Ignorant Influence Diagrams

Marco Ramoni
Cognitive Studies in Medicine
McGill Cognitive Science Centre
McGill University, Montreal, Canada

Abstract

Influence Diagrams (IDS) are a graphic formalism able to provide a compact representation of decision problems. IDs are based on the axioms of probability and decision theory, and they define a normative framework to model decision making. Unfortunately, IDs require a large amount of information that is not always available to the decision maker. This paper introduces a new class of IDS, called *Ignorant Influence* Diagrams (iiDs), able to reason on the basis of incomplete information and to improve the accuracy of their decisions as a monotonically increasing function of the available information, IIDS represent a net gain with respect to the traditional IDs, since they are able to explicitly represent lack of information, without loosing any capability of traditional IDs when the required information is available. Furthermore, I IDs provide a new method to assess the reliability of the decisions by replacing the traditional sensitivity analysis with a single analytical measure.

1 Introduction

Bayesian Belief Networks (BBNS) [Pearl, 1988] are a well-known formalism to reason under uncertainty and they have been successfully applied to a variety of problems in different domains. A BBN is a direct acyclic graph in which nodes represent stochastic variables and arcs represent conditional dependencies among variables. From a probabilistic point of view, they provide a straightforward way to represent independence assumptions among variables, thus making easy the representation and the acquisition of knowledge. BBNs are particularly appealing since they are based on a sound probabilistic semantics and they easily extend into a complete decision theoretic formalism, called *Influence Diagrams* (IDs). IDs [Horviti *et al*., 1988] provide a compact representation for decision problems and their sound probabilistic semantics guarantees the normative character of their decisions, IDS are an appealing complement to more traditional methods for representing decision problems, such as tables of joint probability distributions or decision trees, because they exploit the ability of BBNs to express conditional independence assumptions in graphical terms, thus dramatically reducing the amount of information needed to specify a decision problem.

Nonetheless, a BBN still requires a *fixed* and potentially very large amount of probabilistic information, that is not always available to the decision maker: the number of conditional probabilities needed to specify a conditional dependency grows exponentially with the number of its parent variables. Current propagation algorithms require that all the conditionals probabilities defining a conditional dependency among variables have to be known, as well as all prior probabilities for the states of the root variables, before any reasoning process can start. Furthermore, these probability measures have to be assessed as point-valued probabilities, even when the decision maker is not completely sure about them. This requirement, called *Credal Uniqueness* [Stiling and Morrel, 199l], is one of the most controversial points of Bayesian probability and decision theory [Levi, 1980; Kyburg, 1983], and it is the reason why BBNs require such a large amount of probabilistic information: in order to specify a unique probability distribution over the stochastic variables of a BBN, we need to know the conditional probabilities relating it with all its parents in the network.

This limitation becomes even more apparent in the development of an ID: when the decision maker is not able to specify a single probability, he is nonetheless forced to provide point-valued probability measures, and then to perform a costly and tiring analysis to assess the sensitivity of the resulting decisions to all the possible combinations of his imprecise assessment.

To overcome this limitation, and maintain the appealing features of probabilistic soundness and graphical nature of BBNs, we have developed a class of BBNs, called ignorant *Belief Networks* (iBNs) [Ramoni and Riva, 1994], able to relax the Credal Uniqueness assumption and to reason on the basis of incomplete probabilistic information. IBNs implement an inference policy, largely wished in the literature about probabilistic reasoning systems, called *incremental refinement* policy [Horviti, 1989], able to improve the accuracy of the solutions as a monotonically increasing function of the allocated resources and the available information.

The aim of this paper is to extend IBNs into a complete decision theoretic formalism called *Ignorant* Influence Diagrams (iIDs), and to show how I IDs can be useful to model decision making when the information required by traditional IDs is not available. The reminder of this paper will briefly outline the theory and the properties of the IBNs. Then, it will describe the way in which an IBN can be extended into an IID and which decision procedures are needed when the available information is not sufficient to specify point-valued probability measures. It will also outline a new method, provided by IIDs, to assess the reliability of decisions without the costly sensitivity analysis required by traditional IDs. A simple example will illustrate the properties of IIDs and a brief comparison with some related works will be provided.

2 Ignorant Belief Networks

The representation and use of incomplete information is a long standing challenge for AI researchers. During the past decade, they have developed a class of reasoning systems, called *Truth* Maintenance *Systems* (TMSs) [McAllester, 1990], which incrementally record justifications for beliefs and propagate Boolean truth values along chains of justifications. TMSs are independent reasoning modules which incrementally maintain the beliefs for a general problem solver and enable it to reason on the basis of temporary assumptions and incomplete information. TMSs able to propagate probabilistic rather than binaries truth-values are called Belief Maintenance *Systems* (BMSs) [Falkenhainer, 1986; Laskey and Leaner, 1989]. IBNS are belief-maintained BBNs: they exploit a BMS based on probabilistic logic, and therefore called *Logic-based* BMS (LBMS) [Ramoni and Riva, 1993].

2.1 Belief Maintenance

A LBMS uses a standard propositional language $\mathcal{L}$ defined by a set of *atomic propositions* $\mathcal{S} = \{a_1, a_2, \ldots, a_n\}$ and by the standard Boolean operators $\neg$, $\vee$, $\wedge$, $\supset$, and $\equiv$. A *literal* l is an atomic proposition a_i or its negation $\neg a_i$. An atomic proposition a_i is a *positive* literal and the negation of an atomic proposition $\neg a_i$ is a *negative* literal. A *clause* C is a finite disjunction of literals $\bigvee_{i=1}^{s} l_i$. A *Conjunctive Normal Form* (CNF) *formula* f is a finite conjunction of clauses $\bigwedge_{i=1}^{t} C_i$. Any legal formula of a propositional language can be converted into a CNF formula.

The LBMS relaxes the Credal Uniqueness assumption by replacing the standard point-valued probability function $p(f)$ with a convex set of these functions. The resulting evaluation function $P(f)$ of the language $\mathcal{L}$ will define a real interval between two probability functions $p_1(f)$ and $p_2(f)$ such that

$$P(f) = \{p^{\alpha}(f) | p^{\alpha}(f) = \alpha p_1(f) + (1-\alpha) p_2(f)\} \qquad (1)$$

where α is a real number and $0 \le \alpha \le 1$. The interval assigned by the function $P(f)$ to the formulas in $\mathcal{L}$ is called the *label* of f and we will denote with $P_*(f) = \pi_*$ and $P^*(f) = \pi^*$ the lower and the upper bounds of the interval $P(f) = [\pi_*, \pi^*]$, respectively. A label $P(f) =$ [Π, *Π']* is satisfied for any subinterval of [π, π*].

The LBMS uses a forward chained unit-resolution style algorithm called *Epistemic Constraint Propagation* (BCP) to propagate labels over a network of propositions! formulas, BCP can be regarded as an extension to intervals of the Boolean Constraint Propagation (BCP) algorithm [McAllester, 1990] used by the TMSs based on the propositions! calculus. The intuition behind the BCP is simple and elegant. The algorithm starts converting any formula in CNF, that is, a set of clauses. Bach clause acts as a constraint on the truth-values of the literals occurring in it. To be satisfied, a clause must contain at least one literal labeled as true. A clause is *violated* when all the literals occurring in it are labeled as *false,* thus producing a *contradiction.* When all literals but one in a clause are labeled as *false* BCP forces the unlabeled literal to be true.

In order to extend the BCP from Boolean to probabilistic truth-values, we derived, from the theory of *probabilistic entailment* [Nilsson, 1986], a probabilistic interpretation of disjunction able to define which constraints are imposed by a clause over the (probabilistic) truth-values of the literals occurring in it [Ramoni and Riva, 1995].

The first constraint, imposed by a clause over the literals occurring in it states that the label of a literal /, in clause $\bigvee_{i=1}^{s} l_i$ is bounded by:

$$P_*(l_i) \ge P_*(\bigvee_{i=1}^{s} l_i) - \sum_{j \ne i}^{s} P^*(l_j) + \mathcal{F} \qquad (2)$$

where

$$\mathcal{F} = \sum_{\lambda_1=0}^{1} \cdots \sum_{\lambda_s=0}^{1} P_*(\bigwedge_{i=1}^{s} l_i^{\lambda_i}) \cdot \Delta_C(\bigwedge_{i=1}^{s} l_i^{\lambda_i})$$

and the function Δ_c is defined as

$$\Delta_C(\bigwedge_{i=1}^{s} l_i^{\lambda_i}) = \max\{0, (\sum_{i=1}^{s} \lambda_i) - 1\}$$

The second constraint states that the label of the literal l_1 is bounded by:

$$\begin{aligned} P_*(l_1) &\ge \sum_{i=1}^{2^{s-1}} (1 - P^*(\neg l_1 \vee C_i)) \\ P^*(l_1) &\le 2^{s-1} - \sum_{i=1}^{2^{s-1}} P_*(\neg l_1 \vee C_i) \end{aligned} \qquad (3)$$

where {I1,...,!,} is a set of literals, and {C1, ...,C2—»} are the clauses built from {/1,..., 1,} with all the possible combinations of the negated and unnegated literals in the set {/a,...,l,}.

The propagation of these constraints is performed by a version of Welti's propagation algorithm [Walts, 1975] extended to intervals [Davis, 1987]: each proposition is labeled with a set of possible values, and the constraints are used to restrict this set. This property, which is implicit in the form of the inequalities 2 and 3, implies a monotonic narrowing of the labels, thus ensuring the incrementality of BCP.

The most important feature of BCP is the ability to reason from any subset of the set of clauses representing a joint probability distribution, by bounding the probability of the propositions within probability intervals, and incrementally narrowing these intervals as more information becomes available. Furthermore, BCP is *sound:* it never excludes from its intervals any probability value that could be derived by standard probability theory from the available information. Even if incomplete in general, BCP is complete with respect to the clauses representing a joint probability distribution. This means that, with respect to this subclass of the language, BCP returns the tightest entailed interval. The incompleteness with respect to other clauses of the language is the result of a compromise between expressivity and efficiency, since theoretical analysis and empirical results show that the BCP propagation runs to completion in linear time with respect to the number of clauses [Ramoni and Riva, 1995], thus making easy the estimation and the trading off of the computational effort.

2.2 Representation

IBNS are belief-maintained BBNs based on the LBMS. The IBN acts as a knowledge representation formalism expressing the assumptions of conditional independence in the domain of application and communicates the available conditional probabilities to the LBMS. These conditional probabilities are transformed into clauses relating the propositions of the LBMS which represent states of the stochastic variables of the IBN. In this way, any computation is left to the LBMS and the IBN can exploit the incremental character of BCP. The elements of a BBN can be easily translated into a LBMS network.

Nodes In a BBN, a node represents a stochastic variable. A stochastic variable is a set of mutually exclusive and exhaustive states. Therefore, the probability values assigned to the states in a variable have to sum to unit. In an IBN, when a variable is denned, each state is communicated to the LBMS as an atomic proposition. Moreover, a set of clauses is installed to ensure that the states of the variable are mutually exclusive and exhaustive. For all propositions $a_1, \dots, a_n$ in the LBMS representing the states of the variable, the disjunction $a_1 \vee \dots \vee a_n$ and all the conjunctions $\neg(a_i \wedge a_j)$ (with $i \neq j$) are asserted as true in the LBMS. When a probability value is assigned to a proposition a_1 representing a state of the variable, the LBMS receives the conjunction $P(a_1 \wedge \neg a_2 \wedge \dots \wedge \neg a_n) = P(a_1)$

Arcs In a BBNs, arcs represent conditional dependencies among nodes. A conditional dependency defines a dependency relation between a set of parent nodes and a child node. A conditional dependency is denned by the conditional probabilities $P(a_i|a_k, \dots, a_s) = [\pi_*\ \pi^*]_i$ where $a_h, \dots, a_s$ is a combination of states of the stochastic variables represented by the parent nodes of the dependency and *at* is a state of stochastic variable represented by the child node. Conditional dependencies can be propagated both ways over an IBN, thus emulating the two main operations involved in the evaluation of a BBN: node removal and arc reversal.

2.3 Inference

There are two main operations involved in the evaluation of an ID: node removal and arc reversal. Node removal corresponds to propagating probability values along the direction of the arcs in the graph, while arc reversal corresponds to flowing backward the direction of the arcs and assessing the posterior probability of parent nodes in a dependency.

Node Removal In a BBN, node removal corresponds to marginalization. When the probability values of all states represented by the propositions $a_j \dots a_s$ is assigned, the two different clauses resulting from the application of the De Morgan's laws to $(a_i \wedge a_j \dots a_s)$ and $(\neg a_i \wedge a_j \dots a_s)$ are communicated to the LBMS algorithm. $P(a_i \wedge a_j \dots a_s)$ and $P(\neg a_i \wedge a_j \dots a_s)$ are calculated according to a version of the chain rule extended to intervals:

$$P_*(a_i \wedge a_j \dots a_s) = \prod_{h=j}^{s} P_*(a_h) P_*(a_i | a_j, \dots, a_s)$$

$$P^*(a_i \wedge a_j \dots a_s) = \prod_{h=j}^{s} P^*(a_h) P^*(a_i | a_j, \dots, a_s)$$

$$P_*(\neg a_i \wedge a_j \dots a_s) = \prod_{h=j}^{s} P_*(a_h)(1 - P^*(a_i | a_j, \dots, a_s))$$

$$P^*(\neg a_i \wedge a_j \dots a_h) = \prod_{h=j}^{s} P^*(a_h)(1 - P_*(a_i | a_j, \dots, a_s))$$

The resulting conjunctions are converted into clauses by the LBMS before propagating them through BCP.

Arc Reversal From a probabilistic point of view, the capability of performing arc reversal in IBNs is provided by the well-known Bayes' Theorem:

$$p(a_j \dots a_s | a_i) = \frac{p(a_i | a_j \dots a_s) \prod_{h=j}^{s} p(a_h)}{p(a_i)} \quad (4)$$

where $i \notin \{j, \dots, s\}$. When provided with new evidence $\underline{p}(a_i)$, we can apply the chain rule to the conditional probability calculated by Formula 4 and $p(a_1)$ to obtain the conjunction:

$$p(a_i \wedge a_j \dots a_s) = p(a_j \wedge \dots \wedge a_s | a_i) \underline{p}(a_i).$$

Since $p(\neg a_i | a_j \dots \wedge a_s) = 1 - p(a_i | a_j \wedge \dots \wedge a_s)$, we can also derive:

$$p(\neg a_i \wedge a_j \wedge \dots \wedge a_s) = p(a_j \wedge \dots \wedge a_s | \neg a_i) \underline{p}(\neg a_i).$$

When converted into clausal form, these conjunctions turn out to be the same clauses that were generated during marginaliiation. Therefore, arc reversal results in an

updating of the probabilities of a set of already existing clauses. Extending this definition to intervals, the arc reversal formula for IBNs is obtained as follows:

$$P_*(a_i \wedge a_j \ldots a_s) = \frac{P_*(a_i)}{P^*(a_i)} P_*(a_i | a_j, \ldots, a_s) \prod_{h=j}^{s} P_*(a_h)$$

$$P^*(a_i \wedge a_j \ldots a_s) = \frac{P^*(a_i)}{P_*(a_i)} P^*(a_i | a_j, \ldots, a_s) \prod_{h=j}^{s} P^*(a_h)$$

It is worth noting that the resulting posterior probability set is still convex, since it is the result of a linear mapping of the extreme values of a probability interval, and BCP was proved to preserve convexity during propagations [Ramoni and Riva, 1993]. Furthermore, in IBNS, the inversion formula shares with the chain rule the appealing property of propagating only the available information, thus preserving the incremental refinement policy of **BCP.**

2.4 Properties

From the theory of TMSS, the LB MS inherits the concept of *consumer* [de Kleer, 1986]. A consumer is a forward-chained procedure attached to a proposition, that is fired when the truth-value of the proposition is changed, that is, the probability interval associated with a state is narrowed. Using consumers, IBNs do not perform any computation themselves, but rather act as a high-level knowledge representation language, while the propagation of probabilities is performed by the LBMS.

There are some properties of the IBNs that will be crucial in the development of the IIDs. First of all IBNs converge toward point valued probabilities, and when all the conditionals defining a joint probability distribution, they behave as standard BBNS, returning point-valued probabilities. Furthermore, an IBN will infer the tightest intervals from any subset of conditional probability in a conditional dependency, since the LBMS is compete for clauses representing joint probability distributions and the IBN simply minimises and maximise the standard rules of node removal and arc reversal.

Finally, it is worth noting that the LBMS both performs and drives the propagation, since consumers are attached to the propositions of the LBMS and are fired according to the changes occurring in their labels. Therefore, the computational cost of a propagation grows linearly in space and time with respect to the number of conditional probabilities, even if the number of conditional probabilities needed to specify a conditional dependency grows exponentially with the number of parent nodes in the dependency. However, the incremental character of inference policy implemented by the IBNs will allow the decision maker to trade execution time with precision of solutions, since an IBN will propagate only those conditional probabilities explicitly assessed by the decision maker.

3 Influence Diagrams

IDS [Horvits *et a*/., 1988] are a natural extension of BBNs. They allow the formulation of a decision problem into the sound and compact formalism of BBNS. In this section, we will illustrate how IBNs can be easily extended to a complete decision formalism, thus creating a new class of IDS called *Ignorant* Influence Diagrams (IIDS). IIDs inherit from the IBNs the ability to reason on the basis of incomplete information and to incrementally refine the accuracy of their decisions as more information becomes available.

Figure 1: A simple *Ignorant Influence Diagram.*

3.1 Representation

IDs are BBNs containing three different kinds of nodes: *chance* nodes (also called *state* nodes), *decision* nodes, and *value* nodes (also called *preference* nodes). These nodes are related by standard conditional dependencies. The resulting ID can be transformed into a BBN following the method proposed by Cooper [1988]. On this view, the decision problem can be solved by determining the instantiations of the decisions which maximise the expectations of the decision maker.

Chance Nodes A chance node represents a state of the world. It is basically a standard stochastic variable. BBNs are usually defined as "influence diagrams containing just chance nodes" [Horvits *et a*/., 1988]. In Figure 1, chance nodes are depicted as oval nodes in the graph.

Decision Nodes A decision node identifies a set of possible alternative actions available to the decision maker. In the IID of Figure 1, the decision node is depicted as a square node. A set of actions representing a possible solution for a decision problem is called *strategy* or *policy.*

Value Nodes Different strategies lead to different outcomes. Value nodes represent preferences or utilities of the decision maker for alternative outcomes. A decision problem may be represented as the problem of finding the strategy which maximises the preferences expressed by the decision maker over the possible outcomes of the problem. In Figure 1, the value node is depicted as a diamond.

3.2 Decision Criteria

A standard ID is solved by choosing the strategy with maximal *expected utility* of its outcomes. In the example depicted in Figure 1, the expected utility of a strategy is the value assigned by the network to the value node. The optimal strategy s_i will be the strategy with the highest expected utility $u(s_i)$. As IIDs propagate convex sets of probability distributions, they will also lead to convex sets of expected utilities for each strategy $U(s_i)$. We will denote with $U_*(s_i)$ and $U^*(s_i)$ the lower and upper bounds of the convex set of expected utilities for the strategy s_i, respectively. There are several criteria to rank convex sets of expected utilities.

The most conservative criterion is called *Stochastic Dominance* [Kyburg, 1983] and dictates that a strategy s_i has to be preferred to a strategy s_j if and only if $U_*(s_i) > U^*(s_j)$ (i.e. the expected utility intervals do not overlap). This criterion is extremely safe, because it provides a decision if and only if there will be no chance that any additional information will change the ranking order among s_i and s_j. However, this criterion is often unable to discriminate in a variety of decision situations.

When the Stochastic Dominance criterion fails, we have to resort to a weaker condition of admissibility, able to discriminate among competing strategies when the expected utility intervals are not disjoint [Levi, 1980]. Pittarelli [1988] proposes to adopt a generalised version of the well-know *Hurwics* criterion. This criterion suggests to rank the possible strategies according to a weight average of their minimum and maximum expected utility, using a constant α , with $0 \leq \alpha \leq 1$. Therefore, the strategy s_i is preferred to the strategy s_j if and only if

$$U_*(s_i)\alpha + U^*(s_i)(1-\alpha) > U_*(s_j)\alpha + U^*(s_j)(1-\alpha)$$

The constant α, called *Hurwics* value, can be thought as a *boldness* index representing the daring attitude of the decision maker. When $\alpha = 1$, the Hurwics criterion reduces to the well-known *maximin* criterion. The *maximin* criterion prescribes to select the strategy having the highest minimum expected utility. Hence, a strategy s_i is preferred to a strategy s_j if and only if $U_*(s_i) > U_*(s_j)$. This criterion reflects the behavior of a cautious decision maker, who wants to be sure that, even if an unfavorable state of the world occurs, there is a known minimum payoff below which he cannot fall.

When $\alpha = 0$, the Hurwics criterion collapses on the *maximax* criterion. The *maximax* criterion adopts the opposite point of view than the maximin criterion. It consider only the maximum expected utility of the strategies and select the strategy with the highest. Then a strategy s_i is preferred to a strategy s_j if and only if $U^*(s_i) > U^*(s_j)$. This criterion reflects the standpoint of a daring gabler who cares just about the maximum payoffs of his strategies and can afford to stand possible losses.

3.3 Sensitivity Analysis

Traditional decision theoretic formalisms require the assessment of point-valued probability measures, even when the decision maker is not completely confident about them. Since the decision maker cannot specify confidence intervals, he will have to perform a tiring and costly analysis of the sensitivity of the decisions to all the probability values he does not feel sure about. IIDs reduce the cost of this procedure by solving the diagram just once and producing a sort of "simultaneous" multivariate sensitivity analysis. An advantage of this approach is that it provides a straightforward way to evaluate the sensitivity of decisions not only to prior probabilities but also to conditional probabilities, which are usually difficult to modify and test using traditional sensitivity analysis methods.

Probability	1	2
$P($ [cause=yes] $)$	[0.5 0.7]	0.6
$P($ [effect=yes] | [cause=yes] [action=yes] $)$	[0.3 0.5]	0.4
$P($ [effect=yes] | [cause=yes] [action=no] $)$	[0.5 0.8]	0.8
$P($ [effect=yes] | [cause=no] [action=yes] $)$	[0.1 0.3]	0.1
$P($ [effect=yes] | [cause=no] [action=no] $)$	—	0.2
$P($ [cost=yes] | [action=yes] $)$	[0.8 0.9]	0.9
$P($ [cost=yes] | [action=no] $)$	0.01	0.01

Table 1: The probability measures defining the IID of the example.

Furthermore, the adoption of the Hurwics criterion to discriminate between competing strategies introduces a natural measure of the reliability of the decisions. When maximin and maxima* criteria conflict, the Hurwics value itself becomes a measure of reliability. We can analytically identify the decision threshold for the Hurwici value in order to assess the sensitivity of the discrimination between the strategy s_1 and the strategy S_j to the boldness attitude of the decision maker, using to the following formula:

$$\tau = \frac{U_*(s_i) - U_*(s_j)}{U^*(s_j) - U_*(s_j) + U_*(s_i) - U^*(s_i)} \qquad (5)$$

Intuitively, this formula identifies for which Hurwics value the decision maker will change his policy, and therefore the robustness of the decision reached so far. The lower is the decision threshold, the lower will be the chance that new information will change the preferred strategy, and therefore, the higher will be the quality of our decision. The value r will be an Hurwics value (that is, $0 < T < 1$) as long as one utility interval will be subset of the other, and therefore the maximin and the maximax criteria will conflict. When $r < 0$, both criteria will select the same strategy.

The optimality of these results is guaranteed by the ability of IBNS to return the tightest interval for any state from any subset of the conditional probabilities defining a conditional dependency and by the soundness of the method developed by Cooper [1988] to transform an ID into a BBN.

4 Example

In a standard ID, all the conditional probabilities that make up a conditional probability distribution are

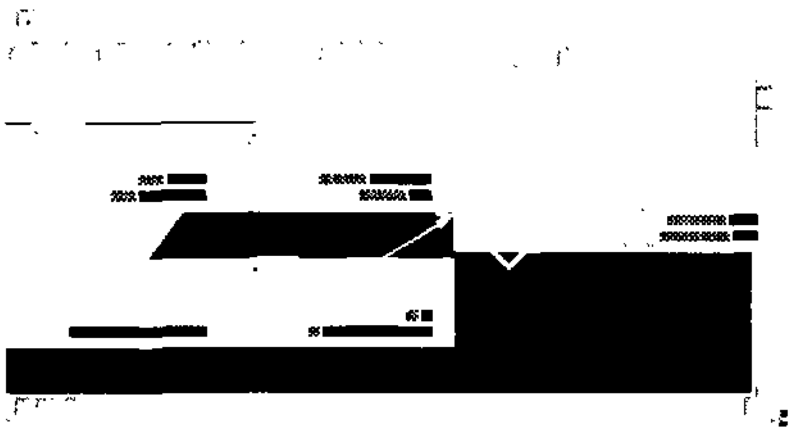

Figure 2: The IID defined by the probability measures reported in the first column of Table 1.

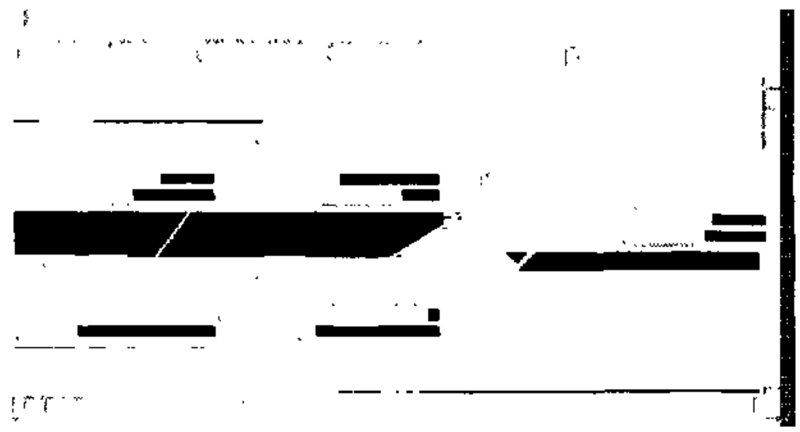

Figure 3: The IID defined by the probability measures reported in the second column of Table 1.

needed before any reasoning process can start. We will now show with an example how an IID is instead able to reason from incomplete conditional probability distributions. Figure 1 shows the IID used as an example. It represents the decision problem of taking a certain action in order to prevent a certain effect of a particular cause. The action has a foreseeable cost, and the problem consists in trading-off the value gained by prevention of the effect with the cost of the action.

The example is composed of two steps. In the first step, the system receives only the probability measures listed in the first column of Table 1, with the preference function over the outcomes defined as: v([cost=yes] [effect=yes]) = 0, v([cost=yes] [effect=no]) = 0.8, v([cost=no] [effect=yes]) = 0.2, and v([cost=no] [effect=no]) = 1. Figure 2 shows the graphical representation of the network generated in the first step of the example. The pop-up windows over the nodes graphically report the probability interval associated to each one of their states. In each bar, the area between 0 and $P_*(a_1)$ is black, the area between $P^*(a_i)$ and 1 is white, and the area between $P_*(a_i)$ and $P^*(a_1)$ is gray. Thus, the width of the gray area is proportional to the ignorance about the probability.

The decision node reports its value in a strategy. The chances nodes report the probability values of each of their states for the same strategy. It is worth noting that the conditional P([effect=yes] | [cause=no] [action=no]), not reported in the first column of Table 1, is actually not included in the definition of the first IID, and therefore its probability value is not even propagated. The value node value reports the expected utility intervals for the possible strategies [action=y»s] and [action=no], namely: U([action=yes]) = [0.344 0.768] and U([action=no]) = [0.2767 0.7993].

It is apparent that the Stochastic Dominance criterion is unable to select only one strategy, and remains undecided. The weaker Hurwici criterion is able to pick up a unique decision. However, the solution can change according to the Hurwici value. Using formula 5, we can estimate the robustness of the decision by calculating the decision threshold for the Hurwici value. In our case, the decision threshold is r — 0.6825.

When all the probability measures required by a standard ID are specified, as in the second column of Table 1, the IID behaves as a standard ID in returning point-valued probabilities, as shown in Figure 3. Since all the criteria adopted collapse on the maximum expected utility criterion when they have to rank point-valued expected utilities, also the solution provided by the IID is identical to the one suggested by a standard ID.

5 Related Work

There are at least two line of research trying to relax the Credal Uniqueness assumption within a coherent Bayesian framework and to develop automated decision making systems able to reason on the basis of interval rather than point-valued probability functions.

The first line of research is based on the concept of *probabilistic* database, which generalises the standard relational model by replacing the characteristic function of a relation with a finite probability distribution function. Using this concept, Pittarelli [1994] proposed a model able to represent probability intervals and provided methods for decision making based on it. The analogy between a probabilistic database and a table of joint probability distributions is apparent. Therefore, IID s improve the probabilistic database model as IDs improve the traditional decision theoretic methods based on tables of joint probability distributions: IIDs explicitly represent conditional independence assumptions in the domain of application, thus reducing the amount of probabilistic information needed to specify a decision problem.

Closer to the aim of IIDS are the efforts addressed by Breeze and Pertig [1990] to develop *Interval Influence Diagrams.* They describe procedures for node removal and arc reversal in IDS where lower bounds of probability intervals are stored at each node in the ID. The appealing feature of this method is its effort to preserve both the probabilistic soundness and the graphical nature of standard IDs. Unfortunately, the probability bounds calculated by this method quickly degradatc during the propagation, thus resulting in the assignment of too wide probability intervals and jeopardising the normative character of their decisions: an analog of the classical Dutch

Book can be made against an agent making decisions on the basis of the probability distributions erroneously allowed by too wide probability intervals.

6 Conclusions

IDs are a powerful formalism to mechanise decision making, and they have been successfully applied to a wide range of problems. However, they require a large amount of information that is not always available to the decision maker. The acquisition of this large amount of information, either from human experts or from statistical analyses of databases, usually represents one of the major challenges in the process of developing decision support systems based on IDS. This paper introduced a new class of IDs, called IIDs, able to reason on the basis of incomplete information, and to incrementally refine the accuracy of their decisions as more information becomes available. However, when they are provided with complete probabilistic information, IIDs behave as standard IDS.

Therefore, IIDs represent a net gain with respect to the traditional IDS, since they are able to explicitly represent the actual lack of information, without loosing any capability of the traditional IDs when the required information is available. Furthermore, by relaxing the Credal Uniqueness assumption, I IDS provide a new method to assess the reliability of the decisions by replacing the costly and tiring sensitivity analysis with a single analytical measure. Finally, the monotonic, incremental character of the refinement process in the IIDs provides a way to trade-off the amount of computational time and available information with the accuracy of the decisions. This feature makes IIDs a suitable formalism for real-time, resource-bounded decision tasks.

Acknowledgments

This research was supported in part by the AIM Programme of the Commission of the European Communities (A2034). I would like to thank Alberto Riva, Greg Cooper, Carlo Beriuini and Riccardo Bellassi for their helpful suggestions, and Mario Stefanelli for his continuous support.

References

[Breeze and Fertig, 1990] J. Breese and K.W. Fertig. Decision making with interval influence diagrams. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence,* pages 122-129, 1990.

[Cooper, 1988] G.F. Cooper. A method for using belief networks as influence diagrams. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence,* pages 55-63, 1988.

[Davis, 1987] E. Davis. Constraint propagation with interval labels. *Artificial Intelligence,* 32:281-331,1987.

[de Kleer, 1986] J. de Kleer. Problem solving with the ATMS. *Artificial Intelligence,* 28:197-224, 1986.

[Falkenhainer, 1986] B. Falkenhainer. Towards a general purpose belief maintenance system. In *Proceeedings of the 2nd Workshop on Uncertainty in AI,* pages 71-76, 1986.

[Horvits *et al.,* 1988] E.J. Horviti, J.S. Breese, and M. Henrion. Decision theory in expert systems and artificial intelligence. *International Journal of Approximate Reasoning,* 2:247-302, 1988.

[Horvits, 1989] E.J. Horvits. Reasoning under varying and uncertain resource constraints. In *Proceedings of the International Joint Conference on Artificial Intelligence,* pages 1121-1127. Morgan Kauffman, San Mateo, CA, 1989.

[Kyburg, 1983] H.E. Kyburg. Rational belief. *Behavioral and Brain Sciences,* 6:231-273, 1983.

[Laskey and Lehner, 1989] K. B. Laskey and P. E. Lehner. Assumptions, beliefs and probabilities. *Artificial Intelligence,* 41(l):65-77, 1989.

[Levi, 1980] I. Levi. *The Enterprise of Knowledge. An Essay on Know ledge, Credal Probability and Chance.* MIT Press, Cambridge, MA, 1980.

[McAllester, 1990] D. McAllester. Truth maintenance. In *Proceedings of the National Conference on Artificial Intelligence,* pages 1109-1115, 1990.

[Nilsson, 1986] N.J. Nilsson. Probabilistic logic. *Artificial Intelligence,* 28:71-87, 1986.

[Pearl, 1988] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of plausible inference.* Morgan Kaufmann, San Mateo, CA, 1988.

[Pittarelli, 1988] M. Pittarelli. Decision making with linear constraints on probabilities. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence,* pages 283-290, 1988.

[Pittarelli, 1994] M. Pittarelli. An algebra for probabilistic databases. *IEEE Transactions on Knowledge and Data Engineering,* 6(2):293-303, 1994.

[Ramoni and Riva, 1993] M. Ramoni and A. Riva. Belief maintenance with probabilistic logic. In *Proceedings of the AAAI Fall Symposium on Automated Deduction in Non Standard Logics,* Raleigh, NC, 1993. AAAI.

[Ramoni and Riva, 1994] M. Ramoni and A. Riva. Belief maintenance in bayesian networks. In *Proceedings of Tenth Conference on Uncertainty in Artificial Intelligence,* pages 204-212, 1994.

[Ramoni and Riva, 1995] M. Ramoni and A. Riva. Logic-based belief maintenance. Submitted for publication, 1995.

[Stiling and Morrel, 1991] W. Stiling and D. Morrel. Covex bayesian decision theory. *IEEE Transactions on Systems, Man, and Cybernetics,* 21:173-183,1991.

[Walts, 1975] D. Walts. Understanding line drawings of scenes with shadows. In P.H. Winston, editor, *The Psychology of Computer Vision,* pages 19-91. McGraw-Hill, New York, 1975.