Towards Combining Ontologies and Uncertain Knowledge

Joana Hois
SFB/TR8 Spatial Cognition, University of Bremen

Abstract

In this paper, we analyze why and how formal ontologies and probability theories should be combined and give an example of such a combination in a spatial system.

1 Introduction

In the field of artificial intelligence, ontologies are basically defined as a set of concepts and relations among them following certain axiomatic rules that can give further information about the semantics of the concepts and their relations. Ontologies can be formalized in different sorts of logic, such as DL, FOL, modal logic or spatial-temporal logic, therefore, they differ in their expressiveness, decidability, and tools for representation and reasoning support [7].

Uncertainty theories, such as Bayes theory [2], Dempster-Shafer theory [6] or fuzzy set theory [8], define how to deal with uncertain, insecure or vague knowledge. They provide a representation formalism for uncertain information (usually expressed by a value between 0 and 1) and reasoning strategies.

There are only few approaches that try to combine ontologies and probability theories: The ontology language OWL[1] can be extended by uncertain values that are based on Bayesian networks shown in [1], a first exploration in [3] combines OWL ontologies and the Dempster-Shafer theory, and the more expressive language SWRL[2] has been extended to represent fuzzy degrees [5]. While the first two approaches are introducing specific concepts, that are meant to represent the uncertain values, on a pure semantic level, the latter approach adds a fuzzy value to each class membership in order to represent the fuzziness of the type of an instance.

In this paper, however, we would like to concentrate on more general aspects of the combination of both theories. We will point out why they should be combined at all and what exactly should or should not be *uncertain* in the ontology.

2 Need for Combination

Ontologies in philosophy represent what is in the world, and in a more specific view in AI they represent what is in the domain of the application. In both cases, they are designed in order to represent strict but not uncertain information. That an ontology reflects things that can exist, their attributes, relations and dependencies, i.e., a meta-description of all concepts, directly implies that there cannot exist uncertain values, as everything is known because it is well-defined.

But this seems to be a drawback for concrete applications: Within a real world system, different sources for uncertain or vague information influence an agent's knowledge base. It can refer to incomplete knowledge about the environment of an agent, as it might not have enough sensori-motor-activity modules to explore all of its surroundings - and even if the agent had enough exploration facilities, the measurements could be wrong or imprecise. If the system can interact with other agents and gets information from foreign resources, this data might be unreliable and doubtful (again uncertain). Results of future plans, forthcoming events, or expectations of executable actions are not always predictable and therefore uncertain as well. Further sources of errors can be caused by wrong reasoning or ontological representation.

In general, however, an agent benefits from accessing an ontology as background knowledge of its domain. It provides a formal representation of the world, defines possible conditions, and gives a clear specification of any relevant term. But the agent also has to store uncertain information about its current knowledge. This strong intersection between ontological and uncertain parts indicates the combination of both theories.

3 Considerations for Combination

As we have seen, there are basically two layers that make up an agent's knowledge: strict and well-defined background knowledge and imprecise knowledge of a situation. In case of an ontology this reflects the distinction between the TBox and the ABox. The TBox stores terminological knowledge, i.e. the statements, namely concepts and relations, about the domain. The ABox stores assertions, i.e. the facts, namely instantiations of concepts and relations, about the domain.

Therefore, the probability values of imprecise knowledge should only affect the ABox of an ontology and not the TBox. This distinction is illustrated in Fig. 1. Although the TBox layer has to define the way probabilities are instantiated and the concepts and properties that might be uncertain, the TBox

[1] http://www.w3.org/2004/OWL/

[2] http://www.daml.org/rules/

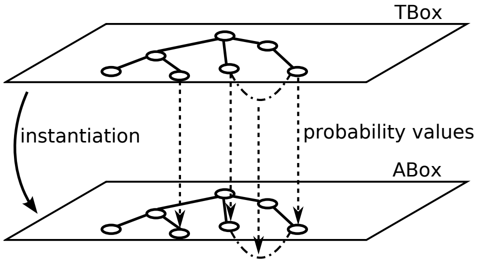

Figure 1: Relationship between certain concepts in the TBox and uncertain instances in the ABox

does not include probabilities itself. Such probabilities instead can indicate either uncertainties about a class membership of an instantiation or uncertainties about relations between instances. Accordingly, the instance of a class can additionally be specified with a probability (depending on the probability theory used), as well as the instance of a property, because it indicates either an attributes or a relation between instances. The representation of such an additional value is similar to the approach in [5].

4 Example

An example of an application using an ontology together with uncertain knowledge is a spatial system that has to perceive and recognize its surrounding entities, interact with its environment, reason about the current situation, and perform its tasks. Certain knowledge about this environment is given in the ontology of the agent about types of objects that can occur, about physical conditions, about possible actions, or about spatial relations between objects and their linguistic expressions. Imprecise knowledge can result from insufficient object recognition processes or wrongly measured distances, on the one side, and reasoning under uncertainties as in case of natural language interaction, on the other.

In the example shown in Fig. 2 an agent has to perform spatial tasks and interact with its natural environment. It has to determine spatial relations between several entities (item A, B, C, and a human user, that have to be recognized as well), not only with respect to quantitative distances but also qualitative, linguistic descriptions in order to interact with the user. These natural language expressions can differ in many respects, as described in [4], and can also suffer from uncertainty: While the relations between the agent and the entities from the perspective of the agent can be generated directly, the relations between two entities are not as clear. Such relations depend on the perspective, i.e. from the viewpoint of the user or from an intrinsic front of the entity or with respect to a third entity, and they depend on further attributes, i.e. certain entity types cannot act as a reference to other entities when they differ in size, function, or other characteristics (it is more common to say that "a mouse is to the right or left of a laptop" than "a laptop is to the right/left of a mouse"; in some specific cases this might be more applicable, though).

Therefore generation and processing of such expressions have to be analyzed with respect to object types, perspectives, aims, or user specific knowledge, and uncertain information

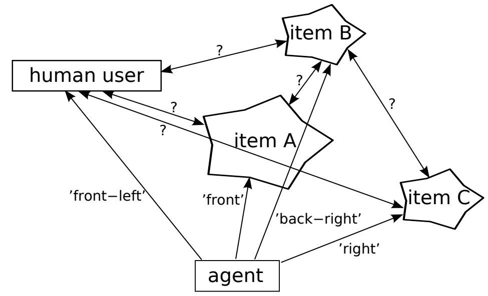

Figure 2: An agent in a spatial environment surrounded by other entities, arrows indicate spatial relations

can occur in any of these factors.

5 Conclusion

A lot of investigations on the combination between ontologies and probabilities still has to be done. In this paper, we have pointed out some of their general considerations, which parts of the ontology might actually be uncertain, and why this combination should be developed at all. Future work will have to address implementation aspects with respect to (editing) tools and reasoning support for such ontologies.

Acknowledgments

The Collaborative Research Center for Spatial Cognition (SFB/TR8) of the university of Bremen and the university of Freiburg is funded by the Deutsche Forschungsgemeinschaft (DFG), whose support we gratefully acknowledge.

References

[1] Zhongli Ding, Yun Peng, and Rong Pan. A Bayesian Approach to Uncertainty Modeling in OWL Ontology. In *Proc. of the Int. Conf. on Advances in Intelligent Systems - Theory and Applications*, page 9. IEEE, 2004.

[2] John A. Hartigan. *Bayes Theory*. Springer-Verlag, 1983.

[3] Joana Hois, Kerstin Schill, and John A. Bateman. Integrating uncertain knowledge in a domain ontology for room concept classifications. In *Research and Development in Intelligent Systems XXIII, Proc. of AI-2006*, pages 245–258. Springer-Verlag, 2007.

[4] Stephen C. Levinson. *Space in Language and Cognition*. Cambridge University Press, 2003.

[5] Jeff Z. Pan, Giorgos Stoilos, Giorgos Stamou, Vassilis Tzouvaras, and Ian Horrocks. f-SWRL: A Fuzzy Extension of SWRL. *Data Semantics, special issue on Emergent Semantics*, 4090:28–46, 2006.

[6] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.

[7] John F. Sowa. *Knowledge Representation - Logical, Philosophical, and Computational Foundations*. Brooks/Cole, 2000.

[8] Lotfi A. Zadeh and Janusz Kacprzyk. *Fuzzy logic for the management of uncertainty*. Wiley, 1992.