Chapter 2.8
Semantics for the
Semantic Web:
The Implicit, the Formal, and the Powerful

Amit Sheth
University of Georgia, USA

Cartic Ramakrishnan
University of Georgia, USA

Christopher Thomas
University of Georgia, USA

ABSTRACT

Enabling applications that exploit heterogeneous data in the Semantic Web will require us to harness a broad variety of semantics. Considering the role of semantics in a number of research areas in computer science, we organize semantics in three forms—implicit, formal, and powerful—and explore their roles in enabling some of the key capabilities related to the Semantic Web. The central message of this chapter is that building the Semantic Web purely on description logics will artificially limit its potential, and that we will need to both exploit well-known techniques that support implicit semantics, and develop more powerful semantic techniques.

INTRODUCTION

Semantics has been a part of several scientific disciplines, both in the realm of Computer Science and outside of it. Research areas such as information retrieval (IR), information extraction (IE), computational linguistics (CL), knowledge representation (KR) artificial intelligence (AI), and data(base) management (DB) have all addressed issues pertaining to semantics in their own ways.

Most of these areas have very different views of what "meaning" is, and these views are all built on some meta-theoretical and epistemological assumptions. These different views imply very different views of cognition, of concepts, and of meaning (Hjorland, 1998). In this chapter, we organize these different views to three forms of semantics: implicit, formal, and powerful (a.k.a. soft). We use these forms to explore the role of semantics that go beyond the narrower interpretation of the Semantic Web (that involve adherence to contemporary Semantic Web standards) and encompass those required for a broad variety of semantic applications. We advocate that for the Semantic Web (SW) to be realized, we must harness the power of a broad variety of semantics encompassing all three forms.

IR, IE, and CL techniques primarily draw upon analysis of unstructured texts in addition to document repositories that have a loosely defined and less formal structure. In these sorts of data sources, the semantics are *implicit*.

In the fields of KR, AI, and DB, however, the data representation takes a more formal and/or rigid form. Well-defined syntactic structures are used to represent information or knowledge where these structures have definite semantic interpretations associated with them. There are also definite rules of syntax that govern the ways in which syntactic structures can be combined to represent the meaning of complex syntactic structures. In other words, techniques used in these fields rely on *formal semantics*.

Usually, efforts related to formal semantics have involved limiting expressiveness to allow for acceptable computational characteristics. Since most KR mechanisms and the relational data model are based on set theory, the ability to represent and utilize knowledge that is imprecise, uncertain, partially true, and approximate is lacking, at least in the base/standard models. However, there have been several efforts to extend the base models (e.g., Barbara, Garcia-Molina, & Porter, 1992). Representing and utilizing these types

of more powerful knowledge is, in our opinion, critical to the success of the Semantic Web. Soft computing has explored these types of powerful semantics. We deem these *powerful (soft)* semantics as distinguished, albeit not distinct from or orthogonal to *formal* and *implicit semantics*.

More recently, semantics has been driving the next generation of the Web as the Semantic Web, where the focus is on the role of semantics for automated approaches to exploiting Web resources. This involves two well-recognized, critical enabling capabilities—ontology generation (Maedche & Staab, 2001; Omelayenko, 2001) and automated resource annotation (Hammond, Sheth, & Kochut, 2002; Dill et al., 2003; Handschuh, Staab, & Ciravegna, 2002; Patil, Oundhakar, Sheth, & Verma, 2004), which should be complemented by an appropriate computational approach such as reasoning or query processing. We use a couple of such enabling capabilities to explore the role and importance of all three forms of semantics.

A majority of the attention in the Semantic Web has been centered on a logic-based approach, more specifically that of description logic. However, looking at past applications of semantics, it is very likely that more will be expected from the Semantic Web than what the careful compromise of expressiveness and computability represented by description logic and the W3C adopted ontology representation language OWL (even its three flavors) can support. Supporting expressiveness that meet requirements of practical applications and the techniques that support their development is crucial. It is not desirable to limit the Semantic Web to one type of representation where expressiveness has been compromised at the expense of computational property such as decidability.

This chapter is not the first to make this above observation. We specifically identify a few. Uschold (2003) has discussed a semantic continuum involving informal to formal and implicit to explicit, and Gruber (2003) has talked about informal, semi-formal, and formal ontolo-

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/semantics-semantic-web/27433

Related Content