Fuzzy Evidence in Bayesian Network

Ali Ben Mrad
University of Sfax, ENIS
ENIS, CES
Sfax, Tunisia
benmradali2@gmail.com

Mohamed Amine Maalej
University of Sfax, FSS
ENIS, CES
Sfax, Tunisia
m.maalej@gmail.com

Veronique Delcroix
Université Lille Nord de France, F-59000, Lille, France
UVHC, LAMIH FRE CNRS 3304, F-59313,
Valenciennes, France
Veronique.Delcroix@univ-valenciennes.fr

Sylvain Piechowiak
Université Lille Nord de France, F-59000, Lille, France
UVHC, LAMIH FRE CNRS 3304, F-59313,
Valenciennes, France
Sylvain.Piechowiak@univ-valenciennes.fr

Mohamed Abid
University of Sfax, ENIS
ENIS, CES
Sfax, Tunisia
mohamed.abid_ces@yahoo.fr

***Abstract*— The Bayesian Networks are graphical models that are easy to interpret and update. These models are useful if the knowledge is uncertain, but they lack some means to express ambiguity. To face this problem, we propose Fuzzy Evidence in Bayesian Networks and combine the Fuzzy Logic and Bayesian Network. This has allowed to benefit from mutual advantages of these two approaches, and to overcome the problem of data and observation ambiguity. This paper proposes an inference algorithm which uses the Bayesian Network and Fuzzy Logic reliability. This solution has been implemented, tested and evaluated in comparison with the existing methods.**

Keywords- Artificial Intelligence, Bayesian Networks, Fuzzy Logic, Uncertain Evidence.

I. INTRODUCTION

Bayesian Network (BN) [1, 2] is a model of uncertain knowledge representation and reasoning which is based on probabilities and graph theory. It was formulated by J. Pearl [3], and it was often used for disposing of the uncertain information in Artificial Intelligence.

With the development of the structure construction and inference algorithm of BN, it has been successfully used [4] in medical diagnosis [5], image recognition, language understanding, data mining, genetics, and search algorithms.

However, context inference using BNs has a limitation that is the inability to deal with the diverse information effectively. Since most of practical tools dealing with BNs require discrete inputs, loss of information might happen and it may not reflect the context appropriately. This limitation has been overcome by utilizing the fuzzy system.

Fuzzy system [6, 7] is derived from the conception proposed by Pearl [8]. In a fuzzy system, a variable's state is represented by a set of fuzzy values (FVs). Because fuzzy systems do not force the classical model to discrete continuous state, they are often more robust in the face of noise.

Integrating fuzzy theory with Bayesian network, this paper discusses an effective method of uncertain knowledge representation, reasoning it in details.

The purpose of this work is three folds: to introduce the problem of Fuzzy information in BN, then model some systems by BN with fuzzy evidence and finally to develop an inference algorithm exploiting the wealth of these models.

This paper tries to solve some problems noticed during the diagnosis with classic Bayesian network. These problems concern essentially the modeling which turns out to be difficult in the presence of fuzzy knowledge. However, the algorithms of inference and diagnosis are not generic and adapted to the fuzzy context.

II. BAYESIAN NETWORK AND THEIR INFERENCE

A. Bayesian Network

A Bayesian network is represented by a directed acyclic graph (DAG) which is composed of a set of nodes, each of which representing a variable. Each edge connects two nodes, and has a direction assigned to it ([2] for more details).

An edge between two nodes indicates a relation between the nodes and the direction indicates the causality. If a node has a known value, it is said to be an evidence node. The joint distribution of the whole set of variables is given by the product of local distribution on each variable. The local distribution of each variable is a function of all its parents and is stored as a conditional probability table (CPT).

Inference in a Bayesian network involves determining the probabilities of the query variables, given the state of the evidence variables. When we get new information about variables in the network, we update the conditional probability tables to reflect it. This updating is known as evidence propagation [9]. Once all the beliefs are updated, the conditional probability tables contain the most recent beliefs in any variable and can be queried like a simple database to evaluate probabilities.

There are two main approaches for computing posterior probabilities in a Bayesian network: approximate inference and exact inference. The Junction tree algorithm is one of the most popular exact inference algorithms.

B. *Junction tree algorithm*

This algorithm is applicable in all types of networks, tree or non-tree. First and foremost, it transforms the graph into a junction tree, then it initializes its potential, next it uses the method message passing for propagation messages and compute posterior probabilities.

The algorithm behaves in the following way [10]:

- Graphical Transformation: Transform the directed acyclic graph (DAG) of a belief network into a join tree structure.
- Initialization: Initialize potentials, clusters and separators.
- Global Propagation: Perform an ordered series of local manipulations, called message passes, on the join tree potentials. The message passes rearrange the join tree potentials so that they become locally consistent, thus, the result of global propagation is a consistent join tree.
- Marginalization: From the consistent join tree compute the probability distribution P(V) for each variable of interest V.

The classical observation in BN touches only one state of the observed node, in other words, if we have evidence in a node X, only one of its states will be observed. So, we associate the value 1 with this state and 0 with the others.

However, we cannot represent the case where our evidence is not exact: that is if our observation is uncertain and concerns at the same time several states according to well defined percentages.

III. UNCERTAIN EVIDENCE

Regular evidence, called also hard evidence, is an observation of a random variable, having a particular value.

However, it is not always possible to observe the value of a variable in a particular case, or to have a complete trust on a claimed observation, thus bringing uncertainty to the evidences.

Evidences may be uncertain for various reasons. A reported observation may not be totally trusted due to errors or noise in the observation or reporting process; it may be biased due to the observer's preference; it may not hold when the time or location is different.

In fact, we focus on two types of uncertain evidences, virtual evidence which reflecting the uncertainty one has about a reported observation and soft evidence which reflecting the uncertainty of an event one observes.

A. *Virtual Evidence*

Virtual evidence (VE), first introduced by Pearl [3], offers a principled and convenient way of incorporating external knowledge into Bayesian networks as indicate in [11, 12]. Thus, there are several reasons why VE be valid.

For example, there are scenarios where we can reason only about ratios of likelihoods rather than the absolute values of the likelihoods themselves. Since the ratios are all that matters for computation, we are comfortable when this occurs.

VE can be interpreted as evidence with uncertainty, and is represented as a likelihood ratio. Virtual evidence utilizes a likelihood ratio to represent the observer's strength of confidence toward the observed event. Likelihood ratio *L*(*Y*) is defined as:

$L(Y)=P(ob(y_{(1)})|y_{(1)}) :P(ob(y_{(2)})|y_{(2)}) :\ldots:P(ob(y_{(n)})|y_{(n)})$,

where $y_{(1)}, y_{(2)}, \ldots, y_{(n)} \in Y$ are all instantiations of Y,

$ob(y_{(i)})$ denotes the event that we observed $Y = y_{(i)}$ is *True*, and $P(ob(y_{(i)})|y_{(i)})$ is interpreted as the probability we observe if *Y* is indeed in state $y_{(i)}$.

In the example detailed in [3], we have a BN with three initial nodes H for Burglary, S for Alarm Sound and W for Watson's Testimony, as indicated in Figure 1.

The ultimate goal is to compute the probability of H (a burglary occurred) given knowledge about H, S or W.

However, the only thing that is known is that Mr. Holmes receives a call from another neighbor, Mrs. Gibbons, saying she may have heard the alarm of his house being triggered, and concludes that the chance that Mrs. Gibbons did hear the alarm triggered is four times more.

Thus, the virtual evidence appears with added the variable G, which represents Gibbon's Testimony, and the directed edge S→G, where the value g of G corresponds to the virtual event, and the Conditional Probability Table of G is assigned such that P(g | s) = 4* P(g | not(s)).

For example, we can assign P(g | s) = 0.4 and P(g | not(s)) = 0.1. Then, we assert the observation *G* = g in the belief network, thus we can easily compute the posterior probabilities of any nodes.

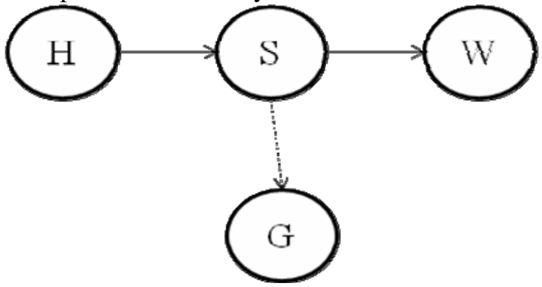

Figure 1. Example for Belief Updating with Virtual Evidence

Unlike virtual evidence, which can be interpreted as evidence with uncertainty, and is represented as a likelihood ratio, soft evidence can be interpreted as evidence of uncertainty, and is represented as a probability distribution of one or more variables.

B. *Soft Evidence*

Soft evidence was introduced by Valtorta [13]. This type of uncertain observation is characterized by the probability distribution R(Y), Y ⊆ X, where Y is a subset of the set of variables X.

Thus, it can be useful in several cases. As mentioned in [12], one may not be able to observe the precise state of a variable but we can know the probability distribution of its states.

Also, sometimes it is more important to know the probability distribution of variable states then its precise state. In addition, when two BNs interact with each other, the information exchanged between them is often in the form of probability distribution of the shared variable.

Concerning soft evidence, there is uncertainty in the choice of the observed state Xi, but we are sure of the probability distribution R(Xi) which will be reserved for updating belief.

However, soft evidence will be treated the same way as hard evidence. In fact, regular evidence which requiring (Xi = a) is a special case of soft evidence (R(Xi = a) = 1, R(Xi = b) = 0 for all states b ≠ a).

Unlike hard evidence, which is the observation of one of node states, soft evidence introduces the uncertainty of the observation. Hence, this concept consists to propagating quantified values, as indicated in Figure 2, for every state in the observed node.

To better explain this concept, we can illustrate an example shown in Figure 2 and detailed in [14].

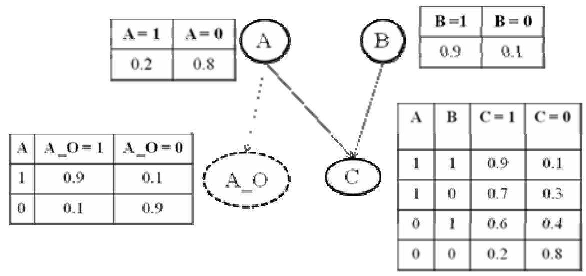

Figure 2. Example for Belief Updating with Soft Evidence

Thus, if observation is made on node A with 90 % for the state 1 and 10 % for the state 2 then the quantified values behaves as follows. Let Q(X) be the distribution resulted from updating P(X) by R(Y) , Y ⊆ X :

Q(A=1) = Q(A=1|Ao=1)
= [P(Ao=1|A=1) * P(A=1)]
/ [P(Ao=1|A=1) * P(A=1) + P(Ao=1|A=0) * P(A=0)]
= [0.9 * 0.2] / [0.9 * 0.2 + 0.1 * 0.8]
= 0.6923.
Q(A=0) = Q(A=0|Ao=1)
= [P(Ao=1|A=0) * P(A=0)]
/ [P(Ao=1|A=1)*P(A=1) + P(Ao=1|A=0)* P(A=0)]
= [0.1 * 0.8] / [0.9*0.2 + 0.1 * 0.8]
= 0.3077.

So far, we can classify uncertain evidence into two categories: virtual evidence and soft evidence. However, there are cases where the evidence is ambiguous. In this context we propose a new type of uncertain evidence called Fuzzy evidence.

IV. OUR ALGORITHM FOR THE FUZZY EVIDENCE IN BAYESIAN NETWORK AND THE SOFTWARE SOLUTION

A. *Fuzzy Evidence*

Soft evidence is based on the probability theory however fuzzy evidence is based on the fuzzy logic theory. In fact these two theories are different although both describe a notion of doubt and uncertainty as shown in table 1.

TABLE 1. COMPARISON BETWEEN PROBABILITY THEORY AND FUZZY LOGIC

Probability	Fuzzy logic
Uncertainty about the occurrence of an event	Ambiguity over the nature of an event
$P(A \cup not(A)) = P(U) = 1$ $P(A \cap not(A)) = P(\emptyset) = 0$	$\mu(A \cup not(A))$ is not necessarily $= \mu(U)$ $\mu(A \cap not(A))$ is not necessarily $= \mu(\emptyset)$

Thus, probability expresses the uncertainty about the occurrence of an event while fuzzy logic is the ambiguity over the nature of an event.

Mathematically, probability of the union of a set with its complement is equal to the probability of the universe and equal to 1, while probability of their intersection is equal to the probability of the empty set and equal to 0.

In fuzzy logic, the membership degree (μ) of the union of a set with its complement is not necessarily equal to the membership degree of the universe and the membership degree of their intersection is not necessarily equal to the membership degree of the empty set.

Thus, we can classify again the uncertain evidences into two categories: lack of precision and randomness. On the one hand, the problem of the randomness that is caused by unpredictable events is solved by soft evidence. On the other hand, to address the problem of lack of precision or ambiguity that is caused by the less defined concept of observation, we use rather fuzzy evidence.

To develop an algorithm of propagation of fuzzy evidence we use a junction tree algorithm combined with modified soft evidence.

However, we can summarize the algorithm in following steps:

1. For every state i of the observed node, we make a classic inference by observing the state i.
2. For every node N in the BN
    For every state j of the node N
    State j = ∑ Value of the state j stemming from the inference i * membership degree of observed value in state i
    End For
  End For

B. *The software solution*

The application of BN that we develop is a flexible and user-friendly graphical interface. It is implemented in C#, therefore the framework is easy to maintain and extend.

This application provides us with saving options and opening a constructed network, as well as it choosing the inference algorithm and format of saving network.

We have presented above the steps of junction tree algorithm in our editor. Firstly, we begin with the creation of the initial graph on which we are going to apply our algorithm.

When we finish the creation of graph with the available tools in our editor, we admit the probability on every node. These probabilities are described by means of an expert or from an information source.

Then, we select our pieces of evidence if they exist, and choose the inference algorithm.

At this step, the page, in which the algorithm is going to take place, closes our opened network to allow seeing changes brought on our graph by the algorithm.

Thus, we can see two opened networks, one containing initial graph and the other taking place in the algorithm.

Finally we can see the results of this algorithm after realizing the steps: Moralization, Triangulation, Junction tree and Marginalization.

C. *Representation of fuzzy data*

To spread the use of our editor to the Bayesian Network with Fuzzy Evidence, we can reach the module of variables representation (see Figure 3).

This interface includes a table containing the various data connected to the fuzzy variable such as its name, its function, the interval of function definition and the color which will be chosen in its graphic representation. In this table, we can add, insert or delete lines.

Concerning fuzzy variable functions, they include several types of mathematical functions in particular the negative numbers, the brackets, the square roots, the powers, the sine, the cosine, the tangent, +, -, * and /.

To better visualize the fuzzy subsets, we have a "Graphic Representation" accessible by "Range", "Screen" and "Graph".

Finally, when we have evidence, we can seize the observed value "Observed Value" in the field and obtain the membership degrees of every fuzzy variable in the second table.

V. EXAMPLE

To illustrate the process of building a Fuzzy Evidence in Bayesian Network, consider the problem of detecting credit-card fraud detailed in [15].

We begin by determining the variables to model. One possible choice of variables for our problem is Fraud (F), Gas (G), Jewelry (J), Age (A), and Sex (S), representing whether or not the current purchase is fraudulent, whether or not there was a gas purchase in the last 24 hours, whether or not there was a jewelry purchase in the last 24 hours, and the age and sex of the card holder, respectively.

In a realistic problem, we could model the states of one or more of these variables at a finer level of detail. For example, we could let Age be a continuous or a fuzzy variable as indicated in [15].

However, the variable "Age" has three states < 30, between 30 and 50, and > 50 as presented in [15]. In this case, observed "Age" values within the same interval in edges or in intervals centers will be treated in the same manner. Hence, the use of membership degrees can be an interesting solution.

Unfortunately, classic BN does not contain this type of variable, thus the importance of our editor appears.

At this step, we intervene to surmount ambiguity by representing the variable Age as fuzzy variable, and defining the functions of specific memberships of the fuzzy subsets "Young", "Adult" and "Old" characterized by well determined functions which can be seized in the specific table of the fuzzy variable (see Figure 3).

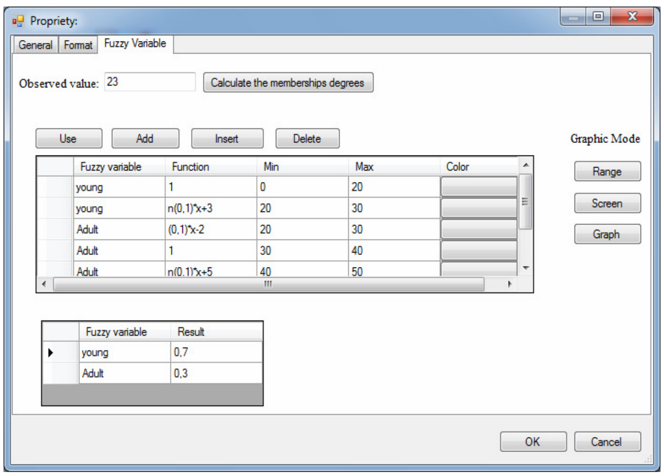

Figure 3. Representation of the fuzzy variable Age

Let us suppose that the observed age is 23 years, we can seize this value directly in the specific field, and then we obtain the membership degrees of the observed value in each of the fuzzy subsets as indicated in Figure 3.

Finally, we obtain the final results supplied by this algorithm. So, we can move towards the node chosen in order to show its probability as a posteriori (see Figure 4).

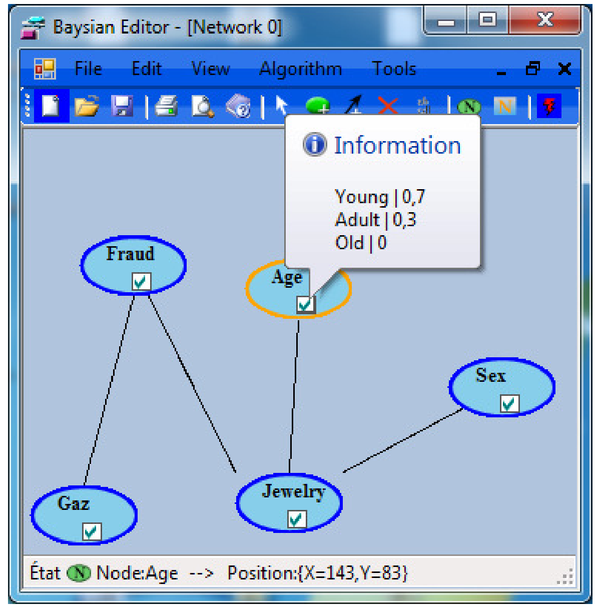

Figure 4. *Final Result*

This last section presents our editor dedicated to the presentation and the inference in BN with fuzzy evidence. The objective of this editor is to illustrate the progress of a diagnosis from the modeling by way of the inference and reaching the results.

The modeling of the fuzzy node appeals in particular to the membership degrees of fuzzy subsets, concerning inference; it is based on the algorithm JLO [16, 9] by taking into account the memberships degrees of the observed value. These two aspects are presented through an example, which is detecting credit-card fraud problem.

VI. CONCLUSION

This paper extends traditional Bayesian networks into Bayesian networks with fuzzy evidence in which fuzzy variables and classical ones may appear.

Moreover, in this study, we shed to light on this theory by detailing its theoretical foundation and its advantages.

Inspired by the results of all the research work carried out, we were able first to model some systems by BN with fuzzy evidence, which allowed us to develop a graphic editor for BN with an additional module for fuzzy variable representation.

Besides developing an algorithm exploiting the wealth of this recent theory, we test and compare this algorithm which we developed by illustrative examples and in view of other work.

REFERENCES

[1] P. Naïm, P. Wuillemin, P. Leray, O. Pourret, and A. Becker, "Réseaux bayésiens", 2004.

[2] F. Jensen, "An Introduction to Bayesian Networks", Springer, New York, 1996.

[3] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Morgan Kaufmann, San Mateo, CA, 1988.

[4] O. Pourret, P. Naim and B. Marcot. Bayesian Networks: "A Practical Guide to Applications". Chichester, UK: Wiley. ISBN 978-0-470-06030-8, 2008.

[5] S. Russell and P. Norvig. "Artificial Intelligence: A modern approach". Prentice Hall, 1995.

[6] D. Dubois and H. Prade, "An introduction to fuzzy systems" Clin Chim Acta, vol. 270, pp. 3-29, 1998.

[7] R. Kruse, J. Gebhardt, F. Klawonn, "Foundations of Fuzzy Systems, ISBN: 978-0-471-94243-6, 1994.

[8] L. Zadeh, "Fuzzy sets", Information and Control, Vol.8, 1965, 338-353.

[9] S. L. Lauritzen and D. J. Spiegelhalter. "Local computations with probabilities on graphical structures and their application to expert systems". International Journal of the Royal Statistical Society, pages 157.224, 1988.

[10] C. Huang and A. Darwiche. "Inference in belief networks: A procedural guide". International Journal of Approximate Reasoning, 15(3) : 225-263, 1996.

[11] X. Li, "On the Use of Virtual Evidence in Conditional Random Fields", Empirical Methods in Natural Language Processing-EMNLP, pp. 1289-1297, 2009.

[12] Y. Peng, S. Zhang, and R. Pan, "Bayesian Network Reasoning with Uncertain Evidences". International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2010.

[13] M. Valtorta, Y. Kim and J. Vomlel, "Soft evidential update for probabilistic multiagent systems", International Journal of Approximate Reasoning 29(1) (2002) 71–106.

[14] H. Tu, J. Allanach, S. Singh, K. R. Pattipati, and P. Willett. "The Adaptive Safety Analysis and Monitoring System". Electrical and Computer Engineering Department, University of Connecticut, 2004.

[15] D. Heckerman. "Bayesian Networks for Data Mining. Data Mining and Knowledge Discovery", Springer Netherlands. November 2004.

[16] Jensen, F., Lauritzen, S., and Olesen, K.. "Bayesian updating in recursive graphical models by local computations". Computational Statistical Quaterly, 4 :269–282, 1990.