Chapter 7.3
Uncertainty Representation and Reasoning in the Semantic Web

Paulo Cesar G. Costa
George Mason University, USA

Kathryn Blackmond Laskey
George Mason University, USA

Thomas Lukasiewicz
Oxford University Computing Laboratory, UK

ABSTRACT

This chapter is about uncertainty representation and reasoning for the Semantic Web (SW). We address the importance, key issues, state-of-the-art approaches, and current efforts of both the academic and business communities in their search for a practical, standard way of representing and reasoning with incomplete information in the Semantic Web. The focus is on why uncertainty representation and reasoning are necessary, its importance to the SW vision, and the major issues and obstacles to addressing uncertainty in a principled and standardized way. Although some would argue that uncertainty belongs in the "rule layer" of the SW, we concentrate especially on uncertain extensions of ontology languages for the Semantic Web.

DOI: 10.4018/978-1-60566-112-4.ch013

WHY CARE ABOUT UNCERTAINTY?

After some years of SW research, the subject remains rife with controversy, and there is still some disagreement on how uncertainty should be handled in SW applications. Thus, it is no surprise that little was said on the subject in previous chapters of this book. A major reason for the present state of affairs is that the most popular technologies employed in SW applications are rooted in traditional knowledge representation formalisms that have historically ignored uncertainty. The most compelling examples are Frame Systems (Minsky, 1975), and Description Logics, which evolved from the so-called "Structured Inheritance Networks" (Brachman, 1977), and form the logical basis for the ontology language OWL.

The spotlight is not on the *status quo*, but on what the future holds. To answer this question, we

begin with a comprehensive analysis of the major challenges to be faced by the SW community, including what kinds of interactions, scenarios, demands, and obstacles must be addressed to make the SW promises a reality. Next, we assess whether protocols that rely only on complete, deterministic information will suffice to address these challenges. Although much progress has been made by tackling problems in which uncertainty is inessential or can be circumvented, addressing the full range of challenges inherent in the Semantic Web vision will require optimal use of *all* available information. In this Chapter, we argue that a principled framework for representing and reasoning with incomplete information is necessary to realizing the SW vision. Because uncertainty is a ubiquitous aspect of most real-world problems, any representation scheme intended to model real-world entities, properties and processes must be able to cope with uncertain phenomena. Current SW technologies' inability to represent and reason about uncertainty in a sound and principled manner raises an unnecessary barrier to the development of new, powerful features for general knowledge application, a limitation that threatens to derail the original vision for the Semantic Web as a whole. In other words, we argue that realizing the SW as envisioned by Tim Berners-Lee (Berners-Lee & Fischetti, 2000) requires a principled framework for representing and reasoning with uncertainty.

The Semantic Web envisions effortless cooperation between humans and computers, seamless interoperability and information exchange among web applications, and rapid and accurate identification and invocation of appropriate Web services. While considerable progress has been achieved toward realization of the Semantic Web vision, it is increasingly apparent that a sound and principled technology for handling uncertainty is an important requirement for continued progress. Uncertainty is an unavoidable factor in knowledge interchange and application interoperability. Different applications have different ontologies, different semantics, and different knowledge and data stores. Legacy applications are usually only partially documented and may rely on tacit usage conventions that even proficient users do not fully understand or appreciate. Furthermore, data that is exchanged in the context of the semantic web is often incomplete, inconsistent, and inaccurate. This suggests that recent work in the application of probability, fuzzy logic, and decision theory to complex, open-world problems could be of vital importance to the success of the Semantic Web. Incorporating these new technologies into languages, protocols, and specifications for the Semantic Web is fundamental to realizing the Semantic Web vision.

Typical Problems Needing Uncertainty Representation and Reasoning. The following web-relevant reasoning challenges illustrate the kinds of problems for which reasoning under uncertainty is important.

- Information extracted from large information networks such as the World Wide Web is typically incomplete. The ability to exploit partial information is useful for identifying sources of service or information. For example, the fact that an online service deals with greeting cards may be evidence that it also sells stationery. It is clear that search tools capable of utilizing probabilistic knowledge could increase search effectiveness.
- Much information on the World Wide Web is likely to be uncertain. Common examples include weather forecasts and gambling odds. A canonical method for representing and integrating such information and the uncertainty associated with it is necessary for communicating such information in a seamless fashion.
- Web information is also often incorrect or only partially correct, raising issues related to trust or credibility. Uncertainty representation and reasoning helps to resolve