Which Role for an Ontology of Uncertainty?

Paolo Ceravolo, Ernesto Damiani, Marcello Leida

Dipartimento di Tecnologie dell'Informazione - Università degli studi di Milano
via Bramante, 65 - 26013 Crema (CR), Italy
{ceravolo,damiani,leida}@dti.unimi.it

Abstract. An Ontology of Uncertainty, like the one proposed by the W3C's UR3W-XG incubator group, provides a vocabulary to annotate different sources of information with different types of uncertainty. Here we argue that such annotations should be clearly mapped to corresponding reasoning and representation strategies. This mapping allows the system to analyse the information on the basis of its uncertainty model, running the inference proccess according to the respective uncertainty. As a proof of concepts we present a data integration system implementing a semantics-aware matching strategy based on an ontological representation of the uncertain/inconsistent matching relations generated by the various matching operators. In this scenario the sources of information to be analyzed according to different uncertainty models are independent and no intersection among them is to be managed. This particular case allows a straight-forward use of the Ontology of Uncertainty to drive the reasoning process, although in general the assumption of independence among the source of information is a lucky case. This position paper highlights the need of additional work on the Ontology of Uncertainty in order to support reasoning processes when combinations of uncertainty models are to be applied on a single source of information.

1 Introduction

Information is hardly ever perfect or certain, specially if it belongs to an unsupervised environment. The theories and the models proposed so far for representing and managing information in a effective way can be successfully applied only in small-scale scenarios. Nature and features of information need to be analyzed taking into account several aspects: context, source of information, temporal location, dependencies and so on. Mathematical models for reasoning with uncertain information has been successfully applied in several situations. But a still open issue is to simultaneously consider different models for managing uncertain information and coordinate the different independent reasoning process by an explicit representation of the different uncertainties present in the knowledge base. Here we argue that the scope of an Ontology of Uncertainty should include this coordination. As a proof of concept of this approach we present, in this paper, a data integration system implementing a semantics-aware matching strategy, which is based on an ontological representation of the matching relations generated by the various matching operators where the semantics of each assertion is also represented explicitly as instances of an ontology. The uncertainty is assigned to each relation using SWRL rules, this allows to divide the knowledge base in sub-parts according with the specific uncertainty. The Ontology of Uncertainty, proposed by W3C's UR3W-XG incubator group, allows an explicit definition of the various types uncertainty. Assigning to each model a reasoner process it then possible to manage different independent sources of information. But the case we present is very particular because the independence of the various source of information. For this reason, our conclusions highlight the need of additional work on the Ontology of Uncertainty in order to support reasoning processes when combinations of uncertainty models are to be applied on a single source of information

2 Uncertain Information Representation and Reasoning

Uncertainty falls at meta-level respect to truth; it arises when the knowledge base does not provide sufficient information to decide if a statement is true or false in the actual situation of the system. Uncertainty can be encoded as the level of certainty of the system about a statement. Nature of uncertainty can be classified as **Epistemic**, if the uncertainty comes from the limited knowledge of the agent that generates the assertion or **Aleatory** if the uncertainty is intrinsic in the observed world. Moreover it is possible to identify two different source of uncertainty: **Objective** if the uncertainty derives from a repeatable observation and **Subjective** if the uncertainty in the information is derived from an informal evaluation. Nature of information can be: **Contingent** if it refers to a particular situation or instant or **Generic** if it refers to situations that summarize trends.Uncertainty moreover can depend on the type of information: **Ambiguous**, **Inconsistent**,**Vague**, **Incomplete** and **Empiric**. Depending on the type of uncertainty to deal with, a certain model is more suitable than another: Fuzzy theories, Probabilistic Theories and Possibility theory.

The need of a unified framework for dealing with gradual truth values and probabilities is arising but, as stated in [5] probability and possibility theories are not fully compositional with respect to all the logical connectives, without a relevant loss of expressiveness. This consideration leads to the consequence that uncertain calculi and degrees of truth are not fully compositional either. Nevertheless some work in this direction has been proposed by imposing restrictions to the expressiveness of the logics. The most relevant studies are: [7,8] where the authors define *probabilistic description logics programs* by combining stratified fuzzy description logics programs with respect to degrees of probabilities in a unified framework. In [4] a definition of possibilistic fuzzy description logics has been proposed by associating weights, representing degrees of uncertainty, to the fuzzy description logic formulas. An extension of the Fuzzy Description Logics in the field of Possibility theory has been presented also in [2]. The models available in literature are then able to deal with uncertainty and the progresses in theories for handling both uncertainty and vague truth values are remarkable. But the uncertainty that these models are able to manage has not been differentiated: as mentioned in Section 3 uncertainty is generated from different situations and has different semantics. For this reason the URW3-XG[1] proposed an ontology (Ontology of Uncertainty) as a generic meta-model for representing the semantics of the uncertainty in various assertions.This ontology is designed for a flexible environment, where different uncertainties can arise in the same knowledge base, so the selection of the correct model for inference is driven by the information in the ontology. But the URW3-XG incubator group did not specify how to deal with situations where more than one model is involved in the inference process. Hybrid theories are considered in the set of possible models as a separate model and new sub concepts of this category can be easily added when a new model appears. The reasoning becomes complex if the result of a inference in a specific reasoning process is dependent to the result of another reasoning process. In [9] the authors propose a framework for sharing information between three different models of uncertainty, where the fuzzy linguistic truth values are propagated through the three models in a nonmonotonic way, by exploiting the extension principle [12] and aggregation of linguistic values. This approach is promising but it is grounded to fixed fuzzy values (linguistic truth) that are used by all the different models and then aggregated according to nonmonotonic rules. In literature we are not aware of hybrid reasoning processes, which can handle a flexible integration of different models. In [11,9,10,1] the interoperability has been studied and defined on a set of selected inference models. Adding new models to the framework can easily result in a revision of the underlying theory.

3 A Use Case in Data Integration

Let us consider the problem of matching heterogeneous data to externally defined types, such as ontology classes. In the case of our matching strategy, the inference process breaks down in two different steps: first step is to divide the knowledge base in sub sets according to the specific model, and the second step is to aggregate the result of the independent inference processes. In our scenario, the various reasoning processes are independent; this important premise allows us to use the Ontology of Uncertainty to partition the various matching relations according to the model used for the reasoning process. The knowledge base containing the necessary information for our matching strategy, it is composed by a set of statements that are represented by the concept `Sentence` in the Ontology of Uncertainty. To each statement the information about the uncertainty is explicitly defined by the concept `Uncertainty` that defines the correct semantics.

In this example, Ontology of Uncertainty is used basically to drive the reasoning process: each type of uncertainty is processed by its specific reasoner and a subsequent process, based on SWRL rules, integrates the results of the various reasoners. In our system we consider a Probabilistic Description Logic reasoner [6], a Fuzzy Description Logic [3].

[1] `http://www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331/`

The first part of the matching strategy is to assign to the various assertions (`Sentence`), the correct information about its uncertainty semantics. This information is classified according to a set of pre-defined SWRL rules that assigns the correct semantics in relation to the matching operator (`Agent`) that has generated the relation; in relation to the presence of a degree of probability and in relation to the level of inconsistency among matching relations. There is one or more rules for each specific uncertainty type, nature, model and derivation.

According to the information that has been provided to each reasoner, the process has to return back to the matching strategy the set of assertions that they believe to be the most trustable ones.

Once the various reasoning processes come to an end, the results are propagated back to the matching ontology by a Reconciliation process. In the case of our matching strategy we make use of SWRL rules to aggregate the results.

4 Conclusions

We presented the idea of extending the scope of the Ontology of Uncertainty for hybrid reasoning under managing different types of uncertainty and showed a simple application of this idea to Schema Matching. The main constraint for the use of the Ontology of Uncertainty it is to make the reasoning process independent to each other, so that no interdependencies between assertions inferred by the reasoner can happen. This way after the various reasoning processes the Reconciliation Process is reduced to handle possible inconsistencies by SWRL rules. The Ontology of Uncertainty does not specifies for each semantics of uncertainty a particular reasoner to use. This limits the use of the Ontology of Uncertainty in real world situations, which differs from the lucky case of our matching strategy. In our opinion, the Ontology of Uncertainty has to provide further information about how the various reasoning processes exchange dependent information.

References

1. Grigoris Antoniou and Antonis Bikakis. Dr-prolog: A system for defeasible reasoning with rules and ontologies on the semantic web. *IEEE Trans. Knowl. Data Eng.*, 19(2):233–245, 2007.
2. Fernando Bobillo, Miguel Delgado, and Juan Gómez-Romero. Extending fuzzy description logics with a possibilistic layer. In *URSW*, volume 327 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007.
3. Fernando Bobillo and Umberto Straccia. fuzzydl: An expressive fuzzy description logic reasoner. In *2008 International Conference on Fuzzy Systems (FUZZ-08)*, pages –. IEEE Computer Society, 2008.
4. Didier Dubois, Jérôme Mengin, and Henri Prade. Possibilistic uncertainty and fuzzy features in description logic. A preliminary discussion. In E. Sanchez, editor, *Fuzzy logic and the semantic web*, pages 101–113. Elsevier, http://www.elsevier.com/, 2006. DMenP001.
5. Didier Dubois and Henri Prade. Can we enforce full compositionality in uncertainty calculi. In *In Proc. of the 11th Nat. Conf. on Artificial Intelligence (AAAI-94*, pages 149–154. AAAI Press / MIT Press, 1994.

6. Pavel Klinov. Pronto: Probabilistic dl reasoning with pellet, 2008.
7. Thomas Lukasiewicz and Umberto Straccia. Description logic programs under probabilistic uncertainty and fuzzy vagueness. In *ECSQARU*, pages 187–198, 2007.
8. Thomas Lukasiewicz and Umberto Straccia. Uncertainty and vagueness in description logic programs for the semantic web, 2007.
9. X. Luo, C. Zhang, and N. R. Jennings. A hybrid model for sharing information between fuzzy, uncertain and default reasoning models in multi-agent systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(4):401–450, 2002.
10. William James Van Melle. *A domain-independent system that aids in constructing knowledge-based consultation programs*. PhD thesis, Stanford, CA, USA, 1980.
11. E. H. Shortliffe and B. G. Buchanan. A model of inexact reasoning in medicine. pages 259–275, 1990.
12. L.A. Zadeh. Fuzzy sets. *Information Control*, 8:338–353, 1965.