Multi-Entity Bayesian Networks for Treasuring the Intangible Cultural Heritage

Giannis Chantas, Spiros Nikolopoulos and Ioannis Kompatsiaris

Information Technologies Institute, Centre for Research and Technology - Hellas, Thessaloniki, Greece
{gchantas, nikolopo, ikom}@iti.gr

Abstract: In this paper, we propose the use of Multi-entity Bayesian networks for modeling the knowledge and analyzing the content pertaining to the domain of Intangible Cultural Heritage (ICH). MEBNs provide a rigorous knowledge representation framework in conjunction with reasoning and probabilistic inference capabilities. There are mainly two reasons motivating the use of MEBNs in the domain of ICH. The first is that MEBNs extend first-order logic with the ability to model uncertainty. The second reason is the capability of MEBN to adapt to specific situations by providing custom, situation specific Bayesian networks. Finally, we use an example to demonstrate the potential efficiency of MEBNs in the domain of ICH.

1 INTRODUCTION

The term intangible cultural heritage (ICH) (UN-ESCO, 2013) refers to valuable traditional art forms and creative practices, such as singing, dancing, craftsmanship, etc. Preserving the knowledge captured by these *human treasures* (as called by UN-ESCO) is considered particularly important and the use of technology to achieve this objective has recently become a popular research topic.

In this paper, we highlight the usefulness of an efficient knowledge representation scheme in facilitating the preservation procedure of ICH. We are motivated by the fact that expert domain knowledge can boost the automated analysis of ICH content and propose the use of a sophisticated tool like Multi-entity Bayesian Networks (MEBNs) for allowing preservation to go beyond the mere digitization of ICH manifestations.

ICH manifestations can be captured in a multimodal fashion, using sensors of various types, such as visual (2D/3D), auditory, motion, etc. Furthermore, the knowledge regarding the ICH domain is diverse, vast and intricate. For the efficient analysis of the captured and digitized content, the expert knowledge plays a crucial role. Thus, this knowledge has to be precisely represented in a digitized environment and in a feasible to exploitation manner. A major difficulty in representing such knowledge is the inherent ambiguity and uncertainty in concepts prevalent, not only the ICH domain, but almost in every real world

environment being modeled. For instance, the steps performed during a traditional dance, although their number is certain, the way that they are followed by the performers may vary with respect to the region, epoch, or even the individual. Moreover, the number of concepts to be detected is rarely *a priori* known. Indeed, the number of dances and dance steps, songs and singers, etc, varies per situation.

Ontologies (Staab and R. Studer, 2009) are used extensively for knowledge representation. An ontology provides knowledge representation of a domain through digital means in a formal language. In this way, the use of this knowledge for the analysis of new content, i.e., reasoning, is feasible. Representation in an ontology is usually based on the description logic (DL), a variant of first-order logic (Davis, 1990) (FOL). However, the expressivity of DL is not sufficient to solve the problem of uncertainty in ICH knowledge.

In order to overcome this obstacle, various works have been proposed that enable logic-based representation schemes to express uncertainty using probability theory. The model we propose for this endeavor in this paper is MEBN (Laskey, 2008). MEBN is the outcome of the cross-fertilization of first-order logic and probability theory. MEBN adopt probability theory as a means to model uncertainty. First-order Bayesian logic (FOBL) is an extension of FOL, enabling MEBN to extend BN by being more flexible to adapt to specific situations that arise. More specifically, with FOBL, a MEBN can be instantiated

to a situation specific Bayesian network (SSBN) that models the situation that has been recorded as evidence in the environment being modeled. This overcomes the rigid and non-flexible nature of Bayesian networks.

In the next section, we provide the necessary theoretical background knowledge about FOL, Bayesian networks (BN) and ontologies. Then, in Section 3 we describe the concept of MEBN as a knowledge representation and analysis model for the ICH domain. Moreover, an example of the application of MEBN is given that demonstrates the advantages of such a choice. In Section 4, existing literature related to ICH domain knowledge modeling is provided. In sections 5 and 6, a discussion and conclusions are given, respectively.

2 BACKGROUND

2.1 First-Order Logic

First-order logic (FOL) (Davis, 1990) is by far the most commonly used, studied, and implemented logical system. FOL is a formal system used to define theories in mathematics, computer science and other scientific fields, and it is a very important concept, both from a theoretical and a practical viewpoint. FOL can also be used as a rigorous foundation of knowledge representation schemes.

A theory in first-order logic consists of the axioms, expressed as sentences in the FOL language, in conjunction with the sentences that are derived from the axioms according to the reasoning rules, i.e., the theorems or valid sentences. In practice, when a FOL theory is implemented in a computer, axioms are stored as data structures and the reasoner that evaluates the truth-value of a sentence has the form of a computer program.

Theories are expressed using the FOL language. The main components of the FOL language are constants, variables, functions and predicates. Variables are placeholders of constants. Functions return a constant given their input arguments. Predicates are an essential component of FOL (the reason that it is also called *predicate logic*), since they are used to define relations between other components, such as variables and constants. For example the predicates *parents(John,Mary,George)*, may state that John and Mary are parents of George. What also makes FOL a strongly expressive language are the rules, such as "*sibling(X,Y):- parent(X,X1,X2) AND(Y,X1,X2)*". This rule states that any two persons *X, Y* are siblings if they have the same parents.

An interpretation of a FOL theory gives a semantic meaning to each constant, predicate and function (formulas). More specifically, an interpretation maps each formula to a specific nominal entity, constraints each predicate to relate with each other entities belonging to a specific set, and relates a function to a function of the domain. A set containing all instantiated formulas of a FOL theory is called an interpretation. Note that, there can be many interpretations, since instantiations vary.

However, FOL does not provide expressivity to model uncertain knowledge, let alone a suitable rigorous reasoning mechanism. This is a consequence of the fact that each interpretation mentioned above shares equal validity with the others. As we will see in 3, the key feature of MEBN is the assignment of a probability to every interpretation. This is achieved with the first-order Bayesian logic (FOBL) language.

2.2 Bayesian Networks

Bayesian networks (BN) are stochastic models that have been applied successfully to problems where uncertainty is prevalent. Bayesian networks are directed, acyclic graphical (DAG) models (Whittaker, 2009). More specifically, a BN consists of random variables, represented by nodes in a DAG, along with their relations determining the structure of the DAG. Moreover, a conditional probability distribution is defined for each variable. Formally, a Bayesian network consists of the following:

- N random variables, denoted by $\mathbf{x} = \{x_1, ..., x_N\}$.

- A graph G (which is a DAG), defined by the set $G = \{E, V\}$, where V is the set of vertices and E the set of directed edges. Each vertex corresponds to a x_i and each edge indicates a parent-child relation between the two connected variables.

- A joint probability distribution of $\mathbf{x}$, given by:

$$p(\mathbf{x}) = \prod_{i=1}^{N} p(x_i|\mathbf{pa}_i). \qquad (1)$$

$p(x_i|\mathbf{pa}_i)$ is the local conditional probability of x_i, $\forall i$ and $\mathbf{pa}_i$ the set of parents of x_i. The joint distribution of $\mathbf{x}$ in written in (1) in its so-called product form. This is possible due to the acyclicity and directivity of G.

When modeling with BNs, we assume that a random variable x_i corresponds to a measurable attribute of a real world entity, e.g., the average temperature of a country in a specific day. Nevertheless, unmeasurable concepts (such as people names, etc.) can be modeled with natural numbers. x_i can take a finite or (un)countably infinite amount of values. Denoting by

D_i the set of possible values, we write that $x_i \in \{D_i\}$. Using this terminology, $p(\mathbf{x})$ is a distribution that assigns a probability to every possible configuration of $\mathbf{x} = \{x_1, ..., x_n\}$.

Inference in Bayesian Networks is a crucial process in order to make them practical and exploit their modeling efficiency. Inference in BN mainly refers to the process of estimating the posterior distribution of a subset of the random variables, given the observed values of another subset. The observations are direct evidence that we obtain from the real world we model with the specific BN. In other words, Bayesian inference is the mechanism that offers estimation of the unknownn variables, given certain, observed values of other variables. Given the observations $\mathbf{x}_o$, the objective of inference is to estimate the posterior probability distribution of a set of unobserved (hidden) random variables $\mathbf{x}_h$, using Bayes rule.

However, the straightforward use of Bayes rule for inference is rarely practical, due to the high computational complexity of the algorithm in real world scenarios. There are many algorithms that offer tractable inference, such as belief propagation (where inference is not necessarily exact), variational approximate inference (Bishop, 2006) (mostly used for continuous variables), etc. Lastly, we should note that inference provides conclusions about the variables being modeled after the presence of new evidence. During training, the conditional probability distributions are found. Of course, there are cases where an expert can provide explicitly such distributions. Typically though, a BN is constructed prior to that, using training data and a suitable training algorithm (e.g., maximum likelihood). The serious disadvantage of BN networks when used for knowledge representation is, once they are designed, their structure is kept fixed. Thus, since in real world scenarios the modeled environment changes, BN are seriously handicapped to adapt to specific situations.

2.3 Ontologies

Ontologies have been extensively used as knowledge representation schemes of a domain, including cultural heritage (Baader et al., 2003). The most notable example is the CIDOC Conceptual Reference Model (CRM) (Doerr, 2003), which is a well-known knowledge representation ontology aiming to facilitate heterogeneous cultural heritage information exchange (tangible or not). It was developed under the guidance of the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM). In September 2006, it was accepted as ISO 21127. CICDOC CRM aims to model

the knowledge regarding cultural objects available at museums, which can be of various forms of art, geographical place, and epoch. The standardized CIDOC CRM ontology is mainly used as a common cultural information exchange framework.

Moreover, an ontology for analyzing traditional dances based on the Laban notational system for human body movement is proposed in (Ioannidis, 2011). Analysis is based on knowledge represented using an ontology. However, the ontology is non-probabilistic, as well as the reasoning performed for the analysis.

Since ontologies are based on classical logic, i.e., the DL language, they enjoy the expressivity for defining complex knowledge and concepts of a domain. However, they lack the ability to express knowledge with probabilistic terms. In the next section, we propose MEBN as a model that overcomes this difficulty.

3 MULTI-ENTITY BAYESIAN NETWORKS

Multi-entity Bayesian networks are the outcome of the combination of Bayesian Networks with FOL. In other words, in MEBN, the capability of BN to model uncertainty is combined with the expressivity of FOL. From a stochastic Bayesian modeling perspective, the objective of using MEBN is to built situation specific BN (SSBN) that are customized according to the snap shot of the environment in this situation. In MEBN, FOL is extended to FOBL, which is used to express and represent knowledge. This overcomes the BN shortcoming of being very rigid and not flexible to model dynamic environments.

More specifically, from the FOL perspective, FOBL extends FOL with the ability to model uncertainty. A MEBN theory, expressed with FOBL, is the probabilistic analogous of a FOL theory. The key feature enabling a MEBN theory to model uncertainty is the treatment of sentences (e.g., predicates/relations) as random variables that can be instantiated to a specific value, belonging to a finite set. For example, the sentence *weatherGood(X,T)* means that at time *T* the weather is good for the *X* activity to be performed, but when treated as random variable it may take one of the following instances: *cloudy, sunny, warm*, according to a probability distribution. In the simplest case, a random variable can be Boolean, e.g., *weatherGood(T)* is a random variable that indicates that the weather is good or not at time *T*. The random variables/sentences represent attributes of entities and relations among them. In MEBN, the random variables comprise a DAG, where each random variable

depends only on its parents. A set with all random variables being instantiated is called an interpretation of the MEBN theory. Thus, the joint probability that combines all the local conditional distributions of the random variables assigns a probability to each interpretation.

From a Bayesian perspective, a MEBN theory (MTheory) is a collection of MEBN fragments (MFrags). An MFrag is the analogous of a BN node, since it includes *resident* nodes, for which a local conditional distribution is defined. These variables are conditioned on the *input* nodes, i.e., their the ancestors. Note that every input node in an MFrag is a resident node in another MFrag, which is called *home* MFrag. Thus, each MFrag is a fragment of the global Bayesian network consisting of all MFrag of the MTheory. Another component of an MFrag is the context terms which are used to specify logical constraints for the instantiations of the ordinary logical variables. We must note that, an ordinary variable can have multiple instantiations for a single MTheory interpretation. Since the random variables have logic variables (or functions) as arguments, the resulting SSBN contains multiple random variables, one for each argument instantiation. Nevertheless, local distributions can be designed with the ability to take into account the volatile number of parents. Note that this is the core mechanism provided by FOBL for the construction of SSBN.

The ultimate goal of modeling with MEBN is inference, which provides us with the ability to analyze observations, taken from the real world and draw conclusions. Inference is performed in SSBN that result based on evidence. There are two steps in SSBN construction. In the first step, new evidence are incorporated in the MTheory by instantiating ordinary variables, which enables the construction of the SSBN that models the specific situation. Other evidence are observations of a subset of the random variables. Thus, before inference, these variables are instantiated to their observations (i.e., they are the *observed* variables). Then, Bayesian inference provides the posterior probability distribution of the unknown random variables we want to estimate. This posterior distribution can also be used to learn the conditional distributions the target random variables.

With this position paper, we aim to highlight MEBN as a knowledge representation and analysis tool for the ICH domain. Specifically, our goal is to recognize ICH concepts, such as dance types, singing styles, etc., in ICH content, based on their formal definition in a knowledge representation schema. This schema has been developed with consultation from the domain knowledge experts and it is expected to

boost the analysis performance. The concepts meant to be detected are manifested in ICH content by performers (singers, dancers, etc.), which content comes with various modalities, e.g., audio, video, motion. Thus, we deal with multimodal analysis.

There are two main reasons we propose MEBN for this task. The first is the uncertainty modeling capability and the second is the situation specific construction of BN that MEBN provides. For example, regarding uncertainty, a movement pattern in a traditional dance is not always deterministically driven, since people may chose to follow a variant of the standard dance movement. Another source of uncertainty arises also from the inaccuracy of the basic concept detection algorithms, e.g., some steps that comprise the dance may not be detected. The second reason is that, usually, it is not *a priori* known how many dancers, singers, or even dances and songs are manifested in a particular timespan of ICH content. SSBN construction using FOBL can be proven very beneficial for dynamic modeling of such situations.

A concise example of MEBN usage for multimodal traditional dance modeling is presented in Figure 1. In this figure, an MTheory is presented consisting of three MFrags. The "Dance" MFrag contains one input random variable, *DanceType*, with its states being the "dance type" instances (Tsamiko, Calus, etc.). The resident random variables of this MFrag are *Steps(d,t1), Steps(d,t2), MusicTempo(d)*. The first two denote a sequence of steps and the third the music tempo accompanying the dance. Due to the recursive definition of the steps, there can be numerous steps, depending on the number of the time variables instances. These instances depend on the evidence being found in the ICH content (i.e., the time instants). Let us suppose a particular variable instantiation in order to demonstrate the mechanism that produces an SSBN. For simplicity of the demonstration, we will work only with the Dance MFrag. Suppose that the following logical entities have been found as evidence: $T0, T1, T2, Dance1$. This means that only one dance is performed. $d=Dance$, starting at time $T0$ and ending at $T2$ ($T0, T1, T2$ is an ordered sequence of time instants). Regarding random variables, the observations are three consecutive steps, resulting in the instantiations *Step(Dance1,T0)=Right, Step(Dance1,T1)=Right, Step(Dance1,T2)=Left*, and one tempo *MusicTempo(Dance1)=Slow*. The partial SSBN produced by these instantiations is depicted in Figure 2.

The instantiated random variables in the SSBN of Figure 2 are the observed variables. The goal is to estimate the posterior distribution of the Dance-Type variable given the observations. This is achieved

through Bayesian inference, based on the local conditional distribution of the *Step* variables, which is depicted in the *Steps* MFrag, and the *MusicTempo* local distribution. All these distributions are conditioned on the DanceType variable. In addition, each *Step* variable is conditioned on the previous, except that for which ($t_1 = T0$). Note that, in a more general example, the SSBN would contain nodes from various MFrags.

4 Existing Literature

A number of models similar to MEBN have been developed that represent probabilistic knowledge. OOBN models (Koller and Pfeffer, 1997) have been proposed as an alternative to standard BN for overcoming the inherent inflexible structure of BN. An OOBN object is an extension of a node in BN and it is a collection of attributes of a domain. This renders OOBN more flexible than BN to be customized to situations that arise. Also, OOBN enjoy many benefits similar to that of object oriented programming, such as model re-use.

Probabilistic relational models (PRMs) (Pfeffer, 2000) extend Bayesian networks more efficiently than OOBN. In particular, they introduce the concept of properties, and relations between them. Like a MEBN theory, a PRM represents a probability distribution over possible interpretations. Practically, the distribution assigns a probability to every possible relation instantiations of the relational schema. Inference algorithms in PRMs can provide, apart from the distribution, also the structure and the uncertainty underlying it. However, OOBN and PRM expressivity is inferior to MEBN, mainly due to the context constraints used to enforce logical contraints on the model variables.

In (Milch and Russell, 2007), a survey about programming languages combining, similar with FOBL, logic and probability can be found. In these languages, inference based on probability is integrated in the reasoning mechanism of this language. However, there are serious limitations in this type of modeling, such as the explicit definition of local distributions, which renders them sub-optimal for our purpose when compared to MEBN.

Ontologies with probabilistic extensions have been used for ICH knowledge representation. For example, in (Mallik et al., 2011), ontology-based mapping of cultural heritage content to ICH concepts is proposed. More specifically, an ontology-based framework is proposed that facilitates the representation of the ICH concepts, their properties and the relationships among them. The concepts are charac-
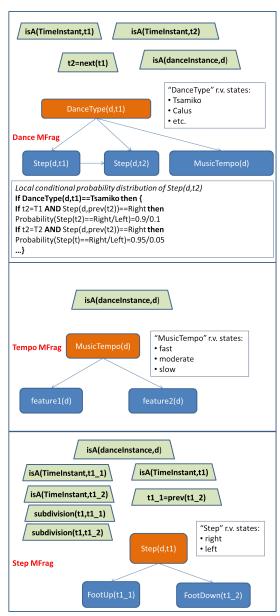

Figure 1: MEBN example for traditional dance knowledge representation: MTheory consisting of three MFrags. Random variable color meaning: orange is "input", blue is "resident", green is "context constraints". Only the local distribution of *Steps(d,t2)* is depicted.

teristics/features (low or high level) that constitute the semantics of the ICH content. Thus, the ontology used in this framework includes descriptions of domain concepts, which descriptions are given formally with terms of related low-level audio-visual features, appearing in the multimedia ICH content. In this way, a convenient semantic interpretation of the multimedia data is enabled. Moreover, a custom language, the Multimedia Web Ontology Language (MOWL),
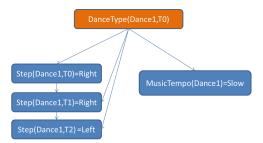
Figure 2: Example partial SSBN for evidence $t = \{T0, T1, T2\}, d = Dance1$.

is used to encode the domain knowledge.

In a closely related work (Mallik et al., 2008), a semi-automatic ontology construction methodology for ICH knowledge modelling is proposed. Particularly, in this work, the aim is to complement the knowledge assisted framework for the semantic analysis of cultural Indian dances presented in (Mallik et al., 2011), i.e., detection of specific dance styles and moves in multimedia with cultural content. Bayesian networks and probabilistic inference for the automatic analysis of the relations among the concepts defined and described by ICH experts is proposed to this end. Note, however, that although the ontology is constructed using probabilistic methods, i.e., as a BN of concepts and relations, the BN remains unchanged. This is a serious modeling handicap, which does not appear in MEBN.

5 DISCUSSION

The main challenges that we expect to encounter by the application of MEBNs for ICH are discussed next. First and foremost, a major challenge is, as stated previously, the inherent uncertainty in ICH knowledge. This stems from the fact that there are many cases where the performers implement an ICH activity (e.g., dance, singing) in a non-determenistic manner. For example, during the peformance of a traditional dance, although the number of the steps and their characteristics, e.g., direction, are predefined, the execution of each step may vary per person and/or occasion, for this specific dance. Moreover, in a dance, figures are performed spontaneously, rather in a predefined pattern. Thus, one aspect of this challenge is the inability of an expert to define rigously some parts of the ICH activity that are inherently uncertain and difficult to be described with deterministic terms. To demonstrate that using the previous example, although the direction of each step can be strictly defined (e.g., left or right), it is sometimes not feasible to describe quantitatively the height of the foot from the ground during a step, the step length, or even a precise motion trajectory of the dancers. With MEBN we can model this uncertain knowlege using the rigorous framework of probabilities and Bayesian networks.

Lastly, the fact that MEBNs incorporate the concept of ontologies into their modeling part enables the systematical collection and formal representation of the multiple concepts of ICH knowledge. Thus, with MEBN we aim to also exploit the useful feature of ontologies to facilitate a convenient and with a non prone to errors developement of the ICH knowledge representation part. Thus, MEBN combine, for this specific problem, the advantages of both worlds, i.e., knowledge expression and representation on the framework of ontologies and modeling uncertainty using probabilistic models (BN).

6 CONCLUSIONS

In this paper, we propose MEBN as an answer to the need of representing ICH domain knowledge and as a theoretical and practical tool to model the concepts and their relation between them. A key feature of MEBN that we find very useful for this case is their built-in capability to express uncertainty that is inherently present to the ICH concepts, as these are defined by the knowledge experts. Moreover, with MEBN we can model the temporal evolution of attributes of an ICH domain, such as the evolution of steps in a dance or the variations of pitch in a musical performance. Lastly, we mention models similar in spirit with MEBN that merge first order logic with probabilistic models.

ACKNOWLEDGEMENTS

REFERENCES

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Patel-Schneider, P. F. (2003). *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc.

Davis, E. (1990). *Representations of Commonsense Knowledge*. San Mateo, California: Morgan Kaufmann.

Doerr, M. (2003). The cidoc conceptual reference model: an ontological approach to semantic interoperability of metadata. *AI Magazine*, 69:(3): 75–92.

Ioannidis, K. E. R. Y. (2011). A labanotation based ontology for representing dance movement. In *9th International Gesture Workshop, Athens, Greece, May*.

Koller, D. and Pfeffer, A. (1997). Object-oriented bayesian networks. In *Uncertainty in Artificial Intelligence: Proceedings of the Thirteenth Conference, San Francisco, CA: Morgan Kaufmann*.

Laskey, K. B. (2008). Mebn: A language for first-order bayesian knowledge bases. *Artificial Intelligence*, 172:(2–3), 140178.

Mallik, A., Chaudhuri, S., and Ghosh, H. (2011). Nrityakosha: Preserving the intangible heritage of indian classical dance. *ACM Journal on Computing and Cultural Heritage*, 4:(3), 11.

Mallik, A., Pasumarthi, A. P., and Chaudhury, S. (2008). Multimedia ontology learning for automatic annotation and video browsing. In *1st ACM International Conference on Multimedia Information Retrieval (MIR'08), New York,*.

Milch, B. and Russell, S. (2007). First-order probabilistic languages: Into the unknown. In *In Proc. of the 16th International Conference on Inductive Logic Programming*, pages 10–24.

Pfeffer, A. (2000). *Probabilistic Reasoning for Complex Systems*. Stanford, CA, Stanford University.

Staab, S. and R. Studer, e. (2009). *Handbook on Ontologies, chapter 23*. International Handbooks on Information Systems, Springer.

UNESCO (2013). Convention of the safeguarding of intangible cultural heritage of unesco, 2003. URL: http://www.unesco.org/culture/ich/index.php.

Whittaker, J. (2009). *Graphical models in applied multivariate statistics*. Wiley Publishing.