Efforts to Blend Ontology with Bayesian Networks: An Overview

Asma S. Larik
Artificial Intelligence Lab, Faculty of Computer Science
Institute of Business Administration
Karachi, Pakistan
asma.sanam@khi.iba.edu.pk

Sajjad Haider
Artificial Intelligence Lab, Faculty of Computer Science
Institute of Business Administration
Karachi, Pakistan
sajjad.haider@khi.iba.edu.pk

***Abstract*— The concept of ontology modeling and engineering has drastically evolved with the emergence of semantic web. Today in a collaborative environment ontologies need to interoperate for efficient data exchange. Due to their deterministic nature, ontologies are unable to capture uncertainty inherent in many real world problems. Efforts have been made to add the dimension of uncertainty by synergizing Bayesian network with ontology - a combination in which domain knowledge and probabilistic information go hand in hand. This paper categorizes different approaches that have been proposed to blend ontology and Bayesian networks. It also discusses the scope of these approaches and their intended application areas. Their relative strengths and weaknesses are also highlighted. The paper contributes significantly by providing a comprehensive comparison which could serve as the starting point for new researchers.**

Bayesian network, ontology, ontology mapping, BayesOWL

I. INTRODUCTION

Ontology generally means a formal representation of entities in a particular domain and their attributes and relationships [1]. It has its basis in classical logic and ontologies representing different domains have been widely contributed by ontology modelers around the globe. As the research in ontology engineering is maturing with the passage of time, it is being realized that this formalism has its limitations that need to be addressed. Firstly, different domain experts model domain ontology with their own view point; thus, making it infeasible for two ontologies belonging to the same domain to talk to each other. Secondly, ontologies lack support for modeling uncertain parameters. This uncertainty could be evident in various forms: the entities in ontology might be uncertain, their relationships might be uncertain or their properties might be uncertain. Thus in a non-deterministic world the deterministic view of ontology might not be expressive enough to address real world problems. In order to solve the above mentioned problems, a mathematical and structural model with inherent support for probabilistic reasoning is required and Bayesian network (BN) [2] seems to be an ideal solution. Mathematically, a Bayesian network is a directed acyclic graph (DAG) in which a set of random variables makes up the nodes in the network. A set of directed links connect pair of nodes and each node has a conditional probability table (CPT) that quantifies the effects of parents on it. Due to their ability to describe complex joint distribution using a collection of local distributions, Bayesian networks have become highly popular and have been used in many applications [3, 4]. At a macro level, the efforts to amalgamate ontology and Bayesian network can be categorized as follows:

1. Ontology mapping enhancement using Bayesian
2. mechanism
3. Semi-automated construction of Bayesian network from domain ontology
4. Probabilistic extension of Ontology Web Language (OWL)
5. networks
   Ontology reasoning enhancement using Bayesian inference

Ontology mapping (category 1 above) is an approach that tries to incorporate two or more ontologies, belonging to the same domain but modeled by different experts, into a single ontology. The two techniques that are discussed in this paper are OMEN [5] (Ontology Mapping ENhancer) and BayesOWL [6-8]. OMEN describes a set of meta- rules that capture the ontology structure, its relationships and then maps simple concept to other concepts in its neighborhood. BayesOWL is an approach that tries to merge ontology model with probabilistic constraints in the OWL file to generate a Bayesian network. The BN thus formed can be utilized in identifying similar concepts. Ontology reasoning via Bayesian inference mechanism (category 2 above) is recently proposed by Andrea and Franco [9]. The approach suggests construction of a two-level Bayesian network by exploiting the information contained in the domain ontology. The approach differs from approaches of category 1 in the sense that an OWL file for probabilistic constraints is not required and it claims that domain ontology would suffice for Bayesian network generation. Semi-automated construction of Bayesian network (category 3) has also been an extensive area of research these days with major contributions made by [10-12]. Language extension of OWL has also been proposed in the form of PR-OWL [13, 14]. PR-OWL differs from other techniques in the sense that it is a complete formalism that not only talks of OWL extension but also has its basis in MEBN logic [15] which itself is a combination of first order logic and Bayesian networks. The focus of this paper is to highlight and categorized the key efforts reported in the literature to integrate ontology and Bayesian networks. To the best of our knowledge, no similar comprehensive study has been done earlier. It is hoped that the work would benefit both practitioners and young researchers who are new into this field.

The rest of the paper is organized as follows. Section 2 describes OMEN and BayesOWL, two popular techniques suggested for ontology mapping. A detailed account of ontology reasoning via BN inference is given in Section 3, while Sections 4 and 5 highlight the techniques for partially mechanized construction of Bayesian network and PR-OWL, respectively. Finally, Section 6 concludes the study and highlights important challenges for future research.

II. Ontology Mapping Enhancement Using Bayesian Networks

Ontology mapping is an approach that tries to overcome the problem of semantic heterogeneity by finding mappings among similar concepts that exist in two or more ontologies. Two techniques that employ Bayesian network to support this process are discussed below.

A. BayesOWL

BayesOWL is a framework that facilitates conversion of an OWL file to a Bayesian network. The purpose of this conversion is the mapping of similar concepts. It primarily answers queries such as the probability of A being a subclass of B, the probability that B is subsumed by A and/or the degree of similarity of A and B where A and B can be any two concepts defined by ontology modelers. The transformation requires encoding of the OWL file with class and probability tags. The structural translation to a DAG is then performed which also generates normal and logical nodes. CPTs for both types of nodes are computed next, and finally reasoning is performed via BN inference mechanism. A summary of these steps is described below. Readers interested in a detailed description should refer to [6-8].

1) *Encoding OWL file:* Ontology concepts and probability information is encoded in the form of XML tags as shown in Table 1.

Annotated OWL file	
XML Tags	***Meaning***
<owl:class>	Root concept
<rdfs:subclassOf>	Representing is-a relationship
<owl:disjointWith>	Two concepts that cannot coexist
<owl:unionOf>	A concept that is a union of two or more concepts
<owl:equivalentClass>	Two similar concepts
<owl:intersectionOf>	A concept that is intersection of two or more concepts
<owl:variable>	Represents a variable
<hasState>	Represents the states that variable can take
<owl:Probability>	Represents the probability of the variable
<PriorProb>	Represents prior probability
<CondProb>	Represents conditional probability
<hasValue>	Represents the value of probability

Table 1: Annotated OWL File

2) *Structural Translation to a DAG:* The encoded OWL file is converted to a DAG based upon the set of translation rules. According to these rules, each concept in the OWL file is mapped to a binary-state node in the Bayesian network. For subclass relationships, directed links are created from child classes to the parent class. To cater for logical operators in the owl file, logical nodes are also created. The basic logical operators as shown in Table 1 are intersection, union, complement, equivalent and disjoint operators. The translation rules for logical nodes are creation of links between the classes based on the logical operator that exists between them.

3) *Construction of CPT:* This phase is divided into two sub phases: (a) the construction of CPTs of the logical nodes and (b) the construction of CPT for concept nodes. For logical nodes the CPT values are computed by using the possible combination of their states and applying the logical operator. These set of rules are easier and can be extended to any number of variables. However, creation of CPTs for the concept nodes is a tedious task and it requires the incorporation of probabilistic constraints defined in the probability.owl file. The BayesOWL approach uses an algorithm, named SD-IPFP (Iterative Proportional Fitting Procedure) that generates the joint probability distribution of a given set of nodes in the Bayesian network. Using this initial distribution as the knowledge base, marginal probabilities are calculated. Next the set of constraints that are either learnt from data or provided by experts are divided by the marginal probability to achieve a ratio of constraints to marginal probability. Lastly the ratio is multiplied with each value in the JPD to obtain a new JPD. This process is iteratively applied until the joint distribution converges. Thus in this way it guarantees that all the previous constraints would be satisfied and the joint would converge to a specific point.

4) *Ontology Reasoning by Bayesian Inference:* The purpose of ontology conversion to Bayesian networks is to test concept satisfiability, concept overlapping and concept subsumption. A simple BN chain rule is used to test concept satisfiability, while the degree of overlap of one concept with the other is computed to test concept overlapping. For concept subsumption, the degree of similarity of one concept with the other, say e and C, is checked by Measure of Similarity of Concept (MSC) based on Jaccard Coefficient defined by:

$$MSC (e, C) = P(e \cap C) / P (e \cup C) \quad (1)$$

To summarize, BayesOWL is a consistent approach with its strength being the ability to perform ontology mapping using Bayesian inference mechanism. A Java based tool for BayesOWL has also been developed that aids a user in understanding the concepts. Some of its limitations are restriction of nodes to binary states and assumption that the probabilistic constraints provided are consistent. If the constraints contradict then BayesOWL would not be able to find a JPD that converges to a point so research in this direction is also being carried out.

B. OMEN

Ontology Mapping ENhancer(OMEN) is a probabilistic ontology mapping tool. The input to OMEN are the source ontologies, a set of initial mappings that are output of some ontology mapping tool and initial probability distribution of matching concepts. OMEN then generates a set of additional mappings utilizing the initial ones For instance, let O1 and O2 be two ontologies then OMEN checks if the initial probability of their match is above the pre-specified threshold, and if yes it creates a node in the Bayesian network and marks it as an evidence node. The mapping pair of concepts, say C1 from O1 and C2 from O2, are also utilized in the creation of nodes. Edges are inserted between the added nodes depending upon the relationships provided in the ontologies. A set of meta rules are defined to compute CPT . For instance, one of the basic rule says if two nodes in the ontology match and so do the arrows coming out of these nodes, then the probability that nodes at the other end of the arrows match as well is increased. Although an intuitive idea but its major weakness is that it is dependent upon other mapping tools for input. Secondly if a concept has 5 parents and other has 7 parents then the total parents would be 35 leading to intractability of knowledge acquisition. Thus heuristic method has been proposed and OMEN is restricted to 10 parents only.

III. Ontology Reasoning Enhancement Using Bayesian Networks

An approach to incorporate Bayesian network in performing probabilistic reasoning via ontology has recently been proposed by Andrea and Franco [9]. It is claimed that the approach is not an extension of owl file, as is the case with BayesOWL, but it is novel in the sense that it only uses information stored in the domain ontology for constructing the corresponding Bayesian network. The hierarchical relationships among entities in the ontology are preserved in TBox, while instances are stored in ABox. The approach generates two level Bayesian networks where the levels are termed as hierarchical level and basic node level. There are two types of nodes in such two level BNs. One is called High Level Nodes (HLN) which corresponds to root concepts in the ontology, while the second is called Low Level Nodes (LLN) which corresponds to subclasses of the root concepts. The proposed approach has three basic steps: (a) construction of a structural two level BN from TBox, (b) construction of CPT from ABox and (c) probabilistic ontological reasoning using BN Inference. All the root concepts in the TBox are converted to High Level Nodes and all the subclasses are converted to Low Level Nodes in the Bayesian network. The HLN are multivalued random variables whereas LLN are Boolean random variables. The relationships in the TBox are used to create links between High Level Nodes and are termed as High Level Relations (HLR). All the is-a relationships are used to create arcs between High Level Nodes and Low Level Nodes and are termed as Low Level Relations (LLR). The instances in the ABox are used to compute conditional probabilities of nodes in BN. There are two kinds of probability tables: Low Level Relations Probability Table (LLRT) and High Level Relations Probability Table (HLRT). The LLRT would contain all the possible combinations of truth values for low level nodes while HLRT would contain combination of truth values for the relation between root concepts that have been transformed to HLN in Bayesian network. After transformation, BN Inference is performed over the compiled two levels BN. The major strength of this technique is that it is very simple and easy to understand, and it relies only upon the domain ontology with no external information. It, however, takes into account only subclass relationships and associative relationships among classes and lacks in other relations such as disjointness, cardinality, etc. Further work is being done to incorporate attributes into it and to reason upon attributes. One shortcoming of this approach is that the generated Bayesian network should possesses polytree property. A poly tree property requires that only single path exists between any two nodes in the network. If this property does not hold then BN inference mechanism would get stuck into a loop and might not converge.

IV. Semi-Automated Construction Of Bayesian Network

The nature of this class of work is similar to the work reported in the previous section of constructing a BN from a given ontology. The focus, however, is different in the sense that this approach argues that complex Bayesian networks are notoriously difficult to construct. This method of semi automated construction of BN would help a domain expert in providing knowledge base in the form of ontology and the process would itself identify variables of interest, their influence on each other and their respective probabilities via some function. The modeler may understand BN inference but does not want to get involved into the nitty-gritty of BN modeling and reasoning. Thus this approach has its target audience as the domain modelers whose sole purpose is to utilize their prior made models for probabilistic reasoning. Attempts have been made to address this issue and different techniques (discussed below) have been proposed. Most of the approaches, however, are case based and there is a need for a generalized technique. One approach presented by Devitt et al. [10] divides the BN construction process into several phases. These phases include identification of variables of interest, values that those variables can assume, relationships among those variables and CPT computation. The case study presented by them, however, is about network-specific ontology extension that seems to make the approach domain specific. The next attempt was made by Zheng et al.[11] In which uncertainty in Clinical Practice Guidelines (CPG) was modeled by first converting CPG, that is available in the form of an Activity Diagram, into domain ontology and then converting the domain ontology to a corresponding Bayesian network. Like the effort reported above, this conversion is also domain dependent and modeling is performed keeping BN in mind. Fenz et al. [12] proposed an approach that is semantically more appealing as compared to the other two approaches mentioned above. The input to this approach is a domain ontology. The first phase of this approach maps the relevant concepts in the ontology to nodes in the corresponding Bayesian network. Next the relationships among the concepts are utilized in creation of arcs in Bayesian network. Axioms in the ontology are mapped to states by defining a three point Likert scale (Low, Medium, and High) and instances in ontology are mapped to findings or evidences in the BN. Lastly, CPT computation is performed by means of some external formula specified by expert.

The strength of this approach is its simplicity of modeling Bayesian network through ontology. The authors of this approach, however, have left the topic of CPT construction to be handled by domain expert which is its major limitation. Although it does provide a method in which external formula can be incorporated into BN but this formula is not provided by the ontology modelers and thus makes this approach a little bit vague. Furthermore, the approach claims to be semi automated and depends upon the quality of ontology being modeled; if the ontology is not modeled properly then human intelligence would again be required to make necessary adjustments/corrections.

V. Ontology Language Enhancement Using PROWL

PR-OWL is a probabilistic extension of the existing OWL. It encompasses a general framework with its basis in MEBN logic which integrates first order logic with Bayesian networks. The focus of this approach is to provide a framework that can deal with uncertainty in a consistent manner not attained by just annotating the OWL file. The MEBN logic is organized in the form of MTheory which itself is composed of multiple MFrags. An MFrag is actually a collection of nodes in a Bayesian network that are logically grouped together. MFrags contain nodes that are categorized as resident, input and context nodes. PR-OWL is defined as an upper ontology and its main elements are depicted in the Fig 1.UnBBayes[1] is a tool that helps in modeling a Multi Entity Bayesian Network for a given domain. The Bayesian network thus formed can be imported into Protégé'[2]. The strength of PR-OWL is its novel idea of incorporating probability in ontology. Its weakness, however, resides in the fact that from the perspective of an ontology modeler it may become very difficult at times to understand the concepts of MEBN theory.

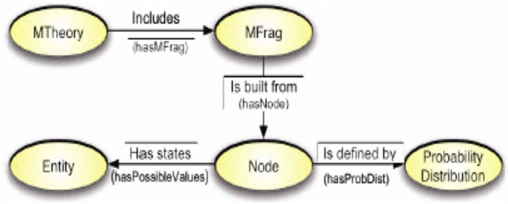

Figure 1: Overview of PR-OWL [14]

VI. Conclusion

This paper is a comparative survey of most of the techniques presented so far to combine the expressive power of ontology with the uncertainty modeling capability of Bayesian networks. It narrated the different ways in which people have tried to merge both modeling techniques. Most of the efforts have been on automated/semi-automated transformation of a given ontology to a Bayesian network. These techniques preserve the domain model and apply probabilities in the form of constraints or probability instances or relationships to make ontologies probabilistic. Extensive efforts have also been made to extend the OWL so as to contain additional semantics that truly capture uncertainty in a consistent manner. Despite many efforts, there are still many open issues that are needed to be addressed. The paper highlighted the strengths and weaknesses of many of the important techniques presented so far thus identifying a need for more generalized approach.

Acknowledgment

The first author would like to acknowledge special support and cooperation from Professor Yun Peng and

[1] UnBBayes is an open source software for modeling probabilistic networks developed at George Mason University http://unbbayes.sourceforge.net/

[2] Protégé is an open source ontology editor developed at Stanford http://protege.stanford.edu/

Shenyong Zhang of University of Maryland Baltimore County.

REFERENCES

[1] T.R. Gruber, "A translation approach to portable ontology specifications," *Knowl. Acquis.*, vol. 5, 1993, pp. 199-220.

[2] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[3] A. Mittal, *Bayesian Network Technologies: Applications and Graphical Models*, IGI Publishing, 2007.

[4] O. Pourret, P. Na?m, and B. Marcot, *Bayesian Networks: A Practical Guide to Applications*, Wiley, 2008.

[5] P. Mitra, N. Noy, and A. Jaiswal, "OMEN: A Probabilistic Ontology Mapping Tool," *The Semantic Web – ISWC 2005*, 2005, pp. 537-547.

[6] Y. Peng, "A Bayesian Network Approach to Ontology Mapping," *The Semantic Web – ISWC 2005*, 2005, pp. 563-577.

[7] Zhongli Ding and Yun Peng, "A probabilistic extension to ontology language OWL," *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, 2004, p. 10 pp.

[8] D. Zhongli, "PhD. dissertation,BayesOWL: A Bayesian Approach to Ontology Mapping," University of Maryland Baltimore County, 2005.

[9] B. Andrea and T. Franco, "Extending ontology queries with Bayesian network reasoning," *Intelligent Engineering Systems, 2009. INES 2009. International Conference on*, 2009, pp. 165-170.

[10] Ann Devitt and K. Matusikova, "Constructing Bayesian Networks Automatically using Ontologies," *Applied Ontology, IOS Press*, 2006.

[11] H. Zheng, B. Kang, and H. Kim, "An Ontology-Based Bayesian Network Approach for Representing Uncertainty in Clinical Practice Guidelines," *Uncertainty Reasoning for the Semantic Web I: ISWC International Workshops, URSW 2005-2007, Revised Selected and Invited Papers*, Springer-Verlag, 2008, pp. 161-173.

[12] S. Fenz, A. Tjoa, and M. Hudec, "Ontology-Based Generation of Bayesian Networks," *Complex, Intelligent and Software Intensive Systems, 2009. CISIS '09. International Conference on*, 2009, pp. 712-717.

[13] P.C.G. Costa and K.B. Laskey, "PR-OWL: A Framework for Probabilistic Ontologies," *Proceeding of the 2006 conference on Formal Ontology in Information Systems: Proceedings of the Fourth International Conference (FOIS 2006)*, IOS Press, 2006, pp. 237-249.

[14] P.C. Costa, K.B. Laskey, and K.J. Laskey, "PR-OWL: A Bayesian Ontology Language for the Semantic Web," *Uncertainty Reasoning for the Semantic Web I: ISWC International Workshops, URSW 2005-2007, Revised Selected and Invited Papers*, Springer-Verlag, 2008, pp. 88-107.

[15] K.B. Laskey, "MEBN: A language for first-order Bayesian knowledge bases," *Artif. Intell.*, vol. 172, 2008, pp. 140-178.