



Towards Probabilistic Ontology based on Bayesian Networks

ANOUAR TADLAOUI Mouenis¹; KHALDI Mohamed²; AAMMOU Souhaib³

^{1,2,3}LIROSA / Faculty of Sciences,
Abdelmalek Essaâdi University,
MOROCCO

Abstract: *The increase and diversification of information has created new user requirements. The problems of representation, processing, analysis and reasoning on the information, especially the uncertain information, are still an important research topic. We try to show in this paper how this issue is crucial. It is essential to propose new approaches to formally represent uncertain information to help machines understand and infer new knowledge. At first, we will start by defining the Bayesian networks, the way of its construction. Then we will focus on probabilistic ontologies and their relationship with Bayesian Networks. Then, we will show how we can integrate a Bayesian network into a probabilistic ontology citing the translation rules which is the first step of the process of integration. Finally we conclude with perspectives of our work.*

Keywords: *Bayesian networks, ontology, probabilistic ontology, uncertainty.*

I. Introduction

In previous years, various researchers ([1], [2], [3] and [4]) tried to suggest approaches to incorporate uncertainty in ontologies, which is an interesting line of research in which this work is situated. One of the characteristics of Bayesian networks is their ability to represent uncertain information in the form of a probabilistic model.

But unfortunately Bayesian networks are not a formal representation and the machines are not capable of understanding it. For this reason, we try to propose a method to formally represent a Bayesian network as a probabilistic ontology using a standard representation of OWL (Web Ontology Language).

This formal representation can be used later as a support to operations of reasoning in different contexts such as research papers on the Web or semantic indexing web pages.

II. Theoretical approach

In this section we will begin by defining our theoretical framework, we will start with the definition and construction of Bayesian Networks, the arcs and their orientation and usage in different fields.

A. Bayesian Networks

1) Definition

A Bayesian network [5] is a directed acyclic graph in which the nodes correspond to the variables (user properties) and links represent probabilistic relations of influence. These variables can belong both to the field of knowledge, knowledge base and / or cognitive model. Each node represents a belief system about the possible values (levels, states) of the variable. Thus, the conditional probability distribution must be specified at each node. If the variables are discrete, they can be represented as a table.

The graph is also called the "structure" of the model, and the probability tables its "parameters". The structure and parameters can be provided by experts, or calculated from data, although in general, the structure is defined by experts and the calculated parameters from experimental data.

2) Construction of a Bayesian network

As we have seen in the definition, the complete specification of a Bayesian network requires one hand to specify its structure (directed acyclic graph that underlies) and other parameters (the probability tables). To do this, two approaches are possible and can be combined: the collection of expertise and machine learning, which is one of the appeals of Bayesian networks.

In the case of collection of expertise, the definition of the network structure begins with the identification of possible nodes and the distinction between informational variables (inputs) or hypothetical (unobservable) [6]. The existence of an arc can be analyzed in terms of influence of one variable on another, but its orientation is more difficult. Traditionally, an arc is directed from A to B if A is a cause of B, but we will see that this interpretation is not as simple in the case of the learner modeling. The parameters are in turn attached in an approximate manner by using frequentists or qualitative information.

Since Bayesian network is a probability distribution, we can use maximum likelihood as statistical learning parameters criterion. The result is as a Bayesian network whose structure is fixed and E which is a comprehensive basis of example, the maximum likelihood is achieved if the parameters of the Bayesian network

are equal to the frequencies of the same features observed in E. statistical learning structure requires for its development test to determine whether or not the random variables are conditionally independent.

3) *Arcs Orientation*

The definition of the structure of a student model based on Bayesian network from expert knowledge, is often done from a representation of domain knowledge, we do not discuss here the validity of the skeleton (graph undirected) resulting in the type of links that are taken into account (mainly partitive or generic). Therefore, obtaining the model structure from the skeleton of the orientation requires arcs. These guidelines determine the diagnosis obtained, since as they play a key role in the relationship dependency between variables.

a) *The influence of the orientation of the arcs on the diagnosis*

We point out an example of the influence on the arcs orientation of the diagnosis obtained. “Figure. 2” shows a simplified version of the model Hydrive, where the performances of the electronics learner can be observed. The set of variables representing the different skills of the learner is affected by this observation. Thus, if the student is found to be competent in electronics, for example because of their initial training, the diagnostic model is the same in all other disciplines.

Consider “Figure. 2” a graph having the same skeleton but with other orientations of the arcs. The spread of the information here is much more limited.

The diagnosis we got depends not only on arcs orientation’s choices, but also it seems that the orientation depends on the expertise level of the learner. Indeed, it seems reasonable to consider the network of “Figure. 1” for a subject having followed all the training (for which we can assume a homogeneity especially in some skill levels), the network 3 seems more appropriate for a about the beginning of training, which can be very proficient in a particular area because of its course without the need to master all the skills.

b) *Which choice for the orientation of arcs?*

In literature, the choice is massively in favor of a focus node of the domain layer to those of the task layer.

The question remains open regarding the arcs linking the nodes in the domain layer. It is common to present this choice as an alternative choice between a general orientation to the individual or its opposite. We find in literature examples of these two choices ([7] and [8]), even if the justification given for this choice is not always totally convincing, even if sometimes it doesn’t exist.

Figure 1. Flow of information in the learner model of Hydrive.

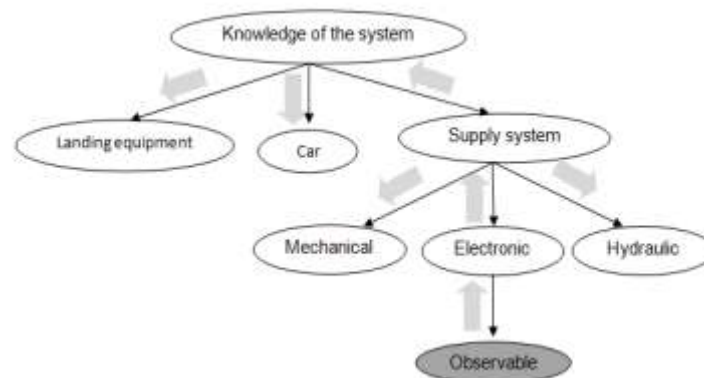
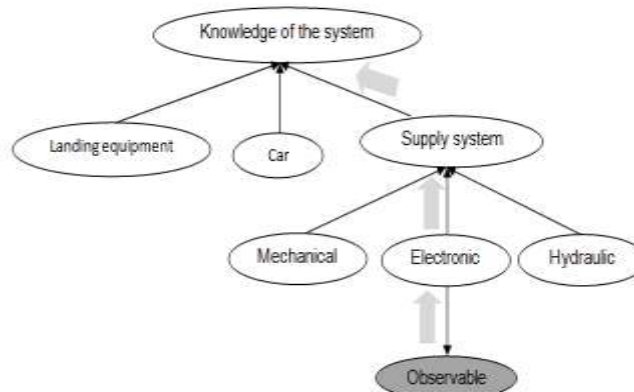


Figure 2. Flow of information in the modified learner model of Hydrive.



The orientation of the arcs from the general to the particular is suitable when the skills of the learner have certain homogeneity because of dependencies of this orientation. The orientation on the other side is most appropriate for learners with diverse skills. Moreover, if we take the findings on the relationship between network structure and level of expertise of the learner [9], we conclude that the model must allow this structural change.

B. Probabilistic Ontologies

1) Ontologies

The definition of ontologies is inherited from a philosophical tradition that focuses on the science of being. Today, it means the "science of beings" that is to say all of the articles recognized as existing in a domain. According to [1], ontology is "an explicit and formal specification of a shared conceptualization". It is a representation that is a set of concepts and relationships describing a particular area. One of the major shortcomings of the ontology is their inability to represent and reason about uncertainty.

The ontology is used for several years in the Knowledge Engineering and Artificial Intelligence to structure the concepts of a domain. The concepts together and they are considered the building blocks for expressing domain knowledge it covers.

According to [3], Ontologies are useful to share knowledge, build consensus, and build knowledge-based systems. Many projects are implemented ontologies such as the Semantic Web. The fundamental problem is to respect the diversity of languages and representations of the world, while enabling the exchange of information. One can characterize ontology as structuring concepts of a domain. These concepts are combined to provide the building blocks and express the knowledge available in this field.

2) Probabilistic ontologies and Bayesian networks

In the other hand, Probabilistic Ontologies are used to describe knowledge about a rational, structured and sharable field, ideally in a format that can be read and processed by a computer and incorporate uncertainty in knowledge.

A Bayesian Network is a probabilistic model that is able to handle uncertainty. It is defined by a directed acyclic graph whose represent random variables in a domain, the edges indicate conditional dependencies between the peaks and the conditional probabilities are used to quantify the dependencies between the nodes. It is based on graph theory to represent the conditional dependencies between variables of a system under study, and on the theory of probability to mathematically define these dependencies [11].

Various researchers have tried to combine Bayesian networks with the ontologies to represent and reason with uncertainty ([3], [10] and [12]) proposed extensions to the standard OWL formalism which are BayesOWL, OntoBayes and PR-OWL. Boths [13] and [14] have proposed new probabilistic extensions of probabilistic description logics (LDP), which are a family of representation languages of probabilistic ontological knowledge and are well adapted to the web.

We have integrated the Bayesian network in the ontology in order to remedy the defects of each one of them (Bayesian network is unable to formally represent the uncertain information and ontology is unable to handle uncertainty) and combine the Bayesian inference (update probabilities) with the ontological inference (verification instances, consistency, etc.). The Probabilistic Ontology obtained can be subsequently used in aid to the decision that must be able to reason with uncertain knowledge applications.

III. Methodologie

In this section, we will present our representation model in a comprehensive way by citing all the steps necessary to realize the transformation of a Bayesian Network to ontology, and then we will define the translation rules which are the first step of the process of this transformation

A. Representation Model

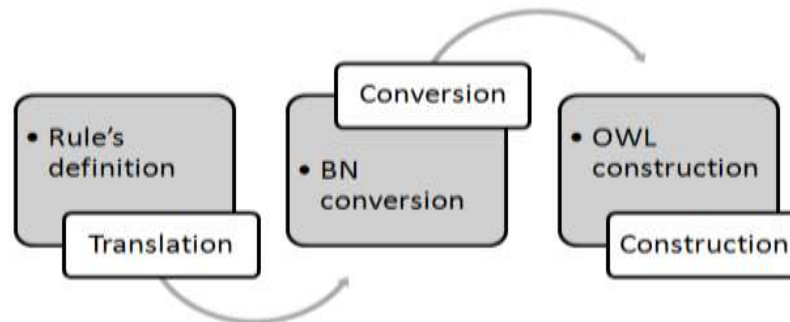
A major advantage of the Bayesian Network is their ability to manage uncertainty: it allows representing uncertain information with a very clear and readable manner and provides powerful techniques for Bayesian inference that is to propagate one or more data (probabilistic values) in the network to deduce how this works on the probability of other variables. This inference is called "update" of probabilities.

To formalize the representation of this model and make it understandable by machines, we propose a formal modeling method of Bayesian Network using standard OWL ontology formalism. The aim of this method is not only formally represent information from Bayesian Network as a Probabilistic Ontology but also combine some highlights of Bayesian Network with those of ontology reasoning tools ontology (subsumption, instance checking, consistency, ...) with the tools offered by the Bayesian Network (representation of uncertain information, inference, ...) to provide mechanisms for representing uncertain information and infer implicit knowledge from explicit knowledge stored in a probabilistic ontology.

"Figure.3" represents the three steps which allow us to transform the Bayesian networks into OWL. In this work we will propose the translation rules in which we can represent information from Bayesian Network as a Probabilistic Ontology. After the translation, the next step will be the conversion of the Bayesian networks into

ontologies concepts and property based on our translation rules. After that, the final step is the construction of the ontologies based on the two previous steps.

Figure 3. The transformation process from Bayesian Network toward ontology



B. Translation rules

Q is a Bayesian network which is defined by a graph $G = (A, S)$, with $A = \{A_1, A_2, \dots, A_m\}$ is the set of dependency's relations size m and $S = \{S_1, S_2, \dots, S_n\}$ is the set of vertices of size n .

The structural similarity is observed between the Bayesian Network graph and the ontology graph. Bayesian network and ontology are defined by a graph. These two graphs have a set of edges and a set of vertices. This facilitates the task of similarity transformation of a Bayesian Network Probabilistic ontology that may be made using a set of translation rules which are:

- **Nodes represented by concepts:**

The nodes shown in the diagram Bayesian calculation must be represented by the ontology concepts, which are relevant to the problem considered. The n nodes of the Bayesian network graph are transformed into a set of concepts $C = \{c_1, c_2, \dots, c_n\}$, where n is the number of ontology concepts.

- **Values represented by instances of concepts:**

The possible values for each node are converted into set of instances of concepts. $I_i = \{I_1, I_2, \dots, I_n\}$ and $I_j = \{i_1, i_2, \dots, i_k\}$, where I_j is the whole concept instance j and k is the number of instances of concept j .

- **theoretically probability by Data Property:**

The theoretically probability $P(A)$ is converted into a probability value (between 0 and 1) of a property DataProperty type, named "TheoreticalProbability C_i " of concept C_i .

$C_i \in R_t$ With $R_t = \{c_1, c_2, \dots, c_l\}$ is the set of roots of Bayesian Network (R_t subset of C), and I is the number of nodes without parents (roots). $c_i \in R_t$ if and only if $C_i \in C$ and the parent of $(C_i) = \emptyset$.

- **The dependency relationship by "N-air"**

The presence of an arc between two nodes of Bayesian Network has a dependency relationship between these two nodes. This dependence will be transformed into an N-Aire relation (N-Aire relationship is a relationship with more than two arguments or additional attribute.), Called "depends-on- j ", which is characterized by property kind DataProperty named "CondProb- j ".

"CondProb- j " "property expresses the conditional probability $P(A | B)$, which indicates the probability for each value of A , given the combinations of values of the parents of A (B).

So the m arcs are converted into a set of N-Aire relationships between concepts $R = \{R_1: \text{dependent on } 1, R_2: \text{dependent on } 2, \dots, R: \text{dependent on } m\}$. Each relation R on R is characterized by a property kind DataProperty "CondProb- j ".

IV. Conclusion and perspectives

In this paper, we proposed a method to semantically represent a Bayesian Network as a Probabilistic Ontology. This work represent the first step into the integration of a Bayesian network into ontology, we have defined in this work the rules of translation from a Bayesian network into in ontology.

The next step in our future work will be the conversion and the construction of an ontology based on a Bayesian Network. We will try to use this Probabilistic Ontology in the experiments we performed in previous work for the management of learner modeling.

References

- [1] Nottelmann H. & Fuhr N. (2006). Adding probabilities and rules to OWL Lite subsets based on probabilistic Datalog, Int. J. Uncertain. Fuzz. pp 17–42.
- [2] Pool M. & Aikin J. (2004). KEEPER and Protégé: An elicitation environment for Bayesian inference tools, in: Proceedings of the Workshop on Protégé and Reasoning held at the IPC.
- [3] da Costa P.C.G. et al. (2005). da Costa P.C.G. et Laskey K.B. Laskey K.J., PR-OWL: A Bayesian ontology language for the Semantic Web, in: Proceedings URSW, pp. 23–33.
- [4] Heinsohn J. (1994). Probabilistic description logics, in: Proceedings UAI-1994, Morgan Kaufmann, pp. 311–318.
- [5] Pearl J. (1988). Probabilistic Reasoning in Intelligent Systems, Morgan Kaufmann.

- [6] M. Anouar Tadlaoui, M. Khaldi, S. Aammou (2014) Bayesian Networks for Learner Modeling, International Journal of Basic Sciences and Applied Computing 1 (1), 5-9.
- [7] Mislevy R., Gitomer D. (1996). The role of probability-based inference in an intelligent tutoring system, User-Modeling and User- Adapted Interaction, n°5, pp 253-282.
- [8] Zapata-Rivera J.-D., Greer J.(2000) « Inspecting and Visualizing Distributed Bayesian Student Models », Intelligent Tutoring Systems 00, pp. 544-553.
- [9] M. Anouar Tadlaoui, M. Khaldi, S. Aammou (2014) TOWARDS A LEARNING MODEL BASED ON BAYESIAN NETWORKS, EDULEARN14 Proceedings, pp. 3185-3193.
- [10] Studer R. et al. (1998). Studer R. Benjamins, V.R., Fensel, D., Knowledge engineering: Principles and methods. Data & Knowledge Engineering 25, 161–197.
- [11] BECKER, A., NAÏM, P., Les réseaux Bayésiens, modèles graphiques de connaissances, Eyrolles, 1999.
- [12] Philippe L. (2006). Réseaux bayésiens : apprentissage et modélisation de systèmes complexes, HABILITATION.
- [13] Yang Y. & Calmet J. (2005). Ontobayes: An ontology-driven uncertainty model. In: (IAWTIC'05). IEEE Computer Society (2005) 457-464.
- [14] Ding Z. & Peng Y. (2004). A probabilistic extension to ontology language OWL, in: Proceedings HICSS.