HAL open science

A knowledge based system for the management of a time stamped uncertain observation set with application on preserving mobility

Véronique Delcroix, Emmanuelle Grislin-Le Strugeon, François Puisieux

▶ **To cite this version:**

Véronique Delcroix, Emmanuelle Grislin-Le Strugeon, François Puisieux. A knowledge based system for the management of a time stamped uncertain observation set with application on preserving mobility. International Journal of Approximate Reasoning, 2021, 134, pp.53-71. 10.1016/j.ijar.2021.04.003 . hal-03396161

HAL Id: hal-03396161

https://uphf.hal.science/hal-03396161v1

Submitted on 9 May 2023

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.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S0888613X21000475
Manuscript_201caa60ebb4e984bae91ebf01432996

A knowledge based system for the management of a time stamped uncertain observation set with application on preserving mobility

Véronique Delcroix, Emmanuelle Grislin-Le Strugeon

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

{vdelcroi,egrislin}@uphf.fr

François Puisieux

Pôle gérontologie, Hôpital gériatrique les Bateliers, CHU de LILLE, Lille, France
Francois.PUISIEUX@chru-lille.fr

Abstract

The aim of this study is to maintain up-to-date information about the current state of elderly people that are medically followed for risks of fall. Our proposal consists of an individual information database management system that can provide information on-demand on various variables. Such a system has to deal with several sources of uncertainty: lack of information, evolving information and reliability of the information sources. We consider that the features of the person may evolve with time causing uncertainty due to obsolete information. Our context includes new information received bit by bit, with no possibility to collect all required information at once. This paper establishes a first proposal to manage a set of uncertain observations, in order to reduce erroneous and obsolete information while keeping the benefit of previously collected information. We propose an architecture of the system based on a probabilistic knowledge model about the characteristics of interest, a set of decay functions that help to evaluate the confidence degree in previous observations, and a reasoning module to manage new observations, maintain the compatibility and the quality of the observation set. We detail the algorithms of the reasoning module, and the algorithm to update the confidence degree of the observations.

Preprint submitted to International Journal of Approximate Reasoning *April 15, 2021*

© 2021 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

Keywords: information aging, uncertain information, information quality, probabilistic model, knowledge based system, reasoning with uncertainty, health application

1. Introduction

The temporal dimension is naturally present in medical reasoning, since situations change, events happen and most properties vary with time [1]. For these reasons, the temporal dimension is a source of uncertainty in elderly persons fall prevention. In this domain, recommendations are made by specialists on the basis of information related to the health and environment of the person [2, 3]. The reasoning to identify adequate recommendations is based on a detailed "snap-shot" of the person and her situation, such as loss of muscle mass, bone fragility, sedentary life style, fear of falling, previous falls, living alone, etc. These are high level static variables that summarize useful temporal patterns. However, it is generally not realistic to collect all this information just at the time it is required. The solution is to use information collected in the past. But this information may have become obsolete, because of aging or changes in the life of the person. We thus have a situation where information aging and incomplete observations generate uncertainty. We also have to deal with uncertainty due to the natural variability of human characteristics, and incomplete information and knowledge. Finally, an additional source of uncertainty comes from the information sources, since a value can be misread, or a dysfunctioning sensor can provide erroneous data.

The problem is thus to maintain an uncertain observation set with an incomplete data collection over time, in order to provide on-demand reliable information regarding the person's state. This requires to evaluate whether the value (or state) of a variable has changed since it was observed. In this problem, time is involved together with uncertainty since only few properties are time-invariant, and because of incomplete and uncertain observation over time and information aging.

In order to illustrate the problem of dealing with an uncertain observation set, consider the following two pieces of information about a given person: when we know that she lives alone, we set $A = a$, else $A = \bar{a}$; when we know that she lives in a retirement home, we set $B = b$, else $B = \bar{b}$. Since living in a retirement home implies not to live alone, the configuration $(A = a, B = b)$ is impossible. This is translated in a general probabilistic model by $P(A = a, B = b) = 0$, also denoted by $P(a, b) = 0$. Now let's consider two uncertain observations about that person: the first uncertain observation o_A is given by the likelihood vector $\pi_A = (0.9, 0.1)$, stating that it is nine times more likely to observe o_A when $A = a$ than when $A = \bar{a}$. The second observation o_B is given by the likelihood vector $\pi_B = (0.2, 0.8)$. Since these observations are collected separately, they do not provide information of the likelihood of the four possible combinations of A and B. From the observation set $\{o_A, o_B\}$, the combination (a, b) cannot be excluded whereas it is excluded in the general probability distribution P. This example illustrates the question of the *compatibility* of an uncertain observation set with a general probabilistic model of knowledge. In this article, we propose the notions of *total* and *partial compatibility* in order to model different kinds of possible situations regarding this question. The objective is to maintain the compatibility in a knowledge based system for the management of an observation set when integrating new and uncertain observations.

This problem lies at the intersection of the domains of uncertain information fusion [4], dynamic, iterated, belief revision [5, 6] and temporal reasoning [7, 1]. The problem is to aggregate observations made at different times with static generic knowledge. Three types of belief change are generally distinguished to address the problem of adding new information to a logical theory [8, 6]: revision, update and arbitration. Our problem context relates to these questions, with additional dimensions of uncertainty, needing other approaches than logical theory.

Information aging has recently acquired a specific interest in the domain of web information. In this domain, the way events and information spread in social networks and news pages is particularly studied. For instance, *aging*

theory is used to model the variations in the quantity of publications related to a topic or an event [9, 10, 11]. To address this problem that shows similarities with ours, the temporal decreases are modeled by decay functions. Decay functions are also used in the domain of sensor data to manage the aggregation of data with different ages: the older is the data, the lighter is its weight in the aggregated data summary [12].

Temporal reasoning is addressed by several formal theories of time, including temporal logics, temporal constraints and temporal reasoning techniques [7]. About temporal reasoning with uncertainty, a well-known model is the qualitative algebra of time, based on intervals, proposed by Allen [13, 14]. This model allows to represent imprecise temporal events and is suitable when the variables have a specific duration, but it does not allow to express uncertainty in events or their temporal relationships [15]. In our context, the variables related to the state of a person can keep stable for years and progressively or suddenly change (e.g., *leg strength weakness*, *goes out*, *drives* or *alcohol* variables)

An interesting alternative is proposed by the probabilistic graphical models, which offer diverse solutions to the problem of temporal reasoning with uncertainty. These models have proven their utility and feasibility in the medical domain, see the detailed presentation of temporal extensions of Bayesian networks used in clinical domain in [16]. Dynamic Bayesian networks (DBNs) [17] are temporal probabilistic graphical models that model complex multivariate time series or sequences. However, DBNs are not adapted to our problem since the variables that we consider evolve according to different time scales, making hard the choice of a temporal granularity. Using a small discrete time step would lead to a DBNs with a huge number of slices, making the model intractable. In Continuous Time Bayesian Networks (CTBN) [18], the states of variables evolve continuously over time, and the evolution of each variable depends on the state of its parents in the graph. Although possibly well-adapted to our problem, CTBNs construction requires a level of resource that would be too high regarding the size of the targeted model (this is discussed in the last section of the paper). Irregular-Time Bayesian Networks (ITBN) [19] are another extension

of Bayesian Networks that includes a representation of time. ITBNs generalize Dynamic BNs to model irregularities in the time intervals between observations. They allow to remove the constraint of a constant time distance between consecutive observations that is required by the Dynamic BNs. Like CTBNs, ITBNs can represent variables with different time slices. Unlike CTBNs, ITBNs are well adapted to model variables having a continuous state space [20]. However, in the current status of our works, we consider only discrete state spaces.

Finally, temporal Bayesian classifiers have been introduced for medical applications [21, 22]. They augment the Bayesian network classifiers with a second root node, whose state is associated with discrete time points. Their inabilities to consider time with different granularities make them inappropriate for our problem.

Since none of these models is fully adequate for our problem, we propose the use of additional mechanisms to maintain a time-stamped uncertain observation set based on a static Bayesian network.

The next section describes the real case context and states the problem, followed by the proposal of an architecture and some specific operations to maintain a set of uncertain observations with an illustrative example. We discuss the possibility of using continuous time Bayesian networks in the last section of this paper, together with other perspectives.

2. The uncertain observation set management problem

In that section we first present the application-related objective and the main research question, then we explain the need to find a compromise between quality and quantity in the observation set. We propose some definitions and the problem statement.

2.1. Application-related objective

Preserving the mobility of an elderly person requires to gather specific information about this person in order to provide adequate recommendations. We

observed and analyzed the functioning of the service of the hospital center of Lille specialized in fall prevention. Each patient is served during a complete day by a medical team that collects a set of specific individual pieces of information. **This information is focused on the evaluation of risk factors for fall that regroup incidence factors, related with fall causes, and severity factors, related with fall gravity. It concerns in particular environmental factors, behavioral factors, precipitating and predisposing factors such as medication, diseases, physical and mental state, and the number of falls during the past six months.** On this basis, a group of experts identifies a small number of adequate recommendations to reduce the risk of fall. These recommendations have to be regularly re-evaluated and must evolve together with the person's situation (abilities, environment, context, etc.). However, each person can not frequently benefits of such a complete consultation since it requires plenty of time, the presence of specialists and specific material to achieve examination. The follow-up is generally made by general practitioners, by whom only little information can be collected at each appointment, making impossible to obtain a complete up-to-date set of information. As a consequence, the reasoning has to be conducted on the basis of incomplete information.

The application-related objective is to develop a system able to provide reliable individual information related to the risk of fall, in the aim to assist practitioners or to be linked to other systems like the Assessment of Risk of Adverse Events proposed in [23] for instance. However, some information may be missing and there is uncertainty about previously collected information.

2.2. Main question and general approach

The main question can be formulated as: what information can be provided by the system on a requested variable when there is no available observation on that variable in the observation set, or when the available observation may be out of date, and/or when the

observation comes from an unsure source? In order to answer this question, information to consider comes both from specific observations realized on the patient and from generic knowledge to complete as much as possible the information that is provided.

***Generic knowledge* concerns the dependency relations between the variables, their variability in the considered population, and their evolution over time, partly provided by experts and partly learned from related data. The data file that we used for the construction of the general knowledge model consists of information about 435 variables for 1173 patients who attended the Lille University Hospital Falls Clinic between January 2005 and December 2014. We conducted a study of these data in collaboration with Lille hospital's experts on fall prevention. This study resulted in a taxonomy of the fall risk factors, represented by an ontology [24], and the selection of 45 relevant variables associated with the characteristics of the person, the main risk factors for fall and possible causes or consequences of them. From this result, we built and evaluated a first model of the generic knowledge, embedded in a Bayesian network [25].**

***Specific observations* concern one patient in particular and are made of collected data about this specific person. The observation set includes uncertain information about any subset of variables of the knowledge model. Among the 45 selected variables, some are frequently observed, meaning that corresponding information is frequently present in the observation set, whereas other variables are rarely observed. Furthermore, the proportion of observed variables is very different from one individual to another when considering the whole population of elderly.**

In this work, we consider that generic knowledge is well confirmed and we do not call it into question, even when some observations are in contradiction with it. Conversely, specific observations are more or less reliable, and when a set of observations is not compatible with the generic knowledge, some of them are

reconsidered. A first step is to maintain the observation set according to criteria of quality and quantity, which is the subject of that paper. A second step, that is beyond the scope of this paper, is to take benefit from the observation set and the generic knowledge model to provide information about a requested variable.

2.3. A compromise between quality and quantity

To provide answers, the system must include as much information as possible but also the most reliable information. This concerns the information quantity and quality.

Data quality, or information quality, is a complex concept which is characterized by multiple dimensions such as accuracy, availability, completeness, consistency, reliability, reputation, timeliness, understandability, validity and others [26, 27, 28, 29, 30, 31]. Since there is no general agreement on data quality dimensions [30], the aspects to be considered depend on the information consumers requirements [32]. Reputation, reliability, accuracy and timeliness are particularly related to the requirements of our applicative context. To represent these aspects, we propose to qualify the information quality according to a confidence degree defined at two levels of granularity: for a piece of information and for an observation set. Data quality assessment is often based on rules [33, 34], which can also handle uncertainty [35, 36]. We propose to represent the knowledge model by a Bayesian network that embeds data quality rules (e.g. about the domain of variables, possible or impossible combinations between subsets of variables and their values). The idea is to maintain the quality of the observation set by continuously (re)evaluating the confidence degree of each piece of information, whether new or not, and maintain the compatibility with the probabilistic generic knowledge model.

Information quantity is another important aspect in our context. The quantity is related to the number of observations in the set.

Choosing quantity rather than quality would lead to accept any new information, whatever the source, and to keep observation as long as the set of observations remains compatible with the general knowledge model even if this augments the risk of keeping obsolete information in the observation set. At the opposite, choosing quality rather than quantity leads to accept a new observation only when its source is really sure, and to remove previous observations as soon as they may have become obsolete, but this choice would result in a set with very few observations.

We thus have to propose a compromise between quality and quantity in the observation set. To that end, the principle is to remove from the set the observations whose quality is no longer sufficient, and the observations that cause a compatibility problem.

2.4. Definitions and problem statement

The information to consider comes both from generic knowledge about the dependency relations between the variables and their variability within a population, and from current and past observed states of the specific patient. Both information types can be managed using a Bayesian network approach [37] provided that some additional mechanisms are added to address the time-related uncertainty.

Before defining Bayesian networks and associated definitions, we give some notations:

- X, X_k denote random variables. Capital letters are used to represent random variables.
- $Dom(X)$ denotes the domain of the variable X. When needed, we use the detailed notation $Dom(X) = \{x^1, \ldots, x^m\}$, but most often, we denote a value of X by $x \in Dom(X)$. Lower-case letters represent the values of a variable.

- $\mathbf{X} = \{X_1, \ldots X_n\}$ denotes a set of variables. Bold capital letters correspond to sets of variables.

- $\mathbf{x} = (x_1, \ldots x_n)$ denotes a possible value (or assignment) of $\mathbf{X}$, with each $x_i \in Dom(X_i)$, and $\mathbf{x} \in Dom(\mathbf{X}) = Dom(X_1) \times \ldots \times Dom(X_n)$. A bold lower-case letter represents the value of a set of variables.

- $\mathbf{X}_K = \{X_{k_1}, \ldots, X_{k_K}\} \subseteq \mathbf{X}$ denotes a subset of variables, with $K = \{k_1, \ldots, k_K\} \subseteq \{1, \ldots, n\}$.

- $\mathbf{x}_K \in Dom(\mathbf{X}_K)$ denotes one of the possible values of $\mathbf{X}_K$. We refer to $x_k \in Dom(X_k)$ as an element of the vector $\mathbf{x}_K$.

- $P(X)$ is the probability distribution $(P(X = x^1), \ldots, P(X = x^m))$.

- $P(\mathbf{X}) = P(X_1, \ldots, X_n)$ is the joint probability distribution on the set $\mathbf{X}$.

- $P(x)$ denotes $P(X = x)$ and $P(\mathbf{x}_K)$ denotes $P(\mathbf{X}_K = \mathbf{x}_K)$.

A *Bayesian network* $\mathcal{B} = (\mathcal{G}, P)$ consists in a set of random variables $\mathbf{X} = \{X_1, \ldots, X_n\}$, a directed acyclic graph $\mathcal{G} = (\mathbf{X}, \mathbf{E})$, and a set of parameters Θ including the local conditional 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 $P(\mathbf{X} = \mathbf{x}) = \prod_{X_i \in \mathbf{X}} P(X_i = x \mid pa(X_i))$, where $pa(X_i) \subset \mathbf{X}$ is the set of the parents (immediate predecessors) of X_i in the graph $\mathcal{G}$.

In our proposal, all the variables of the observation set are in the Bayesian network, but some variables of the Bayesian network may never be observed. We consider the possibility of *uncertain observations*. For example, an observation about the diagnosis of a disease can be uncertain because the specialist declares that she is not completely sure about it. The uncertain observation is specified by a likelihood vector or *potential* as proposed by Pearl [37, 38]. The potential of an uncertain evidence e on a variable X represents the relative likelihoods for the possible observation of every state of the variable $\pi = P(e \mid X = x^i)$. This kind of evidence is easily propagated in a BN by Pearl's method of virtual evidence. A *potential* on a variable X is a measure $\pi : Dom(X) \longrightarrow \mathbb{R}_+$. We

assume that, for any potential on a variable X, at least one value of its potential must be non zero: there exists $x \in Dom(X)$ such that $\pi(x) > 0$. In order to simplify the notation, we also use π to denote the vector $(\pi(x^1), \ldots, \pi(x^m))$.

The most probable values of X in a potential π is defined as follows: $\text{MPV}(\pi) = \{x \in Dom(X), \pi(x) = max_{x \in Dom(X)}(\pi(x))\}$. The set $\text{MPV}(\pi)$ includes at least one element, and at most $|Dom(X)|$ elements. Let $\mathbf{X}_K = \{X_{k_1}, \ldots, X_{k_K}\} \subseteq \mathbf{X}$ a subset of variables and $\{\pi_{k_1}, \ldots, \pi_{k_K}\}$ a set of associated potentials. The product $\Pi_K = \pi_{k_1} \times \ldots \times \pi_{k_K}$ is a measure from $Dom(\mathbf{X}_K)$ in $\mathbb{R}_+$ such that $\Pi_K(x_{k_1}, \ldots, x_{k_K}) = \pi_{k_1}(x_{k_1}) \cdot \ldots \cdot \pi_{k_K}(x_{k_K})$. As previously, we also use Π_K to denote the vector whose components are the images of the elements of $Dom(\mathbf{X}_K)$.

Example 1. ***Potential of a set of observations***

Consider two potentials:

$\pi_1 : \{x_a, x_b\} \to \mathbb{R}_+$ *defined by the vector* $(\pi_1(x_a), \pi_1(x_b)) = (1, 0)$*, and*

$\pi_2 : \{y_a, y_b\} \to \mathbb{R}_+$ *defined by the vector* $(\pi_2(y_a), \pi_2(y_b)) = (0.3, 0.7)$.

The product of the two potentials $\Pi_{1,2} = \pi_1 \times \pi_2$ *is defined by the vector* $(\Pi_{1,2}(x_a, y_a), \Pi_{1,2}(x_a, y_b), \Pi_{1,2}(x_b, y_a), \Pi_{1,2}(x_b, y_b)) = (0.3, 0.7, 0, 0)$.

The observation set regroups the potentials of the observed variables together with the date of the observation and the sources that provided the information.

Definition 1. ***Time stamped uncertain observation*** *A time stamped uncertain observation o on a variable X is a tuple $o = (X, \pi, t, s)$ where π is a potential on X that represents the uncertain observation on X, t is the time of the observation, and s is the source of the observation.*

In this work, we consider the moment when the state (or value) of a variable is observed and we assume that this moment is very close to the moment when we are trying to add the information in the observation set. The observation does not concern the time at which a variable changes in value, but its value at the moment of the observation. In the following, we often simply use the term *time stamped observation* or just *observation* to refer to a time stamped uncertain observation.

When there is a past observation on the variable X and no current observation is available on that variable, we do not know whether its current state is still one of those provided by the past observation. Since we consider the possibility of evolving variables, an observation made at time t may no longer be reliable at time $t + \Delta t$: the observation can become *obsolete*.

Definition 2. ***Obsolete observation***

A time-stamped observation $o = (X, \pi, t, s)$ is obsolete when the current value of X — even if not observed — is no longer one of the set of the most probable values $\mathrm{MPV}(\pi)$.

An observation o can become obsolete independently of the arrival of a new observation on X, which makes it difficult to detect. More generally, an observation $o = (X, \pi, t, s)$ is *erroneous* when the current value of X — even if not observed — is not one of the most probable ones in π. This includes the case where an observation is erroneous from the beginning, for example due to a poor source of information.

We now consider a set of time stamped uncertain observations in order to store available observations.

Definition 3. ***Time stamped observation set on*** $\mathbf{X}$

A time stamped observation set on $\mathbf{X}$ is a set $\mathbf{O} = \{(X, \pi, t, s), X \in \mathbf{X}\}$ such that for any variable $X \in \mathbf{X}$, there is at most one observation on X in $\mathbf{O}$.

We consider both a generic knowledge model which is given by a Bayesian network, and specific knowledge about ~~an~~ one individual which is given by a set of uncertain observations on a subset of variables of the general model. Each observation has to be compatible with the general knowledge, and in a set of observations, the pieces of information have to be compatible altogether. Incompatibility may occur when the probability distribution P of the Bayesian network includes some zeros, and more precisely when there is a value for which P is null and the potential of an observation is not null. When that situation is avoided, the potential of the observation is said to be dominated by P.

Below, we recall the definition of that *dominance* condition and then we propose the notions of *partial compatibility* and *total compatibility* between an observation and a general model of knowledge.

Definition 4. ***Dominance of measures*** *A measure ν is dominated by a measure μ, which is denoted by $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$ for any measurable A.*

In the following, we consider the dominance of a probability measure onto a potential. The probability measure P is part of the Bayesian network that constitutes the knowledge model associated with a set of random variables $\mathbf{X}$. For simplicity, we also denote by P the marginal probability measures associated with a single random variable $X \in \mathbf{X}$ or with a set of random variables $\mathbf{X}_K \subset \mathbf{X}$. When considering an observation $o = (X, \pi, t, S)$ on a single variable X, the potential π of the observation can be dominated or not by P, more precisely by the marginal probability measure of X defined from P.

Definition 5. ***Totally compatible uncertain observation***

An uncertain observation $o = (X, \pi, t, s)$ on a variable $X \in \mathbf{X}$ is totally compatible with the probability distribution P if π is dominated by the marginal probability measure $P(X)$, which means that $P(X = x) = 0$ implies $\pi(x) = 0$ for every $x \in Dom(X)$.

In others words, when the knowledge model states that a particular value is impossible, an observation stating the contrary is not totally compatible with the model. Such an observation may however be partially compatible with the model of knowledge when at least one value is possible regarding both the observation and the knowledge model.

Definition 6. ***Partially compatible uncertain observation***

An uncertain observation $o = (X, \pi, t, s)$ on a variable $X \in \mathbf{X}$ is partially compatible with the probability distribution P if there exists $x \in Dom(X)$ such that $P(X = x) > 0$ and $\pi(x) > 0$.

By extension, we say that an observation is partially compatible or totally compatible with a Bayesian network.

Proposition 1. *Let $o = (X, \pi, t, s)$ be an uncertain observation on a variable $X \in \mathbf{X}$; if π is totally compatible with a probability distribution P, then π is partially compatible with P.*

Proof. By definition, a potential π includes at least one value non zero; let $x \in Dom(X)$ be a value such that $\pi(x) > 0$. Since π is totally compatible with P, π is dominated by P, and thus we must have $P(X = x) > 0$. We thus have a value x for which both P and π are non zero, which means that π is partially compatible with P. □

The opposite is not true: an uncertain observation may be partially compatible with P but not totally compatible as in the example below.

Example 2. ***An observation partially compatible with P but not totally compatible.***

Let $Dom(X) = \{blue, red, green, yellow\}$, and $P(X) = (\alpha, \alpha, \alpha, 0)$, where α represents a non zero value; this means that yellow is impossible. Let's consider an observation o stating that the color is either red or yellow, but is neither blue not green. The observation o is specified by the potential $\pi = (0, \alpha, 0, \alpha)$. The potential π is partially compatible with P since the value red is possible in both measures, but it is not totally compatible with P since the value yellow is impossible according to P but is possible in the observation.

The next two definitions concern the compatibility of a set of observations with a probability measure. When considering a set of observations $\mathbf{O}_K = \{o_k, k \in K\}$ on a set of variables $\mathbf{X}_K$, the potential Π_K defined as the product of the potentials can be dominated or not by P, more precisely by the marginal probability measure of $\mathbf{X}_K$ defined from P.

Definition 7. ***Totally compatible uncertain observation set***

A time stamped observation set $\mathbf{O}_K = \{(X_k, \pi_k, t_k, s_k), X_k \in \mathbf{X}_K\}$ is totally compatible with the probability distribution P if the product $\Pi_K = \pi_{k_1} \times \ldots \times \pi_{k_K}$

is dominated by the marginal probability measure $P(\mathbf{X}_K)$, which means that for every $\mathbf{x}_K \in Dom(\mathbf{X}_K)$, $P(\mathbf{X}_K = \mathbf{x}_K) = 0$ implies $\Pi_K(\mathbf{x}_K) = 0$.

Proposition 2. *If a time stamped observation set $\mathbf{O}_K$ is totally compatible with a probability distribution P, then any subset $\mathbf{O}_{K'} \subset \mathbf{O}_K$ is totally compatible with P. As a consequence, every observation $o_k \in \mathbf{O}_K$ is totally compatible with P.*

Proof. Consider an observation $o_k \in \mathbf{O}_K$, and let $\mathbf{O}_{K'} = \mathbf{O}_K \setminus \{o_k\}$ and $\mathbf{X}_{K'} = \mathbf{X}_K \setminus \{X_k\}$ the corresponding set of variables. Let $\mathbf{x}_{K'}$ be an assignment of $\mathbf{X}_{K'}$ such that $P(\mathbf{x}_{K'}) = 0$ and let's see whether $\Pi_{K'}(\mathbf{x}_{K'})$ is null. Since $P(\mathbf{x}_{K'}) = 0$, for any value $x_k \in Dom(X_k)$, $P(\mathbf{x}_{K'}, x_k) = 0$. Since $\mathbf{O}_K$ is totally compatible with P, this implies that for any value $x_k \in Dom(X_k)$, $\Pi_K(\mathbf{x}_{K'}, x_k) = 0$. By definition, $\Pi_K(\mathbf{x}_{K'}, x_k) = \Pi_{K'}(\mathbf{x}_{K'}) \cdot \pi_k(x_k)$. We know that any potential has at least a value non zero. Let x_{k_0} be a value such that $\pi_k(x_{k_0}) \neq 0$. Since $\Pi_{K'}(\mathbf{x}_{K'}) \cdot \pi_k(x_{k_0}) = 0$, then $\Pi_{K'}(\mathbf{x}_{K'}) = 0$, which confirms that the subset $\mathbf{O}_{K'} = \mathbf{O}_K \setminus \{o_k\}$ is totally compatible with P. □

The opposite is not true (see Example 3).

Example 3. ***An observation set which is not totally compatible with P whereas each of its observation is totally compatible with P.***

Let $\mathbf{X} = \{X, Y, Z\}$ and $Dom(X) = \{x_a, x_b, x_c\}$, $Dom(Y) = \{y_a, y_b, y_c\}$, $Dom(Z) = \{z_a, z_b, z_c\}$. The joint probability P on $\mathbf{X}$ is given in Table 1. This probability measure includes the following zeros: $P(x_a, y_a, z_a) = 0$, $P(x_c, y_b, z_a) = 0$, $P(x_c, y_c) = 0$.

The observation o_1 states that $X = x_c$; it is specified by $\pi_1 = (0, 0, 1)$. The observation o_2 states that Y is either y_b or y_c, with y_c nine times more probable than y_b but Y cannot be y_a; it is specified by $\pi_2 = (0, 0.1, 0.9)$. From Table 1, the marginal probability distributions of X (respectively Y) has no zero value. As a consequence, the observation o_1 (resp. o_2) is totally compatible with P.

We consider the set of observations $\{o_1, o_2\}$. The marginal probability distribution of (X, Y) is represented by the vector $(\alpha, \alpha, \alpha, \alpha, \alpha, \alpha, \alpha, \alpha, 0)$ as shown

in Table 1. The only zero value corresponds to (x_c, y_c)*. The product of the potentials of the two observations is* $\Pi_{12} = \pi_1 \times \pi_2 = (0, 0, 0, 0, 0, 0, 0, 0.1, 0.9)$*. We thus have* $P((X, Y) = (x_c, y_c)) = 0$ *and* $\Pi_{12}(x_c, y_c) > 0$*, which means that the set of observations* $\{o_1, o_2\}$ *on* $\{X, Y\}$ *is not totally compatible with* P*.*

Table 1: Joint probability distribution on $\{X, Y, Z\}$. The symbol α represents a non zero value.

$P(X,Y,Z)$	x_a			x_b			x_c		
	y_a	y_b	y_c	y_a	y_b	y_c	y_a	y_b	y_c
z_a	0	α	α	α	α	α	α	0	0
z_b	α	α	α	α	α	α	α	α	0
z_c	α	α	α	α	α	α	α	α	0
$P(X,Y)$	α	α	α	α	α	α	α	α	0

When a set of observations is not totally compatible with a probability distribution, it may however be partially compatible. This new definition is given below.

Definition 8. ***Partially compatible uncertain observation set***

A time stamped observation set $\mathbf{O}_K = \{(X_k, \pi_k, t_k, s_k), X_k \in \mathbf{X}_K\}$ *is partially compatible with the probability distribution* P *if there exists at least one assignment* $\mathbf{x_K} \in Dom(\mathbf{X}_K)$ *such that* $P(\mathbf{X}_K = \mathbf{x_K}) > 0$ *and* $\Pi_K(\mathbf{x}_K) > 0$*, where* Π_K *is the product of the local potentials* π_k*, for* $k \in K$*.*

Proposition 3. *If a time stamped observation set* $\mathbf{O}_K$ *is partially compatible with a probability distribution* P*, then each subset* $\mathbf{O}_{K'} \subset \mathbf{O}_K$ *is partially compatible with* P*. As a consequence, each observation* $o_k \in \mathbf{O}_K$ *is partially compatible with* P*.*

Proof. Consider a set $\mathbf{O}_K$ partially compatible with the probability distribution P. By definition, there is an assignment $\mathbf{x_K} \in Dom(\mathbf{X}_K)$ such that (1) $P(\mathbf{X}_K = \mathbf{x_K}) > 0$ and (2) $\Pi_K(\mathbf{x}_K) > 0$. Consider an observation $o_k = (X_k, \pi_k, t_k, S_k) \in \mathbf{O}_K$ and the subset $\mathbf{O}_{K'} = \mathbf{O}_K \setminus \{o_k\}$. Let $\mathbf{X}_{K'} = \mathbf{X}_K \setminus \{X_k\}$ be the associated subset of variables. The assignment $\mathbf{x}_K$ can be written as $\mathbf{x}_K = (\mathbf{x}_{K'}, x_k)$ with $\mathbf{x}_{K'} \in Dom(\mathbf{X}_{K'})$, and $x_k \in Dom(X_k)$. Since $P(\mathbf{x}_{K'}, x_k) > 0$,

Table 2: Joint probability distribution of $\{X, Y\}$, and marginal distributions of X and Y in Example 4. The symbol α represents a non zero value.

$P(X,Y)$	x_a	x_b	$P(Y)$
y_a	α	α	α
y_b	0	α	α
$P(X)$	α	α	

we have $P(\mathbf{x}_{K'}) > 0$. In addition, $\Pi_K(\mathbf{x}_K) = \Pi_{K'}(\mathbf{x}_{K'}) \cdot \pi_k(x_k) > 0$, therefore $\Pi_{K'}(\mathbf{x}_{K'}) > 0$. So the subset $\mathbf{O}_{K'}$ is partially compatible with P. □

The opposite is not true (see Example 4).

Example 4. ***An observation set which is not partially compatible with*** *P* ***whereas each observation is partially compatible****. Let $Dom(X) = \{x_a, x_b\}$ and $Dom(Y) = \{y_a, y_b\}$ and the joint probability distribution P is given in Table 2. It follows that the marginal probability distributions $P(X)$ and $P(Y)$ have no null value. For example $P(x_a) = P(x_a, y_a) + P(x_a, y_b) > 0$ since $P(x_a, y_a) > 0$. As a consequence, any potential on X (respectively on Y) is partially compatible with P. Consider the observations: o_1 on X with $\pi_1 = (1, 0)$, stating that $X = x_a$, and o_2 on Y with $\pi_2 = (0, 1)$, stating that $Y = y_b$. The product of the two potentials $\Pi_{12} = \pi_1 \times \pi_2$ includes a single non zero value which is $\Pi_{12}(x_a, y_b)$. Since there is no value of (X, Y) such that P and Π_{12} are both non zero, the observation set $\{o_1, o_2\}$ is not partially compatible with P whereas separately, each observation is partially compatible with P.*

Proposition 4. *Let $\mathbf{O}_K$ be a set of uncertain observations, if $\mathbf{O}_K$ is totally compatible with a probability distribution P, then it is partially compatible with P.*

Proof. Let $\mathbf{O}_K$ be a set of observations totally compatible with P. By definition, each potential has at least one non zero value. Let $\mathbf{x}_K^0 \in Dom(\mathbf{X}_K)$ be such a value for the potential Π_K. Since $\mathbf{O}_K$ is totally compatible with P, $\Pi_K(\mathbf{x}_K^0) > 0$ implies that $P(\mathbf{x}_K^0) > 0$, which means that $\mathbf{O}_K$ is partially compatible with P. □

The opposite is not true (see the example below).

Example 3 (continued). A set of observations partially compatible with P but not totally compatible.

We have seen that the set $\{o_1, o_2\}$ on $\{X, Y\}$ is not totally compatible with P. Note also that the value (x_c, y_b) is possible both in P and in Π_{12} which means that the set $\{o_1, o_2\}$ is partially compatible with P. Finally, the observations o_1 and o_2 are separately compatible with P, the set $\{o_1, o_2\}$ is partially compatible with P, but not totally compatible with P.

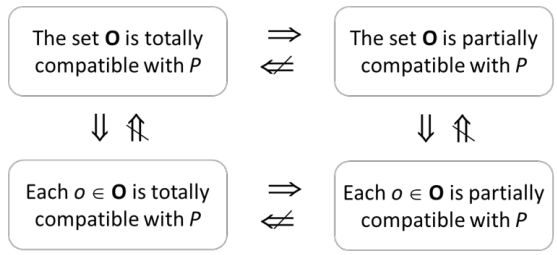

Figure 1: Synthetic view of Propositions 1 to 4, where **O** is an observation set.

Figure 1 summarizes the last four propositions. Note that, in this article, we consider observations sets in which each observation concerns a single variable. Observations on several variables could also be considered and specified by a potential on the product of their domain but we do not consider this possibility in detail.

The potential of a variable in an observation represents a first type of uncertainty that relates to the relative likelihood of each possible value. Two other types of uncertainty can modify the initial information: the uncertainty due to unsure source of information and the uncertainty due to the age of the information: as time has passed since the observation, the value of the variable which is concerned may have evolved. To represent these last two types of uncertainty, we define a *confidence degree* for each observation.

Definition 9. ***Confidence degree***

The confidence degree of a time stamped observation $o = (X, \pi, t, s)$ at the time $t' = t + \Delta t$ is a real value $c(o, t')$ in $[0, 1]$ that represents how much the

observation is reliable at t'.

Note that the confidence degree at the time t of the observation may not be 1; indeed, when the information source is known, the confidence at the time the observation is received depends on the trust in this source.

We extend the definition of the confidence degree to a set of observations. This definition is guided by the two types of decisions that must be taken while maintaining an observation set. (1) When a new piece of information is available, we have to decide whether to add it or not to the observation set. This decision is based both on the compatibility of the new observation with the existing set and on the way it affects the confidence degree of the whole set. **In particular, we want to prevent the addition of new observations with too low a confidence degree since a single erroneous observation can lead to unreliable result. (2) When there is a conflict between the new observation and the existing set, the removal of different subsets of observations may allow to restore compatibility. In order to decide which subset should be removed, the number of observations of each subset can be considered, together with the impact on the confidence degree of the observation set. It is expected that removing the set with the lowest confidence degree leads to the highest confidence degree for the resulting set. Moreover, there should be no counterbalance between low and high confidence degrees: the assessment of the quality of the complete set must not hide a low confidence degree for an observation because of the presence of more reliable ones. For these reasons, we propose the definition below based on the function *minimum*.**

Definition 10. ***Confidence degree of an observation set***

The confidence degree of a non empty observation set $\mathbf{O} = \{o_1, \ldots, o_k\}$ at the time t is a real value in $[0, 1]$ that represents how much the observation set is reliable at t. It is defined by the confidence degree of the less reliable piece of information of the set: $c(\mathbf{O}, t) = min(c(o_1, t), \ldots, c(o_k, t))$.

The properties of this function that are interesting for the management of an observation set are given below: whatever the observation o, the observation set $\mathbf{O}$ and two subsets $\mathbf{O}' \subset \mathbf{O}$, $\mathbf{O}'' \subset \mathbf{O}$,

1. $c(\{o\}, t) = c(o, t)$
2. *Adding an observation can only decrease the confidence degree of a set or let it stable*: $c(\mathbf{O} \cup \{o\}, t) \leq c(\mathbf{O}, t)$. As a consequence, the addition of a new observation to the set should be done with caution.
3. *Removing an observation can not decrease the confidence degree of a set* $c(\mathbf{O} \setminus \{o\}, t) \geq c(\mathbf{O}, t)$
4. *If necessary, to decide which subset among* $\mathbf{O}'$ *and* $\mathbf{O}''$ *should be removed from* $\mathbf{O}$ *depending on the confidence degree comparison:*
   a. if $c(\mathbf{O}', t) < c(\mathbf{O}'', t)$ then $c(\mathbf{O} \setminus \mathbf{O}', t) \geq c(\mathbf{O} \setminus \mathbf{O}'', t)$ **(the subset with the lowest confidence degree should be removed)**,
   b. if $c(\mathbf{O}', t) = c(\mathbf{O}'', t)$ then $c(\mathbf{O} \setminus \mathbf{O}', t) = c(\mathbf{O} \setminus \mathbf{O}'', t)$.

Note that the *mean* function would not be suitable to define the confidence degree of an observation set since it does not verify the properties (2), (3) and (4) (see Example 5).

Example 5. ***The mean function cannot be used for the confidence degree of an observation set.*** *Consider the following three observation sets:* $\mathbf{O} = \{o_1, o_2, o_3, o_4\}$ *with* $c(o_1, t) = 0.1$, $c(o_2, t) = 0.7$, $c(o_3, t) = 0.6$ *and* $c(o_4, t) = 0.8$, $\mathbf{O}' = \{o_2\}$ *and* $\mathbf{O}'' = \{o_3, o_4\}$. *The mean of the confidence degrees in* $\mathbf{O}'$ *(resp.* $\mathbf{O}''$*) equals 0.7. Let's compare the means of* $\mathbf{O} \setminus \mathbf{O}' = \{o_1, o_3, o_4\}$ *and* $\mathbf{O} \setminus \mathbf{O}' = \{o_1, o_2\}$. *The mean of the first set is* $(0.1 + 0.6 + 0.8)/3 = 0.5$ *whereas the mean of the second one is* $(0.1 + 0.7)/2 = 0.4$, *which contradicts the property (4.a).*

When the confidence degree of an observation becomes too low, the observation is likely to have become obsolete. In the same way, when the confidence degree of an observation set becomes too low, it is likely that the set includes an observation that has become obsolete.

We can now define more precisely the problem that we tackle. Let $\mathcal{B}$ be a Bayesian network, $\mathbf{X}$ the corresponding set of discrete random variables, and $\mathbf{O}$ a time stamped observation set on $\mathbf{X}$. Over time, the proportion of obsolete information grows when nothing is made to maintain the observation set. When new observations are received, they may contribute to increase the quantity of observations in the set, but also reveal incompatibilities in the observation set relatively to the general knowledge model.

In that context, we propose to periodically update the observation set in order to (1) take into account recently collected observations by making a compromise between the quantity and the quality of the information, (2) decrease the presence of obsolete or erroneous information and (3) maintain compatibility. These operations are presented in the following section, with the functions for the observation set management.

3. Observation set management system

This section presents the architecture of the system and details the updating operations.

3.1. The architecture of the observation set management system

The system combines individual information contained in the observation set and generic knowledge, to provide on demand the belief and confidence about a variable according to the architecture in Figure 2. The system provides two outputs with uncertainty: (a) the belief on a variable depends on stochastic uncertainty and incomplete information about the considered person; (b) the confidence degree depends on uncertainty due to information aging and trust in sources. The intelligent modules of the system are dedicated to manage the observation set and to evaluate beliefs about requested variables. This study focuses on the management operations: management of new observations, periodic removal of unsure observations and evaluation of the confidence degree associated with each observation along with time.

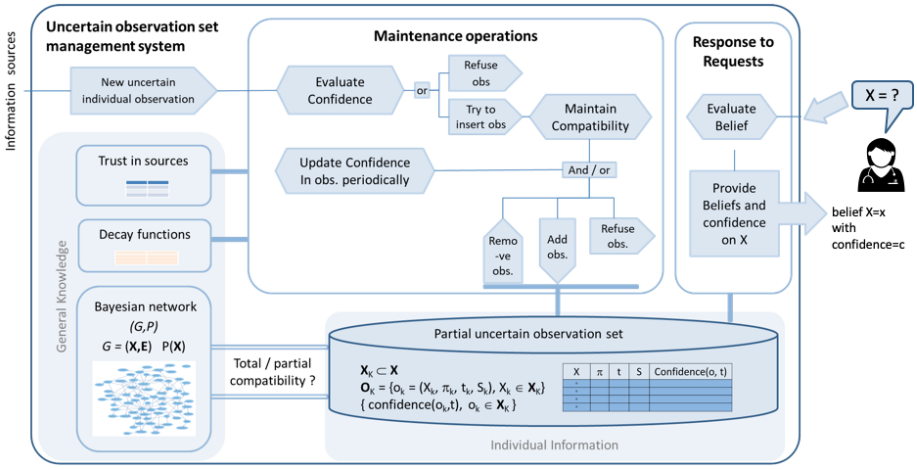

Figure 2: General architecture of the observation set management system.

The Bayesian network represents the main part of the generic knowledge in the system. It is completed by two other models about the decay functions and the trust in the sources.

The confidence in an observation depends both on the trust we have in the source of information and on the length of time since the information has been received. We propose to model the loss of confidence in the past observations by the use of decay functions. These functions are either constant or decreasing with time, and provide the confidence degree in an observation according to time. However these functions have distinct shapes, depending on the variable. The shape of the decay functions can be defined from expert knowledge or learned from data.

Example 6. *The observation "Ms A suffers from depression" which was evaluated as highly reliable at the time of the observation, can be less reliable two months later, and even less reliable four months after, etc. Regarding this information, we can assume to represent the loss of confidence by a decreasing linear function, with a slope to be defined. In contrast, "Ms A lives alone" can be as-*

sumed to remain stable on a long-term perspective, which would be represented by a quasi-constant confidence degree.

Information about the sources of observation is also recorded in order to evaluate the confidence in new observations, at the moment they are received by the system. In order to keep track of the observation sources, learning techniques can be used to adapt the trust level [39, 40].

3.2. Updating operations

The observation set is periodically maintained up-to-date by the two operations described in the `maintain` algorithm below. At first, the observation set is inspected to be cleaned from observations that are likely to be obsolete. When the confidence degree in a past observation falls below a given threshold, the risk for the observation to be obsolete is too high, and thus the observation is removed from the observation set (lines 1-3). Secondly, decisions are made when new observations are received in order to maintain the compatibility of the information base (lines 4-end). These operations call two additional algorithms, `confidence` and `integrateNewObs` described below.

Algorithm 1: maintain($\mathbf{O}$, $\mathbf{O}_{new}$, P, $cType$)

Input: $\mathbf{O}$ ▷ *an observation set*
Input: $\mathbf{O}_{new}$ ▷ *the set of new observations*
Input: P ▷ *a probability distribution*
Input: $cType \in \{$`'total'`, `'partial'`$\}$ ▷ *the type of compatibility*
Output: $\mathbf{O}$ ▷ *the updated observation set*
Data: β ▷ *confidence threshold*

1 **foreach** $o \in \mathbf{O}$ **do**
2 **if** $\mathrm{confidence}(\mathrm{o}, \mathrm{Now}) < \beta$ **then**
3 $\mathbf{O} \leftarrow \mathbf{O} \setminus \{o\}$ ▷ *clean the set by removing obs. likely to be obsolete*
4 **foreach** *new observation* $o' = \langle X, \pi', t', s' \rangle \in \mathbf{O}_{new}$ **do**
5 **if** $\mathrm{confidence}(\mathrm{o}', \mathrm{Now}) \geq \beta$ **then** ▷ *the confidence degree in* o' *is sufficient*
6 $\mathbf{O} \leftarrow \mathrm{integrateNewObs}(\mathbf{O}, \mathrm{o}', \mathrm{P}, \mathrm{cType})$
7 **return O**

The `confidence` algorithm provides the degree of confidence in the observation o at a given time t_c. When the confidence to compute concerns a current

new observation (line 1), the question is: do we already know the observation source? If the response is yes, we can use the trust in the source to provide a confidence degree (lines 2-3). If no, we decide to start with the highest confidence degree (lines 4-5). When the confidence to compute concerns an observation made in the past (lines 6-7), we apply a decay function to compute the current confidence according to the initial confidence degree in the observation and the duration since it was observed, as described in Section 3.

Algorithm 2: confidence(o, t_c)

Input: $o = \langle X, \pi, t, s \rangle$ ▷ *observation π on X provided by the source s at t*
Input: t_c ▷ *time of the confidence degree computing*
Output: c ▷ *the confidence degree in o at time t_c*

Data: **S** ▷ *known input sources*
Data: trust() ▷ *the trust function applied to known input sources*
Data: $\mathrm{decay_X}()$ ▷ *the decay function of the variable X*

```
1 if t_c = t then                                  ▷ the confidence concerns a new observation
2     if s ∈ S then                                ▷ there is knowledge about the observation source
3         c ← trust(s)
4     else                                         ▷ no knowledge about the source s
5         c ← 1                                    ▷ by default, the source is trusted
6 else                                             ▷ the confidence to compute concerns a previous observation
7     c ← decay_X(confidence(o, t), t_c − t)
8 return c
```

The `integrateNewObs` algorithm decides whether to integrate or not a new observation in the observation set and updates the set accordingly. The decision to integrate a new observation depends on (i) the confidence in the new observation, (ii) the compatibility of the observation set according to a general model of knowledge and (iii) the quality of the observation set, given by the confidence degree of the observation set. This decision relies thus on the confidence degree that results from the `confidence` algorithm described above and on the compatibility that is checked by the algorithm `isCompatible`, which is now presented before the detailed description of the `integrateNewObs` algorithm.

The algorithm `isCompatible` checks the compatibility of an observation set regarding the general knowledge represented by the BN joint probability distribution. It may refer to either total compatibility or partial compatibility, as

previously defined in Section 2.4. The algorithm checks whether an observation set is totally or partially compatible with a probability distribution P, depending on the value of the parameter $cType \in \{\text{'total'}, \text{'partial'}\}$. When the probability distribution includes no zero (line 1-2), it follows from Definitions 7 and 8 that any observation set is both totally and partially compatible with P. Else, in order to verify the total compatibility of an observation set (lines 5-12), the assignments of the observed variables that are possible in the observation set, i.e. the assignments $\mathbf{x}_K$ such that $\Pi_K(\mathbf{x}_K) > 0$, are examined until one of them is found with a null probability (line 11), meaning that the observation set is not dominated by the probability P (line 12). When such an assignment is not found, the observation set is totally compatible with P. In order to verify the partial compatibility of an observation set (lines 13-20), the set of the possible assignments of the observed variables is examined until one is found that have a non zero probability. The existence of such an assignment is sufficient to establish the partial compatibility of the set (line 20).

On the base of these operations, we can now described the `integrateNewObs` algorithm in details. The first step is to check whether the observation set includes a previous observation on the variable concerned by the new observation (lines 1-6). When the previous observation is more reliable than the new one, the new observation is refused (line 6). In the other case, the previous observation is removed from the observation set. When the new observation is compatible with the observation set according to the general knowledge model, it can be integrated (lines 7-8). When it is not compatible (lines 9-26), the subsets of observations responsible for the incompatibility are sought and the most relevant of them is selected to be removed in order to add the new observation. At worst, the complete set of previous observations can be considered. The search for the subset of the observations likely to be obsolete (lines 10-20) begins with the singletons (line 10), and when no singleton allows to restore the compatibility (line 15), larger subsets are checked (line 20). When several subsets of the same size are found, we consider the one with the lowest confidence degree (lines 17-19). Finally, the confidence degree of the set of the observations likely to

Algorithm 3: isCompatible($\mathbf{O}$, P, $cType$)

Input: $\mathbf{O}_K = \{(X_k, \pi_k, t_k, S_k), X_k \in \mathbf{X}_K\}$ ▷ *an observation set*
Input: P ▷ *a joint probability distribution*
Input: $cType \in \{$`'total'`, `'partial'`$\}$ ▷ *the type of compatibility*
Output: isC ▷ *a boolean, True when* $\mathbf{O}_K$ *is compatible with* P

```
if P includes no zero then        ▷ P(x_K) > 0 for any assignment x_K ∈ Dom(X_K)
    isC ← True        ▷ any observation set is totally and partially compatible with P
else                                                          ▷ P includes at least one zero
    Π_K ← π_{k_1} × ... × π_{k_K}
    if cType = 'total' then                ▷ aims to test the total compatibility of O
        ▷ try to find an assignment x_K s.t. P(X_K = x_K) = 0 and Π_K(x_K) > 0
        isC ← True
        n ← |Dom(X_K)|
        while isC = True and n > 0 do
            x_K ← next assignment in Dom(X_K)                  ▷ next Π_K(x_K) > 0
            n ← n − 1
            if P(X_K = x_K) = 0 then
                isC ← False
    else                                       ▷ aims to test the partial compatibility of O
        isC ← False
        n ← |Dom(X_K)|
        while isC = True and n > 0 do
            x_K ← next assignment in Dom(X_K)                  ▷ next Π_K(x_K) > 0
            n ← n − 1
            if P(X_K = x_K) > 0 then
                isC ← True
return isC
```

be obsolete is compared with the confidence degree of the new observation and the more reliable is kept whereas the other one is rejected (lines 21-25). When the new observation is refused, the previous observation on the same variable is restored, if it existed (lines 24-25).

When there is some incompatibility, the search for the set of the observations likely to be obsolete is guided by two criteria: first remove the smallest number of observations, in order to minimize the loss of information of the observation set, and second, remove the observation(s) with the lowest confidence degree. The first criterion is known as the principle of minimal change and is frequently used in belief revision [5, 6].

Proposition 5. *Let* $\mathbf{O}'$ *be the observation set updated after the arrival of a new observation by the algorithm* `integrateNewObs`*, with* $\mathbf{O}$ *the previous observation set, o the new observation,* P *the joint probability distribution, and cType the required compatibility ('partial' or 'total'). We have:*

1. *If* $\mathrm{confidence}(\mathbf{O}, \mathrm{t}) \geq \beta$ *and* $\mathrm{confidence}(\mathrm{o}, \mathrm{t}) \geq \beta$ *then* $\mathrm{confidence}(\mathbf{O}', \mathrm{t}) \geq \beta$
2. *If the algorithm aims to test the total compatibility (cType =* `'total'`*) and the set* $\mathbf{O}$ *is totally compatible with* P*, then* $\mathbf{O}'$ *is totally compatible with* P*.*
3. *If the algorithm aims to test the partial compatibility (cType =* `'partial'`*) and the set* $\mathbf{O}$ *is partially compatible with* P*, then* $\mathbf{O}'$ *is partially compatible with* P*.*

Proof. Two main updating operations are performed by `integrateNewObs`: the removal of some observations (line 4 and 22) and the addition of the new observation (line 8 and 22).

(1) The definition of the confidence degree of an observation set states that the removal of any observations from $\mathbf{O}$ can only increase its confidence degree. Thus the confidence degree of the modified set after the removal of an observation (lines 4, 22) remains greater or equal to β. Regarding the addition of a new

Algorithm 4: integrateNewObs($\mathbf{O}$, o, P, $cType$)

Input: $\mathbf{O}$ ▷ *an observation set*
Input: $o = \langle X, \pi, t, s \rangle$ ▷ *a new observation with a sufficient confidence degree*
Input: P ▷ *a joint probability distribution*
Input: $cType \in \{$`'total'`, `'partial'`$\}$ ▷ *the type of compatibility*
Output: $\mathbf{O}$ ▷ *the updated observation set*

```
if ∃ o' = ⟨X, π', t', s'⟩ ∈ O with t' < t then
    ▷ there is a previous observation o' on X
    prevObs ← True
    if confidence(o', Now) ≤ confidence(o, Now) then ▷ confidence in the old
      observation is lower than in the new one
        O ← (O \ {o'})                ▷ remove the old information before managing
          compatibility
    else
        return O                      ▷ quit without any change on O
if isCompatible(O ∪ {o}, P, cType) then
    ▷ the new observation does not break the compatibility with P
    return O ∪ {o}
else
    ▷ find a subset of observations that restore compatibility with P
    s ← 1                             ▷ begins to check singletons
    found ← False
    while found = False and s ≤ size(O) do
        ▷ browse the subsets of O
        foreach S ⊂ O of size s do
            cMin ← 1                  ▷ maximum degree of confidence
            ▷ keep the subset with the lowest confidence degree among subsets of
              size s
            if isCompatible((O \ S) ∪ {o}, P, cType) then
                found ← True
                if confidence(S, Now) < cMin then
                    O' ← S
                    cMin ← confidence(S, Now)
        s ← s + 1
    ▷ a subset O' has been found
    if confidence(O', Now) < confidence(o, Now) then
        ▷ keeps the new obs when it is more reliable than O''
        O ← (O \ O') ∪ {o}
    else
        ▷ the observation set is kept unchanged
        if prevObs = True then
            O ← O ∪ {o'}              ▷ the previous observation is restored
    return O
```

observation o, either o is less reliable than $\mathbf{O}$, then $c(\mathbf{O} \cup \{o\}, t) = c(o, t) \geq \beta$; or o is more reliable than $\mathbf{O}$ then $c(\mathbf{O} \cup \{o\}, t) = c(\mathbf{O}, t) \geq \beta$.

(2) Let $\mathbf{O}$ be totally (resp. partially) compatible with P; from Proposition 2 (resp. Proposition 3), any subset of $\mathbf{O}$ is also totally (resp. partially) compatible with P, thus the removal of some observations have no impact on the compatibility of the resulting set. Concerning the addition of a new observation in the set, this is done only when the resulting set is totally (resp. partially) compatible with P (lines 7 and 15).

□

The following example illustrates the use of these operations to manage the observation set.

3.3. Illustrative example

We illustrate the process with an observation set related with a general model knowledge about fall prevention [3, 24]. We consider a scenario about Ms. A that continues Example 6. The evolution of the observation set is monitored over one year, showing the arrival of new information, their removal and the evolution of the confidence degree of each observation.

We have selected 12 variables from the real data set provided by the fall prevention service at the hospital of Lille (France) (Table 3). These variables include temporal aspects such as the presence of a specific event in the past or the number of repetitions of an event, about behavior and habits, and the current state of the person.

The variables have different temporal granularity, from some months up to several years. The variables have been regrouped in three subsets corresponding to the shape of their decay function. The variables of the first group rarely change and the associated confidence degree decreases very slowly. The life span of a piece of information is generally more than 3 years, when the initial confidence degree is high. At the opposite, in the third group, the decay functions decreases much faster, and the life span of an information on a variable is

Table 3: Description of the variables used in the example.

Group	X	$Dom(X)$	Meaning
1	ALO	(yes;no)	lives alone
	PARK	(no;probable;confirmed)	Parkinson disease
	DEM	(no;probable;confirmed)	dementia
2	ALC	(yes;no)	drinks alcohol
	GAIT	(yes;no)	gait impairment
	LS	(yes;no)	leg strength weakness
	ADL	([0,2];]2,4];]4,6])	autonomy scale
3	PSY	(true;false)	takes at least one psychotropic drug
	DEP	(yes;no)	depression
	UWA	(yes;no)	uses a walking aid
	FEAR	(yes;no)	fear of falling
	AVGO	(yes;no)	avoids going out for fear of falling

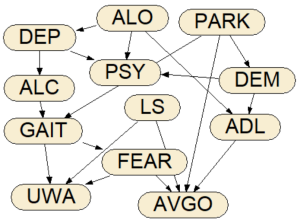

Figure 3: Graph of the Bayesian network used in the illustrative example.

generally less than a year. For those variables, changes may occur several times in a year, and an observation is kept only for a few months in the observation set. The variables in the second group fall between these two extremes.

The structure of the Bayesian network using the variables listed in Table 3 has been proposed for the purpose of this article without strong validation from the experts (Figure 3). However, the network parameters have been learned from a thousand of real cases from the fall prevention service using the counting learning algorithm used in NETICA [41].

In this study, the confidence threshold was set to 0.3, so that any information in the observation set has a higher confidence degree. Several sources of information provide observations about the variables. The two main sources are a doctor who specializes in falls prevention and the treating physician; the trust in what they say is very high (0.95). The third source is a member of the family

Table 4: The observation set with the potentials and the confidence degrees. New observations are in bold.

t	M1		M3		M5		M12	
X	π	$c(o,t)$	π	$c(o,t)$	π	$c(o,t)$	π	$c(o,t)$
ALO					**(1;0)**	**0.95**	(1;0)	0.85
PARK					**(0.9;0.1;0)**	**0.95**	(0.9;0.1;0)	0.85
DEM	**(1;0;0)**	**0.95**	(1;0;0)	0.94	(1;0;0)	0.93	(1;0;0)	0.89
ALC			**(0.9, 0.1)**	**0.4**	**(1;0)**	**0.95**	(1;0)	0.5
GAIT								
LS								
ADL								
PSY	**(0;1)**	**0.95**	(0;1)	0.75	**(1;0)**	**0.95**		
DEP	**(0.8;0.2)**	**0.95**	(0.8;0.2)	0.65	(0.8;0.2)	0.35		
UWA					**(0;1)**	**0.95**	**(0;1)**	**0.95**
FEAR							**(0;1)**	**0.95**
AVGO	**(0;1)**	**0.95**	(0;1)	0.75	(0;1)	0.55		

and is considered as much less trusty (0.4). An additional source of observations is a rumor; the trust in this source is low (0.1) and the related information is not integrated in the observation set.

The observation set is given in Table 4, with its evolution during four slices of time. New observations are written using bold characters. The confidence degree is re-evaluated at each time step.

At month M1, Ms A has an appointment in a specialized consultation since she has fallen with no apparent reason. The doctor inputs the following observations in the system: Ms A has no dementia (DEM = no); she took no psychotropic drug before that day (PSY = false); Ms A is likely to have an episode of depression, but the doctor is not completely sure of his diagnosis. According to him, Ms A is four times as likely to be in that state when she has a depression than when she does not. The doctor thus enters an uncertain observation with the likelihood ratio (0.8, 0.2); Ms A goes out as often has she used to before falling (AVGO = no).

At month M3, a member of family, considered as source of information much less trusty (0.4), reports that it seems very likely that Ms A drinks alcohol. The confidence degree of this new uncertain observation, even if low, is however above the confidence threshold, leading to add the observation in the set.

At M5, the treating doctor clearly remarks that Ms A drinks alcohol. The

confidence degree of this observation is updated in the observation set. The doctor also inputs several other observations about the variables ALO, PARK, UWA and PSY. A previous and different observation on the variable PSY was in the observation set. It is replaced by the new one which has a higher confidence degree.

After that, there is no new observation during several months, but the confidence degree on previous observations is regularly re-evaluated and the observations that are likely to be obsolete are removed. For example, previous observations on DEP, UWA and AVGO are removed since their confidence degree has become too low.

At M12, new observations coming from the doctor are entered about variables FEAR and UWA. An additional observation coming from a rumor states that Ms A no longer goes out. It is not integrated because of the very low level of trust associated to that source (0.1).

At any time, the observation set management system can provide information about Ms A about any variable of the model. When the question concerns a variable for which an observation is available in the observation set, the system provides the most probable value(s) in the potential of the observation, together with the confidence degree of that information. For example, at M12, the system can states with a high confidence degree that Ms A lives alone, has neither Parkinson disease nor dementia and that she does not use a walking aid and has no fear of falling. Another information is available about her alcohol consumption from a previous observation but the confidence degree is much lower (0.5) (see Table 4). When the question concerns a variable for which no observation is available, the system uses the general probabilistic knowledge embedded in the Bayesian network to update the belief on that variable based on available observations on other variables. It returns the posterior probability distribution on the variable.

In addition, the system can also return the marginal probability distribution of the same variable without any observation to allow the comparison of the specific case with the baseline population. For example, during the whole sce-

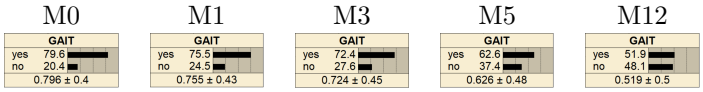

Table 5: Evolution of $P(GAIT \mid \mathbf{Y}_i = \mathbf{y}_i)$ where $\mathbf{Y}_i = \mathbf{y}_i$ represents the observation available at month Mi.

nario, there has been no observation about the variables GAIT, LS and ADL, and concerning the variables PSY, DEP and AVGO, some observations have been received but they have been removed from the observation set.

Table 5 displays the posterior probability distribution of the variable GAIT given the observations for Ms A at each time step of the scenario described above. When no observation is available about Ms A (at M0), the initial belief that she has gait impairment is 79.6%. This rate is the marginal probability of gait impairment in the population represented by the Bayesian network model[1]. At M12, the belief that Ms A has gait impairment has decreased down to 51.9%. This is the posterior probability $P(\text{GAIT} = yes \mid ALO = yes, o_{PARK}, DEM = no, ALC = yes, UWA = no, FEAR = no)$ where o_{PARK} is the uncertain observation on the variable PARK specified by the potential $\pi_{PARK} = (0.9, 0.1, 0)$. The comparison of the initial belief (79.6 %) with the current belief (51.9 %) allows to state that Ms A has much less risk to have gait impairment than the baseline population, but that the risk is however present.

This example illustrates the interest of the observation set management system. It allows to provide information about any variable of the model for a specific person, even when no observation has been collected about that variable. It also provides information about the confidence degree related with previous observations collected about that person.

[1]The Bayesian network model has been learned on the basis of a data set of a thousand of patients received in a department specialized in fall prevention. This population has a high rate of gait impairment

4. Discussion and related works

In this section, we discuss about different approaches of modeling and reasoning with uncertain, incomplete or inconsistent knowledge. First we start with the use of Dempter-Shafer theory for uncertainty modeling, then we continue with several works about uncertain information fusion, belief changes and temporal reasoning, and third, we briefly present Continuous Time Bayesian Networks and discuss the interest of using this knowledge model for the management of a time stamped uncertain observation set.

The last part of the section presents the steps toward a practical implementation of a knowledge based system for the management of an observation set with application on preserving mobility.

4.1. Representation and reasoning with uncertain knowledge

The question of modeling and reasoning with knowledge and uncertainty can be handled by several theories of uncertainty, among which the Bayesian probability theory and the Dempster-Shafer theory of belief functions [42, 43, 44]. According to Cobb [45], *"both frameworks have roughly the same expressive power (...) and can be utilized to model knowledge and evidence of varying types"*. However, *"computationally, D-S belief networks are more expensive to evaluate than Bayesian"*. Uncertain evidence can be specified in the theory of belief functions through masses assigned to the subsets of the domain of a variable. In this article, the domain knowledge is modeled using Bayesian network models, and uncertain evidence. In Definition 1, a time stamped uncertain observation is specified in the form of a likelihood vector, as proposed by Pearl [37].

Beyond the framework used for modeling and reasoning, we present below some works dealing with problems similar to ours.

In *belief revision*, an agent changes his beliefs about a static world in the light of new information, whereas in our problem, the world represents the state

of a person and its context, and it changes over time. In *belief update*, an agent keeps his beliefs up to date within an evolving world. This is also the case in our problem but the updating is made on the basis of time stamped uncertain observations and the focus is set on how to maintain that uncertain observation set. *Belief merging* studies how to aggregate multiple belief bases into a coherent one [46]. Our problem is related to multi-source information fusion and it raises the question of the compatibility of the observation set with a well-established knowledge base. **Several works propose the use of a confidence degree to manage the reliability of different sources [47] or to fuse several distributions of possibility from different sources with some contradictions [48]. In our proposal, the confidence degree is attached with a piece of information and evolves with time.** Information correction and fusion for belief functions has been studied when the information items may be irrelevant, and sources may lie [49]. In order to merge information on a same variable coming from several sources, the authors consider both the notion of truthfulness and the notion of relevance regarding a given question of interest, assuming that uncertain meta-knowledge on the source's relevance and truthfulness is available. The questions of compatibility concern the information brought by sources and assumptions regarding the epistemic state of the agent about the source state. In our work, we also assume that meta knowledge is available about the trust in sources. However, we consider information collected over a long period of time, while taking into account the fact that this information may have become obsolete over time. We evaluate the confidence that can be placed in old information and check the compatibility of specific uncertain information with a general probabilistic knowledge model.

Our system provides two kinds of information about a requested variable depending on whether an observation on the variable is in the observation set or not. The confidence degree relates to the trust we can have in an old observation, whereas the updated belief, thanks to the Bayesian network, is based on general probabilistic knowledge and specific observations. When a trusted observation of the variable is present, this is the information that is provided; when no trusted

observation exists, the other observations and the general knowledge are used. Combining these two pieces of information in a single one is challenging since the semantics are different. It would require to define how to use the confidence degree in the aggregation of the potential of an observation and the posterior probability of the observed variable.

In this paper, we propose to combine decay functions and a BN to re-evaluate the confidence in a piece of information collected in the past. Continuous Time Bayesian Networks (CTBNs) [18, 50] seem to constitute an interesting alternative to gather BN and decay functions in a single model. In CTBNs, the states of variables evolve continuously over time, the evolution of each variable depending on the state of its parents in the graph. The key difference between BNs and CTBNs is that CTBNs model the distribution over the variable's trajectories, considering a distribution over the dynamics of a variable $X(t)$ given the values of the parents of X at time t — the Conditional Intensity Matrices (CIM)—, whereas BNs use parameters which indicates the conditional distributions of a node X given its parents. The elements of the CIM apart from the diagonal denotes the rate departing from a state and arriving in another state in a given period. The time during which a variable stays in a state is usually assumed to be exponentially distributed. Each transition-rate matrix of a CTBN is associated with a time unit that is the reference to define the transition rates. Several algorithms of inference in CTBNs have been proposed [51, 52, 53, 54, 55], associated classifiers have been proposed with specific algorithms [56], and a non parametric approach is proposed in [57] to learn continuous-time Bayesian network in relational domain. A reasoning and learning engine for CTBN is also available[2] [58]. Different applications use CTBN [59]. The drawbacks that we see concern the difficulty to define the CTBN since it requires to identify for each variable of interest the variables that influence its evolution, and the initial BN because of the high number of parameters to adjust: a CTBN with n nodes, each of them having p parents and a domain of size m corresponds to

[2] http://rlair.cs.ucr.edu/ctbnrle/

$n.2^p.m(m-1)$ transition rates. Finally, Hybrid time BN is another new kind of probabilistic graphical model for dynamic systems [60]. It aims to reduce the amount of information required to build the model regarding temporal dimension. However, the model construction may become infeasible for large networks such as in CTBN.

These models could however help to decide whether an observation $X = x_i$ collected at time t should be kept or not in the information database. Indeed, when the value of the parents of X are known, the transition rate of the variable X from the value x to another value can be obtained from the CIM matrix of X by adding up the values of the row of x except the value on the diagonal. When the transition rate is below a threshold, we can decide to keep the information $X = x_i$ in the database, whereas in the opposite case, the risk for that information to be obsolete is too high and the information should be removed.

The proposal presented in this paper is a first approach guided by the desire to get a solution that can actually be implemented. This proposal can be considered as a particular simple case of a future proposal using CTBNs. Indeed, the Bayesian network model with the variables of the observation set of our actual proposition could be the first component of a CTBN as the initial distribution of the process. The decay functions proposed in the current approach can be seen as conditional intensities matrices (CIM) of a CTBN drastically simplified. Indeed, when the transition rates of a variable X do not depend on its parents' value, nor on the values of X, all CIMs of the variable X are identical, and composed of a single parameter (except on the diagonal). From this parameter, we can derive the probability that the value of X changes over a given amount of time, which is the same information provided by a decay function when it is linear. In the current paper, we propose a simpler approach since its implementation is based on a Bayesian network. However, when trying to respond effectively to the problem of falls prevention, the Bayesian network involves a large number of features and its construction requires important time for knowledge extraction and lots of data to learn the model. A validation of the current proposal with a Bayesian network involving a large number of characteristics is

a first step before considering the use of CTBNs whose size may quickly cause it to become intractable [50].

4.2. Toward practical implementation

The development of the system such as presented in Figure 2 requires several upstream steps, related to the construction of the Bayesian network, the knowledge about the decay functions according to the variables and about the trust in the sources. Most of these steps require the collaboration of domain experts to be achieved. After the definition of an ontology of the risk factors for fall [24], we selected the variables that constitute the support of the personal information base. These variables have been used to build the Bayesian network [25]. The definition of the decay functions is underway: the current step is to assign the variables to different categories regarding the evolution of the variables with time. The implementation of the algorithms proposed in this paper is a short-term perspective for which we have identified a small subset of variables to focus on.

The first intended users of our system are general practitioners since they are key players in the prevention of fall. They will provide new observations about their patients and benefit from the output of the system for unobserved variables and variables for which previous information may have changed. The development of the first prototype will thus be made in conjunction with some general practitioners related with the PREMOB network, whose main objective is the prevention of falls [61].

5. Conclusion

This study focuses on real-life situations where it is not possible to instantly gather all the up-to-date information needed to make a decision.It considers information collected over time and from a variety of sources, while taking into account the fact that some information may have become obsolete over time.

We address a real world problem in health context which is to get quickly the adequate information about a person, with sufficient quality and quantity to achieve specific purpose such as fall prevention. Indeed, the application objective is to manage a set of patient specific observations in complement to the generic knowledge embedded in a Bayesian network.

Our contribution includes three main points:

- **the proposition of a *knowledge based system architecture and algorithms* to manage information quality in a time stamped uncertain observation set;**
- **the definition of the *confidence degree* of a piece of observation (respectively of an observation set) to reduce the risk of erroneous information in a personnal observation set;**
- **the definition of the partial / total *compatibility of an uncertain observation* (respectively an observation set) with a joint probability distribution to maintain the consistency of the observation set regarding a probabilistic model of knowledge.**

We present a part of our application context about fall preventing among elderly people, to illustrate how to maintain a set of observations about a subject, with only a few new inputs at each time step. The compromise between quality and quantity allows to benefit from old observations and to remove or reject those in which remains too little confidence.

Our proposal is guided by the criterion of feasibility in order to implement a first prototype as part of a real system dedicated to fall prevention. **Some work remains to be done, regarding both theoretical and applicative aspects. Theoretical future work concerns: how the system can learn to adjust its trust in the information sources; the merging of several pieces of uncertain information for the same variable, coming from different sources at different time, is also an interesting perspective,**

instead of keeping only the information with the highest confidence degree. CTBN is an alternative to be explored provided that the construction of the model is feasible. In addition, uncertain evidence in CTBN [62] could be interesting to manage the uncertainty coming from the trust in sources, requiring additional hypotheses as explained above.

Applicative future work concerns the knowledge embedded in the system and more specifically decay functions. The implementation of the `maintain` algorithm may raise computational questions related with the total or partial compatibility. The development and the evaluation of the maintenance system of an observation set for a real application will involve final users.

References

[1] C. Combi, E. Keravnou-Papailiou, Y. Shahar, Temporal Information Systems in Medicine, Springer, Boston, MA, 2010.

[2] American Geriatrics Society, British Geriatrics Society, American Academy of Orthopaedic Surgeons Panel on Falls Prevention, Guideline for the prevention of falls in older persons, Journal of the American Geriatrics Society 49 (5) (2001) 664–72.

[3] H. Bourdessol, S. Pin, Prevention of falls in the elderly living at home, Good Practice Guide, collection Référentiels, INPES, Saint-Denis, France, 2008, original French version published in 2005.

[4] D. Dubois, W. Liu, J. Ma, H. Prade, The basic principles of uncertain information fusion. An organised review of merging rules in different representation frameworks, Information Fusion 32 (PartA) (2016) 12–39.

[5] O. Papini, Knowledge-base revision, Knowledge Engineering Review 15 (4) (2000) 339–370.

[6] D. Dubois, S. Moral, H. Prade, Belief change rules in ordinal and numerical uncertainty theories, in: D. Gabbay, P. Smets (Eds.), Belief Change,

(D. Dubois, H. Prade, eds.), Vol. 3 of the Handbook of Defeasible Reasoning and Uncertainty Management Systems, Kluwer Academic Publishers, Dordrecht, 1998, pp. 311–392.

[7] M. Fisher, D. Gabbay, L. Vila, Handbook of Temporal Reasoning in Artificial Intelligence, Elsevier, The Netherlands, 2005.

[8] P. Z. Revesz, On the semantics of theory change: Arbitration between old and new information, in: Proceedings of the Twelfth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, PODS'93, ACM, New York, NY, USA, 1993, pp. 71–82.

[9] C. Chen, Y. Chen, Y. Sun, M. Chen, Life cycle modeling of news events using aging theory, in: Proceedings of 14th European Conference of Machine Learning (ECML'03), 2003, pp. 47–59.

[10] J. Paik, Parameterized decay model for information retrieval, ACM Transactions on Intelligent Systems and Technology 7 (3) (2016) 27:1–27:21.

[11] M. Sendi, M. Omri, M. Abed, Discovery and tracking of temporal topics of interest based on belief-function and aging theories, Journal of Ambient Intelligence and Humanized Computing (2018) 3409–3425.

[12] G. Cormode, S. Tirthapura, B. Xu, Time-decaying sketches for sensor data aggregation, in: Proceedings of PODC'07, ACM symposium on Principles of distributed computing, Portland, Oregon, USA, 2007, pp. 215–224.

[13] J. Allen, Maintaining knowledge about temporal intervals, Communications of the ACM 26 (11) (1983) 832–43.

[14] J. Allen, Towards a general theory of action and time, Artificial Intelligence 23 (2) (1984) 123–54.

[15] M. van der Heijden, P. Lucas, Describing disease processes using a probabilistic logic of qualitative time, Artificial Intelligence in Medicine 59 (3) (2013) 143–155.

[16] K. Orphanou, A. Stassopoulou, E. Keravnou, Temporal abstraction and temporal Bayesian networks in clinical domains: A survey, Artificial Intelligence in Medicine 60 (3) (2014) 133–149.

[17] K. Murphy, Dynamic Bayesian networks: Representation, inference and learning, Ph.D. thesis, UC Berkeley, Computer Science Division (2002).

[18] U. Nodelman, C. Shelton, D. Koller, Continuous time Bayesian networks, in: Proceedings of the 18th Conference in Uncertainty in Artificial Intelligence (UAI'02), University of Alberta, Edmonton, Alberta, Canada, 2002, pp. 378–387.

[19] M. Ramati, Y. Shahar, Irregular-time Bayesian networks, in: Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI'10), Catalina Island, California, USA, 2010, pp. 484–491.

[20] M. Liu, F. Stella, A. Hommersom, P. Lucas, L. Boer, E. Bischoff, A comparison between discrete and continuous time Bayesian networks in learning from clinical time series data with irregularity, Artificial Intelligence in Medicine 95 (2019) 104–117.

[21] A. Tucker, V. Vinciotti, X. Liu, D. Garway-Heath, A spatio-temporal Bayesian network classifier for understanding visual field deterioration, Artificial Intelligence in Medicine 34 (2) (2005) 163–177.

[22] A. Tucker, P. 't Hoen, V. Vinciotti, X. Liu, Temporal Bayesian classifiers for modelling muscular dystrophy expression data, Intelligent Data Analysis 10 (5) (2006) 441–455.

[23] L. Cattelani, F. Chesani, L. Palmerini, P. Palumbo, L. Chiari, S. Bandinelli, A rule-based framework for risk assessment in the health domain, International Journal of Approximate Reasoning 119 (2020) 242–259.

[24] V. Delcroix, F. Essghaier, K. Oliveira, P. Pudlo, C. Gaxatte, F. Puisieux, Towards a fall prevention system design by using ontology, in: atelier INOVIVE, IC, Plate-Forme PFIA, Toulouse, France, 2019.

[25] G. Sihag, V. Delcroix, E. Grislin, X. Siebert, S. Piechowiak, F. Puisieux, Prediction of risk factors for fall using bayesian networks with partial health information, in: AIdSH: International Workshop on AI-driven Smart Healthcare, IEEE Global Communications Conference, GLOBECOM 2020, Virtual Event, Taiwan, 2020.

[26] C. Batini, M. Scannapieco, Data and Information Quality - Dimensions, Principles and Techniques, Data-Centric Systems and Applications, Springer, 2016.

[27] G. De Tré, R. De Mol, A. Bronselaer, Handling veracity in multi-criteria decision-making: A multi-dimensional approach, Information Sciences 460-461 (2018) 541–554.

[28] J. Olson, Data Quality: The Accuracy Dimension, Morgan Kaufmann Publishers, 2003.

[29] L. Pipino, Y. Lee, R. Wang, Data quality assessment, Commun. ACM 45 (2) (2002) 211–218.

[30] P. Shamala, R. Ahmad, A. H. Zolait, M. Sedek, Integrating information quality dimensions into information security risk management (ISRM), Journal of Information Security and Applications 36 (2017) 1–10.

[31] M. Ziad, R. Y. Wang, Y. W. Lee, Data Quality, Kluwer Academic Publishers, 2000.

[32] R. Nouvel, M. Zirak, V. Coors, U. Eicker, The influence of data quality on urban heating demand modeling using 3d city models, Computers, Environment and Urban Systems 64 (2017) 68–80.

[33] Z. Wang, S. Dagtas, J. Talburt, A. Baghal, M. Zozus, Rule-based data quality assessment and monitoring system in healthcare facilities, in: F. Y. Lau, J. A. Bartle-Clar, G. Bliss, E. M. Borycki, K. L. Courtney, A. M. Kuo, A. Kushniruk, H. Monkman, A. V. Roudsari (Eds.), Improving Usability, Safety and Patient Outcomes with Health Information Technology

- From Research to Practice, Information Technology and Communications in Health Conference, ITCH 2019, Victoria, BC, Canada, 14-17 February 2019, Vol. 257 of Studies in Health Technology and Informatics, IOS Press, 2019, pp. 460–467.

[34] M. Zolnoori, M. D. Williams, W. B. Leasure, K. B. Angstman, C. Ngufor, A systematic framework for analyzing observation data in patient-centered registries: Case study for patients with depression, Journal of Medical Internet Research (JMIR), Research Protocols 9 (10) (2020) e18366.

[35] B. Heinrich, M. Klier, A. Schiller, G. Wagner, Assessing data quality – a probability-based metric for semantic consistency, Decision Support Systems 110 (2018) 95–106.

[36] Y. Timmerman, A. Bronselaer, Measuring data quality in information systems research, Decision Support Systems 126 (2019) 113–138, perspectives on Numerical Data Quality in IS Research.

[37] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, Morgan Kaufmann, 1988.

[38] A. Ben Mrad, V. Delcroix, S. Piechowiak, P. Leicester, M. Abid, An explication of uncertain evidence in Bayesian networks: likelihood evidence and probabilistic evidence - uncertain evidence in Bayesian networks, Applied Intelligence 43 (4) (2015) 802–824.

[39] H. Nunoo-Mensah, K. Boateng, J. Gadze, The adoption of socio- and bio-inspired algorithms for trust models in wireless sensor networks: A survey, International Journal of Communication Systems 31 (7) (2018).

[40] B. Israelsen, N. Ahmed, "Dave...I can assure you ...that it's going to be all right ..." A Definition, Case for, and Survey of Algorithmic Assurances in Human-autonomy trust relationships, ACM Computing Surveys 51 (6) (2019) 113:1–113:37.

[41] Norsys, Netica Application, (Version 5.12) [Computer software], Norsys Software Corp., Vancouver, BC, Canada (2014).
URL http://www.norsys.com

[42] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist. 38 (2) (1967) 325–339.

[43] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976, 246 pages.

[44] P. Smets, R. Kennes, The transferable belief model, in: R. Yager, L. Liu (Eds.), Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 693–736.

[45] B. R. Cobb, P. P. Shenoy, A comparison of bayesian and belief function reasoning, Information Systems Frontiers 5 (4) (2003) 345–358.

[46] J. Ma, W. Liu, A framework for managing uncertain inputs: An axiomization of rewarding, International Journal of Approximate Reasoning 52 (7) (2011) 917–934.

[47] W. Xu, J. Yu, A novel approach to information fusion in multi-source datasets: A granular computing viewpoint, Information Science 378 (2017) 410–423.

[48] F. Delmotte, P. Borne, Modeling of reliability with possibility theory, IEEE Trans. Syst. Man Cybern. Part A 28 (1) (1998) 78–88.

[49] F. Pichon, D. Dubois, T. Denoeux, Relevance and truthfulness in information correction and fusion, International Journal of Approximate Reasoning 53 (2) (2012) 159–175.

[50] L. Perreault, J. W. Sheppard, Compact structures for continuous time bayesian networks, International Journal of Approximate Reasoning 109 (2019) 19–41.

[51] U. Nodelman, D. Koller, C. Shelton, Expectation propagation for continuous time Bayesian networks, in: Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence (UAI'05), Edinburgh, Scotland,, 2005, pp. 431–440.

[52] Y. Fan, J. Xu, C. Shelton, Importance sampling for continuous time Bayesian networks, Journal of Machine Learning Research 11 (2010) 2115–2140.

[53] E. B. Celikkaya, C. R. Shelton, W. Lam, Factored filtering of continuous-time systems, in: Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence (UAI'11), Barcelona, Spain, 2011, pp. 61–68.

[54] V. Rao, Y. Teh, Fast MCMC sampling for Markov jump processes and continuous time Bayesian networks, in: Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence (UAI'11), Barcelona, Spain, 2011, pp. 619–626.

[55] I. Cohn, T. El-Hay, N. Friedman, R. Kupferman, Mean field variational approximation for continuous-time Bayesian networks, Journal of Machine Learning Research 11 (2010) 2745–2783.

[56] F. Stella, Y. Amer, Continuous time Bayesian network classifiers, Journal of Biomedical Informatics 45 (6) (2012) 1108–1119.

[57] S. Yang, T. Khot, K. Kersting, S. Natarajan, Learning continuous-time Bayesian networks in relational domains: A non-parametric approach, in: Proceedings of the 13th AAAI Conference on Artificial Intelligence, Phoenix, Arizona, USA, 2016, pp. 2265–2271.

[58] C. Shelton, Y. Fan, W. Lam, J. Lee, J. Xu, Continuous time Bayesian network reasoning and learning engine, Journal of Machine Learning Research 11 (2010) 1137–1140.

[59] J. Xu, C. Shelton, Intrusion detection using continuous time Bayesian networks, Journal of Artificial Intelligence Research 39 (2010) 745–774.

[60] M. Liu, A. Hommersom, M. van der Heijden, P. J. F. Lucas, Hybrid time bayesian networks, International Journal of Approximate Reasoning 80 (2017) 460–474.

[61] L. Mirakovska, F. Bloch, P. Chassagne, C. Levasseur, E. Greboval, P. Lescure, C. Delecluse, M. Welter, N. Chastan, C. Chavoix, P. Pudlo, S. Pelayo, A. Thevenon, V. Tiffreau, F. Puisieux, PREMOB : Réseau d'investigation clinique de l'interrégion nord-ouest sur la thématique : "prévenir la perte de mobilité et les chutes chez les personnes âgées ", Neurophysiologie Clinique 49 (6) (2019) 446–447, XXVIe congrès de la Société Francophone Posture Équilibre et Locomotion (SOFPEL) à Montréal, 4 et 5 décembre 2019.

[62] L. Sturlaugson, J. Sheppard, Uncertain and negative evidence in continuous time Bayesian networks, International Journal of Approximate Reasoning 70 (2016) 99–122.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: