION THEORY

For achieving their goals, intelligent agents, natural or artificial, have to select a course of actions among many possibilities. That is, they have to take decisions based on the information they can obtain from their environment, their previous knowledge and their objectives. In many cases, the information and knowledge is incomplete or unreliable, and the results of their decisions are not certain, that is they have to take decisions under uncertainty. For instance: a medical doctor in an emergency, must act promptly even if she has limited information on the patient's state; an autonomous vehicle that detects what might be an obstacle in its way, must decide if it should turn or stop without being certain about the obstacle's distance, size and velocity; or a financial agent needs to select the best investment according to its vague predictions on expected return of the different alternatives and its clients' requirements. In all these cases, the agent should try to make the *best* decision based on limited information and resources (time, computational power, etc.). How can we determine which is the best decision?

Decision Theory provides a normative framework for decision making under uncertainty. It is

DOI: 10.4018/978-1-60960-165-2.ch001

based on the concept of *rationality*, that is that an agent should try to maximize its utility or minimize its costs. This assumes that there is some way to assign utilities (usually a number, that can correspond to monetary value or any other scale) to the result of each alternative action, such that the best decision is the one that has the highest utility. For example, if we wanted to select in which stock to invest $1000 for a year, and we knew which will be the price of each stock after a year, then we should invest in the stock that will provide the highest return. Of course we can not predict with precision the price of stocks a year in advance, and in general we are not sure of the results of each of the possible decisions, so we need to take this into account when we calculate the value of each alternative. So in decision theory we consider the *expected utility*, which makes an average of all the possible results of a decision, weighted by their probability. Thus, in a nutshell, a rational agent must select the decision that maximizes its expected utility.

Decision theory was initially developed in economics and operations research (Neumann & Morgenstern, 1944), but in recent years has attracted the attention of artificial intelligent (AI) researchers interested in understanding and building intelligent agents. These intelligent agents, such as robots, financial advisers, intelligent tutors, etc., must deal with similar problems as those encountered in economics and operations research, but with two main differences.

One difference has to do with the size of the problems, which in artificial intelligence tend to be very *large*; with many possible *states* of the environment and in some cases also a large number of actions or decisions for the agent. Although in economics they also have to deal with big problems, they usually abstract them as they do not have the tools to deal with a large number of states. For example, consider a mobile robot that is moving in a large building and wants to decide the best set of movements that will take it from one place in the building to another. In this case the world state could be represented as the position of the robot in the building, and the agent's actions are the set of possible motions (direction and velocity) of the robot. So the problem can be formulated as the selection of the best motion for each position in the building to reach the goal position (minimizing distance or time, for instance). In this case the number of states and actions are in principle infinite, or very large if we discretize them. The size of the problems in terms of states and actions imply a problem of computational complexity, in terms of space and time, so AI has to deal with these issues in order to apply decision theory to complex scenarios.

The other main difference has to do with knowledge about the problem domain, that is having a *model* of the problem according to what is required to apply decision theory techniques to solve it. This means, in general, knowledge of all the possible domain states and possible actions, and the probability of each outcome of a decision and its corresponding utility. In many AI applications a model is not known in advance, and could be difficult to obtain. Returning to the robot navigation example, the robot might not have a precise model of the environment and might also lack a detailed model of its dynamics to exactly predict which will be its position after each movement. So AI researchers have to deal also with the problem of knowledge acquisition or learning.

The research on decision-theoretic models in artificial intelligence has focused on these two main issues: computational complexity and model acquisition; as well as in incorporating these theoretical advances in different applications of intelligent agents. In the rest of the book we will explore the theoretical and practical developments of this interaction between artificial intelligence and decision theory; but first we will review the basis of decision theory.

Decision Theory: Fundamentals

The principles of decision theory were initially developed in the classic text by Von Neuman and Morgensten, *Theory of Games and Economic Behavior* (Neumann & Morgenstern, 1944). They established a set of intuitive constraints that should guide the preferences of a rational agent, which are known as the axioms of utility theory. Before we list these axioms, we need to establish some notation.

In a decision scenario there are four elements:

Alternatives: are the choices that the agent has and are under his control. Each decision has at least two alternatives (e.g. to do or not do some action).

Events: these are produced by the environment or by other agents; are outside of the agent's control. Each random event has at least two possible results, and although we do not know in advance which result will occur, we can assign a probability to each one.

Outcomes: are the results of the combination of the agents decisions and the random events. Each possible outcome has a different preference (utility) for the agent.

Preferences: these are established according to the agent's goals and objectives and are assigned by the agent to each possible outcome. They establish a value for the agent for each possible result of its decisions.

As an example consider a virtual robot that navigates in a grid environment, see Figure 1. The virtual robot is in a certain cell, and it can move to each of the 4 adjacent cells –up, down, left, right; its objective is to arrive to the goal cell. In this case the alternatives are the four possible actions that the robot can take in each cell. The events could be a failure in the robot or another agent that gets in its way; both can prevent the robot from reaching the desired location. According to how likely these events are, there will be a certain probability that the robot reaches the desired cell or another cell. The outcomes are the result of the combination of the robot actions and the events. For example, if the robot moves to the cell to its right, the possible outcomes could be the that it arrives to the right cell, it stays in the same cell or it arrives to another adjacent cell. The preferences will be established according to the robot objectives. In this case, the robot will prefer the goal cell over all the other empty cells in the environment, and these over the cells that have obstacles.

=9.0cm

In utility theory, these type of scenarios are called *lotteries*. In a lottery each possible outcome or *state*,, has certain probability and an associated preference to the agent which is quantified by a real number, . For instance, a lottery with two possible outcomes, with probability and with probability, will be denoted as:

If an agent prefers than it is written as, and if it is indifferent between both outcomes is denoted as . In general a lottery can have any number of outcomes; an outcome can be an atomic state or another lottery.

Based on these concepts, we can define utility theory in an analogous way as probability theory, by establishing a set of reasonable constraints on preferences for a rational agent, these are the axioms of utility theory:

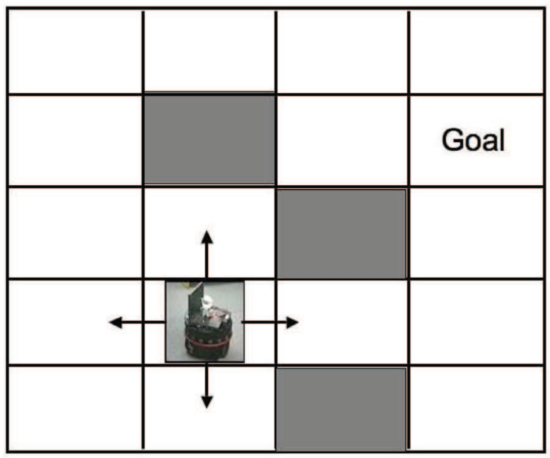
Order: Given two states, an agent prefers one or the other or it is indifferent between them.

Transitivity: If an agent prefers outcome to and prefers to, then it must prefer to .

Continuity: If, then there is some probability such that the agent is indifferent between getting with probability one, or the lottery .

Substitutability: If an agent is indifferent between two lotteries, and, then the agent is indifferent between two more complex lotteries that are the same except that is substituted for in one of them.

Figure 1. An example of a decision problem: A robot in a grid environment. The robot can move to the neighboring cells (as indicated by the arrows) an must select the best moves to avoid the obstacles (in gray) and reach the GOAL.



Monotonicity: There are two lotteries that have the same outcomes, and . If the agent prefers, then it must prefer the lottery in which has higher probability.

Decomposability: Compound lotteries can be decomposed into simple ones using the rules of probability.

Then, the definition of a utility function follows from the axioms of utility.

Utility Principle: If an agent's preferences follow the axioms of utility, then there is a real-valued utility function such that:

1. if an only if the agent prefers over,
2. if an only if the agent is indifferent between and .

Maximum Expected Utility Principle: The utility of a lottery is the sum of the utilities of each outcome times its probability:

Based on this concept of a utility function, we can now define the expected utility (EU) of certain decision taken by an agent, considering that there are possible results of this decision, each with probability:

The principle of Maximum Expected Utility states that a rational agent should choose an action that maximizes its expected utility.

Although it seems straight-forward to apply this principle to determine the best decision, as the decision problems become more complex, involving several decisions, events and possible outcomes, it is not as easy as it seems; and a systematic approach is required to model and solve complex decision problems. One of the earliest modeling tools developed for solving decision problems are decision trees (Cole & Rowley, 1995).

Decision Trees

A decision tree is a graphical representation of a decision problem, which has three types of elements or nodes that represent the three basic components of a decision problem: decisions, uncertain events and results.

A *decision node* is depicted as a rectangle which has several *branches*, each branch represents each of the possible alternatives present in this decision point. At the end of each branch there could be another decision point, an event or a result.

An *event node* is depicted as a circle, and has also several branches, each branch represents one of the possible outcomes of this uncertain event. These outcomes correspond to all the possible results of this event, that is they should be mutually exclusive and exhaustive. A probability value is assigned to each branch, such that the sum of the probabilities for all the branches is equal to one. At the extreme of each branch there could another event node, a decision node or a result.

The *results* are annotated with the utility they express for the agent, and are usually at the end of each branch of the tree (the leaves).

Decision trees are usually drawn from left to right, with the root of the tree (a decision node) at the extreme left, and the leaves of the tree to the right. An example of a hypothetical decision problem (based on an example in (Borrás, 2001)) is shown in Figure 2. It represents an investment decision with 3 alternatives: (i) Stocks, (ii) Gold, and (iii) No invest. Assuming that the investment is for one year, if we invest in stock, depending on how the stock market behaves (uncertain event), we could gain \$1000 or loose \$300, both with equal probability. If we invest in Gold, we have another decision, to have insurance or not. If we get insurance, then we are sure to gain \$200; otherwise we win or lose depending if the price of the gold is up, stable or down; this is represented as another event. Each possible outcome has a certain value and probability assigned, as shown in Figure 2. What should the investor decide?

=12.0cm

To determine the best decision for each decision point, according to the maximum expected utility principle, we need to *evaluate* the decision tree. The evaluation of a decision tree consists of determining the values of both types of nodes, decision and event nodes. It is done from right to left, starting from any node that has only results for all its branches:

- The value of a decision node is the maximum value of all the branches that emanate from it:
- The value of an event node is the expected value of all the branches that emanate from it, obtained as the weighted sum of the result values multiplied by their probabilities:

Following this procedure we can evaluate the decision tree of Figure 2:

Event 1 - Market Price: .
Event 2 - Gold Price: .
Decision 2 - Insurance: .
Decision 1 - Investment: .

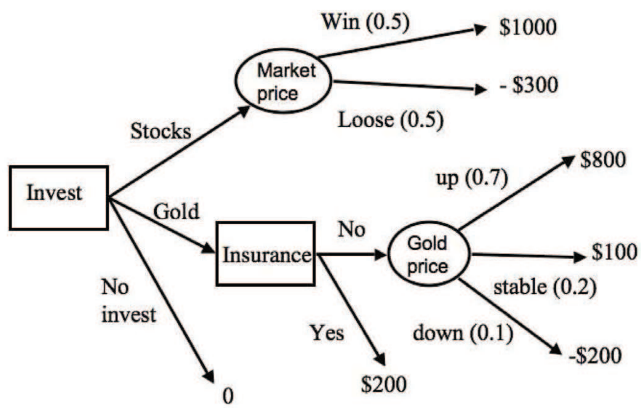
Thus, in this case the best decisions are to invest in Gold without insurance.

Decision trees are a tool for modeling and solving sequential decision problems, as decisions have to be represented in sequence as in the previous example. However, the size of the tree (number of branches) grows exponentially with the number of decision and event nodes, so this representation is practical only for *small* problems. An alternative modeling tool is the *Influence Diagram* (Howard & Matheson, 1984, Shachter, 1986), which provide a compact representation of a decision problem. Influence diagrams are introduced in Chapter 2.

Decision trees and influence diagrams are techniques for solving *simple* decision problems, where a few decisions are involved (usually less than 10). For more complex problems, in particular dynamic decision problems that involve a series of decisions in time (as the robot navigation example), there are other techniques, such as *Markov Decision Processes* (MDPs) and *Partially Observable Markov Decision Processes* (POMDPs). These are reviewed in Chapter 3.

In many domains the transition model is not known in advance and an agent must learn a policy

Figure 2. An example of a decision tree (see text for details)



from experience. For this case, several approaches have been proposed by the reinforcement learning community. These are reviewed in Chapter 4.

OVERVIEW

The book is organized in 3 parts: fundamentals, concepts, and solutions.

Fundamentals

The first part, fundamentals, provides a general introduction to the main decision-theoretic techniques used in artificial intelligence.

Probabilistic graphical models provide a framework for a compact representation of a joint probability distribution, and are incorporated in several advanced decision models, such as influence diagrams and factored Markov decision processes. Chapter 2 gives an overview of probabilistic graphical models and introduces influence diagrams. First, a brief review of probability theory is included. A general introduction to graphical models is given, and a more detailed description of certain types of graphical models is presented, in particular Bayesian networks and dynamic Bayesian networks. Then influence diagrams are introduced, including some of the more common solution techniques. Finally, a brief introduction to dynamic decision networks is presented, and their relation with Markov decision processes.

Chapter 3 provides an introduction to fully and partially observable Markov decision processes as a framework for sequential decision making under uncertainty. It reviews Markov decision processes (MDPs) and partially observable Markov decision processes (POMDPs), describes the basic algorithms to find good policies and discusses modeling and computational issues that arise in practice.

In many applications a decision model, such as an influence diagram or MDP, is not available, so an alternative is to learn directly a decision *policy* from experience using *Reinforcement Learning*. Chapter 4 provides a concise and updated introduction to reinforcement learning from a machine learning perspective. It gives the required background to understand the chapters related to reinforcement learning in the book, and includes an overview of some of the latest trends in the area.

Concepts

Part II presents recent theoretical developments that extend some of the techniques in Part I, such as influence diagrams, Markov decision processes and reinforcement learning, to deal with computational and representational issues that arise in artificial intelligence. In some applications of artificial intelligence, it is not only important to make good decisions, but also to explain how the system arrived at these decisions; a novel method for automatically generating explanations for MDPs is also described.

Semi-Markov Decision Processes (SMDPs) are an extension of MDPs that generalize the notion of time by allowing the time intervals between states to vary stochastically. SMDPs are used to formulate many control problems and play a key role in hierarchical reinforcement learning. Chapter 5 shows how to translate a decision making problem into a form that can instead be solved by inference and learning techniques. It establishes a formal connection between planning in semi-MDPs and inference in probabilistic graphical models.

Multi-stage Stochastic Programming relies on mathematical programming and probability theory to solve complex decision problems. Chapter 6 presents the multi-stage stochastic programming framework for sequential decision making under uncertainty. It describes the standard techniques for solving approximately stochastic dynamic problems, which are based on a discretization of the disturbance space called a *scenario tree*. It also shows how supervised learning techniques can be used to evaluate the quality of an approximation.

Chapter 7 describes a technique to explain policies for factored Markov decision processes by populating a set of domain-independent templates, and a mechanism to determine a minimal set of templates that, viewed together, completely justified the policy. This technique is illustrated in two domains: advising students in their course selection and assisting people with dementia in completing the task of hand washing.

Dynamic limited-memory influence diagrams (DLIMIDs) are novel type of decision support models whose main difference with other models is the restriction of limited memory, which means that the decisions are based only on recent observations. Chapter 8 introduces DLIMIDs and presents several algorithms for evaluating them. The application of DLIMDs is illustrated in a real-world medical problem, including a comparison with other formalisms.

Chapter 9 introduces an approach for reinforcement learning based on a relational representation. The basic idea is to represent states in the domain as sets of first order relations; actions and policies are also represented over those generalized representations. The approach is demonstrated in two applications: a flight simulator and a mobile robot.

Solutions

Part III describes a wide sample of applications of decision-theoretic models in different areas of artificial intelligence, including: intelligent tutors and intelligent assistants, power plant control, medical assistive technologies, spoken dialog systems, service robots, and multi-robot systems. As well as illustrating the practical aspects of using decision theory in AI, it also presents several developments that extend the techniques described in Part I and Part II. These include: extending the expressive power of the models based on relational representations; making the solution techniques more efficient using abstraction and decomposition; and developing decentralized models for distributed applications, among others.

Chapter 10 describes a decision-theoretic tutor that helps students learn from analogical problem solving. The tutor incorporates a novel example–selection mechanism that tailors the choice of examples for a given student based on dynamic Bayesian networks. An empirical evaluation shows that this selection mechanism is more effective that standard approaches for fostering learning.

Chapter 11 presents an intelligent tutor based on dynamic decision networks applied to an undergraduate Physics scenario, where the aim is to adapt the learning experience to suit the learner's needs. It employs *Probabilistic Relational Models* to facilitate the construction of the models, such that the tutor can be easily adapted to different experiments, domains and student levels.

An intelligent assistant for power plant operators is introduced in Chapter 12. *AsistO* provides off-line training and on-line guidance to power plant operators in the form of ordered recommendations. These recommendations are generated using Markov decision processes over an approximate factored model of a power plant. It also describes an automatic explanation mechanism that explains the recommendations to the user.

Chapter 13 presents a general decision–theoretic model of interactions between users and cognitive assistive technologies for various tasks of importance to the elderly population. The model is a POMDP whose goal is to work in conjunction with the user towards the completion of a given task. It is applied to four tasks: assisted hand washing, stroke rehabilitation, health monitoring, and wheelchair mobility.

Spoken dialog systems present a classic example of planning under uncertainty. Thus, decision theory, and in particular partially–observable Markov decision processes, are an attractive approach to building spoken dialog systems. Chapter 14 traces the history of POMDP–based spoken dialog systems, and sketches avenues for future work.

Chapter 15 describes a novel framework based on functional decomposition of MDPs, and its application to task coordination in service robots. In this framework, a complex task is divided into several subtasks, each one defined as an MDP and solved independently; their policies are combined to obtain a global policy. Conflicts may arise as the individual policies are combined, so a technique for detecting and solving different types of conflicts is presented.

Chapter 16 introduces problems related to the decentralized control of multi–robot systems. It presents *Decentralized Markov Decision Processes* (DEC-MDPs) and discusses their applicability to real world multi-robot applications. Then, it introduces OC-DEC-MDPs and 2V-DEC-MDPs which have been developed to increase the applicability of DEC-MDPs.

FINAL REMARKS

Sequential decision making is at the core of the development of many intelligent systems but it has been only recently that a large number of developments have been proposed and real world applications have been tackled. This book provides a general and comprehensive overview of decision theoretic models in artificial intelligence, with an overview of the basic solution techniques, a sample of more advanced approaches, and examples of some recent real world applications.

REFERENCES

Borrás, R. (2001). *Análisis de incertidumbre y riesgo para la toma de decisiones*. Orizaba, Mexico: Comunidad Morelos. (in Spanish)

Cole, S., & Rowley, J. (1995). Revisiting decision trees. *Management Decision*, *33*(8), 46–50. doi:10.1108/00251749510093932

Howard, R., & Matheson, J. (1984). Influence diagrams . In Howard, R., & Matheson, J. (Eds.), *Readings on the principles and applications of decision analysis*. Menlo Park, CA: Strategic Decisions Group.

Neumann, J. V., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Shachter, R. D. (1986). Evaluating influence diagrams. *Operations Research*, *34*(6), 871–882. doi:10.1287/opre.34.6.871