



HAL
open science

An early guidance system for a general knowledge-based aiding framework using probabilistic interventions

Véronique Delcroix, Pierre-Henri Wuillemin

► To cite this version:

Véronique Delcroix, Pierre-Henri Wuillemin. An early guidance system for a general knowledge-based aiding framework using probabilistic interventions. 13èmes Journées d'Intelligence Artificielle Fondamentale (JIAF 2019), Jul 2019, Toulouse, France. pp.81-90. hal-02302950

HAL Id: hal-02302950

<https://hal.science/hal-02302950v1>

Submitted on 1 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An early guidance system for a general knowledge-based aiding framework using probabilistic interventions

Véronique Delcroix¹

Pierre-Henri Wuillemin²

¹ Univ. Polytechnique Hauts-de-France, CNRS, UMR 8201 - LAMIH, F-59313 Valenciennes, France

² Sorbonne université, CNRS, UMR 7606 - LIP6, F- 75252 Paris cedex 05, France
Veronique.Delcroix@uphf.fr Pierre-Henri.Wuillemin@lip6.fr

Résumé

Un problème de décision courant se répète de nombreuses fois, avec le même type d'alternatives et le même ensemble de critères, mais avec une situation de décision différente à chaque occurrence du problème. Dans ce type de problème, le conseil en amont vise à faciliter la sélection d'un sous ensemble d'alternatives satisfaisantes pour le cas de décision considéré, sans demander à l'utilisateur d'avoir des connaissances sur le problème. Cet article propose un système de conseil en amont basé sur un modèle de connaissances du problème de décision courant. Pour commencer, l'article présente la construction d'un réseau bayésien pour embarquer la connaissance dans le système. Ensuite, le concept d'intervention dans un réseau bayésien proposé par Pearl est étendu aux interventions probabilistes pour des variables simples et des ensembles de variables. Enfin, la procédure de conseil en amont pour un problème de décision courant est présentée, sur la base du modèle de connaissance et en utilisant les interventions probabilistes pour fixer l'écosystème de la personne, même lorsque le cas de décision n'est que partiellement observé.

Abstract

A common decision problem repeats a lot of time with the same kind of alternatives and the same set of criteria, but with a different decision case in each occurrence. The objective of early guidance in this kind of problem is to facilitate the selection of a subset of satisfactory alternatives for each new decision case, without asking the user any knowledge of the problem. This article proposes an early guidance system based on a model of knowledge of the common decision problem. It first presents the construction of a Bayesian network for a common decision problem to embed the knowledge in the aiding framework. Second, the concept of intervention proposed by Pearl is extended to probabilistic interventions for a single variable and for a set of variables. Finally the early guidance procedure is presented on the basis of the Bayesian network and using a probabilistic intervention to set a decision case even though it is

partially observed.

1 Introduction

A common decision problem occurs repeatedly in different decision cases, corresponding to different actors, needs, constraints and priorities, but including each time the same type of alternatives and the same set of decision criteria. An example is the problem of study and career guidance for students. This question, what to do after the baccalaureate, concerns a high number of young people, but each of them has its own characteristics (location, level, project, ambition, constraints, etc.). Another example of a common decision problem is the choice of a manual wheelchair that has been analyzed in [21].

The problem of early guidance in a common decision problem consists in providing advice at a very early step of a decision process, when the actors of the decision have not yet seriously thought about the problem. Early guidance precedes the whole decision process. At this step, the persons concerned by the consequences of a potential or future decision are considered to be merely passive.

Since the person concerned by the consequences of the choice is not obligatory the user of the general aiding framework, the early guidance system can be used by different persons involved in a decision case of a common decision problem. It can also be used as an automatic system, without any human user.

In a common decision problem, the whole decision process is most often achieved without the aid of any expert, and the actors involved generally gather only a small part of the relevant knowledge to guide the decision. For example, they do not necessary know what kind of alternative would best match their own situation. Because of that, it is important that an early guidance system does not appeal to the

actors knowledge or preference. The only certain available information is a partial description of the personal ecosystem of the person concerned by the consequences of the decision, including some of the constraints and preferences that could influence the decision.

As a consequence, the early guidance system has to embed general knowledge about the common decision problem, in order to provide some relevant guidance. This knowledge is about : (1) decision criteria to be considered, (2) alternative description in terms of characteristics and their evaluation (3) the personal ecosystem, *i.e.* any elements about the person and the context that could influence the choice.

In this paper, we address the new challenge that concern the design of an early guidance system, as part of a general aiding framework. Once defined, it should be possible to use the system easily for any new decision case, with the constraints presented above about the *ignorant* user. The objective of our early guidance system is to facilitate the selection of a small subset of satisfactory alternatives on the sole base of the available information about the decision case. Depending on the objective, this subset can be the final result, such as in the display of targeted car advertisement for a given internet user, or it could be used to choose an alternative among this subset, which can be done directly or by using a classic decision aiding system.

In the following of this paper, we define what is a common decision problem and set out the early guidance problem. Second, we present the knowledge based aiding framework and we explain the construction of the Bayesian network model of knowledge for a common decision problem. Then we describe the early guidance system based on the Bayesian network models. In that section, we explain the satisfactory situation principle, the definition and the implementation of probabilistic intervention, and the early guidance procedure. In the last section, we briefly present and discuss some related works.

2 Common Decision Problem and Early Guidance

In this section, we precisely define a common decision problem and the early guidance problem. We also provide a brief comparison with classical multi-criteria decision systems.

2.1 A Common Decision Problem

A *common decision problem* is a decision problem (such as choice, sorting or ranking) that repeats a lot of times, each time in a specific *ecosystem*, whose observed part is called a *decision case*. It involves alternatives whose evaluation regarding the decision criteria may also depend on the decision case.

The main components of a common decision problem are :

the characteristics of the alternatives ; The set of alternatives is potentially defined through the domain of the set of attributes. An alternative is an assignment of the set of characteristics. Some aberrant combinations of values are excluded due to some inherent constraints between some subsets of characteristics of an alternative.

the ecosystem of the person concerned by the decision ; it regroups any parameters related to the person and his/her context that may influence the decision. A decision case is a partial instantiation of the ecosystem, corresponding to the available observations of that specific case.

the set of decision criteria, each of them being characterized by :

the level of importance of the criterion in the decision case.

the criterion evaluation index that represents the evaluation (or the quality) of an alternative according to the criterion and the decision case.

the level of satisfaction brought by an alternative on the criterion in that decision case. It depends on the importance and the evaluation such that (1) when an alternative presents a very good quality for a criterion, the satisfaction is very good whatever the importance of the criterion ; (2) the higher is the quality of the alternative, the higher is the satisfaction, but the lower is the importance, the easier it is to get a good level of satisfaction ; (3) when the criterion is indifferent according to a decision case, the satisfaction is good whatever the quality of the alternative according to that criterion.

A *common decision problem* is a triple $(\mathbf{X}_{\text{Alt}}, \mathbf{X}_{\text{eco}}, \mathcal{S})$ where \mathbf{X}_{Alt} and \mathbf{X}_{eco} are finitely-valued variables and \mathcal{S} is a set $\{(Imp_g, Ind_g, S_g), g \in G\}$ with $G = \{1, 2, \dots, |G|\}$ and Imp_g, Ind_g, S_g being probabilistic functions such that :

Imp_g is defined on $\mathbf{X}_g^{\text{imp}} \subset \mathbf{X}_{\text{eco}}$ and $Imp_g(\mathbf{x}) = P(\cdot | \mathbf{x})$ with $\mathbf{x} \in \text{Dom}(\mathbf{X}_g^{\text{imp}})$.

Ind_g is defined on $\mathbf{X}_g^{\text{ind}} \subset \mathbf{X}_{\text{eco}} \cup \mathbf{X}_{\text{Alt}}$ and $Ind_g(\mathbf{x}) = P(\cdot | \mathbf{x})$ where $\mathbf{x} \in \text{Dom}(\mathbf{X}_g^{\text{ind}})$.

S_g only depends on Imp_g and Ind_g .

and

\mathbf{X}_{Alt} is the set of characteristics of the alternatives, \mathbf{X}_{eco} is the set of characteristics of the ecosystem, $\mathbf{a} \in \mathcal{A}$ is an alternative,

$\mathcal{A} \subset \times_{X \in \mathbf{X}_{\text{Alt}}} \text{Dom}(X)$ is the set of alternatives,

Let G a set of criteria and $(\mathbf{X}_{\text{Alt}}, \mathbf{X}_{\text{eco}}, \mathcal{S})$ a common decision problem for G . A *decision case* for $(\mathbf{X}_{\text{Alt}}, \mathbf{X}_{\text{eco}}, \mathcal{S})$

is a couple $(\mathbf{X}_{DC}, \mathbf{x}_{dc})$, where $\mathbf{X}_{DC} \subset \mathbf{X}_{eco}$ and $\mathbf{x}_{dc} \in \times_{Y \in \mathbf{X}_{DC}} Dom(Y)$.

In order to simplify the notation, we use the same notation to denote the probabilistic functions Imp_g , Ind_g , S_g and the resulting random variables. The variables Imp_g , Ind_g and S_g represent respectively the level of importance, the criterion evaluation index and the level of satisfaction.

2.2 The problem of early guidance

Given the set \mathcal{A} of potential alternatives, and a decision case, the problem of early guidance is to make easier the selection of a subset of satisfactory alternatives in the considered situation. The satisfaction refers to the level of satisfaction S_g provided by an alternative regarding each decision criterion

A naive approach would be to evaluate the satisfaction of each alternative according to each criterion, in the considered situation, and try to exploit this huge matrix to classify the alternatives. This is not relevant since the size of the set of alternatives \mathcal{A} is exponential in the number of attributes of the alternatives \mathbf{X}_{Alt} . This approach is possible when an initial small set of relevant alternatives has first been selected. It corresponds to the approach of Multi-criteria Decision Analysis that we discuss further. In our approach we focus on *satisfactory* alternatives, on the basis on the local level of satisfaction for each criterion, without trying to maximize any form of global satisfaction.

Since a common decision problem concerns a wide range of actors, we consider the possibility of "ignorant" actors, or with low involvement; we also consider decision cases where the selection a subset of satisfactory alternatives for a given decision case is achieved without any participation of the person concerned by the decision, as in targeted advertisement. Those reasons underline the interest of an early guidance system.

The repetitiveness of a common decision problem makes interesting to capitalized general knowledge about the decision problem and identify what could be reused in each new decision case. Thus, we propose a knowledge based system. The modeling of the common decision problem is made once. It embeds general probabilistic knowledge about the parameters that should be taken into account and the way they influence the decision. Each use of the early guidance system concerns a specific decision case of the common decision problem, meaning specific persons, situations, objectives, etc. It takes place before the decision process as usually understood in multi-criteria decision analysis (MCDA).

The next section provides a quick comparison between the early guidance problem and MCDA.

2.3 Position of early guidance with respect to MCDA

In multi-criteria decision analysis (MCDA), each decision problem is usually unique (not repeated). The methods developed for MCDA [9] are based on the hypothesis that the decision maker is strongly involved in the decision process, and that he/she is guided by an analyst. At the opposite, we consider that a case of decision of a common decision problem generally occurs without the presence of an analyst or a specialist, meaning that the actors may not be aware of all the aspects of the decision problem. Consequently, an early guidance system can not be based on preferences, whereas most MDCA methods are based on pairwise comparisons of alternatives regarding each criterion.

Another point concerns the definition of a criterion. In [20], a criterion is a tool allowing to compare alternatives according to a particular significance axis or a point of view. More precisely, in MCDA, a criterion is a real-valued function on the set of alternatives, such that it appears meaningful to compare two alternatives a and b according to a particular point of view on the sole basis of the two numbers $g(a)$ and $g(b)$ [2]. In a common decision problem, the evaluation of an alternative according to a criterion may also depend on the decision case, since what is suitable for a decision case may not be suitable for another one. As a consequence, the evaluation of an alternative may be different depending on the decision case.

Another difference is about the set of alternatives : in our approach, the alternatives are potentially defined through the domain of the set of attributes, whereas in MCDA, a problem is usually defined by the so called decision matrix that gathers the performances of n alternatives a with respect to m criteria, in addition with the weights of the criteria. The decision matrix is the start point in MCDA, whereas the subset of satisfactory alternatives is the objective of our proposed early guidance system.

However, the early guidance problem partially meets the objective of the screening problem in MCDA. A screening procedure aims to reduce a large set of alternatives to a smaller set that most likely contains the best choice. As mentioned in [8], screening "should eliminate alternatives that are unlikely to be chosen, so that later effort can be focused on the more attractive options." However, screening is based on preference information provided by the decision maker, whereas we only have a partial description of the ecosystem of the person concerned by the consequences of the decision.

3 A knowledge based aiding framework for a common decision problem

In this section, we present a general aiding framework for a common decision problem and we describe the construction of the knowledge models embedded.

3.1 Overview of the framework

Figure 1 shows the structure of an aiding framework for a common decision problem.

The early guidance system is only part of the knowledge based aiding framework

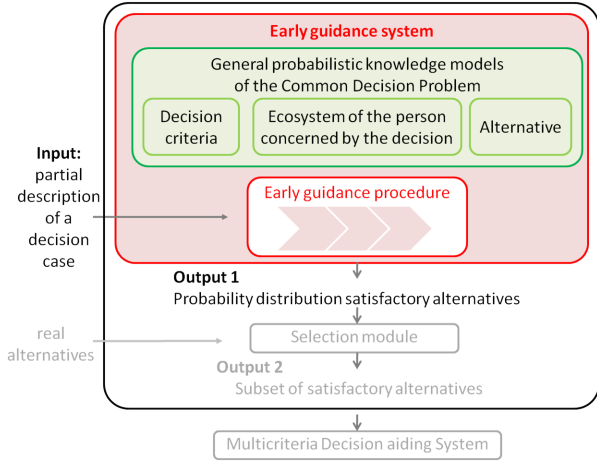


FIGURE 1 – General knowledge based aiding framework for a common decision problem.

The early guidance system is composed of a set of models of probabilistic knowledge, based on Bayesian networks. They concern decision criteria, the ecosystem of the person and the alternatives. Those models are linked together through the evaluation of the alternative regarding the criteria and the ecosystem of a decision case, and the associated satisfaction. Those models are built once for all, since they embed general probabilistic knowledge, which are relatively stable. Based on those models, the early guidance procedure is to be used in each new decision case, meaning in each situation where a decision has to be taken, according to its specific features. The only input is a partial description of the decision case, which can be given by the actors of the choice, or obtained thanks to automatic information collector. The output of the early guidance procedure is a probability distribution of satisfactory alternatives. This output can be exploited through different algorithms in order to get a subset of satisfactory alternatives. In the MCDA literature, this subset is the starting point of the decision process. In common decision problem, we claim that there is often a real need to provide early guidance to get such a subset.

In this article, we focus on the early guidance procedure.

3.2 The Bayesian network model(s) construction

The general knowledge model embedded in the aiding framework is constructed once for all. We propose to em-

bed general probabilistic knowledge about the alternatives (their characteristics) and the ecosystem of the person concerned by a common decision problem in two local Bayesian networks. Those features are next used in the local Bayesian network models that we propose for each criterion.

A *Bayesian network* \mathcal{B} consists in a set of random variables $\mathbf{X} = \{X_1, \dots, X_n\}$, a simple directed acyclic graph $\mathcal{G} = (\mathbf{X}, \mathbf{E})$, and a set of parameters Θ including the local probability distributions associated with each variable in \mathbf{X} . Together, \mathcal{G} and Θ define a probability distribution P over \mathbf{X} which factorizes as $P(\mathbf{X} = \mathbf{x}) = \prod_{X_i \in \mathbf{X}} P(X_i = x_i | pa(X_i))$, where $pa(X_i) \subset \mathbf{X}$ is the set of the parents (immediate predecessors) of X_i in the graph \mathcal{G} . Similarly, $ch(X)$ denotes the set of children of X in the graph. Let $Dom(X)$ represents the finite domain of definition of the variable X . Nodes X , Y , and Z form an *unshielded collider* in a DAG if X and Y have a common child Z and there is no arc between X and Y . In this paper, we consider a *causal* Bayesian network, meaning that each directed edge represents a direct causal influence between a cause and an effect.

We propose the construction of the following three Bayesian networks.

The **knowledge model of the alternatives** is a Bayesian network \mathcal{B}_{Alt} whose nodes are the characteristics \mathbf{X}_{Alt} of the alternatives. The graph structure and the local probability distributions embed pieces of knowledge about the alternatives, such as the (in)dependence between the attributes, internal constraints between some features, the state of the market or the actual distribution of some characteristics regarding some parameters. The set of alternatives \mathcal{A} is composed of any configuration of the variables of \mathbf{X}_{Alt} having a non zero prior probability in the Bayesian network $\mathcal{B}_{Alt} : \mathcal{A} = \{\mathbf{a} \in \times_{X \in \mathbf{X}_{Alt}} P(\mathbf{a}) > 0\}$.

The **knowledge model of the ecosystem of the person** concerned by a common decision problem is a Bayesian network \mathcal{B}_{eco} regrouping all the parameters that may influence the decision process, except the characteristics of the alternatives. The set of nodes is \mathbf{X}_{eco} and the graph reflects statistic and expert knowledge about those variables. A decision case \mathbf{dc} is a partial instantiation of the ecosystem, $\mathbf{X}_{DC} = \mathbf{dc}$, where $\mathbf{X}_{DC} \subset \mathbf{X}_{eco}$.

We also propose a **knowledge model for each decision criterion**. In the context of common decision problem, a *criterion* is a way to evaluate an alternative according to a decision case and a particular significance axis. A criterion given by the triplet (Imp_g, Ind_g, S_g) is represented in a causal Bayesian network by an unshielded collider as shown in Figure 2. The parents of the unshielded collider are the importance index Imp_g and the evaluation index Ind_g whose ancestors are some characteristics of the alternative in \mathbf{X}_{Alt} and possibly some characteristics of the decision case in \mathbf{X}_{eco} . the node S_g associated with the level of satisfaction of the criterion g has exactly two parents since the para-

eters involved in the evaluation of an alternative for the criterion are ancestors of the node Ind_g .

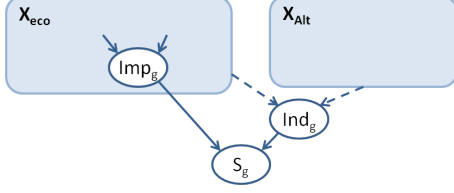


FIGURE 2 – Graph structure of the local Bayesian network of a decision criterion. Dotted arrows represent sets of arrows from any node of the set \mathbf{X}_{Alt} or \mathbf{X}_{eco} to the node Ind_g .

The modeling of a decision criterion by an unshielded collider in a Bayesian network can be found for example in [5, 14] but it was not formalized.

The complete graph of the **Bayesian network of a common decision problem** is obtained by combining the local graphs of each criterion with the graphs of \mathcal{B}_{Alt} and \mathcal{B}_{eco} . Combining those graphs consists in merging the nodes that represent the same parameter when several criteria depend on it. This operation does not bring to add a new parent for any node, and thus no conditional probability tables have to be modified.

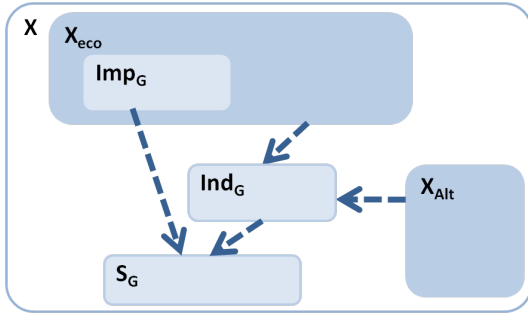


FIGURE 3 – The Bayesian network graph of a common decision problem. Bold dotted arrows represent sets of arrows between any node of the connected sets of nodes.

Figure 3 shows the general structure of the graph of the Bayesian network associated with a common decision problem. The set of nodes is $\mathbf{X} = \mathbf{X}_{eco} \cup \mathbf{X}_{Alt} \cup \mathbf{S}_G \cup \mathbf{Ind}_G$, where :

$\mathbf{S}_G = \{S_g, g \in G\}$ is the set of variables associated with the level of satisfaction of each criterion.

$\mathbf{Imp}_G = \{Imp_g, g \in G\}$ is the set of variables associated with the level of importance of each criterion,

$\mathbf{Ind}_G = \{Ind_g, g \in G\} \cup \mathbf{I}$ is the set of variables associated with the criterion evaluation index of each criterion, augmented with the set \mathbf{I} of additional intermediate variables defined to evaluate a criterion.

\mathbf{X}_{eco} is redefined as the set of the characteristics of the ecosystem, augmented with the set \mathbf{Imp}_G , since the level of importance of a criterion is a derived characteristic of the ecosystem.

Once defined, this Bayesian network can be used for early guidance in any new decision case.

4 A Bayesian network based early guidance system

When a new decision case occurs in a common decision problem, it corresponds to a partial instantiation of the ecosystem of the considered situation, because the situation is only partially observed. We thus have a subset $\mathbf{X}_{DC} = \mathbf{dc}$, with $\mathbf{X}_{DC} \subset \mathbf{X}_{eco}$. Remark that the composition of the subset \mathbf{X}_{DC} differs from one situation to another. The uncertainties about the ecosystem can be reduced by updating the unobserved elements of \mathbf{X}_{eco} given \mathbf{dc} in the Bayesian network \mathcal{B}_{eco} , meaning computing $P(\mathbf{X}_{eco} | \mathbf{dc})$.

The objective of early guidance lead to focus on the characteristics of the alternative for a specific decision case, guided by the desire of a good level of satisfaction of criteria. It's about fixing the ecosystem during the reasoning about the characteristics of the alternative and criteria. We thus make an action that consists in setting the ecosystem as a whole. Indeed, without that action, the reasoning could lead to some belief changes on unobserved variables of the decision case, which is not relevant.

Such an action corresponds to the concept of intervention in a Bayesian network, with the difference that we consider a set of variables, and we do not have a complete instantiation of that set for the intervention. We thus propose definitions for a *probabilistic intervention* on a variable $do(R(X))$ and for a set of variables $do(R(\mathbf{X}))$, that extend the definition of an intervention $do(X = x)$.

In the next section, we first explain the principle of *satisfactory situation* that guides the early guidance procedure. Then we define probabilistic interventions and we propose the early guidance procedure based on these two points.

4.1 The satisfactory situation

Given a common decision problem and a decision case, we first define a *satisfactory alternative* in an informal way as an alternative providing a *sufficient* level of satisfaction regarding each criterion, meaning that we have $S_g \geq s$ for each criterion $g \in G$, where $s \in [0, 1]$ represents the minimum level of satisfaction to be satisfactory. Remark that we consider the level of satisfaction and not the quality index of the criterion, which make relevant to consider the

same value s for all the criterion. Following this, we define a *satisfactory situation* as a situation where a satisfactory alternative has been found.

In order to find a subset of *satisfactory alternatives* for a given decision case we propose the following two steps :

1. Imagine a satisfactory situation in which the person concerned by the consequence of a decision has found a satisfactory alternative ;
2. Compute the probability distribution of the alternative that *explain* that satisfactory situation in the considered decision case.

Remark that we are not interested in the alternative that best explains the satisfactory situation since we do not look for a single alternative, but we look for a subset of satisfactory alternatives.

As introduced above, this result cannot be obtained by a simple inference (updating) in the Bayesian network \mathcal{B}_{eco} because of the unobserved variables in \mathbf{X}_{eco} .

We have to compute $P(\mathbf{X}_{Alt} | \mathbf{S}_G \geq s, do(R(\mathbf{X}_{eco})))$, where $\mathbf{S}_G \geq s$ represents the observation of a satisfactory situation : $S_g \geq s$, for each $g \in G$ and $do(R(\mathbf{X}_{eco}))$ is a probabilistic intervention that set the ecosystem with the considered decision case. Thanks to this approach, we obtain a probability distribution of the alternatives that *explains* the satisfaction.

With that objective, we know introduce Probabilistic intervention.

4.2 Probabilistic intervention

An intervention $do(X = x)$ in a causal Bayesian network [18] consists in making an action on the variable X that makes the variable being in the state x . An intervention (or manipulation) forces a variable into a certain value or state. As a consequence, incoming arrows have to be removed, since we remove the direct influences by intervening.

Intuitively, a probabilistic intervention $do(R(X = x))$ on a variable X in a causal Bayesian network consists in an action on X that makes $R(X)$ stand fix, such that further reasoning takes place in a restricted world where $R(X)$ stands. This means that the belief on X must not be modified by future observations on other nodes of the Bayesian network. In order to integrate the influence of $R(X)$, the Bayesian network is temporarily modified.

Let consider a causal Bayesian network $\mathcal{B} = (\mathcal{G}, P)$, with $\mathcal{G} = (\mathbf{X}, \mathcal{E})$.

A *probabilistic intervention* $do(R(X))$ on a variable $X \in \mathbf{X}$ in is an action that integrate $R(X)$ in the probability distribution of a modified Bayesian network $\mathcal{B}' = (\mathcal{G}', P')$ obtained as follows :

1. Remove the node X and the associated arcs,
2. Define the graph \mathcal{G}' as the subgraph of \mathcal{G} induced by the set $\mathbf{X}' \subset \mathbf{X} \setminus \{X\}$ including the children of X and all connected nodes.

3. The probability distribution P' on \mathbf{X}' is defined by :

$$P'(\mathbf{X}') = \prod_{Y \notin ch(X)} P(Y | pa(Y)) \prod_{Z \in ch(X)} P'(Z | pa_{\mathcal{G}'}(Z))$$

where $P'(Z | pa_{\mathcal{G}'}(Z))$ is defined as follows, supposing without loss of generality that Y is the only parent of Z in \mathcal{G}' , meaning that X and Y were common parents of Z in \mathcal{G} (see Figure 4) :

$$P'(Z | Y) = \sum_x P(Z | X = x, Y)R(X = x) \quad (1)$$

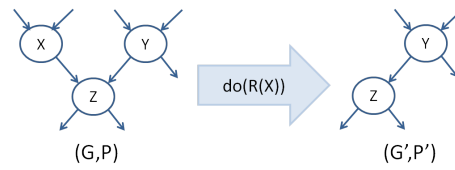


FIGURE 4 – Configuration of the graph around a node X before and after a probabilistic intervention $do(R(X))$.

Figure 4 shows the modification of the graph of the Bayesian network after the intervention $do(R(X))$. The children of X in the initial graph \mathcal{G} are the only nodes whose local conditional probability distribution is modified by the absorption of the node X in \mathcal{G}' .

Remark that when the probability distribution $R(X)$ is a dirac on the value x , the definition of $do(R(X))$ and $do(X = x)$ are the same.

We extend the definition of probabilistic intervention for a subset of variables in the configuration of the graph described on the left part of in Figure 5. We consider a Bayesian network $\mathcal{B} = (\mathcal{G}, P)$ whose set of nodes is $\mathbf{X} \cup \mathbf{Z}$, with $\mathbf{X} \cap \mathbf{Z} = \emptyset$ and there is no arrow from a node in \mathbf{Z} toward a node in \mathbf{X} .



FIGURE 5 – Modification of the graph of a Bayesian network after a probabilistic intervention on a subset of variables. The bold dotted arrow represents the set of arrows from any node of the set \mathbf{X} to any node of the set \mathbf{Z} .

A *probabilistic intervention* $do(R(\mathbf{X}))$ on a set \mathbf{X} is an action that integrates $R(\mathbf{X})$ in the probability distribution P'

of the modified Bayesian network $\mathcal{B}' = (\mathcal{G}', P')$ obtained as follows :

1. Define the graph \mathcal{G}' as the subgraph of \mathcal{G} induced by the set \mathbf{Z} ;
2. Define the subset $\mathbf{X}_{\text{out}} \subset \mathbf{X}$ composed of the nodes of \mathbf{X} with an outgoing arc toward \mathbf{Z} . By construction the children of a node in \mathbf{X}_{out} are in $\mathbf{Z} \cup \mathbf{X}_{\text{out}}$.
3. Define the probability distribution P' on \mathbf{Z} :

$$P'(\mathbf{Z}) = \prod_{Y \notin \text{ch}(\mathbf{X}_{\text{out}})} P(Y | \text{pa}(Y)) \prod_{Z \in \text{ch}(\mathbf{X}_{\text{out}})} P'(Z | \text{pa}_{\mathcal{G}'}(Z))$$

where $P'(Z | \text{pa}_{\mathcal{G}'}(Z))$ is defined below ; for that, let $\mathbf{X}_{\text{pa}(Z)}$ be the set of parents of a node $Z \in \text{ch}(\mathbf{X}_{\text{out}})$ in \mathbf{X}_{out} , meaning that $\mathbf{X}_{\text{pa}(Z)} = \text{pa}(Z) \cap \mathbf{X}_{\text{out}}$. We suppose without loss of generality that Y is the only parent of Z in \mathcal{G}' , meaning that Y and $\mathbf{X}_{\text{pa}(Z)}$ were common parents of Z in \mathcal{G} (see Figure 6). Remark that it doesn't matter whether $Y \in \text{ch}(\mathbf{X}_{\text{out}})$ or $Y \notin \text{ch}(\mathbf{X}_{\text{out}})$, since we compute the new conditional probability distribution of Z to incorporate in it the influence of $R(\mathbf{X})$:

$$P'(Z | Y) = \sum_{\mathbf{x}} P(Z | \mathbf{X}_{\text{pa}(Z)} = \mathbf{x}, Y) R(\mathbf{X}_{\text{pa}(Z)} = \mathbf{x}) \quad (2)$$

The probability $R(\mathbf{X}_{\text{pa}(Z)} = \mathbf{x})$ is a marginal from $R(\mathbf{X})$.

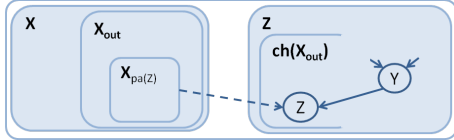


FIGURE 6 – Detail of the parents of a node Z whose probability is re defined in the modified Bayesian network after the probabilistic intervention on a subset of variables \mathbf{X} .

A direct consequence of this definition is that there is no distinction between the probabilistic interventions $do(R(\mathbf{X}))$ and $do(R(\mathbf{X}_{\text{out}}))$; they produce exactly the same Bayesian network \mathcal{B}' .

Remark also that when the set \mathbf{X} is a singleton, the definition of the probabilistic intervention $do(R(\mathbf{X}))$ and $do(R(X))$ are the same.

We can now present the early guidance procedure.

4.3 The early guidance procedure

The early guidance procedure follows the idea of the *satisfactory situation* and it is based on the Bayesian network

of the common decision problem. The input of the procedure is a partial description of the ecosystem of the new decision case. The output is the posterior probability distribution of satisfactory alternatives. It can be later used in different ways to provide a list of relevant alternatives.

The three main steps of the early guidance procedure are presented in Figure 7.

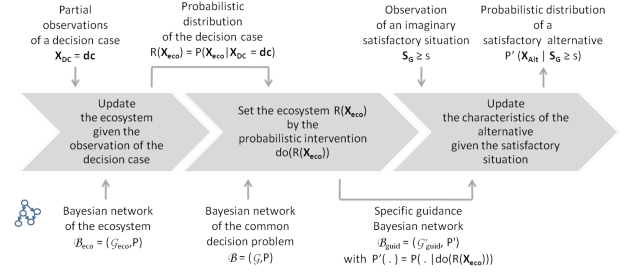


FIGURE 7 – The early guidance procedure.

Here is the description of the early guidance procedure :

Input :

$\mathbf{X}_{\text{DC}} = \mathbf{dc}$, with $\mathbf{X}_{\text{DC}} \subset \mathbf{X}_{\text{eco}}$ a partial description of a decision case,

\mathcal{B} , the Bayesian network of the common decision problem,

s , minimum level of satisfaction for the criteria.

Output : $P(\mathbf{X}_{\text{Alt}} | \mathbf{S}_{\text{G}} \geq s, do(R(\mathbf{X}_{\text{eco}})))$ probability distribution of the attributes of the satisfactory alternatives for the given decision case.

step 1 - Compute $R(\mathbf{X}_{\text{eco}}) = P(\mathbf{X}_{\text{eco}} | \mathbf{X}_{\text{DC}} = \mathbf{dc})$: use the Bayesian network of the ecosystem \mathcal{B}_{eco} to update the belief about the unobserved variables of the ecosystem from a partial set of observations $\mathbf{X}_{\text{DC}} = \mathbf{dc}$ of the decision case.

step 2 - Make the probabilistic intervention $do(R(\mathbf{X}_{\text{eco}}))$ in the Bayesian network \mathcal{B} to set the ecosystem with $R(\mathbf{X}_{\text{eco}})$. The result is a new Bayesian network $\mathcal{B}' = (\mathcal{G}', P')$ to be used for early guidance in that specific decision case. Note that the condition to make a probabilistic intervention on the set \mathbf{X}_{eco} is fulfilled since there is no arcs from the rest of the graph towards \mathbf{X}_{eco} (see the left part of Figure 8); the Figure shows the graph before and after the probabilistic intervention. The set of nodes of \mathcal{B}' is $\mathbf{X}' = \mathbf{S}_{\text{G}} \cup \mathbf{Ind}_{\text{G}} \cup \mathbf{X}_{\text{Alt}}$.

The new probability distribution P' on \mathbf{X}' incorporates $R(\mathbf{X}_{\text{eco}})$. It is defined following equation (3), by redefining the local conditional probability distributions of nodes of \mathbf{S}_{G} and \mathbf{Ind}_{G} .

step 3 - Compute $P'(\mathbf{X}_{\text{Alt}} | \mathbf{S}_{\text{G}} \geq s)$ to propagate an imaginary satisfactory situation $\mathbf{S}_{\text{G}} \geq s$ in the Bayesian network \mathcal{B}' .

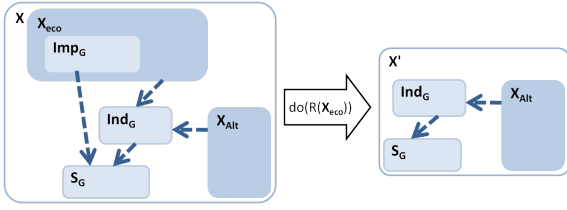


FIGURE 8 – The graph resulting of the probabilistic intervention $do(R(\mathbf{X}_{eco}))$ in the Bayesian network \mathcal{B} .

This procedure requires inferences in steps 1 and 3, to compute joint probability distributions on the sets \mathbf{X}_{eco} and \mathbf{X}_{Alt} . Since each of these set corresponds to a local Bayesian network, it is sufficient to compute and store the set of local conditional probability distributions associated to each node given its parents, following the partial order induced by the graph. By this way, the required size in memory is decreased. It only depends on the number of parents of the nodes and the size of the domains.

Moreover, thanks to the remark above about equivalent probabilistic interventions, Step 1 can be made easier by considering only the nodes of \mathbf{X}_{eco} that are parents of a node in \mathbf{X}' . Let $\mathbf{X}_{pa(\mathbf{X}')} \subset \mathbf{X}_{eco}$ be this set : $\mathbf{X}_{pa(\mathbf{X}')} = \text{Img}_G \cup (pa(\text{Ind}_G) \cap \mathbf{X}_{eco})$. The probabilistic interventions $do(R(\mathbf{X}_{eco}))$ and $do(R(\mathbf{X}_{pa(\mathbf{X}')}))$ are equivalent. So, in Step 1, it is sufficient to compute $R(\mathbf{X}_{pa(\mathbf{X}')})$, where $\mathbf{X}_{pa(\mathbf{X}'})$ is composed of the nodes of importance of the criteria and the variables of the ecosystem involved in the evaluation of the quality of an alternative regarding a criterion.

The result of this procedure is the posterior probability distribution of satisfactory alternatives for a decision case. As shown in Figure 1, this result is a first step to make easier the selection of a subset of satisfactory alternatives. With this aim, different algorithms could be proposed depending on the common decision problem and the context of use of the aiding framework. For example, in the choice of a manual wheelchair [21], an interesting objective is to propose a very short list of three real wheelchairs such that all of them are relevant for the considered person and his/her situation, and such that each wheelchair is as different as possible from the other two. Each of the three proposal could then be evaluated in detail according to the most important criteria for this decision case, which would allow the person to understand the advantages and the relative drawback of each wheelchair. For the final selection of a subset of alternatives, several elements can be very different depending on the objective : what is the desired size of the final set of alternatives? Do we desire all the most satisfactory alternatives, or just a panel? Can we have a subset of characteristics, and / or an ordered subset of characteristics? Is it possible (desirable) to involve the actors

of the choice in this step? and what kind of information could be waited? Finally, the complexity of this final procedure has also to be considered. In this article, we provide a general framework that can suit different common decision problems, and different kind of applications. We let for a future work the selection of a subset of alternatives from the posterior probability distribution that results from the early guidance procedure proposed in this paper.

5 Related works and conclusion

This last part is dedicated to a brief comparison of probabilistic intervention with another proposal that also aim to fix a local conditional probability distribution in a Bayesian network.

5.1 A brief comparison with soft evidence and AEBN

The idea to keep fixed a local probability distribution in a Bayesian network has been approached through *Soft evidence* [1, 22, 15]. It refers to a local probability distribution that defines a constraint on the belief after this information has been propagated; it describes the state of beliefs “all things considered” and can not be modified by the propagation of other findings. Actually, soft evidence may correspond to the second meaning of the term “uncertain input” in the probabilistic framework [4] : *The input is a partial description of a probability measure; the uncertainty is part of the input and is taken as a constraint on the final cognitive state. The input is then a correction to the prior cognitive state.* It must be clearly distinguished from the cases where the uncertainty bears on the meaning of the input, which is usually referred as *virtual evidence* or *likelihood evidence* [17] in the context of Bayesian networks. Soft evidence were first introduced in the context of Agent Encapsulated Bayesian networks (AEBN) [1, 22, 13, 11]. In an agent based model using AEBNs, the belief of a receiver agent is updated following the transmission of a soft evidence sent from a publisher agent. Several algorithms have been proposed to manage soft evidence in Bayesian networks. Most of them are based on the Iterative Proportional Fitting Procedure (IPFP), adapted to the factorized form of the joint probability distribution to be used with Bayesian networks [19, 22, 10, 16, 12], including cases with inconsistencies [23]. However, those algorithms are rarely proposed in Bayesian network software tools¹. This kind of algorithm has been used in Markov networks in a complex application for item planning and capacity management in the automotive industry [6].

Even if both probabilistic intervention and the revision of a probability distribution leads to define a new proba-

1. In 2014, Bayesialab was the only Bayesian network software to propose fixing the probabilities of some selected nodes to the current marginal probabilities [15].

bility distribution, they are deeply different. First because an intervention is an action performed on a *causal* model, and it leads to modify the graph since the action performed move aside any other cause of the value resulting from the action. At the opposite, the revision of a joint probability distribution by a local one has nothing to do with causality, and it keeps unchanged the structure of the graph.

Another point of comparison is worth mentioning : the propagation of several pieces of evidence including at least one piece of soft evidence is not commutative. As a consequence, as long as a *soft evidence* stands, each additional piece of evidence requires to apply again the algorithm to maintain the constraint. Using a probabilistic intervention leads to define temporarily a new Bayesian network in which several pieces of evidence can be easily taken into account.

5.2 Conclusion and perspectives

In this article, we define what is a common decision problem, and introduce the concept of ecosystem of the person concerned by the decision as an explicit set of parameters to be considered. we provide a brief comparison between the early guidance problem and multi-criteria decision analysis. We explain the need of early guidance to facilitate the selection of a subset of satisfactory alternatives, before the decision process really starts, and without requiring any specific knowledge from the person regarding the problem. We propose an early guidance system that embeds general probabilistic knowledge about a common decision problem and redistributes it appropriately in each decision case as relevant guidance. Our proposition helps the actors of a common decision problem and takes in input only a partial description of their decision case (who, what, where, when, why, how important, etc.).

We propose a new definition of a decision criterion that gathers the importance of the criterion according to the ecosystem of the person, the evaluation of the alternative according to ecosystem, and the satisfaction associated with an alternative in a given decision case. We explain the construction of the Bayesian network for a common decision problem, embedding knowledge about the alternative, the ecosystem and using an unshielded collider in the graph for each decision criterion. We introduce the principle of the satisfactory situation as our starting point in the reasoning towards satisfactory alternatives. In that reasoning, we explain the need of a probabilistic intervention in the Bayesian network to set the ecosystem even though it is partially observed. We extend the concept of intervention proposed by Pearl to the concept of probabilistic intervention. for a single variable, and for a set of variables in a specific configuration of the graph. Finally, we propose the early guidance procedure based on the knowledge model and using probabilistic intervention. The output of the

early guidance system is a probability distribution of satisfactory alternatives that is a basis for the selection of a subset of relevant alternatives.

In future works, we plan to complete this proposal with an additional algorithm to provide a subset of satisfactory alternatives based on the posterior probability distribution resulting from the early guidance procedure. With that aim, we will consider a specific common decision problem and a precisely defined context of use of the aiding framework. We also plan to implement and test probabilistic intervention in pyAgrum²[7], and to implement the early guidance procedure on a specific common decision problem as the one presented in [3].

6 Acknowledgments

This research has been supported by the project EL-SAT 2020, the International Campus on Safety and Intermodality in Transportation; the Hauts-de-France Region; the European Community; the Regional Delegation for Research and Technology; the Ministry of Higher Education and Research; and the National Center for Scientific Research. The authors gratefully acknowledge the support of these institutions.

7 References

Références

- [1] Bloemeke, Mark: *Agent encapsulated Bayesian networks*. Ph.D. thesis, Department of Computer Science, University of South Carolina, 1998.
- [2] Bouyssou, Denis: *Readings in Multiple Criteria Decision-Aid*, chapitre Building criteria : a prerequisite for MCDA, pages 58–80. Springer Verlag, Heidelberg, 1990.
- [3] Delcroix, Véronique: *Bayesian Network Model of a Criterion in a Repeated Multi-Criteria Decision Problem*. Dans *Actes des Journées de l'Intelligence Artificielle Fondamentale (JIAF)*, Caen, France, july 2017.
- [4] Dubois, Didier, Serafin Moral et Henri Prade: *Belief change rules in ordinal and numerical uncertainty theories*. Dans Gabbay, D.M. et Ph. Smets (éditeurs) : *Belief Change, (D. Dubois, H. Prade, eds.), Vol. 3 of the Handbook of Defeasible Reasoning and Uncertainty Management Systems*, pages 311–392. Kluwer Academic Publishers, Dordrecht, 1998.
- [5] Fenton, Norman et Martin Neil: *Making Decisions : Bayesian Nets and MCDA*. *Knowledge-Based Systems*, 14(7) :307–325, 2001.

2. <http://agrum.gitlab.io/>

- [6] Gebhardt, J., C. Borgelt et R. Kruse: *Knowledge Revision in Markov Networks*. Mathware & Soft Computing, 11(1) :93–107, 2004.
- [7] Gonzales, Christophe, Lionel Torti et Pierre-Henri Wuillemin: *aGrUM : A Graphical Universal Model Framework*. Dans *Advances in Artificial Intelligence : From Theory to Practice - 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE, Proceedings, Part II*, pages 171–177, Arras, France, June 2017.
- [8] Hobbs, Benjamin F. et Peter Meier: *Energy Decisions and the Environment A Guide to the Use of Multicriteria Methods*. Springer US, 2000, ISBN 978-0-7923-7875-4.
- [9] Ishizaka, A. et P. Nemery: *Multi-criteria Decision Analysis : Methods and Software*. Wiley, 2013, ISBN 9781118644928.
- [10] Kim, Young Gyun, Marco Valtorta et Jirí Vomlel: *A Prototypical System for Soft Evidential Update*. Applied Intelligence, 21(1) :81–97, 2004.
- [11] Langevin, Scott: *Knowledge representation, communication, and update in probability-based multiagent systems*. Thèse de doctorat, University of South Carolina, Columbia, SC, USA, 2011, ISBN 978-1-124-64713-5. AAI3454755.
- [12] Langevin, Scott et Marco Valtorta: *Performance Evaluation of Algorithms for Soft Evidential Update in Bayesian Networks : First Results*. Dans *SUM*, pages 284–297, 2008.
- [13] Langevin, Scott, Marco Valtorta et Mark Bloemke: *Agent-encapsulated Bayesian networks and the rumor problem*. Dans *AAMAS '10 Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, tome 1, pages 1553–1554, 2010.
- [14] Leicester, Philip A., Chris I. Goodier et Paul Rowley: *Using a Bayesian Network to evaluate the social, economic and environmental impacts of community renewable energy*. Dans *Clean Technology for Smart Cities and Buildings (CISBAT)*, 2013.
- [15] Mrad, Ali Ben, Véronique Delcroix, Sylvain Piechowiak, Philip Leicester et Mohamed Abid: *An explanation of uncertain evidence in Bayesian networks : likelihood evidence and probabilistic evidence - Uncertain evidence in Bayesian networks*. Appl. Intell., 43(4) :802–824, 2015.
- [16] Pan, Rong, Yun Peng et Zhongli Ding: *Belief Update in Bayesian Networks Using Uncertain Evidence*. Dans *ICTAI*, pages 441–444, 2006.
- [17] Pearl, Judea: *Probabilistic reasoning in intelligent systems : networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988, ISBN 0-934613-73-7.
- [18] Pearl, Judea: *Causality : Models, Reasoning, and Inference*. University Press, New York, USA, 2000, Second ed., 2009.
- [19] Peng, Y., S. Zhang et R. Pan: *Bayesian Network Reasoning with Uncertain Evidences*. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 18(5) :539–564, 2010.
- [20] Roy, B.: *Multicriteria Methodology for Decision Aiding*. Kluwer Academic, Dordrecht, 1996.
- [21] Sedki, K., V. Delcroix, F. X. Lepoutre, E. Adam, A. P. Maquinghen-Godillon et I. Ville: *Bayesian network model for decision problems*. Dans Klopotek, M.a., M. Marciniak, A. Mykowiecka, W. Penczek et S.t. Wierzchon (Ed.) (rédacteurs) : *Intelligent Information Systems, new approaches*, pages 285–298, Publishing House of University of Podlasie, Siedlce, Poland, June 2010. ISBN 978-83-7051-580-5.
- [22] Valtorta, Marco, Young Gyun Kim et Jirí Vomlel: *Soft evidential update for probabilistic multiagent systems*. International Journal of Approximate Reasoning, 29(1) :71–106, 2002.
- [23] Vomlel, Jirí: *Integrating Inconsistent Data in a Probabilistic Model*. Journal of Applied Non-Classical Logics, 14(3) :367–386, 2004.