Domain Ontology Construction using Formal Concept and Relational Concept Analysis

M.Farida Begam
Department of Information Science and Engineering
CMR Institute of Technology
Bengaluru, India
farida.b@cmrit.ac.in

Abstract - **Domain ontology construction is an important task in knowledge management applications. Knowledge representation in appropriate form is mandate requirement using which extraction of functional facts and applying them in various business operations is feasible. Semantic web technologies are boon to the technical community that develops applications based on knowledge management. Ontology is the mean by which knowledge can be captured and queried in well-defined manner. Developing domain ontologies and investigation of domain knowledge from huge data set/corpus is the tedious task. Formal Conceptual Analysis (FCA) and Relational Concept Analysis(RCA) are data analysis methods that can be applied to domain ontology construction and information retrieval. The concept lattices generated are used for domain ontology construction. We proposed a semi-automated methodology for generating concept lattices based on FCA and RCA techniques for Tree data set. Data about the for various trees found in the world are taken into consideration and their attributes are collected from Internet. We obtain concept lattices and association rules from FCA. Modeling based on RCA has been carried out and resultant concept lattices are generated.**

Index Terms - *Ontology, Lattice, Concept Analysis, FCA, RCA, Pruning, Association rule mining .*

I. INTRODUCTION

Profound applications of ontology have been identified in the fields like knowledge management, information retrieval, information integration, bioinformatics Software Engineering and e-learning. Ontology is the way to represent the knowledge that can be shared and reused for various knowledge management applications. Domain ontology is a collection of vocabularies and the specifications of the conceptualization of a given domain[1]. Domain ontology is suitable for the applications which need artificial reasoning and information retrieval. It provides deep understanding of the concepts, axioms and relationships among them. Manual construction of domain specific ontology incurs more cost. This work can be automated by extracting the terms and their relationships from text corpus or data set. Domain ontology construction can be well facilitated with the help of Formal Concept Analysis (FCA) and Relational Concept Analysis(RCA). Concept Analysis with the generalization and specialization relationships between them forms a concept lattice with inherent mathematical structure. FCA is a mathematical field helps analysis of concepts in terms of object-attribute form.

Knowledge acquisition technique which is based on FCA and RCA can be devised to automate this ontology construction purpose. It focuses on creating implied relationship among objects and has widespread usage in various knowledge management processes such as knowledge representations, knowledge acquisitions, and visualization. It defines relationships between set of objects and set of attributes. It groups the objects based on their attributes. It performs actually a kind of classification based on common attributes. It produces two outputs: i) Concept Lattice which can be represented as a collection of formal concepts in an ordered hierarchy of subconcept - superconcept relation ii) Attribute Implications that describes the dependency. FCA is a algebraic way of knowledge processing and one of the powerful tools in conceptual processing. It integrates three ideas of data and Knowledge, i.e., Finding the concepts in data, 2. Reasoning with dependencies in data, 3. Visualization of data, concepts, and dependencies with generalization and specialization capabilities. Relational concept analysis (RCA) supports exploratory search in related data set. RCA differs from FCA in terms handling multi-relational datasets. It deals with various objects which are defined with its' own set of attributes and relationships among the objects. RCA builds concept lattices for objects using interactive approach.

In this paper, we devised a semi-automatic technique for constructing the domain specific ontology for Tree data set. We obtained the implied relationships among trees especially Indian Trees. We have also verified the ontology using Description logics. We manually applied pruning and refactoring method for verification and validation of the domain ontology created

In section 2 We deal with the related work happened in this field and, in section 3, we discuss the theoretical background and basic concepts of FCA and RCA, in Section 4 the actual methodology for domain ontology construction has been devised. Section 5 deals with the concepts of how the generated lattices can be converted to ontologies. Section 6 is meant for conclusion.

II. RELATED WORK

In this section, we will discuss about the various works that have been already implemented in this field. In paper[2], FCA based ontology construction method is implemented. The integrated framework is developed for structural and