

Thesis: BiTRDF: Extending RDF for BiTemporal Data

Abstract: The Internet is not only a platform for communication, transactions, and cloud storage, but it is also a large knowledge store where people as well as machines can create, manipulate, infer, and make use of data and knowledge. The Semantic Web was developed for this purpose. It aims to help machines understand the meaning of data and knowledge so that machines can use the data and knowledge in decision making. The Resource Description Framework (RDF) forms the foundation of the Semantic Web which is organized as the Semantic Web Layer Cake. RDF is limited and can only express a binary relationship with the format of <subject predicate object>. However, expressing higher order relationships requires reification which is very cumbersome. Naturally, time varying data is very common and cannot be represented by only binary relationships. We first surveyed approaches that use reification or extend RDF for higher order relationships. Then we proposed a new data model, BiTemporal RDF (BiTRDF), that incorporates both valid time and transaction time explicitly into standard RDF resources. We defined the BiTRDF model with its elements, vocabulary, semantics and entailment, and the BiTemporal SPARQL (BiT-SPARQL) query language. We also explored different approaches to implement the BiTRDF model and evaluated their performances based on memory consumption, time for loading, and time for searching. We concluded this thesis with potential research directions. This thesis lays the foundation for a new approach to easily embed any or more dimensions, such as temporal data, spatial data, probabilistic data, confidence levels, etc.

Committee:

- Professor Abdullah Uz Tansel, Baruch College
- Professor Sarah Zelikovitz, College of Staten Island
- Professor Susan P. Imberman, College of Staten Island

Outside member:

- Professor Wlodek Zadrozny, University of North Carolina at Charlotte