International Journal of Approximate Reasoning 96 (2018) 1–17

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Decision analysis networks

Francisco Javier Díez *, Manuel Luque, Iñigo Bermejo

Dept. Artificial Intelligence, Universidad Nacional de Educación a Distancia (UNED), Juan del Rosal 16, 28040 Madrid, Spain

A R T I C L E   I N F O

Article history:
Received 7 July 2017
Received in revised form 15 December 2017
Accepted 21 February 2018
Available online 27 February 2018

Keywords:
Decision analysis
Decision trees
Influence diagrams
Probabilistic graphical models
Asymmetric decision problems

A B S T R A C T

This paper presents decision analysis networks (DANs) as a new type of probabilistic graphical model. Like influence diagrams (IDs), DANs are much more compact and easier to build than decision trees and can represent conditional independencies. In fact, for every ID there is an equivalent symmetric DAN, but DANs can also represent asymmetric problems involving partial orderings of the decisions (order asymmetry), restrictions between the values of the variables (domain asymmetry), and conditional observability (information asymmetry). Symmetric DANs can be evaluated with the same algorithms as IDs. Every asymmetric DAN can be evaluated by converting it into an equivalent decision tree or, much more efficiently, by decomposing it into a tree of symmetric DANs. Given that DANs can solve symmetric problems as easily and as efficiently as IDs, and are more appropriate for asymmetric problems—which include virtually all real-world problems—DANs might replace IDs as the standard type of probabilistic graphical model for decision support and decision analysis. We also argue that DANs compare favorably with other formalisms proposed for asymmetric decision problems. In practice, DANs can be built and evaluated with OpenMarkov, a Java open-source package for probabilistic graphical models.

1. Introduction

The two formalisms most widely used for the representation and analysis of decision problems are decision trees (DTs) [31] and influence diagrams (IDs) [15]. DTs have the advantage of almost absolute flexibility, but also have three drawbacks: their size grows exponentially with the number of variables, they cannot represent conditional independencies, and they require in general a preprocessing of the probabilities [15,4]; for example, medical diagnosis problems are usually stated in terms of direct probabilities, namely the prevalence of the diseases and the sensitivity and specificity of the tests, while DTs are built with inverse probabilities, i.e., the positive and negative predictive values of the tests. Even in cases with only a few chance variables, this preprocessing of probabilities is a difficult task.[1] In contrast, IDs have the advantages of being very compact, easily representing conditional independence, and using direct probabilities; but they can only represent symmetric decision problems. Following partially the definitions of Bielza et al. [4] and Jensen et al. [17], we say that there is *domain asymmetry* when the value taken by a variable restricts the values that other variables can take; there is *information asymmetry* when a variable Y is observed for some values of a variable X but not for others; and there is *order asymmetry* when the decisions can be made in different orders.

* Corresponding author.
E-mail addresses: fjdiez@dia.uned.es (F.J. Díez), mluque@dia.uned.es (M. Luque), ibermejo@dia.uned.es (I. Bermejo).

[1] When there are two or more chance variables between two decisions in a DT, their order can be changed without altering the results of the evaluation, but when there is a decision between two chance nodes, the order of the two cannot be changed. For this reason it is often necessary to preprocess the probabilities.

https://doi.org/10.1016/j.ijar.2018.02.007
0888-613X/

In practice, virtually all real-world problems are asymmetric, in particular all those that involve the possibility of getting additional information at a cost; for example, the variable that represents the result of a test can take some values only when the decision is to do the test. Several formalisms have been proposed for representing and solving asymmetric decision problems, but all of them have drawbacks; for example, unconstrained IDs [18] cannot represent domain asymmetry nor information asymmetry, and sequential IDs [17] may need redundant links with complex labels, as we discuss in Section 3. In this paper we present a new formalism, called *decision analysis networks* (DANs), which can represent all symmetric problems as easily as IDs and, as we argue below, can also represent real-world asymmetric decision problems more naturally than other existing formalisms. We developed it when trying to solve a complex medical decision problem: the mediastinal staging of non-small cell lung cancer [26].

The DAN for that medical problem, together with other DANs for the most famous asymmetric decision problems proposed in the literature, can be found at www.ProbModelXML.org/networks; they are encoded in ProbModelXML, an open format for probabilistic graphical models [3]. OpenMarkov, an open-source tool for probabilistic graphical models, can be used to view, create, edit, and evaluate DANs. It is available at www.openmarkov.org. Its tutorial explains how to build and evaluate DANs and, when possible, compares them with equivalent IDs.

The rest of the paper is structured as follows. First, we introduce the n-test problem, which will serve to illustrate the properties of DANs and to discuss different formalisms. Then the paper consists of two main parts. The first deals with knowledge representation: Section 2 presents the definition of DANs and Section 3 compares them with other formalisms. The second deals with inference: Section 4 describes several algorithms for the evaluation of DANs and Section 5 analyzes the efficiency of those algorithms. Finally, Section 6 contains the conclusions and some proposals for future work.

Example 1. The *n-test problem* consists in deciding how to treat a patient that may suffer from a certain disease. After an initial examination of the symptoms, the doctor may order one or more of n available tests, each one having a cost. Every test can be performed once at most and its result is known immediately. The doctor has to decide which tests to perform and in which order.

In the simplest version of the problem there is only one symptom and all variables are dichotomous, i.e., the disease and the symptom are either present or absent and the result of each test is either positive or negative. The diabetes problem [10] is an instance of the two-test problem.

2. Definition of a DAN

In this section we define DANs by describing the elements that compose them and the ways of indicating the availability of information; we also discuss the meaning of the different types of links. We make several remarks about causality and availability of information, which may be useful for those who wish to apply DANs to real-world problems. However, a reader interested only in an axiomatic definition of DANs can skip those remarks without reducing the logical consistency of the presentation.

2.1. Graph and variables of a DAN

In this paper we represent variables with capital letters (X) and their values with lower-case letters (x). A bold upper-case letter ($\mathbf{X}$) denotes a vector of variables and a bold lower-case letter ($\mathbf{x}$) represents a configuration of them, i.e., the assignment of a value to each variable in $\mathbf{X}$. When $\mathbf{Y} \subseteq \mathbf{X}$, $\mathbf{x}^{\downarrow \mathbf{Y}}$ denotes the projection $\mathbf{x}$ onto $\mathbf{Y}$, which is the configuration of $\mathbf{Y}$ compatible with $\mathbf{x}$.

The set of nodes of a DAN, $\mathbf{V}$, can be partitioned into three disjoint subsets: chance variables ($\mathbf{C}$), decisions ($\mathbf{D}$), and utilities ($\mathbf{U}$). Chance variables represent properties or events that are not under the direct control of the decision maker, decisions correspond to actions that are under their direct control, and utilities represent payoffs (the decision maker's values). A DAN also has an acyclic directed graph such that each node represents a variable; hence we speak indifferently of nodes and variables. When the graph has a link $X \to Y$, we say that X is a parent of Y and Y is a child of X. The set of parents of a node X is denoted by $Pa(X)$, and $pa(X)$ represents a configuration of them. When there is a directed path from X to Y, we say that X is an ancestor of Y and Y is a descendant of X. In this paper we assume that utility nodes do not have children, but this requirement can easily be relaxed. We also assume that all the decisions and chance variables are discrete.

The DAN in Fig. 1 contains 3 decisions, drawn as rectangles, 4 chance variables, drawn as rounded rectangles, and 3 utility nodes, drawn as hexagons.

2.2. Restrictions

A *restriction* associated to a link $X \to Y$, such that X and Y are chance or decision variables, is a pair (x, y), where x is a value of X and y is a value of Y. It means that variable Y cannot take the value y when X takes the value x. In OpenMarkov the restrictions between X and Y are represented by a table with a column for each value of X and a row for each value of Y; when the values x and y are compatible, i.e., when there is no restriction (x, y), the corresponding cell contains a 1 and is colored in green; otherwise it contains a 0 and is colored in red—see Fig. 2. When all the values of Y are incompatible