SEMANTIC WEB BASED KNOWLEDGE REPRESENTATION FOR NUCLEAR REACTOR DOMAIN

By

N. Madurai Meenachi

Enrolment Number: ENGG02200704015

Indira Gandhi Centre for Atomic Research, Kalpakkam, India

A thesis submitted to the

Board of Studies in Engineering Sciences

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

January 2014

Recommendations of the Viva Voce Board

As members of the viva voce board, we certify that we have read the dissertation prepared by **N. Madurai Meenachi** entitled "Semantic Web Based Knowledge Representation for Nuclear Reactor Domain" and recommend that it may be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

____ Date:
Chairman - **Dr. T. Jayakumar**

____ Date:
Guide / Convenor - **Dr. M. Sai Baba**

____ Date:
Member 1 - **Dr. B. K. Panigrahi**

____ Date:
Member 2 - **Dr. B. P. C. Rao**

____ Date:
External Examiner - **Prof. G. Sivakumar**

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it may be accepted as fulfilling the dissertation requirement.

Date:

Place: Indira Gandhi Centre for Atomic Research, Kalpakkam

Dr. M. Sai Baba
Guide

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under the rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author.

N. Madurai Meenachi

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

N. Madurai Meenachi

List of Publications arising from the thesis

Journals

1. "A Survey on Usage of Ontology in Different Domains", N. Madurai Meenachi and M. Sai Baba, *International Journal of Applied Information Systems*, **2012**, *9.*, 46-55.

2. "Web Ontology Language Editors for Semantic Web - A Survey", N. Madurai Meenachi and M. Sai Baba, *International Journal of Computer Applications*, **2012**, *53.*, 12-16.

3. "Quick Mapping Evaluator – Application for Ontology Mapping Evaluation: A Case Study on the Nuclear Reactor Domain", N. Madurai Meenachi, N.Hari Prasad and M. Sai Baba, *International Journal of Nuclear Knowledge Management*, **2013**, *6.*, 143-154.

4. "A Semantic Web based Knowledge Portal for Fast Breeder Nuclear Reactor Domain (KMNuR)", N.Madurai Meenachi and M. Sai Baba, *DESIDOC Journal of Library and Information Technology* ,**2014**, 34(5)., 426-434.

5. "Matrix Rank Based Ontology Matching: an Extension of String Equality Matching", N. Madurai Meenachi and M. Sai Baba, *International Journal of Nuclear Knowledge Management* (in press).

Conferences

1. "Knowledge Management Security for Nuclear Reactor", N. Madurai Meenachi, M. Sai Baba, B.Babu, B.Anandapadmanaban and V.Ramanathan, Presented at eINDIA 2009, Hyderabad, 25-27 August 2009.

2. "Knowledge Management and Ontology Representation For Fast Breeder Test Reactor", N. Madurai Meenachi, M. Sai Baba, B.Babu, B. Anandapadmanaban and V.Ramanathan, Presented READIT, Kalpakkam, 29-30 December 2009.

3. "Intelligent Agents for Nuclear Domain", D. Persis, N. Madurai Meenachi and M. Sai Baba, Presented at READIT, Kalpakkam, 29-30 December 2009.

4. "Semantic Digital Library - Migration of Dublin Core to RDF", E.Soundarajan, N. Madurai Meenachi and M. Sai Baba, Presented at IEEE International Conference on Signals and Image Processing, RMD Engineering College, Chennai, 250 – 254, 15-17 December 2010.

5. "Ontology Based Association Rule Mining For Library Domain", N.Madurai Meenachi and M. Sai Baba, Presented at READIT, Kalpakkam, 27-29 December 2011.

6. “Implementing Semantic Search in Library Web Portal”, N.Madurai Meenachi and M. Sai Baba, International Conference on Trends in Knowledge and Information Dynamics, Bangalore, 10-13 July 2012.

7. “Ontology Based Knowledge Sharing and Ontology Matching Algorithm”, N.Madurai Meenachi and M. Sai Baba, Presented at READIT, Kalpakkam, 17-19 December 2013.

8. “Methodology for Semantic Web Based Domain Knowledge Implementation”, N.Madurai Meenachi and M. Sai Baba, Presented at READIT, Kalpakkam, 17-19 December 2013.

9. “An Analysis of Semantic Agent for Implementing Library Domain Portal”, N.Madurai Meenachi and M. Sai Baba, Presented at SACTiL, VECC, Kolkatta, 28-30 January 2014.

10. “Pareto based Ontology Matching for Nuclear Reactor Domain”, N. Madurai Meenachi and M.Sai Baba, NACLIN 2014, Pondicherry, 9-11 December 2014.

N. Madurai Meenachi

DEDICATED

To

SHIRDI SAI BABA

ACKNOWLEDGEMENT

A journey is easier when we travel together. This thesis is the result of many years of work whereby I have been accompanied and supported by many people. It is a pleasant aspect that I have now the opportunity to express my gratitude to them.

I bestow my gratitude to the thesis supervisor, **Dr. M. Sai Baba** for his profound concerted guidance, encouragement and active support throughout the course of this doctoral work. His trust and scientific excitement inspired me in the most important moments of making right decisions and I am glad to work with him.

Special thanks go to **Dr. P. Swaminathan**, Former Group Director, EIG who provided me with the necessary guidance, motivation and support to register for Ph.D.

I am extremely thankful to my doctoral committee Chairman **Dr.T.Jayakumar** and committee members **Dr. B. K. Panigrahi** and **Dr. B. P. C. Rao** for their support and encouragement during my research work.

My sincere thanks go to **Dr. P. R. Vasudeva Rao**, Director, IGCAR for providing excellent environment to carry out research work.

Foremost thanks to **Prof. S. Thiruvengadam,** Retired Professor of Electrical Engineering and Dean (Admin), NIT, Calicut, for his overly enthusiasm, meticulate ideas and integral view on research which stirred me to achieve this milestone.

I place on record, my sincere thanks to **Dr. Vidya Sundararajan**, Head, SPS, for her support and encouragement.

I am deeply grateful for the support of my friends **Ms.E.Radha, Mr.S.Siva Kumar, Dr.S.Seyezhai, Dr.Kiruthika, Mr.K.Ravi Kumar, Ms.T.Sathiya Sheela, Mr.D.Sunil Kumar, Ms.M.M.Shanthi, Mr.G.Jagannathan, Mr.S.Muthu Kumar, Ms.Leena Mohan, Ms.K.Suganthi, Ms.C.P.Deepa.** Furthermore, I would like to thank my colleagues at ROMG and RMG for their support and encouragement.

The episode of acknowledgement would not be complete without the mention of my family. I owe the largest debt of gratitude to my loving mother who was the pillar of strength behind all my achievements in life. Heartful thanks to my husband and my father-in-law for giving me confidence and a new source of inspiration to my life.

N. Madurai Meenachi

TABLE OF CONTENTS

SYNOPSIS

Knowledge management is the process through which organizations generate value from their intellectual and knowledge based assets. The system that is developed to capture the knowledge involves acquiring, creating, synthesizing, representing, contributing and utilizing available experience and expertise to achieve the organizational goal. Sharing of knowledge would be helpful in designing a sound knowledge management system. It facilitates preserving of both tacit and explicit knowledge and converting the available data into a reusable knowledge repository for the future use. Explicit knowledge includes reports, which can be easily described and coded for storage and dissemination. The tacit knowledge on the other hand is highly personal, context specific, and can be manifested only through sharing, applying and practicing and cannot be easily coded.

The effective knowledge representation is achieved by overcoming the lack of precision. In the present era, it can be done by enabling the machine to process the available knowledge. As it is not an easy task to teach machines to comprehend natural languages, the way the information is provided to the machine has to be analysed. This is realisable by semantic web, which is an extension of current web where machine understands the concept and processes the knowledge. In this respect, languages used to represent the knowledge like RDF (Resource Description Framework), OWL (Web Ontology Language), XML (eXtensible Markup Language), F-logic, ObjectLogic etc. are used to feed the collected requisite knowledge about a particular domain to the machine.

Ontology is an explicit formal specification of the terms in the domain and relations among them. It is also defined as a set of assertions and relations among the objects for specifying the concepts involved in the specific domain. Ontology is used to capture the knowledge by using controlled vocabulary of words from the corresponding domain. A controlled vocabulary is a set of restricted words, used for describing resources or discovering data for any particular domain. When representing knowledge for a given domain, controlled vocabulary prevents misspellings and avoids the use of arbitrary, duplicate or confusing words that cause inconsistency. The major advantage of the use of ontology is that it will provide a globally unique identifier for the concepts. It helps to share common understanding of the structure of information among the users and to reuse and analyze the domain knowledge. It enables to merge already existing knowledge, thereby, expanding it further. Ontology for any domain is developed by formulating a set of questions that the envisioned knowledge-based agent should be able to answer.

While sharing the domain knowledge, ontology is used to make the knowledge interoperable and also reusable thereby having seamless exchange of information. Ontology is developed to provide the common semantics for agent communication so that it acts as a bridge when two or more agents need to communicate or exchange information.

Motivation, Objectives and Scope of the Thesis

The aim of the work carried out in this thesis is to semantically search a nuclear reactor domain by offering greater functionality and interoperability for automatically extract the knowledge by the machine. The objective is to find out an optimised methodology for semantic heterogeneity problems. Ontology matching is solution for semantic heterogeneity. Ontology versioning, duplication, inconsistency between data providers and data users are the root causes for hindering the functioning of ontology matching. Research contributions of the thesis are: QME algorithm, matrix rank based ontology matching algorithm, Pareto optimisation and development of knowledge management portal. Development of Quick Mapping Evaluator tool for ontology mapping is used to avoid duplication in the ontology matching. An algorithm for string based ontology mapping, by using matrix rank based technique is also demonstrated. Using this matrix ranking algorithm, partial overlap (or) duplicate (or) unique ontology is determined. The result of the ranking algorithm decides whether to eliminate or reuse or share the knowledge. In addition, a Pareto based optimisation technique is carried out for the ontology matching, which is used to find the matching in large ontologies. The process of matching involves calculation like Kullback divergence, Cosine structure divergence, string equality measure, Levenshtein distance and Largest Common Subsequence. Based on the above analysis, the optimised solution for matching the ontologies is evaluated.

Unlike domains like medicine, education, information technology, ontology representation in nuclear domain is rather limited. Hence, as an initiative, Fast Breeder Test Reactor system is taken for knowledge representation. As a test bed for ontology matching, Knowledge management portal has been developed for FBTR and christened as KMNuR portal. In the portal RDF, OWL, UML and graph representation of knowledge are embedded. Based on a literature survey for finding the best suited representation ontology tool, for knowledge management system development, Protégé tool, is found to be suitable and employed in the present work.

Proposed Chapters of the Thesis

The work carried out as part of the thesis are organized and divided in to five chapters.

Chapter 1. Introduction: This chapter gives introduction about knowledge management system, need for knowledge management system, survey of nuclear related knowledge management portals, knowledge representations, semantic web technologies, ontology construction for nuclear reactor domain.

Chapter 2. Semantic Web based Knowledge Representation Schemes and Tools: In this chapter, literature survey on ontology application in various domains and survey on ontology editor tools are presented.

Chapter 3. KMNuR: Semantic Web Based Knowledge Management Portal for Nuclear Reactor Domain: In this chapter the need for knowledge management portal, a brief introduction about Fast Breeder Test Reactor, system architecture of KMNuR web portal, representing OWL format in KMNuR, representing OWL-GRAPH in KMNuR, representing OWL_UML in KMNuR, knowledge representation for nuclear power plant, its subsystems and components are discussed. Challenges faced while developing ontology is also discussed.

Chapter 4. Development of Algorithms for Ontology Mapping: In this chapter**,** introduction about ontology mapping and string similarity measures, implementation of Quick Mapping Evaluator algorithm and algorithm for rank of a matrix for ontology matching are discussed. A Pareto based optimisation for ontology matching technique is also explored.

Chapter 5. Conclusions and Scope for Future Work: The work carried out in the thesis is summarized and scope of work to be carried out in future is discussed.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The requisite for knowledge management and knowledge representation are presented in this introductory chapter. Semantic web based knowledge representation portals and machine languages used for implementing the same are explained. The process of capturing the knowledge about a particular domain using ontology and its editor tools are briefly described. Types of nuclear reactor and review of available knowledge management portals for nuclear reactor are summarised. Knowledge management portal for Fast Breeder Test Reactor, Kalpakkam, various algorithms like Quick mapping evaluator, matrix rank based ontology matching algorithm and a Pareto optimisation are also described.

1.1 Introduction

Nuclear energy generated from nuclear reactors plays a significant role in producing electricity which in turn, meets development goals of the country. For constructing future reactors and upgrading the existing plant, knowledge preservation is essential. Preservation of the knowledge pertaining to nuclear reactor is a vital task for safe operation of the plant. Building a knowledge-based system for nuclear power plant is complex, as it requires collection of knowledge from multiple sources and must be integrated to share. Knowledge derived from different fields has its own distinctive terminology.

Knowledge sharing is a systematic process for creating, acquiring, synthesizing, receiving, contributing and utilizing available experience and expertise to achieve the organizational goal. Sharing of knowledge would be helpful in designing a sound knowledge management system for any organization. Knowledge management is the process through which an organization generates value from their intellectual and knowledge based assets. The process of development of knowledge management system involves developing tools, which would capture the knowledge for machine processing. There are different ways to represent the knowledge that machine will understand and process.

The ability to share and re-use knowledge is exactly what is needed for large scale internet applications. The ontology is used to store relations that are connecting the knowledge and hence they are useful for semantic web applications. Domain areas can often be organized in a tree structure composed of a super-concept and relevant sub-concepts. The tree structure for the knowledge of ontology-based search engine facilitates the search of web knowledge simply by adding structure to the largely unstructured web.

1.2 Knowledge Management

Knowledge management is the process through which organizations generate value from their intellectual and knowledge based assets. Capturing the knowledge available in each domain and preserving it for the future is one of the prime objectives of knowledge management. According to Oracle, knowledge management system helps to reduce research time, increase resolution accuracy, reduce training time, increasing service volumes, creation of service insight etc [1].

The system that is developed to capture the knowledge, involves identifying, creating, representing and finally distributing. Provision to share and re-use of knowledge is what is needed for any knowledge based system.

It is a challenging task for any organization to achieve the same knowledge management that is being adopted in many domains [2 - 4]. The effort is to employ a suitable mechanism for obtaining and integrating all the expertise, experience and knowledge available within the organization.

It facilitates preserving of both tacit and explicit knowledge and converting the available data into a reusable knowledge repository for the future use. Explicit knowledge includes reports, which can be easily described and coded for storage and dissemination. The tacit knowledge on the other hand is highly personal, context specific, and can be manifested only through sharing, applying and practicing and cannot be easily coded. Obtaining knowledge from both tacit and explicit is vital. Knowledge Management tools are used to capture and transfer knowledge from the experienced members to the new workforce [5]. Knowledge management for a specific domain is implemented by adopting different technologies like knowledge discovery, knowledge sharing, knowledge representation, knowledge engineering, knowledge refinement, and knowledge acquisition [6].

1.3 Knowledge Representation

Information integration has always been a challenging topic and one can expect this challenge to continuously trigger research and development. As one important step, physical exchange of data is not an issue any longer; being part of the global internet is a commodity for most enterprises. Knowledge representation is a branch of artificial intelligence aimed at modelling what a domain is about by creating schemas that can access and reason about information with a set of interference rules. Research in the field has developed a number of knowledge representation languages, each with its own set of features and tradeoffs. These languages differ in the way that knowledge is acquired, the extent of the descriptions they provide, and the type of inferences that they make [7].

A knowledge representation is used to enable an entity to determine consequences by thinking rather than acting, that is, by reasoning about the world rather than taking action in it [8]. Knowledge representation aims at designing computer systems that reason about a machine-interpretable representation of a domain, similar to human reasoning. The domain of interest can be any part of the real world or any hypothetical system about which one desires to represent knowledge for computational purposes. Knowledge representation offers the methodology for capturing expert knowledge, and it helps in solving the problem of improving document retrieval [9]. It is the process of expressing the analyzed knowledge in an understandable and usable form without ambiguity for enhancing communication between the expert and the knowledge engineer [10]. A knowledge-based system maintains a knowledge base which stores the symbols of the computational model in a form of statements about the domain, and it performs *reasoning* by manipulating these symbols. Applications can base their decisions on domain-relevant questions posed to a knowledge base. The effective knowledge representation is achieved by overcoming the lack of precision [11]. In the present era it can be done by enabling the machine to process the available knowledge. As it is not an easy task to teach machines to comprehend natural languages, the way the information is provided to the machine has to be analysed. To expose the represented knowledge to the outside world and to the community users, knowledge portals are used and details about it are presented in the following section.

1.4 Requirements for Knowledge Portals

The aim of knowledge portals is to make knowledge accessible to users and allow them to exchange the knowledge. The portal should provide different spaces of knowledge sharing, synchronous and asynchronous communications media, document storage and retrieval etc**.** The portals provide views of domain specific information on the World Wide Web, thus helping their users find relevant, domain-specific information. They also specialize in a certain topic to offer deep coverage of the domain of interest and address a community of users. [12]. The portals serve a knowledge management function by "dealing with information glut in an organized fashion" [13]. It acts as a gateway for content manipulation of the web so that the existing information can be updated, or new information can be added and by deleting old and out dated information. Thus it acts as intermediaries for knowledge access and knowledge sharing on the World Wide Web.

1.4.1 World Wide Web

The **W**orld **W**ide **W**eb (WWW) is a system of interlinked, hypertext documents accessed via the internet and serves as a powerful resource for application to application communication. The WWW is recognized for its function of facilitating information sharing by providing a man-machine interacting mechanism enabling users to go online to search for valuable facts or knowledge. The current web is written in natural language and it is a web of linked pages, instead of a web of facts. It is human understandable and machine processable and not machine understandable. Users have to look for possible keywords so that search engines can obtain the desired information. The returned data is very voluminous which motivated an approach to turn the web of pages into a web of knowledge, so that web users can query the information of interest directly. More specifically, the approach superimposes a web of knowledge over the web of pages, allowing users to query the information. The web of knowledge can be realized by transforming web pages into populated conceptual models. This is realisable by semantic web, which is an extension of current World Wide Web where machines understand the concepts and process the knowledge [14].

1.5 Semantic Web

A web service is a software system designed to support machine to machine interaction over the internet. The web of documents is morphed into a web of data. The semantic web embraces three stages of internet growth. The first stage, Web 1.0, is about connecting information and getting on the net. Web 2.0 is about connecting people putting the “I” in user interface and the “we” into a web of social participation. The next stage, Web 3.0, is about representing meanings, connecting knowledge and putting into work, thereby making internet more relevant and useful. The Web 3.0 would be something similar to a “read-write-execute” web, where data isn’t owned but instead shared. With the expectation for knowledge share, the next generation of web would be semantic web, which is designed for machine understanding.

In semantic web, the information is given well-defined meaning for enabling computers and people to work in cooperation. Semantic is a step towards web intelligence [15]. Classified hierarchies and advanced content based searching in semantic frameworks will increase knowledge management capabilities. Information structuring and access methods which are crucial in semantic web will eventually enable semantic indexing and retrieval. One of the rules about the semantic web is that “anyone can say anything about any topic” [16]. Ontologies are needed for the semantic web to function, because ontologies bring structure to information, such as attributes (descriptive data), classes, sub classes, and relations among entities. Inference rules in ontologies allow knowledge reasoning and help to make automatic processing. It provides a structured vocabulary that describes concepts and relationships between them as well as a specification of the meaning of terms used in the vocabulary.

For the semantic web to be effective, ontologies have to be comprehensive. In Berners-Lee's concept, they would exist in the form of *metadata* which are information included in the code for web pages that is invisible to humans, but readable by machines [14]. The software agents could utilize metadata, ontologies, and logic to carry out its tasks. Hyperlinks connect the pages and documents in the current web which enables the user to reach the desired information easily. However, from a machine or a software agent perspective, a web page is nothing but pure **H**yper **T**ext **M**arkup **L**anguage (HTML) code, which does not give any clue about the *meaning* of the content. Thus, the automatic agents cannot browse the web and collect information as easily. Semantic web is proposed to overcome these difficulties. It gives meaning to documents and the entities in that document. Entities can be uniquely identified and have their own set of properties. Two entities can be related to each other via these properties allowing data integration, data reuse and automation.

1.5.1 Semantic Web based Search and Search Engines

Semantic search refers to the process where in the semantic annotation algorithm extracts concepts or instances from domain ontology, annotates document pool for domain resource repository, and generates semantic index repository. A search program takes the input from the user and performs search task from semantic index repository and return the search results with semantic characteristic [17]. Semantic web allows machine to intelligently search, combine and process the web content, based on the meaning of content in the domain [18]. The Semantic web utilizes conceptual relations between various resources denoting real world objects, people, places and events [19].

Semantic search engines are expected to handle the concept based query and deliver the quality search results [20]. Thereby enabling the applications to search and exploit their index repositories directly. There are five main categories of semantic search engine, semantic search methods, hybrid semantic search engine, semantic XML (extensible markup language) search engine, semantic ontology search engines and semantic multimedia search engine [21]. In general, the process of semantic search engine is as listed below:

a) User question is interpreted, extracting the relevant concepts from the sentence

b) Using the set of concepts ontology is queried

c) The final results are presented to the user [22].

Simply searching for keywords is not enough anymore due to incorrect and incomplete keyword declarations and ambiguous words that exist in our natural languages. Some best-known search engines are Google, Yahoo, and Bing etc. Each search engine has unique characteristics and uses different algorithms to produce the information expected by the user. Search engines rely on syntactic means for content matching with user queries. Keyword matching is based on direct comparison of query keywords and the words that appear directly in the web documents. While searching, the return results are not very accurate due to ambiguity among natural language words and properties like synonymy or polysemy and syntactic methods. When a user types keywords or phrases into the search engine, it will simply return all the pages containing the given keywords or phrases.

The same term can have different meanings and different terms can mean exactly the same thing [23]. The main reason for this overloading problem is that a search engine simply returns all websites containing the string, a user has searched for, even though the resultant website listed may not be relevant. Moreover the display order of search results is irrelevant and confusing. One of the biggest problems of current search engine is user being flooded with large volume of search results. In addition to this ambiguity, nature of words also affects the search results. In this context a semantic search for searching is recommended.

Since the semantic search engine is aware of the meaning of the search query, the user can be offered relevant results related to the context. For instance, by showing related information to the search query of the user, it becomes very easy to guide a user to find the information user is looking for. The key to implementation of semantic web is structured data which is machine readable. The final goal of semantic web is to structure the meaningful contents of unstructured data thereby enabling more sophisticated knowledge modelled management systems.

Irrelevant results from search engines are due to ‘semantic gap’ existing between the meanings of terms used by the user and those documented by the search engines [24]. In semantic web, each of the websites is annotated with ontology. Hence the whole web consists of agglomerations of domain-specific ontology. In semantic web, the user query is analyzed at knowledge level and will be answered by performing logical inferencing using ontology.

In the present work semantic web based knowledge management is taken for implementation. To produce domain specific ontology, semantic web requires ontological languages such as **W**eb **O**ntology **L**anguage (OWL), **R**esource Description **F**ramework (RDF) so that information can be reasoned by the machines and make new conclusion and not just match the keywords.

1.5.2 Semantic Layers and Languages

In semantic networks each concept is represented by a node in a graph, and semantically related concepts are connected by arcs, implying that each concept is connected to other concepts [25]. Information representing and reasoning and formal structural of data are the things to be taken care of. In modern database, knowledge representation focus is on management and access of stored complex data through integrated logic-based modelling languages and reasoning capabilities. Similarities and differences among pieces of data collected from the different domains have to be sorted out in order to create integrated and more meaningful information in the semantic level [26]. Semantic integration helps in data interpretation and data understanding without human involvement, and therefore achieves more accurate and efficient data integration. Semantics web is to provide a common framework that allows data and knowledge to be shared and reused across applications, enterprises and communities by making the web documents' meaning explicit [27]. It aims at developing methodologies for representing large amount of knowledge in web accessible form. Challenges in automatic extraction are the unstructured nature of texts and large volume of texts rendering full-edged natural language processing methods infeasible. Most existing knowledge representations are still not compatible with each other and hence research on automated development of ontologies from texts has become increasingly important.

Semantic web components can be described as being similar to layers of a cake, with the upper layers, representing semantic languages, serve to add meaning and functionality to the lower, foundational layers [28]. The lower layers create the ontology framework while the upper layers work to infer information from the ontology. Description of the various meta languages such as XML, RDF, and OWL etc developed for encoding ontology of the domain are given in the next section.

1.5.2.1 URL and XML

One of the key aspects of making an ontology modular is to create a reference set of concepts for a resource by means of **U**niform **R**esource **I**dentifiers (URIs), **U**niform **R**esource **L**ocator (URLs) such that they are globally available and unique. URI identifies resource items like human beings, corporations and bound books in a library, just not "network retrievable" ones. A URI for a book, for example, would be the ISBN number. The URL relates a resource (a text document) to a unique identifier. In generic terms the web is information space and URIs are the points in that space. More exactly, URIs is the naming/addressing technology. URIs are short strings that make resources available under a variety of naming schemes and access methods such as HTTP, FTP, and Internet mail addressable in the same simple way.

Structured data are often considered similar to template-based format such as shape files, Digital Elevation Model data, etc. There is another type of structured data, one that makes explicit the relationship between the declarative meaning of the data and the data itself, rather than relying on an implicit relationship. This data format is called the e**X**tensible **M**arkup **L**anguage (XML) which is a simple and universal meta-language derived from **S**tandard **G**eneralized **M**arkup **L**anguage (SGML) and builds upon HTML. XML based metadata can be understood by humans and hence it allows users to create their own customized tags, enabling the definition, transmission, validation, and interpretation of data between applications. Scripts, or programs, can make use of these tags, but the script writer has to know what the page writer uses each tag for. To make XML files more amenable to automated processing by web browsers, XML schemas were created to impose a standard set of terminology for the types of information. Even in XML schemas the interpretation of the meaning of this data is still left to the external program that reads the data. Hence, the schema approach also limits what can be in the file in the interest of imposing the advantages of a structure on the concept-based knowledge. Hence both XML and XML-schema do not provide a mechanism to deal with the semantics (the meaning) of data. Meaning is expressed by RDF [29].

1.5.2.2 Resource Description Framework

The Resource Description Framework (RDF) is the third component of the semantic web, which uses a simple relational model that allows structured and semi structured data to be mixed, exported, and shared across different applications. While XML provides interoperability within one application using a given schema, RDF provides interoperability across applications. RDF is intended for representing metadata about web resources, such as the title, author, and modification date of a web page, copyright and licensing information about a web document, sitemap, keywords that search engines look for and the web page's intellectual property rights etc. Developed under the guidance of the World Wide Web Consortium, RDF is designed to allow developers to build search engines that relay on the metadata and to allow internet users to share web site information more readily. The goal of RDF is to make work easier for autonomous agents and automated services by supplying a rudimentary semantic capability. The basic unit in RDF is an RDF statement which consists of triples in the form of subject, predicate, object. The subject denotes a resource, predicate denotes aspects of that resource and expresses a relationship between the subject and the object which can be a resource identified by a URI. Each RDF statement can therefore be modeled as a graph with two vertices (subject, object) connected by a directed relation arc (predicate). Even though each statement within the RDF body expresses its own unique meaning, any part of the entity can be connected to other entities.

The usage of URI in RDF model breaks the constraints that documents and statements have to be physically localized to be aggregated. Moreover, being a logical graph model with specified logical relations and constraints, the machine-readable RDF can be managed and analyzed using query and inference tools in an automatic manner [30].

1.5.2.3 RDF Schema

RDF being a simple data model does not have significant semantics. In order to overcome this **RDF** **S**chema (RDFS) standard provides additional modelling primitives to define classes, subclass relationships between classes, properties, sub property relationships between properties, and restrictions on property domains and ranges etc. Eventhough, RDFS provides simple functions to build vocabularies for RDF statements thereby associating metadata to each other, they did not prove sufficient to handle real world modelling needs. For instance, they cannot model class disjointness and intersection relationships, property symmetry, cardinality, etc. This is one of the reasons for the development of more expressive languages - Web Ontology Language (OWL) [31].

1.5.2.4 Web Ontology Language

The main requirements of an ontology language for machine-processing are a well-defined syntax, a formal semantics, an efficient reasoning support system and sufficient expressive power. The OWL is designed to formally define the terminology used on web, and thus to facilitate machine interpretability. OWL is developed based on the ontology languages, **S**imple **H**TML **O**ntology **E**xtensions (SHOE), **D**ARPA **A**gent **M**arkup **L**anguage (DAML) and **O**ntology **I**nference **L**ayer (OIL). These web ontology languages were designed for web page authors to annotate their web pages with formal knowledge representation semantics. OWL adds to RDF and RDF-S an additional set of specified predicates, objects, and their properties that OWL interpreters will know how to use and to make additional inferences, such as logic properties and restrictions and has the same XML based syntax [32]. The OWL is partially mapped on description logic which is a subset of predicate logic for which efficient reasoning support is possible.

OWL is applied to capture the domain knowledge using structured vocabulary such that it can be used by both humans and machines. More importantly, OWL defines and relates concepts to one another through the relationships which may have constrains placed on them. OWL provides a mechanism for creating components for ontology, such as classes, instances, properties, and axioms. Classes can have super or sub classes. The property component defines the relationships between instances or instances to data type values. Axioms specify the information about classes and properties, such as the relationship of two classes or a range of a property. By defining the relationships and adding constraints, OWL provides a mechanism for reasoning and inheritance of properties, which is the key to semantic web to allow machine interoperability.

OWL Full: OWL Full uses all the primitives and supports maximum expressiveness and fully compatible with RDF and RDFS syntax and semantics, but has no computational guarantees. In OWL Full, a class can be treated simultaneously as a collection of individuals and as an individual in its own right. The disadvantage of OWL Full is that it is impossible to perform automated reasoning and therefore cannot provide complete or efficient reasoning support.

OWL DL: **O**ntology **W**eb **L**anguage **D**escriptive **L**ogic (OWL DL) is a sublanguage of OWL Full that restricts how the constructors from OWL and RDF can be used. Description logics are a decidable fragment of first order logic, and are therefore amenable to automated reasoning and supports strong expressiveness while retaining computational completeness and decidability. It is possible to automatically compute the classification hierarchy and check for inconsistencies in an ontology.

OWL Lite: OWL Lite excludes enumerated classes, disjointness statements, and arbitrary cardinality and supports classification hierarchy and simple constraints. It is simpler for tool support and provides a quick migration path for taxonomies. The choice between OWL Lite and OWL DL depends on the extent of requirements for expressive constructs. To use ontology, a reasoner and query engine is needed. A popular query engine for RDF and RDF-S is called SPARQL [33].

1.5.2.5 SPARQL

The **S**imple **P**rotocol **A**nd **R**DF **Q**uery **L**anguage (SPARQL) is a SQL-like language for querying RDF data. In SPARQL a query is written as a sequence of triple patterns, conjunctions, disjunctions, and optional relations composing as a **B**asic **G**raph **P**attern (BGP). A triple pattern is like an RDF triple, but with the option of a variable in place of RDF terms (i.e., URIs, URLs, literals or blank nodes) in the subject, predicate or object positions. BGP allows applications to make queries where the entire query pattern must match for there to be a solution. If the optional part does not match, it creates no bindings but does not eliminate the solution. In order to realise semantic web ontology is a pre-requisite for knowledge representation which is elaborated in the subsequent section.

1.6 Ontology-based Knowledge Representation

The word “Ontology” is originated from Greek philosophers Aristotle and Plato, and is used to describe the existence of being or “the study of the state of being”. Onto means “being” and logos means “treatise”. Gruber defines ontology as “the formal explicit specification of conceptualizations, used to help programs and humans share knowledge” [34]. A conceptualization is an abstract, simplified view of the world that user wish to represent. Ontologies revolve around properties and relationships that are associated with a group of objects or concepts. Explicit means that the type of concepts used and the constraints on their use are explicitly defined. Formal means the machine should be able to interpret the information provided unambiguously. Shared reflects the idea that ontology should capture consensual knowledge accepted by the communities.

The following seven distinct definitions are collected and analyzed in [35-37].

i. Ontology as Philosophical discipline.
ii. Ontology as in informal conceptual system.
iii. Ontology as a formal semantic account.
iv. Ontology as a specification of conceptualization system.
v. Ontology as a representation of a conceptual system characterized by specific formal properties and only by its specific purposes.
vi. Ontology as the vocabulary used by logical theory.
vii. Ontology as a specification of logical theory

The heart of all semantic web based knowledge representation is the use of ontology. Ontology is hierarchically structured and is used to capture the knowledge by using controlled vocabulary of words from the corresponding domain. By using such structures, a knowledge base is built [38]. It helps to share common understanding of the structure of information among the users and to reuse and analyze the domain knowledge. A controlled vocabulary is a set of restricted words, used for describing resources or discovering data for any particular domain [39]. When representing knowledge for a given domain, controlled vocabulary prevents misspellings and avoids the use of arbitrary, duplicate or confusing words that cause inconsistency. The major advantage of the use of ontology is that it will provide a globally unique identifier for the concepts. It enables to merge already existing knowledge, thereby, expanding it further. Using ontology the knowledge created is shared and re-used for domain applications [40-41]. Ontology can also be thought of as machine interpretable dictionaries since they provide formal definitions for domain concepts [42]. Ontology for any domain is developed by formulating a set of questions that the envisioned knowledge-based agent should be able to answer [43].

Ontology can be defined as

$O = (C; T; R; A; I; V; \leq C; \leq T; \sigma R; \sigma A; iT; iR; iA)$.

It consists of disjoint sets of concepts (C), types (T), relations (R), attributes (A), instances (I), and values (V). The partial orders $\leq C$ (on C) and $\leq T$ (on T) define a concept hierarchy and a type hierarchy respectively. The function σR: $R \rightarrow C^2$ provides relation signatures (i.e., for each relation, the function specifies which concepts may be linked by this relation), while σA: $A \rightarrow C \times T$ provides attribute signatures. Finally, there are functions iC: $C2^I$ the assignment of instances to concepts, iT: $T2^V$, the assignment of values to types, iR: $R \rightarrow 2^{I \times I,}$ which instances are related by a particular relation, and iA: $A \rightarrow 2^{I \times V}$. This is the value of each attribute for each instance [44].

There are multiple layered concepts in the knowledge system. The hierarchy of the general concepts is more significant. In addition, the ontology can be indexed to further facilitate efficient searches.

Ontology is created from scratch by extracting information from domain experts and merging already existing ontology into the new ontology. For annotating or tagging content, syntactic and semantic metadata can be used. Ontology is developed to provide the common semantics for agent communication so that it acts as a bridge when two or more agents need to communicate or exchange information [45].

In order to develop ontology the steps are followed

- Domain and scope of the application should be defined
- Important terms in that concepts should be identified
- Classes and class hierarchy for the concepts should be defined
- Properties of classes and constraints should be defined
- Instances of classes should be created [46].

The simplest solution is not to teach computers to behave as humans to "understand" natural languages, but to change the way information is presented to the user [47]. The ultimate vision for a semantic web is to create knowledge representation so that machine can understand and navigate [48]. This includes definitions and an indication of how concepts are inter-related and collectively impose a structure on the domain and constrain the possible interpretations of terms [49]. Ontologies specify the set of physical and/or conceptual characteristics of resources that have been deemed relevant for a particular community of users [50]. Some of the reasons to create ontology are to share common understanding of the structure of knowledge, to enable reuse of domain knowledge, to make domain assumptions explicit, to separate domain knowledge from operational knowledge [51].

However, the construction of ontology is a time-consuming process requiring expert involvement from both ontology engineering and domain of interest. With the continuous evolvement of knowledge and application needs, ontologies also need to change accordingly which impose a key challenge on ontology maintenance.

Ontology development [52] enables:

- Sharing common understanding of knowledge among people and machine
- Reusing of the domain and expert knowledge
- Making the domain assumptions explicit
- Increasing interoperability among various domains
- Increasing the scalability

Ontologies have often been constructed in one of the following two ways. Building ontologies by domain experts has two disadvantages: 1) it is time consuming to construct ontologies by hand, and 2) view point of the expert, and are subjective, and limited to some degree by the available expertise. The other approach to building ontologies is using text mining techniques on numerous web pages. In this approach, extracted concepts are prone to errors and relationships among concepts are hard to be extracted. As of now, ontologies for various domains are developed manually. Some of the issues involved in the design and development of ontologies are the requirement of expert knowledge of the domain, extensive group discussions in understanding the view point of the domain, and incremental modifications to the ontology.

Ontology should be sharable, adaptable, reusable and interoperable. Interoperability among the machines can be achieved by translating between different methods, languages and software tools. Ontology is used as metadata for searching and serving as an index into a repository of information. The main role of ontology is ambiguity reduction. There are several classes of ontologies like upper ontology, core ontology and domain ontology details of which are presented below.

1.7 Types of Ontology

Artificial intelligence classifies ontology by using different criteria. One classification is based on conceptualization 1) amount and type of structure of the conceptualization and 2) subject of conceptualization [53]. According to Natalya Fridman Noy and Carole D. Hafner, ontology is created for general or for specific domain. Ontology is classified as generic, core, domain, task and application [54].

Ontology is also classified as formal and informal ontology. A formal ontology is specified by a collection of names for concept and relation types organized in a partial ordering by the type subtype relation. Informal ontology is a catalogue of types that are either undefined or defined only by statements in a natural language [55]. Based on concepts, ontology is classified as upper ontology, core ontology and domain ontology.

Upper ontology, top-level ontology or foundation ontology, defines general base concepts that are similar across all domains and supports ontology development and facilitates common-sense, human-like understanding and reasoning [56].

Core ontology should be more than the upper level of a terminology. It is a basic and minimal ontology, consisting only of the minimal concepts required to understand other concepts. The goal of core ontology is to provide a global and extensible model into which data originating from distinct sources can be mapped and integrated [57].

Domain ontology defines the terminology and concepts relevant to a particular topic or area of interest. As systems that rely on domain ontologies expand, there is often a need to merge domain ontologies into a more general representation. Different ontologies in the same domain can also arise due to different perceptions of the domain, based on cultural background, education, ideology or because a different representation language is chosen.

Availability of editor tools is a pre-requisite in order to define ontology for any domain. In the following section various tools existing in the literature are briefly discussed.

1.8 Ontology Editor Tools

Some of the available tools for creating ontology are: Protégé, Model Futures OWL Editor, TopBraid Suite, OntoLingua, OntoEdit, WebODE, KAON, ICOM, DOE, WebOnto, Medius Visual Ontology Modeler, LinKFactory Workbench, K-Infinity and OntoStudio [58]. Many of these are available as freeware tools. Based on the result of analysis which will be discussed in Chapter 2, Protégé IDE is used for ontology development.

1.8.1 Protégé

The Protégé system is an environment for knowledge-based systems development which began as a small application designed for a medical domain (protocol based therapy planning), but has evolved into a much more general-purpose set of tools. More recently, Protégé has developed a world-wide community of users, who themselves are adding to Protégé's capabilities, and directing its further evolution. The original version of the Protégé software was an application that took advantage of structured information to simplify the knowledge-acquisition process. Through four distinct releases, the Knowledge Modeling Group at Stanford Medical Informatics, Stanford University, has worked to turn Protégé into a general-purpose environment for knowledge modeling. Protégé-2000 is far more general than the original version, yet it maintains the original focus on the use of meta-knowledge to create usable knowledge-acquisition tools [59]. It is a Java based open source stand alone application. It enables the users to load and save OWL and RDF ontologies, edit and visualize classes, properties and SWRL rules. Reasoner like RACER, FaCT++, Pellet or KAON2 are supported using application program interface.

1.8.2 TopBraid

TopBraid composer is used for creation and maintenance of ontologies. It is built using Eclispse platform and uses Jena as API (application program interface). It supports Pellet reasoner. Apart from the creation of OWL and RDF files, it supports import of databases, XML-Schemas, UML, spreadsheets. It also supports SWRL and Jena rules [60].

1.8.3 OntoLingua

Ontolingua provides a distributed collaborative environment to browse, create, edit, modify, and use ontologies. This tool is an Ontology library and server, which supports a WWW interface and translation into various formats. It provides a suite of ontology authoring tools and a library of modular, reusable ontologies. The environment is available as a World Wide Web service and has a substantial user community. The tools in Ontolingua are oriented toward the authoring of ontologies by assembling and extending ontologies obtained from the library [61].

1.8.4 SWOOP

SWOOP is an open-source hypermedia-based OWL ontology editor. It is designed as a native OWL editor. It uses Pellet reasoner.

1.8.5 OntoStudio

It is a commercial product of Ontoprise. It supports direct creation of rules. It uses Eclipse platform for development with all advantages of using plugin cocept. OntoStudio is coupled with F-Logic. It offers a graphical and a textual rule editor as well as debugging features and also support form based query editor [62].

1.8.6 OntoEdit

It is part of OntoStudio, based on IBM Eclipse framework [63]. OntoEdit is a development environment for ontology design and maintenance. It supports multilingual development, and the knowledge model is related to frame-based languages. OntoEdit is based on an open plug-in structure. Every plug-in provides other features to deal with the requirements an ontology engineer has. Data about classes, properties and individuals may be imported or exported via different formats, such as OXML, F-Logic, RDF/RDFS, OWL [64].

1.8.7 OilEd

It is an ontology editor allowing the user to build ontologies using DAML+OIL, the language that inspire the actual OWL standard [65]. The current versions of OilEd do not offer a full ontology development environment, but provides enough functionality to allow users to build ontologies and to demonstrate how we can use the FaCT reasoner to check those ontologies for consistency. Data can be imported from DAML+OIL, OWL RDF/XML and OIL text formats. OilEd can save ontologies as DAML+OIL documents only. OilEd is available as an open-source Java project under the GPL license.

1.8.8 Model Futures OWL Editor

Model Futures have developed a free OWL Editor Tool [66]. The editor is tree-based and has a “navigator” tool for traversing property and class-instance relationships. It can import XMI (the interchange format for UML) and Thesaurus Descriptor and EXPRESS XML files. The software runs on Windows.

1.8.9 KAON 2

KAON2 is an infrastructure for managing OWL-DL, SWRL and F-Logic ontologies. It was produced by the joint effort of the following institutions: Information Process Engineering at the Research Center for Information Technologies, Institute of Applied Informatics and Formal Description Methods at the University of Karlsruhe, Information Management Group at the University of Manchester. KAON2 is a successor to the KAON project (often referred to as KAON1). The main difference in KAON1 is the supported ontology language: KAON1 used a proprietary extension of RDFS, whereas KAON2 is based on OWL-DL and F-Logic, reasoners, such as FaCT, FaCT++, RACER, DLP or Pellet [67]. A module for extracting ontology instances from relational databases. KAON2 has been fully implemented in Java. Queries can be formulated using SPARQL.

1.8.10 Integrated Collaboration Object Model (ICOM)

ICOM provides a simple, freeware conceptual modelling tool that demonstrates the use of, and stimulates interest in, the novel and powerful knowledge representation based technologies for database and ontology design. ICOM can express: the standard Entity-Relationship data model, enriched with “IsA” links, disjoint and covering constraints, full cardinality constraints, and definitions attached to entities and relations by means of view expressions over other entities and relationships in the schema. ICOM tool is written in standard Java, and it is being used on Linux and Windows machines. ICOM communicates via a CORBA protocol with the FaCT description logic server [68].

1.8.11 Differential Ontology Editor (DOE)

DOE is a simple ontology editor which allows the user to build ontologies according to the methodology. It supports RDFS, OWL, DAML+OIL, OIL and CGXML (a language to specify conceptual graphs). DOE is not intended as a full ontology development environment. It will not actively support many activities that are involved traditionally in ontology construction, such as advanced formal specification dealt with by tools like Protégé 2000. It is rather a complement of other editors, offering linguistics-inspired techniques which attach lexical definition to the concepts and relations used, and justify their hierarchies from a theoretical, human-understandable point of view [69].

1.8.12 WebOnto

WebOnto is an ontology library system developed by the Knowledge Media Institute of the Open University (UK) [70]. It is designed to support the collaborative creating, browsing and editing of ontologies. It provides a direct manipulation interface displaying ontological expressions and also an ontology discussion tool called Tadzebao, which could support both asynchronous and synchronous discussions on ontologies. It provides a web-based visualisation, browsing and editing support for developing and maintaining ontologies and knowledge models specified in OCML. WebOnto is a Java applet coupled with a customised web server which allows users to browse and edit knowledge models over the web.

1.8.13 LinkFactory

LinkFactory is the formal ontology management system, developed by L&C, used to build and manage the medical linguistic knowledge base LinKBase **[**71**]** Different views on the semantic network are implemented as Java beans. The LinKFactory® framework is implemented in Java code. LinKFactory is a tool that can be used to build large and complex language-independent formal ontology's. LinKFactory is also platform independent and works in Windows, Solaris, Unix and Linux.

1.8.14 Medius Visual Ontology Modeler

Medius Visual Ontology Modeler is a UML-based ontology modeling tool that enables component-based ontology development and management. Features include:

- A multi-user, network-based environment for ontology development using a graphical notation.
- A set of ontology authoring wizards that create and maintain the required UML model elements for the user.
- Automated export facilities in XML schema, RDF, DAML and other formats.
- Integration with a commercial, scalable object database with OKBC and CORBA-compliant access mechanisms.

The Visual Ontology Modeler implements Sandpiper's UML Profile for Knowledge Representation, which extends UML to enable modeling of frame-based knowledge representation concepts such as class, relation, function and individual frames, as well as the slots and facets that constrain those frames [72].

1.8.15 K-Infinity

K-Infinity is the tool developed by Intelligent Views, German company, for the creation, maintenance and the use of knowledge network. It is a knowledge editor based support for object-oriented knowledge modeling. RDF format is supported in this tool [73].

In the following section, implementation of the above mentioned editor tools for various application domains practiced worldwide is given.

1.9 Ontology Application in Domains

A survey of application of ontology in various domains is based on the available research papers, referred journals, reports in the respective domains, scholarly articles available on the advantages of ontology application and usage.

This survey has been carried out in domains like agriculture, education, medicine, defense and where the usage of ontology has been proved to be helpful. A broad picture of ontology applications in various domains practiced today are surveyed and described in Chapter 2.

Enhanced development of ontology would aid in the evolution of semantic web leading to complete sharing of knowledge in a given domain. It can also be inferred that the ontology development is a continuous process and success could be achieved by participation of the domain experts and users. As the development of ontology is limited in the field of nuclear energy, focus is given to create knowledge representation in the nuclear reactor domain. This is discussed in detail in the next section.

1.10 Nuclear Energy and Reactor

Nuclear energy which is an alternative to chemical energy has attracted researchers since the early 20^{th} century. Intense research and development during the last few decades has brought Nuclear power as the most important source for production of electricity. Fundamental reactor designs have progressed continuously so as to maximize efficiency and safety.

1.10.1 Nuclear Reactor

The principal part of nuclear reactor is the core which contains the nuclear fuel. The essential ingredient of a nuclear fuel is a fissionable material i.e. a substance that readily undergoes fission when struck by neutrons. The solid fuel material which are in shapes of plates, pellets, pins etc are usually put together and packed as sub-assemblies. A reactor core may contain from tens to hundreds of these fuels sub-assemblies held in fixed geometrical pattern. Moderators such as ordinary water, heavy water, graphite, beryllium etc are used in thermal reactor to slow down the neutrons as the fuel has high fission cross-section for low energy neutrons. In some reactors the fuel materials and moderator materials are intimately mixed together. The same material used for moderator can be used for the reflectors in the case of thermal reactors. The reflector reduces the leakage of neutrons by reflecting back the neutrons escaping from the core. In the fast reactors where fast neutrons are utilized for fission, nickel, molybdenum and stainless steel reflectors are used. In a reactor, in order to control the amount of heat produced coolants such as heavy and light water, gases, and liquid metals such as sodium, lithium, potassium etc., are generally used. The heat produced depends upon the number of fissions taking place per second in the reactor, which in turn depends upon the number of neutrons present in the reactor. The control system is designed to control the number of neutrons, thus control the rate of the chain reaction and power level. This system includes a number of devices, sensing elements that measure the number of neutrons in the reactor, control rods – containing strong neutron absorbers such as cadmium or boron, and other devices to regulate the position of the control rods. These neutron absorbing control rods when lowered into the reactor absorb the neutrons to reduce the neutron population and when raised allow the rise in number of neutrons. To protect the persons working near the reactor from harmful radiations like α, β and γ which is accompanied in fission reactions, the reactor is enclosed in steel and concrete.

1.10.2 Types of Reactors

Nuclear reactor may be classified as thermal or fast reactors according to the velocities of the neutrons which cause fission.

a. Thermal reactors: A reactor where the fission is mainly caused by the capture of thermal i.e., slow neutrons of energies up to 0.025 eV. To slow down the neutrons, some moderator is used.
b. Fast reactors: A reactor where fission is brought about by fast neutrons with energies more than 1000 eV. Moderator is not used in fast reactors, since the fission is caused by the fast neutrons with high energy.

Reactors are also classified according to secondary features such as the type of moderator, coolant, etc.

i. Graphite reactors and heavy-water reactors - they use graphite and heavy-water as moderators.

ii. Liquid metal cooled reactor –liquid metals like sodium, potassium etc., are used as coolants.

iii. Gas cooled reactor – CO_2, Helium etc., are used as coolants.

Reactor classification based on purpose:

(a) Power reactors – The reactors used for generating electrical power. Example Madras Atomic Power Station, Tarapur Power Station, Rajasthan Atomic Power Station etc.

(b) Research reactors – These reactors are mainly used for research purpose. Example **F**ast **B**reeder **T**est **R**eactor (FBTR), Kamini, Apsara etc.

India embarked on a three-stage nuclear power program (Figure 1.1), envisaged by Dr. Homi Bhabha. Stage one involves the deployment of natural uranium pressurized heavy water reactors. It will be followed by FBRs burning plutonium to breed U^{233} from thorium. The FBRs are to be followed, by Advanced Heavy Water Reactors capable of utilizing India's abundant thorium resources [74]. Nuclear energy is an inevitable source to meet the fast growing energy demands of India. India's present nuclear installed capacity is 4,780 MW, almost entirely based on PHWRs. With two imported LWRs, each of 1000 MW and one 500 MW PFBR to be commissioned shortly, the installed capacity would reach 7,280 MW. Department of Atomic Energy is also developing a 300 MW AHWR using U^{233} as fuel. Some of the typical reactors used in India for generation of electrical power are:

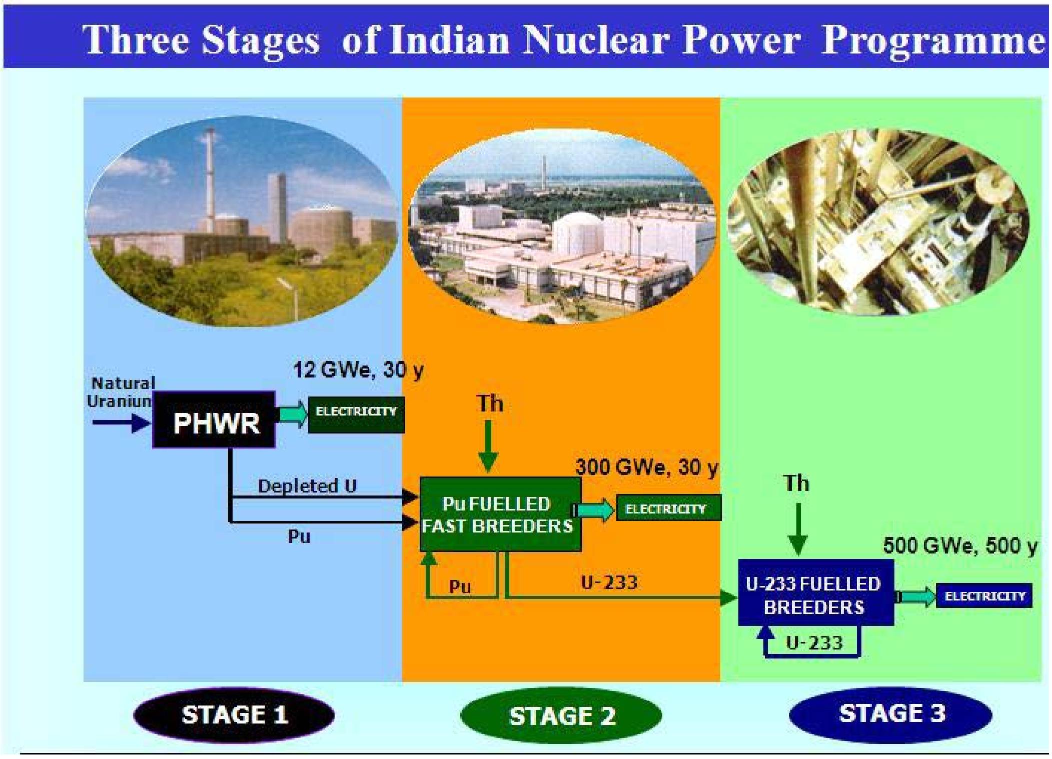


Figure 1.1 Three stages of nuclear power programme
Pu: Plutonium; U: Uranium; Th: Throium; GWe: Giga-Watt electrical; y: year;

1.10.2.1 Boiling Water Reactor (BWR)

In this thermal reactor the enriched uranium oxide is used as the fuel and light water is used as the coolant and moderator. The water is circulated by a pump and the water boils in the reactor core itself and the steam produced is fed directly to a turbine. The reactor pressure vessel is a strong concrete containment vessel to prevent hazard from the failure of the pressurised circuit. The exhaust steam from turbine is condensed and the condensate is sent back to the reactor core through a feed pump. Another pump is used for recirculating the coolant in the reactor vessel before converting to steam. In Tarapur Atomic power station two BWR’s are used for power generation of 210 MWe each [75].

1.10.2.2 Fast Breeder Test Reactor

Fast breeder reactors are energy systems which breed more fissionable material than they consume while producing power. The fast neutrons which cause the fission allows fast reactors to increase the energy yield from natural uranium by a factor of sixty to seventy compared to thermal reactors. A fast reactor does not use a moderator to slow down the neutrons produced during fission and the fuel used is fissioned directly by high energy or fast neutrons emitted during the fission process. The reactor uses a fairly high concentration of fissionable isotope, either U^{235} or Pu^{239}. However, Pu^{239} has a distinct advantage when used as fuel since the number of neutrons produced in Pu fission is high, making sufficient number of excess neutrons to produce more Pu from U used as a blanket surrounding the fuel core [76]. The excess of neutrons thus generated is the key parameter in the nuclear fission scenario, which is the measure of quality of the fissile element with respect to breeding, production of more fissile material for nuclear reactors. The fuel burn-up or depletion is offset to some extent by breeding (i.e. by converting fertile to fissile material). A reactor system in which this has been realized is called a 'fast breeder reactor'. This category of reactors can sustain energy production without any external feed of fissile material; but accumulate extra fissile material in the reactor, which can be used for feeding a new reactor after reprocessing. In advanced FBRs, it is possible to achieve high breeding ratios. DAE had started FBR programme as early as 1965 and a preliminary design of a 10 MWe experimental fast reactor was initiated through effective collaboration with France. FBR is also an enabling technology to make a transition to large-scale utilisation of thorium resources available in the country.

FBTR is a sodium cooled 40 MWt/13.2 MWe reactor with an unique plutonium rich carbide fuel (70% Pu and 30% U) (Figure 1.2). The high power generated in the compact core of the fast reactor necessitates the use of a liquid metal like sodium as the reactor coolant. Hence it is imperative to use efficient heat-transfer fluid as coolant, which should also possess favorable nuclear characteristics of low neutron moderation / absorption. Liquid metals, and among them liquid sodium meet almost all the requirements of a fast reactor coolant with its high thermal conductivity, reasonable specific heat, low neutron moderation, absorption and high boiling point, giving a large operating temperature range at near atmospheric pressure. It serves as a test bed for irradiation of fuel materials and provides experience in large scale sodium handling and reactor operation. It has two primary and secondary loops and a common steam water circuit with once through steam generator supplying super heated steam to the condensing turbine.

There are two steam generators per loop and are located in the common casing. The heat transportation circuit has been divided in to two loops so that in case of non availability of one loop, the other loop is available for removing the decay heat from the core. Heat generated by the reactor is removed by these two primary sodium loops, and transferred to corresponding secondary sodium loops through intermediate heat exchangers.

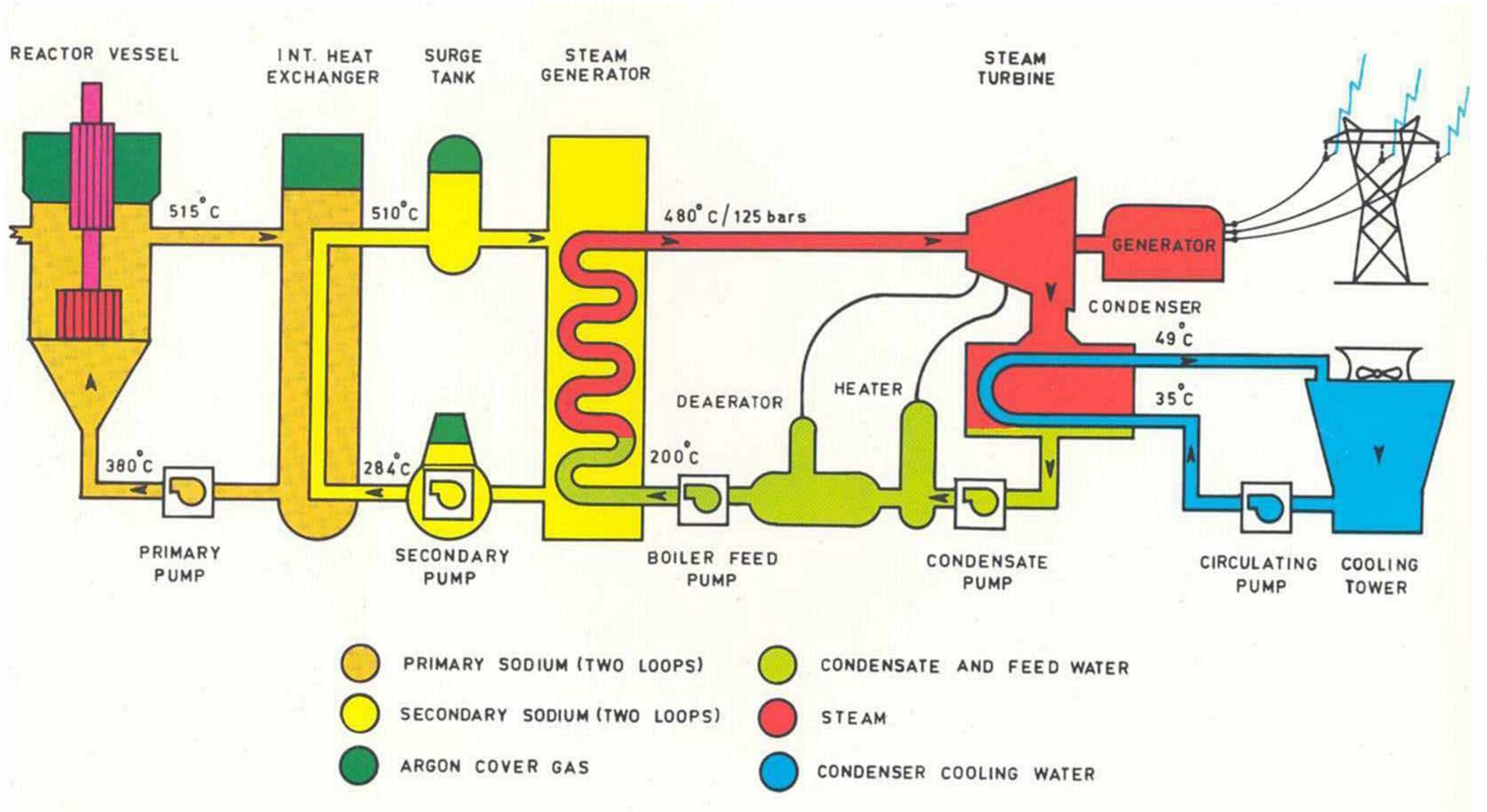


Figure 1.2 Schematic of fast breeder test reactor

1.10.2.3 World Wide Fast Reactors

Fast **F**lux **T**est **F**acility (FFTF) in USA, the small size prototype fast reactor in the United Kingdom, the prototype Phénix in France, the BN-350 in Kazakhstan, the BN-600 in Russia, Monju in Japan, China Experimental Fast Reactor CEFR, the Superphénix in France, etc. have provided an operational experience base of about 400 reactor-years. The Russian BN-800 and the **P**rototype **F**ast **B**reeder **R**eactor (PFBR) in India, both are at advance stage of construction.

The **I**nternational **A**tomic **E**nergy **A**gency (IAEA) which is a worldwide intergovernmental nuclear energy organization, examines and advices on current and evolving safety issues. The IAEA's Basic Safety Standards for radiation protection and the **N**uclear **S**afety **S**tandards (NUSS) for nuclear power plants have been adopted entirely or in part by many member states as the basis for national regulations. IAEA has strengthened its safety evaluation services for nuclear plant operations, by programmes such as IAEA-IRS which is an incident reporting system providing an exchange of nuclear plant operations experience in member states on safety related issues, to draw out the lessons learned, and to disseminate information among participants. Also ASSET: Assessment of Safety Significant Events Teams – to provide plant operators and regulators with independent analysis and guidance regarding specific events that have occurred, their causes and safety implications, and corrective actions that were taken for operational safety.

Trends in the nuclear power field also show that emphasis is shifting from design and construction to plant operation. Efforts are certainly being put into improving current nuclear plants designs as well as developing new concepts. There are also efforts under way to streamline construction methods and procedures in order to reduce construction time and investment costs, and to improve quality. More and more emphasis is being placed on plant operating performance as well. Safety, reliability and quality are the principal aspects where improvements are being promoted and achieved worldwide by the nuclear industry. Activities relating to plant operations, personnel qualification, man-machine interface, quality assurance, and especially operational safety are receiving increased attention. The nuclear industry will have to face the alternatives of plant life extension or decommissioning and this is another area that will gradually be receiving greater attention. For such requirement and advancement, knowledge preservation about the nuclear reactor becomes necessity.

So, for the nuclear reactor knowledge management system, knowledge is obtained from design, operational, maintenance, safety, quality assurance personnel etc. With the passage of time, the retiring people involved in the design and commissioning of nuclear systems would be replaced with new man power. Hence, the focus for the development of the portal is to make the skill / expertise available to new workforce. Because knowledge portals focus on particular domains, ontologies appear ideally suited to support knowledge sharing and reuse between knowledge portal providers and the users of the portal.

In the next section, knowledge management portals for nuclear system available in literature are described.

1.11 Review on Nuclear Related Knowledge Management Portals

IAEA aims to identify major nuclear reactors, radiation effects, waste and transport safety knowledge domains and then to capture and share the critical knowledge in each domain. IAEA provides guidelines and framework for designing the portal for knowledge management systems, which cover plant policies, its strategy, operation, safety, management and performance information etc [77-78]. This would lead to preserving the institutional memory and stimulate new knowledge for current as well as future generation of scientists, engineers and technicians. It also covers emergency response, country specific information, nuclear installation, thematic knowledge, etc. [4]. IAEA also conducts various technical meetings and conferences to collect knowledge from various nuclear power plants.

The Gesellschaft für Anlagen- und Reaktorsicherheit (GRS), Germany carries out research in reactor safety, radioactive waste management, radiation and environmental protection. An intranet portal has been set up by GRS to prevent the knowledge loss. In that, Ontoprise semantic tool is used for ontology representation [79]. The portal developed by GRS is embedded with document management, yellow pages, announcements, data collections, support information, suggestion box, new sticker for work scope news, message boards for department, knowledge representations of skill areas etc [80].

The **Krško** **N**uclear **P**ower **P**lant (Krško NPP) is a pressurised light water reactor power system of Slovenia. The intranet portal (named IntraNEK) at Krško NPP, allows the user to access various plant applications and links. In this portal, equipment details, structures, documents, human, regulatory requirements and commitments, non conformances, failure analysis, domestic and industry operating experience and corrective actions are covered. It establishes guidance on the effective and efficient use of operating experience information to improve plant/personnel safety, plant reliability and commercial performance [81].

Kazakhstan Atomic Energy Committee developed a portal to serve their day-to-day activities, to support working processes and to manage the documentation. It also facilitates to create knowledge base thereby providing an archive for nuclear knowledge [82].

NuArch project developed in Italy Trieste School of nuclear knowledge management, uses a web crawler that will identify and harvest the nuclear information resources from the Internet. The harvested information is automatically indexed and stored in a high-volume archive with version control and finally makes them accessible to the user [83].

As referred from IAEA technical report, KOZLODUY NPP, Bulgaria portal is developed based on FrameWork 1.1 and DotNetNuke. This portal covers plant operation, safety, system data, training and human resources etc. Apart from this, online technical parameters of the nuclear power plant units are also highlighted [84].

National Nuclear Energy Commission (CNEN), Brazil developed a nuclear knowledge portal for licensing and controlling the nuclear activities [85]. This portal defines the knowledge tree about licensing, control, legislation, regulation, training and documentation as main classes. Fuel cycle installation, radioactive installation, nuclear waste management, administrative rules and resolutions are defined as subclasses of the knowledge tree. It also provides an opportunity to share information and knowledge in real time among the collaborators.

Asian **N**etwork for **E**ducation in **N**uclear **T**echnology (ANENT) web portal is developed by Korean Atomic Energy Research Institute, Nuclear Training

Center, Korea. This portal is being used to share information about nuclear education and training information materials with the members of its institutions [86].

Japan **A**tomic **E**nergy **R**esearch **I**nstitute (JAERI) constructed a knowledge management framework for nuclear energy policy. This framework covers up to date important intellectual assets of JAERI [87].

Knowledge management portals listed above are developed for their respective institutes by employing semantic web. In most of the nuclear power plants, the web portals deployed are intranet based and hence obtaining data structure information about their portal is not easy. Moreover, the methodology adopted to implement them is also not available in the published literature.

The present work covers the development of the knowledge management for nuclear reactor portal at IGCAR.

1.12 Scope of the Thesis

The aim of the present work is to develop a knowledge management portal for nuclear reactor domain. A semantic web based knowledge representation is designed for FBTR, Kalpakkam and christened as **K**nowledge **M**anagement for **Nu**clear **R**eactor (KMNuR) portal. One of the major issue encountered which developing the portal is to avoid duplication / overlap in the ontology submitted. The editor tools for employing the portal are already discussed. To overcome the same an algorithm quick mapping evaluator has been postulated.

A matrix rank based ontology matching algorithm is also developed. Pareto optimization is used for finding the matching algorithm for getting optimized matching. The brief introduction about the same are described below.

1.12.1 Knowledge Management for Nuclear Reactor Portal

The portal involves the creation of knowledge base by collecting all the requisite knowledge available about the nuclear reactor. The content of the portal is organized in such a way that the adopted ontology would be known to users, so as to make an effective use of the same. The portal is developed for fast breeder test reactor at Kalpakkam. The knowledge relating to FBTR has been represented in the portal.

FBTR is a loop type experimental reactor consisting of reactor assembly, reactor core, control rod drive mechanism, primary sodium system, secondary sodium system, steam water system, steam generator system, fuel handling, reactor protection system, safety analysis, emergency core cooling system, auxiliary system etc. Creating KMNuR knowledge management portal would help in making the accumulated knowledge about FBTR system available for ease of reference.

The KMNuR web portal is a client/server architecture having user interface in the front end and database at the backend. The front end of the user interface developed using java web application, allows the user to get the requisite knowledge about the reactor from MySQL database at the backend. *Net Bean IDE* is used for developing the *Java* web based application, and *GlassFish* server for publishing the web application. The portal developed is pertaining to FBTR. The knowledge pertaining to FBTR is obtained from data sources like journals, books, internal reports, existing data in IAEA, open archives and by taking inputs from nuclear experts.

KMNuR comprised of knowledge represented in different semantic formats like RDF, OWL and UML to enable the web crawler to share and reuse nuclear knowledge. Systems and parameters of the nuclear reactor, whose knowledge is represented in the KMNuR portal includes reactor core, neutron energy, reactor dynamics, neutron fission, steam generator, electrical system, primary sodium system, secondary sodium system etc.

Knowledge assimilated about FBTR systems is represented semantically in the KMNuR portal. In addition to the sub systems of FBTR, other nuclear characteristics like nuclear flux, gamma ray source, reactor steady state, shielding materials are also covered in the portal. The overall aim of the KMNuR portal is to integrate and infer the semantic knowledge needed by the users in performing the nuclear reactor domain tasks. The methodologies used for developing components of portal for the FBTR domain are discussed in details in the forthcoming chapter. When the portal is developed, in order to integrate the new information submitted by the users, ontology matching algorithm has been developed in the present work.

1.12.2 Ontology Matching

Ontologies are developed for specific domain application. Relating ontology is very important for many ontology based applications. It is inevitable that similar ontologies are constructed and unifying these ontologies has to be done. This is done to make use of the knowledge available in one ontology in combination with other ontology. For matching, one tries to find two corresponding entities. These do not necessarily have to be the same. A correspondence can also be in terms of a lock and the fitting key. A certain degree of similarity along some specific dimension is sufficient, e.g., the pattern of the lock/key. Whereas combining allows many different relations at the same time, matching implies one specific kind of relation. A typical scenario for matching is web service composition, where the output of one service has to match the corresponding input of the next service. Any schema matching or ontology matching algorithm may be used to implement the Match operator. Matching corresponds to our definition of general alignment, however, where a fixed relation between the aligned entities expresses the kind of match. Classifications of ontology matching techniques are shown in the Figure 1.3.

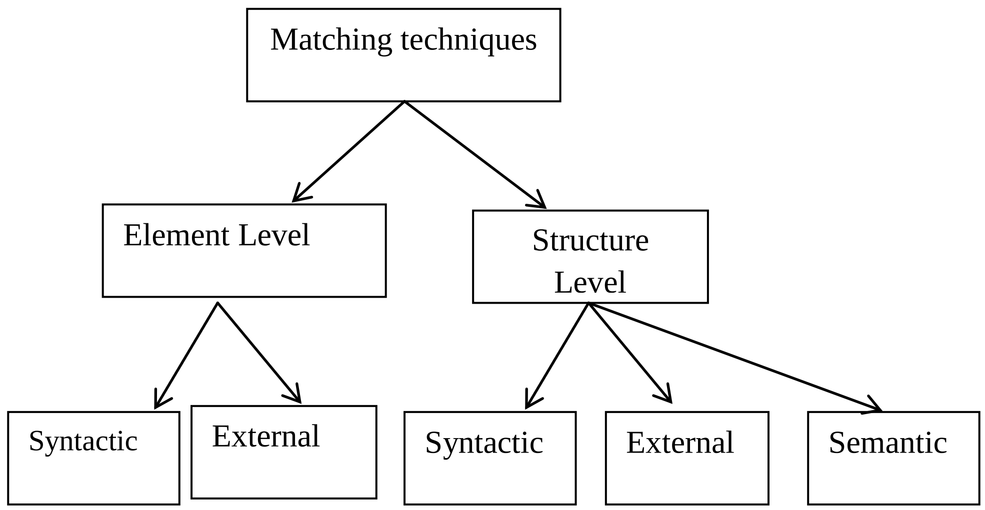


Figure 1.3 A fragment of classification of matching techniques

Element level

Element level matching techniques compute correspondences by analyzing entities or instances of those entities in isolation, ignoring their relations with other entities or their instances [88].

Structure Level

It computes correspondences by analyzing how entities or their instances appear together in a structure.

Syntactic Level

The key characteristic of the syntactic is that they interpret the input following some stated algorithm.

External Level

For matching technique resources of a domain and common knowledge are used to interpret the input.

Semantic Level

In this technique some formal semantics are employed for finding the matching.

1.12.2.1 Related Works

We have summarized the ontology matching algorithms mentioned by Shaviko and Euzenat [204] briefly. Among the several system that have appeared in recent years they selected which are repeatedly participated in ontology alignment evaluation initiative campaigns.

Falcon, RiMOM, Anchor flood and AgreementMaker are developed for generic matchers. SAMBO and ASMOV were developed for biomedical ontologies. The algorithms proposed mostly used background knowledge base like UMLS for that particular domain. Many of these algorithms don't have graphical user interface. Most of them produced results by either combined the terminological and structural results or aggregate them and output as 1:1 alignment. The best match is the one which maximizes all the similarity measures. However, in practice there may be conflicting values amongst the similarity measures thus impeding the matching process. To resolve this issue, there is a need for optimization. Optimization means finding a solution which is most appropriate. Pareto trade off analysis is a method where a multi-objective optimization is performed. This does not give a single solution but a set of solutions called pareto set of minimal elements defined by the dominance relation.

1.12.3 Development of Ontology Matching Algorithms

When two ontologies differ in representation, they lead to disorder and influence the interoperability of ontologies, because they provide a relatively equal description (or) represent the different viewpoints of the same domain. Entities of ontologies (matched ontology) are used for ontology merging, query answering, data translation, domain knowledge sharing, and also navigating in the semantic web. In addition, ontology merging, integration and alignment can lead to ontology reuse. In continuation with our work of nuclear knowledge management system for nuclear reactor domain a Quick Mapping Evaluator, which is an application program for ontology mapping and matrix rank based ontology were developed. They are briefly described below.

1.12.3.1 Quick Mapping Evaluator (QME)

QME allows the user to choose ontology alignment algorithm like String Equality, Longest Common Subsequence (LCS) and Levenshtein distance for extracting the shared knowledge. The users are allowed to change the type of match making algorithm. The StringEqual algorithm has a very high precision, a very poor recall and also a very low F-Measure irrespective of the thresholds. The LCS algorithm steadily increases in precision with increasing threshold, but decreases in recall. The overall F-Measure of the LCS algorithm increases with threshold. The Levenshtein distance also has a steady increase in the precision and decrease in the threshold. The overall F-Measure remains almost consistent at a high value. The Levenshtein distance algorithm has the highest F-Measure and hence gives the best accuracy.

1.12.3.2 Matrix Rank based Ontology

Mapping/Alignment techniques are used to find proximity between the entities of two ontologies. Matrix rank based ontology, proposed in this thesis is an extension of String based ontology matching to calculate the ontology matching. Partial overlap (or) duplicate (or) unique ontology are determined by using matrix ranking methods thereby easing inclusion or exclusion of ontology required for reuse. Two approaches are possible one by merging the ontologies to create a single coherent ontology and the second one is to align the ontology by establishing link between them and allowing them to reuse information from one another.

Each entity of ontology O is matched with other ontology O' and the result of comparison is inserted into a matrix. If entity is exactly matched then the matrix is filled with '1' else the matrix is filled with '0'. String based ontology matching technique is found to be faster and stable. While finding the rank of the two ontologies, a value of '0' signifies that ontologies are unique. If rank value lies between 0 and m (where m is the size of the matrix), it indicates that the two ontologies are related. Identity or unit matrix implies that it is a duplicate ontology.
The algorithm proposed aids the agent program to filter and collect the knowledge for the search, reuse and share, thereby giving relevant answers to the user query.

1.12.3.3 Pareto Optimisation

Pareto optimality is a domain independent property that can be used to coordinate distributed ontology. A Pareto rank learning technique is proposed for enhancing multi objective evolutionary optimisation. The goal of multi objective optimisation function is to generate various feasible solutions which are closer to Pareto front from which the best solution could be selected. A Pareto optimal solution cannot be uniquely determined. Usually there exists a set of solutions that all satisfy Pareto optimality which form the Pareto front in the solution space [89].

The optimisation refers to finding the best possible solution to a problem given a set of limitation or constraint. Finding the mapping between two ontologies, involves computing ontologies at lexical level and conceptual level. In lexical level comparison, string based equality, Leventhein distance and Least common subsequence algorithms are used. Structural level comparison is achieved through probability distribution techniques like Kullback Divergence, Cosine method. Then the Pareto ranking is done to find the optimised solution for ontology matching. Before calculating Kullback Leibler divergence have to calculate the conditional probability and frame the probability mass vector function. For this, pre defined thesaurus files are used to count instances where attributes are associated with other attributes.

The decision to merge is determined by number of similar items in each ontology and thesaurus file. The user sets an acceptable level of similarity and the decision to merge is made by machine. Once the Kullback Leibler, Cosine divergence, Lexical similarity (using string equality, Leventhein distance, Least common subsequence) are calculated. If the results meet the acceptable level, the ontologies are suitable for merging. This is a semi automatic method of merging ontologies and users should give the threshold value. The minimum value is greater than the threshold then the alignment is accepted or else it is rejected. An algorithm is proposed for finding the matching using Pareto optimization technique and implemented in FBTR.

The work carried out are organized and divided in to four chapters.

Chapter 2. Semantic Web based Knowledge Representation Schemes and Tools: In this chapter, literature survey on ontology application in various domains, survey on ontology editor tools is covered. Based on a literature survey conducted in the present work, Protégé IDE is found to be best suited ontology tool for knowledge management system representation.

Chapter 3. Semantic Web Based Knowledge Management Portal for Nuclear Reactor Domain: In this chapter the need for knowledge management portal and introduction about Fast Breeder Test Reactor are presented. The developed semantic web based portal is christened as Knowledge Management for Nuclear Reactor domain (KMNuR). It involves the creation of knowledge base by collecting all the requisite knowledge available about the nuclear reactor. The content of the portal is organized in such a way that the adopted ontology would be known to users, so as to make an effective use of the same. The portal is developed for fast breeder test reactor at Kalpakkam. The knowledge relating to (Fast Breeder Test Reactor) FBTR has been represented in the portal. Reactor system consists of various sub systems like Primary Sodium System, Secondary Sodium System, Reactor Assembly, Reactor Core, Steam Water System, Steam Generator System, Control Rod Drive Mechanism, Fuel Handling, Reactor Protection System, Safety Analysis, Auxiliary System etc. The representation formats utilized like OWL, OWL-GRAPH, OWL_UML are elaborately discussed.

Chapter 4. Development of Algorithms for Ontology Mapping: In this chapter**,** introduction about ontology mapping and string similarity measures are discussed. To enhance the process of ontology matching a Quick Mapping Evaluator algorithm tool is introduced. Another outcome of the study is based on the string based ontology mapping by using matrix rank based technique. Using this matrix ranking algorithm, partial overlap (or) duplicate (or) unique ontology is determined. The result of the ranking algorithm decides whether to eliminate or reuse or share the knowledge. Also a Pareto based optimisation technique employed for ontology matching to find the optimised algorithm is presented.

Chapter 5. Conclusions and Scope for Future Work: The work carried out in the thesis is summarized and scope of work to be carried out in future is discussed.

CHAPTER 2

SEMANTIC WEB BASED KNOWLEDGE REPRESENTATION SCHEMES AND TOOLS IN VARIOUS DOMAINS

A survey of development ontology in various domains is described. The tools required for development, ontology language used for representation, programming language, database, reasoner etc., employed in the various domains are discussed. A comparative study on the various ontology editor tools like OntoStudio, Protégé, SWOOP and TopBraid tool are also discussed. Based on this, the appropriate tool required for ontology development is selected.

2.1 Introduction

Ontology is essentially annotated taxonomy of the world one wish to describe for sharing data and for interoperability. For example, ontology about a nuclear reactor would contain information about reactor core, coolant system, steam generator, detector, protection system, fission reaction, etc. Hence the necessary resources about the nuclear reactor have to be generated through proper semantics, so that a meaningful ontology exists about the reactor to the user community. In this connection a restricted set of semantic about the nuclear reactor are generated and represented through a proper ontology. The review of application of ontology in various domain is carried out based on the available research papers, referred journals, reports in the respective domains, scholarly articles etc [90]. A broad picture of ontology applications in various domains practised is surveyed and described (Figure 2.1).

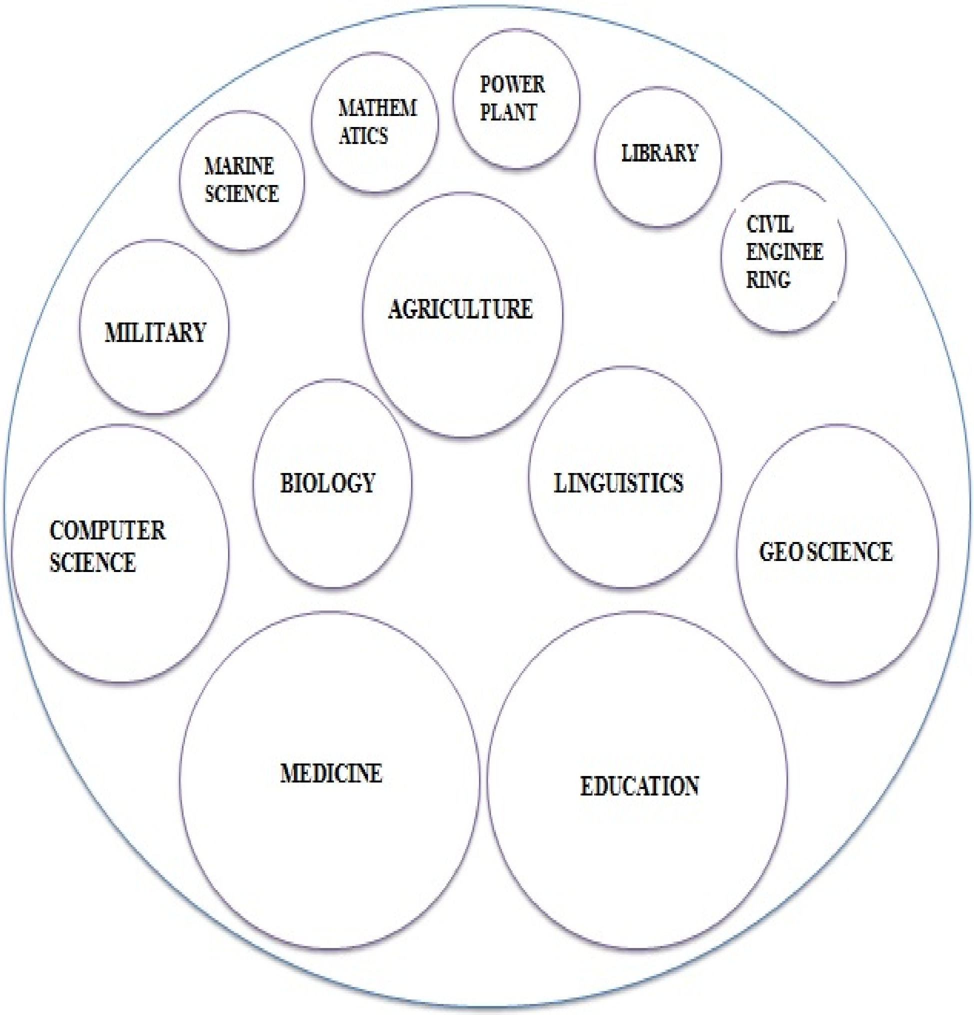


Figure 2.1 Ontology applications developed in various domains

Enhanced development of ontology would aid in the evolution of semantic web leading to complete sharing of knowledge in a given domain. It can also be inferred that the ontology development is a continuous process and success could be achieved by participation of the domain experts and users. As the development of ontology is limited in the field of nuclear energy, emphasis is given to create knowledge representation in the nuclear reactor domain. In order to represent the domain knowledge in ontology, integrated development tools like Protégé, Model Futures OWL Editor, TopBraid Suite, OntoLingua, OntoEdit, WebODE, KAON, ICOM, DOE, WebOnto, Medius Visual Ontology Modeler, LinKFactory Workbench, K-Infinity and OntoStudio are required. Some of these are available as freeware tools. A survey of the ontology tools is also done to identify a suitable tool to represent the knowledge in nuclear reactor domain. Protégé, SWOOP (open source tools) and OntoStudio, TopBraid (commercially available) are taken for analysis and their features are compared in the following section.

2.2 Survey on Ontology Application Domain

The survey was carried out to analyze the progress of ontology development in various domains. This is carried out using the available information in the web site, journal article, reports etc. Need of the ontology, its development tool, algorithm employed for processing, language supported, platform supported etc are also reviewed. The survey paved the way to develop ontology for nuclear reactor domain.

2.2.1 Ontology Developed in the Domain of Agriculture

Several institutions and organizations provide educational resources on the agricultural topics, some of them openly available through the web. Promotion of ecological practices and sustainable agriculture requires much effort in terms of education. Avoiding pesticides and promoting organic farming helps to preserve earth’s natural fertility, these practices have come after sustained efforts and database collection. However, locating those resources with conventional search engines is complicated, mainly due to noise in the results of common input terms. It is possible to get richer additional information about the agriculture domain by browsing, navigating and searching for educational resources by utilizing the formal annotation based on ontologies. The food and agriculture organization of the United Nations is recognized as an information and knowledge base organization, whose activities comprise in capture and analyze, disseminate and share, localize and provide information and knowledge about the agricultural field. Some of the semantic web based knowledge management portals, in areas such as food and safety issues, soil-plant-nutrient processes are discussed below.

AGROVOC is used for searching of vocabulary associated with agriculture in various systems throughout the **F**ood and **A**griculture **O**rganization (FAO) [91]. It is represented in RDF/SKOS-XL linked format and accessed through SPARQL. AGROVOC enables the machine and the users to access the structure and standardize the agricultural, forestry, fisheries, food and other related domains in agriculture terminology in multiple languages [92 - 93]. It contains over 32,000 concepts organized in a hierarchy, each concept having labels in twenty languages: Arabic, Chinese, Czech, English, French, German, Hindi, Hungarian, Italian, Japanese, Korean, Lao, Persian, Polish, Portuguese, Russian, Slovak, Spanish, Thai, Turkish. Four more language versions are under development.

Food Safety Semantic Retrieval System is an ontology-based semantic retrieval system which includes all aspects of knowledge about food safety. This system helps the users to access the accumulation of the knowledge in the food safety domain. It is developed using *JDK* in *MyEclipse* integrated development tool and *MySQL* as database [94].

Agriculture Literature Retrieval System captures the concepts from the encyclopedia of Chinese agriculture and catalogue of ancient Chinese literature on agriculture. The ontology extracted from this domain is described in XML, RDF, OWL formats. There are more than 10,000 keywords extracted from the research papers of Chinese agricultural history [95].

Citrus Water and Nutrient Management System is defined for water and nutrient balance processes in citrus production with 700 symbols and 500 equations. This includes concepts like block, soil cell, soil profile, soil layer, root distribution, irrigation system, weather etc. All models of dynamic systems in agriculture and natural resources are defined by a set of mathematical equations. *Mathematica* and *Simile* are used to design and build models at mathematical level. *Java* language based program is used for the finding similarities and difference between equations in different models [42].

OntoSim-Sugarcane is an ontology based application which represents hydrology, nutrient cycling, plant growth, soil moisture, crop growth on organic soils and nutrient uptake in sugarcane production in Southern Florida. All the process identified for simulating sugarcane growth on Florida organic soils have been represented as mathematical equations. This collection consists of 195 equations and 247 symbols. Equations and symbols are stored internally as ontology objects. To run a simulation, *JAVA* based code is used to generate objects automatically from equation and symbols. Then the simulated mode is debugged for errors [96]. The domain ontology of agriculture is shown in Figure 2.2.

2.2.2 Ontology Developed in the Domain of Aviation

In civil aviation domain, concepts and role are defined for aircraft, non-aircraft, emergency-in-aircraft, emergency-in-nonaircraft etc in order to improve the reasoning ability of civil aviation knowledge by developing ontology.The ontology is described using **D**escription **L**ogic (DL) language which facilitates checking of rationality by defining concepts, roles and instances. Pellet reasoner is used to check the consistency of ontology and individual retrievals are done using *Jena* reasoner [97].

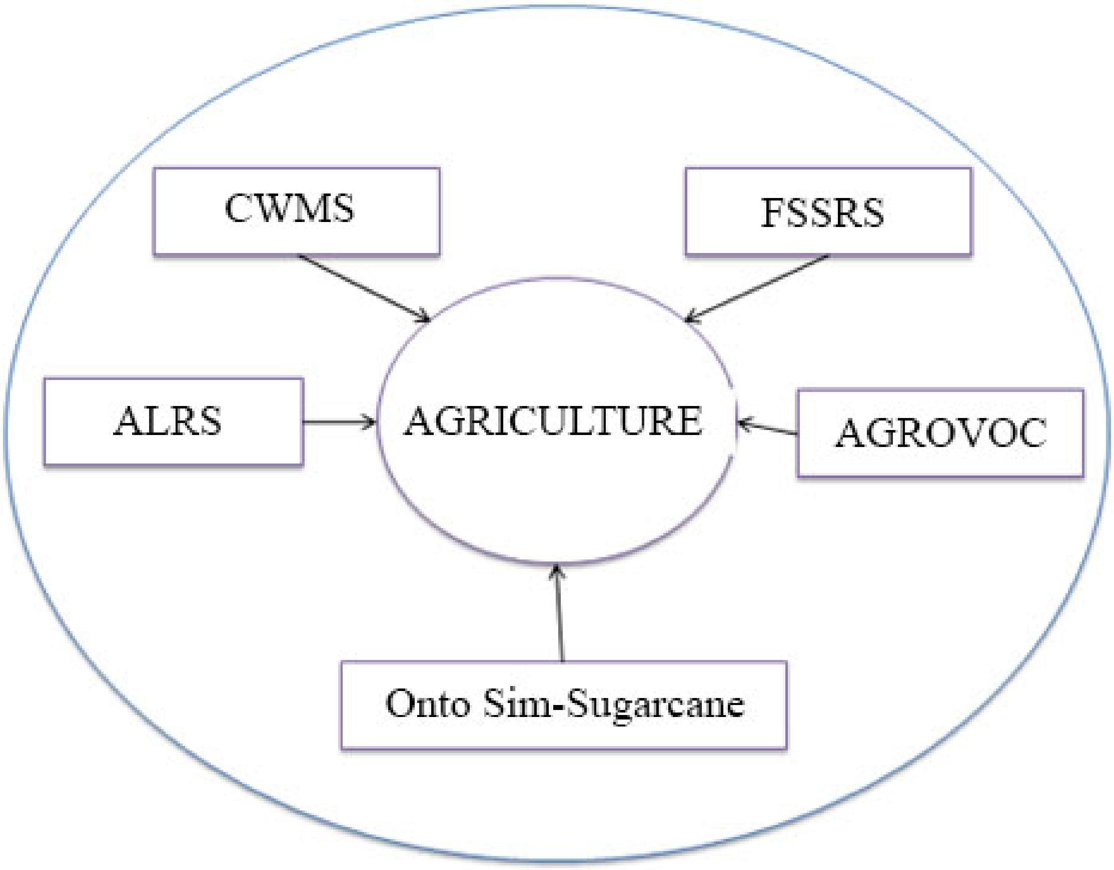


Figure 2.2 Ontology in the field of agriculture domain
FSSRS: Food Safety Semantic Retrieval System; ALRS: Agriculture Literature Retrieval System; CWMS: Citrus Water and Nutrient Management System.

2.2.3 Ontology Developed in the Domain of Biology

The plant ontology is a community resource designed to fulfill the need for uniform terminology to describe plant structure and developmental stages. The plant ontology consortium builds upon the work by the gene ontology consortium by adopting and extending its principles, existing software and database structure. The driving force of these consortiums is that they allow the fruits of research in one plant species to be more easily used in the study of other species, leading to a greater understanding of plant biology. With the advancement of computer graphics, the results on the works on simulating the appearance and growth of the plants can be used to forecast the production using computer based simulation. As the existing plant models are difficult to satisfy the needs of knowledge sharing, the need for plant ontology is gaining importance. The plant ontology may provide new ways to ensure the clarity integrity, expandability of representation of botany knowledge sharing. With the available software tools for querying the database, it is possible to get huge amount of biochemical and molecular information on all classified enzymes. Some of these ontology tools developed in the field of biology are discussed.

Plant Ontology Database describes the controlled vocabulary (ontology) for plants. It is a collaborative effort of model plant genome database developer and plant researchers, to create and maintain the database. It also implements a semantic framework to make meaningful cross-species and database comparisons [98]. Plant anatomy consists of 30087 terms defining a controlled vocabulary of plant's morphological and anatomical structures. It represents organs, tissues, cell types and their biological relationships based on spatial and developmental organization. Stamen, gynoecium, petal, parenchyma, guard cell, plant structure etc are defined in this ontology. In plant structure, ontology consists of a controlled vocabulary of growth and developmental stages in various plants and their relationships. This is an internet based application and the programming is done using *Perl* language. It has a collection of more than 1600 unique genes [99].

Plants Domain Ontology has the collection of knowledge in respect of botany and environment. The knowledge in botany consists of the geometry and topology. The knowledge of geometry is used to describe the shape of the plants and three-dimensional geometry of organs. The topology knowledge is used to describe the location relationships to generate three-dimensional structure of plants. Protégé tool is used to describe the plant domain ontology [100].

Transparent Access to Multiple Bioinformatics Information Sources (TAMBIS) uses ontology to enable biologists to retrieve knowledge by querying the multiple external databases. It has unrelated applet containing two versions having approximately 1800 concepts of bio informatics and molecular biology. An applet linked to external resources has 250 concepts and relationships related to protein which is used to browse the model and answer the query. Collection Programming Language (CPL) which supports multi database language is used for storing the data resources of TAMBIS. Java programming language is used to interface the components of the TAMBIS, GRAIL query processor is used for processing the query to retrieve data from the CPL program [101].

BRaunschweig ENzyme DAtabase (BRENDA) is maintained at the Technology University, Germany. It allows the users to search simultaneously the whole of biochemically relevant ontology. It contains approximately 3400 terms on tissues, organs and cell types. **F**ull **R**eference **E**nzyme **DA**ta (FRENDA) and **A**utomatic **M**ining of **EN**zyme **DA**ta (AMENDA) are additional databases created by text-mining procedures. Biochemical and molecular properties of enzymes such as classification and nomenclature, reaction and specificity, functional parameters, organism-related information, enzyme structure, isolation and preparation, literature references, application and engineering, enzyme–disease relationships are also available. BRENDA contains ontology definition for 3300 enzymes in 500 different organisms [102]. The domain ontology of biology is shown in Figure 2.3.

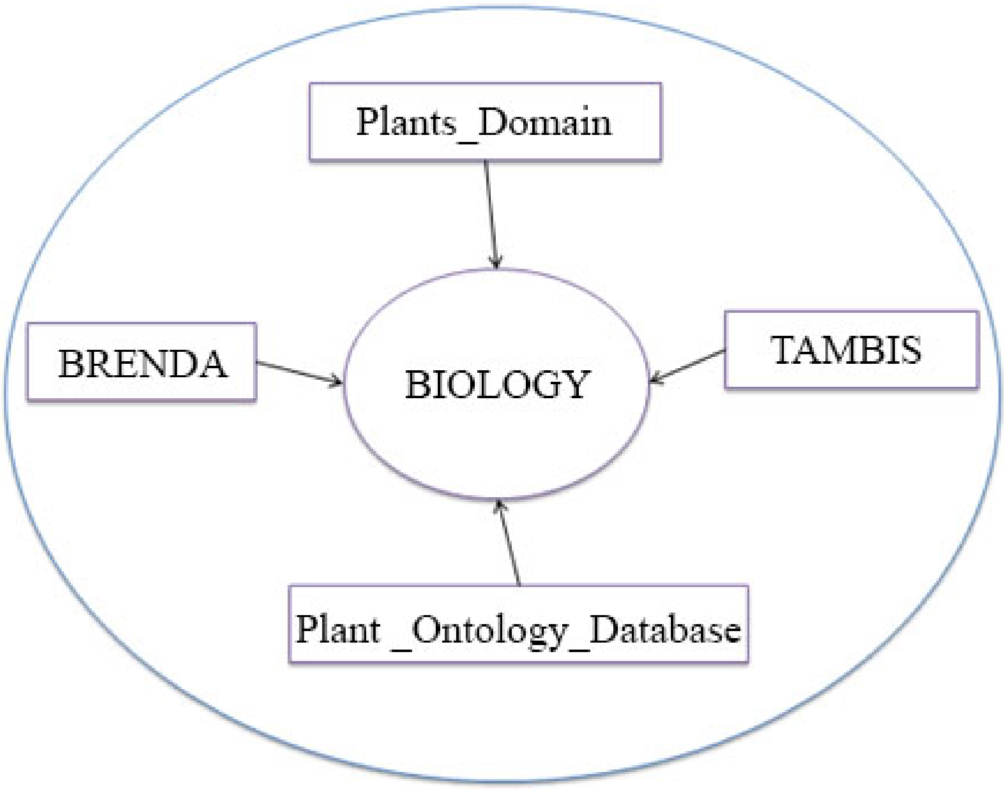


Figure 2.3 Ontology in the field of biology domain
BRENDA: BRaunschweig ENzyme Database; TAMBIS: Transparent Access to Multiple Bioinformatics Information Sources.

2.2.4 Ontology Developed in the Domain of Chemistry

Ontology in the field of chemistry domain is being used in several applications. It is an information-retrieval system that lets users consult and access, in their own language, the knowledge contained in the chemistry domain. Other chemistry ontologies acts as a teacher broker that allows the students to learn chemistry in a simplified way by providing the necessary domain knowledge. These domains help to test student’s skills. The molecular entities which represent both natural and synthetic products in the area of living organisms demand a database in order to understand the metabolism of organisms on a systematic level. The requisite ontologies developed in response to these are discussed in this chemical ontology.

In the **chemistry domain**, ontology has two main categories: chemical-elements and chemical-crystals. Chemical-elements have 16 concepts, 103 instances, 3 functions, 21 relations and 27 axioms. Chemical-crystals have 19 concepts, 66 instances, 8 relations, and 26 axioms. Chemicals also include public Ontolingua ontologies, such as standard-units, standard-dimensions, and knowledge interchange format lists. The chemical ontologies represented in Spanish in the Ontogeneration (information retrieval system) allow the user to consult and access the knowledge contained [103].

Chemical Entities of Biological Interest (**ChEBI**) is a dictionary of molecular entities in the 'small' chemical compound. Number of entities in ChEBI is approximately 18000, whereby the relationships between molecular entities or classes of entities and their parents and children are specified. ChEBI text search allows the user to search all the data based on the ChEBI identifiers, names, database link, formula etc. Open source *OrChem* project is used to develop chemical structure searching in ChEBI database. The Ontology defined in ChEBI has more than 400,000 molecules [104].

2.2.5 Ontology Developed in the Domain of Civil Engineering

Research in the area of healthy housing has drawn much attention in recent years. Hence the ontology in the field of construction domain has already gained importance. This is due to the fact that healthy housing, is an important concept on construction field, which enhances the status of health, fundamental resident qualifications, and also safe, convenient, comfortable and healthy inhabitancy. Some of the ontologies developed in the field of housing and national building infrastructure are discussed.

Healthy Housing is an important concept in construction field. It defines the regulation of vocabularies characterizing the residency from the knowledge of healthy housing, creating the terminologies and relations for the core vocabularies of the field. OWL language is used for defining these terminologies and relations. **M**ethod of **A**nalysis, **D**esign, **R**epresentation and **E**valuation (MADRE) graphic language is used to acquire specific subclass, instance and conceptual elements [105].

Semantic web ontology project in **Finland ONTOlogy** (**FinnONTO**) is taken up to build a national infrastructural foundation for the semantic web. This is carried out by establishing a large research consortium representing universities, public organizations, companies working in collaborative mode. It contains more than 10,000 resources that define the meaning of individual persons, organizations, locations, artifact types, actions etc [106]**.** The civil engineering domain ontology is shown in Figure 2.4.

2.2.6 Ontology Developed in the Domain of Computer Science

The field of computing has evolved to encompass a number of distinct sub disciplines. However, many of these sub fields have still overlapping nature. New programs are being developed in the information technology and software engineering to join the standards in computer engineering, computer science and information systems. These various programs make it difficult for the general public to understand the goals and content of the program. The source of these problems is the absence of an explicit and fully developed conceptual model to formalize them. Ontologies are a way to overcome these obstacles. Particularly in the field of software business only little and vague guidance is available for software product management. Also interoperability problem in heterogeneous network is also solved nowadays through domain ontologies. Some of the ontologies developed in the field of computer domain are discussed.

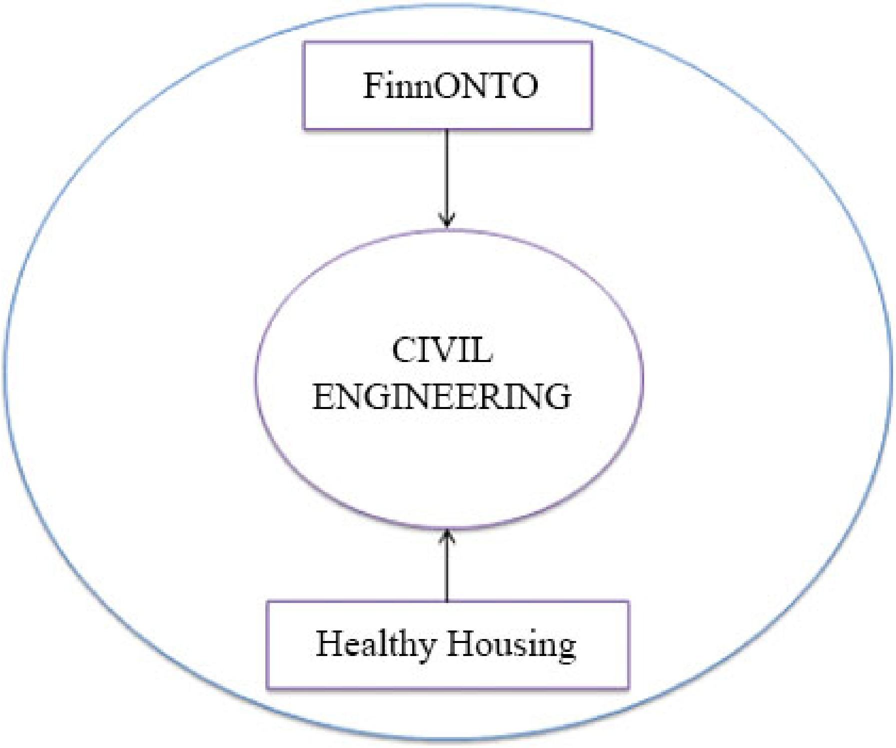


Figure 2.4 Ontology in the field of civil engineering domain
FinnONTO: Finland Ontology

The **Innovation and Technology in Computer Science Education (ITiCSE)** ontology describes various disciplines, topics, subtopics that belong to the domain of computer sciences. It is a collection of computing topics, early methods, devices, machines, hardware (non-system), hardware-software systems, software, theory, people, institutions and milestones. OWL file format is used for ontology representation [107].

Standard Ontology for Ubiquitous and Pervasive Applications (SOUPA) is designed to model and support pervasive computing applications. Ontology is expressed using the OWL. It includes modular component vocabularies to represent intelligent agents with associated beliefs, desires and intentions, time, space, events, user profiles, actions and policies for security and privacy.

Description logic reasoner like Racer and FaCT are used for reasoning the ontologies. The SOUPA ontology is used in **C**ontext **Br**oker **A**rchitecture (CoBrA) to facilitate knowledge sharing and ontology reasoning [108].

Software System Ontology is a combination of domain ontology and class diagram ontology. Domain ontology is domain vocabulary which is built by domain experts, while class diagram ontology is automatically populated from source code to represent the knowledge in the code. It also includes method of class diagram to ontology transformation and algorithm of ontology combination. Description Logic is a knowledge representation formalism used for representing the ontology [109].

Software Product Management domain aims to identify the recent domain-specific research on software product management. It extracts text corpus with respect to terms, concepts, hierarchical relations of the concepts and the non-hierarchical relationships between the concepts used for ontological learning process. RapidMiner data mining software is used to extract terms from the ontology and TextToOnto ontology learning system is used in this domain [110].

Open Mind Indoor Common Sense (OMICS) is a collection of commonsense data consisting of 152098 items by 1009 users for indoor mobile robots. Based on the concepts, there exists four types of relations: hierarchical, semantic, sequential and coherent. Initially the relatively semantic granularity of concepts hierarchies are measured. The measured semantics are converted to relative probabilities among concept hierarchies and then ranked them according to probability. Finally, the probabilities to relative weights of relation using Bayesian networks are calculated. Thus the reasoning is done based on the calculated stochastic weights for the relations. [111].

Video Indexing and Retrieval consists of two types of ontology namely object ontology and shot ontology. In object ontology users are allowed to query a video collection using semantically meaningful concepts without the need for performing manual annotation of visual information. But in shot ontology, users are allowed to retrieve the video by submitting either single or multiple keywords queries. A segmentation algorithm is used in the video indexing and retrieval system. For indexing the large video databases, unsupervised spatiotemporal algorithm is employed [112].

In the **Software Engineering Lifecycle**, ontology is defined for each phase from analysis, design, requirement engineering, component reuse, implementation, integration, testing till documentation etc. RDF, OWL, UML are used for ontology representation. Protégé ontology editor is used for development of software engineering lifecycle [113].

In **Telecommunication Management Network Model**, Ontology is introduced to fix the interoperability problem of the network and its equipment. In the domain of network ontology, concepts like tangible router interface, intangible border gateway protocol parameters, network objects and management operations are defined. OWL and description language are used for ontology representation. Ontologies are constructed using Protégé and ontology mapping is implemented using *Java*. Approximately 250 Cisco commands and 200 Novel commands were analyzed [114].

Image Classification Using Neural Networks is used to classify objects from an image. The network takes an image as input and gives classification as output which is processed by ontology to discover the relationships among objects. Image segmentation algorithm is used for finding individual objects in the image. Pruning algorithm is used for descending order sorting the concepts based on the ranking [115]. The domain ontology of computer science is shown in Figure 2.5.

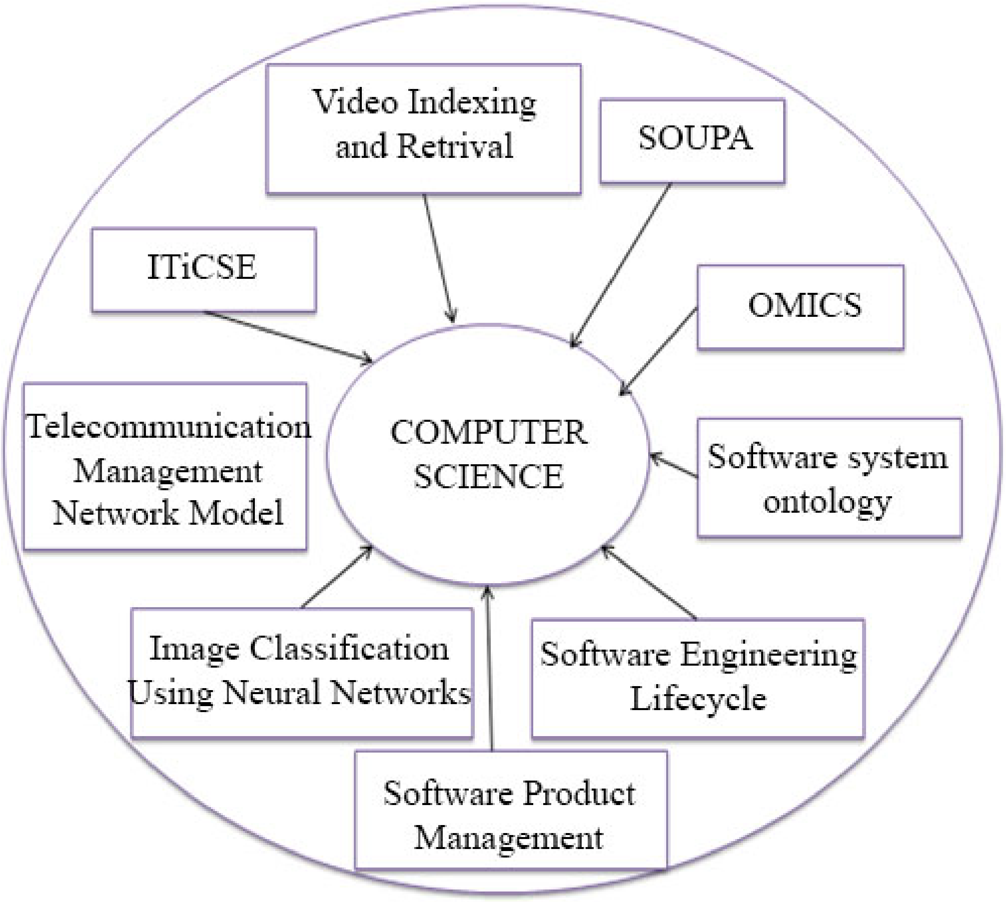


Figure 2.5 Ontology in the field of computer science domain
ITiCSE: Innovation and Technology in Computer Science Education; SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications; OMICS: Open Mind Indoor Commonsense.

2.2.7 Ontology Developed in the Domain of e-Business

Information and document exchange among people is possible through web. However, the same is not true when information is exchanged among software applications. Ontology based approaches have found a solution for this.

Symbolic Ontology XML-based management system (**SymOntoX**), is a software for the OntoPrivacy, and supports a legal database for the protection of personal information. and Interoperable minimal harmonise ontology. **B**usiness and **E**nterprise **O**ntology (BEO) represents the core of an ontology-based platform for business games. SymOntoX has added advantage of multi lingual support. OWL, SHOE and XOL are also employed. This service is available in internet and java language is used for achieving interoperability and platform independent [116].

2.2.8 Ontology Developed in the Domain of Education

Computer aided education has an important role to play in the developing countries, particularly in the area of higher education. Web based teaching has now become popular, hence, learning resources available over the network programs and servers have to be properly integrated so that they can be retrieved and shared by the users. Architectures to support interoperability among various web-based educational information systems are of current interest. Furthermore, to have an automated, structured and unified authoring support for their creation is the other challenge. With the development of network educational resource, the way the people learn have changed a lot from traditional teaching to resource based teaching and learning. With this view point, an overview of the ontologies developed in the field of education application is presented. Sharable content object reference model is a collection of standards and specifications for web-based e-learning technology [117].

Topic maps for e-Learning provides support for creating and using ontology-aware topic maps-based repositories of online materials. It includes research papers, special issues of journals, books, projects, software, conference papers, workshop proceedings, mails, research labs and working groups having a unique URI. Topic map editor and Topic map viewer are tools used for standalone application. Topic map editor supports topic maps merging. Topic map viewer is used to get graph view, text view and tree view [118].

Web-Based educational systems, defines communication ontology, communication content ontology and interaction protocol ontology. XML is used to represent the information, agent communication language, **K**nowledge **I**nterchange **F**ormat (KIF) and *Prolog* are used in this architecture. **S**imple **O**bject **A**ccess **P**rotocol (SOAP) is a standard lightweight protocol used for exchanging information [119].

Ontologies for Education (O4E) consist of concepts like technological perspective and application perspective. Technological perspective defines the knowledge organization, knowledge inference, information, information vizualisation, information navigation, information querying, subject domain ontology, instructional knowledge. Application perspective defines sub concepts in knowledge construction, knowledge externalization, knowledge communication and architectural knowledge [120].

Ontologies for the Use of digital learning Resources and semantic Annotations on Line (**OURAL)** defines ontology in the e-learning domain, which includes problem-situation, problem solving, critical analysis, case study, debate, cyber quest, project, exercise etc. The resources can be digital or non digital merging with part of the learning context. Protégé editor is used to represent the ontology [121].

Learning Resource metadata for the description of the content of learning resources provides a range of standard and universal method. An educational learning resource ontology semantic network defines concepts like ontology learning resources, context, ontology learning resource structure. IEEE **L**earning **O**bject **M**etadata (LOM) binding with RDF is used for representation of education learning resources [122].

Network Education Resource Library, uses *JSP* and *Java bean* language in its architecture. OWL based ontology representation is used for storage DB2/MySQL/ORACLE database is utilized [123].

Semantic web **ontologies for e-learning systems in higher education** consists of user profile ontology, the person ontology, the contact ontology and the activities ontology. This is implemented using Protégé-OWL ontology editor [124].

European credit vocational system uses ontologies in the construction of the educational resources library, to provide a common access to the information regarding the qualification systems of nine European countries. The Dutch system of **Secondary Vocational Education**, which entails **I**nternational **S**tandard **C**lassification of **ED**ucation (ISCED-97) for educational levels 3 and 4, is designed to impart professional training to students of age 14 to 18 and adults at **I**nternational **S**tandard **C**lassification of **O**ccupations (ISCO-88) at skill level 2. Using Protégée, the ontology of core duty, specific information, core part, qualification profile, competence, core assignment, certifiable unit, criteria for mastery, exit differentiation etc. are defined [125].

National School of Commerce and Management defines **Ontologie_US_ENCG** consists of administrative documents, project documents, financial information, legislation, institutional life, procedures, pedagogical document, exploration, capitalization, management, examination, reports of meeting and general policy classes. Protégé 2000 editor is used for development [126].

Virtual Lab Ontology isdeveloped using Protégé. Ontology "VLabResources" is defined to include all resources needed for any practical activity in an engineering education program. The classes like subjects, competence and tasks are defined to perform the practical activities of the Virtual Lab in a virtual learning environment. Standard reasoner tools like Pellet or FaCT++ for validation are used [127].

Economic ontologies are designed to represent the structure of economic knowledge in Croatia to define taxonomy of economics. It represents institutional curricula, academic discipline, documenting the data and metadata, meta data about learning and management systems, online resources for training materials and teaching [128].

Cultural Artefacts in Education (CAE) ontology is defined for countries like China, UK and Ireland. It consists of interrelated sub-ontologies, authority, group language, lesson and data [129].

Sahayika is used for building knowledge structures in education domain in India and its interface is available in both English and Bengali language. This deals with school education domain which covers subjects like biology, geography, physics, chemistry, history etc. [130].

Remote education system is classified as teaching practice and teaching management. Teaching practice ontology covers browsing of course, learning, online examination and online direct learning, electronic courseware management etc. [131].

The purpose of the **Information and Communication Technologies (ICT) Education ontology** is to provide a central repository of classified knowledge in ICT education. ICT ontology consists of concepts like ontology of ICT curriculum, ontology of ICT job, ontology of ICT skill and ontology of ICT research [132]. The domain ontology of education is shown in Figure 2.6.

2.2.9 Ontology Developed in the Domain of Electronics

Sensor networks deploy heterogeneous sensing nodes for capturing environmental data. These sensors help to enhance the search task and obtain value knowledge that is unreachable using classical information retrieval techniques.

Sensor networks domain is implemented through **Suggested Upper Merged Ontology (SUMO)**. It contains approximately 25,000 terms and 80,000 axioms about CPU processing power, memory, power supply, radio and sensor modules. The SUMO ontology comprises various domains such as computing services (networks, systems, and services), finance, geography, time, economy and transportations [133].

2.2.10 Ontology Developed in the Domain of Geoscience

Geoscience information is the key to effective planning and decision-making in a variety of application domains. Literature survey indicate that lot of efforts has been undertaken to increase the software tools for web searches in respect of earth science data and information through semantic web.

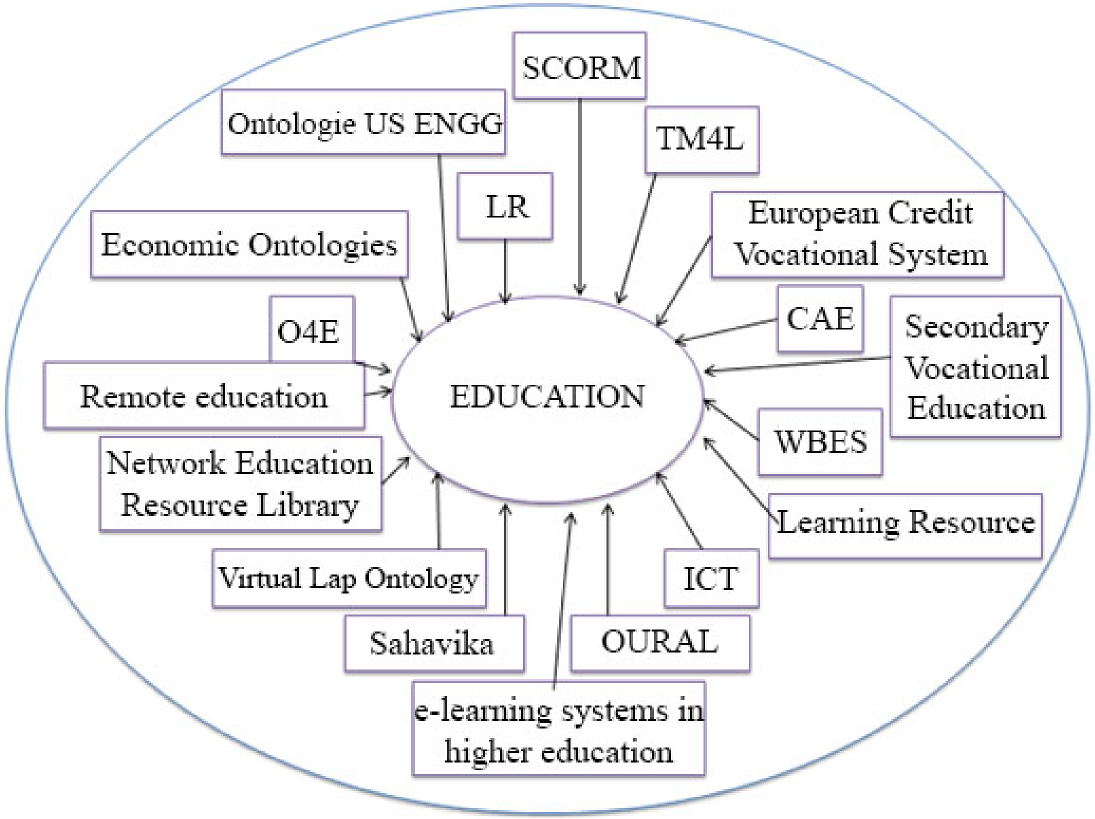


Figure 2.6 Ontology in the field of education domain
SCORM: Sharable Content Object Reference Model); TM4L: Topic Maps for e-Learning; WBES: Web-Based Educational Systems; O4E: Ontologies for Education; LR: Learning Resource; OURAL: Ontologies for the Use of digital learning Resources and semantic Annotations on Line; CAE: Cultural Artefacts in Education; ICT: Information and Communication Technologies.

The different kinds of geoscience data which is documented at different locations and distributed in many formats have to be made interoperable. Multi-level ontologies are a pre-requisite for semantic integration to exchange and discover this vast amount of data on geoscience. Some of the ontologies developed world wide to handle the geoscience data is surveyed. Geoscience domain ontology which played a main role in NASA project and other ontologies in the field of geosciences such as knowledge shared in Earth and planetary ontology, geological hazard ontology, digital geospatial metadata are discussed.

Semantic Web for Earth and Environmental Terminology (SWEET) is a project by NASA for developing domain ontologies to describe earth science data and knowledge. It includes the earth realm, non-living element, living element, physical property, units, numerical entity, temporal entity, spatial entity, phenomena and human activities ontologies [134]. There are 6000 concepts in 200 separate ontologies defined in SWEET [135].

Earth and Planetary ONTology (**EPONT**) is a domain level ontology for sharing data among geoscientists. It uses existing community-accepted high level ontologies such as semantic upper Ontology (SUO): IEEE endorsed, SWEET and **N**orth **A**merican geological **D**ata **M**odel (NADM) [136].

Federal Geographic Data Committee (FGDC) content standard for digital geospatial metadata is developed to describe all possible geospatial data [137].

The Bremen University Semantic Translator for enhanced retrieval combines ontology-based metadata with an ontology-based search. This ontology is used to find the geographic information services for estimating potential storm damage in forests [138].

Geological hazard Ontology is a hierarchical framework, which defines the ontology concept of hazard geology, such as earthquakes, landslip, landslides, debris flow and other hazard [139]. The domain ontology of geoscience is shown in Figure 2.7.

2.2.11 Ontology Developed in the Domain of Human Resources

Human resource management is a crucial factor to enhance the economic development of any organization. Its function consists of tracking personal records of each employee, payroll records etc. Employment services, online job exchange services, human resource advisors, and workforce mobility are of strategic importance for any organization.

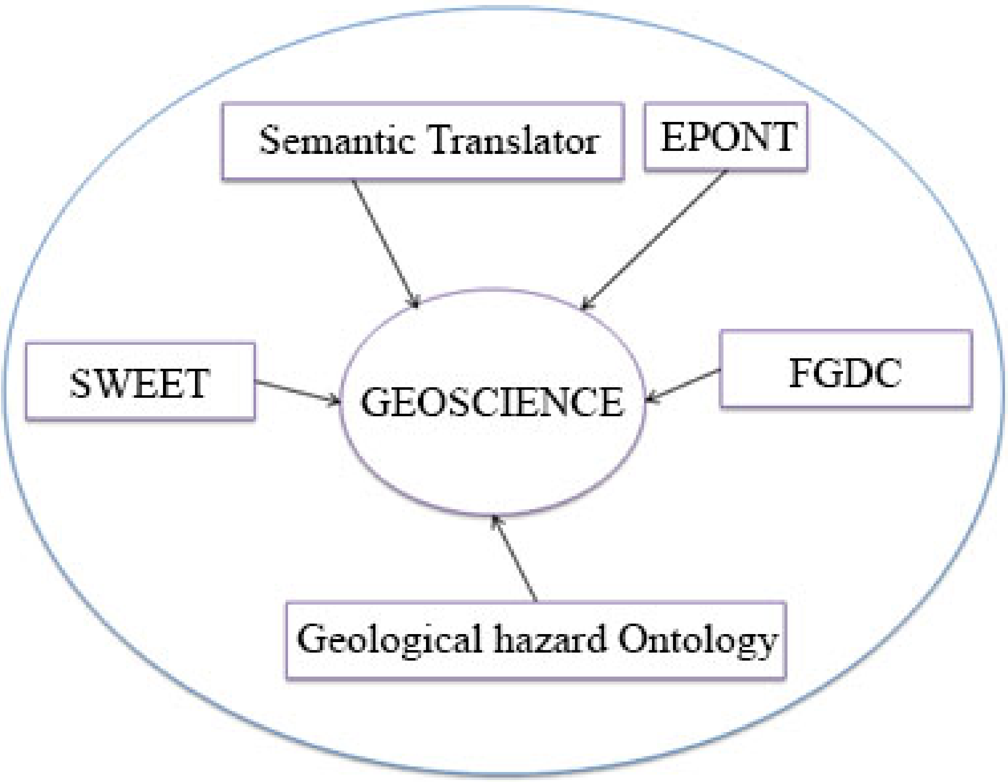


Figure 2.7 Ontology in the field of geoscience domain
SWEET: Semantic Web for Earth and Environmental Terminology; EPONT: Earth and Planetary ONTology; FGDC: Federal Geographic Data Committee.

Ontology has been developed for existing human resources management standards and systems classifications like compensation ontology is based on the ISO 4217. Occupation ontology based on the ISCO-88 has 609 concepts. Education ontology based on the ISCED has 130 concepts. Geography ontology based on the ISO 3166 has 490 concepts. Skill ontology based on European dynamics skill classification has 291 skills. DAML ontology is used for defining the concepts [140].

2.2.12 Ontology Developed in the Domain of Linguistics

Sharing and reusing knowledge management of a domain through ontology can be made only by building a list of structured vocabulary through a language. In terms of resource availability of the data about a particular domain English language is the best suited and elaborately defined. Different disciplines such as agriculture, medicine, automotive etc can be best represented and utilized if the concerned vocabulary is also done in local language of interest to the user community. Several works have already been aimed to improve technological aspects of ontology, like representation of languages and inference mechanism. Thus the use of ontology in the field of natural language processing has become a necessity in exploiting the information for an efficient and useful management of knowledge. *Iban* is one of the divergent Dayak ethnic groups in Sarawak. Sarawak is one of two Malaysian states on the island of Borneo. ***Iban* WordNet** (**IbaWN**) for agricultural domain ontology is developed using Iban as the main language [141]. **SOLAT**-based ontology involves the Al Qur'an, the authentic Hadith, and books that focus on the Shafie's school of thought. It involves the types and characteristics of Solat, hukm, purification such as ghusl, wudu and tayammu. It also includes Qurani verses in Arabic language, images and video. There are 48 concepts, 51 properties and 282 instances [142].

Chinese ancient poetry learning system provides the high knowledge relevance among poems, poet, allusion, genre etc., and presents knowledge according to the user's preference and educational level. The system collects about 270000 of ancient poems and 10000 of allusions [143].

ENGOnto, integrates multiple relevant ontologies for personalized agents to deal with dynamic changes of learner’s learning process. It also interacts between instructor and learner and learning resources in the environment of English language education. This ontology consists of people ontology, language ontology, pedagogy ontology, curriculum ontology and knowledge-point ontology, for individual personalized learning of English [144]. The representation of linguistic domain ontology is shown in Figure 2.8.

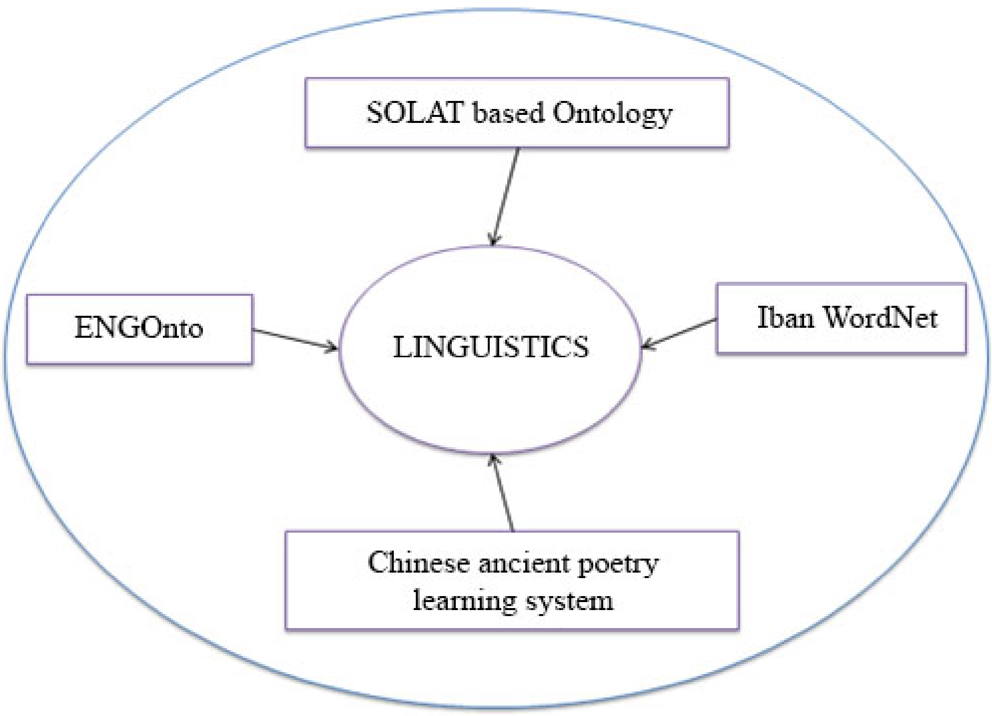


Figure 2.8 Ontology in the field of linguistic domain

2.2.13 Ontology Developed in the Domain of Library

Digital libraries have seen an enormous growth during the last two decades. Ontologies-based schema will enable information resource to enhance technologies, standards and management in digital libraries [145-146].

Ontology based **Chinese Digital Library** resources consist of ontology of bibliographic relations, ontology-based digital library metadata schema, MARC format and thesaurus. It also involves mapping data from MARC to the ontology, and reasoning about the data to establish the relationships [147].

Document Classification System (DCS) consists of four modules: keyword extraction, ontology construction, document classification and document searching. In this system formal concept analysis method is used for the analysis of data. Nearly 525 documents in the area of information management are retrieved from the electronic theses and dissertations system. Amongst these, 360 documents act as the training document and 165 documents for testing purpose [148]. The library domain ontologies are shown in Figure 2.9.

2.2.14 Ontology Developed in the Domain of Marine Sciences

In the naval operations environment the ability to automatically integrate information from multiple sources is a complex task. Ocean researchers have many valuable documents such as observational data and experimental results which help to produce a dynamic, comprehensive and accurate picture of naval conditions. The integration and utilization of these heterogeneous resources are crucial in knowing about ocean eco system and maritime awareness. Few of the ontologies developed in connected with these are discussed

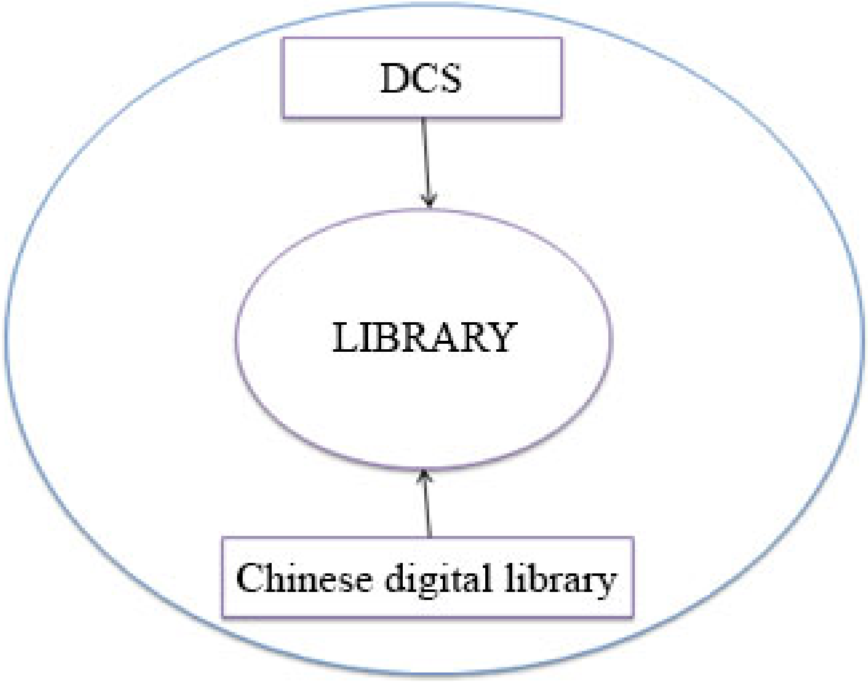


Figure 2.9 Ontology in the field of library domain
DCS: Document Classification System

.

Marine Metadata Interoperability (**MMI**) develops a web based marine metadata vocabulary for the users [149]. **Maritime Domain Awareness** (**MDA**) integrates information from multiple sources in a complex and evolving scenario to produce a dynamic, comprehensive, and accurate picture of the naval operations environment. This would aid in identification of intrusion by suspicious ships [150].

Marine Biology Ontology, has relevant knowledge about oceanic food chain and biodiversity protection. There exist approximately 200 000 kinds of marine life. Marine biology ontology include concepts like halobios, plankton, neuston, nekton, benthos, phytoplankton, zooplankton, bacterioplankton, pleuston, epineuston hypoeneuston, vertebrates, molluscs, crustaceans, zoobenthos, phytobenthos etc. Approximately 160 terms are available in this domain [151]. The marine science ontologies are shown in Figure 2.10.

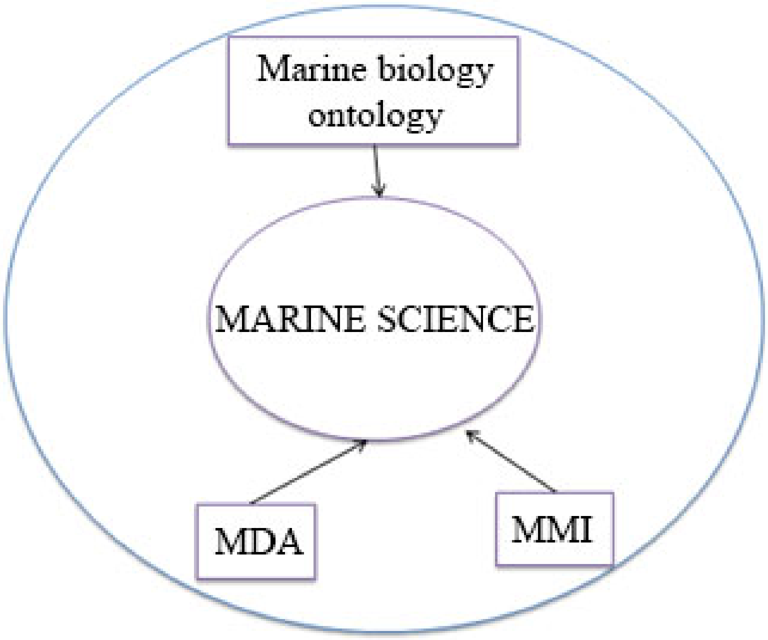


Figure 2.10 Ontology in the field of marine science domain
MMI: Marine Metadata Interoperability; MDA: Maritime Domain Awareness

2.2.15 Ontology Developed in the Domain of Mathematics

Semantic relatedness measures the closeness or likeness between concepts in natural language processing. It is implemented in lexical ontologies such as WordNet. The feature of this method include a unique approach to the weighted edge measure. Each edge is weighted based on applying a concept probability algorithm to a multiset composed of ontology property [152].

Open Mathematical Document (**OMDoc**) is used as an ontology language. It is a content-based markup which focuses on the semantic mathematical formulae. **Learning Style Ontology** (**LSO**) consists of cognitive processing and modality perception. Cognitive processing includes attributes like analytical and global, whereas modality perception is comprised of four attributes like visual, verbal, auditory and tactile-kinesthetic [153]. The domain ontologies of Mathematics is shown in Figure 2.11.

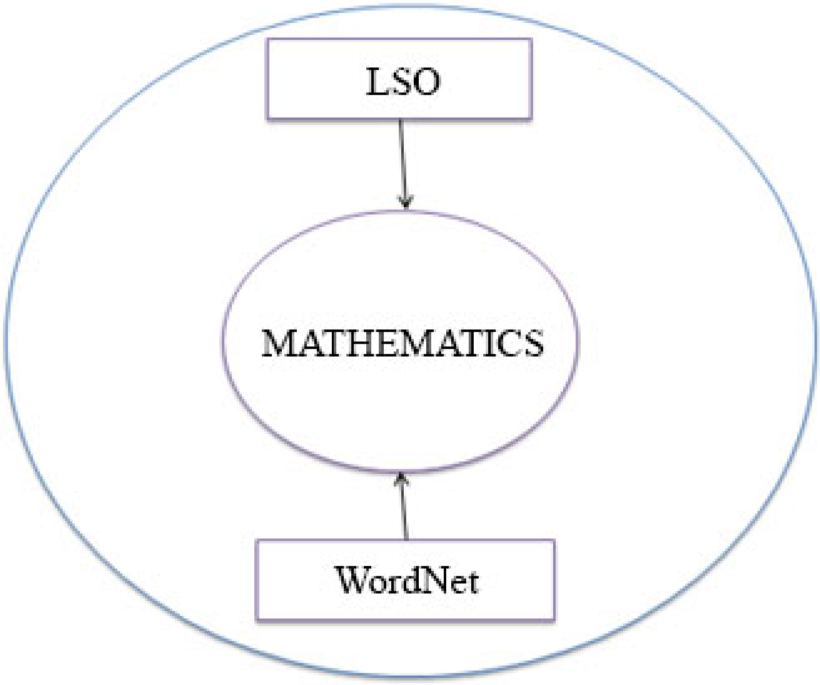


Figure 2.11 Ontology in the field of mathematics domain
LSO: Learning Style Ontology

2.2.16 Ontology Developed in the Domain of Medicine

Medicine is a broad name covering different areas of specialization. Information technology and its widespread availability over the internet lead to proliferation of huge amounts of data related with human health in areas such as gene products and sequences, protein, neuro, heart, cancer, thoracic radiology, drug description, clinical trials, human anatomy etc. In each of the domain specified above, knowledge is subjective by nature and concepts are poorly systematisized. World-wide efforts are on these health sciences domains to bring about a consensus vocabulary and consequently sharing and reuse of knowledge data through specific ontologies. Some of which are discussed here.

The **Gene ontology** contains structured, controlled vocabularies and classifications for several domains of molecular and cellular biology and is freely available in the annotation of genes, gene products and sequences. It has 22684 biological processes, 2987 cellular components and 9375 molecular functions [154].

UBERON is a multi-species metazoan anatomy ontology. This is created to support translational research by allowing comparison of phenotypes across species and provide logical cross-product definitions for gene ontology biological process terms. The current version of the ontology has 2808 terms, 5110 links between terms, 9339 links out to other anatomical ontologies, more than 1643 wikipedia cross-references and has been referenced in 682 gene ontology cross-products [155].

The **Microarray Gene Expression Data Ontology (MGED Ontology)** defines all aspects of a microarray experiment. It also analyzes the data to describe the design of the experiment and array layout, by preparation of the biological sample and the protocols used to hybridize the RNA (Ribonucleic acid). There are 233 classes, 143 properties and 681 individuals defined in this ontology [156].

Mouse Genome Database (MGD), a model for studying human biology and disease, integrates genetic, genomic and phenotypic information about the laboratory mouse. It also includes comprehensive characterization of genes and their functions, standardized descriptions of mouse phenotypes, extensive integration of **D**eoxyribo **N**ucleic **A**cid (DNA) and protein sequence data, normalized representation of genome and genome variant information including comparative data on mammalian genes [157].

<u>C</u>linical <u>B</u>ioinformatics <u>O</u>ntology (CBO), is a semantic network describing clinically significant genomics concepts. It includes concepts appropriate for both molecular diagnostics and cytogenetics [158]. It contains approximately 8155 concepts, 18946 relationships, 4304 facets and 13341 terms.

Thoracic Radiology contains knowledge of anatomy and imaging procedures. In this a total of 138 classes, including radiology orderable procedures, procedure steps, imaging modalities, patient positions, and imaging planes are available. Radiological knowledge was encoded as relationships among these classes [159].

<u>S</u>ystematized <u>N</u>omenclature <u>O</u>f <u>MED</u>icine--<u>C</u>linical <u>T</u>erms (SNOMED-CT) is a terminology system developed by the college of American pathologists. It contains over 344,000 concepts and was formed by restructuring of SNOMED RT (Reference Terminology) and the United Kingdom National Health Service clinical terms [160].

<u>U</u>nified <u>M</u>edical <u>L</u>anguage <u>S</u>ystem (UMLS) is a repository of biomedical vocabularies developed by the US National Library of medicine. The UMLS integrates over two million names for some 900,000 concepts from more than 60 families of biomedical vocabularies, as well as twelve million relations among these concepts [161].

GoMiner is an application that organizes lists of under and over expressed genes from a microarray experiment for biological interpretation in the context of the gene ontology. GoMiner achieves a computational resource that automates the analysis of multiple microarrays and integrates results across all of the microarrays [162].

The **FungalWeb** ontology supports the data integration needs of enzyme biotechnology from inception to product roll out. It serve as a knowledge base for decision support, to link fungal species with enzymes, enzyme substrates, enzyme classifications, enzyme modifications, enzyme retail and applications [163]. It contains 3667 concepts, 12686 instances and 157 properties [164].

Protein Mutation Impact Ontology conceptualizes impacts and the mutations associated with them. To design the mutation impact ontology, information text elements, biological entities and entity relations are also required. OWL format is used to define the relations between these entities [165].

Personalized Information Platform for health and life Services (PIPS) deals with medical knowledge, food and nutrition knowledge, about patients, their clinical records, products and treatments. Food ontology deals with the development process that describes 177 classes, 53 properties and 632 instances [166].

Neuro-pediatric Physiotherapy is an area that includes diagnosis, treatment and evaluation of babies by the physiotherapist in order to observe the progress of treatment. Neuro-pediatric ontology is composed of 100 classes and subclasses, 30 properties and 200 axioms [167].

Cardiovascular Medicine Ontology in domain of **Mechanical Circulatory Support Systems** (**MCSS**) is designed to avoid lack of uniformity in the information available in the field. There are 30 different types existing in this domain [168].

Cystic Fibrosis is a subset extracted from a large **Medical Literature Analysis and Retrieval System Online (MEDLINE)** collection. There are 1239 files with 821 concepts and the average number of concepts assigned to a document is around 3 [169].

The **National Center for Biomedical Ontology** (**NCBO**), California, maintains a bioPortal, an open library of more than 200 ontologies in biomedicine. The aim of this portal is to provide support to a researcher to browse and analyze the information stored in these diverse resources [170].

The **National Cancer Institute (NCI) Thésaurus** developed in U.S, is a public domain description logic-based terminology for bioinformatics **cancer Common Ontologic Representation Environment (caCORE)** distribution. It contains 26,000 concepts and 71,000 terms divided among 24 taxonomies. The final OWL ontology is made up of approximately 450,000 triples in a file that is over 33 MB [171].

The **Foundational Model of Anatomy Ontology (FMA**) is a domain ontology that represents a coherent body of explicit declarative knowledge about human anatomy. It is a frame-based ontology and there are 148 relationship types, 70,000 anatomical concepts interrelated by over 580,000 relationship instances [172]. The domain ontology of medicine is shown in Figure 2.12.

2.2.17 Ontology Developed in Military Domain

Information age has brought about a dramatic change in the way in which the military activities are organized. A suitable knowledge infrastructure in the military domain would have a strong premise of transforming information superiority, so that the combat power is improved. Ontologies developed in this area are discussed.

Military intelligence domain Ontology is referred as the ontology structure in HowNet and WordNet, and it stores the characteristics of the military information [173].

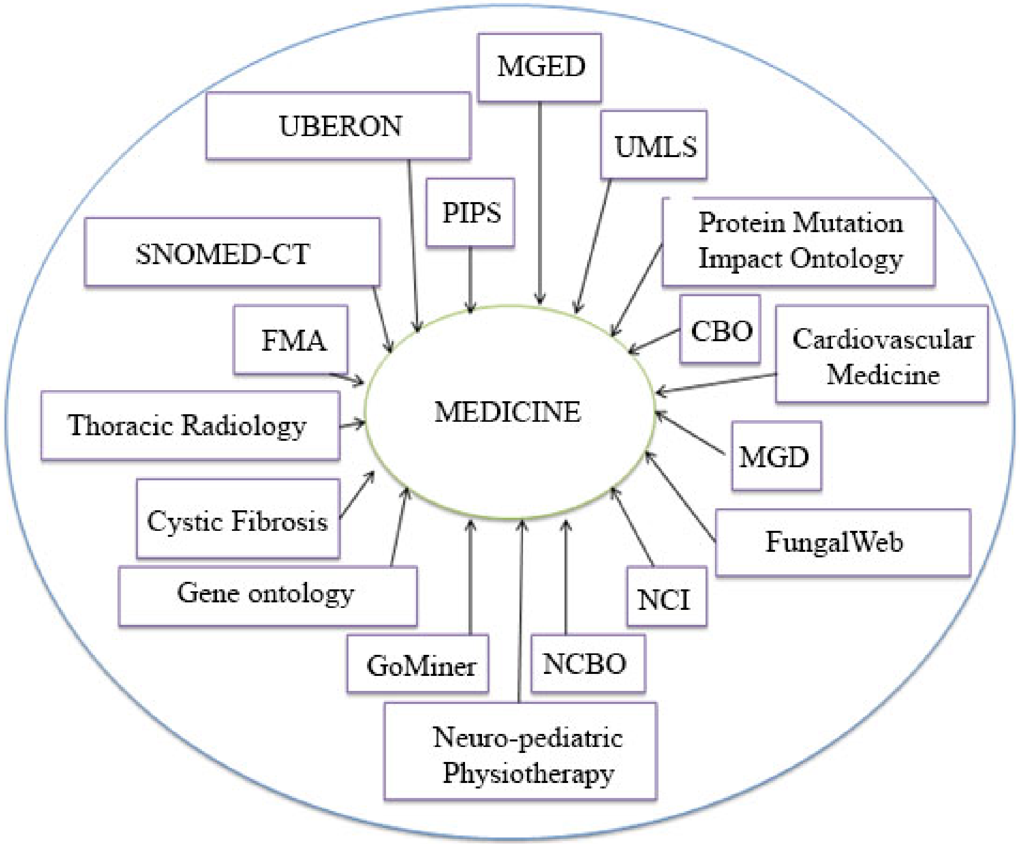


Figure 2.12 Ontology in the field of medicine domain
MGED: Microarray Gene Expression Data: MGD: Mouse Genome Database; CBO: Clinical Bioinformatics Ontology; SNOMED-CT: Systematized Nomenclature of Medicine-Clinical Terms; UMLS: Unified Medical Language System; PIPS: Personalized Information Platform for health and life Services; NCBO: National Center for Biomedical Ontology; NCI: National Cancer Institute; FMA Foundational Model of Anatomy Ontology.

5W1H-based conceptual modeling framework for domain ontology is proposed, which is used to analyze domain concepts and relations from six aspects like who, when, where, what, why and how. According to this framework, the conceptual model of **Science and Technology Project Ontology (STPO)** in science and technology domain is designed. From the analysis, real world model is designed using the 5W1H conceptual modeling framework by mapping the class model in the object-oriented method [174].

Collaboration of **Military Domain Ontology Construction Approach (MDOCA)**, **S**ituation **O**ntology **C**onstruction **A**pproach (SOCA) and **M**ilitary **R**ule **O**ntology **C**onstruction **A**pproach (MROCA) are used to construct **S**ituation **O**ntology (SO) and **M**ilitary **R**ule **O**ntology (MRO) [175]. The domain ontology of military is shown in the Figure 2.13**.**

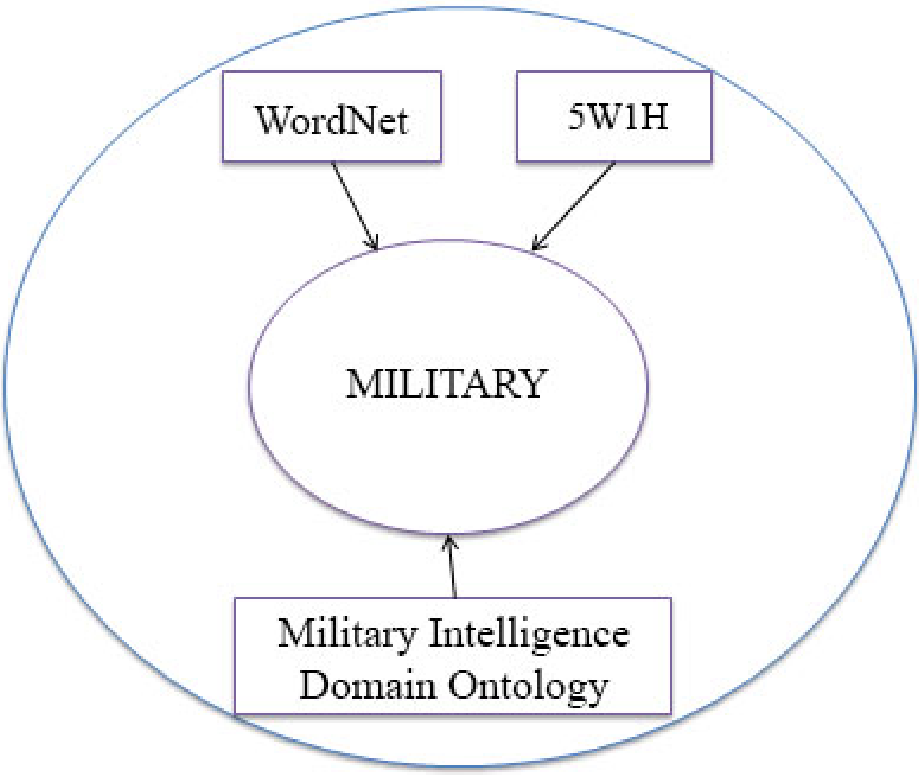


Figure 2.13 Ontology in the field of military domain

2.2.18 Ontology Developed in the Domain of Nuclear Weapons

Nuclear weapon non-proliferation organizations have a stupendous task in the nuclear activities, materials, facilities, equipment etc. Ontology based efforts to a effective safeguard activity have attracted attention in the nuclear field. In the domain of **nuclear weapons**, radionuclide concepts are defined. It has sub concepts like data products, laboratory managers, programs, facilities and data managers [57].

2.2.19 Ontology Developed in the Domain of News

In the **news domain**, information extraction doesn't rely on the page structure but the result of this information extraction cooperates with the pre-defined ontology. The web pages downloaded with the use of .NET's web browser component are formed into a **DO**cument **M**odeling (DOM) tree. Ontology of news domain consists of following sub concepts like navigation page, seed page, content page, navigation page marker path, content page marker path, title, time, picture and content [176].

2.2.20 Ontology Developed in the Domain of Power Plants

Power plant safety depends on several factors whose characteristics influence the reliability and accuracy of the assessment. Identification and nature of occurrence of equipment fault in power plant is complicated. Hence ontology based knowledge management systems act as a tool for fault diagnostic maintenance system for power plants.

A Safety Assessment Management Information System for Power Plants is developed on a client server model. It has been used in power plant of Datang Group Corporation in China and reported to be satisfactory. Knowledge of equipment fault was captured by knowledge transformation, collecting original literature and data, identifying relations among basic glossaries which contain complex information, determining the rules [177].

Steam turbine ontology is created by integrating and merging with existing databases. It enables sharing the knowledge through a shared ontology for the maintenance of a steam turbine [178]. The power plant ontology is shown in Figure 2.14.

2.2.21 Ontology Developed in the Domain of Transport

Manufacturing systems are faced with growing complexity and depends crucially on the transport systems. The need for flexibility and agility in this domain needs structured database were ontology gains importance.

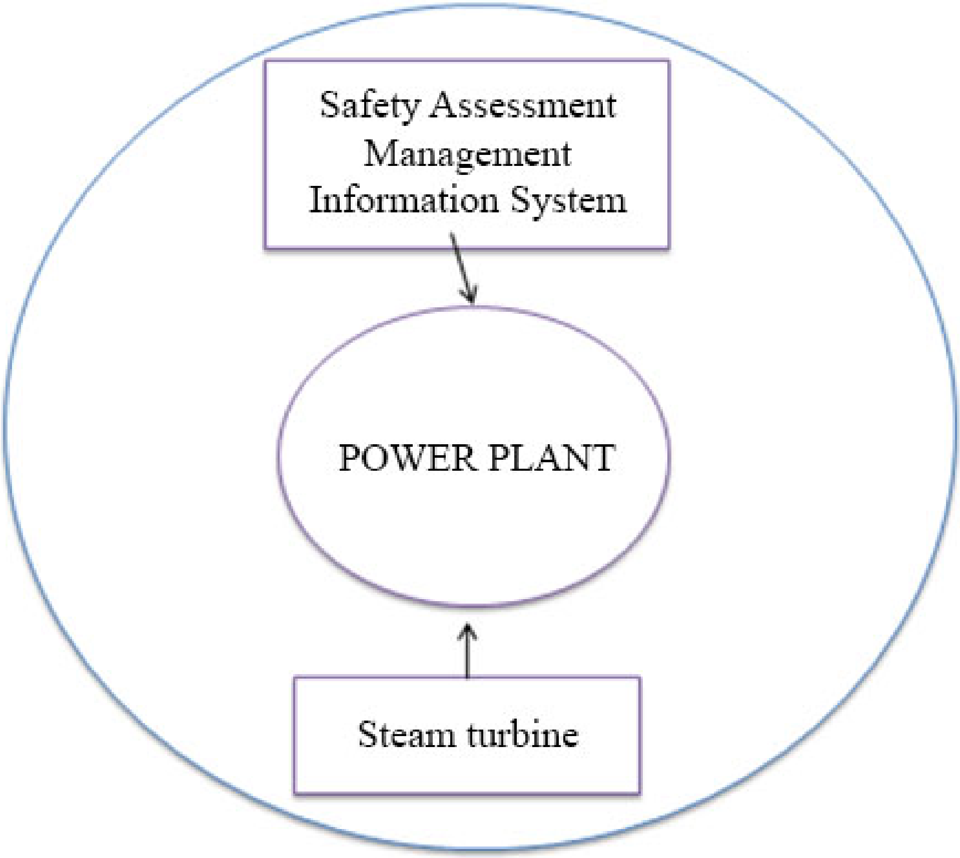


Figure 2.14 Ontology in the field of power plant domain

The main components of the pallet transfer system are **conveyor belts** which deliver items from one place to another, index stations, identification units (RFID) for identification of passing pallet units, and intersection units. This is used in ontology to represent locations whose attributes provide the details of locations reachable by it [179].

2.3 Comparison of Application Domain

The list of the domains surveyed, development tools, ontology language used, query language, language support, ontology language, programming language, database, reasoner are summarized in Table 2.1. It is seen that for programming, *JAVA* language is preferred. *Perl or .Net* or *C#* being other language used. From the survey, Protégée integrated development tool is found to be used by most of the application domains for developing ontology. For query language, SPARQL is used by majority of the application and in some application domains RDQL, Mouse Genome Informatics Batch Query Tool, new **R**acer **Q**uery **L**anguage (nRQL) etc are used. Pellet or FaCT++ or Racer is the widely used reasoners in the domain. For storage, *DB2 or MySQL or ORACLE* is mostly utilized in the domain. It is seen that English language used extensively in applications of domain ontology. Agrovoc supports Arabic, Chinese, Czech, English, French, German, Hindi, Hungarian, Italian, Japanese, Korean, Lao, Persian, Polish, Portuguese, Russian, Slovak, Spanish, Thai, Turkish languages. Ontology defined in the field of Chemistry domain supports Spanish, Sahayika - English and Bengali language, IbanWordNet used Iban, Chinese ancient poetry learning system and Chinese Digital Library used Chinese language, SymOntoX supports multi languages.

Table 2.1 List of tools, query language, supported language, ontology language, programming language, database, reasoner used by ontology application domain

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Agriculture	Agrovoc	Protégé	SPARQL	RDF / SKOS-XL OWL	Not Specified	Not Specified	Not Specified
	Food Safety Semantic Retrieval System	Not Specified	Not Specified	Not Specified	JDK1.6 using My Eclipse 7.5	MySQL	Not Specified
	Agriculture Literature Retrieval System	Protégé	Not Specified	XML, RDF, OWL	Jena	Not Specified	Not Specified
	Citrus Water and Nutrient Management System	Not Specified	SPARQL	OWL	JAVA	Not Specified	Not Specified
	OntoSim-Sugarcane	Not Specified	Not Specified	Not Specified	JAVA	Not Specified	Not Specified
Aviation	Aviation	Not Specified	Not Specified	Description Logic	Jena	Not Specified	Pellet
Biology	Plant Ontology database	Not Specified	Not Specified	OWL	Perl	MySQL	Not Specified
	Plants domain Ontology	Protégé	Not Specified	OWL	Not Specified	Not Specified	Not Specified
	Transparent Access to Multiple Bioinformatics Information Sources	Not Specified	GRAIL	Description Logic	JAVA applet	Collection Programming Language for Multi database support	FaCT
	BRaunschweig ENzyme DAtabase	Not Specified	Not Specified	Not Specified	SOAP based web service API	Not Specified	Not Specified

88

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Chemistry	Chemistry Domain	Methontology	Not Specified	Not Specified	Visual Basic	Not Specified	Not Specified
	Chemical Entities of Biological Interest	Not Specified	Not Specified	OBO	JAVA	Oracle	Not Specified
Civil	Healthy Housing	Protégé	Not Specified	OWL	MADRE	Not Specified	Not Specified
	FinnONTO	OntoViews	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
Computer Science	Innovation and Technology in Computer Science Education	Not Specified	Not Specified	OWL	Not Specified	Not Specified	Not Specified
	Standard Ontology for Ubiquitous and Pervasive Applications	Not Specified	Not Specified	OWL	Not Specified	Not Specified	RACER and FaCT
	Software System Ontology	Protégé	Not Specified	Not Specified	JAVA	Not Specified	Not Specified
	Software Product Management	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	KAON
	Open Mind Indoor Common Sense	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Video Indexing and Retrieval	Not Specified	Query By Example	Not Specified	Not Specified	Not Specified	Not Specified
	Software Engineering Lifecycle	Not Specified	SPARQL	RDF, OWL, UML	Not Specified	Not Specified	Not Specified

89

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Computer Science (continued)	Telecommunication Management Network Model	Protégé	Not Specified	OWL, DL	JAVA	Not Specified	Not Specified
	Image Classification Using Neural Networks	Not Specified	Not Specified	directed acyclic graph	Not Specified	Not Specified	Not Specified
e-Business	Symbolic Ontology XML-based	Not Specified	Not Specified	RDF, OWL, SHOE and XOL	JAVA	Xindice	Not Specified
Education	Sharable Content Object Reference Model	Protégé, OntoEdit	Not Specified	RDF, RDFS	JAVA SCRIPT	Not Specified	Not Specified
	Topic Maps for e-Learning	Topic Map Editor and viewer	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Web-Based Educational Systems	Not Specified	Not Specified	XML, KIF, ACL	Not Specified	Not Specified	Not Specified
	Ontologies for Education	Not Specified	Not Specified	OWL, DL, XML, RDF, XTM	Not Specified	Not Specified	Not Specified
	Ontologies for the Use of digital learning Resources and semantic Annotations on Line	Protégé	Not Specified	OWL	Not Specified	Not Specified	Not Specified
	Learning Resource	Not Specified	Not Specified	RDF	Jena	DB2 / My SQL / ORACLE	Not Specified

90

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Education (continued)	Network Education Resource Library	Not Specified	Not Specified	OWL	JSP and Java bean language	DB2 / My SQL / ORACLE	Not Specified
	e-learning systems in higher education	Protégé	Not Specified	OWL	Not Specified	Not Specified	Not Specified
	Secondary Vocational Education	Protégé	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Ontologie_US_ENCG	Protégé	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Virtual Lab Ontology	Protégé	SPARQL	DAML+ OIL,OWL	Not Specified	Not Specified	Pellet or FaCT++
	Economic Ontologies	Protégé	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Cultural Artefacts in Education	CAE_L Ontology Framework	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Sahayika	OntoEdit	Not Specified	Not Specified	JAVA Beans	ORACLE/ MYSQL	Not Specified
	Remote education	OntoLearning	RDQL	RDF	Not Specified	Not Specified	Not Specified
	Information and Communication Technologies	Protégé	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
Electronics	Suggested Upper Merged Ontology	Protégé	RDQL	RDF	Not Specified	Not Specified	Not Specified
Geoscience	Semantic Web for Earth and Environmental	Not Specified	Not Specified	DAML+ OIL, OWL	JAVA and Perl	Postgres	Not Specified

91

92

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Geoscience (continued)	Earth and Planetary ONTology	Not Specified	Not Specified	OWL	.Net and JAVA	PostgreSQL	Not Specified
	Federal Geographic Data Committee	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	The Bremen University Semantic Translator	BUSTER	Not Specified	XML, DL, RDF, OWL	Not Specified	Not Specified	RACER
	Geological hazard	Protégé	Not Specified	OWL	Not Specified	Not Specified	Not Specified
Human Resources	Human Resources Management	Methontology	Not Specified	DAML	Not Specified	Not Specified	Not Specified
Linguistics	IbanWordNet	Methontology	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	SOLAT	TopBraid	SPARQL	RDF, OWL, DAML+ OIL	Not Specified	Not Specified	Not Specified
	Chinese ancient poetry learning system	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
Library	Chinese Digital Library	Not Specified	Not Specified	RDF	Not Specified	Not Specified	Not Specified
	Document Classification System	Not Specified	Not Specified	XML	Not Specified	Not Specified	Not Specified
Marine Science Marine	Marine Metadata Interoperability	Protégé and SWOOP	Not Specified	OWL	Not Specified	Not Specified	Pellet
	Maritime Domain Awareness	Not Specified	Situation Specific Bayesian Network	OWL,UML	Not Specified	Not Specified	Multi Entity Bayesian network reasoner

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Science (continued)	Marine Biology Ontology	HOZO	Not Specified	RDF,DAML+ OIL,OWL	Not Specified	Not Specified	Not Specified
Mathematics	Semantic relatedness	OntoNL	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	Open Mathematical Document	Not Specified	Not Specified	RDF, OWL	Not Specified	Not Specified	Not Specified
Medicine	Gene ontology	AmiGO browser	Not Specified	Not Specified	Perl ,JAVA	MySQL	Not Specified
	Microarray Gene Expression Data Ontology	OilEd	Not Specified	RDF, DAML+ OIL, OWL	Perl ,JAVA	Not Specified	Not Specified
	Mouse Genome Database	Mouse Gbrowse, Mouse BLAST	Mouse Genome Informatics Batch Query Tool	Not Specified	Not Specified	Sybase	Not Specified
	Clinical Bioinformatics Ontology	Not Specified	Not Specified	DL	Not Specified	Not Specified	Not Specified
	Unified Medical Language System	Not Specified	Not Specified	Not Specified	JAVA	SQL	Not Specified
	GoMiner	AmiGO, DAG-Edit	Not Specified	Not Specified	JAVA	JDBC, MySQL	RACER
	FungalWeb	Protégé	new Racer Query Language	DL,OWL	Not Specified	Not Specified	RACER

93

94

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
Medicine (continued)	Protein Mutation Impact Ontology	Not Specified	SPARQL	OWL	Not Specified	Not Specified	Not Specified
	Personalized Information Platform for health and life Services	Protégé	Not Specified	DL,OWL	Not Specified	Not Specified	RACER, Pellet
	Neuro-pediatric Physiotherapy	Methontology, On-To-Knowledge	Not Specified	DL,OWL	Not Specified	Not Specified	RACER
	Cardiovascular Medicine Ontology	Protégé	Not Specified	Not Specified	JAVA	Not Specified	Not Specified
	Cystic Fibrosis	Webocrat	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified
	National Center for Biomedical Ontology	Not Specified	Not Specified	Not Specified	Not Specified	MySQL	Not Specified
	National Cancer Institute	Not Specified	Not Specified	RDF,OWL	Not Specified	Not Specified	Not Specified
	Foundational Model of Anatomy Ontology	Protégé	Not Specified	Not Specified	Not Specified	MySQL	Not Specified
Military	5W1H-based conceptual modeling framework	Protégé	Not Specified	OWL,UML	Not Specified	Not Specified	Not Specified
	Military Domain Ontology Construction Approach	Protégé	Not Specified	XOL, SHOE, UML, RDFS, OIL,SWRL	Not Specified	Not Specified	FaCT and RACER
Nuclear Weapons	Nuclear Weapons	Protégé	Not Specified	RDF,OWL,	JAVA	Not Specified	Not Specified

Domain Name	Ontology Name	Integrated Development Tools	Query Language	Ontology Language	Programming Language	Database	Reasoner
News	News Domain	Not Specified	Not Specified	XML DOM	.Net,C#	Not Specified	Not Specified
Power Plants	Safety Assessment Management Information System for Power Plants	Protégé	Not Specified	OWL,UML	Jena	SQLServer	Not Specified
	Steam turbine	Protégé	Not Specified	RDF,OWL	JAVA	JDBC	JESS, RACER
Transport	Conveyor Belts	Protégé	Not Specified	OWL	Not Specified	Not Specified	JESS

95

In Table 2.2, algorithms and protocols used in application domains are summarized. It can be seen in several domains complete details are not made available.

Table 2.2 Algorithm and protocols used in application domains

Domain Names	**Algorithm and Protocols**
Standard Ontology for Ubiquitous and Pervasive Applications	CoBrA
Video Indexing and Retrieval	Block Matching Algorithm
Software Engineering Lifecycle	Ontology Definition Metamodel, Dhurv
Image Classification Using Neural Networks	Neural Networks
Web-Based Educational Systems	SOAP
Ontologies for Education	SOAP
Learning Resource	SOAP
Marine Biology Ontology	OASIS(Ontology_Aware System for Information Searching)
Systematized Nomenclature of Medicine--Clinical Terms	Term modelling, LinK Classifier
Nuclear Weapons	Semantic Homology Model
Conveyor Belts	Pallet Agent

From this survey it is seen that extensive work has been reported in the medical domain. In this domain work has been done in defining ontology from micro organism to macro organism. The ontology development covers concepts related to database of patients and diseases with multilingual support. In the domain of education, efforts have been directed towards creating systems that aid learning process along with formal teaching. From kindergarten to higher education, ontology has been defined to achieve person independent knowledge based system.

In the computer domain, the ontology development has evolved considerably, in defining concepts relating to hardware and software systems, image processing, videos and audios and neural networks. The survey indicated that in some domains like power plants and nuclear energy, the application of ontology has been rather limited. This study helped in getting an overall idea of domain concepts, implementation and tools required for the development of the domain [180].

In the next section, the available tools for ontology development like Protégé, SWOOP and OntoStudio, TopBraid for ontology development are compared.

2.4 Literature Survey on Ontology Editor Tools

2.4.1 OntoStudio Tool

OntoStudio is a commercially available engineering work bench and can be installed at ease. It consists of classes, properties, rules, queries and mapping. Provision to define properties, range, minimum and maximum values are available. Apart from this ObjectLogic source, view of ontology shows, classes and properties are also available. OWL, RDF and ObjectLogic formats are supported in this editor [181]. While creating a project, users have a choice to select the storage type like file based internal repositories or collaboration server. It has the provision to get data from database and also facilitates with chart and report generation. Extraction of the report data in different formats like xml, html, doc, ppt, pdf, postscript is also possible. It is user friendly for connecting SQL database with a query builder SQL facility, so that the queries can be created easily.

OntoBroker tool is an added package available along with OntoStudio. It is used for establishing network connection and for setting up the startup configuration. In the OntoBroker ontology construction, tools are available to convert OWL to object logic, RDF/XML, Turtle, N-triples, n3, ANSI thesaurus etc. OntoBroker consists of ObjectLogic query and SPARQL for querying the graph. SPARQL is a **W**orld **W**ide **W**eb **C**onsortium (W3C) standard. It is a query language for RDF. OntoBroker only implements a subset of SPARQL. Inturn SPARQL only supports a subset of the ObjectLogic query features. One important feature of OntoBroker is that it can interface with programming languages like *.net*, *java*. It is also a multilingual editor and languages like French, English, and German can be used for defining equivalent synonyms.

2.4.2 Protégé Tool

Protégé runs on a variety of platforms like Windows, Linux, MacOSX, Sun, Solaris, HPUX and IBM. It allows the users to customize interface extensions, incorporates the open knowledge-base connectivity knowledge model and there by interacts with the standard storage formats like relational databases, XML and RDF [182]. OWL extends RDFS to allow for the expression of complex relationships between different RDFS classes and of more precise constraints on specific classes and properties (for example, cardinality relations, equality, enumerated classes, characteristics of properties etc).

This tool allows formats like RDF/XML, OWL/XML, OWL Functional syntax, Manchester OWL syntax, OBO 1.2 flat file, KRSS2 syntax, Latex and Turtle (Terse RDF Triple Language). Protégé uses reasoners like FACT++, Hermit and RACER to fix the inconsistency. The main services offered by reasoner are to test whether or not one class is a subclass of another class. By performing such tests on all classes, it is possible for a reasoner to compute the inferred ontology class hierarchy. Another reasoning service is “consistency checking” which is to check whether it is possible or not for the class to have any instance. The ontology can be sent to the reasoner automatically, to compute the classification hierarchy, and also for checking the logical consistency of the ontology. In Protégé, the manually constructed class hierarchy is called the asserted hierarchy and if automatically computed by the reasoner it is called inferred hierarchy. Maintaining relational database at the backend enables Protégé to process Ontologies which are too large to reside in the memory. Finally, through Protégé’s plug-in mechanism, user can generate their own custom import or export plug-ins to work with custom or specialized formats [183]. OWLGrEd is a unified modeling language style graphical ontology editor for OWL and allow graphical ontology exploration and development including interoperability with Protégé [184].

2.4.3 SWOOP Tool

Semantic web ontology editor -SWOOP is a tool for creating, editing, and debugging OWL and also acts as a web ontology browser and editor. It is developed and released by the MIND lab at University of Maryland, USA in the year 2007, and is currently an open source project with contributors from all over the world. It supports PELLET, and RDFS-like reasoner. Advanced features of SWOOP include, possibility of repairing the ontology and recalculating the rank, splitting the ontology, pellet query, version control and also partitioning the ontology automatically [185]. It is developed using java and can be executed both in Linux and Windows operating system. Multimedia markup plug-ins that facilitate image and video annotations are supported [186]. While Protégé supports loading of OWL files that are developed in SWOOP, the files developed in Protégé cannot be loaded in SWOOP.

2.4.4 TopBraid Tool

TopBraid is a collection of integrated semantic solution product. All the components of the TopBraid work within an evolving open architecture platform and are implemented adhering to W3C semantic web standards. This software is released in 2011 and is available in three forms like free, standard and maestro [60]. It is found that free version has several limitations. Maestro version has many advanced features like its own internal web server for testing application development. It has a flexible and extensible framework with an application program interface for developing semantic client/server or browser-based applications. Some of the concepts are similar to Protégé, like generation of schema based forms for data acquisition. Many features are available like graphical editor which can be used for designing the ontology easily. Moreover cloning of classes and subclasses is also supported.

2.4.5 Comparison of the Integrated Development Environment Tools

A comparison of characteristics of IDE tools like Protégé, OntoStudio, TopBraid and SWOOP is summarized in Table 2.3. All the four IDEs can run both in Windows and Linux operating systems. Protégé runs in Mainframe, Apple and Sun system, indicating its platform independence support.

Table 2.3 Comparisons of Integrated development environment (IDE) tools Protégé, OntoStudio, TopBraid and SWOOP

Characteristics	Protégé	OntoStudio	TopBraid	SWOOP
Supporting platform	Windows, Linux, MacOSX, Sun, Solaris, HPUX, IBM	Windows,Windows Server 2003, SUSE Linux 10.x	Windows, Linux, Macintosh	Windows, Linux
Supporting File Formats	RDF/XML, OWL/XML, OWL Functional syntax, Manchester OWL syntax, OBO 1.2 flat file, KRSS2 syntax, Latex and Turtle (Terse RDF Triple Language)	OWL, RDF, F-logic ObjectLogic	OWL, RDF, turtle, n-triple, xml	OWL, RDF, XML, text, SWOOP ontology object files
Graphical representation	OntoGraf,OwlViz, OWLGrEd for UML	Supported	Supported	Not available
Query Supports	DL query	SPARQL, ObjectLogic query, Query Builder	SPARQL, SPIN are supported	Pellet Query
Reasoner	Fact,Fact++,Hermit ,Pellet	Pellet	Pellet, Jena, TopSPIN, Jena built in reasoned, SwiftOWLIM	Pellet, Racer, Fact
Programming languages interface	PROTÉGÉ API, Protégé Script Tab	Interface with .net, Java program	Adobe Flex, HTML and Javascript: used with SPARQL Web Pages (SWP), Java Web service API is implemented using SPARQLMotion	Interface with Java program

101

Characteristics	Protégé	OntoStudio	TopBraid	SWOOP
Conversion from one form to another	RDF/XML,OWL/XML	ObjectLogic to F_logic, ontology to object logic, RDF/XML, Turtle, N-triples, n3, ANSI thesaurus	RDF, spread sheet, share point, Graphs are supported	Not available
Availability	Freeware	Licensed version	Standard, Maestro versions are available	Freeware
Multi user support	Supported	Not Supported	Supported	Not Supported
Consistency check	Supported	Supported	Supported	Supported
Language supported to define synonyms	English	English, French and German	English	English
RIF (Rule Interchange Format)	Not Supported	Supported	Supported Jena rule engine	Not Supported
Report generation support	Not Available	Available.Using Business Intelligent Reporting Tool	Available	Not available
Chart Generation support	Not Available	Available	Available	Not available
Graphical editor support	Not Available	Available	Available	Not available
Implemented using	Java	Eclipse IDE	Eclipse, Jena, Apache	Java
Ontology version comparison facility	Available	Not available	Available	Available

102

Characteristics	Protégé	OntoStudio	TopBraid	SWOOP
Developed and Supported by	Stanford university	OntoPrise	TopQuadrant	MINDSWAP research group
Database support	Not Supported	SQLServer ,Oracle, DB2, JDBC	Supports Oracle database	Not Supported
Ontology difference	PromptDiff plugin-to compare the ontology version	Not Available	Not Available	Not Available
Split ontology	Not Available	Available	Available	Available
UML representation	Using OWLGrEd plug-ins	Supported	Supported	Not Supported
Root node	Thing	owl:Thing	owl:Thing	owl:Thing
Version control	Not Availbale	Available	Available	Available

103

Supporting file format implies that languages in which ontology data are *read in* and *write out*. In Protégé *read in* and *write out* file formats are starting from flat file, RDF, OWL, Latex to Turtle format whereas in others OWL, RDF, F-logic and objectlogic are supported. OntoGraf, OwlViz plug in are used in Protégé to view the graphical format of the ontology and also to get image of the ontology in jpg, png, gif format. Other integrated development environment tools use their own built in software for graphical representation. Any application program developed using *Java* language supports interfacing with OntoStudio, TopBraid and SWOOP tools. Protégé tool is having its own **A**pplication **P**rogram **I**nterface (API) to integrate with any given application. Multiple language support is available in OntoStudio. Inference engine is used to recognize any ambiguity about the meaning of the terminology used in the ontology [187]. Pellet reasoner is supported by all the IDE. Many reasoners use first-order predicate logic to perform reasoning. One of the difficulties of IDE is that the OWL file created in one editor tool is not compatible with other editor tools even though the file types are the same.

2.5 Summary

A survey on usage of ontology in different domains based on the available literature is given. It is seen from the survey, that ontology has been applied in several domains encompassing areas like agriculture, education, medicine, defense, aviation, computer science and linguistics etc.

Among the four tools available for development of ontology described above TopBraid is user-friendly. OntoStudio is a commercially available tool with additional features like multi language support, database interface, reports and charts generation and has more flexibility. Dynamic help facility, syntactic validation and cloning concepts are supported in TopBriad. SWOOP and Protégé are open source editors. Protégé tool is also flexible, easy to use, with open editor and widely used by the semantic group. Based on the above analysis, Protégé has been employed for ontology development.

In order to develop the requisite ontology for the Fast Breeder Test Reactor, a semantic web based knowledge representation is designed and christened as **K**nowledge **M**anagement for **N**uclear **R**eactor (KMNuR) portal by collecting the knowledge from the nuclear domain experts. This is discussed in details in Chapter 3.

CHAPTER 3

SEMANTIC WEB BASED KNOWLEDGE MANAGEMENT PORTAL FOR NUCLEAR REACTOR DOMAIN

This chapter covers the system architecture and implementation of Knowledge Management for Nuclear Reactor (KMNuR). It also explains the methodology adopted for developing the components of portal for the FBTR. The RDF, OWL and UML representation are done for each system and parameter of FBTR.

3.1 Need for Knowledge Management Portal

One of the objectives of the knowledge management portal is to capture the knowledge existing in a domain and preserve it for the future. To achieve this, efforts are required to employ a suitable mechanism for obtaining and integrating the collective knowledge in the domain. This knowledge is obtained from design, operational, maintenance, safety, quality assurance personnel etc. The content of the portal is organized in such way that the adopted ontology would be available to users to make an effective use of the same. Considering the life cycle of nuclear plant, the creation of knowledge base not only aids to maintain the existing nuclear reactors but also in the design, construction, commissioning, operation and decommissioning of reactors in general. In the present work, a semantic web based knowledge representation is designed and christened as **K**nowledge **M**anagement for **Nu**clear **R**eactor (KMNuR) portal (refer to as KMNuR in the rest of the thesis) [188]. It involves the creation of knowledge base by collecting all the requisite knowledge available about the nuclear reactor. KMNuR is comprised of knowledge represented in different semantic formats like RDF, OWL and **U**nified **M**odeling **L**anguage (UML) to enable the web crawler to share and reuse nuclear knowledge [189]. The portal is developed for **F**ast **B**reeder **T**est **R**eactor (FBTR) (refer to as FBTR in the rest of the thesis) at Kalpakkam and a brief description about FBTR is given in section 1.10.2.3. The knowledge relating to FBTR has been represented in the portal.

3.2 System Architecture of KMNuR Web Portal

The system architecture of KMNuR is represented in Figure 3.1. Protégé tool is used for developing the requisite ontology. The KMNuR web portal is a client/server architecture having user interface in the frontend and database at the backend. The frontend of the user interface developed using *Java* web application allows the user to get the requisite knowledge about the FBTR from *MySQL* database at the backend. *Net Bean IDE* is used for developing the *Java* web based application and *GlassFish* server for publishing the web application. The knowledge pertaining to FBTR system is obtained from data sources like journals, books, internal reports, existing data in IAEA, open archives and by taking inputs from experts [190]. The data is stored in the *MySQL* database in formats like RDF, OWL, UML and graph. Users can select the particular format to retrieve the knowledge about the nuclear system of their choice.

Important basis of RDF, OWL, UML are summarized below. The most important requirement is to share or reuse the knowledge, which is achieved by having universal method of expression, uniform format for data-exchange, syntactic and semantic interoperability. Universal method of expression and syntactic interoperability can be achieved by employing XML. However, XML has limitation when it comes to semantic interoperability. RDF is defined as syntax independent and its advantage over XML: The object-attribute structure provides natural semantic units as each objects are independent entities. OWL is standardized and broadly accepted language for semantic web. OWL is recommended format by OAEI (Ontology Alignment Evaluation Initiative) for ontology matching. For the representation of ontology UML is a good tool which represents static knowledge using class diagram and dynamic knowledge using UML dynamic and state diagram. UML representation is easily discernible model for building systems in object-oriented programming languages and has been widely adopted by the software engineering community. Hence in the development of portal we have included RDF, OWL and UML.

The methodology adopted for developing the components of portal for the FBTR domain involves the following steps [191].

i. Ontology is created by defining the classes, subclasses, object properties, data properties for each system by receiving the knowledge from available sources.
ii. The created ontology is verified and validated by checking the inferred knowledge to remove any logical inconsistency. The verification is carried out by starting any one of the reasoners like Fact++, Hermit and Pellet.

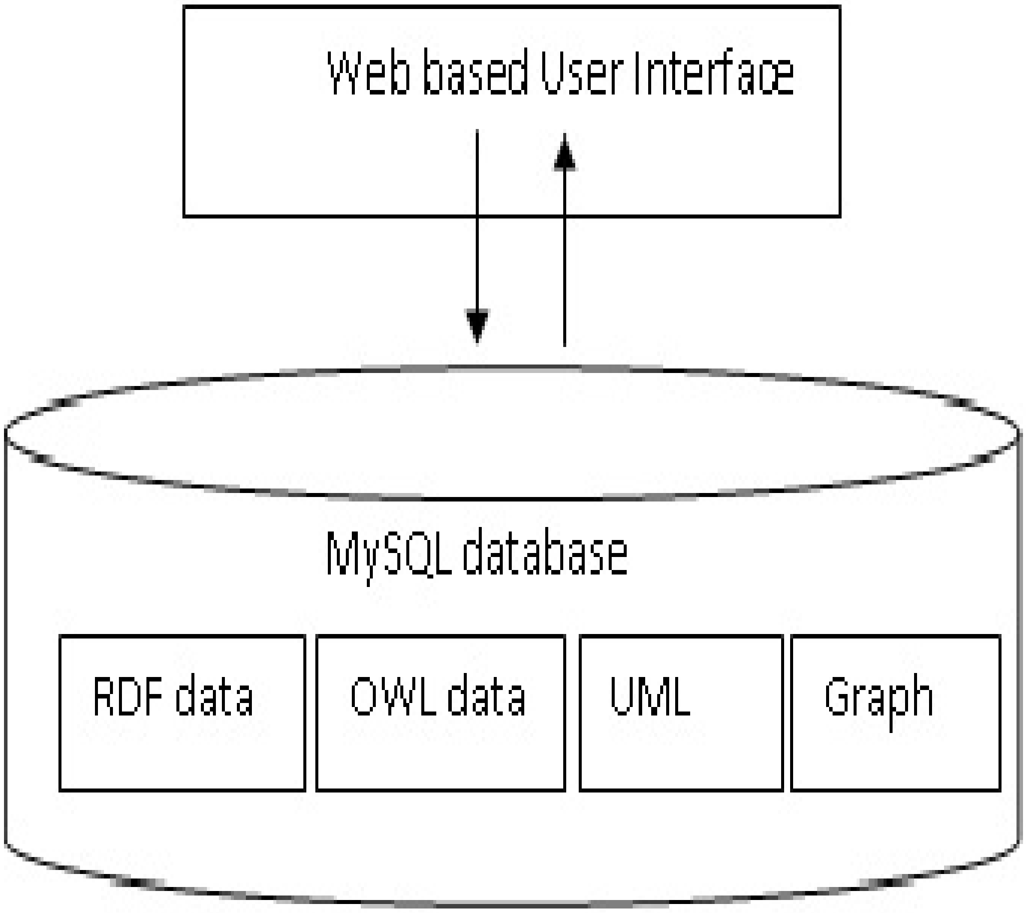


Figure 3.1 System Architecture of Knowledge Management Portal for Nuclear Reactor

iii. RDF/XML and OWL/XML rendering file options available in the Protégé tool, are utilized to generate the RDF, OWL files for the given ontology. These data are captured and stored in the database. This would aid the machine processing by allowing the crawler to crawl the web to extract the knowledge embedded in the system.

iv. The graphical representation of the knowledge is created using Onto_graph and the image is saved in jpeg or gif file format. This is an add_on tool available in the Protégé.

v. OWLGrEd is another program which when executed simultaneously with Protégé tool would enable the user to export the ontology to construct a UML graph.

vi. The collected knowledge and information of each system is stored in the database for the user and agent program to retrieve process and display the knowledge.

The screen shot of KMNuR portal main page (Figure 3.2) is expanded and shown in Figure 3.3 and Figure 3.4 The left side frame of the web portal lists out the names of systems such as reactor core, control rod drive mechanism, primary sodium system and the parameters such as nuclear flux, gamma ray source, reactor steady state etc. related to nuclear reactor domain. Systems and parameters of the nuclear reactor, whose knowledge is represented in the KMNuR portal, are listed out in Table 3.1. When the user selects a particular system it will provide information about it on the right side frame. RDF or UML graph or OWL buttons are made available in the portal thereby allowing the user to view the acquired knowledge. Since the UML images are complex in nature, zoom facility is offered to get an enlarged image. The overall aim of the KMNuR portal is to integrate and infer the semantic knowledge needed by the users, in performing the nuclear reactor domain tasks.

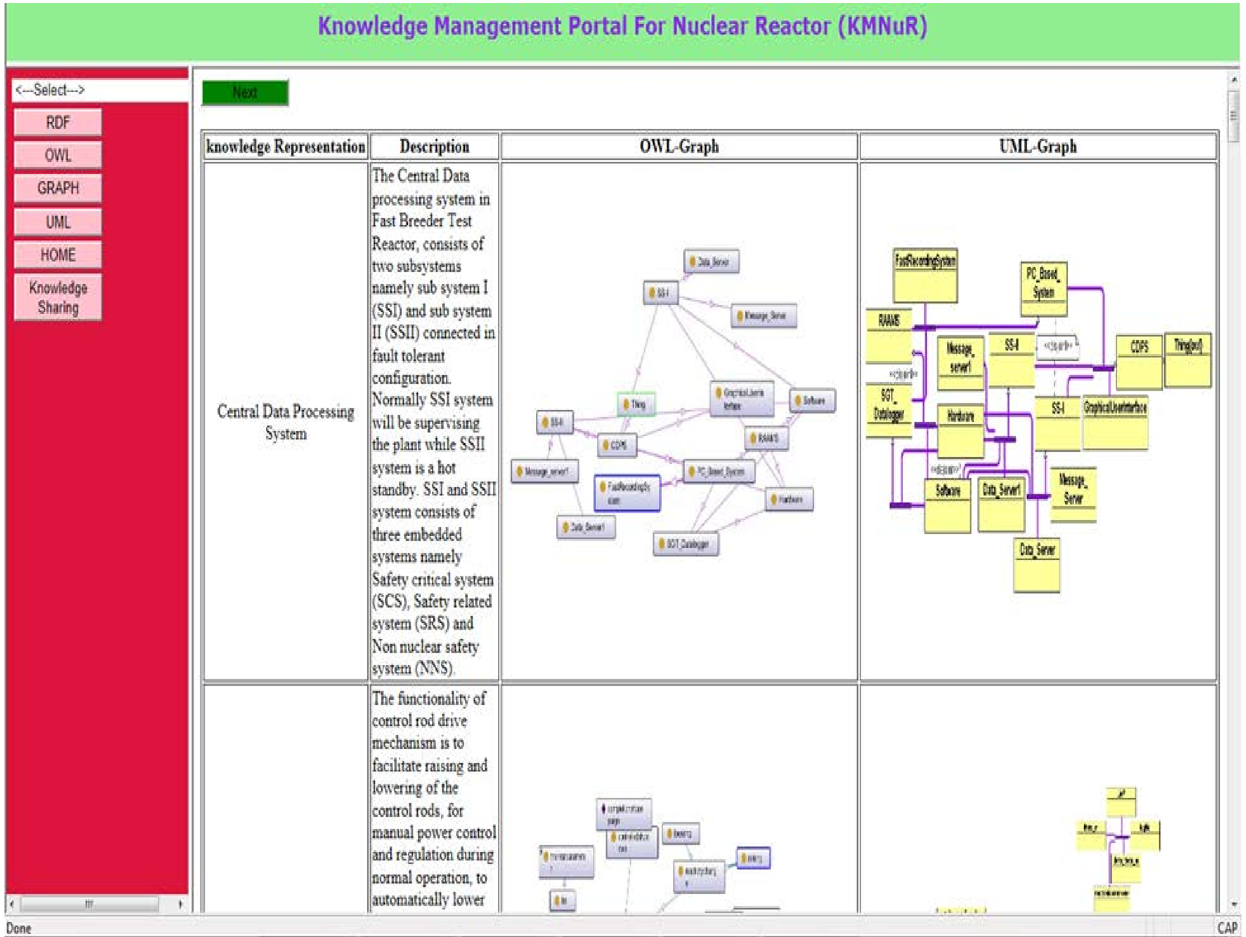


Figure 3.2 Main screen of KMNuR showing the knowledge representation of each system, its description, OWL_Graph and UML_Graph

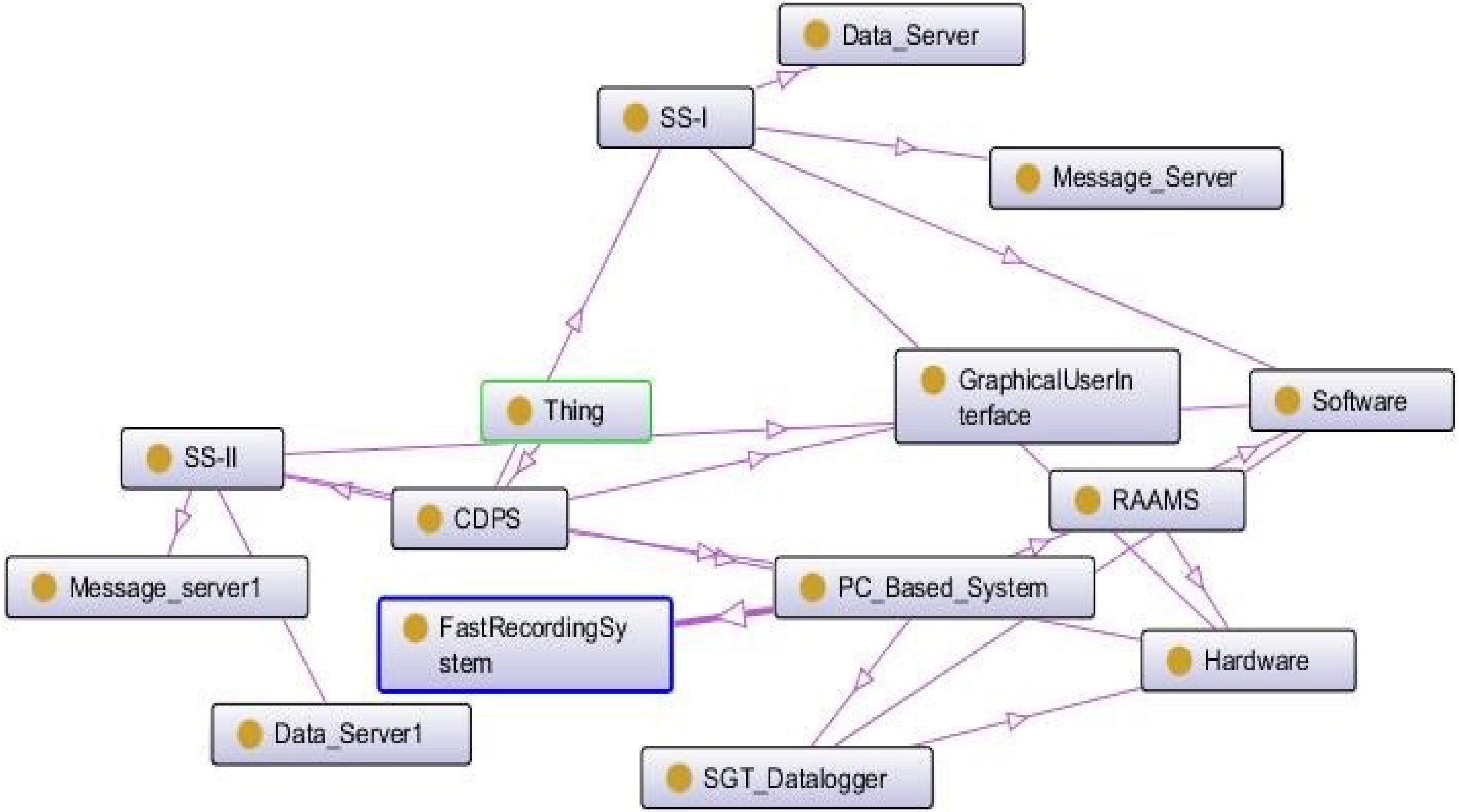


Figure 3.3 OWL-GRAPH enlarged image of Figure 3.2 for CDPS

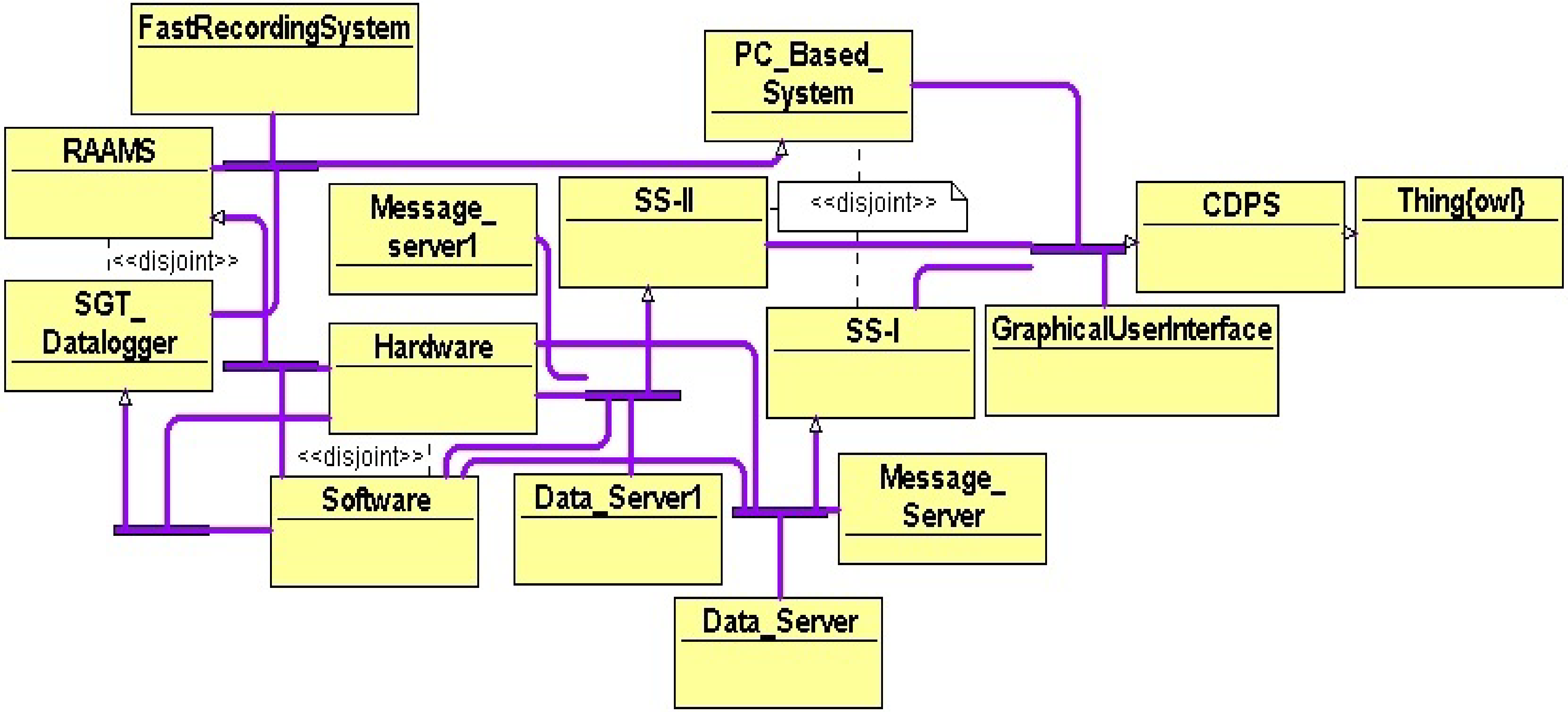


Figure 3.4 UML-graph enlarged image of Figure 3.2 for CDPS system

Table 3.1 Systems and parameters of nuclear reactor in KMNuR portal

Sl.No	**Name of the System**	**Sl.No**	**Name of the System**
1	Central Data Processing System	25	Primary Sodium Circuit
2	Control Rod Drive Mechanism	26	Radiation
3	Electrical System	27	Reactivity Feedback
4	Emergency Core Cooling	28	ReactorAccident
5	FbtrOffsitePowerSupply	29	ReactorAccidentalCondition
6	Flux	30	Reactor Assembly
7	FuelFabricationOxidetoCarbide	31	Reactor core
8	FuelFabricationOxidetoMetal	32	ReactorDynamics
9	GammaInteractionWithMaterials	33	ReactorFueLHandling
10	GammaRaySource	34	ReactorShielding
11	LOR Parameter	35	ReactorShutdownMechanism
12	Neutronic_Channels	36	Reactor Start Up Condition
13	Ne_InteractionwithMaterials	37	ReactorState
14	NetReactivity	38	ReactorSteadyState
15	NeutronEnergy	39	Reactor Transients state
16	NuclearFission	40	Safety Analysis
17	NeutronicInstrumentation	41	Scram Parameter
18	Nuclear_Reaction	42	Secondary Sodium System
19	NuclearData	43	ShieldingMaterials
20	NuclearReactorDesign	44	ShieldwithinReactor
21	NuclearReactorDesignTest	45	Special_Transport_Equation
22	Nuclear Power Plant	46	Spent_Fuel
23	NuclearReactorFBTR	47	SteamGenerator
24	Permissivecircuits	48	SteamWater

Table 3.2 Database structure of KMNuR in MySQL database

Field	Type	Key
Id_no	integer(10)	Primary Key
Kr_name	varchar(30)	-
Owl_graph	longblob	-
Uml_graph	longblob	-
Kr_desc	varchar(1000)	-
Rdf_desc	varchar(2000)	-
Owl_desc	varchar(2000)	-

3.3 KMNuR Database Structure

KMNuR database structure (Table 3.2) consists of fields like unique identification number (Id_no), name of the system (Kr_name), its description (Kr_desc), OWL-Graph, Rdf_desc, Owl_desc and UML graph. All the main and auxiliary systems of the FBTR are collected and loaded in the database. As an example, **C**entral **D**ata **P**rocessing **S**ystem (CDPS) is taken for the analysis to discuss the representation of the KMNuR portal and details are given below.

3.4 Central Data Processing System

CDPS system is a computer based system used to monitor, control and take safety action in the FBTR system. Figure 3.3 depicts the enlarged image of the CDPS in OWL graph format. CDPS in FBTR consists of two subsystems namely SubSystem I (SSI) and SubSystem II (SSII) connected in fault tolerant configuration. SSI and SSII systems consist of three embedded systems, namely, safety critical system (SCS), safety related system (SRS) and non nuclear safety system (NNS). It also consists of radiation and air monitoring systems (RAAMS) for radiation related data processing, data server for acquiring nuclear data from the embedded system, message server system for retrieving fault messages from embedded system, fast recording system for recording analysis of safety parameters etc. Figure 3.4, depicts the enlarged image of the CDPS in UML-GRAPH format.

3.4.1 Representing RDF in KMNuR Portal

RDF is a metadata defined in the W3C family and is used for illustrating the relationships between the various object resources [192]. It is a URI based syntax data representation which allows a secure and reliable mechanism for metadata exchange between web applications [193]. RDF files for each of the system is generated using Protégé and are stored in the database. When a particular system in the nuclear domain is selected from the list, its corresponding matching RDF information is displayed in the main window as shown in Figure 3.5. As an example, when CDPS is selected in RDF format, it lists out the classes and their relationship.

3.4.2 Representing OWL in KMNuR Portal

The OWL is an extension of RDF and RDF-S. It facilitates greater machine interpretability of the web content than that supported by XML, RDF and RDF-S by providing additional vocabulary along with formal semantics. It describes properties like disjointness, cardinality, equality, symmetry, transitivity, functional, inverse and enumerated relations between classes. Using the OWL language, components like concepts, instances, properties (or relations) and axioms of ontology are created [194]. *OWL Lite*, *OWL DL* (Description Language) and *OWL Full* are three different versions of the OWL sub languages [195]. It aims to bring the reasoning power of description logic to the semantic web. By clicking the OWL button in the KMNuR portal, the corresponding OWL file for the CDPS system is listed out as shown in Figure 3.6.

Figure 3.5 Part of screen shot of RDF file for the CDPS system.

118


```
<?xml version="1.0"?>


<!DOCTYPE Ontology [
    <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
    <!ENTITY xml "http://www.w3.org/XML/1998/namespace" >
    <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
    <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]>


<Ontology xmlns="http://www.w3.org/2002/07/owl#"
     xml:base="http://www.semanticweb.org/ontologies/2012/2/cdps.owl"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:xml="http://www.w3.org/XML/1998/namespace"
     ontologyIRI="http://www.semanticweb.org/ontologies/2012/2/cdps.owl">
    <Prefix name="" IRI="http://www.w3.org/2002/07/owl#"/>
    <Prefix name="owl" IRI="http://www.w3.org/2002/07/owl#"/>
    <Prefix name="rdf" IRI="http://www.w3.org/1999/02/22-rdf-syntax-ns#"/>
    <Prefix name="xsd" IRI="http://www.w3.org/2001/XMLSchema#"/>
    <Prefix name="rdfs" IRI="http://www.w3.org/2000/01/rdf-schema#"/>
    <Declaration>
```

Figure 3.6 Part of screen shot of OWL file for the CDPS system

3.4.3 Representing OWL-GRAPH in KMNuR Portal

Properties and relationships of nuclear knowledge are viewed in graphical representation by using the onto_graph plug-in available in Protégé. The same can be stored in the graphical formats like *jpeg, gif, emf, bmp* etc. In KMNuR database, OWL_GRAPH in *jpg* file format is used for storing. The snapshot of the CDPS system is shown in OWL graph format in Figure 3.7.

3.4.4 Representing OWL_UML in KMNUR

UML is a general-purpose modeling language used to capture the information about different views of systems, like static structure and dynamic behavior. UML and OWL have identical characteristics for defining classes, associations, properties, packages, types, generalization and instances etc [196].
OWLGrEd is UML style based graphical notation editor for OWL. Each individual tool of OWLGrEd is created through a specially designed **T**ransformation **D**riven **A**rchitecture (TDA) tool for storing information like types, styles, constraints and relationships among elements [197]. Both Protégé and TDA tool, should be executed parallely, to convert the OWL to UML. In Protégé tool, “EXPORT to TDA” menu option is chosen to convert OWL to UML. By doing this the OWL file is opened in TDA program automatically. Similarly, to export the ontology from OWLGrEd to Protégé, user has to right-click on the UML diagram and then has to select “Export to Protégé” option.

In the KMNuR Portal, rollover to zoom_in facility is also available to view the enlarged size of the UML image for a particular system as shown in Figure 3.8.

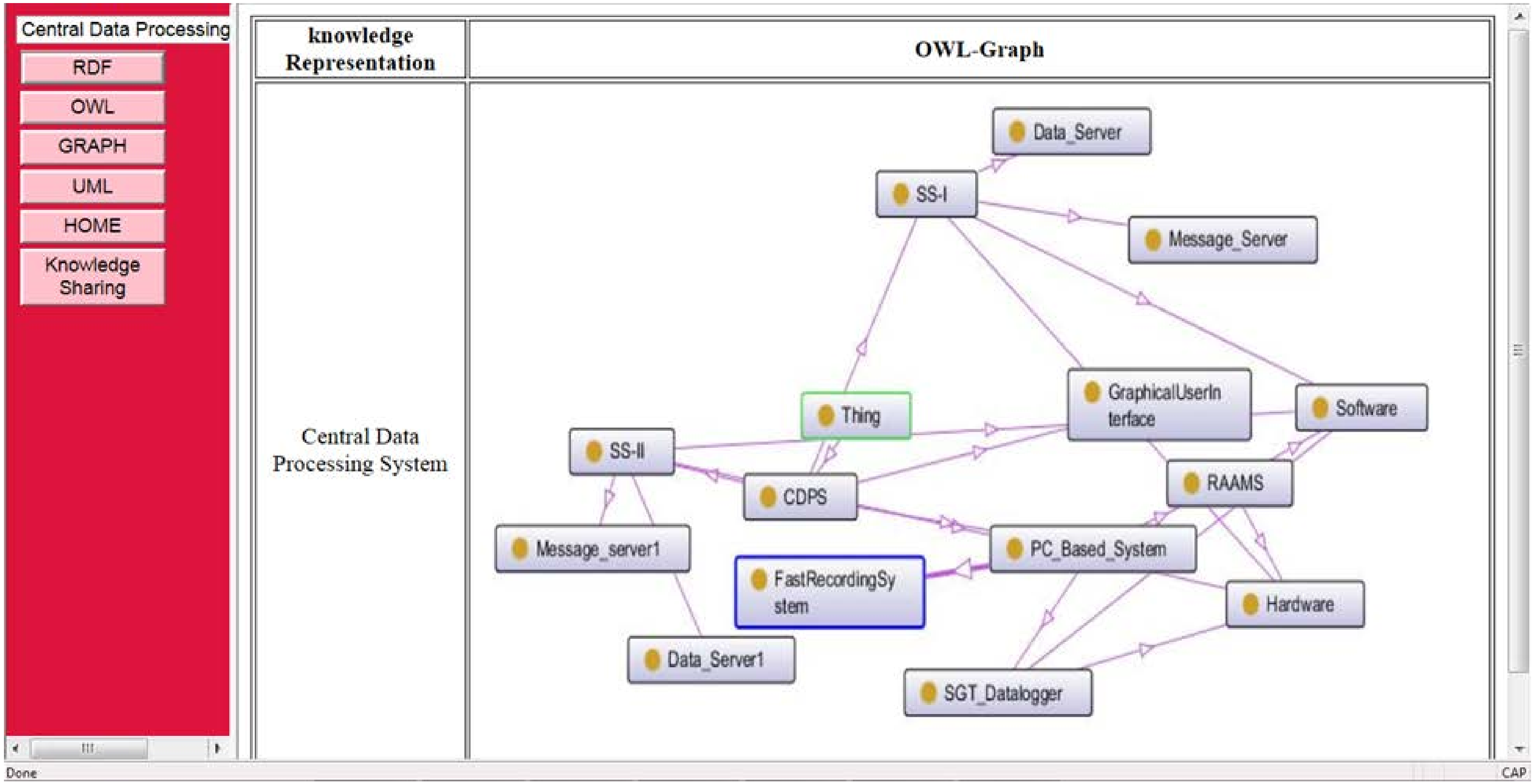


Figure 3.7 Part of screen shot of OWL-GRAPH for the CDPS system.

120

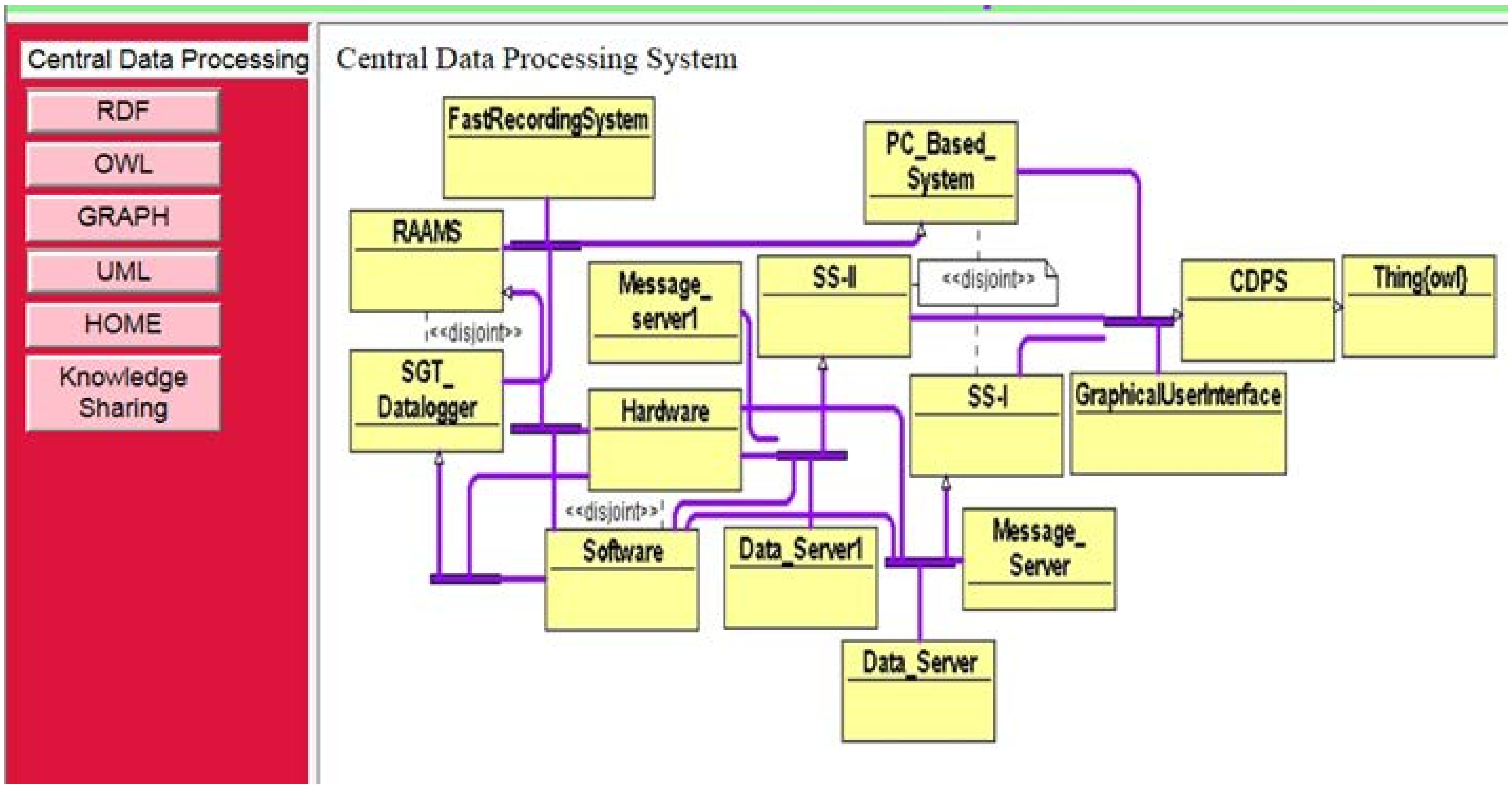


Figure 3.8 Part of screen shot of UML graph for the CDPS system

KMNuR is deployed presently in the intranet to address issues related to nuclear knowledge management of FBTR. The initiative to develop the portal is an attempt to integrate the existing nuclear data and information bases about FBTR in an easily accessible form. Feedback and inputs from the domain experts and operator personnel will further enhance and enrich the available knowledge of KMNuR to maintain the required quality. Finally this would be deployed to share the fast breeder knowledge across the World Wide Web enabling reusing, sharing and processing the nuclear knowledge.

3.4.5 Knowledge Sharing

In the KMNuR portal by pressing the “Knowledge Sharing” menu the users can submit new inforamtion to the system. This is shown in Figure 3.9. The user has to fill in details like name of the submitting knowledge, authors contributing the knowledge, details of the knowledge, RDF, OWL and UML format of the corresponding knowledge. Once the relevant data are entered then it is submitted to the database for approval by pressing the “Submit the Knowledge” button. These data are then stored in the database which then needs to be approved by the administrator. Once the administrator approves, it will be added to the web portal. The knowledge database is shown in the Table 3.3. In order to view the pending knowledge the “view pending knowledge base” button needs to be pressed.

In the next section the systems and parameters using UML and OWL data of the FBTR nuclear power plant are discussed.

Table 3.3 Databases structure of knowledge_base in MySQL database

Field	Type
Km_name	varchar(30)
Km_authors	varchar(100)
Km_desc	varchar(2000)
Km_RDF	varchar(2000)
Km_OWL	varchar(2000)
Km_UML	varchar(2000)
Approval_Status	Char

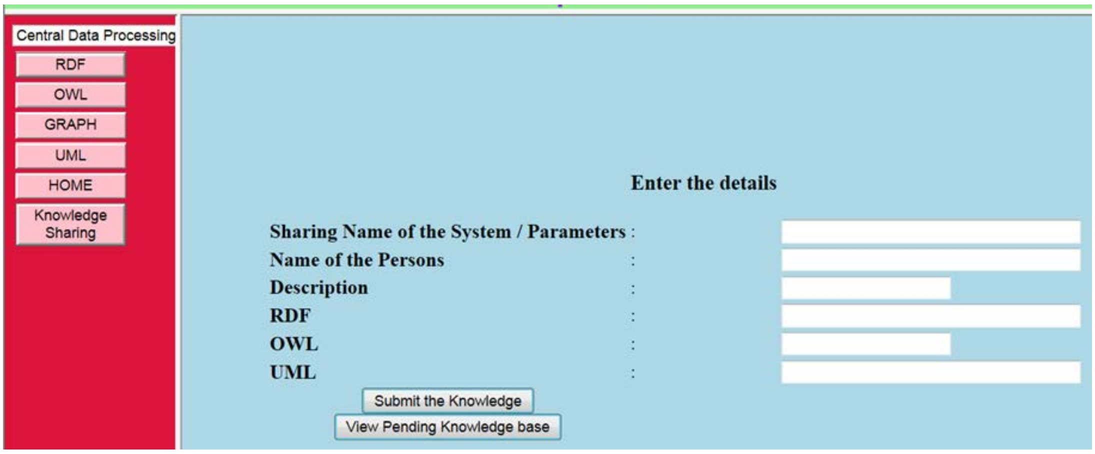


Figure 3.9 Part of screen shot of new information sharing

3.5 Nuclear Power Plant

Nuclear power plant has a number of interconnected subsystems with different functions to perform. OWL and UML representation for nuclear power plant are shown in Figure 3.10 and Figure 3.11.The reactor core, nuclear instrumentation, primary and secondary system, auxiliary system and related parameters are discussed.

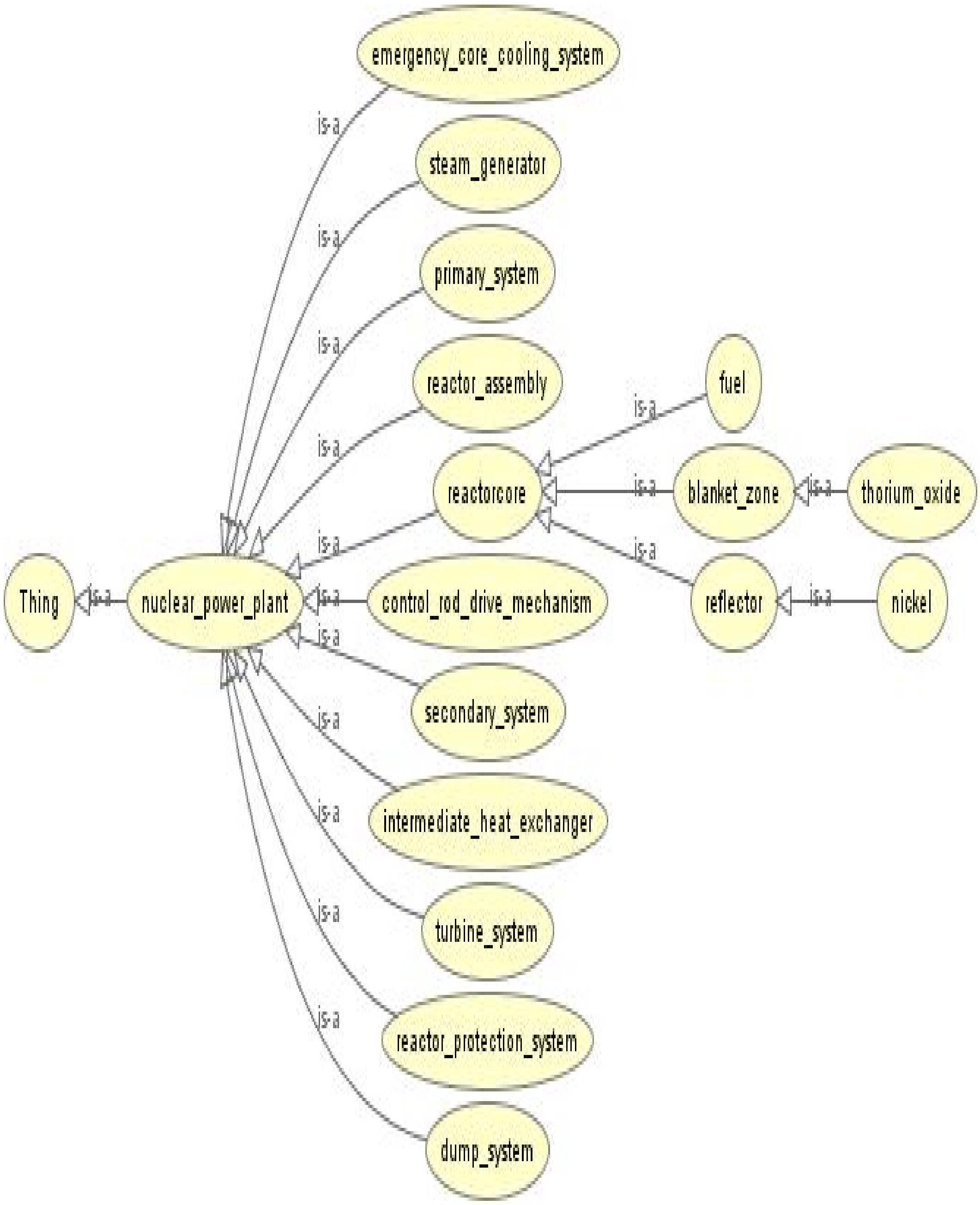


Figure 3.10 OWL representations for nuclear power plant

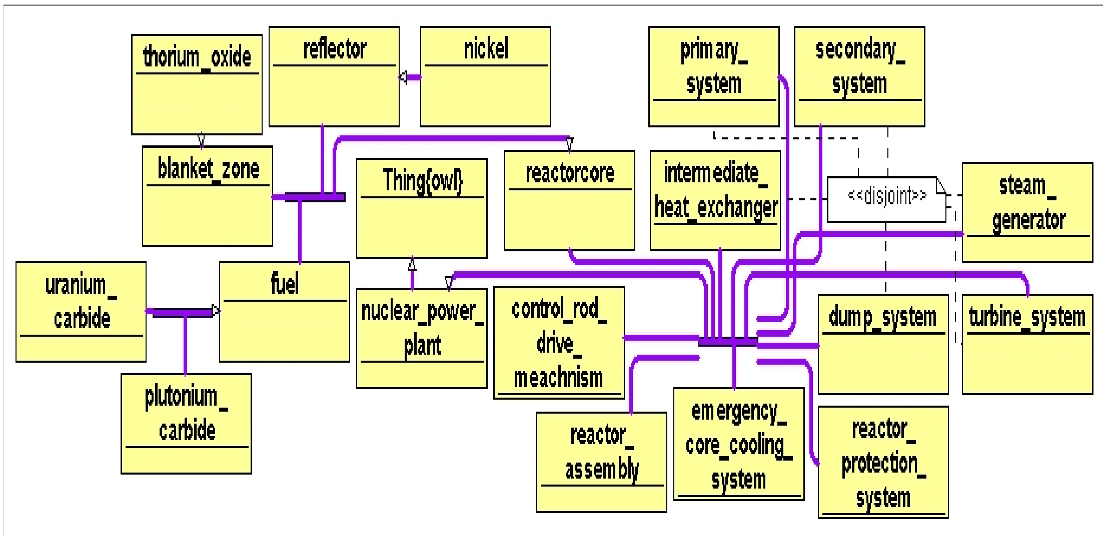


Figure 3.11 UML representations for nuclear power plant

3.5.1 Nuclear Fission

When fissile nuclei capture neutrons, fission takes place; the nuclei divide usually into two parts. In the process of fission about 200 MeV of energy is released, largely in the form of kinetic energy of the two fission fragments and this immediately manifests itself as heat [198]. Fission is also accompanied by the emission of two or three neutrons and gamma-rays, emitted from the fission fragments during their recoil. The products of fission are themselves generally neutron rich and hence radioactive. When some fission fragments have undergone a beta-decay they have sufficient excitation energy to emit a neutron. These neutrons, which are called delayed neutrons, are emitted with half lives of up to 55 seconds and are very important for the control of nuclear reactors. The beta and gamma-ray decay energy of the fragments is emitted over a long period. OWL and UML representation for nuclear fission are shown in Figure 3.12 and Figure 3.13.

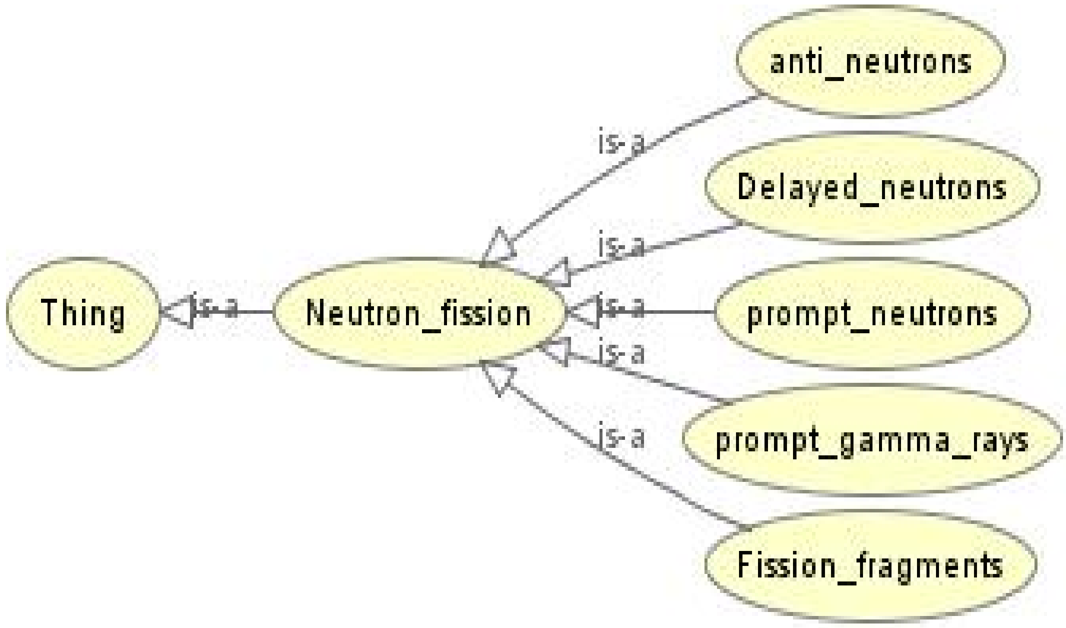


Figure 3.12 OWL representations for nuclear fission

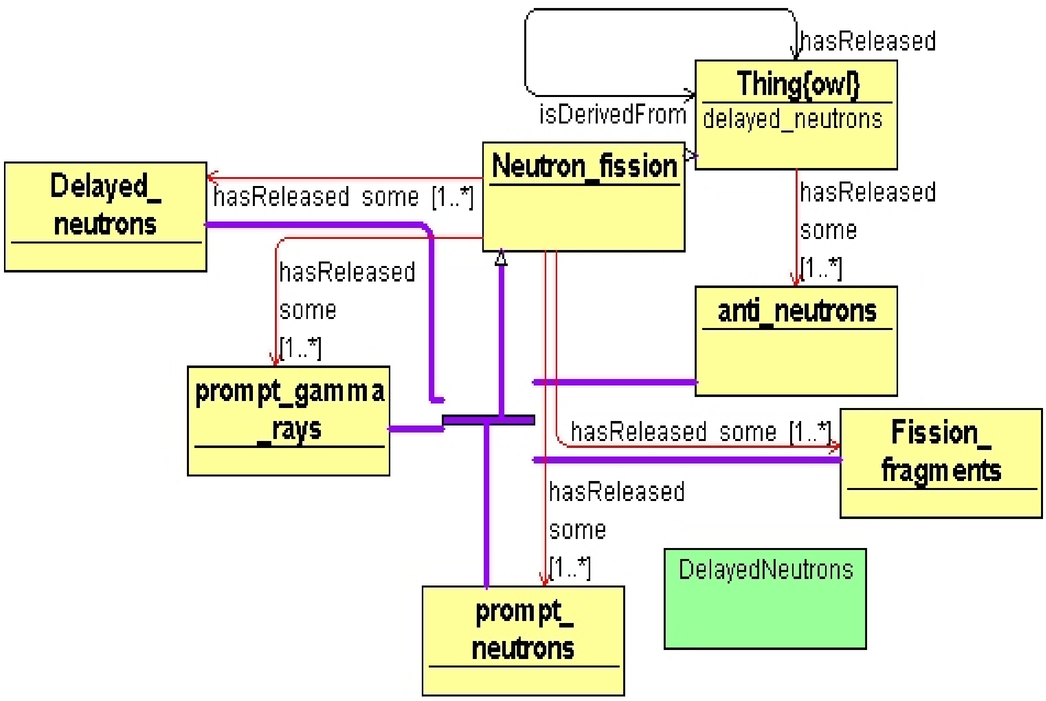


Figure 3.13 UML representations for nuclear fission

3.5.2 Neutron Energy

Neutrons are broadly classified into thermal and fast neutrons. A fast neutron is a free neutron with a kinetic energy close to 1 MeV. A thermal neutron is a free neutron with energy 0.0253 eV. Neutrons are classified based on energy and the method of production. Most fission reactors utilize neutron moderator to slow down, or thermalize the neutrons that are emitted by nuclear fission. OWL and UML representation for neutron energy are shown in Figure 3.14 and Figure 3.15.

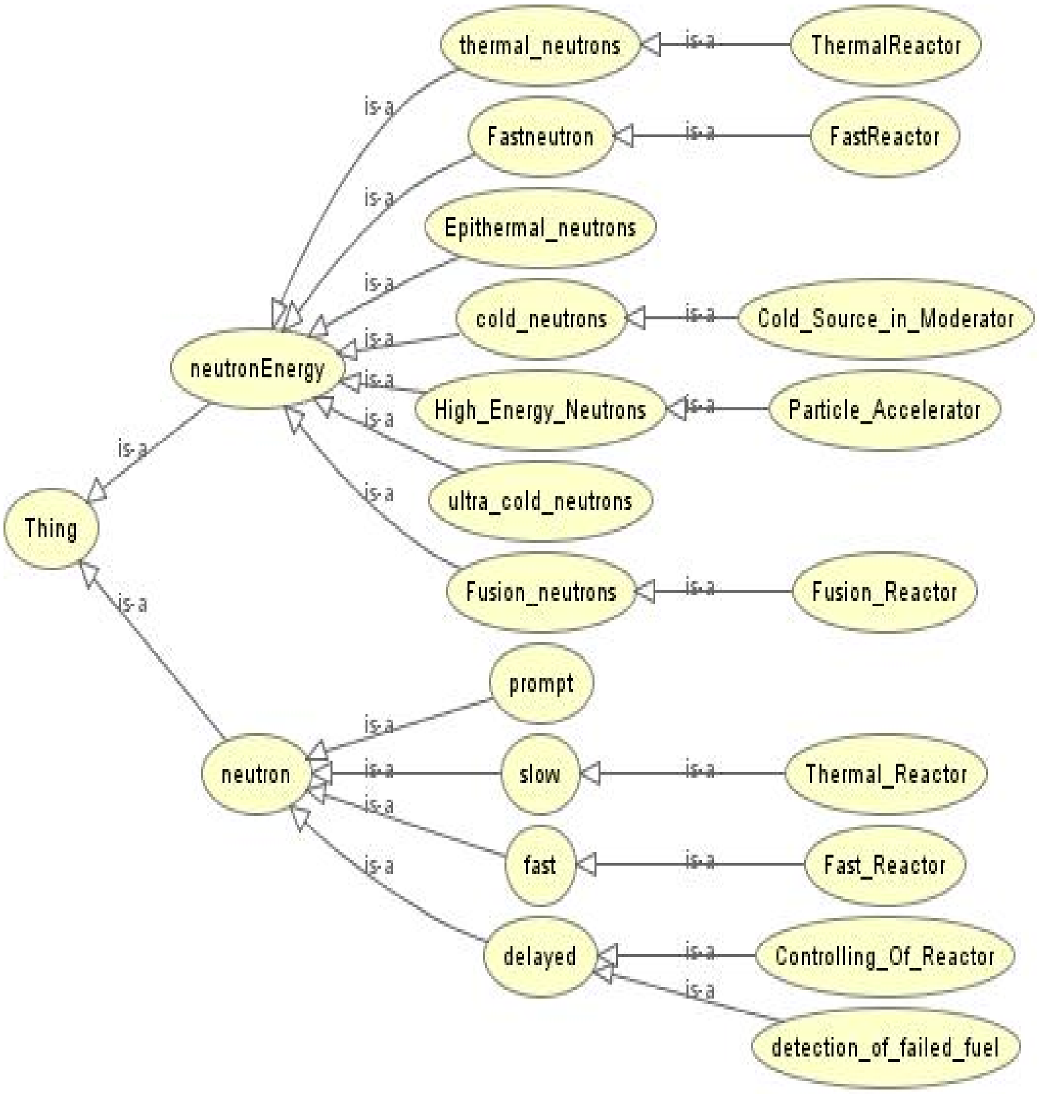


Figure 3.14 OWL representations for neutron energy

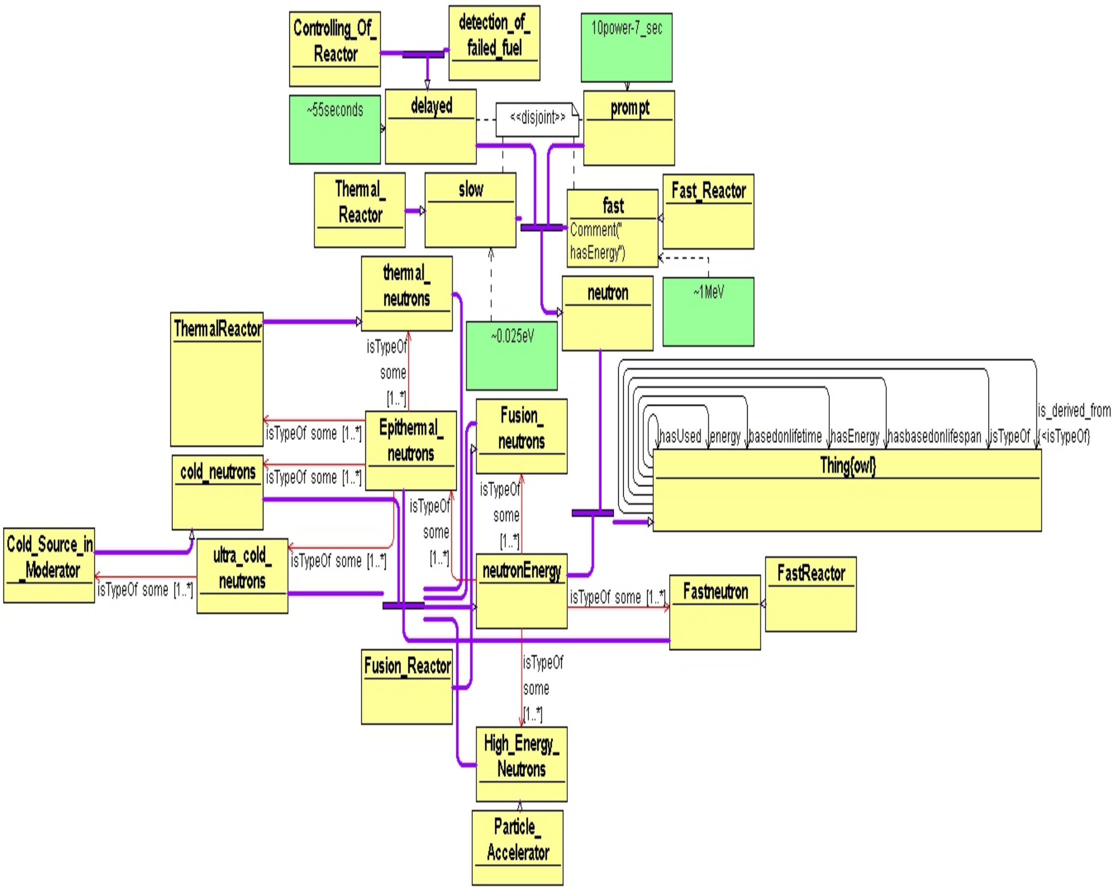


Figure 3.15 UML representations for neutron energy

3.5.3 Nuclear Reaction

A nuclear reaction is semantically considered to be the process in which two nuclei, or else a nucleus of an atom and a subatomic particle (such as a proton, neutron, or high energy electron) from outside the atom, collide to produce one or more nuclides that are different from the nuclide(s) that began the process. Nuclear fission is the process in which the nucleus of an atom (heavy) splits into smaller parts (lighter nuclei). In the conventional nuclear fission, fuel is divided into two: fissile nucleus (U^{233}, U^{235} and Pu^{239}) and fertile nucleus (Th^{232}, U^{238}). OWL and UML representation for nuclear reaction are shown in Figure 3.16 and Figure 3.17.

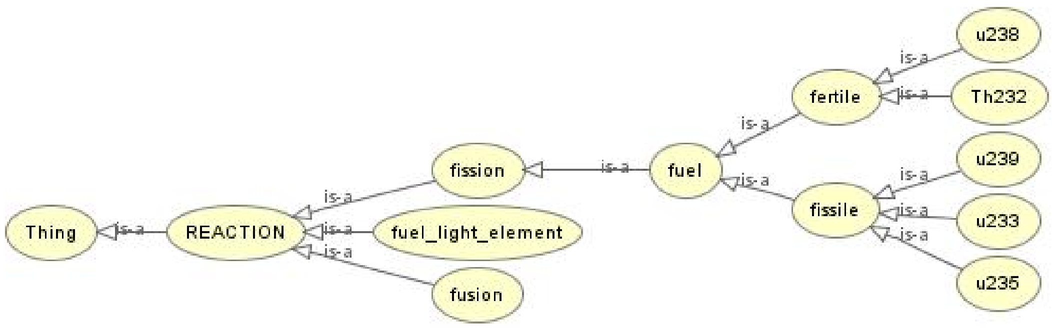


Figure 3.16 OWL representations for nuclear reaction

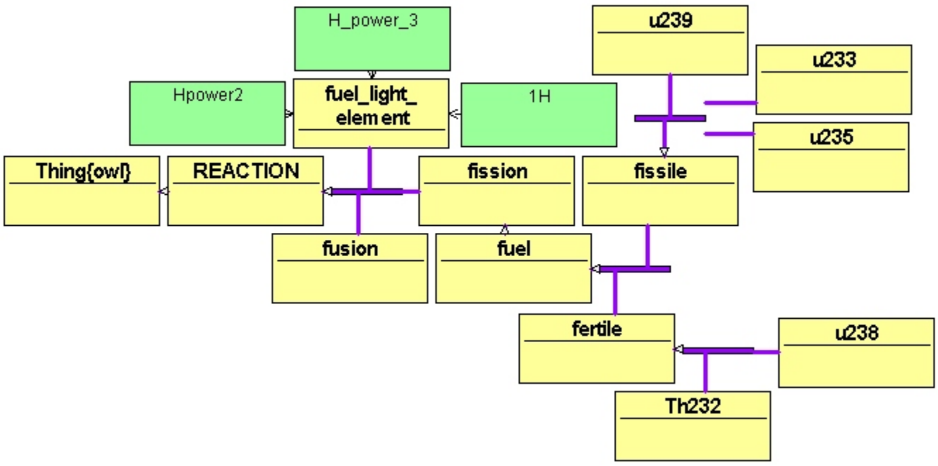


Figure 3.17 UML representations for nuclear reaction

3.5.4 Neutron Interaction with Materials

Neutron interaction with material is broadly classified into two categories, namely, scattering and absorption. In scattering, neutron merely exchange energy and remains free after the interaction. In absorption, neutron is retained in the nucleus and a new particle is produced. Scattering involves elastic and inelastic scattering. In elastic scattering kinetic energy is conserved. In inelastic scattering, the neutron loses some of its kinetic energy to the target nucleus. OWL and UML representation for neutron interaction with materials are shown in Figure 3.18 and Figure 3.19.

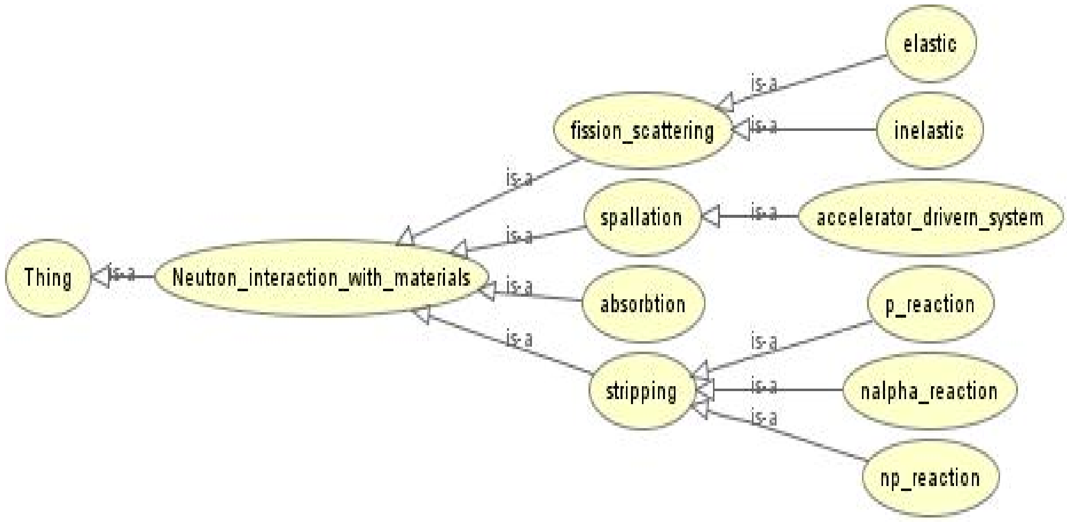


Figure 3.18 OWL representations for neutron interaction with materials

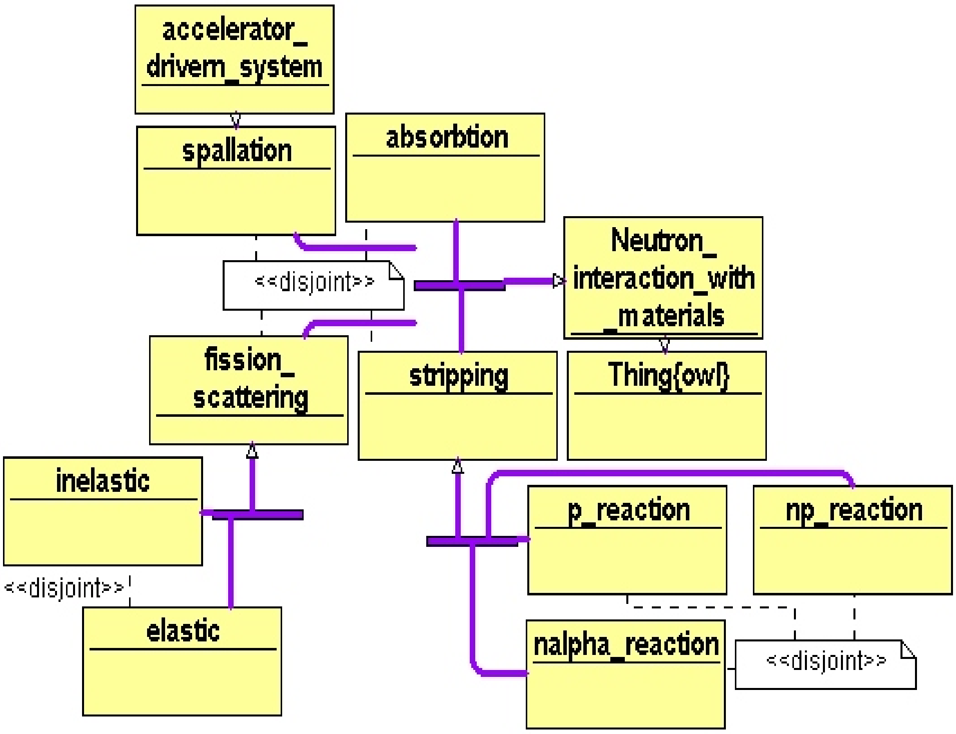


Figure 3.19 UML representations for neutron interaction with materials

3.5.5 Gamma Ray Source

Gamma rays are produced from fission of a nucleus in a reactor. After the fission, prompt gamma rays and delayed gamma rays are produced. Natural sources of gamma rays include gamma decay from naturally occurring radioisotopes and secondary radiation from atmospheric interactions with cosmic ray particles. Gamma rays are also produced by Bremsstrahlung, inverse Compton scattering and synchrotron radiation, in reactors by activation of atoms. OWL and UML representation for gamma ray source are shown in Figure 3.20 and Figure 3.21.

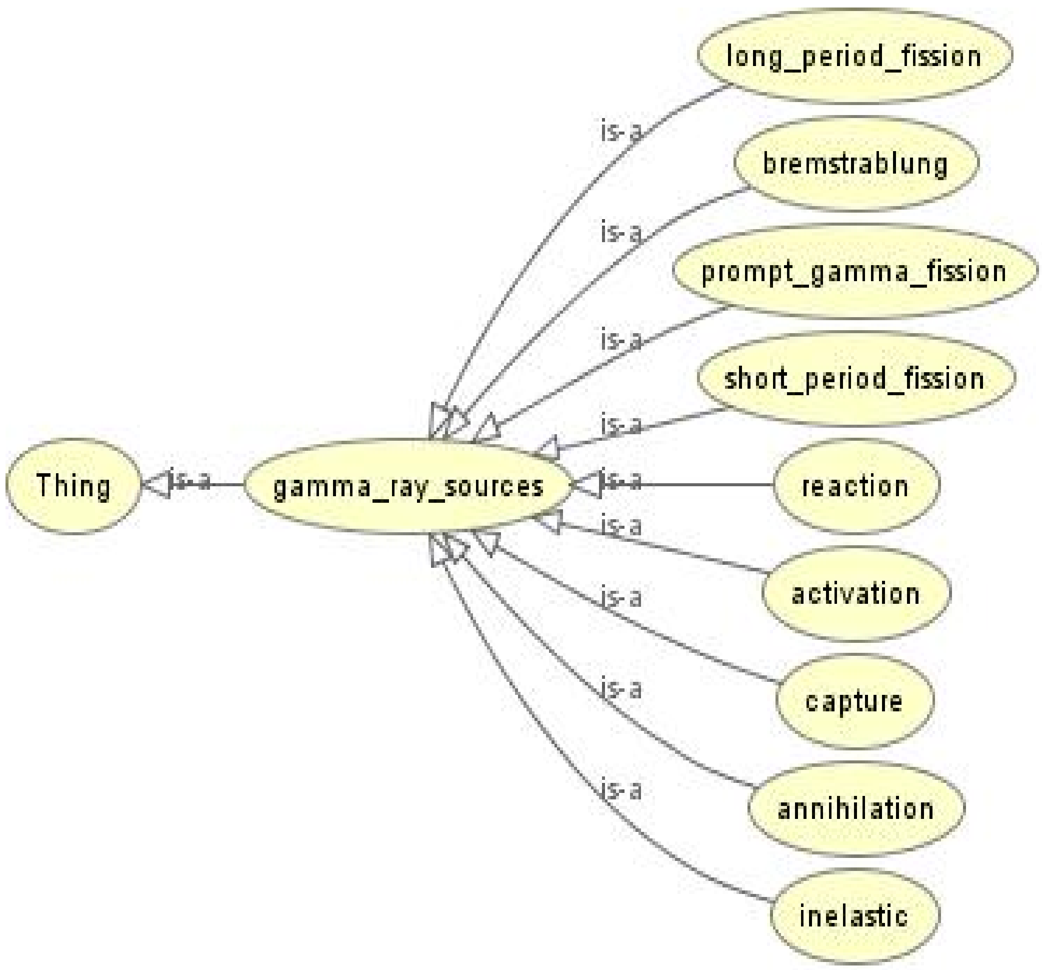


Figure 3.20 OWL representations for Gamma ray source

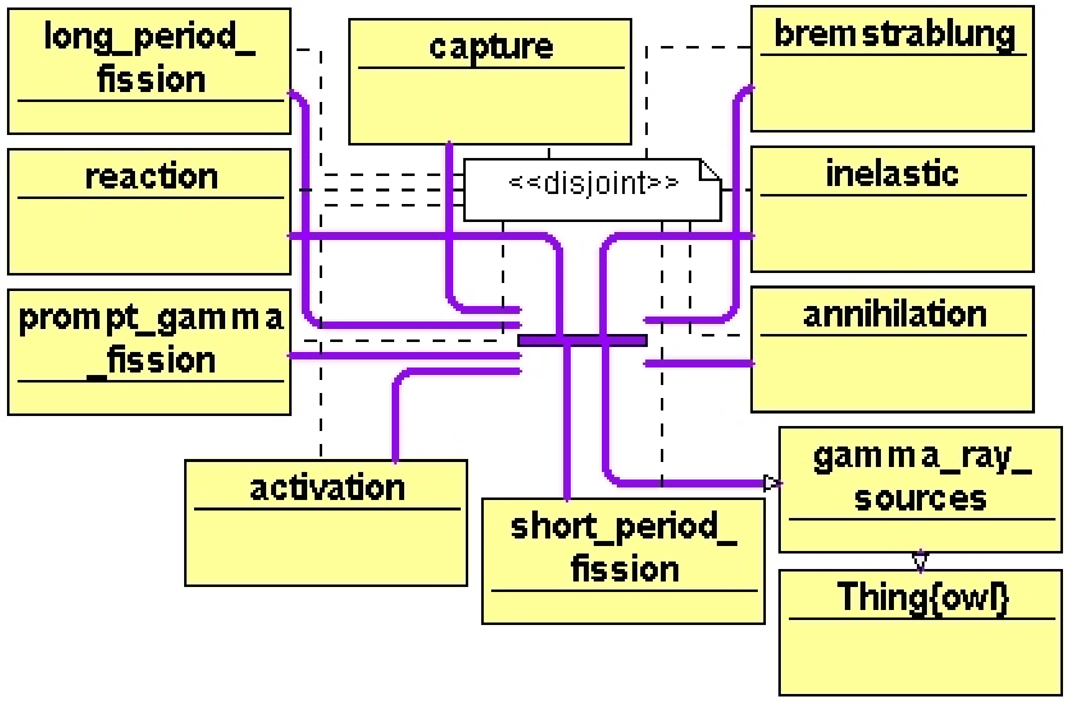


Figure 3.21 UML representations for Gamma ray source

3.5.6 Gamma Interaction with Materials

There are several ways in which gamma rays interact with materials. Among them photo electric pair production and Compton scattering are the important interactions. These gamma interactions occur with atomic electrons. It varies as a function of atomic number and gamma energy. OWL and UML representation for gamma interaction with materials are shown in Figure 3.22 and Figure 3.23.

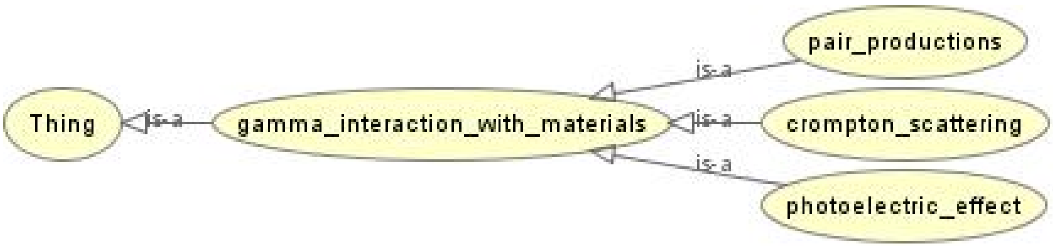


Figure 3.22 OWL representations for Gamma interaction with materials

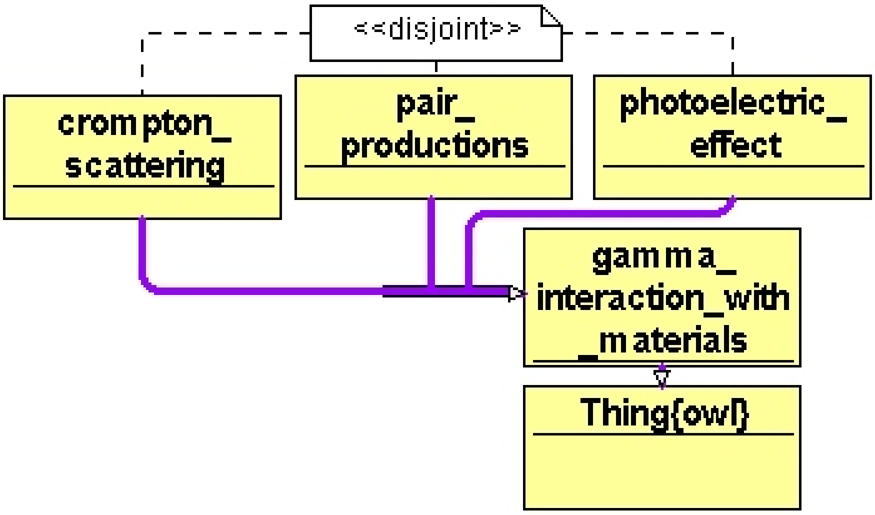


Figure 3.23 UML representations for Gamma interaction with materials

3.5.7 Nuclear Data

Nuclear data is a number which gives quantitative information of an atomic nucleus. Nuclear data is classified into (a) nuclear constants (b) nuclear structure ground & excited states, (c) nuclear decay data half-life, decay energy (d) nuclear reaction data cross sections of particles/radiation ejected. Nuclear data are obtained through experiments and theoretical models. OWL and UML representation for nuclear data are shown in Figure 3.24 and Figure 3.25.

3.5.8 Nuclear Reactor Design

Nuclear reactors are designed for particular power output and for prescribed lifetime. Using nuclear code, the fuel composition and control rod worth is obtained for the given power along with the feedback coefficients, like, temperature, void, burn-up, power and Doppler. OWL and UML representation for nuclear reactor design are shown in Figure 3.26 and Figure 3.27.

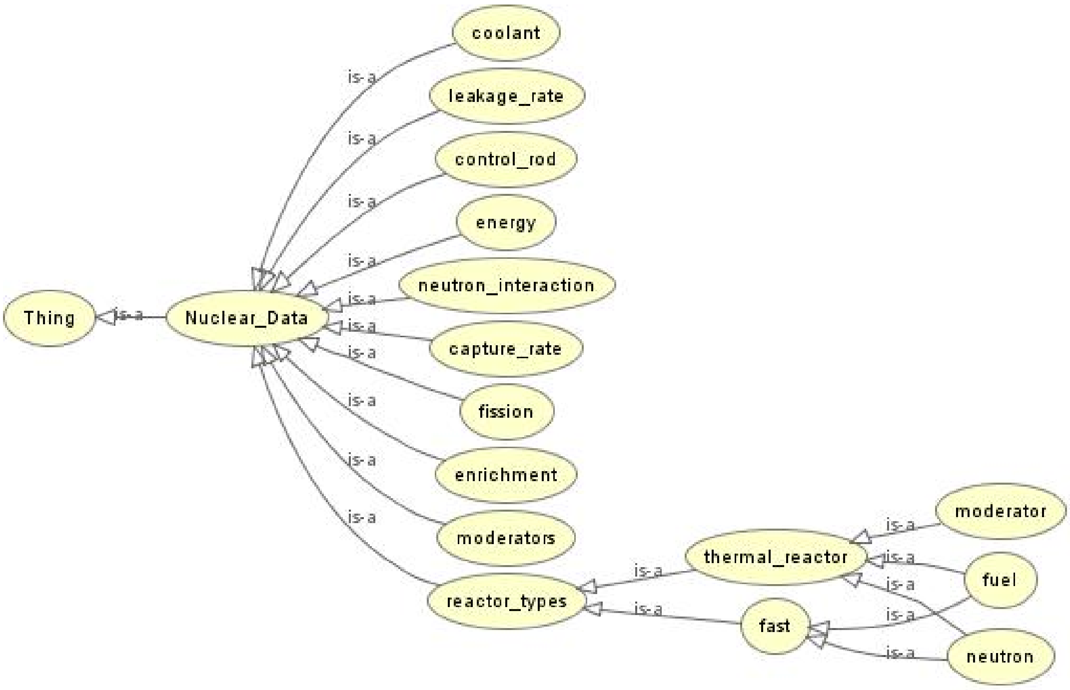


Figure 3.24 OWL representations for nuclear data

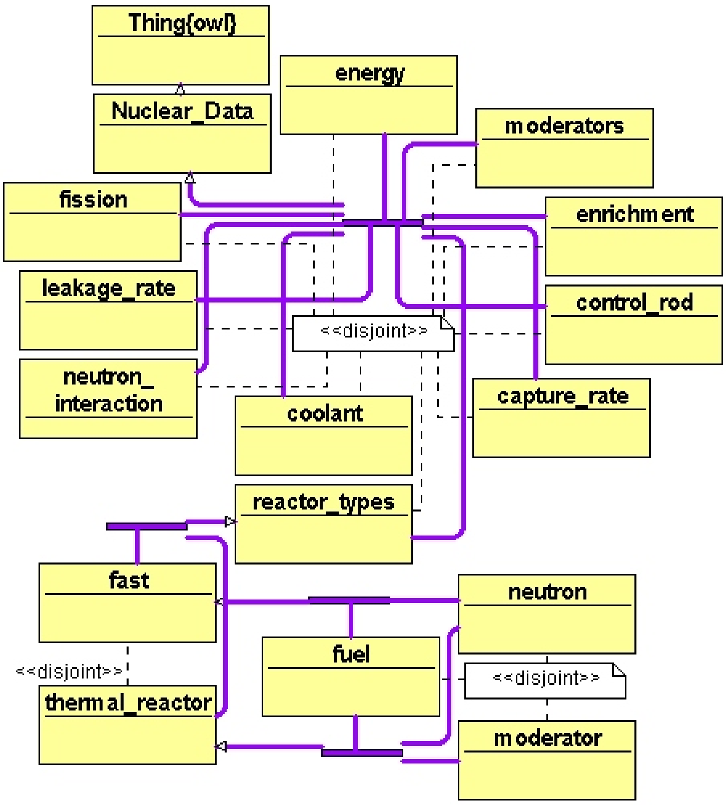


Figure 3.25 UML representations for nuclear data

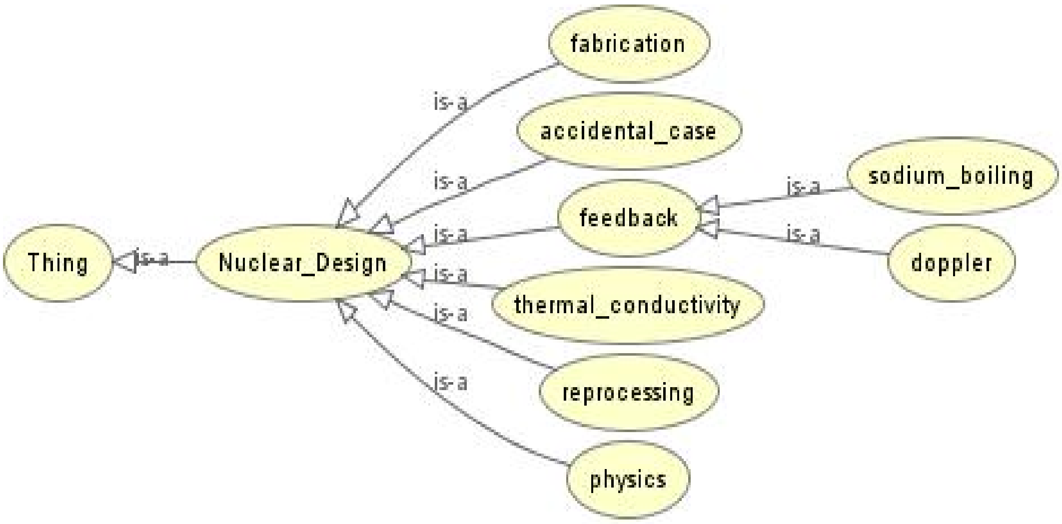


Figure 3.26 OWL representations for nuclear design

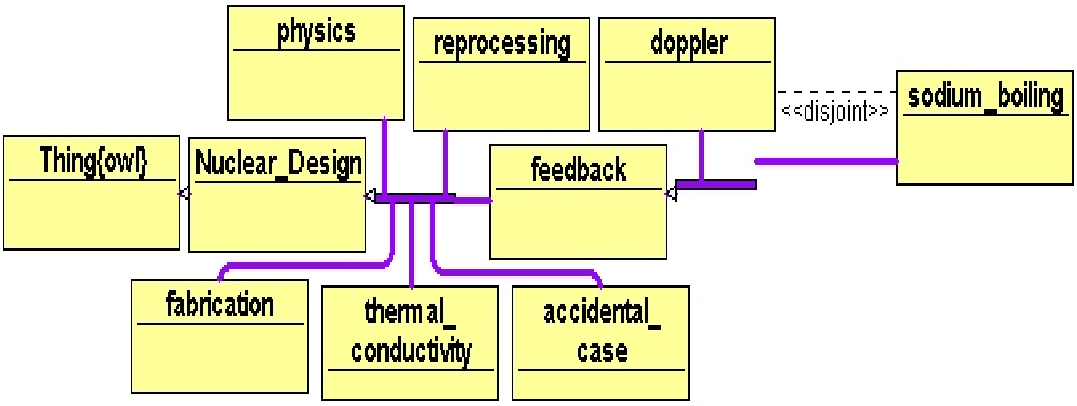


Figure 3.27 UML representations for nuclear reactor design

3.5.9 Nuclear Reactor Design Test

The reactor parameters for a given reactor can be verified using deterministic/transport code. Depending upon the complexity of the design, 1-D, 2-D or 3-D calculation is carried out. OWL and UML representation for nuclear reactor design test are shown in Figure 3.28 and Figure 3.29.

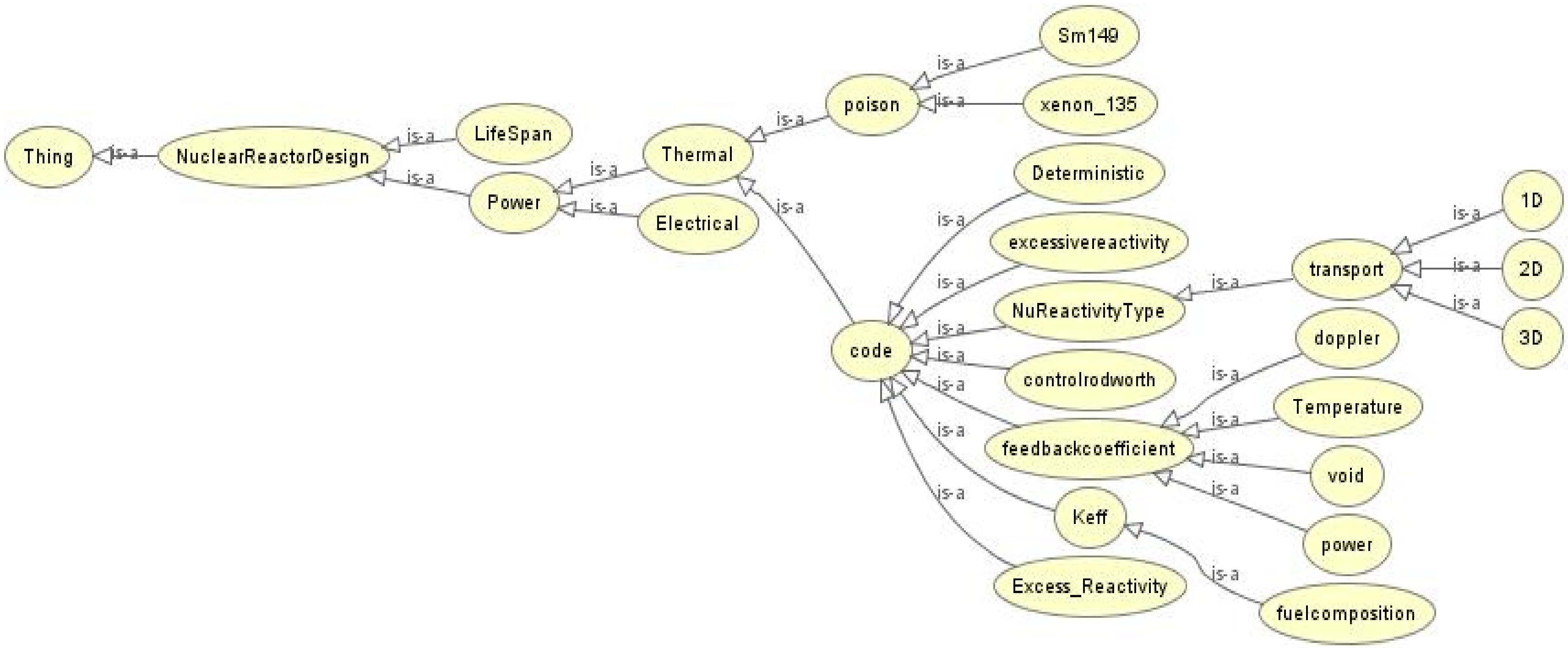


Figure 3.28 OWL representations for nuclear reactor design test

137

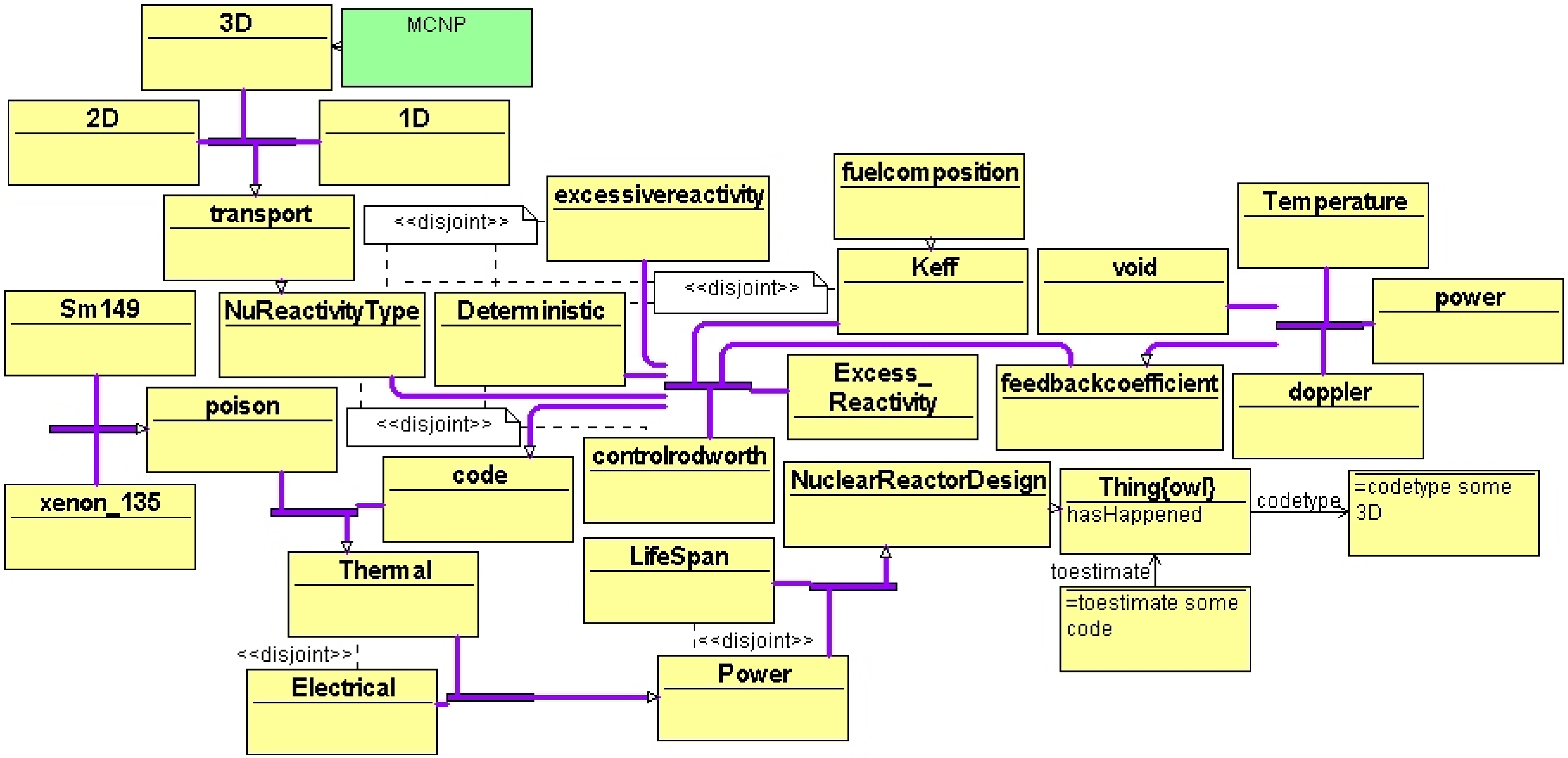


Figure 3.29 UML representations for nuclear design test

3.5.10 Reactor Core

The reactor core (FBTR core) consists of a central zone of fuel subassemblies containing Uranium Carbide and Plutonium Carbide. The nickel reflector zone encircles the fuel region and reflector zone surrounded by the blanket zone containing thorium oxide. The core contains six control rods placed in the fissile zone to control the reactor power. OWL and UML representation for reactor core are shown in Figure 3.30 and Figure 3.31.

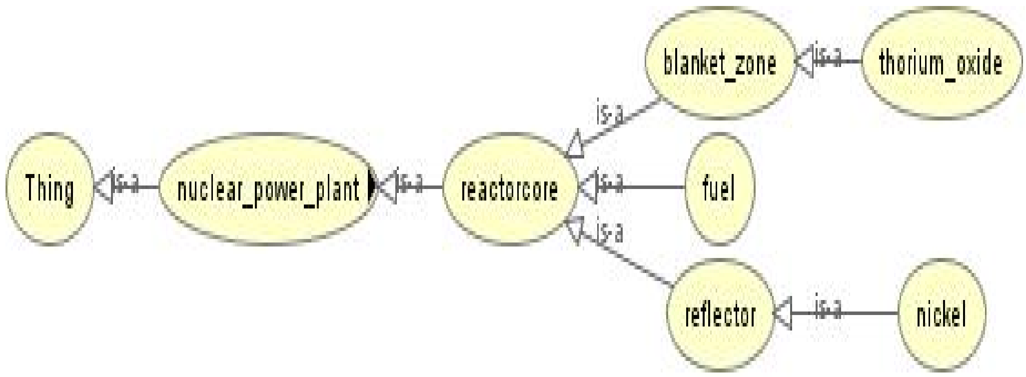


Figure 3.30 UML representations for reactor core

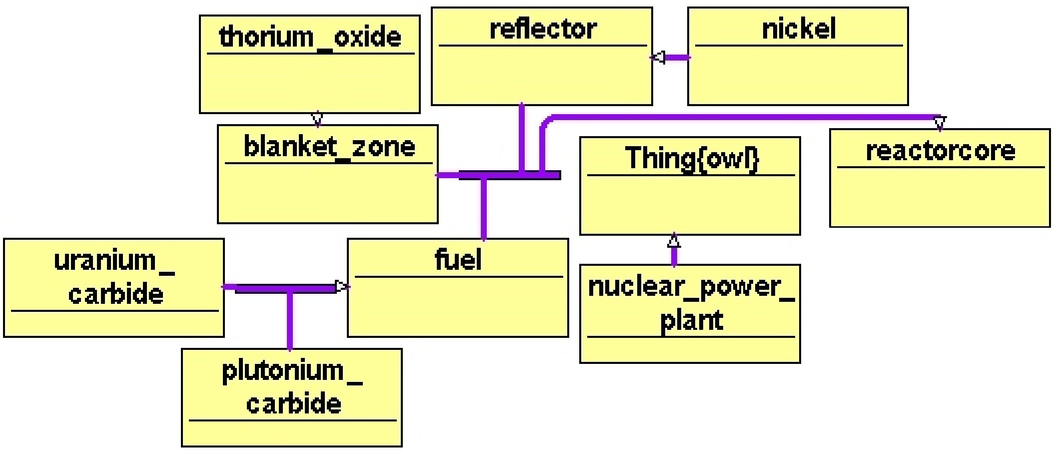


Figure 3.31 UML representations for reactor core

3.5.11 Reactor Assembly

The reactor vessel of FBTR is made of cylindrical stainless steel and houses fuel, blanket, reflector and control rod assemblies which are supported on a grid plate. The neutron shield surrounding these assemblies reduces the neutron gamma flux to the reactor vessel. The thermal shields are provided inside the reactor vessel to reduce the thermal stress. The vessel has a double envelope to contain liquid sodium in the event of any leak. The primary coolant enters the reactor vessel at the bottom, passes through the sub assemblies and leaves through two diametrically opposite outlets which are well above the heads of the sub-assemblies. OWL an UML representation for reactor assembly are shown in Figure 3.32 and Figure 3.33.

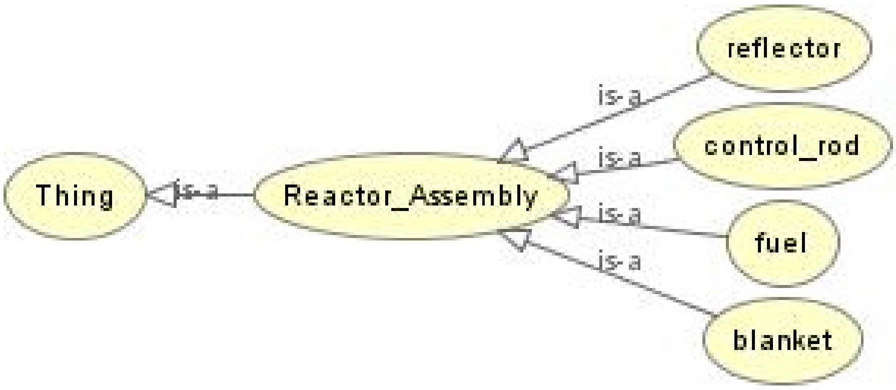


Figure 3.32 OWL representations for reactor assembly

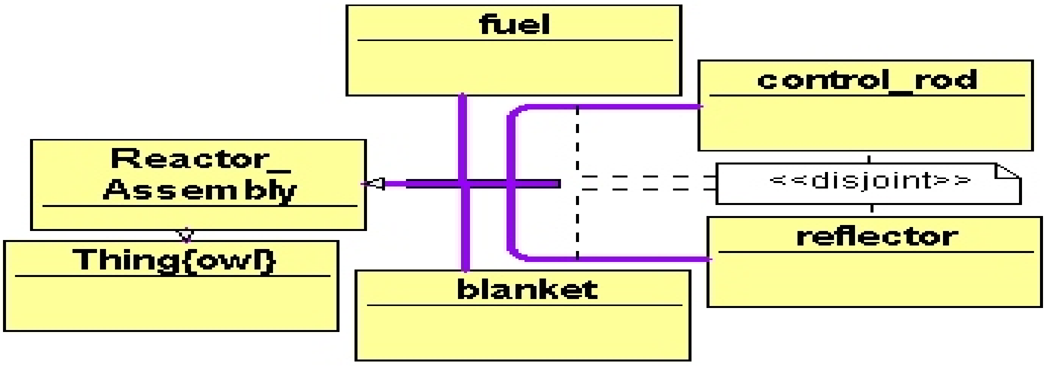


Figure 3.33 UML representations for reactor assembly

3.5.12 Primary Sodium Circuit

The primary sodium circuit carries out the following function to cool the subassemblies and transport useful heat energy from the core to the secondary sodium circuit through the intermediate heat exchanger. This is done to maintain sufficient level of sodium in the reactor vessel under all circumstances so that the fuel subassemblies always remain fully immersed in sodium, to detect the fuel clad rupture by the principle of detection of delayed neutron [199]. OWL and UML representation for primary sodium circuit are shown in Figure 3.34 and Figure 3.35.

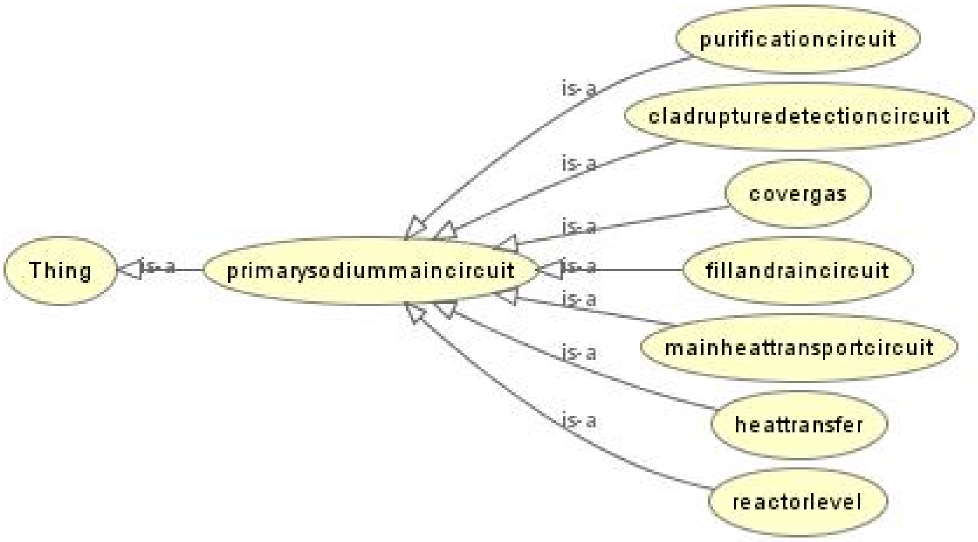


Figure 3.34 OWL representations for primary sodium circuit

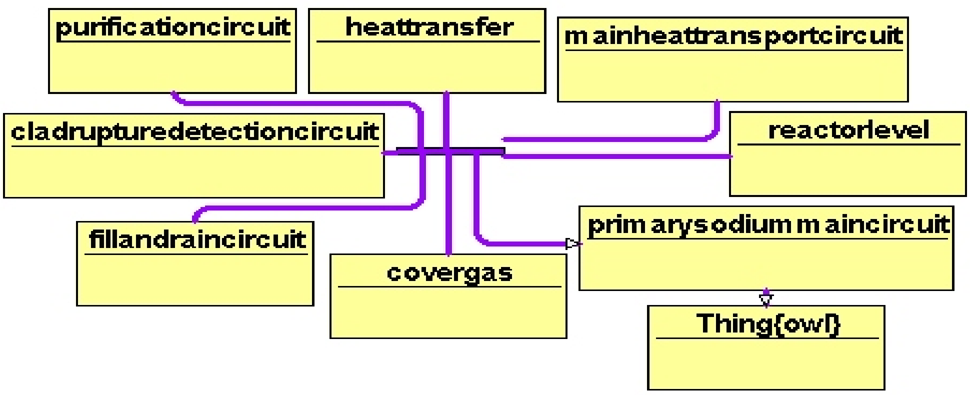


Figure 3.35 UML representations for primary sodium circuit

3.5.13 Secondary Sodium System

Secondary sodium circuit is interposed between radioactive primary sodium circuit and steam water circuit. This system consists of two identical and independent secondary sodium circuits. Each one of the two primary sodium circuits dissipates heat to corresponding secondary sodium circuit in the intermediate heat exchangers. Secondary sodium transfers heat to sodium/water in once through steam generator to produce super heated steam. It acts as a barrier between radioactive primary sodium and steam/water and eliminates the possibility of steam/water leak into primary sodium. It prevents pressure surges caused by sodium water reaction from reaching the reactor core. It serves to remove residual decay heat and also serves to keep the primary sodium hot, in the absence of adequate nuclear heat [200]. OWL and UML representation for secondary sodium circuit are shown in Figure 3.36 and Figure 3.37.

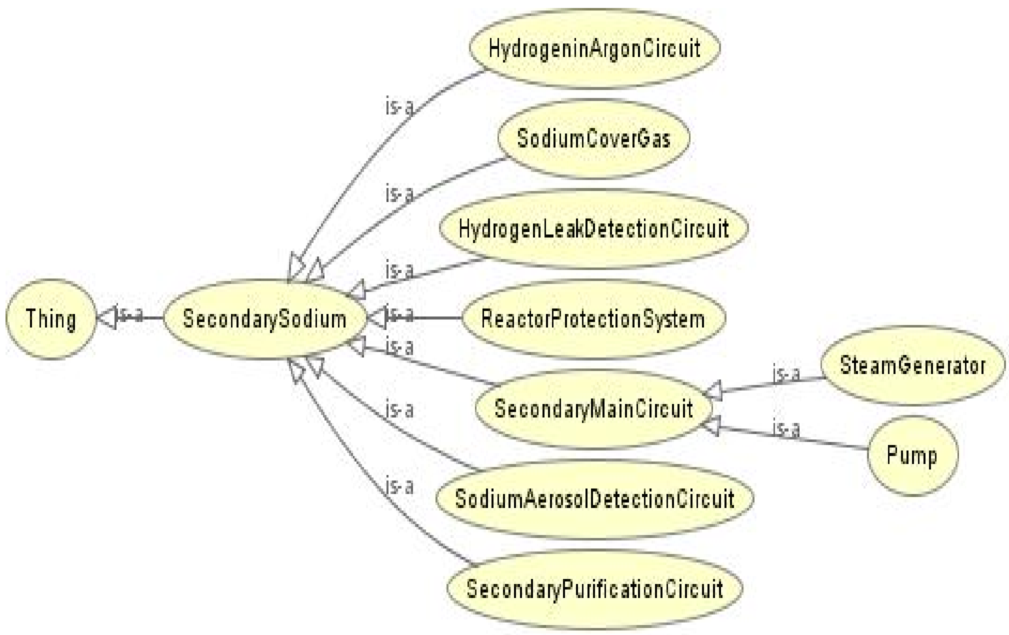


Figure 3.36 OWL representations for secondary sodium circuit

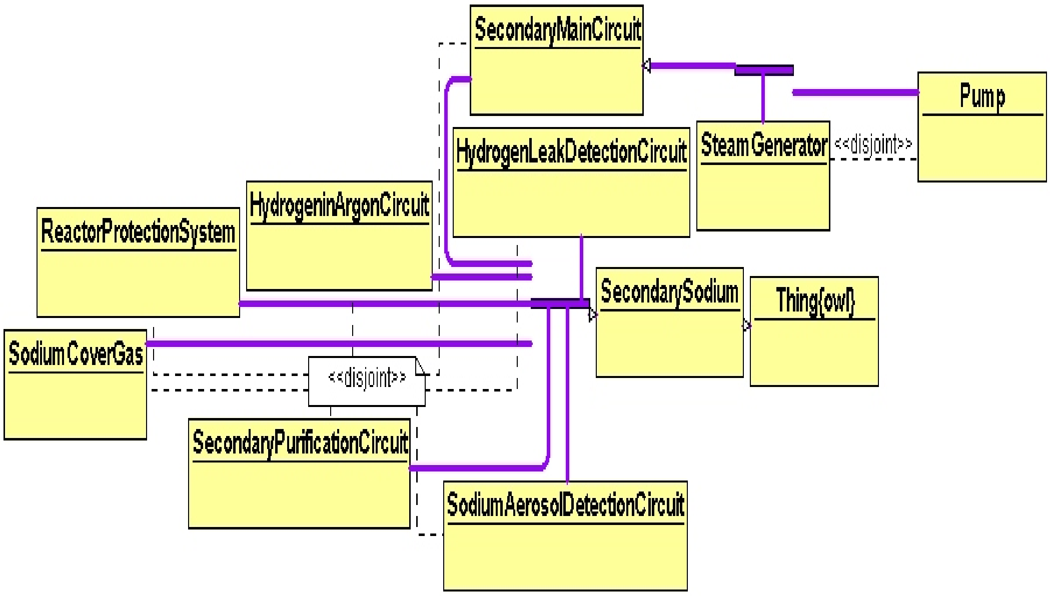


Figure 3.37 UML representations for secondary sodium circuit

3.5.14 Electrical System

Power supply system is divided into class A, class B, class C and class D. FBTR offsite power supply includes two numbers 10 MVA, 33 KV/3.3 KV/6.6 KV incomer transformers. Incomer transformer 1 is fed from 33 KV bus A and incomer 2 is fed from 33 KV bus B. Incomer transformer 1 feeds to 6.6 KV bus A (HTtb100A) and incomer transformer 2 feeds to 6.6 KV bus B (HTtb100B). Normally these two transformers are operating in parallel through bus coupler CBht 110 in 6.6 KV bus. Two numbers 910 HP, 6.6 KV main boiler feed pump motors are fed from 6.6 KV bus. One 19.5 MVA, 16 MW, 6.6 KV alternator is synchronized to grid through breaker CBht 108, when it is operating. It is synchronized via 6.6 KV bus. There are four numbers of 415 V buses, namely, BTsb100, BTsb 200, BTtb 300, BTab 400. All are sectionalized buses. Each section is fed from 1 MVA, 6.6 KV/415 V transformers. Out of these buses BTsb 100 and

BTsb 200 are emergency buses to which two numbers 415 V, 1 MVA Diesel generators are connected to cater to the needs of emergency loads in the event of failure of offsite power supply. OWL and UML representations for electrical system are shown in Figure 3.38 and Figure 3.39.

3.5.15 Neutronic Channels

It consists of instrument channel, control channel, logic channel and protective channel. Instrument Channel: An instrument channel consists of sensor and auxiliary equipment required to generate, transmit and/or display a signal related to the plant parameter monitored by that sensor. Control channel : control channel includes instrument channel and its auxiliary equipment. Control action: is an action initiated by control channel to maintain system parameters within the set operating range. Logic channel: An arrangement of relay contacts actuated by instrument channel and auxiliary equipment, the operation of which will initiate either a control or protective action. Protective channel: includes instrument channel and its logic channel. Protective action: is an action initiated by protective channel when set value is exceeded. Safety action: An action which automatically causes SCRAM or LOR so as to prevent safety limits being exceeded. OWL and UML representation of neutronic channels are shown in Figure 3.40 and Figure 3.41.

3.5.16 Neutronic Instrumentation System

The neutronic instrumentation is to monitor the entire range of neutron flux level encountered from shutdown to full power. It takes safety action in case of any discrepancies by giving necessary inputs to the reactor protection system. OWL and UML representation are shown in Figure 3.42 and Figure 3.43.

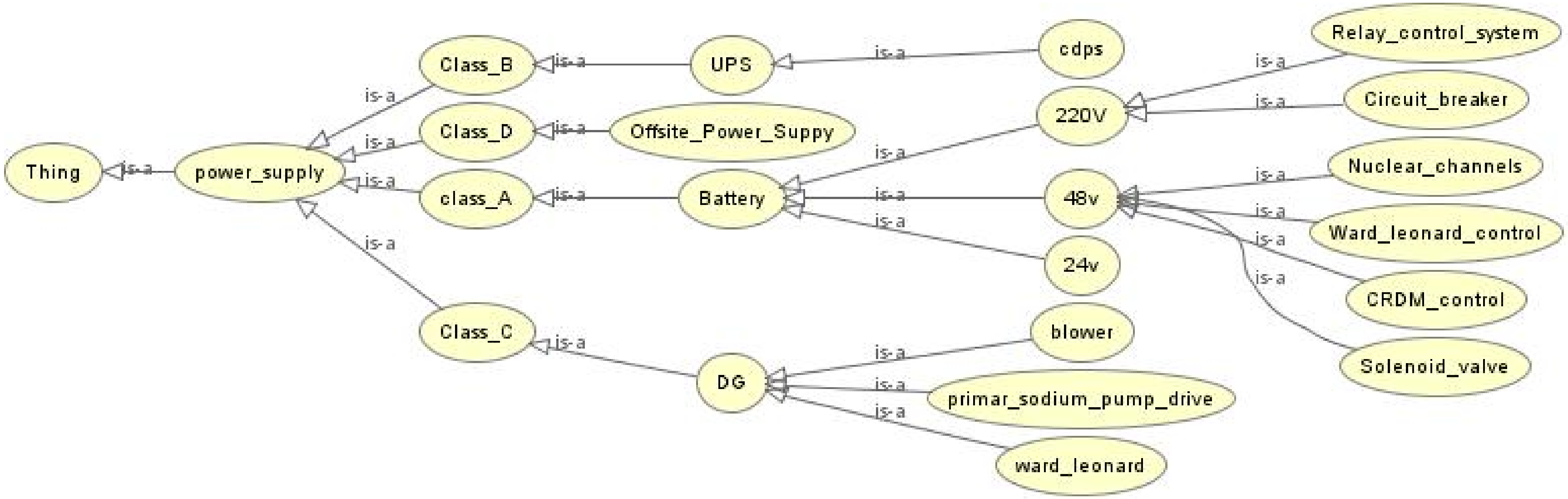


Figure 3.38 OWL representations for electrical system

145

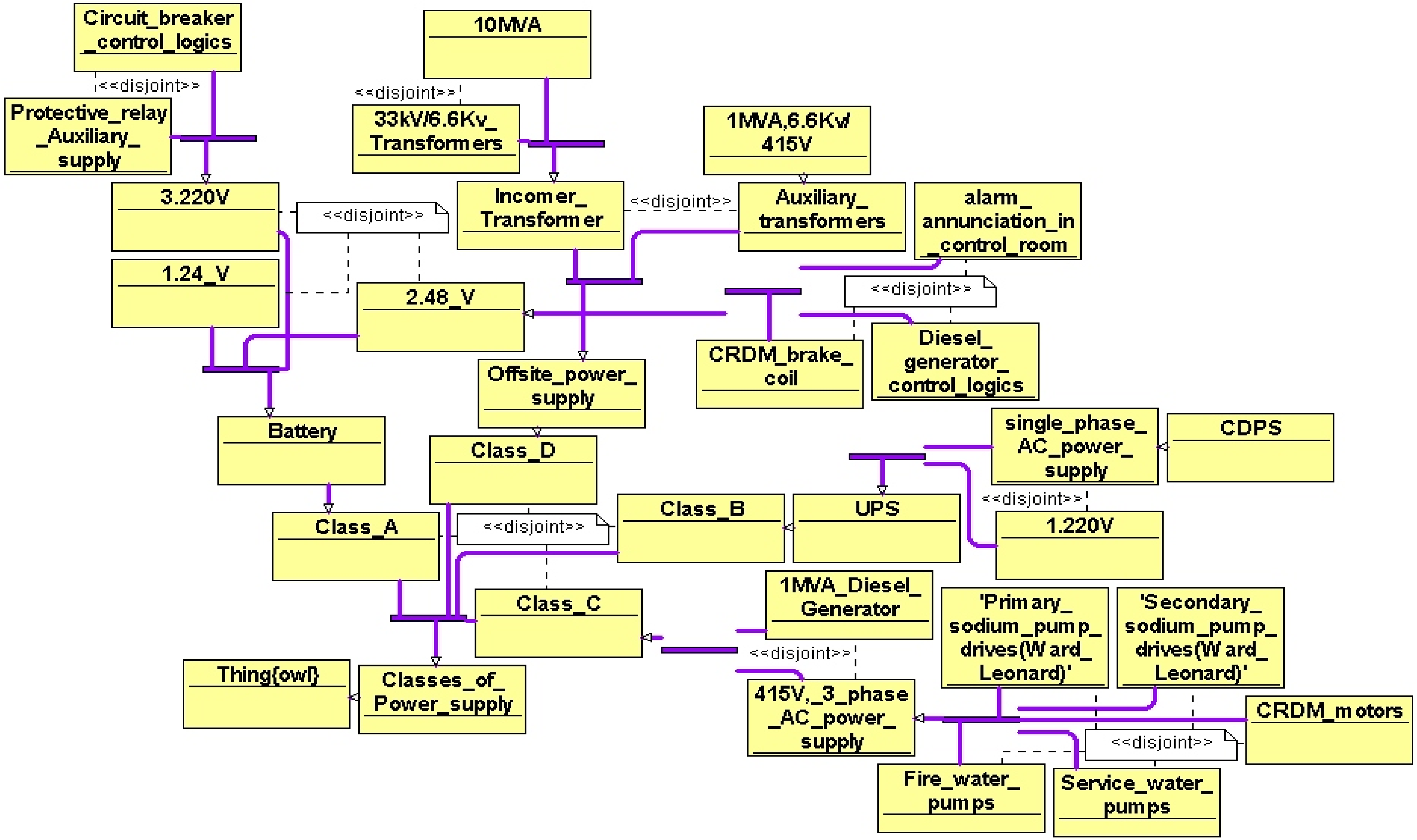


Figure 3.39 UML representations for electrical system

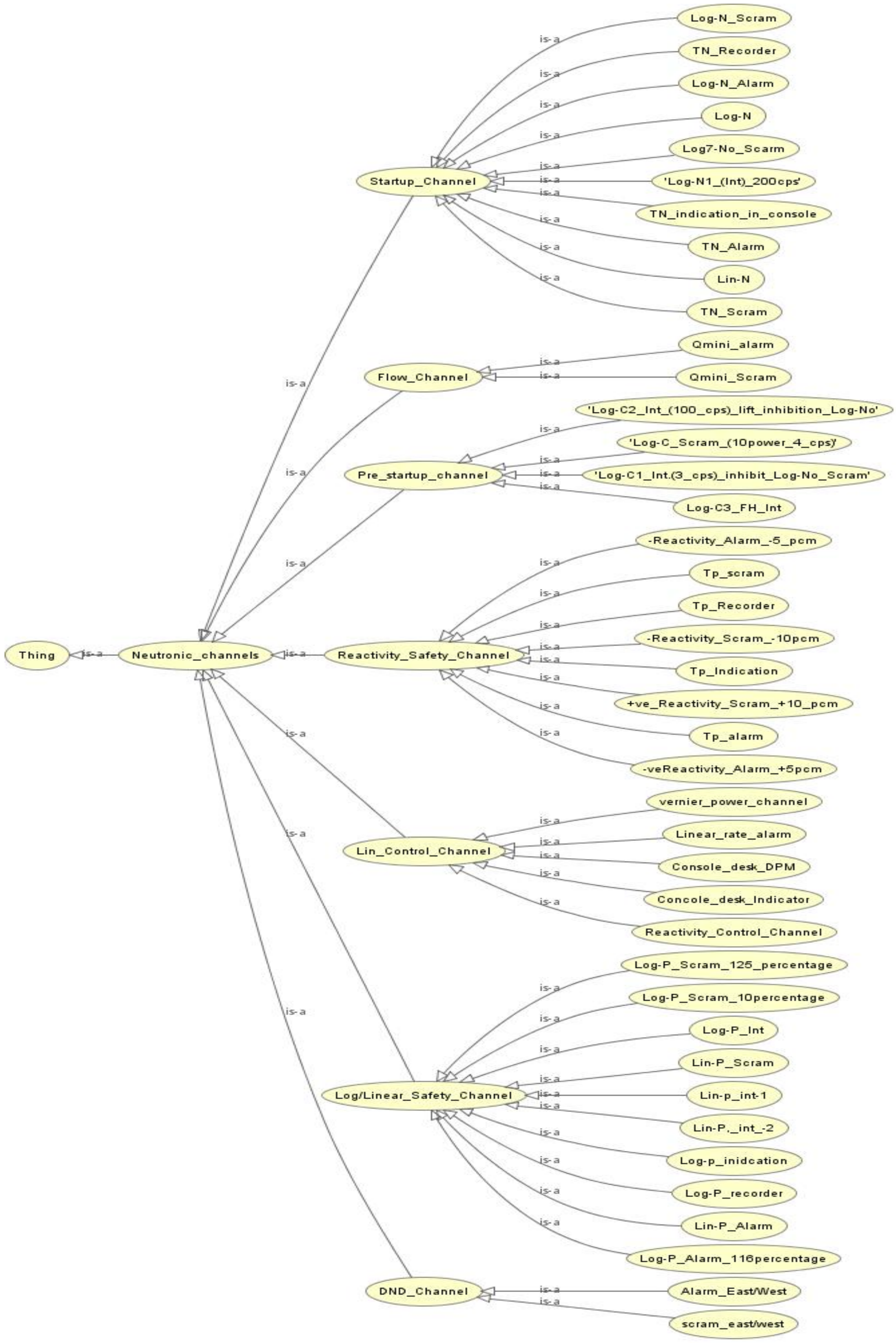


Figure 3.40 OWL representations for neutronic channel

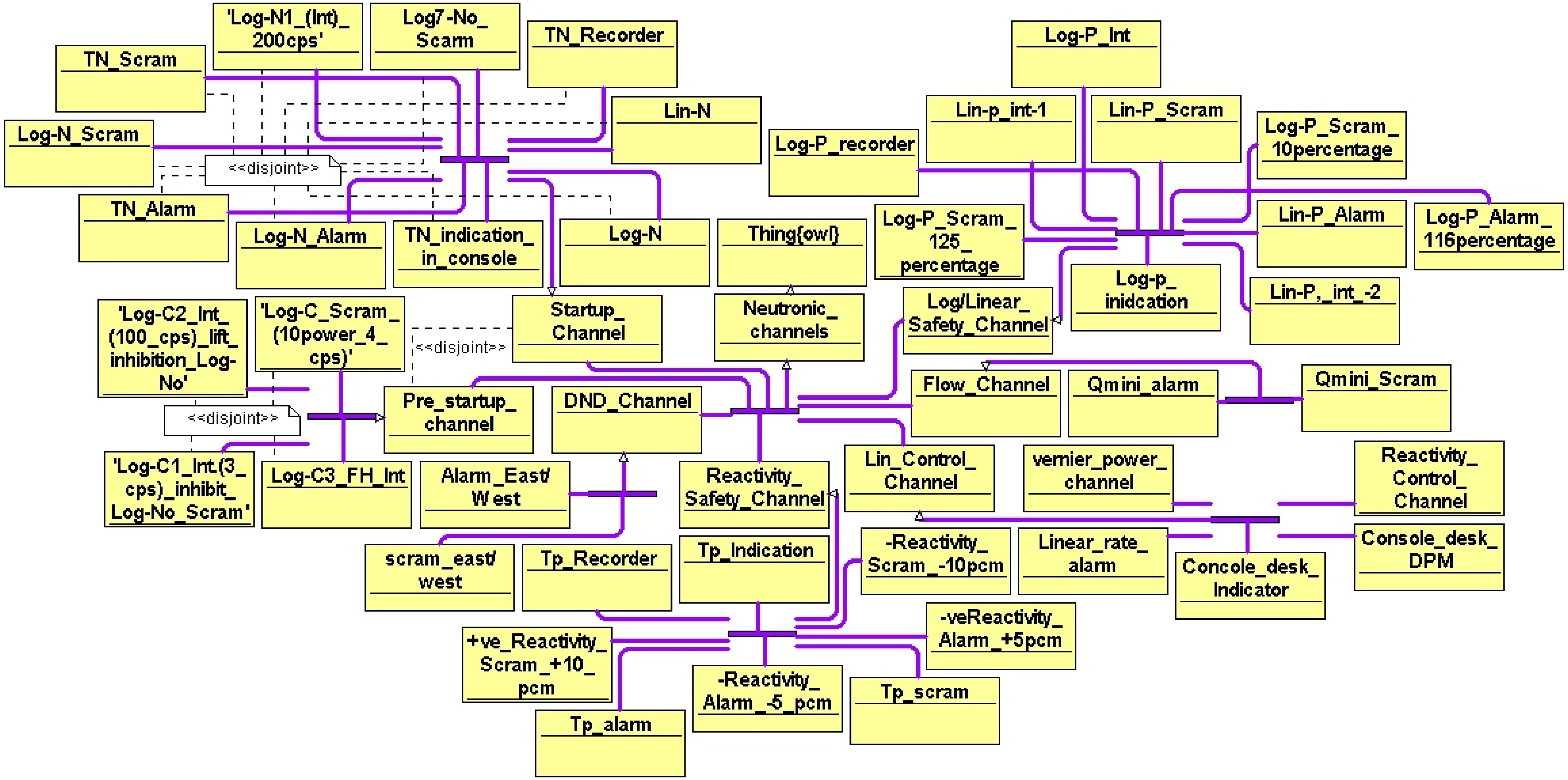


Figure 3.41 UML representations for neutronic channel

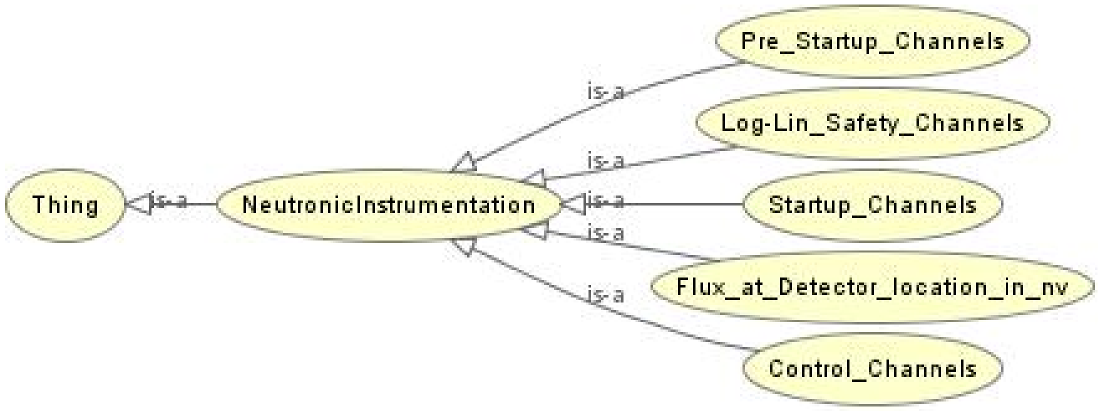


Figure 3.42 OWL representations for neutronic instrumentation system

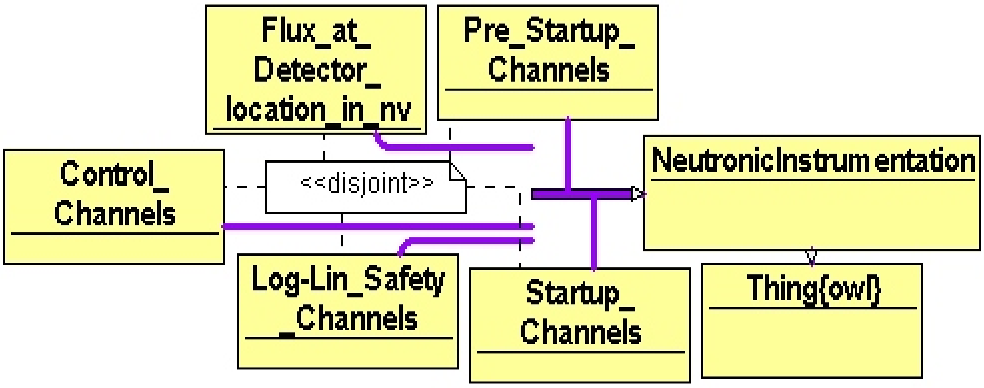


Figure 3.43 UML representations for neutronic instrumentation system

3.5.17 Steam Water System

The function of the steam water system is to extract the nuclear heat through primary, secondary sodium system and convert the feed water into superheated steam during power and to remove decay heat of the reactor core during normal and emergency shutdown states of the plant [201]. OWL and UML representation for steam water system are shown in Figure 3.44 and Figure 3.45.

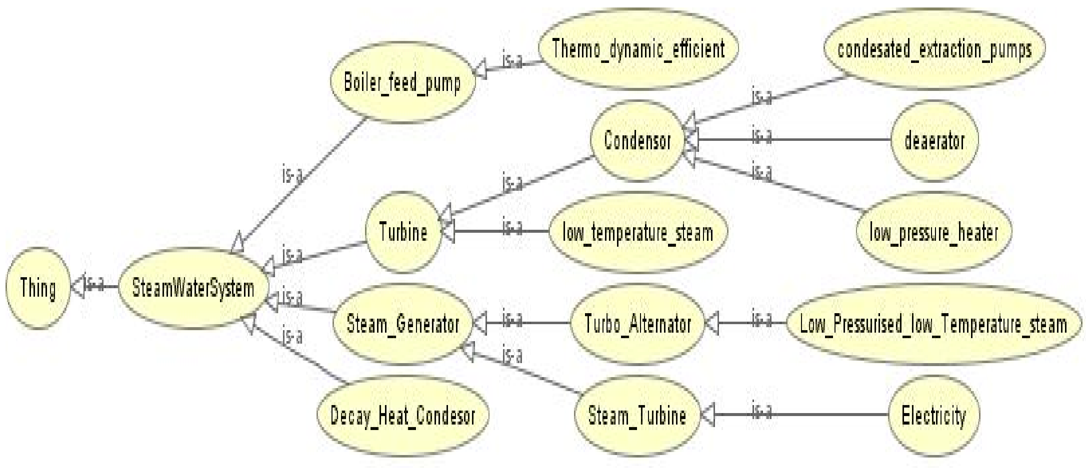


Figure 3.44 OWL representations for steam water system

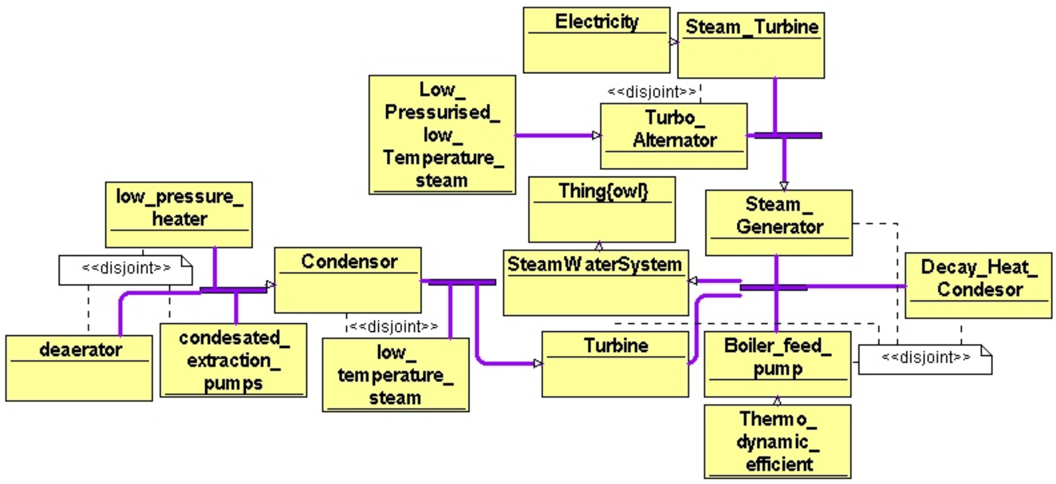


Figure 3.45 UML representations for steam water system

3.5.18 Lowering of Rods (LOR) Parameter

LOR causes a gradual shutdown of the reactor by driving down simultaneously all the six control rods with the help of drive motors. The input signals for LOR are primary loop flow difference high, non availability of dump system, feed water pump tripping, main line steam pressure being high, mean core outlet temperature being high, mean differential temperature across the core being high, primary sodium pump of east loop being out of service, primary sodium pump of west loop being out of service, secondary sodium pump of east loop being out of service, secondary sodium pump of west loop goes out of service, leaking of water into sodium system, manual actuation of LOR. OWL and UML representation for LOR Parameter are shown in Figure 3.46 and Figure 3.47.

3.5.19 Safety Control Rod Activation Mechanism (SCRAM) Parameter

SCRAM circuit causes fast shutdown of reactor by dropping all the control rods into the reactor by de-energising the electromagnet which holds them to the drive out in their position. Inputs for SCRAM are neutron flux high, fast period, high neutron power, high reactivity change in the core, actual core subassembly outlet temperature above calculated outlet temperature for that power, mean differential temperature across the core high, LOR order not carried out due to absence of 415V power supply, delayed neutron in the primary loop high due to rupture of fuel cladding and manual scram. OWL and UML representation for scram are shown in Figure 3.48 and Figure 3.49.

3.5.20 Reactor Shielding

The main purpose of shielding is to protect the people, equipments and structures from harmful effects of radiation. Shield design for reactor involves invessel shielding, bulk shield for reactor vault and top shield, reactor control building cell shielding, shielding for shipping cask, fresh fuel and spent fuel storage bay, complementary shielding for components penetrating top shield. Shields are designed based on shield design criteria specified based on atomic energy regulatory board criteria. OWL and UML representation for reactor shielding are shown in Figure 3.50 and Figure 3.51.

Figure 3.46 OWL representations for LOR parameter

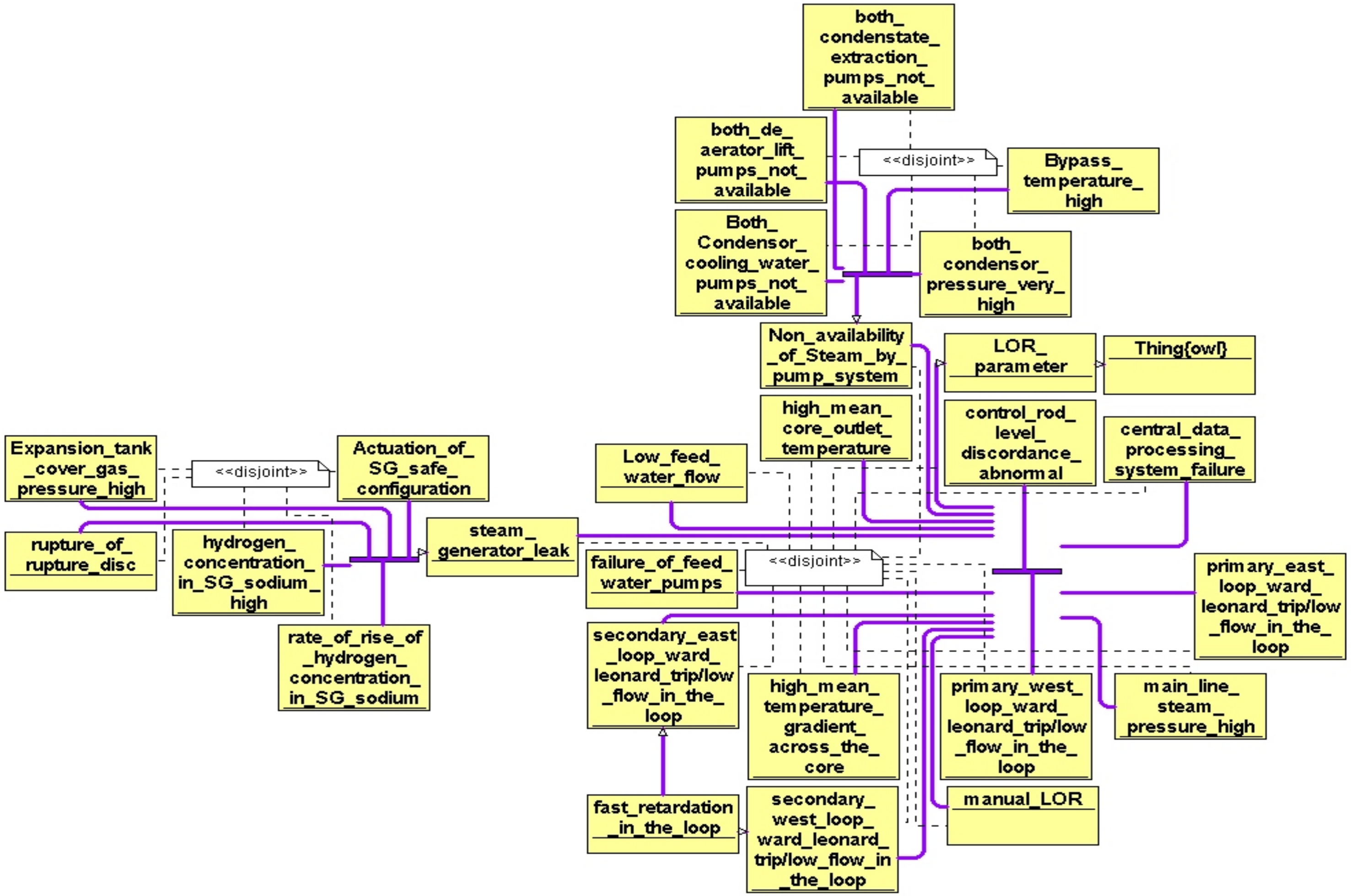


Figure 3.47 UML representations for LOR parameter

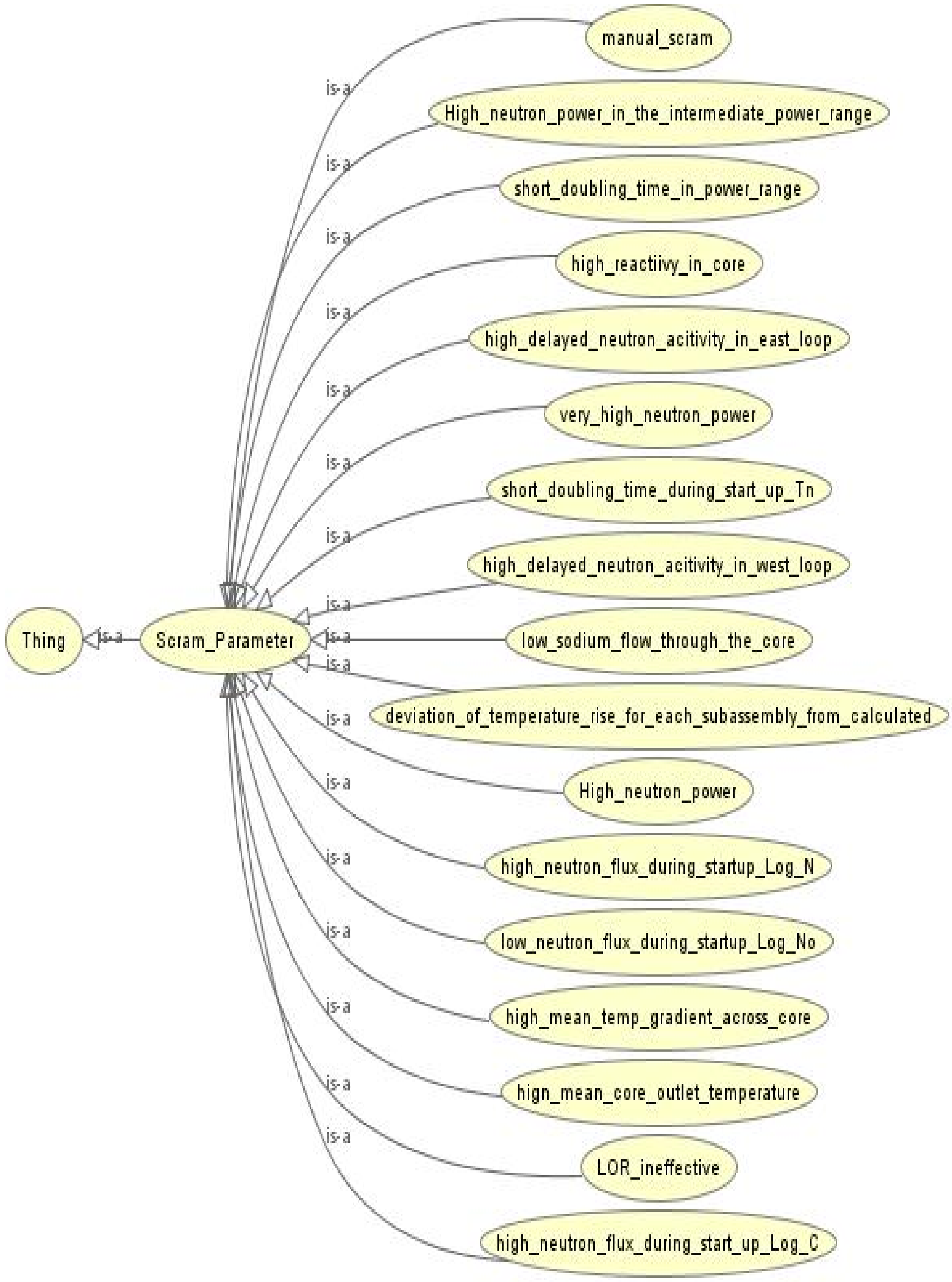


Figure 3.48 OWL representations for SCRAM parameter

154

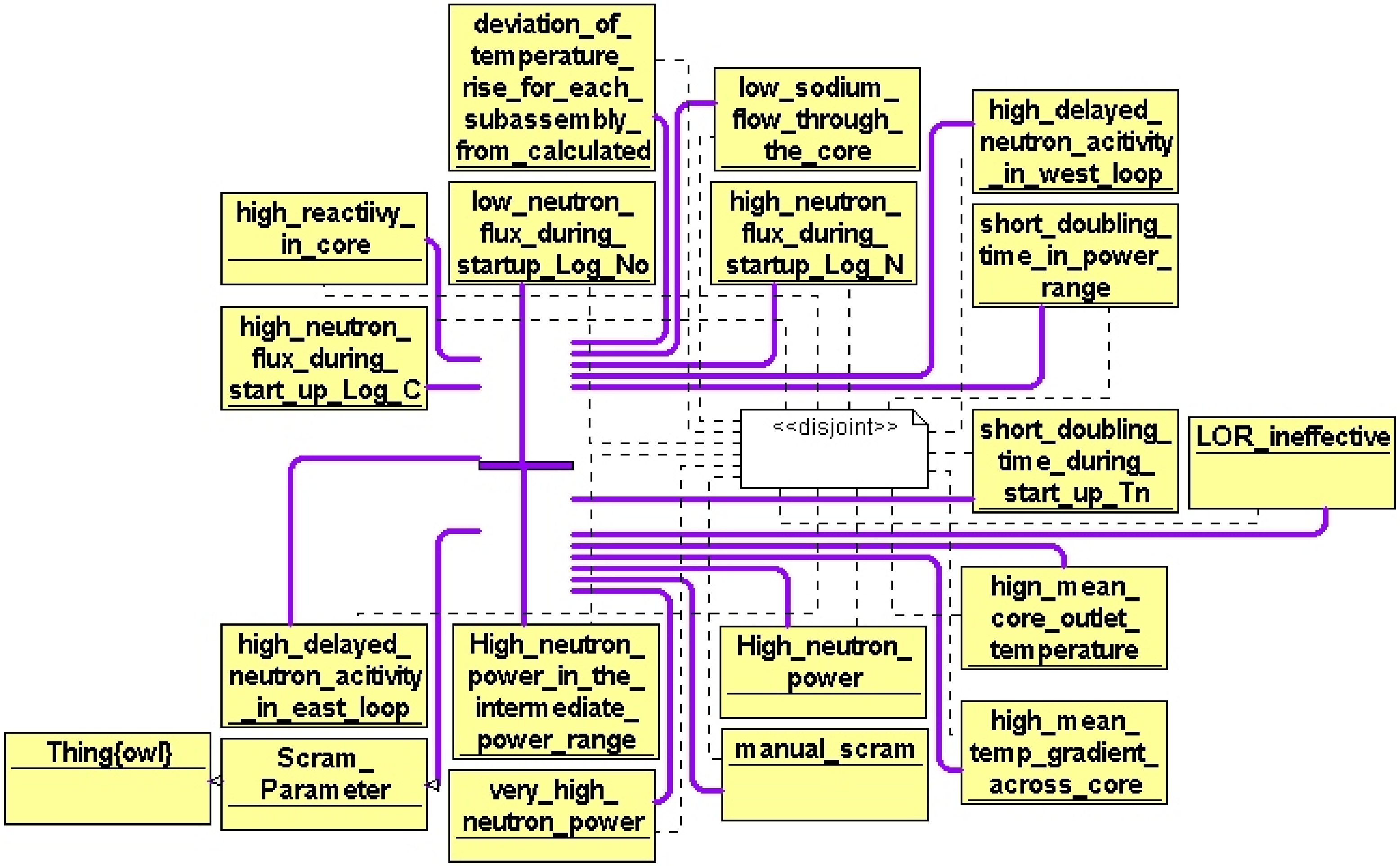


Figure 3.49 UML Representations for SCRAM parameter

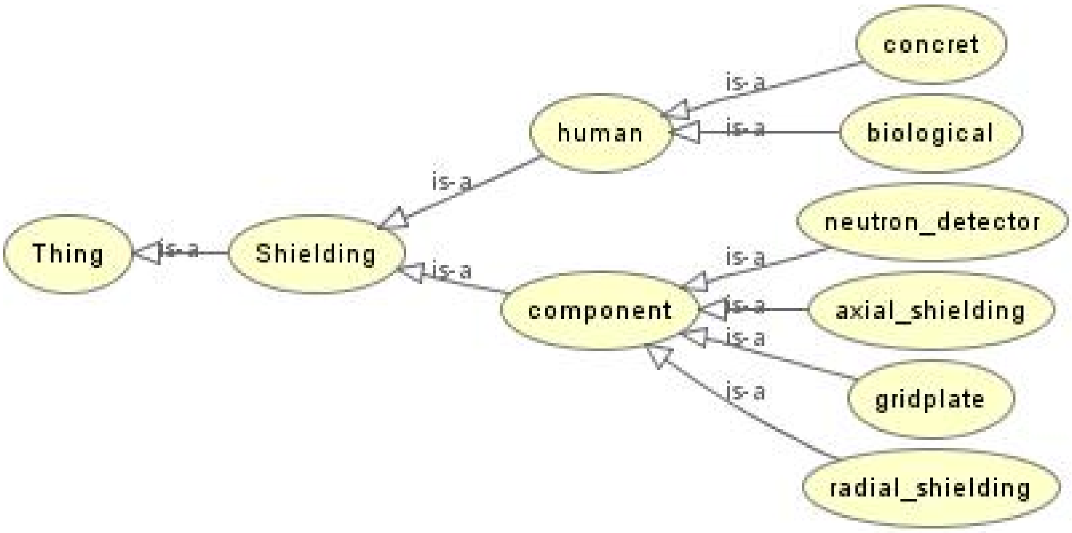


Figure 3.50 OWL representations for reactor shielding

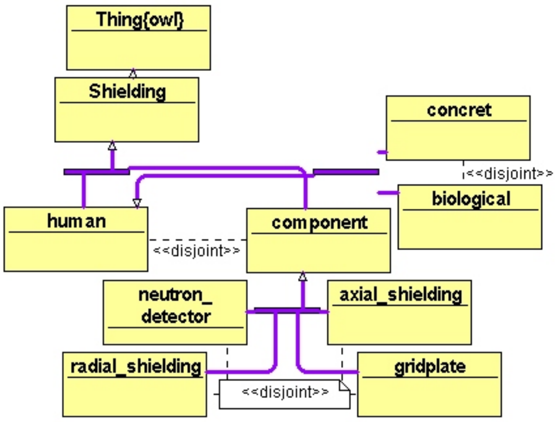


Figure 3.51 UML representations for reactor shielding

3.5.21 Shielding Materials

Hydrogenous materials like polyethylene, water are good neutron shield materials. Boron carbide is a good absorber of thermal neutrons. Steel and concrete also act as good neutron shields. Materials with high density like lead and steel are preferred as gamma shielding materials. Concrete acts as neutron as well as gamma shield. OWL and UML representation for shielding materials are shown in Figure 3.52 and Figure 3.53.

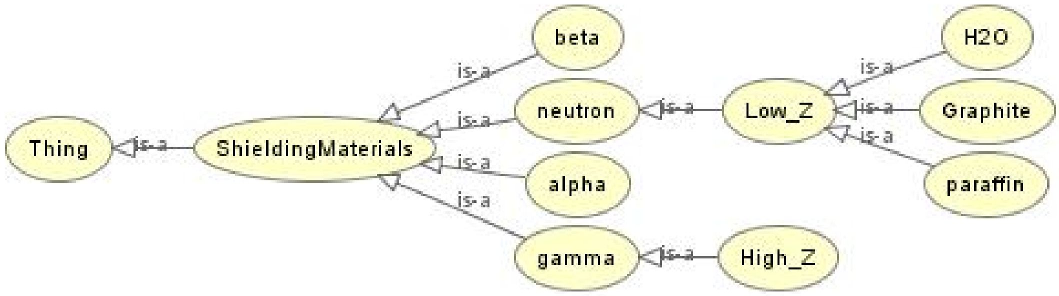


Figure 3.52 OWL representations for shielding materials

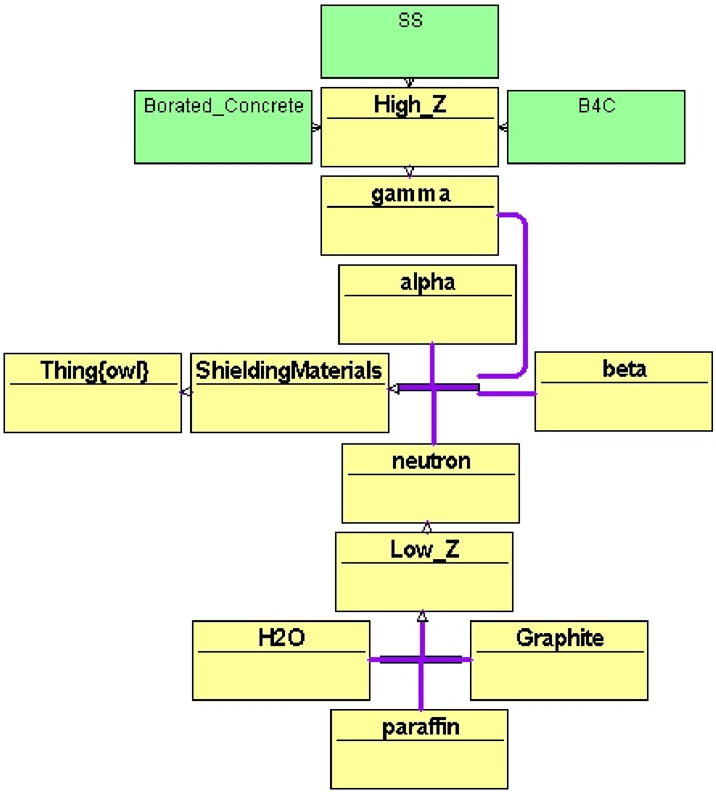


Figure 3.53 UML representations for shielding materials

3.5.22 Shield within Reactor

The invessel radial and axial shields are optimised based on secondary sodium activity in IHX, neutron flux at the control plug detector location, irradiation damage to invessel components, helium production in high temperature components and gamma heating in radial shields. The optimised axial shield consists of 65.5 cm SS followed by 10 cm B_4C. OWL and UML representation for shield within reactor are shown in Figure 3.54 and Figure 3.55.

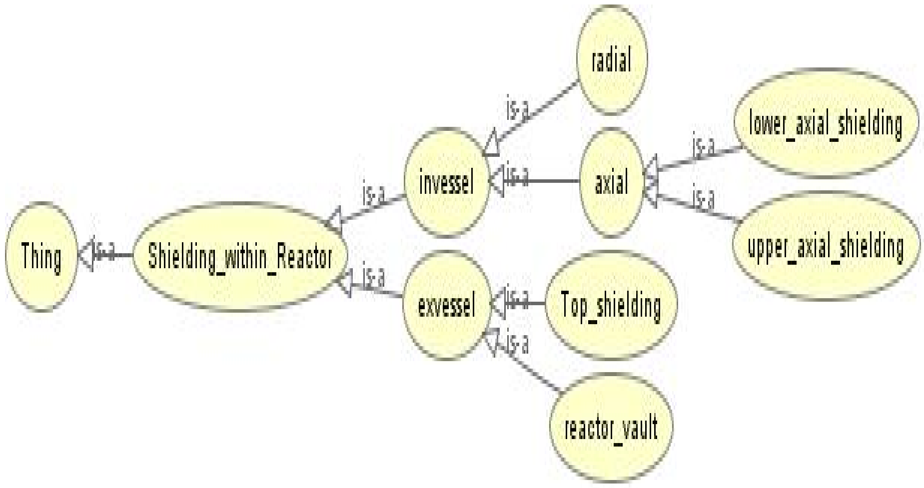


Figure 3.54 OWL representations for shield within reactor

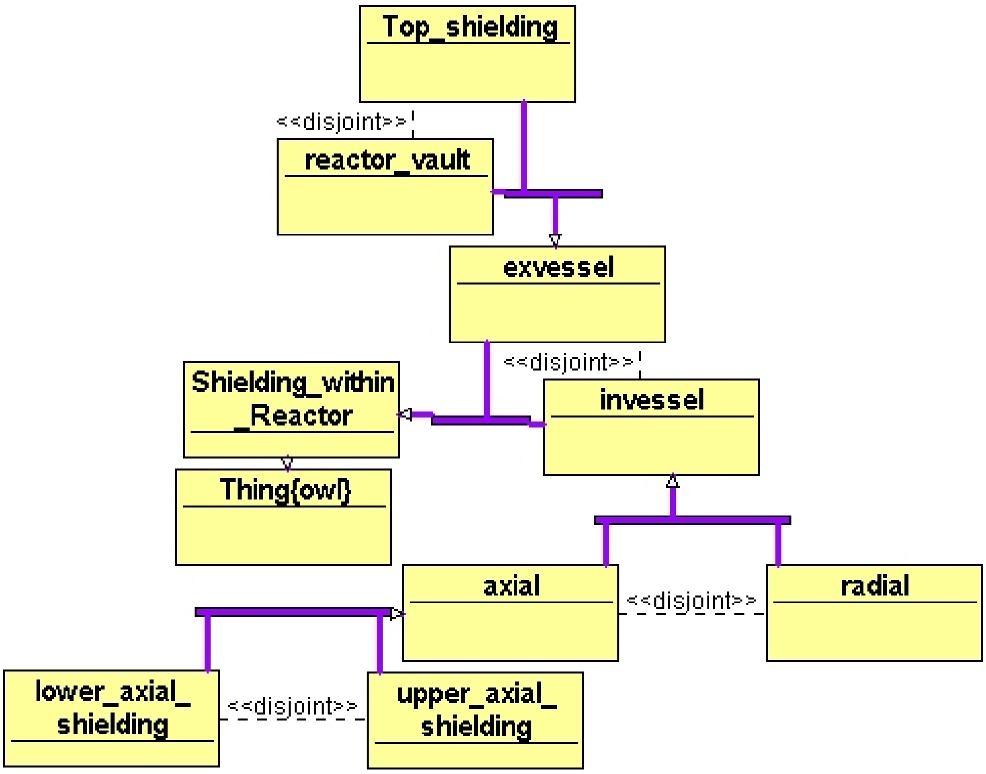


Figure 3.55 UML representations for shield within reactor

3.5.23 Reactor Fuel handling system

Fuel handling system of FBTR serves the purpose of loading the reactor with fresh fuel sub assemblies (SA) and discharging the burnt fuel SA from the reactor. The fuel handling cycle consists of new sub assembly receipt, inspection and storage of new sub assemblies in active building. Access to different subassemblies in the reactor is through the fuel handling canal in the small rotating plug. The subassemblies are loaded into the reactor one by one using a charging flask. Discharging of spent fuel from the reactor is by discharging flask and loading it in transfer pot in discharge pit area. Transportation of the transfer pot containing the spent fuel to irradiated element storage area is done by using secondary flask. OWL and UML representation for reactor start up conditions are shown in Figure 3.56 and Figure 3.57.

3.5.24 Permissive Circuits

The permissive circuits for rector startup and start up of fuel handling get inputs from primary sodium and reactor assembly. The permissive circuits for reactor start up gives authorization to raise the control rod. The fuel handling permissive circuit gives authorization for rotation of the rotating plugs. OWL and UML representation for reactor permissive circuits are shown in Figure 3.58 and Figure 3.59.

159

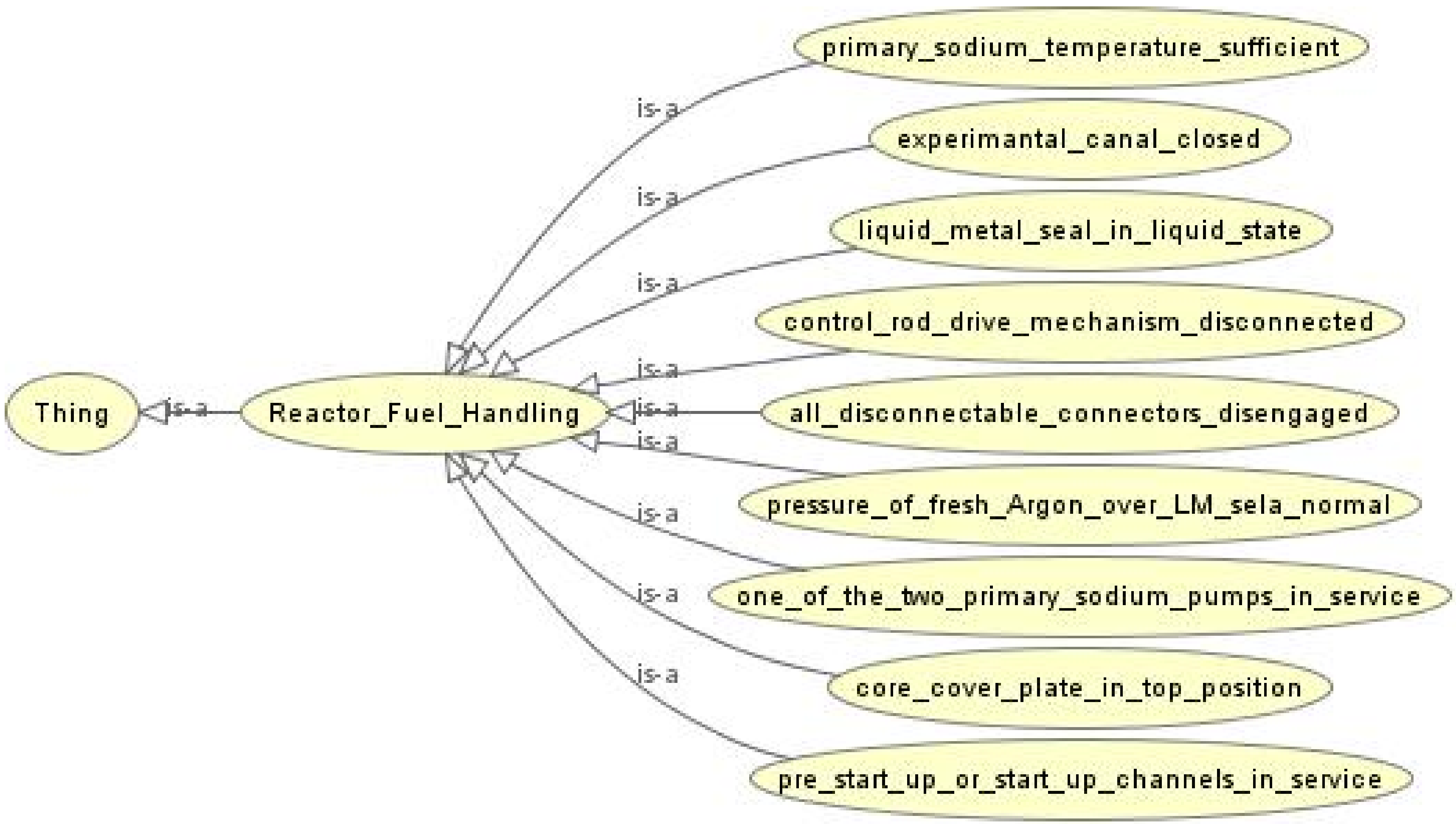


Figure 3.56 OWL representations for reactor fuel handling system

160

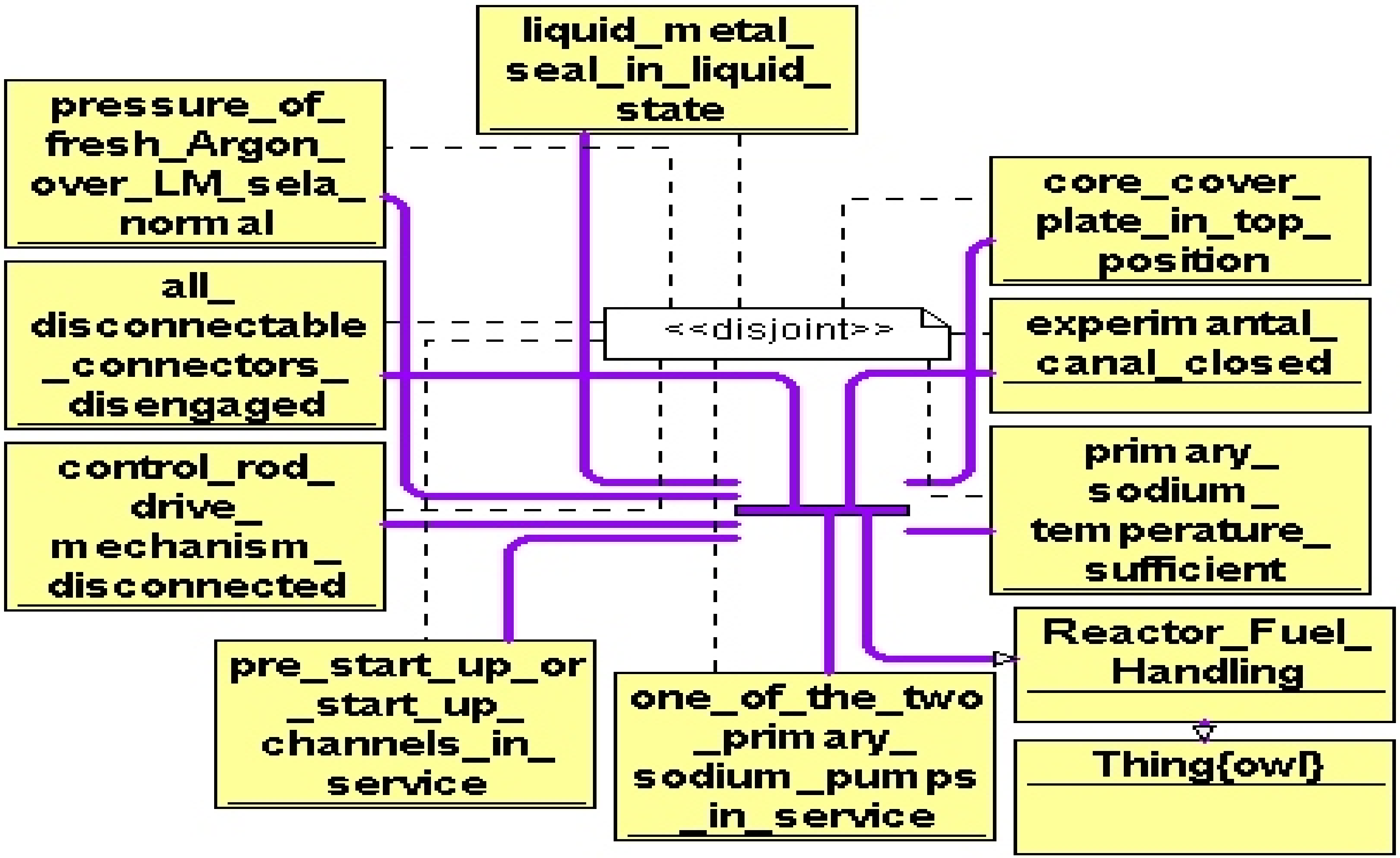


Figure 3.57 UML representations for reactor fuel handling system

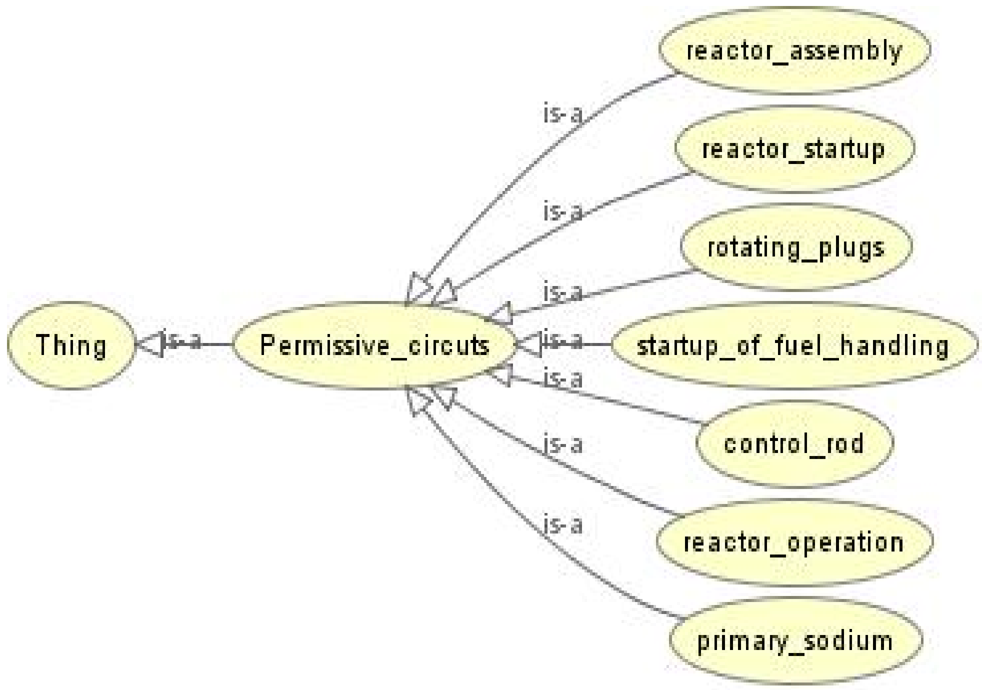


Figure 3.58 OWL representations for reactor permissive circuits

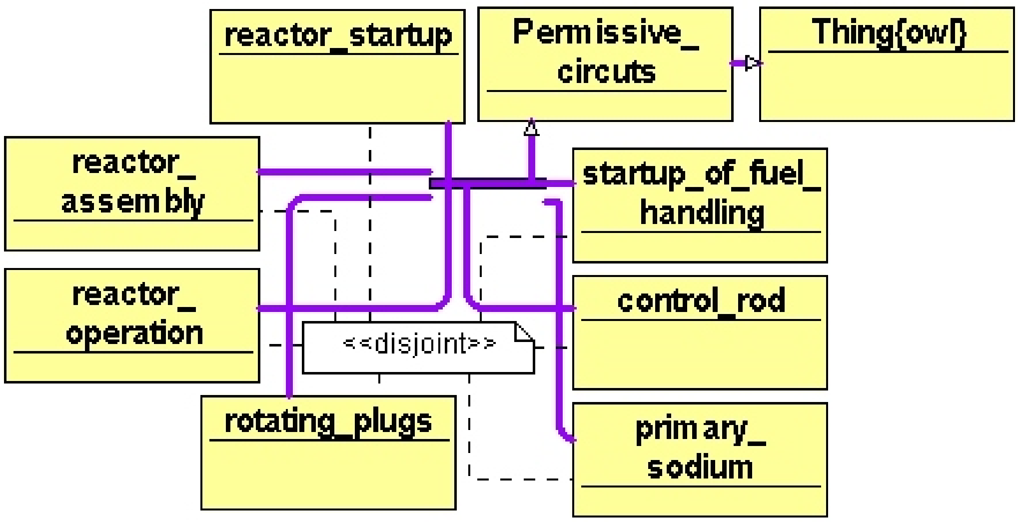


Figure 3.59 UML representations for reactor permissive circuits

3.5.25 Reactor Shutdown Mechanism

Reactor is said to be in shutdown state when all six control rods are fully inserted in the core. FBTR has two types of shutdown systems based on fault severity. LOR (Lowering of Control Rods) will take place when there is any thermal parameter faults. SCRAM order will take place when there are any Neutronic parameters faults.

Lowering **o**f control **R**ods (LOR): During LOR, simultaneous lowering of all six control rods is done at 1mm/s up to their bottom most position.

Safety **C**ontrol **R**od **A**ctuating **M**echanism (SCRAM): During SCRAM automatic release and drop of all six control rods under gravity takes place in response to certain logically analyzed signals. OWL and UML representation for reactor shutdown mechanism are shown in Figure 3.60 and Figure 3.61.

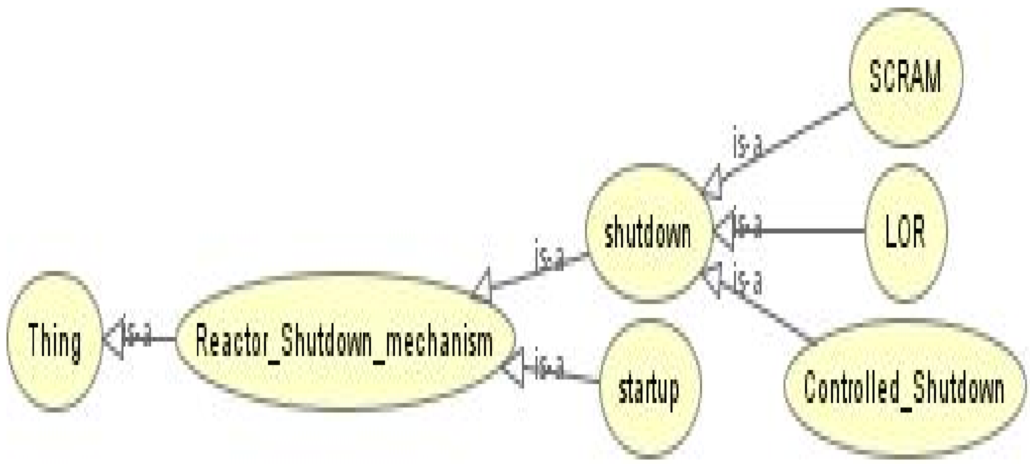


Figure 3.60 OWL representations for reactor shutdown mechanism

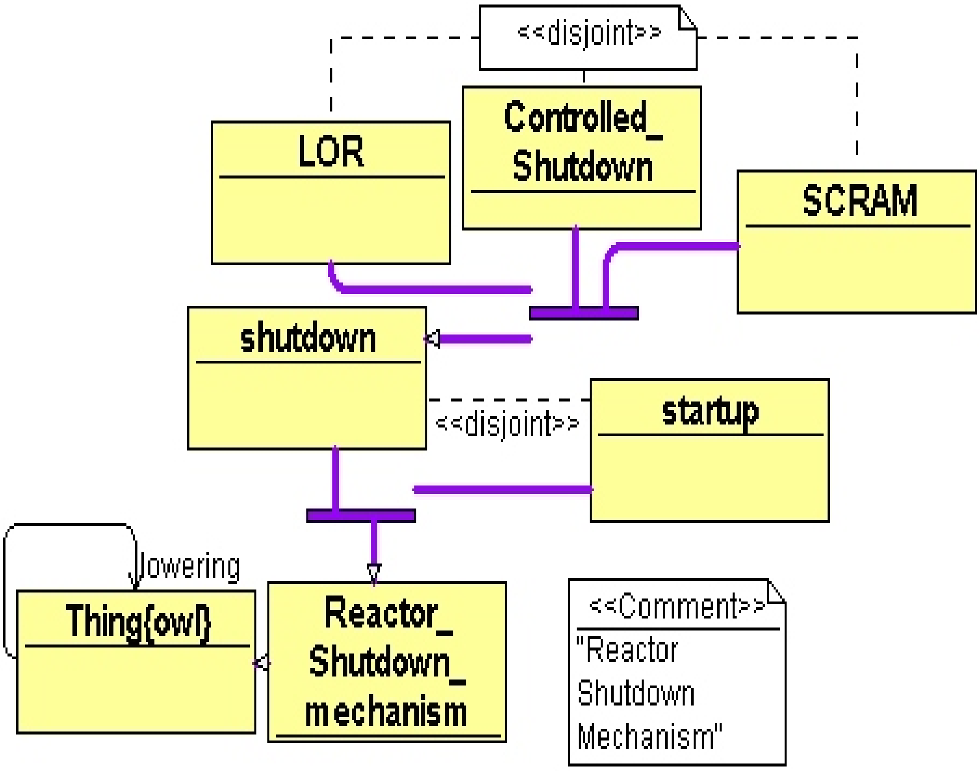


Figure 3.61 UML representations for reactor shutdown mechanism

3.5.26 Control Rod Drive Mechanism

The functionality of control rod drive mechanism is to facilitate raising and lowering of the control rods. During an abnormal condition, all the six rods are lowered simultaneously (LOR) to cause partial or full shutdown of the reactor. An automatic fast shutdown is achieved by dropping all the six rods simultaneously into the reactor in the event of a fault condition in the reactor (SCRAM) [202]. OWL and UML representation for control rod drive mechanism are shown in Figure 3.62 and Figure 3.63.

3.5.27 Reactor State

Reactor in Shutdown State: The reactor is said to be in shutdown state when all six control rods are remain fully inserted in the core.

Start up of the Reactor: The reactor is said to be in this state when all the 27 sub conditions are satisfied and sodium inlet temperature is 180^0C with all six control rods are inserted in the core.

Reactor operation state: Is defined as the reactor is operating at any steady power level within the limits and conditions stipulated in technical specifications.

Reactor in Fuel Handling State: when the core sodium inlet temperature is 180 ^{o}C with only one primary pump and corresponding secondary pump running. All control rods are fully deposited in the core and other necessary operations are carried out on the reactor assembly to facilitate rotation of plugs to carry out handling of subassemblies in the core. OWL and UML representation for reactor state are shown in Figure 3.64 and Figure 3.65.

3.5.28 Reactor Start Up Condition

Before starting the reactor for operation, a check list of system functioning and health are ensured. When all the parameters required for the startup condition are satisfied, then the reactor is ready for start up for operation. OWL and UML representation for reactor start up condition are shown in Figure 3.66 and Figure 3.67.

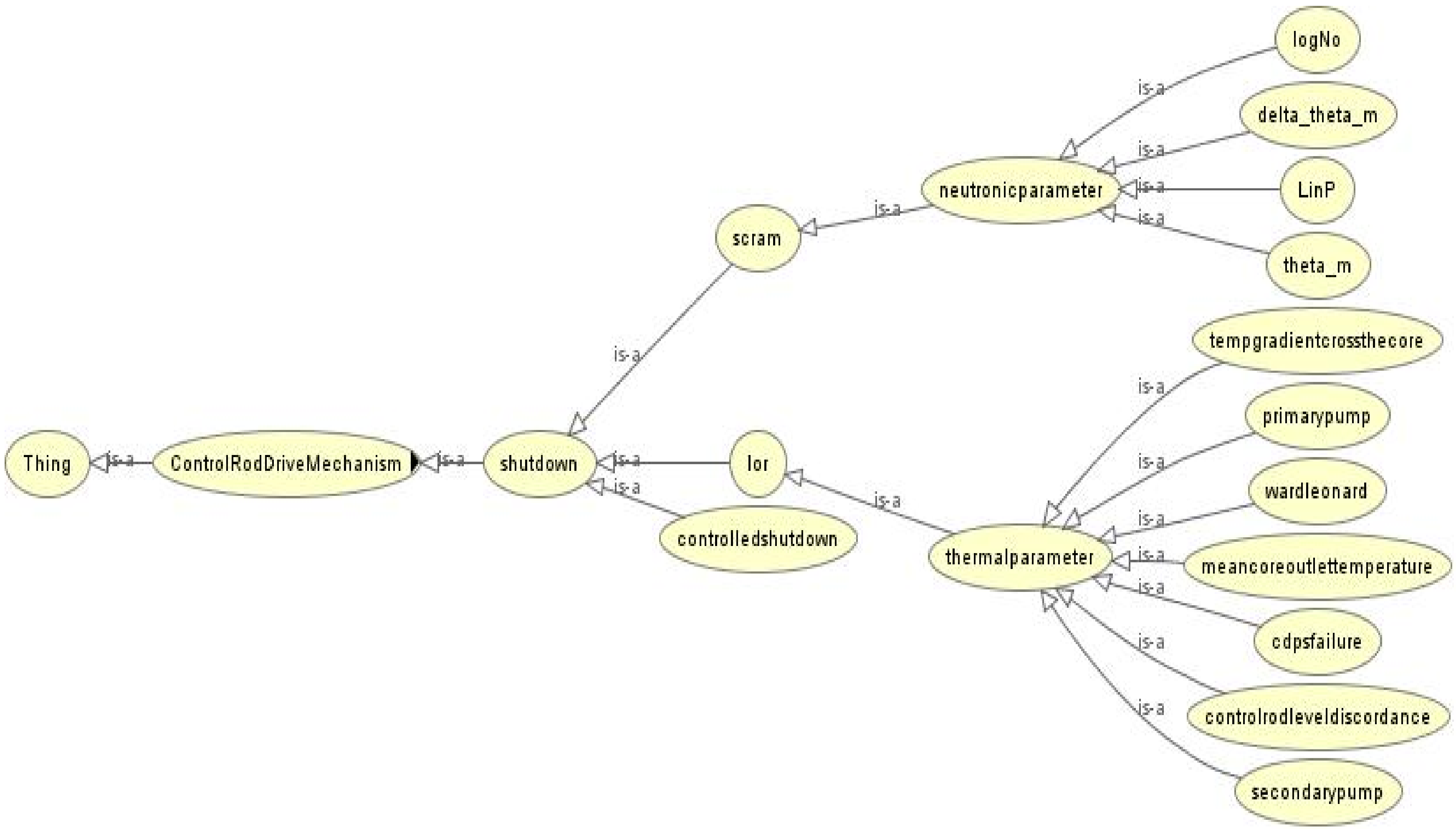


Figure 3.62 OWL representations for control rod drive mechanism

165

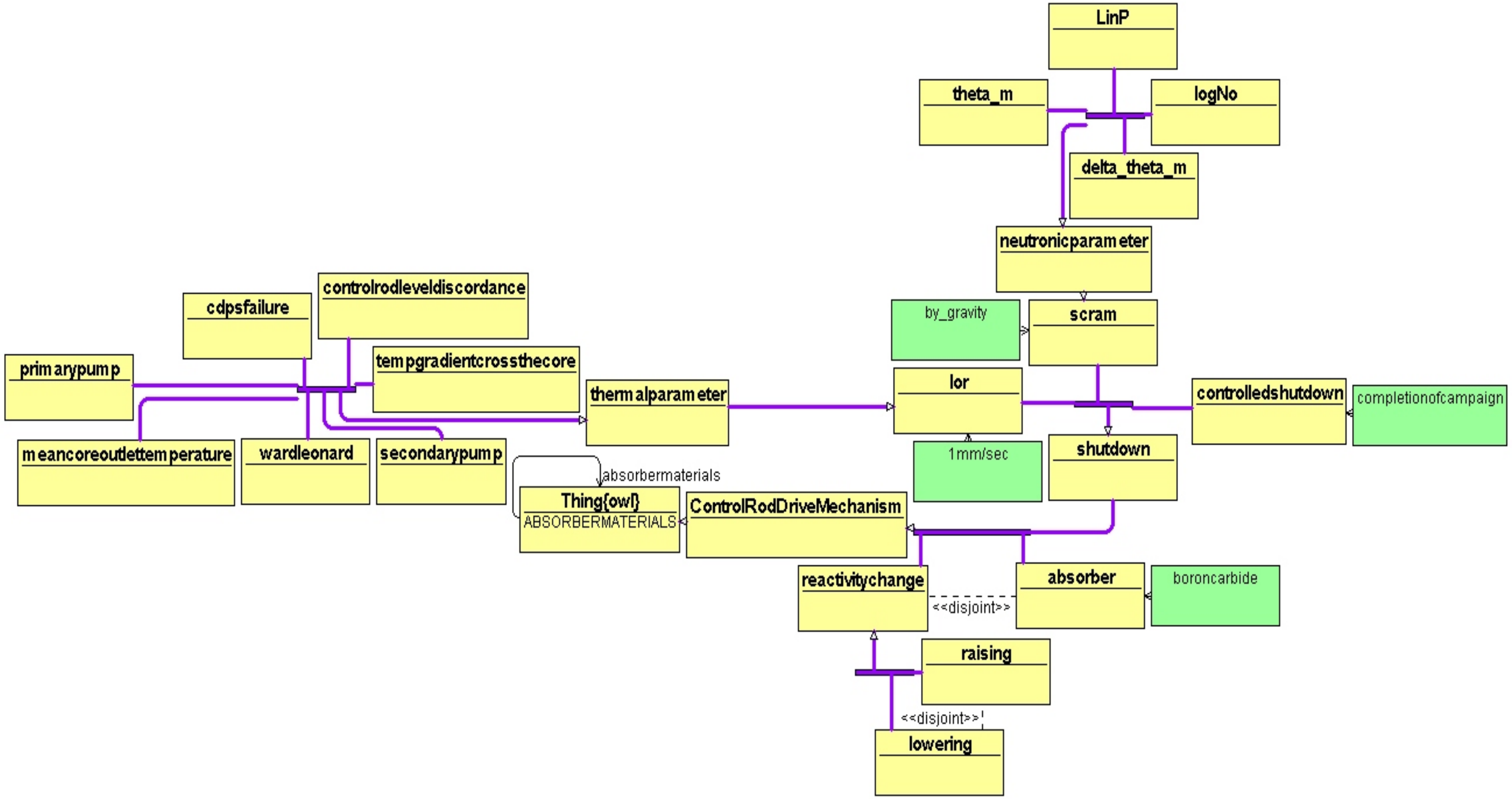


Figure 3.63 UML representations for control rod drive mechanism

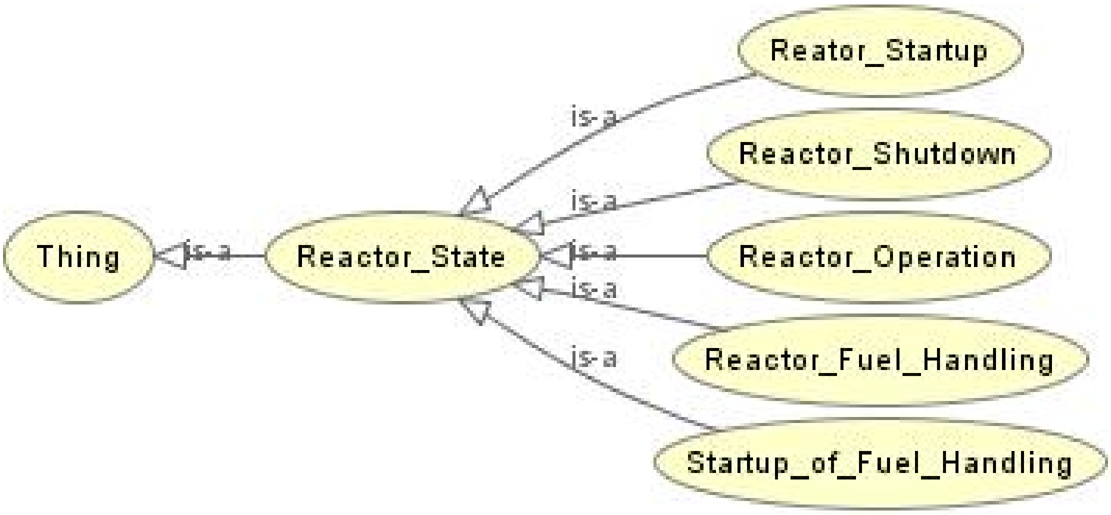


Figure 3.64 OWL representations for reactor state

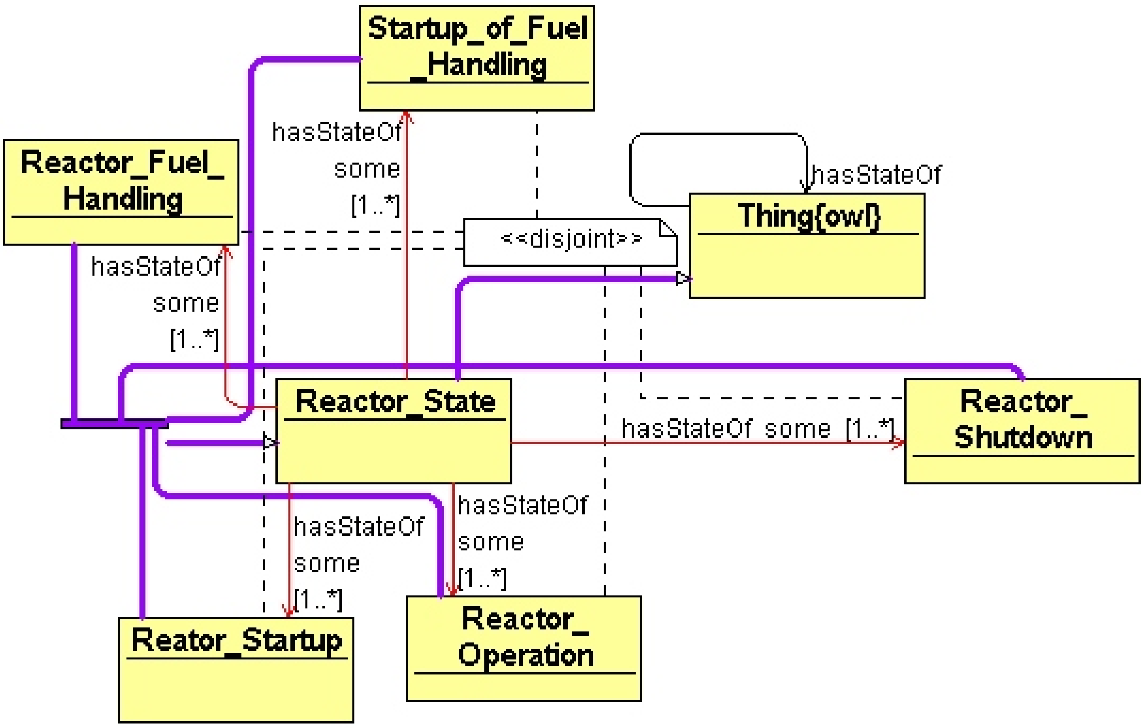


Figure 3.65 UML representations for reactor state

Figure 3.66 OWL representations for reactor start up condition

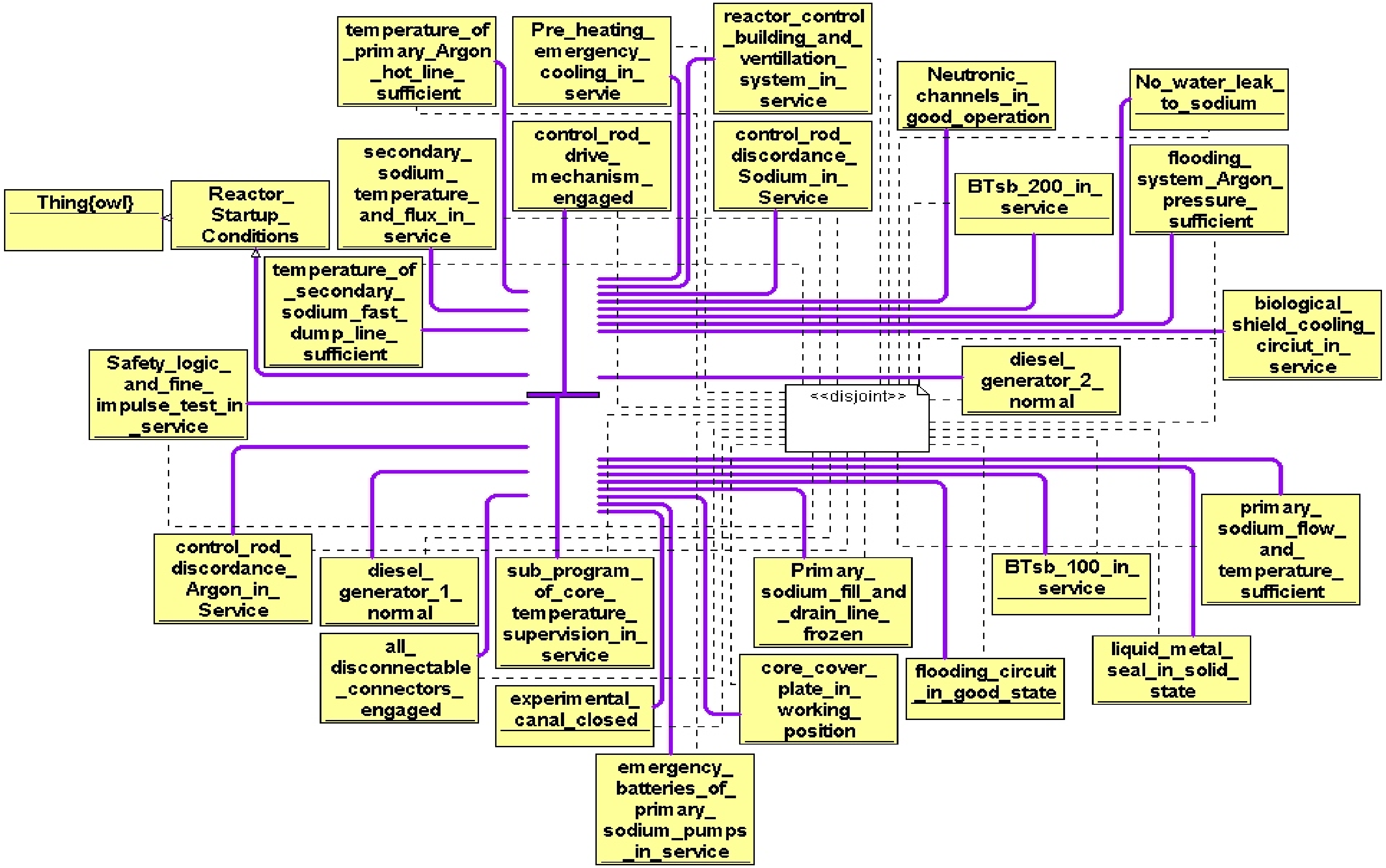


Figure 3.67 UML representations for reactor start up condition

3.5.29 Nuclear Reactor Transient state

Any perturbation in the steady state reactor could initiate transients. Plant protective system could counter act the transients and will bring the reactor to a safe shut down. Though the transients could not be entirely eliminated there are feedbacks available to mitigate the transients and take the reactor to safe conditions. In a typical liquid metal fast breeder reactor unprotected transients are broadly classified as the **U**nprotected **T**ransient **O**ver **P**ower (UTOP) and **U**nprotected **L**oss **O**f **F**low (ULOF). For a given perturbation in external reactivity, that is, for UTOP transients initiated at high power, there are expansion feedbacks available to mitigate the transients. If the total feedback reactivity is good enough to over compensate the external perturbation, net reactivity will become zero, then the reactor will go to another steady state. If the external reactivity is high and if it is not compensated with feedbacks, then net reactivity will become positive. Positive reactivity results in power rise. Subsequently there will be fuel temperature rise, which will lead to melting of fuel. Molten fuel may undergo extrusion or slumping based on the melt fraction and burn up. Dispersion of fuel due to fuel extrusion gives negative reactivity feedback and slumping gives positive reactivity feedback and take the reactor to **C**ore **D**isruptive **A**ccident (CDA).

In case of ULOF, the coolant flow reduces with time, which results in coolant temperature rise. Coolant temperature rise and reduction in coolant density introduces positive reactivity. Further reduction in coolant flow initiates voiding and introduces positive reactivity. This positive voiding reactivity feedback leads to CDA. OWL and UML representation for nuclear reactor transients state are shown in Figure 3.68 and Figure 3.69.

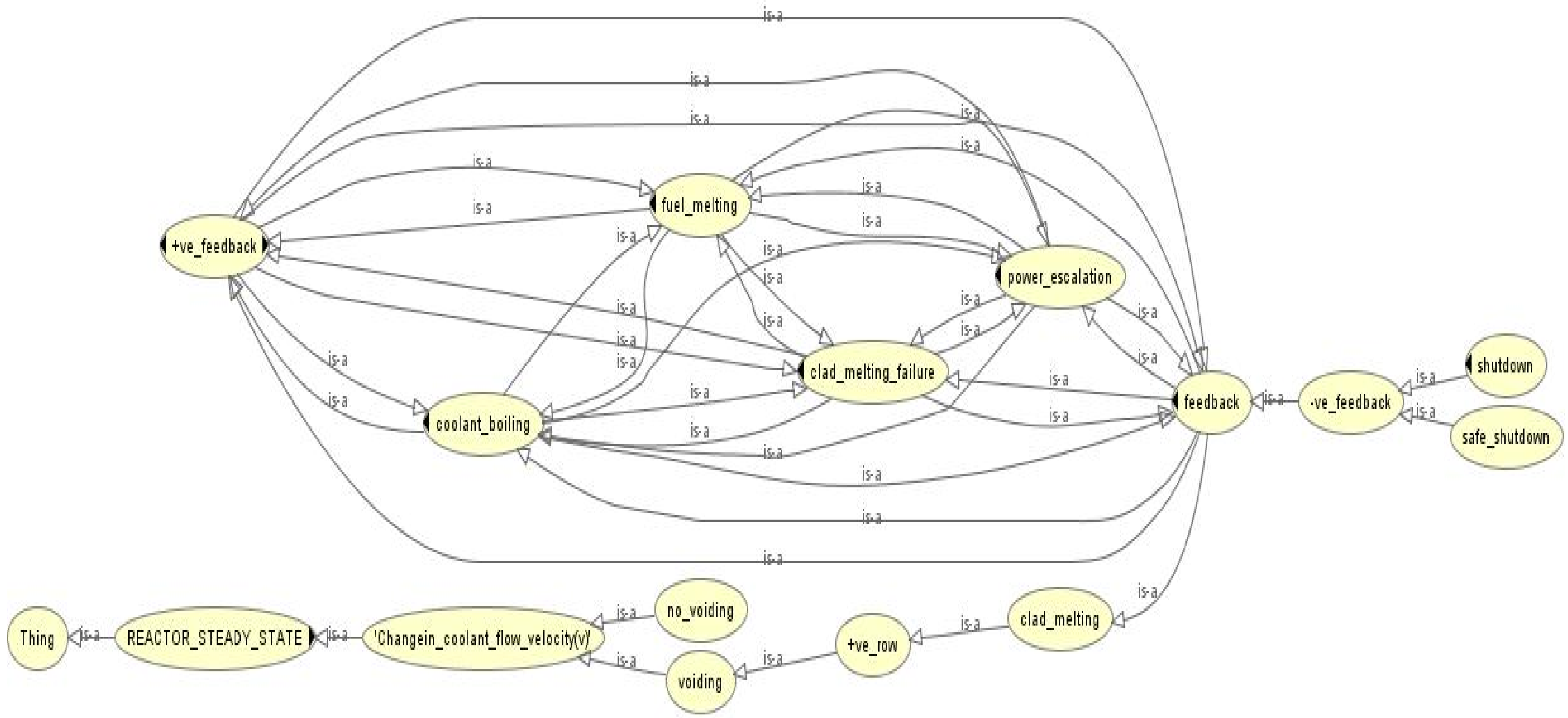


Figure 3.68 OWL representations for nuclear reactor transients state

171

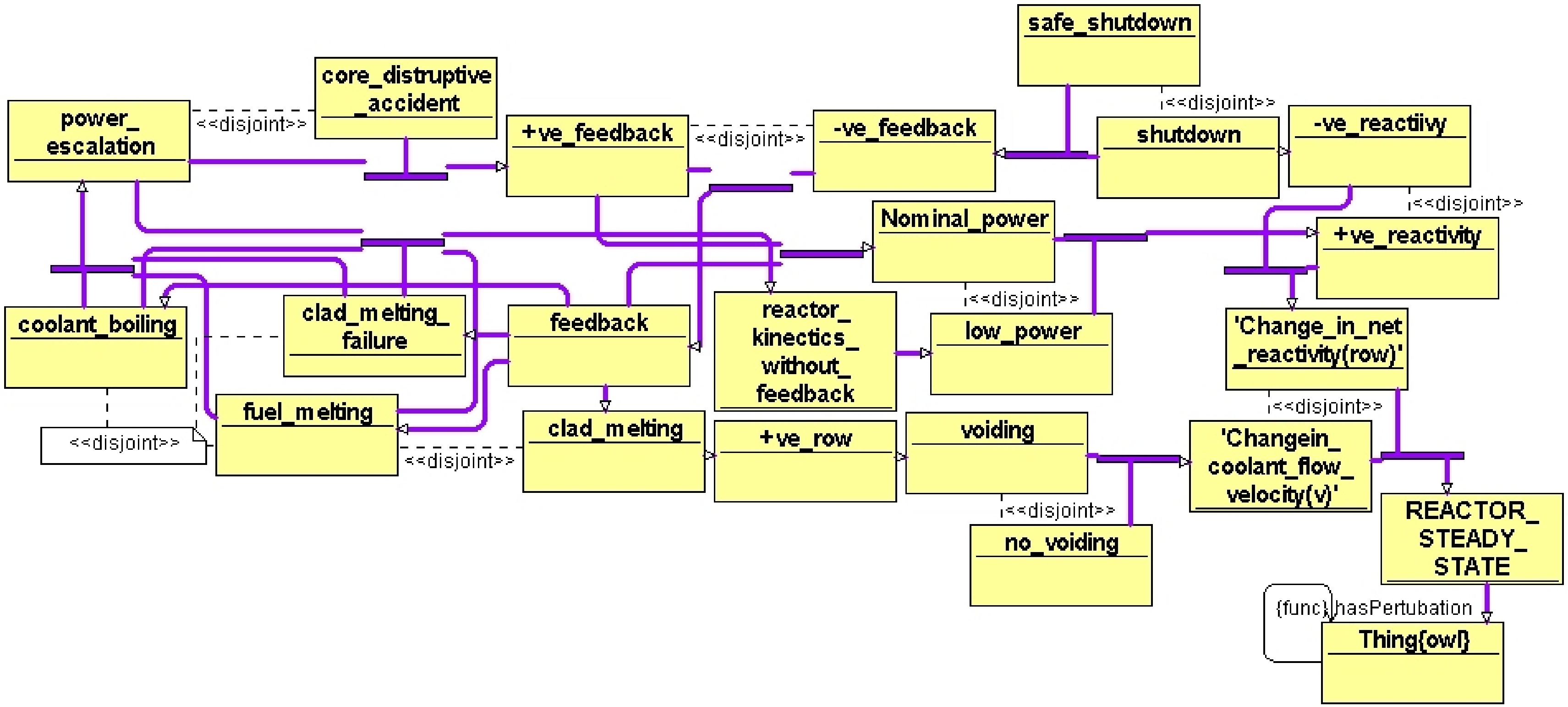


Figure 3.69 UML representations for nuclear Reactor transients state

3.5.30 Reactivity Feedback

A change in power for a reactor operating at high power, generally alters the temperatures of the fuel, moderator, and coolant and also alters the amount of void (if any) in the reactor coolant. A change in temperature of any of these components causes a change in reactivity that, in turn, affects reactor operation (a feedback effect). The reactor control system must therefore continually adjust the reactivity control mechanisms during a demanded power level change to keep the actual power changing at a rate that corresponds to the set point change. Positive feedback (like, clad and coolant expansion) will tend to cause instability whereas large negative feedback (like, axial and radial expansion of core) would oppose any power level change. OWL and UML representation for reactivity feedback are shown in Figure 3.70 and Figure 3.71.

3.5.31 Net Reactivity

Reactivity in the reactor is affected by many factors including coolant/moderator temperature and density, fuel temperature and density and structural materials temperature and density. Control rod movement introduces positive reactivity. Reactivity co-efficient, like temperature, Doppler, void, burn-up and power normally contributes to the negative reactivity addition to the reactor. Net reactivity is the sum of all these. OWL and UML representation for net reactivity are shown in Figure 3.72 and Figure 3.73.

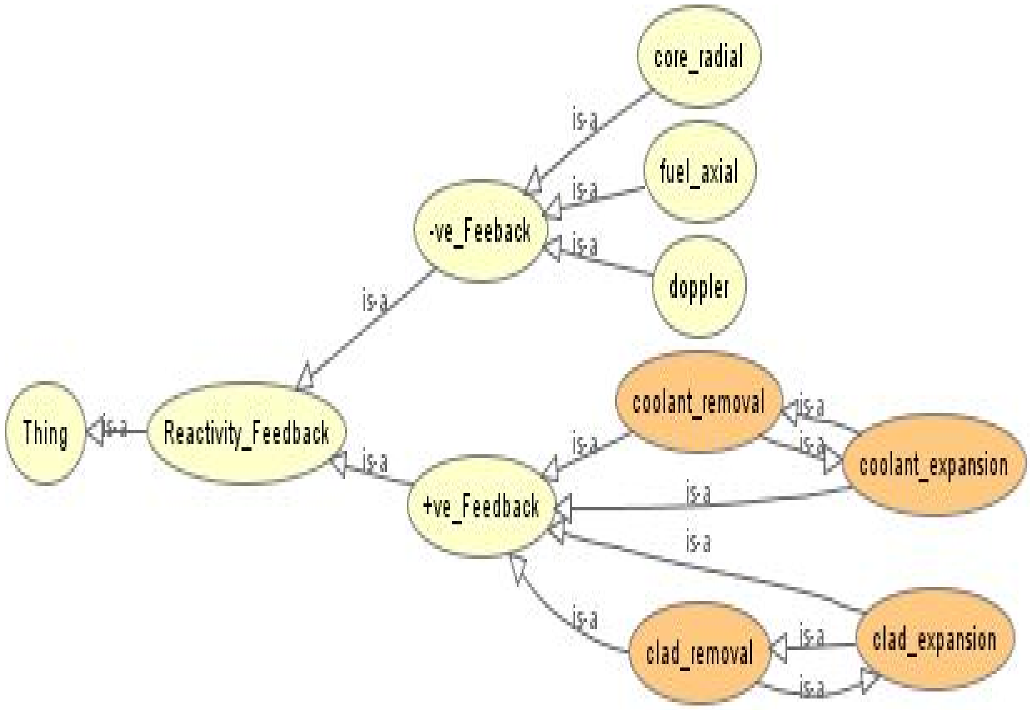


Figure 3.70 OWL representations for reactivity feedback

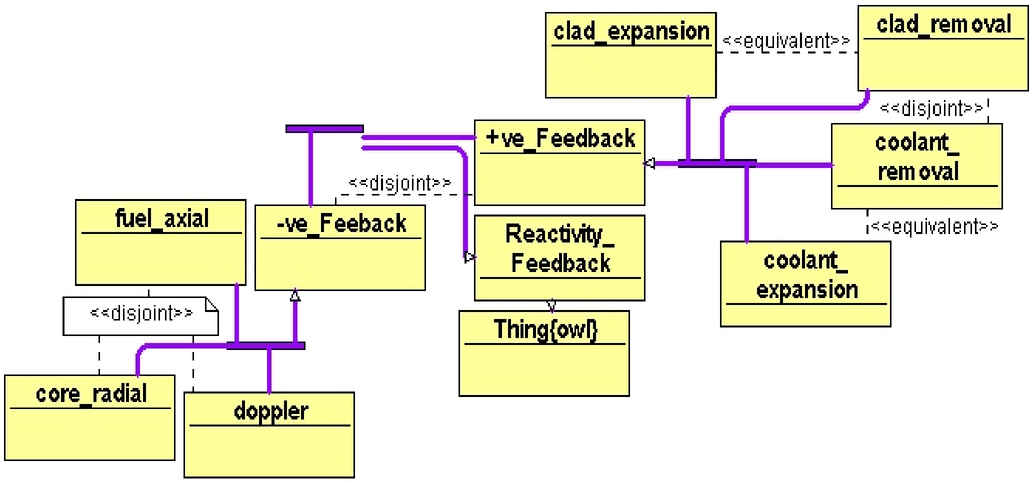


Figure 3.71 UML representations for reactivity feedback

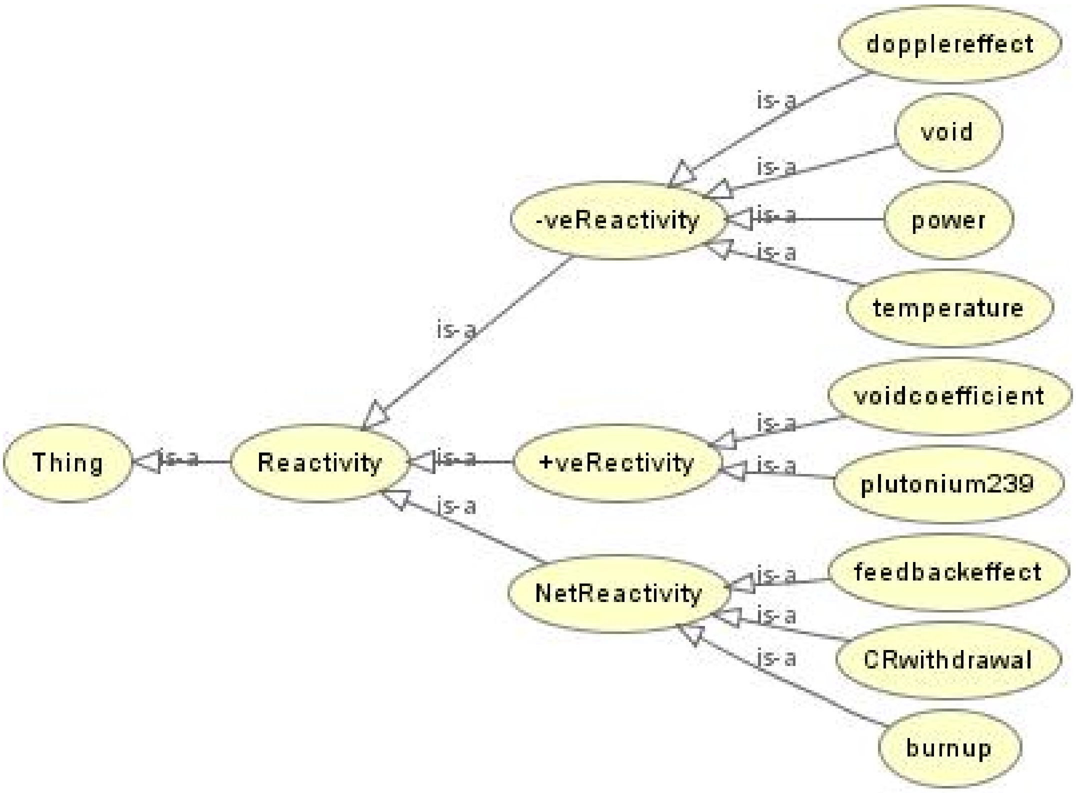


Figure 3.72 OWL representations for net reactivity

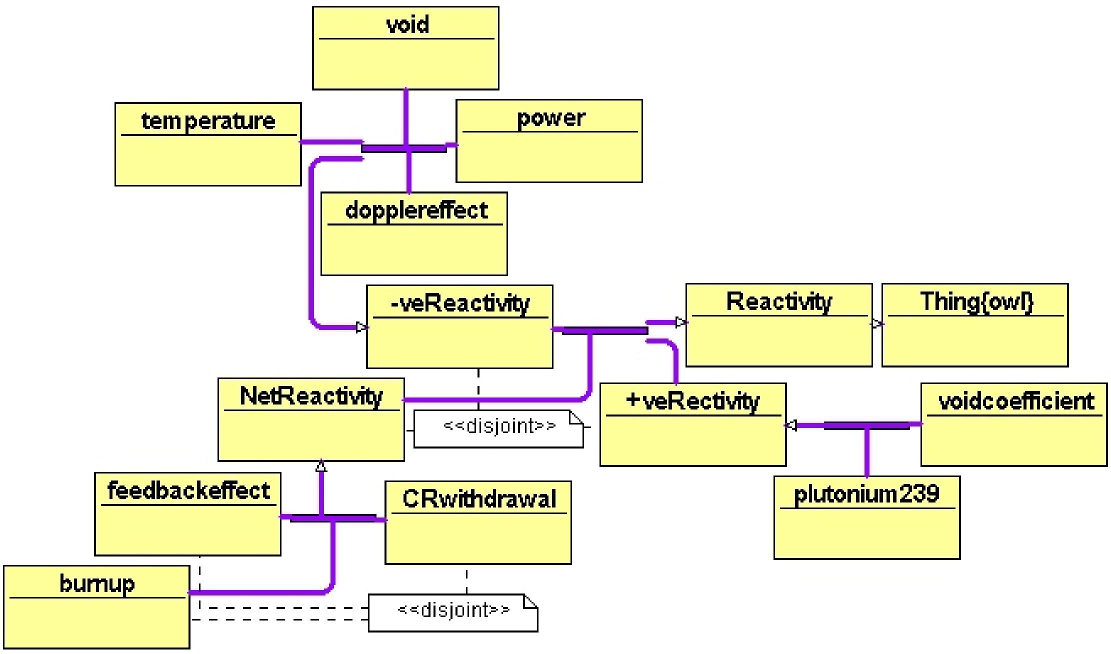


Figure 3.73 UML representations for net reactivity

3.5.32 Safety Analysis

Safety analysis consists of probability safety analysis and deterministic safety analysis. The probability safety analysis is classified into internal and external safety analysis. The external hazards to the system may be due to fire, flood or seismic. The internal hazards may be affected by component failure, system failure, ageing, environment factors etc. OWL and UML representation for safety analysis are shown in Figure 3.74 and Figure 3.75.

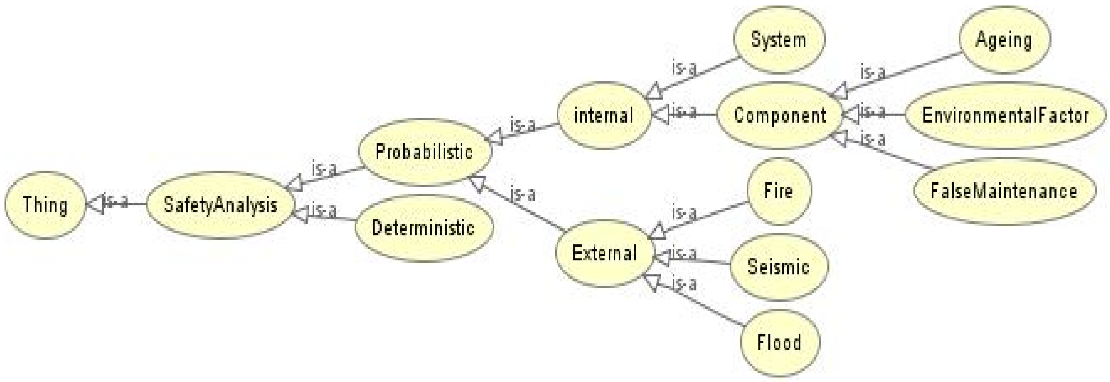


Figure 3.74 OWL representations for safety analysis

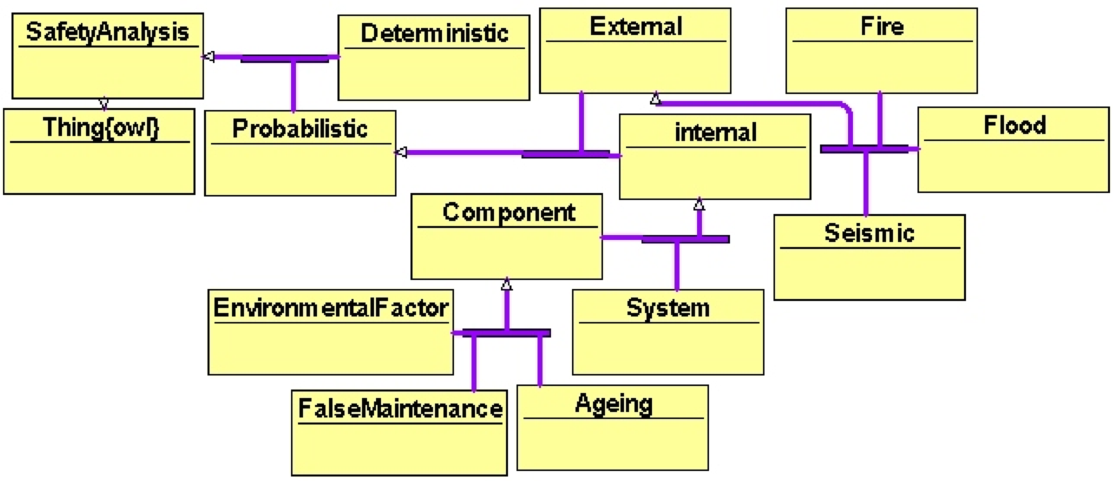


Figure 3.75 UML representations for safety analysis

3.5.33 Reactor Accident

For a given external reactivity perturbation that is for unprotected transient over power transients initiated at high power, there are expansion feedbacks available to mitigate the transients. If the total feedback reactivity is good enough to over compensate the external perturbation, net reactivity will become zero, then the reactor will go to another steady state. The external reactivity is high and if it is not compensated with feedbacks, net reactivity will become positive. Positive reactivity results in power rise. Subsequently there will be fuel temperature rise which will result in fuel melt. Molten fuel may undergo extrusion or slumping based on the melt fraction and burnup. OWL and UML representation for nuclear accident are shown in Figure 3.76 and Figure 3.77.

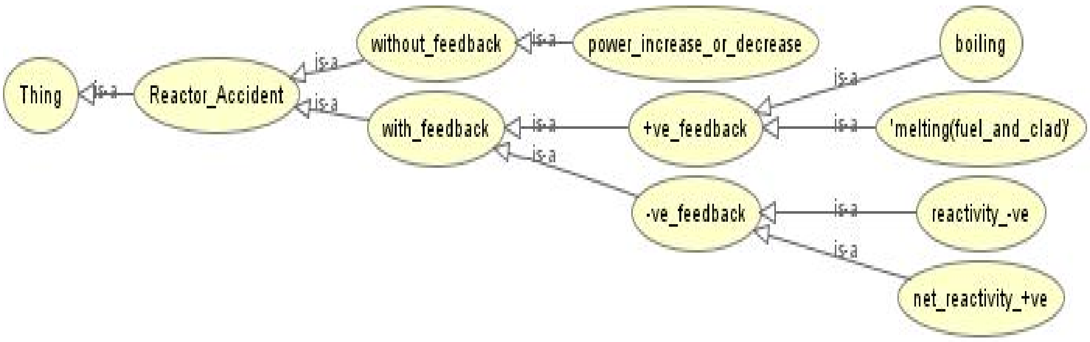


Figure 3.76 OWL representations for nuclear accident

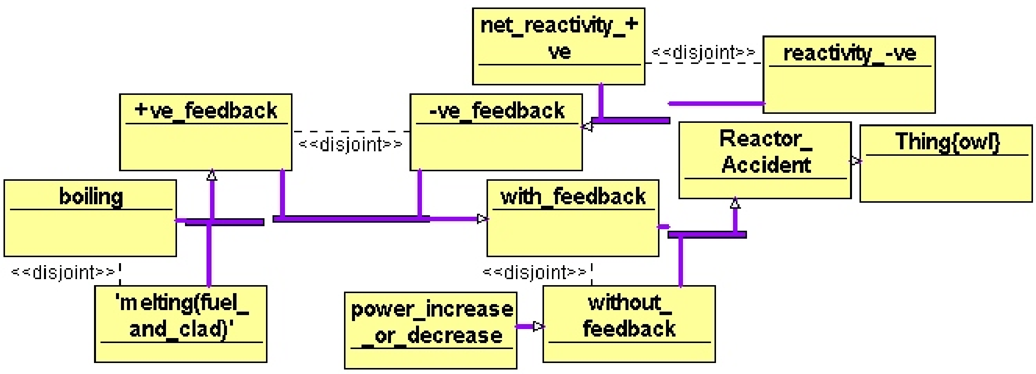


Figure 3.77 UML representations for nuclear accident

3.5.34 Conditions of Reactor Accident

There are three types of reactor accident situations: **T**emperature **O**ver **P**ower **A**ccident (TOPA), **L**oss **O**f **C**oolant **A**ccident (LOCA), **L**oss **O**f **F**low **A**ccident (LOFA). TOPA : Initiated due to uncontrolled removal of control rod with failure of protection system. LOCA : Assuming the coolant loss at full power resulting, fall in reactivity and subsequent reduction in power. Due to this fuel, starts getting overheated on account of due to residual neutron and decay power. In due course of time, fuel sub assembly starts melting and slumping. LOFA: Assuming loss of coolant flow due to pump failure (power/mechanical gripping) it may lead to boiling and vaporization of coolant in the core. In large fast reactors voiding of coolant leads to addition of positive reactivity which will result in increase of power and subsequent fuel melting. OWL and UML representation for nuclear accident are shown in Figure 3.78 and Figure 3.79.

3.5.35 Preheating and Emergency Core Cooling System

This circuit in FBTR is mainly intended for preheating of all components of primary sodium main circuit to 150^{o}C prior to initial filling and for heating the components to maintain the temperature when heat either from reactor core or from secondary sodium circuit is not available. If sodium leaks into A1 cell due to simultaneous rupture of reactor vessel and its double envelope, then sodium level will decrease in the reactor vessel. Under this situation preheating and emergency cooling system cannot be used to remove the decay heat, and sodium is injected into the reactor vessel from the flooding tanks. OWL and UML representation for preheating and emergency core cooling are shown in Figure 3.80 and Figure 3.81.

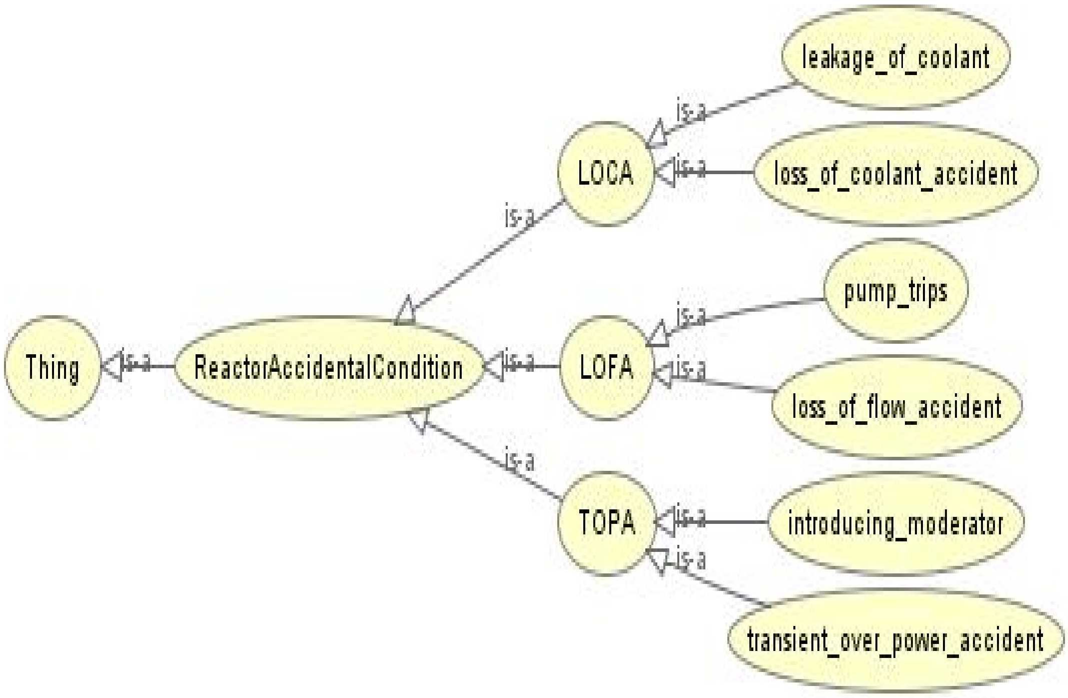


Figure 3.78 OWL representations for nuclear accident condition

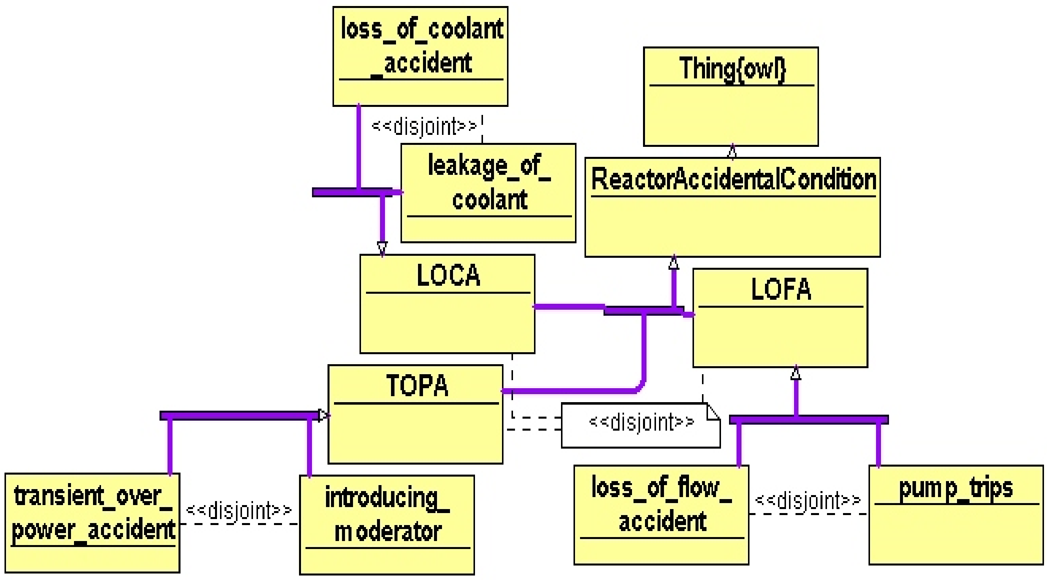


Figure 3.79 UML representations for nuclear accident condition

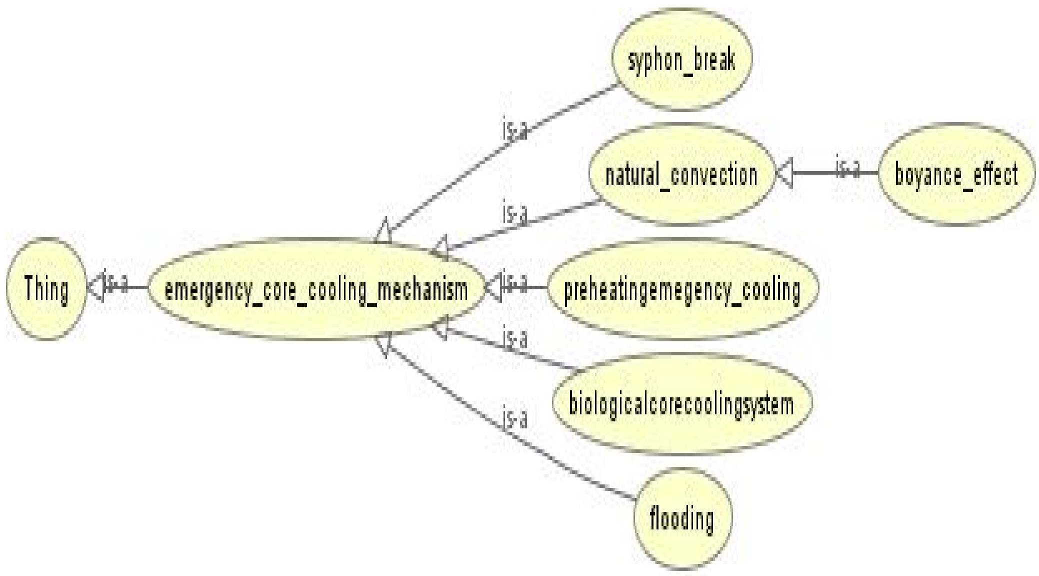


Figure 3.80 OWL representations for preheating and emergency core cooling system

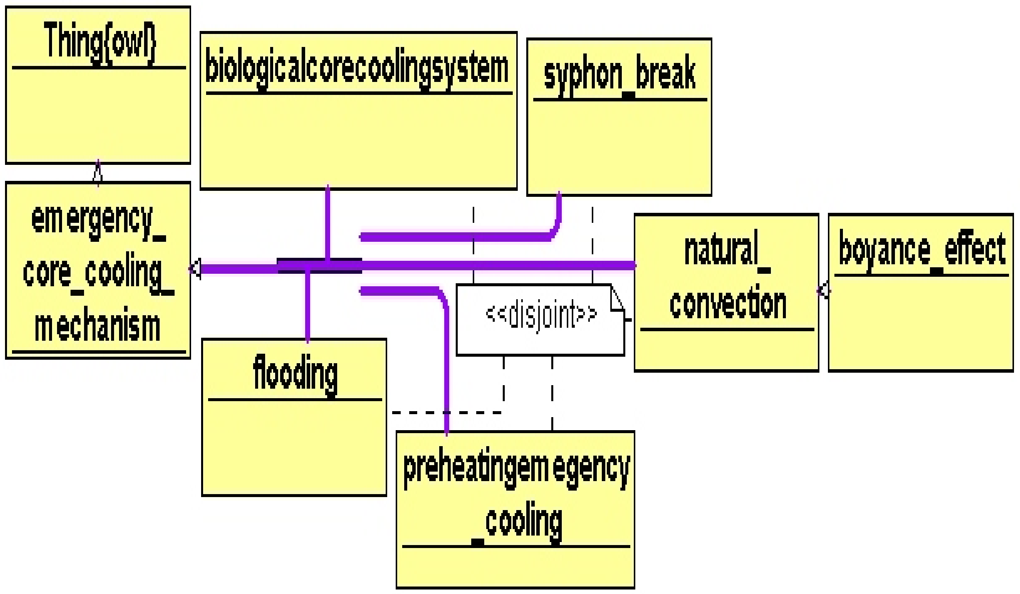


Figure 3.81 UML representations for preheating and emergency core cooling system

3.5.36 Neutron Flux

The neutron flux inside the reactor is a very important parameter. It will vary in all directions (X, Y and Z). The flux shape is the term applied to the density or relative strength of the flux as it moves around the reactor. Typically the higher neutron flux occurs in the middle of the reactor core, becoming lower toward the edges. It is maximum at the core centre and normally a cosine function. The higher the neutron flux the greater the chance of a nuclear reaction occurring as there are more neutrons going through an area. In a fast reactor, the neutron spectrum is very hard and close to the fission spectrum. OWL and UML representation for neutron flux are shown in Figure 3.82 and Figure 3.83.

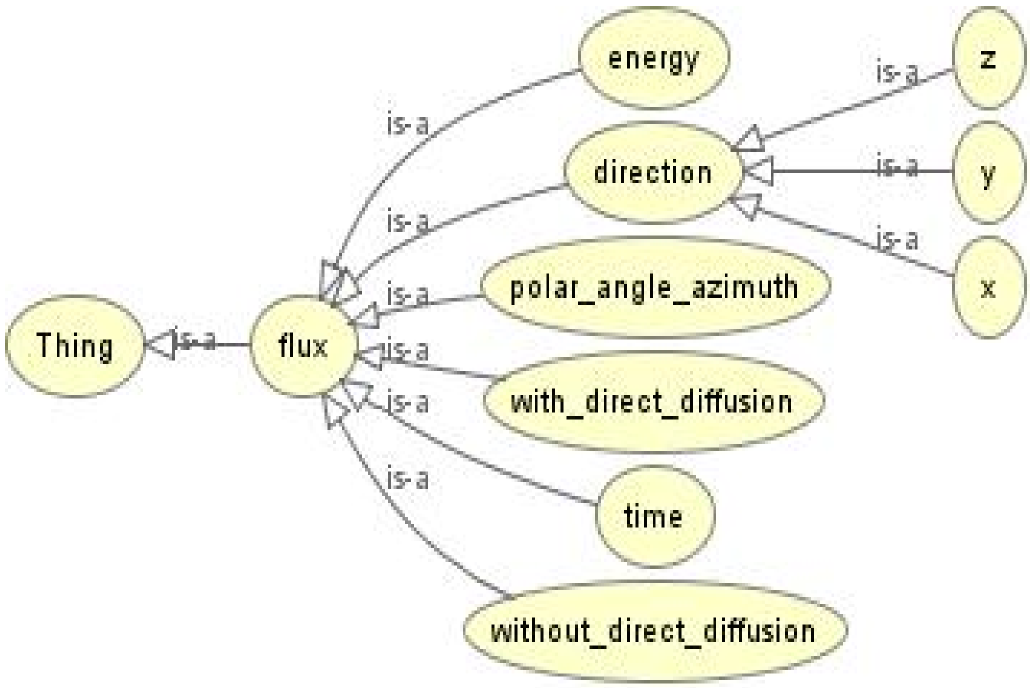


Figure 3.82 OWL representations for neutron flux

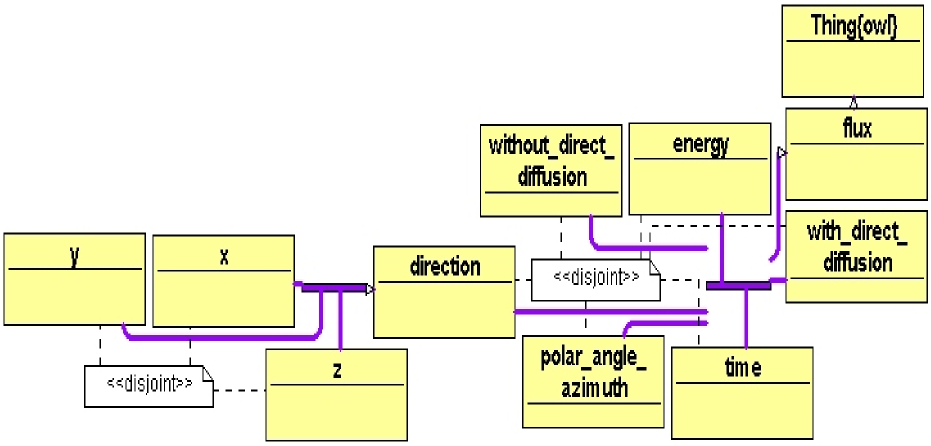


Figure 3.83 UML representations for neutron flux

3.5.37 Radiation

Radiation is a process in which energetic particles or wave travel through a medium or space. It is classified into two types, based on the interaction with matter: ionizing radiation (alpha, beta, gamma and neutron), non-ionizing radiation (heat, visible light, radio waves). Ionizing radiations are harmful and can be detected by ionization chambers, fission counters. Exposure to ionizing radiation can cause biological effects that may be harmful to the exposed individual and the progeny. The biological effects are divided into two types: deterministic, stochastic. Internal exposure of the radiation can be reduced by dilution process, whereas external exposure can be controlled by reducing the time to be spent with the source, increasing the distance from the radiation source and having the shield in front of the source. Radioactivity is measured in Becquerel (Bq) per second. 1 Bq means one disintegration per second. It is also measured in Curie (Ci), 1 Curie = 3.7×10^{10} Bq or disintegrations per second. The radiation absorbed dose is measured in *Gray, rad, rem and Sievert* (Sv). For personnel external monitoring, thermo-luminescent dosimeter (TLD), fast neutron badges (CR-39) and direct reading dosimeter (DRD) are used. External radiation is monitored using GM survey meter, area gamma monitors, RADMON survey meter, rem-response neutron monitor and teletector. OWL and UML representation for radiation are shown in Figure 3.84 and Figure 3.85.

Figure 3.84 OWL representations for radiation

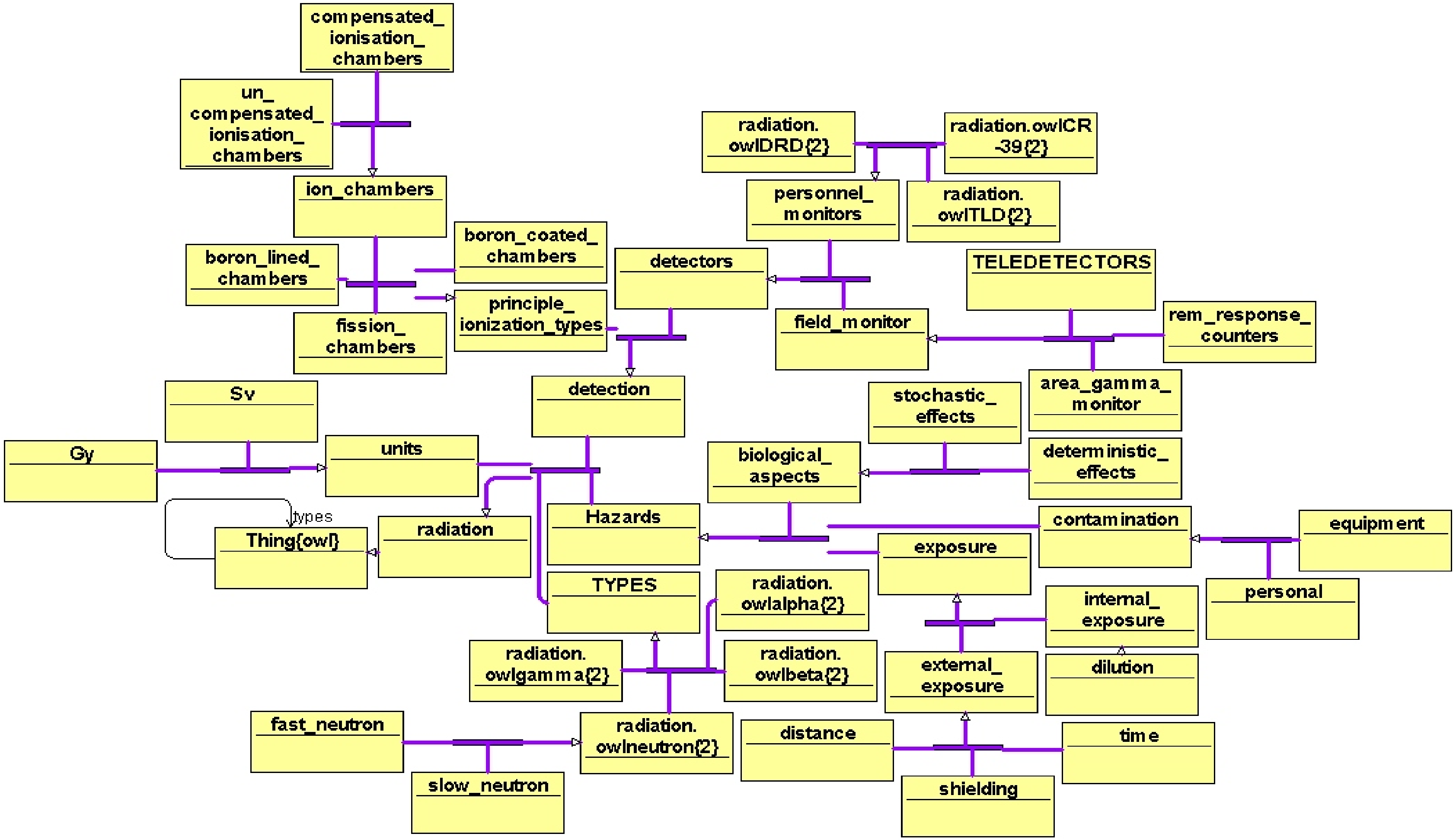


Figure 3.85 UML representations for radiation

3.5.38 Transport Equation

The linear Boltzmann transport equation is an integro-partial differential equation embodying the physics of neutral particle transport. This equation and boundary conditions are required for problems of finite geometric extent. The boundary conditions specify the distribution of particles entering the geometric problem through its exterior boundaries. The Boltzmann equation, together with the appropriate boundary conditions (and an initial condition for time-dependent problems), constitutes a mathematically well-posed problem having a unique solution. The number of independent variables is seven: three spatial variables, one momentum variable (energy or speed), two angles and one time variable. It is a balanced equation between production and loss of neutrons or gammas. Transport equation contains leakage term, interaction term, scattering term and source term. OWL and UML representation for transport equation are shown in Figure 3.86 and Figure 3.87.

186

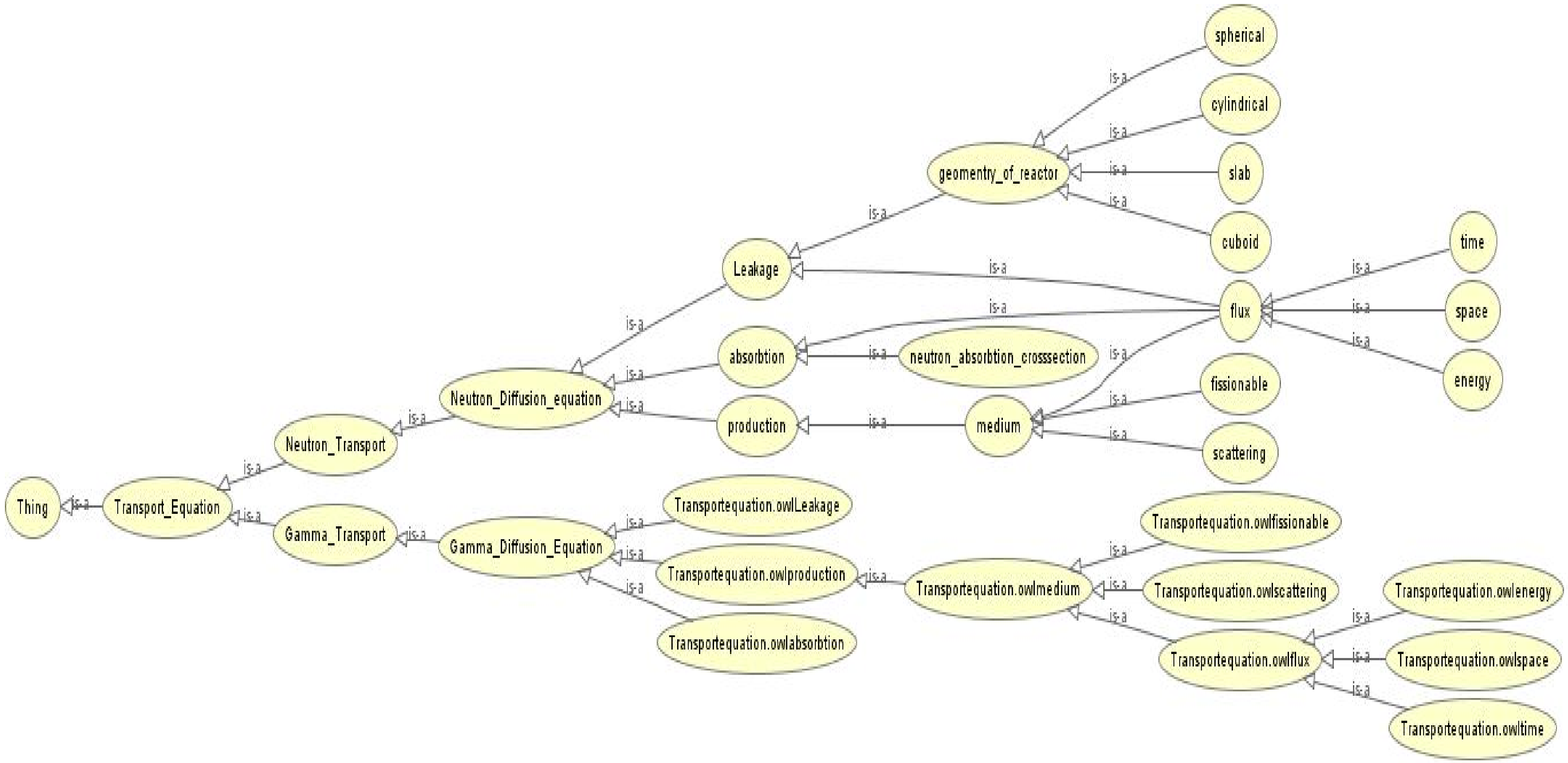


Figure 3.86 OWL representations for transport equation

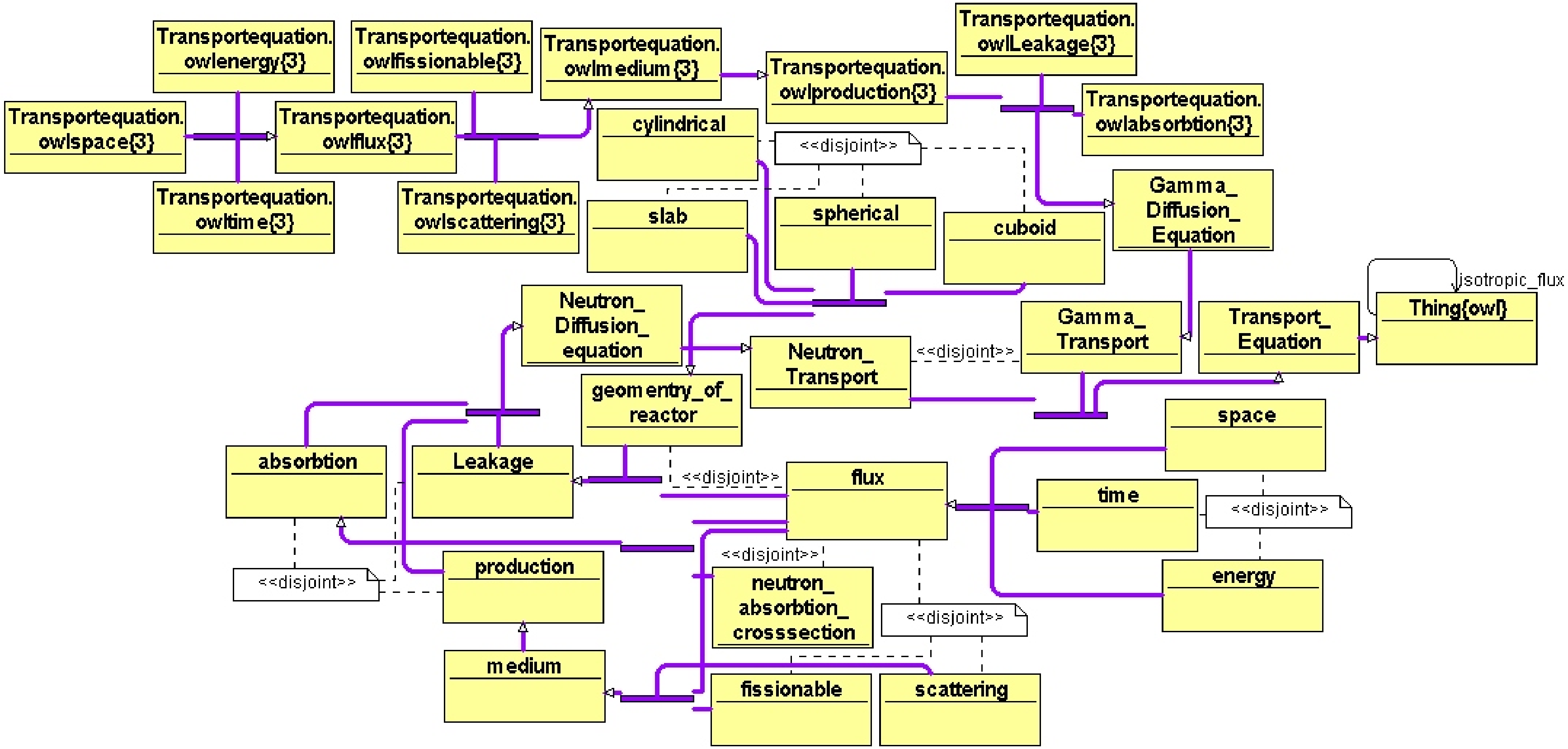


Figure 3.87 UML representations for transport equation

3.6 Challenges faced in the development of Ontology

Nuclear reactor technology being multi-disciplinary in nature is a complex system, hence ontology to be developed has to address the complexity. Ontology dynamics in general, leads to substantial reworking of the extant data. Semantic mismatch among data providers and users, are to be taken care by the ontology developer. A pre-requisite for the development of ontology necessitates the prior knowledge about the system and components and the same is applied to nuclear domain. Many of the systems in the nuclear reactor have many sub classes which consists are critical and needs to be represented in the ontology (viz. transport equation, nuclear channels etc). Interaction with operating personals is of importance to understand the essential intricacies, which needs to be addressed in the development of ontology. Several aspects of knowledge about the FBTR system are collected using the following methodology. In the first phase ontology developer acted as an inter-mediator between the domain expert and the specialist. Subsequently, the information collected from the working expert in their respective field by interviewing them. Then the information accumulated is compiled and converted to machine readable RDF, OWL format and it is uploaded in the portal. In general to carry out the ontology development for any domain the following methodology is to be adopted:

a. List out the knowledge likely to be implemented.
b. Identify the corresponding knowledge resource personals.
c. Cross verifying the same with other resource and operating personals.
d. Finally publishing the knowledge in the portal.

e. Motivating the resource personnel to add or update the knowledge in the portal.

Since nuclear reactor domain has extensive classification or division defining a relevant system and its components is a challenging process. In order to define a class in the ontology, its related object properties, its characteristics like functional, inverse functional, transitive, symmetric, asymmetric properties if required are also to be defined to integrate the knowledge.

The Protégé tool has the following problems:

- For large ontologies viewing of classes / nodes with clarity posed larger issues (eg. radiation, neutronic channels, electrical system, nuclear design test etc)
- Instances, data properties, properties like disjoint, intersection_of, union_of ,is_symmetric are not viewable in the owlviz or OntoGraf.
- Only is-a relationship is visible in the graph format, other formats are not viewable.
- Object properties defined are not viewable, but it is viewable in UML diagram.
- While defining the class for the ontology, space between words are not allowed. For readability, instead of space between words underscore is utilized.

3.7 Summary

A semantic web portal for nuclear reactor domain has been developed and deployed. The client /server architecture is adopted for development of the web portal. In this portal knowledge of FBTR is represented. To develop the portal, requisite knowledge of the nuclear domain is acquired and is represented in various formats like RDF, OWL and UML. The structure of database employed for storing the knowledge of FBTR is also described.

CHAPTER 4

DEVELOPMENT OF ALGORITHMS FOR ONTOLOGY MAPPING

This chapter covers the work carried out in developing the algorithm for ontology matching and optimisation. Quick mapping evaluator algorithm to avoid duplicate information in the portal and matrix rank based algorithm used to optimise the memory usage are discussed. Application of the algorithms is demonstrated by taking FBTR system as an example. Pareto optimisation technique for finding the optimised solution for matching is also described.

4.1 Introduction

When two ontologies differ in representation, they lead to disorder and influence the interoperability of ontologies. Entities of ontologies (matched ontology) are used for ontology merging, query answering, data translation, domain knowledge sharing and also navigating in the semantic web. In addition, ontology merging, integration and alignment can lead to ontology reuse. A Quick Mapping Evaluator (QME) which is an application program for ontology mapping and matrix rank based ontology matching have been developed, in continuation with our work of KMNuR. To eliminate inclusion of redundant or duplicate information QME algorithm is utilized. This is demonstrated by taking steam generator system of FBTR as an example. For matching different ontologies various types of similarity measures are used: terminological, structural, semantic and instance. The best match is the one which maximizes all the similarity measures. In practice there may be conflicting values amongst the similarity measures thus impend the matching process. To resolve this issue there is a need for optimization. Optimization means finding a solution which is most appropriate. A solution is a Pareto-optimal solution if it is not dominated by any other feasible solution. For getting better optimised solution in finding the similarity between two ontologies, Pareto optimisation technique is employed. This is demonstrated by taking three ontologies from nuclear reactor domain. The successful and unsuccessful match between the two ontologies is taken for analysis and the outcome of matching is illustrated. In the Pareto optimisation, a comparative study of different algorithms like Kullback Leibler Divergence, Cosine Divergence, Levenshtein String Similarity, Least Common Subsequence, String Similarity are found out and compared with user defined threshold to evaluate the optimised solution for matching. Terminological similarity measures used are Leventhein distance, Least Common Subsequence, String similarity, WordNet. Structural similarity measures used are Kullback Leibler Divergence, Cosine structure divergence. Algorithms are chosen for their popularity, simplicity and extensively usage in many fields.

4.1.1 Ontology Mapping

Ontology mapping is the process of creating association between two distributed ontologies, provided the sources are consistent with each other. Ontology matching systems have been developed by many researchers. By using semantic relations, the entities in the source ontology, are transformed into the target ontology [203]. Semantic heterogeneity problems are solved by matching the ontologies [204].

Ontology mapping is defined mathematically as described below:

Let a mapping element O be defined as five tuples <id,e,e′,n,R> where id is an unique identifier, e and e′ are entities, n is a confidence measure and R is a relation (for example equivalence, disjointedness). An alignment is a set of mapping elements, and the ontology mapping operation determines an alignment for a pair of ontologies *O* and *O′*[205]. Ontology mapping can be broadly classified into [89]:

i) Element Level

Syntactic – String, Language, Constraint based Technique

External – Linguistic resources, Upper Domain specific formal ontology

ii) Structure Level

Syntactic – Graph based, Taxonomy based

External – Repository of structures

Semantic – Model based Technique

In the present work, an element level syntactic method based on similarity measures is utilized. The main task of ontology mapping is that of finding similarity measures between different ontologies by comparing the names of the entities of the ontologies. This is achieved by using algorithms like String Equality, Longest Common Subsequence and Levenshtein Distance.

4.1.2 Ontology Alignment Tools

PROMPT, Chimerae, Glue, CTXmatch, COMA, MULTIKAT etc. are some of the known ontology alignment tools. These tools are classified as automatic and semiautomatic. PROMPT, an algorithm that provides a semi-automatic approach to ontology merging and alignment is available in Protégé as a plug-in [206]. The algorithms used for finding the ontology matching are described below.

4.1.3 String Similarity Measures

Methods of comparing strings have been used in many applications. Detecting data duplication, databases and software analysis are some of the applications of string similarity [207-209]. The ways of finding such string similarity measures have been classified as: edit distance like functions, token based distance functions, hybrid distance functions and blocking or pruning methods [210].

4.1.3.1 String Equality

String Equality is a strict measure used to compare strings [211]. It returns a similarity measure of 1, if the first string is the same as the second string, otherwise it returns 0. It is used as the baseline for measuring efficiency and effectiveness of different algorithms. Threshold is a number using which similarity (same or not) measure between two strings are determined and it varies between 0 and 1. For example, if the threshold is set to be 0.8, then all pairs of strings with similarity measure ≥ 0.8 are considered to be the same. The value of threshold can be different for different metrics to have different mathematical properties and hence one optimal threshold for one metric will not suit for other metrics [212].

4.1.3.2 Longest Common Subsequence (LCS)

A subsequence of a string is defined as any sequence which can be obtained from the string by deleting one or more characters from the original string. The advantage of Longest Common Subsequence (LCS) [213] is that:

i. It does not require consecutive matches
ii. Results in-sequence matches that reflect sentence level word order and
iii. Results in not having the necessity of predefined length

The similarity measure is represented as:

$$\text{Similarity} = \frac{\text{Length (LCS)}}{(\text{Min (Length (string1), Length (string2))}}$$

For example, ‘kitten’ and ‘sitting’ have LCS ‘ittn’ as common and hence similarity is 4/6 or 0.666667. Least common subsequence finds application in fields like computational biology to data compression, syntactic pattern recognition and file comparison. LCS used to deal with two types of symbols to model the difference in the number of occurrences allowed for each gene [214]. It is also employed for measuring the interleaving relationship between sequences. The common sequences can be viewed as the identical parts of the input sequences, which can help to reconstruct an alignment of the sequences. This finds application in molecular biology research [215]. LCS is an appropriate measure in DNA or protein sequences to find the maximum number of identical symbols among them and also in preserving the symbol order [216-217].

4.1.3.3. Levenshtein distance

Levenshtein distance or Edit distance is a string relation metric, very commonly used in generating string similarity measures. It is the minimum number of edits like substitutions, additions or deletions required to convert one string to another string [218].

For example in ‘kitten’ and ‘sitting’

kitten → sitten → sittin → sitting

Hence the Levenshtein distance between these pair of words is three. The string equality algorithm is implemented using the ‘isEqual()’ function in java whereas the LCS and Levenshtein distance algorithms are implemented using dynamic programming. Levenshtein distance is used in many fields, like computational biology, text processing etc. In the field of information theory and computer science, Levenshtein distance is a metric for measuring the amount of difference between two sequences. It is also used for comparison of on-line signatures authentication [219]. Levenshtein is also used to construct a case-insensitive pass-phrase system that can tolerate zero, one or two spelling errors per word [220]. In plagiarism detection the Levenshtein distance is used to change the likely scarcity, which improves both time and space efficiency [221]. Using these algorithms **Q**uick **M**apping **E**valuator (QME) program is developed and the details are given below.

4.2 Quick Mapping Evaluator

QME is developed as a *JAVA swing* based graphical user interface (GUI) application. Alignment **A**pplication **P**rogram **I**nterface (API) was used for obtaining the output whether ontology is same or different when two sets of ontologies are compared [222]. The users are allowed to change the type of match making algorithm. Apart from this it will also be used to find the metrics [223].

4.2.1 Implementing QME

The development was carried out using three packages – GUI, Algos and Utilities. GUI package contains ontologyMapperGUI.java which contains functions linking the GUI of the application with their corresponding actions. Algos package contains the algorithm for the implementations of String Equality, Longest Common Subsequence and Levenshtein distance. Utilities package contains TextAreaWriter.java which is used to display text in the GUI text areas. It also contains external libraries (.jar files) from the alignment API for parsing input, displaying output in the alignment format and support to easily extend and implement new algorithms.

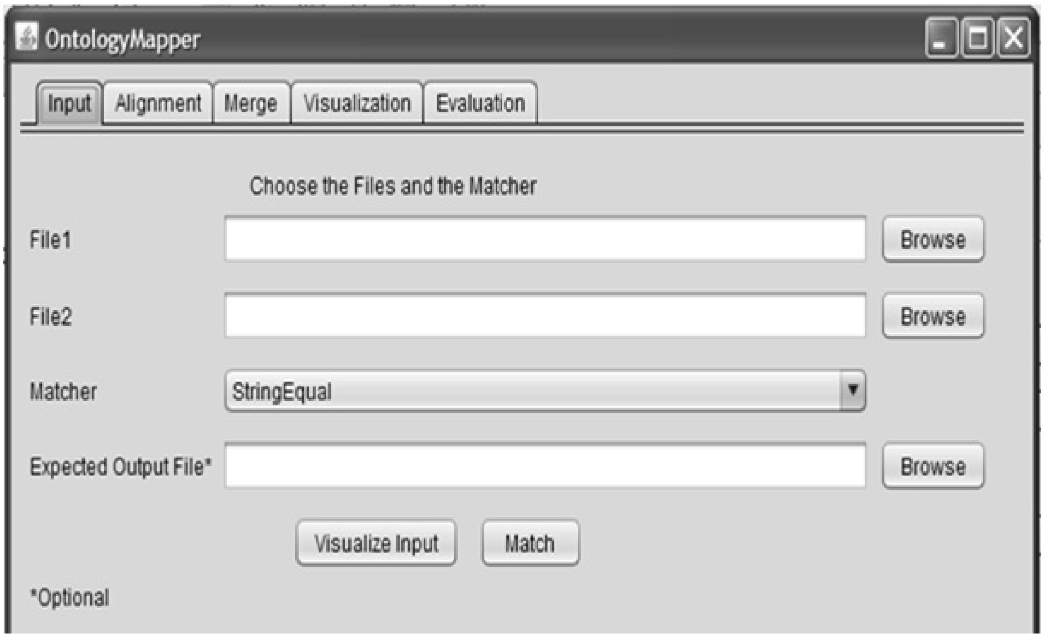


Figure 4.1 The application graphical user interface– input tab

The user interface shown in Figure 4.1 is developed as a multi-tabbed pane with different tabs (panels) displaying different parts of the ontology mapping process. The input panel takes two OWL files and an optional expected output file in RDF alignment format as input. There is also an option to select the matcher from the matcher dropdown box. On clicking the match icon, the algorithm runs on the selected matcher and evaluation (if any) is performed using the expected output file. This is shown in Figure 4.2. Here logics of electrical power supply system and distributed power system of the nuclear reactor system are compared to retrieve the ontology matching. The resultant alignment can be viewed in the alignment tab and the evaluation (if any) can be viewed in the evaluation tab. The input files, electrical power supply.owl and power supply distribution.owl are shown in Figure 4.3 and Figure 4.4 respectively. In the alignment tab, the two OWL files using the specified matcher is displayed in the RDF format as shown in Figure 4.5. The alignment can be visualized using RDF-gravity. A snapshot of the same is shown in Figure 4.6. In the evaluation tab, the matching at different thresholds is displayed in terms of Precision, Recall, F-Measure and the number of cells over which these are calculated and a snap shot of the same is shown in Figure 4.7. One of the major challenges faced during the development of this tool was the absence of an embeddable visualization tool for OWL as well as RDF-alignment format.

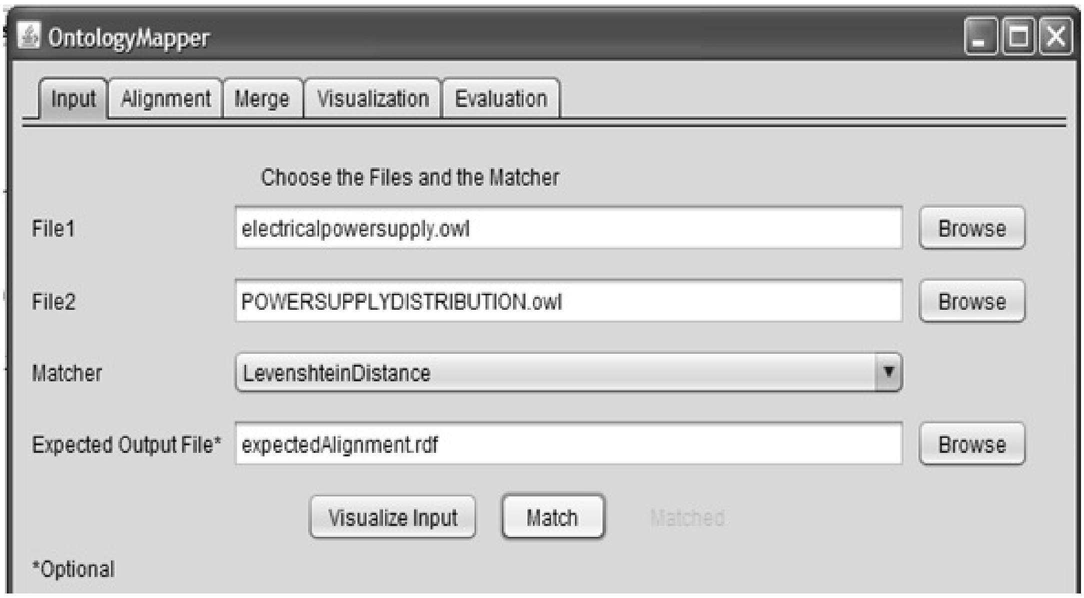


Figure 4.2 Running the mapping algorithm

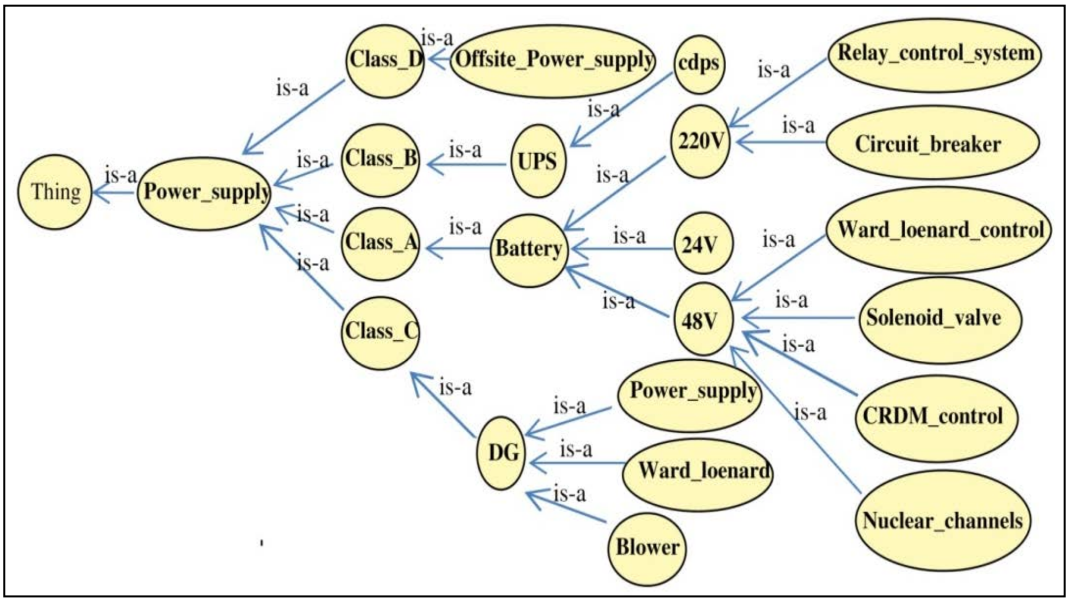


Figure 4.3 View of electrical power supply output from Protégé

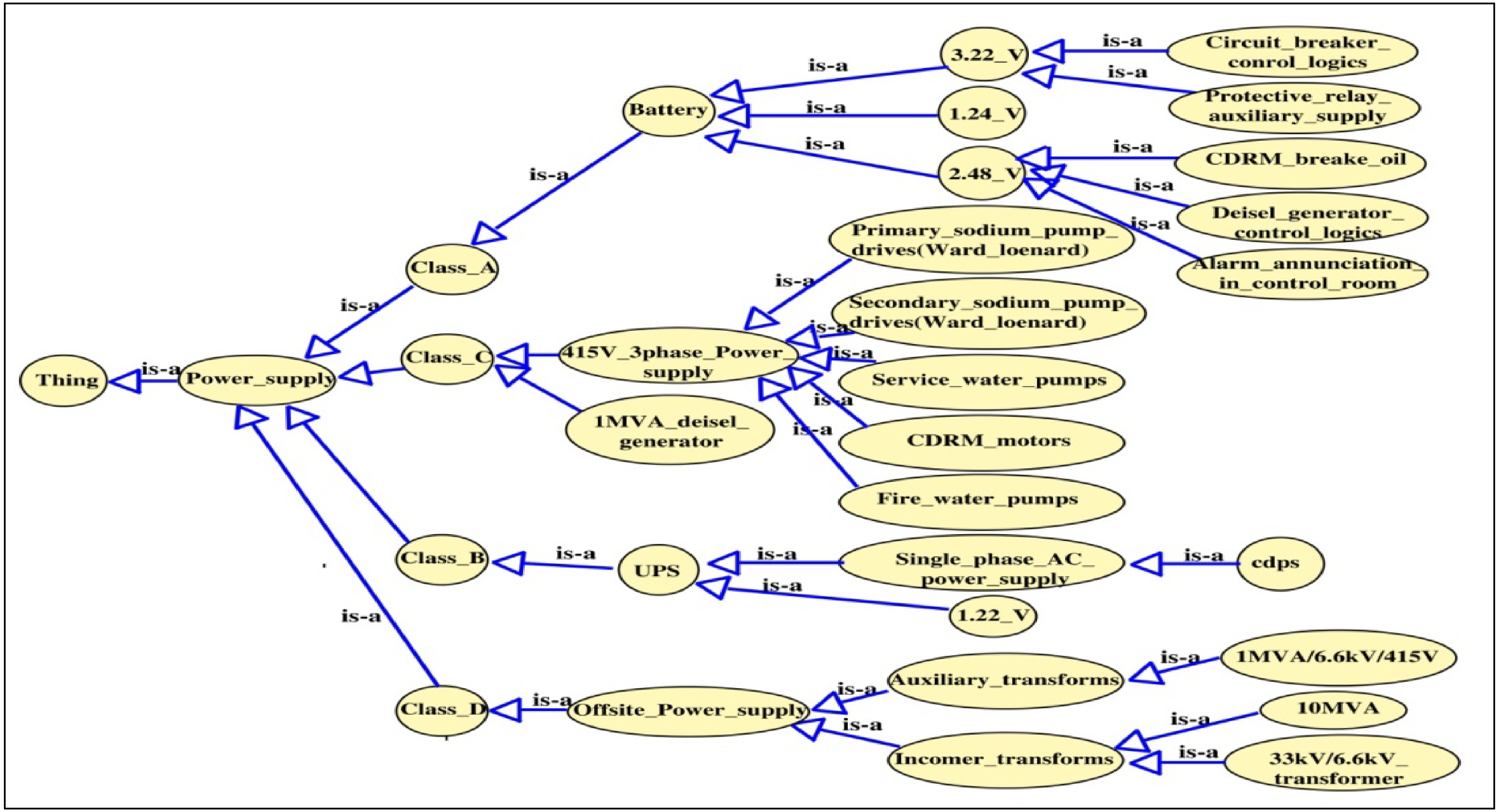


Figure 4.4 View of power supply distribution logic from Protégé

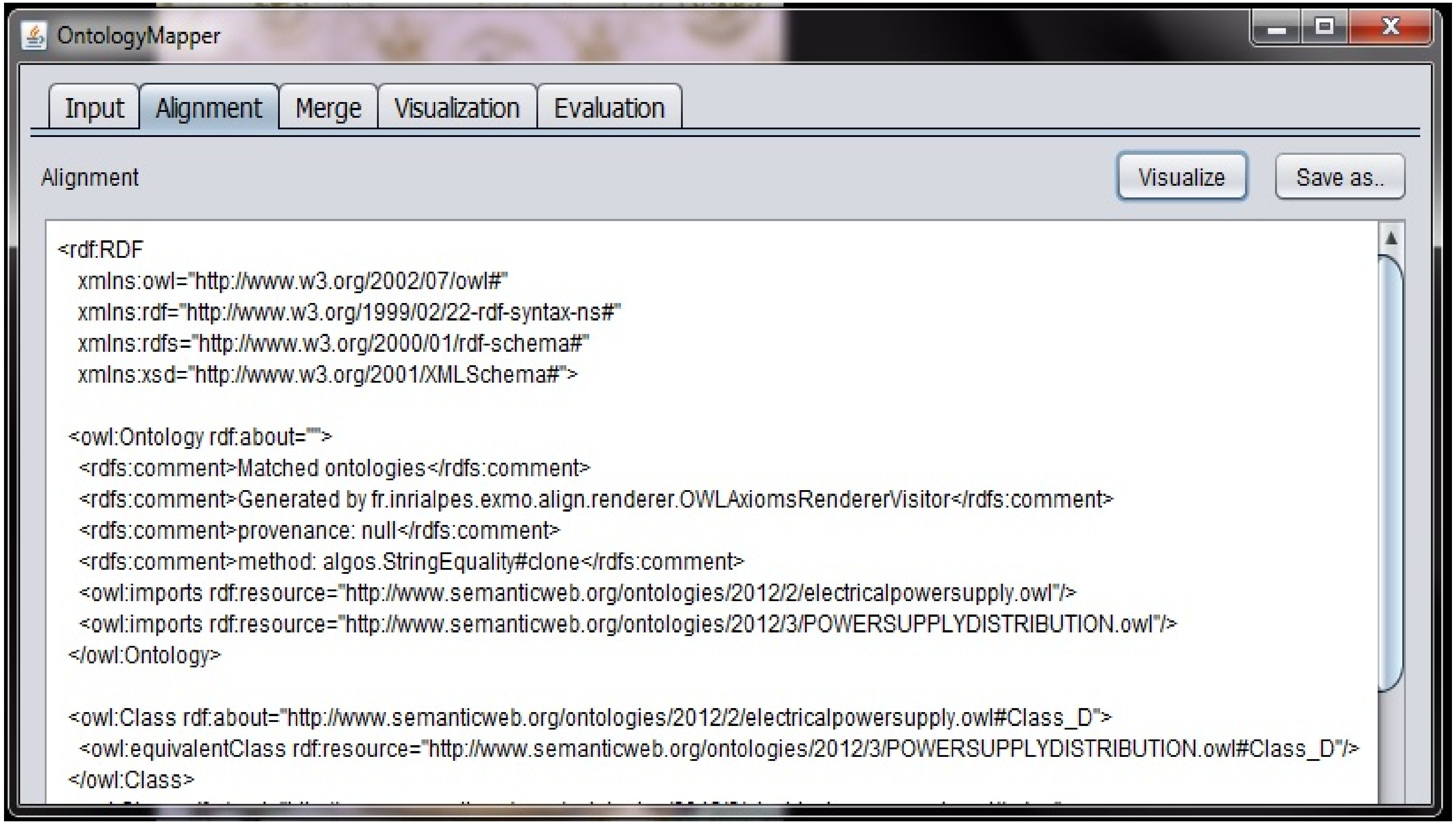


Figure 4.5 The alignment of the ontologies – alignment tab

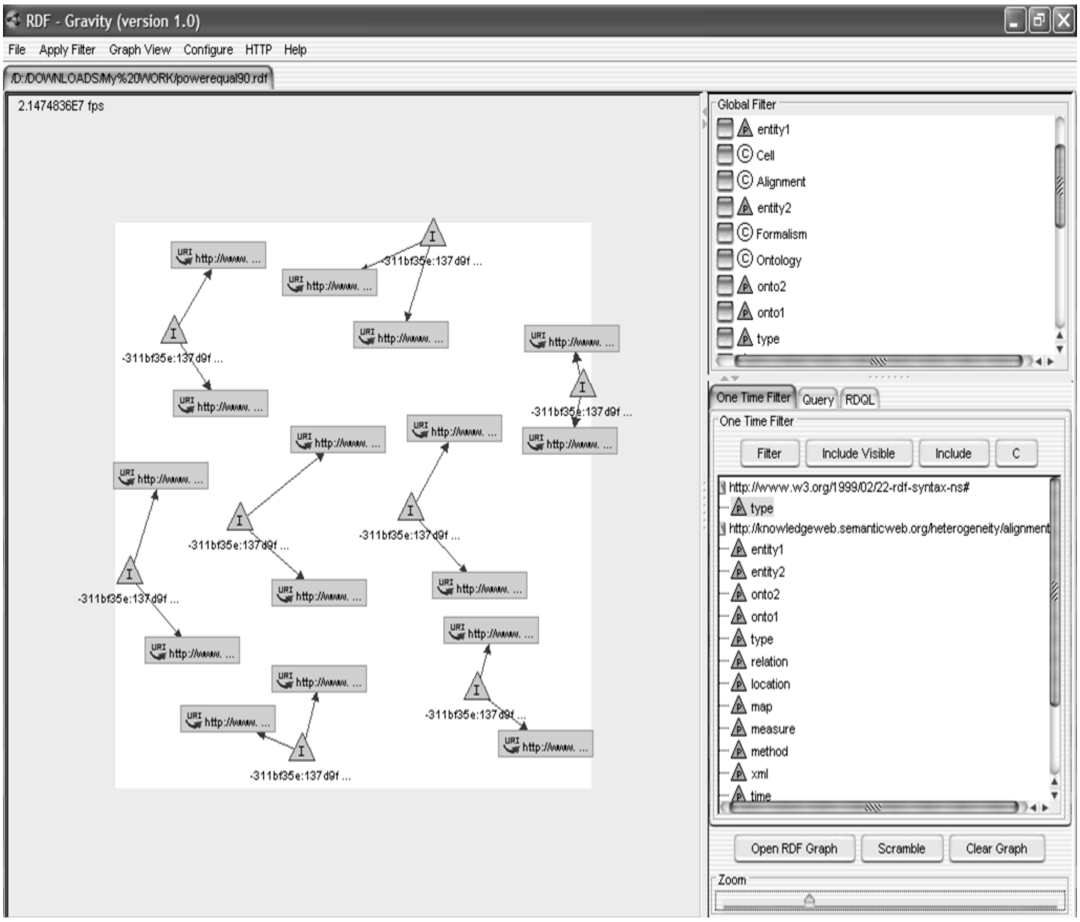


Figure 4.6 Visualization of the alignment using RDF-Gravity

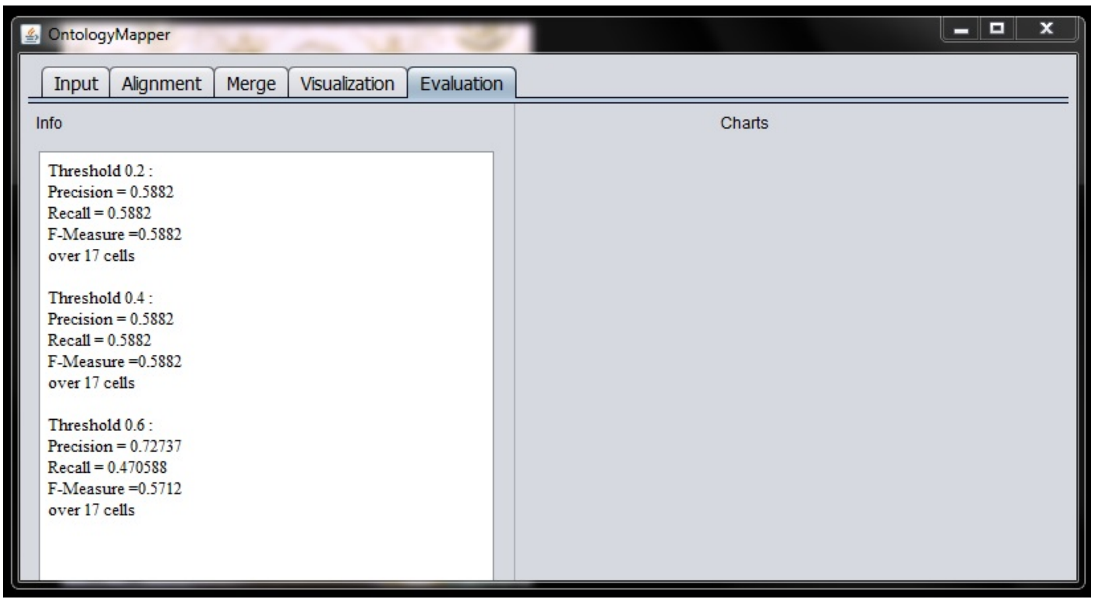


Figure 4.7 The evaluation of the mapping – evaluation tab

4.2.2 Evaluation of the Results

Precision refers to the fraction of retrieved instances that are relevant.

$$Precision = \frac{\text{Number of retrieved instances which are relevant}}{\text{Total number of retrieved instances}}$$

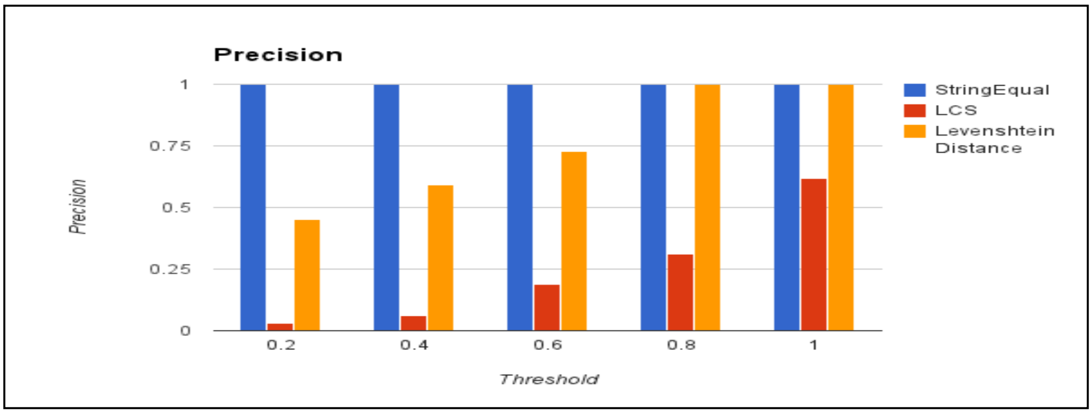
Recall is the fraction of relevant instances that are retrieved.

$$\boldsymbol{Recall} = \frac{\textbf{Number of relevant instances which are retrieved}}{\textbf{Total number of relevant instances}}$$

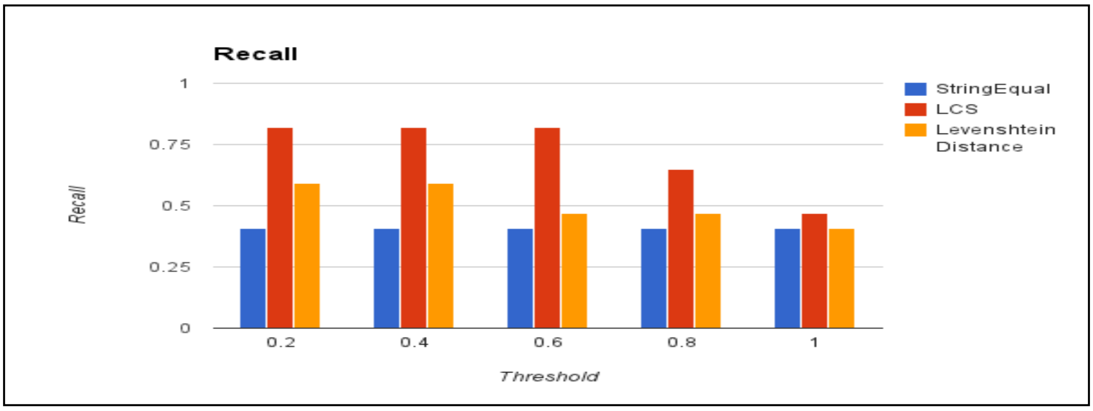
F-Measure is the harmonic mean of Precision and Recall. It gives a measure of accuracy of the mapping.

$$\mathrm{F-Measure} = \frac{\text{2 x Precision x Recall}}{\text{Precision} + \text{Recall}}$$

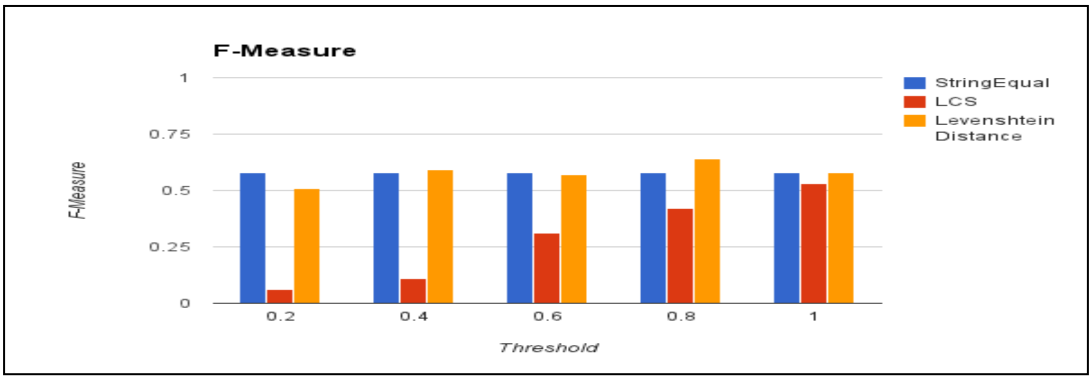
The ontology of electrical power supply and distributed power system logics of the nuclear reactor are matched employing QME, and the metrics (Precision, Recall, F-Measure) with different thresholds are shown in Figure 4.8. The StringEqual algorithm has a very high precision, a very poor recall and also a very low F-Measure irrespective of the thresholds. The LCS algorithm steadily increases in precision with increasing threshold, but decreases in Recall. The overall F-Measure of the LCS algorithm increases with threshold. The Levenshtein distance also has a steady increase in precision and decrease in threshold. The overall F-Measure remains almost consistent at a high value. At a threshold of 0.8, the Levenshtein distance algorithm has the highest F-Measure and hence gives the best accuracy. Since this QME add-on is integrated with the KMNuR, depending on the user's choice, ontology alignment algorithm can be chosen to get the matching results.



(a)



(b)



(c)

Figure 4.8 Results of the comparison of String Equal, LCS and Levenshtein distance using (a) Precision (b) Recall (c) F-Measure metrics

4.2.3 Challenges Faced for Identification of Ontology Duplication

Duplicate information leads to following difficulties: unavailability of any standard leads to different ontologies being defined for the same subject. Also other aspects are different ontology tools that use different ontology languages. Duplicates have merely similar attribute values due to different formats and fuzziness. A challenge of automatic ontology is to identify the duplicate information and redundant annotations across various domains. The knowledge extractor has to find solutions to the presence of duplicate entries to unnecessary processing of alignment task, as they occupy memory space and also decrease the precision [224]. These duplicate pages increase the size of search engine indexes and reduce the data quality [225]. To eliminate inclusion of redundant or duplicate information while sharing the information in the web portal QME algorithm is utilized.

4.2.4 Knowledge Sharing in Web Portal using QME

In KMNuR web portal, whenever the user inputs a new information, it will be cross checked with the existing information for similarity. This is done to avoid duplicate entry in the portal. If the user submitted information is new, then it is added to the web portal for reusing and sharing of knowledge. In place of second ontology (File 2) for ontology mapping, here the OWL knowledge available in KMNuR portal is loaded (Figure 4.9). So that the matching is calculated and the approval_status field in the knowledge database is updated. Thus the QME algorithm aids the portal to avoid duplicate information from different sources.

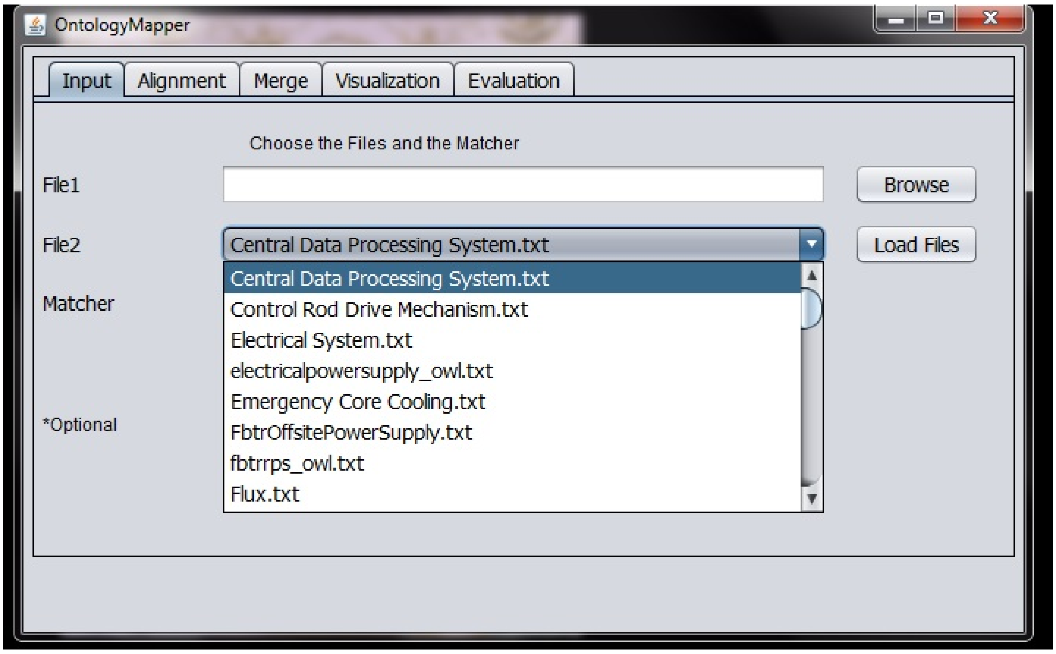


Figure 4.9 A screenshot for matching the ontology with the existing ontology in KMNuR Portal.

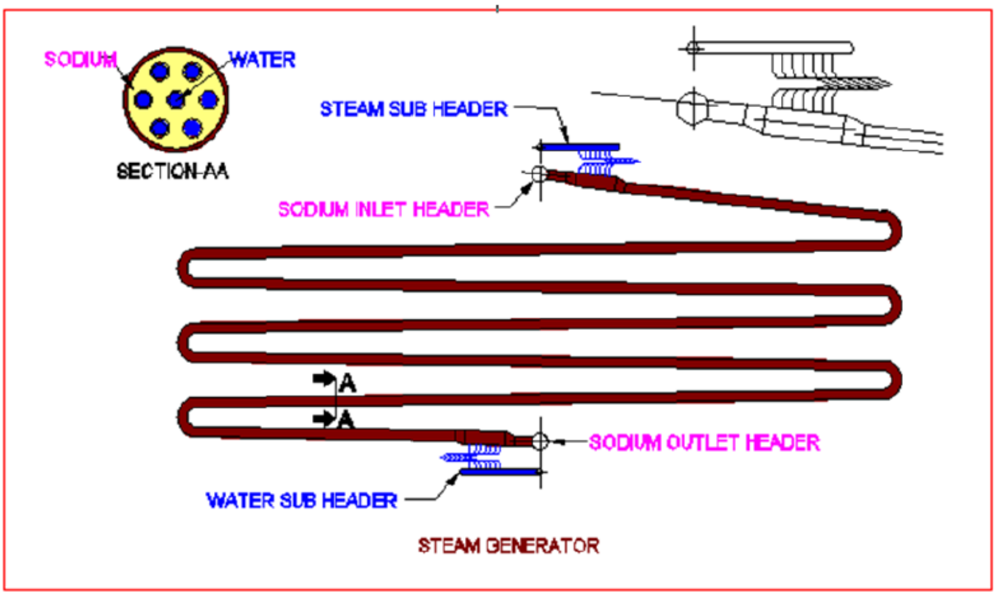


Figure 4.10 Steam generators for fast breeder reactor [taken from the IGCAR internal website]

As a case study steam generator system of FBTR is taken for analysis (Figure 4.10). FBTR has two primary and secondary loops and a common steam water circuit with four once through **S**team **G**enerator (SG) modules, which supply super-heated steam to the condensing turbine. The steam produced is supplied to a condensing turbine coupled to an alternator [226]. Sodium water reaction is exothermal and any leak at incipient stage is to be detected to avoid major damage to the steam generator. The **S**team **G**enerator **L**eak **D**etection **S**ystem (SGLDS) consists of a Nickel Diffuser through which sodium from the steam generators is passed. The shell of the diffuser is kept under ultra-high vacuum and hydrogen diffusing through the nickel tubes is monitored by the Sputter Ion Pump current. The steam generator system for FBTR ontology concept is shown in Figure 4.11.

Case 1: New Information Submission.

When the user submits the new information, it will be added to the KMNuR portal and also by setting the “approval status” field to “Yes”. Thus the new knowledge submitted is added and displayed in the KMNuR portal.

Case 2: Partial Overlapping

Consider the case if a steam generator system submitted by the user describes the property of the machine as shown in the Figure 4.12. When QME algorithm is executed, it is added to the portal by setting the “approval status” field to “Yes”. It is unique and thus the QME algorithm helps to avoid the duplicate submission in the KMNuR portal. This algorithm is further enhanced by employing rank of the matrix which is discussed in the next section.

Case 3: Duplicate Submission

When any user submits an ontology which is identical to the one existing in the KMNuR portal, the QME algorithm when executed will set the "approval status" field to "No" in the Knowledge_base database (as described in Chapter 3 Table 3.3). Then it will not be displayed in the KMNuR web portal. This eliminates the display of the redundant information in the KMNuR portal.

During the above implementation, in order to ascertain the matching of ontology, a matrix rank based ontology is employed. Partial overlap (or) duplicate (or) unique ontology are determined by matrix ranking methods. The result of ranking decides whether to eliminate or reuse or share the knowledge.

In order to optimise the memory usage while matching the ontologies, rank based ontology matching algorithm is proposed in the following section.

Figure 4.11 FBTR steam generator ontology concepts

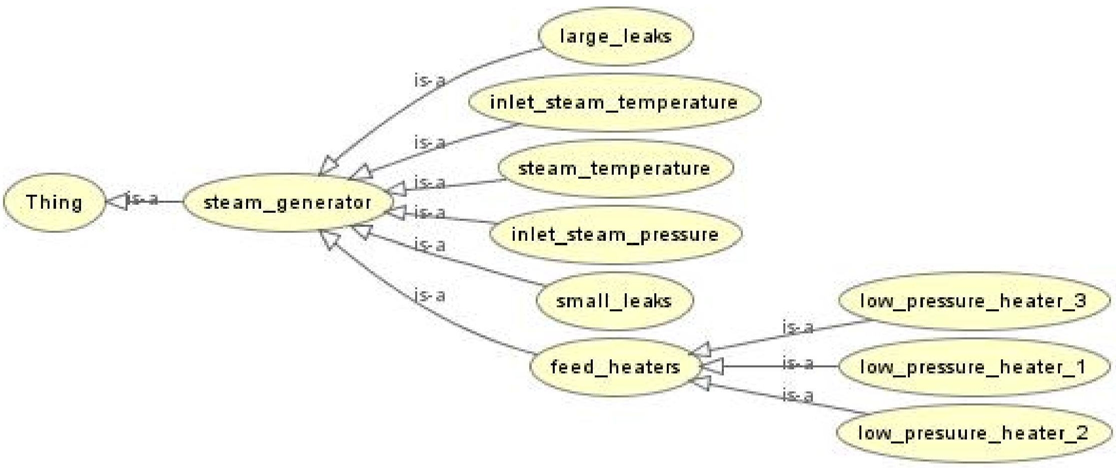


Figure 4.12 Steam generator system with partial overlapping concepts

4.3 Rank of a Matrix for Ontology Matching

4.3.1 Representation of Ontology Matching using Matrix Form

In String based techniques, strings are represented as sequences of alphabets and comparisons are done with respect to the lexical structures. To frame a matrix, string equality algorithm is utilized.

String equality is a similarity defined [205] as:

$$\sigma : S \times S \rightarrow [\,0\,,1]$$

$$such\ that\ \forall\, x, y \in S\,, \sigma\,(x, y) = 1$$

$$and\ if\ x \neq y, \sigma(x, y) = 0$$

Each entity of ontology O is matched with other ontology O' and the result of comparison is inserted into a matrix. If entity is exactly matched, then the matrix is filled with '1' else with '0'. The schematic diagram for matrix rank based ontology matching is shown in Figure 4.13.

4.3.2 Algorithm for Rank of a Matrix for Ontology Matching

Let A be m × n matrix over a field K. The column rank of A is the maximum number of linearly independent columns of A and the row rank of A is the maximum number of linearly independent rows of A [227]. The matrix rank indicates the number of independent rows (or) columns present in the matrix. Let A_1, A_2 and A_{ii} be the concepts defined in the ontology O and B_1, B_2 and B_{jj} defined in ontology O'. Being linearly independent implies that the number of columns (or) rows in ontology is matching. If both O and O' are different, then by finding the rank of the matrix, agent program can conclude whether they are identical (or) overlap (or) not related. Lower number of the matrix column or row decides maximum rank value.

The following algorithm is used for evaluating the nature of ontology, by finding the number of matching rows and columns in the matrix.

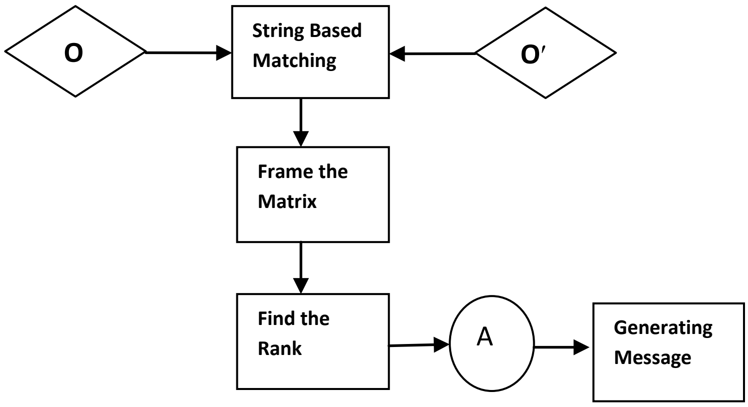


Figure 4.13 Schematic diagram for rank based ontology matching

Step 1. Find the match between each node of ontology 1 with each node of ontology 2 and frame the matrix using string based matching. Matrix is filled with '1' if there is a match, else with '0'

Step 2. Find the rank of the matrix

Step 2.1 Rank = 0 means the ontology is not matched with each other and it can be excluded

Step 2.2 Rank > 0 and Rank < m (where m is the size of the matrix) means partially overlapped or inclusion

Step 2.2.1 Identity matrix (rank=m) implies that it is a duplicate [228].

While finding the rank of the two ontologies, rank=0 signifies that ontologies are unique. If rank value lies between 0 and m (where m is the size of the matrix) it indicates that the two ontologies are related. Identity or unit matrix implies that it is a duplicate ontology.

4.3.2.1 Ontology Inclusion

If ontology O' includes ontology O, then ontology O' is the union of the definitions in O with those specific to O' [229]. When A_2 in ontology O is equal to B_2 in ontology O′ then it is called ontology inclusion. Ontology inclusion allows users to build large knowledge base by matching together a set of ontologies. The resultant matrix is shown in Table 4.1 (A) where A_2, B_2 is 1 and all other entries in the matrix are 0. When the rank is calculated its rank is 1. If rank value lies between 0 and m, then it is partially overlapped or inclusion is possible. This partial inclusion will help to reuse the ontology or share the knowledge of a particular domain.

4.3.2.2 Ontology Exclusion

A matrix is framed by comparing the rows and columns for a case shown in Table 4.1 (B) using string based matching and the rank of the matrix is obtained. Let O be mapped with O' then in this matrix all the rows and columns are zero and rank = 0 for zero matrix. If the rank is zero then it means the ontology O and O' are unique. This is also called exclusion of ontology meaning, a unique concept is defined for that application domain.

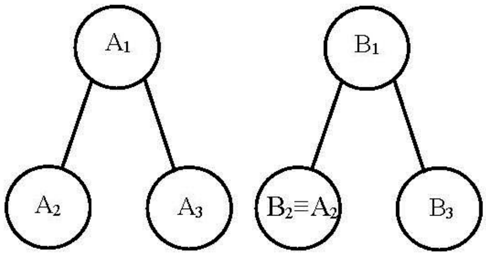
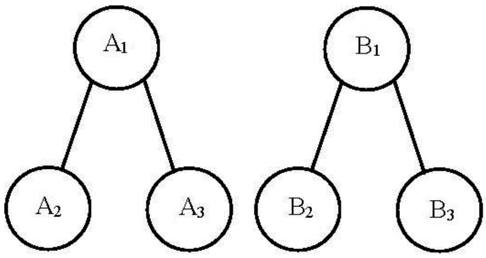
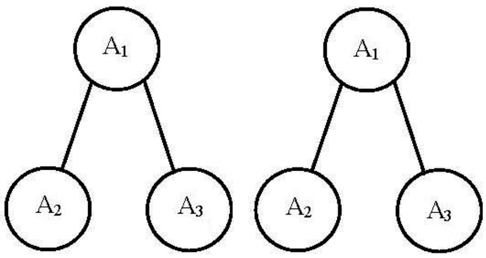
4.3.2.3 Ontology Duplication

If A_1 is equal to B_1, A_2 is equal to B_2, A_3 is equal to B_3 as shown in Table 4.1 (C), then they are duplicate. If rank = m, it is an identity matrix and hence the two ontologies are identical and duplicate. This is another duplicate version of the application domain and it will be rejected. If two versions of same ontology exist, then it will be omitted when the matrix constructed is found to be identity matrix. As a case study this algorithm is demonstrated in FBTR domain and is given below.

4.4 Rank based Ontology Matching Algorithm: A case study of FBTR

As a case study, steam generator and control rod drive mechanism systems are taken for analysis. FBTR has a common steam water circuit with four once through steam generator modules, which supplies super-heated steam to the condensing turbine. In steam generator, sodium and high pressure water/steam are separated by a thin wall boundary [226]. The steam generator produces super heated steam at a pressure of 125 bars and temperature of 480°C [230]. The steam which is at a relatively higher temperature and pressure gives high thermal efficiency of steam cycle [231].

Table 4.1 Rank of a matrix for different type of ontology matching

Type of Matching	**Schematic Diagram**	**Resultant Matrix**
Ontology Inclusion (A)		$\begin{array}{c\|cccccc} & B_1 & B_2 & . & . & B_m \\ \hline A_1 & 0 & 0 & . & . & 0 \\ A_2 & 0 & 1 & . & . & 0 \\ . & . & & & & \\ . & . & & & & \\ A_m & 0 & 0 & . & . & 0 \end{array}$ m=3, n=3, Rank = 1
Ontology Exclusion (B)		$\begin{array}{c\|cccccc} & B_1 & B_2 & . & . & Bm \\ \hline A_1 & 0 & 0 & . & . & 0 \\ A_2 & 0 & 0 & . & . & 0 \\ . & . & & & & \\ . & . & & & & \\ A_m & 0 & 0 & . & . & 0 \end{array}$ m=3, n=3, Rank = 0
Ontology Duplication (C)		$\begin{array}{c\|cccccc} & B_1 & B_2 & . & . & Bm \\ \hline A_1 & 1 & 0 & . & . & 0 \\ A_2 & 0 & 1 & . & . & 0 \\ . & . & & & & \\ . & . & & & & \\ A_m & 0 & 0 & . & . & 1 \end{array}$ m=3, n=3, Rank = 3

Control rod drive mechanism in FBTR consists of six control rod for rapid/control reactor shutdown and regulation of reactor power. This mechanism consists of an upper part and lower part in which the lower part is coupled to the upper part by an electro magnet. When the electromagnet is deenergised, the lower part that holds the control rod will drop due to gravity [232].

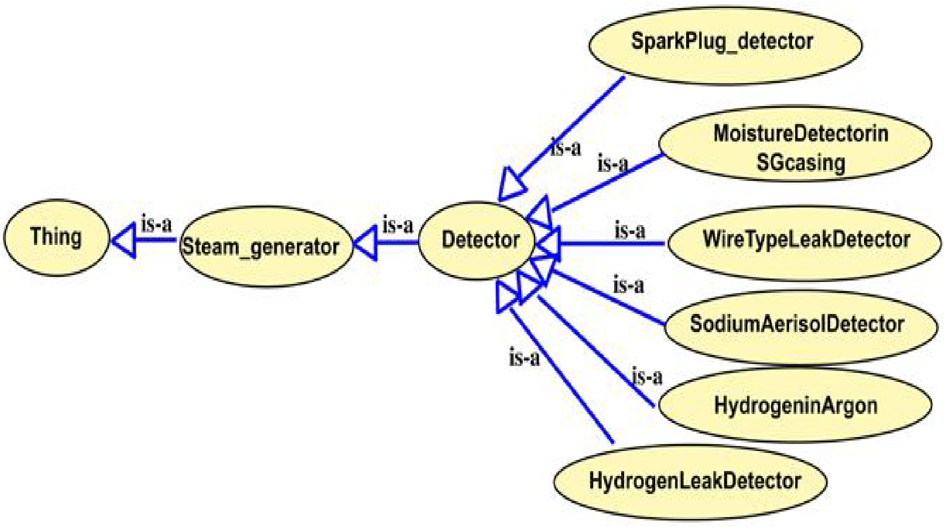


Figure 4.14 Ontology of steam generator detector systems

Case 1: Ontologies of steam generator detectors and steam generator protection systems developed are shown in Figures 4.14 and 4.15 respectively. Using matching algorithm these inputs are processed and an 8 x 16 matrix is constructed (Figure 4.16). It is found that the common matching entity is the 'steam generator' and a value of 1 is set in the matrix. In the matrix all values are '0' except in the first row and first column. When the rank of the matrix is calculated, it will yield value of 1, indicating that there is a partial overlap existing between these two ontologies leading to group and share the knowledge on steam generator.

Case 2: The ontology developed for control rod drive mechanism is shown in Figure 4.19. The ontologies for steam generator detector system and control rod drive mechanism are compared. An 8 x 18 matrix is constructed and all the rows and columns will be '0'resulting in the and there by the rank of the matrix being '0'. This is due to the fact that the ontology of the control rod related information and steam generator system are not directly connected. Hence, these two ontologies are unique dealing with distinct information.

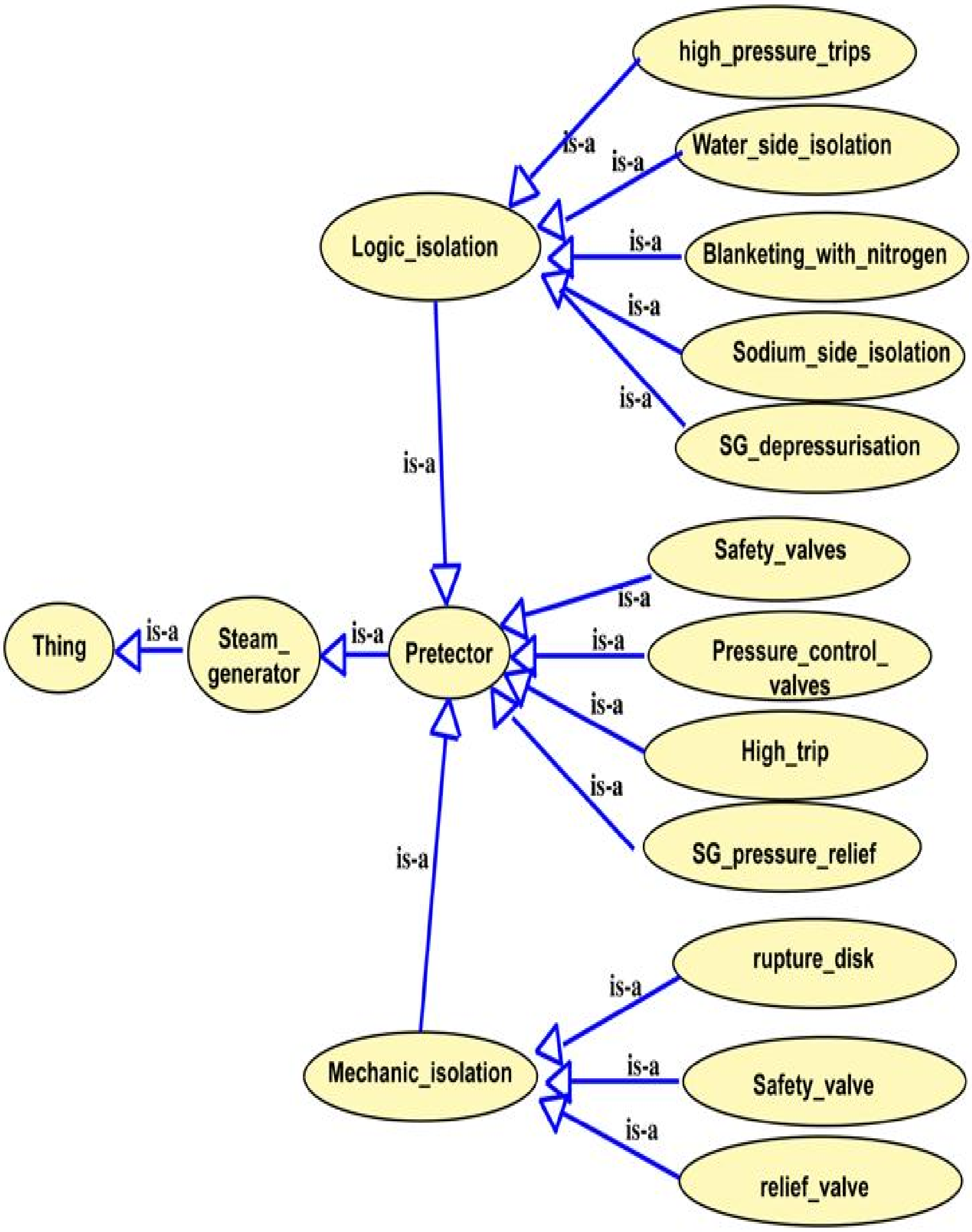


Figure 4.15 Ontology of steam generator protection systems

$$
\begin{array}{c|cccc}
 & B_1 & B_2 & \ldots\ldots\ldots\ldots & B_{16} \\
\hline
A_1 & 1 & 0 & \ldots\ldots\ldots\ldots & 0 \\
A_2 & 0 & 1 & \ldots\ldots\ldots\ldots & 0 \\
 & & & & \\
A_8 & 0 & 0 & \ldots\ldots\ldots\ldots & 1 \\
\end{array}
$$

Figure 4.16 Resultant matrix for steam generator for a case given in Figure 4.14 and Figure 4.15.

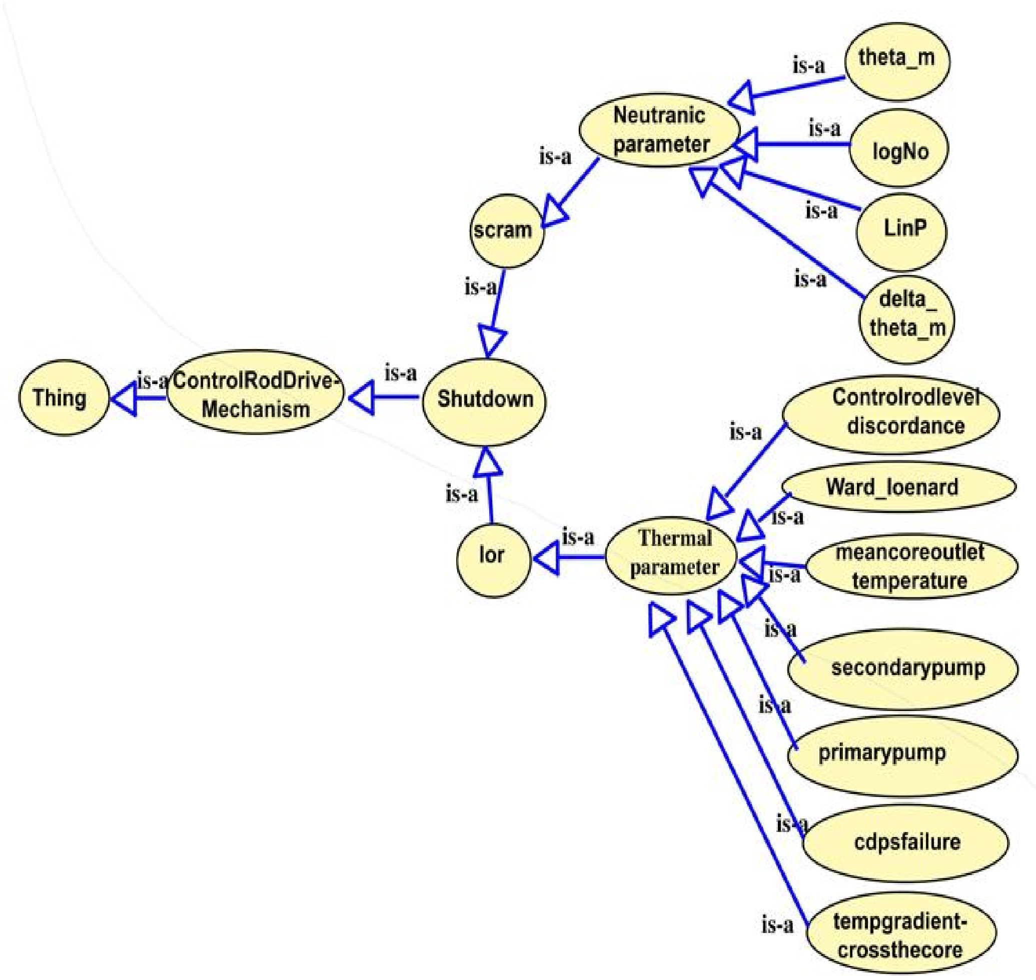


Figure 4.17 Ontology of control rod drive mechanism

$$
\begin{array}{c|l}
 & B_1\ B_2\ \ldots\ldots\ldots\ldots\ B_{18} \\
\hline
A_1 & 0\ \ 0\ \ \ldots\ldots\ldots\ldots\ 0 \\
A_2 & 0\ \ 0\ \ \ldots\ldots\ldots\ldots\ 0 \\
 & \\
A_8 & 0\ \ 0\ \ \ldots\ldots\ldots\ldots\ 0
\end{array}
$$

Figure 4.18 Resultant matrix for control rod drive mechanism for a case given in Figure 4.15 and Figure 4.17.

The matrix rank based algorithm will optimise the search time and retrieve the relevant knowledge. Thus the formulated algorithm is verified and successfully tested with FBTR systems. Usage of existence of matching algorithm with non matching algorithm are compared and shown in the Table 4.2.

Table 4.2 Comparison between the use of matching algorithm and unmatching algorithm

Characteristics	With Matching algorithm	Without Matching algorithm
Correspondence between ontologies	Yes	No
Possibility of ontology merging	Yes	No
Achievement of interoperability	Yes	No
Realization of sharing between heterogeneous system	Yes	No
Possibility of exploiting external resources	Yes	No
Possibility of reuse of knowledge	Yes	No
Possibility of synchronization between Ontologies	Yes	No
Possibility of discovering the hidden background knowledge		No
Existence of ontologies as a group	Yes	No (exist as an island)

In order to find the optimised solution for ontology matching, multi objective optimisation has been adopted. Details of methodology employed are described in the following section.

4.5 Similarity Calculation for Ontology Matching

In large size Ontologies, it is impossible to examine all possible mappings in order to select the best one. So there is a need for some computational method. In ontology matching framework, similarity is defined as measure of resemblance between different concepts. When comparing ontologies, a similarity measure is used to measure the similarity existing between two ontologies. It will determine whether the two ontologies can be merged or not. Merging of ontology will align data to be interpreted in different and meaningful ways. The decision to merge is determined by finding similar item in each ontology. The user can set an acceptable level of similarity between two ontologies and the decision to merge is determined by the machine. Similarity calculation is based on the determination of a family of similarity measures which assess the likelihood of equivalence in the ontology. Similarity measures is classified into lexical similarity and structural similarity.

Lexical similarity measures are concerned about lexical graphical similarity. It does not vary between iterations and therefore calculated during preprocessing. The id, the label, and the comment are the three such lexical features ontologies [233].

Structural similarity measures leverage hierarchical relationships among the concepts. It examines graph structures regardless of nature of concepts. In order to compare two graphs and compare the similarity between them we apply the methods which are employed to reveal similarity in ontology graph to perform alignment operation. Finding a set of coincidences across two ontologies is a key enabler for the success of semantic web. Coincidence of the structures of two ontologies can be represented as typed graphs which will enable to define a mechanism to score mappings across ontologies.

4.5.1 Ontology Alignment Weightage using Coincidence Matching Technique

Ontology matching to the space of matching two graphs was introduced by Haeri et al [234]. It targets scoring the mapping, based on the coincidence of ontologies appears for different mappings graphically. Concepts of ontology and relationships between different concepts are represented by nodes and typed graph in this model. The weightage function is based on the similarity measure across the concepts of the two ontologies which could evaluate a given alignment. The weightage is assigned based on similarity, more similar the mappings to the ideal mapping, more weight they gain. Similarly, if there is no alignment between the two ontologies then the weight assigned is very low. If there exist a partial overlapping then its weight lies in between the maximum and minimum weight.

Coincidence matching offers the possibilities of combining different ideas from different realms of ontology matching. All possible alignments are generated based on measure of the score, and finally opt to select the one having global maximum score. This method has limited application as suffers from exponential runtime and hence suitable only for small ontologies. Given that the phase one of ontology alignment gives us a measure for similarity of concepts across the ontologies, it is considered as an estimate for the distance between each concepts, and use for estimating the extent to which the two coincide. Alongside, an estimate is offered for the extent of coincidence between two edges, and then accumulates all these as final estimate for the coincidence of the two ontologies.

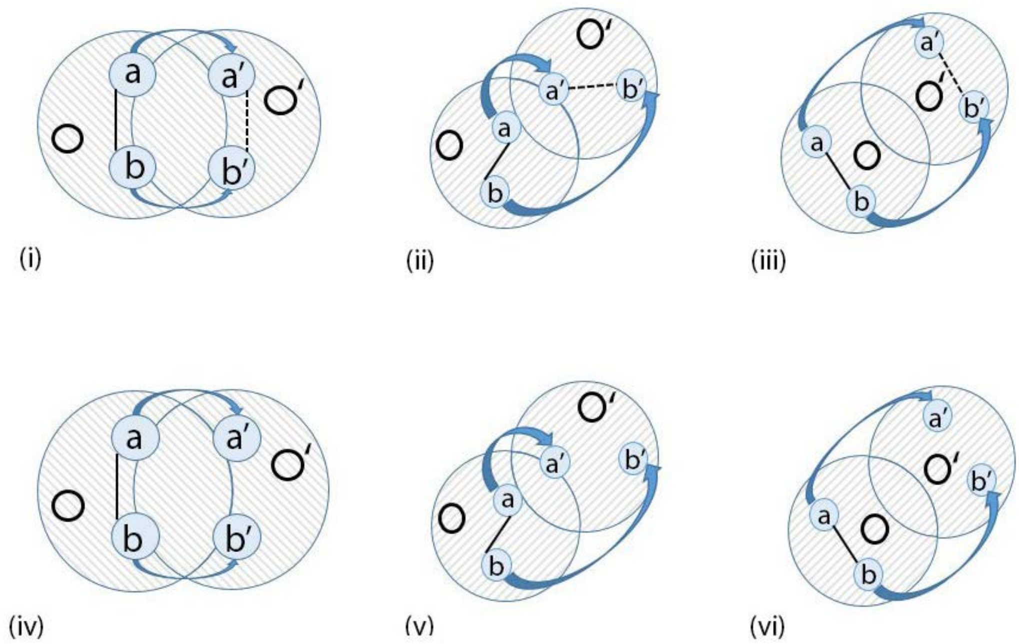


Figure 4.19 Categorization based on coincidence

As described already it is proposed to utilize a weight function for matching. The set of properties are divided into six cases, based upon preservation of the edge and upon the mutual distance between its heads. In all categories of Figure 4.19, O and O' are the input ontologies, a and b will be concepts in O, and, a' and b' concepts in O'. The arrows show mappings. And, the lines – be it solid or dotted – show the edges of graphs. The classified cases are given weightage depending on the coincidence exist between them and this is explained by taking nuclear reactor ontology as an example. Each category is explained by taking two ontologies and its coincidence is estimated and categorised to assign weightage to the attributes. Probability divergence is calculated based on weightage assigned for a set of three ontologies and this is explained in the forthcoming section.

It provides guidelines for assigning values to the concepts for matching. These six categories are listed below.

Category I: a and a' are too close like b and b' meaning that the two edges coincide considerably. The edge between them is also preserved and a greater weightage is assigned. It means that the two ontologies probably describe the same concepts. As they are conceptually describing the same concepts, they have to be assigned higher weight. To elucidate, consider the case when a and b are "nuclear reaction" and "uranium" respectively and a' *and* b' are "fission" and "plutonium". Fission is part of a nuclear reaction and both uranium and plutonium are reactor fuels, both the concepts resemble or meaning the same. The fact that there is an edge between both a and b, a' and b' means very much that the two ontologies are perhaps describing the same.

Category II: Edge between the pairs is preserved but only one of the pairs is closed to each other. Some of the concepts are partially overlapping so it should be assigned a weight less than the category I.

As an example of such case, one can consider O to be describing a “fast breeder reactor” and O' describing a “thermal reactor”. Furthermore, suppose that a and b are “fission” and “fast neutrons”, and a' and b' are “fission” and “slow neutrons”. An interpretation of this is that, a fast breeder reactor and thermal reactor are describing two different concepts and getting coincident from the side of fission concept. Therefore, the weight of such mapping is moderate and smaller than previous category.

Category III: Edge is not preserved while the relevant concepts are so close. In this till some relevant concepts exists between the two ontologies so weight should be assigned less than category II. Consider, for example, when O is describing a “nuclear power plant” while O' is the ontology of a “nuclear weapon unit”. In this respect, a and b could be “fission” and “energy”, and a' and b' also describing the same respectively. Nuclear power plant utilizes nuclear fission in a controlled form and for the production of energy. While in case of nuclear weapon unit, it is an uncontrolled process. The non-preservation of edge is considered as a negative point, but because the vertices coincide, this matching is not penalized that much. This is logical because the closeness of (a, a') and (b, b') means that the edge (a', b') is perhaps mistakenly missed.

Category IV: a and a' are close but b and b′ are far from each other. The edge between a and b is not preserved, while only one side of the edge is too close to what it is mapped to. This mapping is penalized more than previous category because only one end of the pairs is close to each other. The weight should be assigned less than category III. As an example of such a case, the following is considered: O is describing “nuclear power plant” and O' is a submarine. Assume that a is “nuclear reactor” and a' is also a “nuclear reactor”, while b is “stationary” and b' is “mobile”. Nuclear reactor utilized for generation of electricity is a stationary plant in homeland application. Submarines operated based on nuclear reactor. Like category III which is somehow dual of category I, this category can be considered dual of category II.

Category V and VI: In this category edges are far from each other and do not coincide that much. The difference between two ontologies is in edge preservation. Category V will receive a low benefit because of the edge preservation, while category VI will get highly penalized because neither nodes are close nor the edge between them is preserved. For an example, the case when a is “nuclear reaction” and b is “radiation” is considered”, a' is “cosmic rays” and b' is “radiation” (which is natural produced). Therefore, for the category V, it is desirable that the mapping receives a low benefit. In category VI, where a is “nuclear reactor” and b is “recycling” while a' is “waste management” and b' is “recycling”. A mapping of this category will be given very low weightage.

Table 4.3 and Table 4.4 summarize the six categories along with the suggested treatment for each case. The categories are used as the reference to define the user defined file for ontology comparison as a prerequisite. This is primary step to calculate and estimate the values in the user defined thesaurus file for calculating the optimised solution using Pareto optimisation.

Table 4.3 Ontology defined for each categories using nuclear reactor domain

Category	**Ontology O**		**Ontology O′**	
I	a. Nuclear reaction	b.Uranium	a′. Fission	b′. Plutonium
II	Fast breeder reactor		Thermal reactor	
	a. Fission	b. Fast Neutrons	a′. Fission	b′. Slow Neutrons
III	Nuclear power plant		Nuclear Weapon	
	a. Fission	b. Energy	a′. Fission	b′. Energy
IV	Nuclear power plant		Submarine	
	a.Nuclear Reactor	b.Stationary	a′.Nuclear Reactor	b′. Mobile
V	a. Nuclear reaction	b. Radiation Artificial	a′. Cosmic Rays	b′. Radiation Natural
VI	a. Nuclear Reactor	b. Recycling	a′. Waste Management	b′.Recycling

Table 4.4 Weights Assigned using coincidence technique

Proximity ⇨ Type of Edge ⇩	Both Ends Close	One End Close	Neither End Close
Preserved	High Benefit **Category I**	Modest Benefit **Category II**	Low Benefit **Category III**
Not Preserved	Low Penalty **Category IV**	Modest Penalty **Category V**	High Penalty **Category VI**

4.6 Pareto Optimisation

When users provide new information, the administrator of the ontology has to be provided a suitable merging mechanism to match the ontologies with the existing one in order to have proper information fusion of the relevant data. The possibility about merging multiple ontologies into a single viable ontology is proposed. For which, it is needed to focus on issues like, the ability to know when to merge or how successful the merging of ontologies could be. The multi objective optimisation is defined as

$$Minimize\ f(x) = [f1(x), f2(x), \dots fk(x)]^T$$

$F_i(x)$ is also called as objectives like criteria, pay off functions, cost functions or value functions [235].

If $f(s1) = \left(f1(s1), \dots \dots \dots \dots . fm(s1)\right)$ and $f(s2) = \left(f1(s2), \dots \dots \dots ., fm(s2)\right)$ are two solutions.

$s1$ is said to be dominate $s2$, iff $(s1 \succ s2)$ iff $\forall i : fi(s1) \leq fi(s2)$.

The set of n solution is sorted out s= { s1,….,sn} according to their Pareto front rank

f ={ F1,…Fk,… FK} [236]

With the following condition

1. $\forall\, si, sj \in Fk : \; si \nsucc si \wedge sj \nsucc si$
2. $\forall\, si \in Fk + 1, \forall sj \in Fk\text{:}\; si \succ sj$
3. $\forall\, si \in Fk + 1, si \nsucc sj \in Fk\text{:}\; sj \succ si$

There is no single global solution to a problem and it is necessary to determine the set of points that all fit into a predetermined optimum solution. The Pareto optimisation is the concept of dominance and is also called as multi objective optimisation. The potential solution to a problem can be classified into dominated and non dominated solution. A solution is Pareto optimal solution, if it is not dominated by any other feasible solution. The curve or surface composed of the Pareto-optimal solutions is known as the Pareto front.

A multi objective optimisation, has a set of optimal solution as Pareto front. In Pareto front each member is not dominated by others. All the members are optimal from viewpoint of one or more objectives, but none of them is optimal for all the objectives. The choice among the Pareto optimal solution is depend on the users defined threshold [237]. The goal of multi objective optimisation function is to generate various feasible solutions which are closer to Pareto front from which the best solution could be selected [238]. A multi-objective problem is often solved by combining its multiple objectives into one single-objective scalar function. An important issue in the multi objective optimisation is a quantitative comparison of performance of different algorithms with multi objective optimisation problems, and the knowledge about the Pareto-optimal front helps the decision maker in choosing the best compromised solution. That is the resulting Pareto front facilitates the selection of the most suitable ontology matching.

Three ontologies from the nuclear reactor domain are taken for the analysis of Pareto optimisation. The algorithms taken for this analysis are Kullback Leibler divergence, Cosine divergence, Levenshtein distance, Least Common subsequence and String similarity. Already the Levenshtein distance, Least Common subsequence and String similarity are explained initially in section 4.3.

4.6.1 Kullback Leibler Divergence

Kullback **L**eibler (KL) is used in information theory that quantifies the proximity of two probability distribution. It is used to measure dissimilarity. The highest similarity represented by a Kullback Leibler divergence is zero.

KL divergence can be directly defined as the mean of the log-likelihood ratio and it is the exponent in large deviation theory [239]. It is also known as relative entropy between two probability density function. KL divergence is only defined when p(x) > 0 and q(x) > 0 for all values of i and when p(x) and q(x) both sum to 1.

It is calculated using the equation.

$$D(p||q) = \sum_{x \in X} p(x). \log(\frac{p(x)}{q(x)})$$

Kullback Leibler divergence approach is developed for similarity measurement of texture features in the wavelet domain. KL is more appropriate and efficient in similarity measurement for features extracted by statistical models. It achieves a higher retrieval rate, while keeping the same level of computational complexity [240].

4.6.2 Cosine Divergence

The cosine similarity measure extents the keyword overlap to accommodate non binary weights associated with each keyword. Cosine similarity is a normalized metric because its value fall in [0,1]. Normalization avoids score dependency from document length.

$$\mathbf{cos}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}) = \frac{\sum_{x \in X, y \in Y} \boldsymbol{x} . \boldsymbol{y}}{\sqrt{\sum_{x \in X} \boldsymbol{x}^2 \; . \; \sum_{y \in Y} \boldsymbol{y}^2}}$$

If the resulting cosine similarity ranges from −1, meaning exactly opposite to 1, meaning exactly the same, with 0 usually indicating independence, and in-between values indicating intermediate similarity. Values are calculated for a given ontologies and the optimised solution is found out by finding the minimum among them. The property of cosine is that it carries the value “1” when the angle between them is zero. This means the two vectors lies directly on the top of each other. It carries the value “0” when the angle is 90 which mean the two vectors share no dimensions. Cosine similarity calculated in pattern recognition, medical diagnosis, machine learning, decision making and image processing in that similarity between fuzzy set. Efficiency of the cosine similarity measures for pattern recognition in medical diagnosis calculated and found to be reasonable [241].

It is used widely used in high dimensional data in text mining, information retrieval and bioinformatics. It is useful for measuring proximity in high dimensional space. It also measures hold symmetry, triangle inequality, null invariance and cross support properties etc [242].

The above values are calculated for the given ontologies and the optimised solutions are found out by finding the minimum among them. In order to find the optimised solution for ontology matching the Pareto based ontology matching algorithm is utilized.

4.6.3 Pareto based Ontology Matching Algorithm

Finding the best matching among the two ontologies is a difficult task. A Pareto rank learning scheme is utilized for enhancing multi objective evolutionary optimisation on problems that are intensive to compute [243]. The goal of multi objective optimisation algorithm is to generate trade-offs. Exploring all this trade-offs is particularly important because it provides the system designer or operator with the ability to understand and weigh the different choices available to them [244]. To be interoperable, information fusion system need a shared conceptualization of the application domain. Translation uncertainty is a natural consequence of the alignment process. In order to support information fusion and decision making, it is necessary to quantify and qualify such uncertainty. The goal is to qualify the uncertain associated with alignment of two ontological information fusion systems for decision making [245].The flowchart for the Pareto optimisation for ontology matching is shown in Figure 4.20.

A different weight to each probability distribution can be assigned. This makes it particularly suitable for the study of decision problems where the weights could be the prior probabilities. For certain applications such as in the study of taxonomy in nuclear reactor, one is required to measure the overall difference of more than two distributions. The problem of handling decentralized and distributed information about the domain through ontology is a challenging task. This section of the thesis briefly introduces the different approaches followed and their outcomes when comparing similar and dissimilar ontologies for matching.

Suppose it is desirable to merge ontology 1 and ontology 2 using predefined thesaurus files by associating each attribute. In this example for calculating the divergence the weightage assigned for each category is shown below.

For category I values assigned to attribute from range 25 to 30.

For category II range from 20 to 24

For category III range from 15 to 19

For category IV range from 11 to 14

For category V range from 6 to 10

For category VI range from 1 to 5

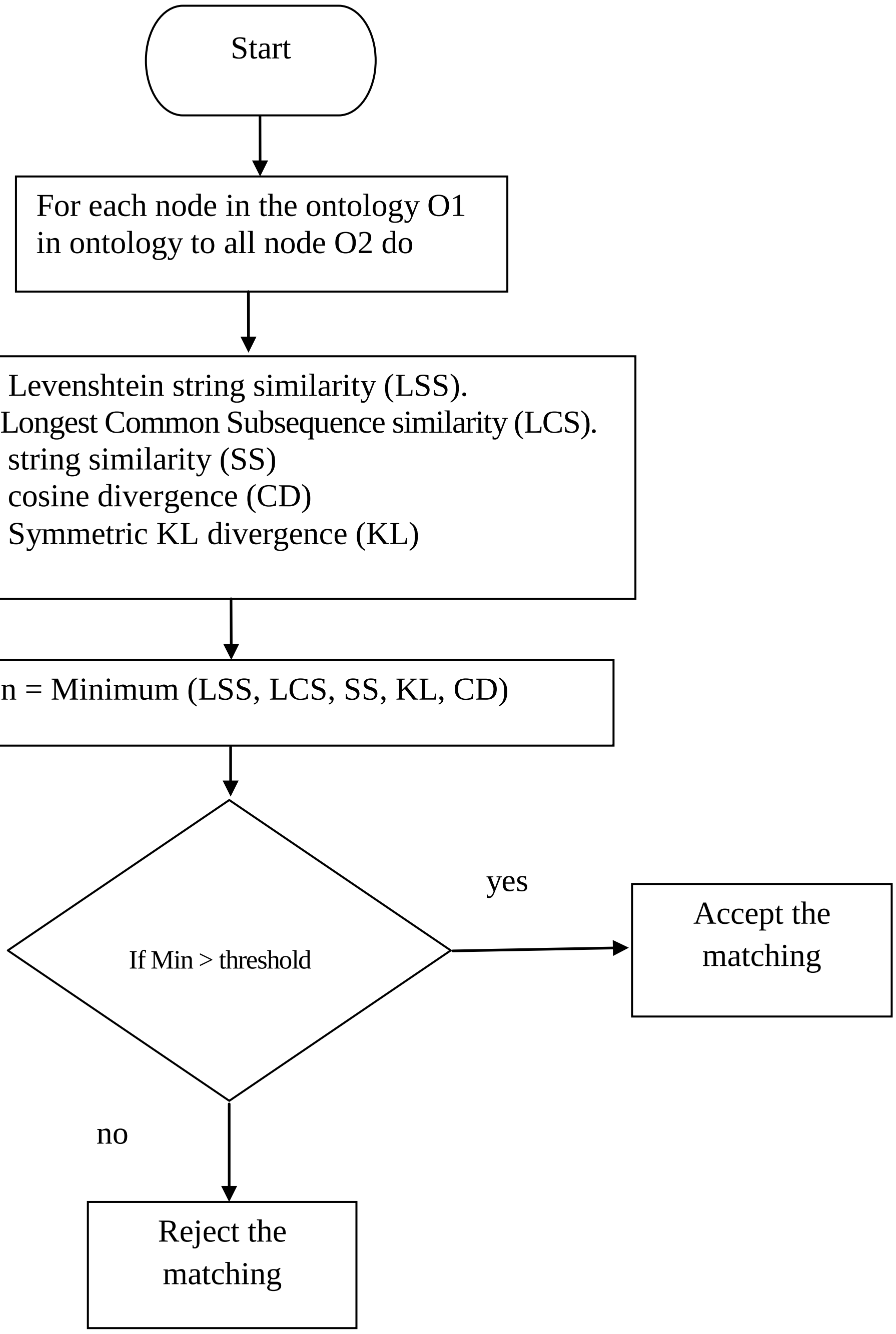


Figure 4.20 Flowchart for ontology matching using Pareto optimisation

To calculate the divergence the conditional probability is calculated using Bayes's theorem. Bayes's theorem could be instrumental in the merging and growth on ontologies. Bayesian learning generates probabilistic attribute-value rules which are based on the assumption of conditional independence between different attributes (A and B). The number of similar items in two different ontologies helps to arrive at a decision to merge them together as per the acceptable level set by the administrator.

$$p\left(\frac{A}{B}\right) = \frac{p\left(\frac{B}{A}\right).p(A)}{p(B)}$$

After calculation of the Bayesian probability, one needs to compute cosine and Kullback Leibler divergence for the two ontologies of interest. Ontologies should be represented in a separate file with information about its structure, terms, associations, roles etc along a separate associated thesaurus file containing information about how the items or terms get related. The probability of each concept (each in their own thesaurus file) is computed for each instance of a word. The higher the concept probability, possibility of merge decision is good. However, when it falls below a specified range the merging probability becomes less. In the merging process it does not necessitate to merge every term between two ontologies. Similar items are combined into one entity, those items not fit for merging are left disjoint. The user shall specify a cut-off level which, if met, will constitute a successful merge.

To demonstrate about merging of two ontolgoies in nuclear domain, a case study chosen is shown in Table 4.5. For simplicity only three attributes are considered in each ontology.

1. Merge is successful (ontology 1 and ontology 2)

2. Merge is unsuccessful (ontology 1 and ontology 3).

Table 4.5 Examples ontologies

Ontology 1	Steam generator Protection Isolation Leakage
Ontology 2	Steam generator system Detector system Protector system
Ontology 3	Radiation detection Detector Radiation monitors

Table 4.6 Associated attribute values using coincidence weight technique

	Steam Generator	Protection Isolation	Leakage	Steam Generator System	Detector System	Protector System
Steam Generator	-	30	18	12	8	21
Protection Isolation	30	-	21	18	6	25
Leakage	18	21	-	3	15	3
Steam Generator System	12	18	3	-	30	26
Detector System	8	6	15	30	-	24
Protector System	21	25	3	26	24	-

4.6.3.1 Successful Merge

Suppose it is desired to merge two ontologies (steam generator and steam generator system), one needs to count instances between the two attributes stored in pre-defined thesaurus files (Table 4.6). Probabilities for each attribute of steam generator have to be computed to each attribute of steam generator system using Bayes equation :

$$p\left(\frac{\text{steam generator}}{\text{steam generator system}}\right) = \frac{12}{\left(\frac{30+18+12+8+21}{1}\right)} .1 = 0.14$$

The conditional probability (P) is computed for all possible combinations between the two ontologies, the results are as given at Table 4.7. By choosing a proper percentage of values ontologies which cross the chosen value can be merged. In order to evaluate the cosine value, the vectors needed for the calculation are to be computed (Table 4.8). Large probability values depicted in Table 4.7 are taken to measure the cosine values for each pair and displayed in Table 4.9. Most of the cosine values are quiet high and one need to define some threshold value for the merge attempt. For deciding about the possibility of a potential merge, Kullback Leibler divergence is to be computed through the probability mass functions (Table 4.10) for each attribute.

Table 4.7 Values for p for steam generator and steam generator system

	Steam Generator	Protection Isolation	Leakage	Steam Generator System	Detector System	Protector System
Steam Generator	-	0.33	0.20	0.14	0.08	0.24
Protection Isolation	0.30	-	0.21	0.18	0.06	0.25
Leakage	0.30	0.35	-	0.05	0.25	0.05
Steam Generator System	0.13	0.20	0.03	-	0.33	0.30
Detector System	0.09	0.07	0.18	0.36	-	0.29
Protector System	0.21	0.25	0.03	0.26	0.24	-

Table 4.8 Attribute vectors for successful merge

Attribute	**Vector**
Protector system	(21,25,3,26,24,0)
Protection isolation	(30,0,21,18,6,25)
Steam Generator System	(12,18,3,0,30,26)
Detector system	(8,6,15,30,0,24)
Leakage	(18,21,0,3,15,3)
Steam generator	(0,30,18,12,8,21)

Table 4.9 Cosine measures for successful merge

Attribute	Cosine measure
cos (protection system, steam generator)	0.795
cos (protection system, protection isolation)	0.767
cos (steam generator system, protection isolation)	0.673
cos (detector system, leakage)	0.389

Table 4.10 Probability mass vectors successful merge

Attribute	Vector
Protector system	(0.21,0.25,0.03,0.26,0.24,0)
Protection isolation	(0.30,0,0.21,0.18,0.06,0.25)
Steam generator system	(0.13,0.20,0.03,0,0.33,0.30)
Detector system	(0.09,0.07,0.18,0.36,0,0.29)
Leakage	(0.30,0.35,0,0.05,0.25,0.05)
Steam generator	(0,0.33,0.20,0.14,0.08,0.24)

Table 4.11 Kullback Leibler divergence values for successful merge

Attribute	Kullback Leibler Divergence
D(protection system||steam generator)	0.13
D(protection system||protection isolation)	0.12
D(steam generator system||protection isolation)	0.18
D(detector system||leakage)	0.43

With the above calculated values we can now compute the KL divergence for each pair in Table 4.7 and the result of which is presented in Table 4.11. For example the value for (steam generator / steam generator system) is calculated. From the data calculated on cosine and Kullback Leibler functions the possibility whether the merge is possible or not can be explored. A minimum value of more than 0.60 is assigned for cosine data to be suitable for merge and for Kullback Leibler the values may be limited to have a maximum upto 0.30. Looking at the data presented in Table 4.9 and Table 4.11 one can see that atleast three of the four pairs are suitable for the merge. Within the specified range set for cosine and Kullback Leibler the same three pairs are found suitable for merge and hence it would be advisable to merge steam generator and steam generator system. One needs to understand that even though the merge is successful, there may be some spurious data since not every attribute met the criteria.

4.6.3.2 Unsuccessful Merge

To illustrate about unsuccessful merge, if the possibility of merging steam generator system ontology with radiation detection ontology is considered. The attribute values for the detector in steam generator and radiation system are not the same, even though the label is identical. Attributed values are calculated and this is shown in Table 4.12. Bayesian probabilities (Table 4.13), attribute vector (Table 4.14) cosine values (Table 4.15), probability mass vector (Table 4.16) and Kullback Leibler data are calculated and displayed for the four highest probabilities in Table 4.17. As per the norms set for higher cosine measure and smaller Kullback Leibler values, only one pair (detector system, detector) meets the stipulated requirement. In this case, the ontologies are not allowed to merge, as only one pair matches the set criteria.

It can be seen clearly from the Table 4.15 and Table 4.17 that whether the two ontologies can be merged or not. One should take note that the cosine value and KL cut off value based on which the merging process is driven, is a user defined parameter. While humans can understand that steam generator and radiation detector system are not common, for a machine this information has to be fed properly through the user defined parameters.

Table 4.12 Associated attribute values for an unsuccessful merge

	Steam Generator System	Detector System	Protector System	Radiation detection	Detector	Radiation monitors
Steam Generator System	-	30	18	2	21	2
Detector System	30	-	21	4	25	4
Protector System	18	21	-	5	2	10
Radiation detection	21	25	5	-	37	30
Detector	2	5	2	30	-	26
Radiation monitors	2	5	10	32	24	-

Table 4.13 Values for *p* for steam generator system and radiation detection

	Steam Generator System	Detector System	Protector System	Radiation detection	Detector	Radiation monitors
Steam Generator System	-	0.41	0.25	0.02	0.29	0.02
Detector System	0.36	-	0.25	0.05	0.30	0.05
Protector System	0.32	0.38	-	0.09	0.03	0.18
Radiation detection	0.18	0.21	0.04	-	0.31	0.25
Detector	0.03	0.08	0.03	0.46	-	0.40
Radiation monitors	0.03	0.07	0.14	0.44	0.33	-

Table 4.14 Attribute vectors for unsuccessful merge

Attribute	**Vector**
Radiation monitor	(2,4,10,30,26,0)
Detector system	(30,0,21,25,5,5)
Radiation detection	(2,4,5,0,30,32)
Detector	(21,25,2,37,0,24)
Protector system	(18,21,0,5,2,10)
Steam generator system	(0,30,18,21,2,2)

Hence by use of Bayesian probability and KL divergence, it is possible for a machine to make correct and valid decisions on merging two ontologies. Interoperability between two user submitted data is a challenging task till their internal structure is clearly understood. Hence the advent of ontology and the process of alignment or merging may be utilized without losing any valuable information about the domain of interest. Thus the Pareto optimisation technique for ontology matching is demonstrated by using KL divergence, cosine divergence, String Equality, Least common subsequence, Levenshtein distance algorithms.

Table 4.15 Cosine measures for unsuccessful merge

Attribute	Cosine measure
cos (radiation monitor, steam generator system)	0.58
cos (radiation monitor, detector system)	0.63
cos (radiation detection, detector system)	0.28
cos (detector, protector system)	0.81

Table 4.16 Probability mass vectors unsuccessful merge

Attribute	Vector
Radiation monitor	(0.03,0.07,0.14,0.44,0.33,0)
Detector system	(0.36,0,0.25,0.05,0.30,0.05)
Radiation detection	(0.18,0.21,0.04,0,0.31,0.25)
Detector	(0.03,0.08,0.03,0.46,0,0.40)
Protector system	(0.32,0.38,0,0.09,0.03,0.18)
Steam generator system	(0,0.41,0.25,0.02,0.29,0.02)

Table 4.17 Kullback Leibler divergence values for unsuccessful merge

Attribute	Kullback Leibler Divergence
D(radiation monitor ||steam generator system)	0.52
D(radiation monitor ||detector system)	0.35
D(radiation detection || detector system)	0.16
D(detector || protector system)	0.40

4.7 Summary

When a user submits knowledge, it is first verified by checking it with already existing knowledge in the KMNuR portal. The QME ontology matching algorithm elucidate whether the new knowledge – is of overlapping nature or duplicate or new knowledge. By setting the acceptance status flag in the database, the new knowledge is considered for acceptance. In heterogeneous system matrix, rank based ontology matching is deployed to find whether the alignment is duplicate or partial overlap or new. For a case of two ontologies existing in an open system, if the matrix mapping yields the rank value between 0 and m (size of matrix), then partial overlap exists between these ontologies enabling to category and share the knowledge. Rank value of 0, implies that two ontologies are unique implying ontology exclusion for that application domain. When rank = m, it is an identity matrix and the matrix constructed would be omitted if two versions of the same ontology exist. The algorithm proposed aids the agent program to filter and collect the knowledge for search, reuse, share and satisfied thereby giving relevant answers to the user query.

By using the Pareto optimisation technique the optimised solution is found for ontology matching. The successful and unsuccessful matches between ontologies cases from nuclear reactor domain are checked. In the Pareto optimisation a comparative study of different algorithms like Kullback Leibler divergence, Cosine divergence, Levenshtein string similarity, Least Common Subsequence, String Equality values are calculated and compare it with user defined threshold to find the optimised solution for matching.

CHAPTER 5

SUMMARY AND SUGGESTIONS FOR FUTURE WORK

5.1. Summary

The objective of the research work presented in this thesis is to semantically search a nuclear reactor domain by offering greater functionality and interoperability for automatic knowledge extraction by machine. The core issue of the work is to find out an optimised methodology for semantic heterogeneity problems and thereby enhancing interoperability. The semantic heterogeneity is solved by matching the ontology. Ontology versioning, inconsistency between data providers and mismatches between data providers and data users are the root causes for hindering the functioning of ontology matching. Other issues like duplication / overlap of ontology are also addressed.

Research contributions of the thesis are: QME algorithm, matrix rank based ontology matching algorithm, Pareto optimisation and development of knowledge management portal. QME is equipped with different algorithms like String Equality, Longest Common Subsequence and Levenshtein Distance for evaluation of the ontology matching at different thresholds. In QME, the users are allowed to change the type of match making algorithm. It also evaluates the results by using metrics like precision, recall and F-measure with different thresholds. User could customize the algorithm and match the ontology to extract the requisite knowledge.

A matrix rank based algorithm is deployed to calculate the ontology matching to an extension of QME. Partial overlap (or) duplicate (or) unique ontology are determined by matrix ranking methods. The result of the ranking decides whether to eliminate or reuse or share the knowledge and also the overall memory usage is optimised. Efficiency of matching is of prime importance in dynamic application. Other outcome of the research is to enhance the matching methodology by employing a hybrid approach of the terminological and structural similarity measures. Pareto based ontology matching algorithm is utilized to find an optimized solution amongst the conflicting values in the similarity measures. Optimization goal is to find a methodology which simultaneously minimizes the cost and execution time. Coincidence weightage technique is employed to provide well defined associated information in the ontology to make semi-automatic extraction of knowledge. Terminological similarity measures are concerned about lexical graphical similarity. It does not vary between iterations and therefore calculated during preprocessing. Structural similarity measures leverage hierarchical relationships among the concepts. It includes algorithm like Kullback divergence, cosine structural divergence (Structural similarity), string equality measure, Levenshtein distance and Longest Common Subsequence (Lexical similarity). It examines graph structures regardless of nature of concepts. Coincidence of the structures of two ontologies can thus be represented as typed graphs which will enable to define a mechanism to score mappings across ontologies. Finding a set of coincidences across two ontologies is a key enabler for the success of semantic web.

In order to understand the application of ontology in various domains, a literature survey is carried out in various domains like agriculture, aviation, biology, chemistry, civil engineering, computer science, medicine, literature, power plant, e-business. The survey revealed that ontology is rather very limited in the field of nuclear reactor domain. Considering the life cycle of any nuclear plant, the knowledge base creation not only helps to sustain the existing nuclear reactors but also in the design, construction, commission and operate the reactor likely to be built in the future. A semantic web based knowledge representation is designed to capture the knowledge existing in nuclear reactor domain and preserve it for future.

As a test bed, Knowledge management portal has been developed for FBTR and christened as KMNuR portal. KMNuR portal would cater to making the accumulated knowledge available as reference to operational personal and also to new force. A client/ server architecture has been adapted for the development of web pages. RDF, OWL, UML and graph formats are represented in the portal. Based on a literature survey, Protégé IDE tool is found to be suitable for RDF, OWL representation. KMNuR portal is to integrate and infer the requisite knowledge needed by the users in performing the nuclear reactor domain. It also allows the user to submit the new knowledge to the system. It is envisaged that domain experts provide information relating to the nuclear reactor and an administrator would facilitate making it available in the machine readable format and uploading in to the portal. Any portal should have a provision to acquire new or update information and this is filtered by using QME ontology matching algorithms.

5.2 Scope for Future Work

This research work contributes to the applications of semantic web and matching algorithms. Having successfully applied the semantic based knowledge management for nuclear reactor domain, there is still lot of scope for future work which are suggested below.

- Future development efforts would be put to develop appropriate tools for automating the process of converting the accumulated knowledge from the domain expert into machine readable format.
- The portal would be deployed across WWW to share knowledge relating to the fast breeder reactor enabling the use, share and process the nuclear knowledge.
- Enhancement and implementation of the portal to cover other forms of nuclear reactors such as thermal, boiled water reactor, pressurised heavy water reactor etc may be thought of.
- Standardizing the procedures for validating ontology format.

One of the challenges of semantic web is the mismatch caused by ontology inconsistency between data providers. Definitions of the same concept may be diverse among those ontologies and, in turn, in the data of those data providers. Even for the data collected and archived by the same organisation, there may be several versions of ontologies on a same topic, so the semantic mismatch may also exist within a single organisation. In general challenges are due to ontology versioning, ontology inconsistency between data providers and data users. Data in the study may vary in style, classification and nomenclature. To fix this problem it is recommended to standardize the procedures for validating ontology format.

- To develop a methodology to get authenticated information from the users.

  To avoid unwanted information submission in the ontology and to get high precision, users may be allowed to submit the information in an authenticated methodology.

- Multi objective optimisation by using vector based Genetic Algorithm to get better solution can be introduced. The performance can be evaluated in terms of Pareto-based solution measures.

- Detection of near duplicate in ontology to filter out false positives.

  To avoid plagiarism of ontology submission and increase the precision near duplicates are to be filtered out. Duplicate pages increase the size of search engine indexes and reduce the quality of search results.

- Integrating with INIS (International Nuclear Information System) thesaurus as background knowledge for ontology matching. INIS is the world's leading information system on the peaceful uses of nuclear science and technology. Currently, the INIS database contains over 3 million bibliographic records and almost 2,00,000 full-text nonconventional documents, consisting of scientific and technical reports and other non-copyrighted information.

- GUI based n:m mapping alignment can also be developed.

REFERENCES

[1] Getting Knowledge Management Right: Five Best Practices for a Better Service Experience
An Oracle White Paper (2011)
http://www.oracle.com/us/5-befit-knowlg-manag-cust-serv-wp-521298.pdf.

[2] Evaluation of Factors Influencing Knowledge Sharing Based on a Fuzzy AHP Approach
H.F. Lin, H.S. Lee and D.W. Wang
Journal of Information Science, 35 (2009) 25-44.

[3] The Role of Tacit Knowledge in Group Innovation
D. Leonard and S. Sensiper
California Management Review, 40 (1998) 112-132.

[4] Knowledge Management for Nuclear Industry Operating Organizations
IAEA-TECDOC-1510, (2006) 1-176.

[5] Context-based Electronic Health Record: Toward Patient Specific Healthcare
W. Hsu, R.K. Taira, S.El-Saden, H. Kangarloo and A.A. Bui
IEEE Transaction on Information Technology in Biomedicine, 16 (2012) 228-234.

[6] Knowledge Management Technologies and Applications—Literature Review from 1995 to 2002
S.H Liao
Expert Systems with Applications, 25 (2003) 155–164.

[7] A Portrait of the Semantic Web in Action
J. Heflin and J. Hendler
IEEE Intelligent Systems, 16 (2001) 54-59.

[8] What is a Knowledge Representation
R. Davis, H. Shrobe and P. Szolovits
AI Magazine, 14 (1993) 17-33.

[9] Knowledge and Software Modeling using UML
W. C. Christine
Software and Systems Modeling, 3 (2004) 294-302.

[10] Web Ontology Segmentation: Analysis, Classification and Use
J. Seidenberg and A. Rector
http://www.ra.ethz.ch/cdstore/www2006/devel-www2006.ecs.soton.ac.uk/programme/files/pdf/4026.pdf.

[11] Introduction: Approaches to Knowledge Representation
G. McCalla and N. Cercone
IEEE Computer, 16 (1983) 12-18.

[12] Knowledge Portals Ontologies at Work
S. Staab and A. Maedche
AI Magazine, 22 (2001) 63-75.

[13] Developing a University Research Web-based Knowledge Portal
N.B. Jones, D. Provost and D. Pascale
International Journal of Knowledge and Learning, 2 (2006) 106-118.

[14] Semantic Web Road Map
T.Berners Lee
http://www.w3.org/ DesignIssues/Semantic.html.

[15] Web Mining Model and its Application for Information Gathering
Y. Li and N. Zhong
Knowledge-Based Systems, 17 (2004) 207-217.

[16] Semantic Web Working Ontologist: Effective Modeling in RDFS and OWL
D. Allemang and J.Hendler
Morgan Kuffmann, 2011.

[17] An Ontology-based Methodology for Semantic Expansion Search
G.Zou, B. Zhang, Y. Gan and J.Zhang
IEEE 5th International Conference on Fuzzy Systems and Knowledge Discovery, (2008) 453-457.

[18] Foundations of Semantic Web Technologies
P. Hitzler, M. Krotzsch and S. Rudolph
Chapman and Hall (2009) 50-70.

[19] Semantic Search
R. Guha, R. McCool and E. Miller
International World Wide Web Conference, (2003) 700-709.

[20] An Analysis of Semantic Agent for Implementing Library Domain Portal
N. Madurai Meenachi and M. Sai Baba
SACTiL, VECC, Kolkatta, 2014.

[21] A Survey in Semantic Web Technologies-Inspired Focused Crawlers
H.Dong, F. K. Hussain and E. Chang
3rd IEEE International Conference on Digital Information Management, (2008) 934-936.

[22] Introduction to Semantic Search Engine
J. M. Kassim and M. Rahmany
International Conference on Electrical Engineering and Informatics, 2009.

[23] Introduction to Semantic Web and Semantic Web services
L.Yu
Chapman and Hall/CRC, (2007) 1-368.

[24] Personalization in a Semantic Taxonomy-Based Meta-Search Agent
W. Kim, L. Kerschberg and A. Seime
International Conference on Electronic Commerce, Austria, 2001.

[25] Semantic Memory
M. Quillian
Semantic Information Processing, (1968) 227–270.

[26] The Semantic Web
M. Daconta, L. Obrst and K. Smith
The Future, Wiley Publishing, 2003.

[27] Implementing Semantic Search in Library Web Portal
N. Madurai Meenachi and M. Sai Baba
International Conference on Trends in Knowledge and Information Dynamics, Bangalore, 2012.

[28] The Semantic Web
T.B. Lee, J. Hendler and O. Lassila
http://www.cs.umd.edu/~golbeck/LBSC690/SemanticWeb.html.

[29] XML
William R. Stanek
PHI Publication, 2002.

[30] Resource Description Framework (RDF):concepts and abstract syntax
World Wide Web Consortiumhttp://hdl.handle.net/10421/2427.

[31] Latest RDF Primer Versions
W3C working Group
http://www.w3.org/TR/rdf-primer/

[32] OWL Web Ontology Language Semantics and Abstract Syntax
P.P. Patel-Schneider, P. Hayes and I. Horrocks
http://www.w3.org/TR/owl-semantics/

[33] SPARQL Query Language for RDF
E. Prud'hommeaux and A. Seaborne
http://www.w3.org/TR/rdf-sparql-query/

[34] Ontologies: Bio-Ontologies: Their Creation and Design
Peter Karphttp
http://www.cs.man.ac.uk/~stevensr/tutorial01/master.ppt

[35] Ontology Evaluation
A. Gomez-Perez
Handbook on Ontologies, International Handbooks on Information Systems, (2004) 251.

[36] Ontologies and Knowledge Bases Towards a Terminological Clarification
N. Guarino and P. Giaretta
Towards Very Large Knowledge Bases, (1995) 25-32.

[37] Formal Ontology, Conceptual Analysis and Knowledge Representation
N.Guarino
The Role of Formal Ontology in the Information Technology, 43 (1995) 625–640.

[38] Toward Distributed use of Large-scale Ontologies
Swartout, R. Patil, K. Knight and T. Russ,
Proceedings of 10th Workshop on Knowledge Acquisition for Knowledge-Based Systems,.
http://www.isi.edu/isd/banff_paper/Banff_final_web/banff_96_final_2.html.

[39] Ease of Interaction Plus Ease of Integration: Combining Web2.0 and the Semantic Web in a Reviewing Site
T. Heath and E. Motta
Web Semantic: Science, Services and Agents on the World Wide Web, 6 (2008) 76-83.

[40] Combining RFID with Ontologies to Create Smart Objects
M. Grüninger, S. Shapiro, M.S. Fox and H. Weppner
International Journal of Production Research, 48 (2010) 2633-2654.

[41] Managing Semantic Content for the Web
A. Sheth, C. Bertram, D. Avant, B. Hammond, K. Kochut and Y. Warke
IEEE Internet computing, 6 (2002) 80-87.

[42] Ontology-based Simulation in Agricultural Systems Modelling
H. Beck, K. Morgan, Y. Jung, S.Grunwald, H.Y. Kwon and J. Wu
Agricultural Systems, 103 (2010) 463–477.

[43] Methodology for the Design and Evaluation of Ontologies
M. Gruninger and M.S. Fox
Workshop on Basic Ontological Issues in Knowledge Sharing, (1995) 1-10.

[44] Getting Started with Protege-Frames
E. Sachs
http://protege.stanford.edu/ doc/tutorial/get_started/getstarted.html# understanding_Ontologies.

[45] Ontology Research and Development: Part 2 - A Review of Ontology Mapping and Evolving
Y. Ding and S. Foo
Journal of Information Science, 28 (2002) 375-388.

[46] Using Explicit Ontologies in KBS Development
G.V. Heijst, A.T. Schreiber and B.J. Wielinga
International Journal of Human-Computer Studies, 46 (1997) 183-292.

[47] Semantic Web for Knowledge Sharing
N.J. Kings and J. Davies
Semantic Knowledge Management, (2009) 103-111.

[48] Knowledge Level Modelling: Concepts and Terminology
M. Uschold
The Knowledge Engineering Review, 13 (1998) 5-29.

[49] Ontologies and the Semantic Web
E.K. Jacob
Bulletin of American Society for Information Science and Technology 29 (2005) 19-22.

[50] Similarity for Ontologies - A Comprehensive Framework
M. Ehrig, P. Haase, M. Hefke and N. Stojanovic
13th European Conference on Information Systems in a Rapidly Changing Economy, Germany, 2005.

[51] A Comparative Study of Ontology Building Tools in Semantic Web Applications
B. Kapoor and S. Sharma
International Journal of Web and Semantic Technology, 1 (2010) 1-13.

[52] Some Observations on Mind Map and Ontology Building Tools for Knowledge Management
B.K.Sarker, P. Wallace and W. Gill
Ubiquity, 9 (2008).

[53] Ontology Development 101: A Guide to Creating Your First Ontology
N. F. Noy and D.L. McGuinness
"http://protege.stanford.edu/publications/ontology_development/ontology101-noy- mcguinness.html.

[54] Ontology
T. Gruber
Encyclopedia of Database Systems. Editor: L.Liu Springer-Verlag, (2009).

[55] Standard Upper Ontology Working Group (SUO WG)
J. Schoening
IEEE P1600.1. http://suo.ieee.org.

[56] Towards a Core Ontology for Information Integration
M. Doerr, J. Hunter and C. Lagoze
Journal of Digital information, 4 (2003).

[57] Ontology Potentials in Increasing Effectiveness of Safeguards and Nuclear Weapon Non-proliferation Activities
S. Laban, A.I.EI. Desouky, A.S. ElHefnawy and A.F. El-Gebaly
Symposium on International Safeguards, Future Verification Challenges, 2010.

[58] Ontology Management
S. Bloehdorn, P. Haase, Z. Huang, Y. Sure, J. Volker, F. Harmelen and R. Studer
Semantic Knowledge Management, (2009) 3-20.

[59] The Evolution of Protégé: An Environment for Knowledge-based Systems Development
J.H. Gennari, M. A. Musen, R.W. Fergerson, W. E. Grosso, M. Crubézy, H.Eriksson, N. F. Noy, and W. Samson Tu
International Journal of Human-Computer Studies, 58, (2003) 89-123.

[60] Top Braid Suite
Top Quadrant, Inc 2011
http://www.w3.org/2001/sw/wiki/TopBraid.

[61] Ontolingua
Stanford University
http://www.ksl.stanford.edu/software/ontolingua/.

[62] OntoStudio
http://semanticweb.org/wiki/OntoStudio.

[63] True Semantics
http://www.ontoprise.de/

[64] Survey on Web Ontology Editing Tools
S.C.Buraga, L.Cojocaru, O.C.Nichifor and P.Politechnica
Transactions on Automatic Control and Computer Science, 2006.

[65] The Semantic Web: Research and Applications
S.Bechhofer, M. Hauswirth, J. Hoffmann, M.Koubarakis
Lecture Notes in Computer Science, 5021 (2008).

[66] The Model Futures OWL Editor
http://www.modelfutures.com/owl/

[67] KAON2
http://kaon2.semanticweb.org/

[68] Integrated Collaboration Object Model (ICOM) for Interoperable Collaboration Services
http://docs.oasis-open.org/icom/icom-ics/v1.0/csprd03/icom-ics-v1.0-csprd03.html.

[69] DOE – The Differential Ontology Editor
http://www.eurecom.fr/~troncy/DOE/index.html.

[70] WebOnto: Discussing, Browsing, and Editing on the Web
J. Domingue, Tadzebao
Proceedings of the 11th Knowledge Acquisition for Knowledge-Based Systems
http://kmi.open.ac.uk/people/domingue/banff98-paper/ domingue.html.

[71] LinKFactory® : an Advanced Formal Ontology Management System
W. Ceusters and P. Martens
http://www.isi.edu/~blythe/kcap.../ LinKFactoryXWhiteXPaperXfinal.doc.

[72] The Visual Ontology Modeller
http://www.sandsoft.com/products.html.

[73] Knowledge Management with K-Infinity
J. Schuemmer
Intelligent Views
http://www.esug.org/data/ESUG2001/schuemmer.pdf.

[74] Shaping the Third Stage of Indian Nuclear Power Programme
http://dae.nic.in/writereaddata/.pdf_32.

[75] Safety Evalution of Indian Nuclear Power Plants BWRs AT Tarapur Atomic Power Station (TAPS-1&2)
http://www.npcil.nic.in/pdf/A1.pdf

[76] Two Decades of Operating Experience with the Fast Breeder Test Reactor
G.Srinivasan, K.V.Sureshkumar and P.V. Ramalingam
International Conference on Research Reactors: Safe Management and Effective Utilization, Australia, 2007.

[77] Development of Knowledge Portals for Nuclear Power Plants,
IAEA Nuclear Energy Series Technical Reports, NG-T-6.2. 2009.

[78] Knowledge Management for Nuclear Research and Development Organizations
IAEA Nuclear Energy Series Technical Reports, IAEA-TECDOC-1675 2012.

[79] Nuclear Knowledge Management: The GRS Realization
D. Beraha, T. Heigl and P. Puhr-Westerheide
International Journal of Nuclear Knowledge Management, 1 (2005) 351-361.

[80] Improved Organizational Performance with a KM System
P. Puhr-Westerheide
Joint ICTP-IAEA School of Nuclear Knowledge Management, Trieste, Italy. 2008,
http://www.iaea.org/nuclearenergy/nuclearknowledge/schools/NKMSchool/archive/2010/ trieste2010/11_Improving_perfor mance-Trieste2010(Kosilov).pdf

[81] Industry Operating Experience Process at Krško NPP
B. Bruno, B. Bozin and R. Cizmek
9th International Conference on Nuclear Option in Countries With Small and Medium Electricity Grids, Zadar, Croatia 2012.

[82] A Knowledge Management and Information Portal for the Kazakhstan Atomic Energy Committee
A. Kosilov, W. Mandl, Z. Pasztory, M. Gladyshev, M. Idrissova and T. Zhantikin
International Journal of Nuclear Knowledge Management, 3 (2009) 382-394.

[83] Preservation of Nuclear Information and Records, School of Nuclear Knowledge Management
A. Tolstenkov
http://www.iaea.org/nuclearenergy/nuclearknowledge/schools/NKMSchool/archive/2010/trieste2010/23_Tolstenkov_Preservatoin_of_Nuclear_Information_2.pdf.

[84] Nuclear Knowledge Portal for Supporting Licensing and Controlling Nuclear Activities in the Brazilian Nuclear Energy Commission
E. Gomes and F. Brage
International Journal of Nuclear Knowledge Management, 1 (2005) 244-254.

[85] NET DB on the ANENT Web Portal, IAEA Regional Workshop on the Development of a Web Portal for the ANENT
E.J. Lee, K.W. Han, Y.T. Kim and J.K. Lee
http://www.iaea.org/km/documents/A4_8.pdf.

[86] Development of ANENT Web Portal, IAEA Regional Workshop on the Development of a Web Portal for the ANENT
K.W. Han, E.J. Lee, Y.T. Kim and J.K. Lee,
Korea, 2005 http://www.iaea.org/ km/documents/A4_7.pdf.

[87] Long-term Nuclear Knowledge Management of Innovative Nuclear Energy Systems (INES) - A Case Study of the Japan Atomic Energy Research Institute (JAERI)
K. Yanagisawa, R.H. Bexdek and T. Sawada
Progress in Nuclear Energy, 50 (2008) 683-687.

[88] Ontology Matching
J. Euzenat and P. Shvaiko
Springer, 2007.

[89] Using Ant Algorithm to Derive Pareto Fronts for Multi-objective Setting of Emergency Service Facilities
N. Liu, B. Huang and X. Pan
Journal of the Transportation Research Board, 1935 (2005) 120-129.

[90] A Survey on Usage of Ontology in Different Domain
N. Madurai Meenachi and M. Sai Baba
International Journal of Applied Information Systems, 9 (2012) 46-55.

[91] Reengineering Thesauri for New Applications: The AGROVOC Example
D. Soergel, B. Lauser, A. Liang, F. Fisseha, J. Keizer and S. Katz
Journal of Digital Information, 4 (2004) 1-23.

[92] Using an AGROVOC-based Ontology for the Description of Learning Resources on Organic Agriculture
S.S. Alonso and M.A. Sicilia
Metadata and Semantics, (2009) 481-492.

[93] The AGROVOC Concept Server Workbench System: Empowering Management of Agricultural Vocabularies with Semantics
M. Sini, S. Rajbhandari, M. Amirhosseni, G. Johannsen, A. Morshed and J. Keizer
13th World Congress Scientific Information, Rural Development, (2010) 1-8.

[94] Study on Food Safety Semantic Retrieval System Based on Domain Ontology
Y. Yang, J. Du and M. Liang
IEEE Cloud Computing and Intelligence Systems, (2011) 40–44.

[95] Domain Ontology-based Construction of Agriculture Literature Retrieval System
L. Cao and L. He
4th International Conference on Wireless Communications, Networking and Mobile Computing, (2008) 1-3.

[96] Ontology-based Simulation of Water Flow in Organic Soils Applied to Florida Sugarcane
H.Y. Kwon, S. Grunwald, H.W. Beck, Y. Jung, S.H Daroub, T.A. Lang and K.T. Morgan
Agricultural Water Management, 97 (2010) 112–122.

[97] The Application Research of Description Logic in Civil Aviation Domain Ontology
W. Hong, G. Siting and W. Jing
International Conference on Management and Service Science, (2011) 1–4.

[98] The Plant Ontology Database: A Community Resource for Plant Structure and Developmental Stages Controlled Vocabulary and Annotations
S. Avraham, C.W. Tung, K. Ilic, P. Jaiswal, E.A. Kellogg, S. McCouch, A. Pujar, L.Reiser, S.Y. Rhee, M.M. Sachs, M. Schaeffer, L. Stein, P. Stevens, L. Vincent, F. Zapata and D. Ware
Nucleic Acids Research, 36 (2008) D449-D454.

[99] Plant Ontology (PO): A Controlled Vocabulary of Plant Structures and Growth Stages
P. Jaiswal, S. Avraham, K. Ilic, E.A. Kellog, S. McCouch, A. Pujar, L.Reiser, S.Y. Rhee, M.M. Sachs, M. Schaeffer, L. Stein, L.Vincent, D. Ware and F. Zapata
Comparative and Functional Genomics, 6 (2005) 388–397.

[100] Research of Plant Domain Knowledge Model Based on Ontology
F. Jing, Z.X. Pei and D.T. Yang
IEEE 3rd International Conf. on Innovative Computing Information Control, 2008.

[101] Transparent access to Multiple Bioinformatics Information Sources
C.A. Goble, R. Stevens, Ng. Gary, S. Bechhofer, N.W. Paton, P.G. Baker, M. Peim and A. Brass
IBM Systems Journal, 40 (2001) 532-551.

[102] BRENDA, AMENDA and FRENDA the Enzyme Information System: New Content and Tools in 2009
A. Chang, M. Scheer, A. Grote, I. Schomburg and D. Schomburg
Nucleic Acids Research, 37 (2009) D588-D592.

[103] Building a Chemical Ontology Using Methontology and the Ontology Design Environment, Mariano
M.F. López, A.G. Pérez and J.P. Sierra
IEEE Intelligent Systems, (1999) 37-46.

[104] Chemical Entities of Biological Interest: an Update
P. De. Matos, R. Alcántara, A. Dekker, M. Ennis, J.Hastings, K. Haug, I. Spiteri, S. Turner and C. Steinbeck
Nucleic Acids Research, 38 (2010) D249–D254.

[105] A Domain Ontology Construction Method Towards Healthy Housing
Y.L. Zhang, T. Li, L. Jie and Yu. ChongChong
2nd International Conference on Information Engineering and Computer Science, 2010.

[106] FinnONTO—Building the Basis for a National Semantic Web Infrastructure in Finland
E. Hyvonen
12th Finnish Artificial Intelligence Conference, 2006.

[107] The Computing Ontology: Application in Education
L.N. Cassel, G. Davies, W. Fone, A.Hacquebard, J. Impagliazzo, R.LeBlanc, J.C. Little, A. McGettrick and M. Pedrona
ACM SIGCSE Bulletin, 39 (2007) 171-183.

[108] SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications
H. Chen, F. Perich, T. Finin and A. Joshi
1st Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, (2004) 258–267.

[109] Developing Application Specific Ontology for Program Comprehension by Combining Domain Ontology with Code Ontology
H. Zhou, F. Chen and H. Yang
8th IEEE International Conference on Quality Software, (2008) 225-234.

[110] Developing a Domain Ontology for Software Product Management
A. Botzenhardt, A. Maedche and J. Wiesner
5th International Workshop on Software Product Management, (2011) 7-16.

[111] Automatically Learning Robot Domain Ontology from Collective Knowledge for Home Service Robots
D. Kang, E. Seo, S. Kim and H.J. Choi
11th International Conf. Advanced Commun. Technology, (2009) 1766-1771.

[112] Real-Time Compressed-Domain Spatio-temporal Segmentation and Ontologies for Video Indexing and Retrieval, Circuits and Systems for Video Technology
V. Mezaris, I. Kompatsiaris, N.V. Boulgouris and M.G. Strintzis
IEEE Transactions on Circuits and Systems for Video Technology, 14 (2004) 606–621.

[113] Applications of Ontologies in Software Engineering
H.J. Happel and S. Seedorf
2nd International Workshop on Semantic Web Enabled Software Engineering, 2006.

[114] Ontology Mapping for the Interoperability Problem in Network Management
A.K.Y. Wong, P. Ray, N.Parameswaran and J. Strassner
IEEE Journal on Selected Areas in Communications, 23 (2005) 2058-2068.

[115] Image Classification Using Neural Networks and Ontologies
C. Breen, L. Khan and A. Ponnusamy
IEEE Conference on Database and Expert Systems Applications, (2002) 98-102.

[116] SymOntoX: a Web-ontology Tool for Ebusiness Domains
M. Missikoff and F. Taglino
4th International Conference on Web Information Systems Engineering, (2003) 343–346.

[117] A Study on Searching and Recommending SCORM CPs by Ontology
J.T. Yang, M.J. Hwang and Y.F. Chu
5th IEEE International Conference on Advanced Learning Technologies, 2005.

[118] The New Challenges for E-learning: The Educational Semantic Web
L. Aroyo and D. Dicheva
Journal of Educational Technology and Society, 7 (2004) 59-69.

[119] Minimalist Approach to Support Ontology-driven Educational Information Systems Communication
D. Dicheva and L. Aroyo
International Workshop on Applications of Semantic, 2004.

[120] Ontological Web Portal for Educational Ontologies
D. Dicheva, S. Sosnovsky, T. Gavrilova and P. Brusilovsky
International Workshop Applications of Semantic Web Technologies for E-Learning, 2005.

[121] Sharing an ontology in Education: Lessons Learnt from the OURAL project
M.Grandbastien, F.Azouaou, C.Desmoulins, R.Faerber, D.Leclet, Q. Joiron
7th IEEE International Conference on Advanced Learning Technologies, 2007.

[122] Study of Learning Source Ontology Modeling in Remote Education
D. Xiaopeng and L. Xu
Proceedings International Conference Multimedia Technology, (2010) 1–4.

[123] Ontology Based Application Framework for Network Education Resources
F. Yang and Y. Chen
2nd International Workshop Education Technology Computer Science, (2010) 423–426.

[124] Designing a Semantic Web Ontology for E-learning in Higher Education
M. Bucos, B. Dragulescu and M. Veltan
IEEE 9th International Symposium Electronics and Telecommunications, (2010) 415–418.

[125] Integration at Vocational Education and Training Level through Mapped Ontologies
C. Cubillos, F. Lamberti and C. Demartini
3rd International Conference on Convergence and Hybrid Information Technology, (2008) 117–122.

[126] An Ontology Based Architecture to Support the Knowledge Management in Higher Education
A. Laoufi, S. Mouhim, E.H. Megder and C. Cherkaoui
International Conference Multimedia Computing and Systems, (2011) 1–6.

[127] Virtual Laboratory Ontology for Engineering Education
J.P. Blázquez, I.G. Torà, J.H. Joancomartí and A.G. Roldán
38th ASEE/IEEE Frontiers in Education Conference, (2008) S2F-1-S2F-6.

[128] Approach to Create Domain Ontologies for Higher Education in Economics
J. Mesaric and B. Dukic
29th International Conf.on Information Technology Interfaces, (2007) 75–80.

[129] Cultural Artefacts in Education: Analysis, Ontologies and Implementation
C. Stewart, K. Chandramouli, A. Cristea, T. Brailsford and E. Izquierdo
IEEE International Conference on Computer Science and Software Engineering, (2008) 706–709.

[130] Sahayika: A framework for Participatory Authoring Knowledge Structures for Education Domain
P.K. Bhowmick, S. Bhowmick, D. Roy, S. Sarkar and A. Basu
International Conference on Information and Communication Technologies and Development, (2007) 1–11.

[131] Research on Remote Education System Using Semantic Web
Z. Wang
International Conference on Computer Science and Network Technology, (2011) 2815-2818.

[132] A Sustainable ICT Education Ontology
K.L. Chin and E. Chang
IEEE 5th International Conference Digital Ecosystems and Technologies, (2011) 350-354.

[133] A Universal Ontology for Sensor Networks Data
M. Eid, R. Liscano and A.El. Saddik
IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2007.

[134] Semantic web for Earth and environmental terminology (SWEET)
R. Raskin
Workshop on Semantic Web Technologies for Searching and Retrieving, Scientific Data, 2003.

[135] Knowledge Representation in Semantic Web for Earth and Environmental Terminology
R.G. Raskin and M. J. Pan
Computers and Geosciences, 31 (2005) 1119–1125.

[136] Semantic Integration in Geosciences
Z. Malik, A. Rezgui, B. Medjahed, M. Ouzzani and A.K. Sinha
International Journal of Semantic Computing, 4 (2010) 301–330.

[137] Ontology Based Semantic Metadata for Geo-Science Data
V. Parekh, J.P. Gwo and T. Finin
International Conference on Information and Knowledge Engineering, (2004) 485-490.

[138] Ontology-based Discovery of Geographic Information Services—An Application in Disaster Management
E. Klien, M. Lutz and W. Kuhn
Journal Computers, Environment and Urban Systems, 130 (2006) 102–123.

[139] Research and Application of Geological Hazard Domain Ontology
G. Liu, Y. Wang and C. Wu
18 th International Conference on Geoinformatics, (2010) 1-6.

[140] An Ontology for Modelling Human Resources Management Based on Standards
A.G. Perez, J. Ramirez and B.V. Terrazas
Lecture Notes in Computer Science, (2007) 534-541.

[141] Challenges in Building Domain Ontology for Minority Languages
P. Talita, A.W. Yeo and N. Kulathuramaiyer
International Conference on Computer Applications and Industrial Electronics, (2010) 574–578.

[142] A Process for Building Domain Ontology: An Experience in Developing Solat Ontology
S. Saad, N. Salim, H. Zainal and Z. Muda
IEEE International Conf. on Electrical Engineering Informatics, (2011) 1-5.

[143] Design and Implementation of Chinese Ancient Poetry Learning System Based on Domain Ontology
R. Yao and J. Zhang
IEEE International Conference on e-Education, e-Business, e-Management and e-Learning, (2010) 460–463.

[144] ENGOnto: Integrated Multiple English Learning Ontology for Personalized Education
F. Zhu, A.W.P Fok, H.S. Horace and J. Cao
IEEE International Conference on Computer Science and Software Engineering, (2008) 210-213.

[145] Semantic Digital Library - Migration of Dublin Core to RDF
E.Soundarajan, N. Madurai Meenachi and M. Sai Baba
IEEE International Conference on Signals and Image Processing, Chennai, (2010) 250 – 254.

[146] Ontology Based Association Rule Mining For Library Domain
N.Madurai Meenachi and M. Sai Baba
READIT-2011, Kalpakkam

[147] Ontology-based Metadata Schema for Chinese Digital Libraries
M.S. Lai and X.D. Yang
“http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.135.7833. 2001.

[148] Automatically Constructing a Domain Ontology for Document Classification
Y. Chang
IEEE 6th International Conference on Machine Learning and Cybernetics, (2007) 1942–1947

[149] A Marine Platforms Ontology: Experiences and Lessons
L.Bermudez, J. Graybeal and R. Arko
The Semantic Sensor Networks Workshop 2006.

[150] Modeling a Probabilistic Ontology for Maritime Domain Awareness
R.N. Carvalho, R. Haberlin, C. Paulo, G. Costa, K. B. Laskey and K.C. Chang
14th International Conference Information Fusion, (2011) 1–8.

[151] Research on Building Ocean Domain Ontology
Y. Hongyan
Second International Workshop on Computer Science and Engineering, (2009) 146–150.

[152] A Meronomic Relatedness Measure for Domain Ontologies Using Concept Probability and Multiset Theory
P. Witherell, S. Krishnamurty, I. Grosse and J. Wileden
28th North American Fuzzy Information Processing Society Annual Conference, 2009.

[153] Mathematical Knowledge Representation for Education Semantic Web Based on Learning Style
I.F.M Mahtar and N.A.M Zin
International Symposium on Information Technology, 2010.

[154] The Gene Ontolgoy (GO) Database and Informatics Resource
M.A. Harris, J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, K.Eilbeck, S. Lewis, B. Marshall, C. Mungall, J. Richter, G.M. Rubin, J.A.Blake, C. Bult, M. Dolan, H. Drabkin, J.T. Eppig, D.P. Hill, L. Ni, M.Ringwald, R. Balakrishnan, J.M. Cherry, K.R. Christie, M.C. Costanzo, S.S. Dwight, S. Engel, D.G. Fisk, J.E. Hirschman, E.L. Hong, R.S. Nash, A.Sethuraman, C.L. Theesfeld, D. Botstein, K. Dolinski, B. Feierbach, T.Berardini, S. Mundodi, S.Y. Rhee, R. Apweiler, D. Barrell, E. Camon, E.Dimmer, V. Lee, R. Chisholm, P. Gaudet, W. Kibbe, R. Kishore,

E.M.Schwarz, P. Sternberg, M. Gwinn, L. Hannick, J. Wortman, M.Berriman, V. Wood, N. Cruz, P. Tonellato, P. Jaiswal, T. Seigfried and R.White
Nucleic Acids Research, 32 (2004) D258-D261.

[155] Uberon: Towards a Comprehensive Multi-Species Anatomy Ontology
M.A. Haendel, G.V. Gkoutos, S.E. Lewis and C.J. Mungall
Nature Proceedings, International Conference on Biomedical Ontology 2009.

[156] The MGED Ontology: A Resource for Semantics-based Description of Microarray Experiments
P.L. Whetzel, H. Parkinson, H.C. Causton, L. Fan, J. Fostel, G. Fragoso, L. Game, M. Heiskanen, N. Morrison, P. Rocca-Serra, S.A. Sansone, C. Taylor, J. White and C.J. Stoeckert,
Bioinformatics, 22 (2006) 866-873.

[157] The Mouse Genome Database (MGD): Mouse Biology and Model Systems
C.J. Bult, J.T. Eppig, J.A. Kadin, J.E. Richardson and J.A. Balke
Nucleic Acids Research, 36 (2008) D724-D728.

[158] The Clinical Bioinformatics Ontology: A Curated Semantic Network Utilizing Refseq Information
H. Hoffman, C. Arnoldi and I. Chuang
Pacific Symposium on Biocomputing, (2005) 139-150.

[159] An Ontology for PACS Integration
C.E. Kahn, D.S. Channin and D.L. Rubin
Journal of Digital Imaging, 19 (2006) 316-327.

[160] Ontology-Based Error Detection in SNOMED-CT
W. Ceusters, B. Smith, A. Kumar and C.Dhaen, M. Fieschi et al. (Eds)
MEDINFO, IOS Press, (2004) 482-486.

[161] The Unified Medical Language System (UMLS): Integrating Biomedical Terminology
O. Bodenreider
Nucleic Acids Research, 32 (2004) D267-D270.

[162] GoMiner: A Resource for Biological Interpretation of Genomic and Proteomic Data
B. R. Zeeberg, W. Feng, G. Wang, M.D. Wang, A.T. Fojo, M. Sunshine, S.Narasimhan, D.W. Kane, W.C. Reinhold, S. Lababidi, K..J Bussey, J.Riss, J.C. Barrett and J.N. Weinstein
Genome Biology. 4 (2003) R28.

[163] The Fungal Web Ontology: Application Scenarios
C.J.O. Baker, R. Witte, A.S. Nejad, G. Butler and V. Haarslev
4th International Semantic Web Conference on Bioinformatics and Genomics, 2005.

[164] The Fungal Web Ontology: Semantic Web Challenges in Bioinformatics and Genomics
A.S. Nejad, C.J.O. Baker, V. Haarslev and G. Butler
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.6919&rep=rep1&type=pdf.

[165] Ontology-Based Extraction and Summarization of Protein Mutation Impact Information
N. Naderi and R.Witte
Workshop on Biomedical Natural Language Processing, (2010) 128-129.

[166] An Example of Food Ontology for Diabetes Control
J. Cantais, D. Dominguez, V. Gigante, L. Laera and V. Tamma
Workshop on Ontology, 2005.

[167] Modeling and Building an Ontology for Neuropediatric Physiotherapy Domain
L.V. Castilho, H.S. Lopes and C.A. Tacla
8th International Conference on Hybrid Intelligent Systems, (2008) 210–215.

[168] A Domain Ontology for Mechanical Circulatory Support Systems
C.De. Lazzari, E. Guerrieri and D.M. Pisanelli
IEEE Computers in Cardiology, (2003) 417–419.

[169] Ontology-based Information Retrieval
J. Paralic and I. Kostial
14th International Conference on Information and Intelligent systems, 2003.

[170] NCBO Resource Index: Ontology-Based Search and Mining of Biomedical Resources
C. Jonquet, P. LePendu, S.M. Falconer, A. Coulet, N.F. Noy, M.A. Musen and N.H. Shah
Web Semantics: Science, Services and Agents on the World Wide Web, 9 (2011) 316-324.

[171] The National Cancer Institute's Thésaurus and Ontology
J. Golbeck, G. Fragoso, F. Hartel, J. Hendler, J. Oberthaler and B. Parsia
Journal of Web Semantics: Science, Services and Agents on the World Wide Web, 1 (2003) 75–80.

[172] A Relation-Centric Query Engine for the Foundational Model of Anatomy
L.T. Detwiler, E. Chung, A. Li, J.L.V. Mejino, A.V. Agoncillo, J.F. Brinkley, C.Rosse and L.G. Shapiro, M. Fieschi
MEDINFO, (2004) 341-345.

[173] Research on Domain Ontology Construction in Military Intelligence
M.Y. Jia, B. Yang, D. Zheng and W. Sun
3rd International Symposium Intelligent Information Technlogy Application, (2009) 116–119.

[174] 5W1H-based Conceptual Modeling Framework for Domain Ontology and its Application on STOP
L. Yang, Z. Hu, J. Long and T. Guo
IEEE 7th International Conference on Semantics, Knowledge and Grids, (2001) 203-206.

[175] Study on Construction and Integration of Military Domain Ontology Situation Ontology and Military Rule Ontology for Network Centric Warfare
J. Song, Z. Ming, W, Xiao and Z. Xu
IEEE International Conference Publication on e-Technology, e-Commerce and e-Service, (2005) 368–373.

[176] Web Information Extraction Based on News Domain Ontology Theory
J. Shi and L. Li
IEEE 2nd Symposium on Web Society, (2010) 416-419.

[177] Construction of Ontology-based Safety Assessment System for Power Plants
Z. Yang, C. Cheng and Z. Feng
IEEE International Conference on Networking, Sensing and Control
(2008) 1092-1096.

[178] A Steam Turbine Diagnostic Maintenance System Based on Evaluative Domain Ontology
M.T. Khadir and S. Klai
International Conference on Machine Web Intelligence, (2010) 360–367.

[179] Application of an Ontology in a Transport Domain
M. Merdan, G. Koppensteiner, I. Hegny and B.F. Bulle
IEEE International Conference on Industrial Technology, (2008) 1-6.

[180] Web Ontology Language Editors for Semantic Web – A Survey
N. Madurai Meenachi and M. Sai Baba
International Journal of Computer Applications, 53 (2012) 12-16.

[181] OntoStudio as a Ontology Engineering Environment
M. Weiten
Semantic Knowledge Management, (2009) 51-60.

[182] The Evolution of Protégé: An Environment for Knowledge-based Systems Development
J.H. Gennari, M.A. Musen, R.W. Fergerson, W.E. Grosso, M. Crubézy, H.Eriksson, N.F. Noy and S.W. Tu
International Journal of Human-Computer Studies, (2003) 89-123.

[183] Protégé A Tool for Managing and Using Terminology in Radiology
D.L. Rubin, N.F. Noy and M.A. Musen
Journal of Digital Imaging, 20 (2007) 34-46.

[184] OWLGrEd: a UML Style Graphical Editor for OWL
J. Barzdins, G. Barzdiņs, K. Cerans, R. Liepiņs and A. Sprogis
http://owlgred.lumii.lv/ 2010.

[185] Swoop: A Web Ontology Editing Browser
A. Kalyanpur, B. Parsia, E. Sirin, B.C. Grau and J. Hendler
Web Semantics: Science, Services and Agents on the World Wide Web, 4 (2006) 144–153.

[186] Hypermedia Inspired Ontology Engineering Environment: SWOOP
A. Kalyanpur, E. Sirin, B. Parsia and J.S. Hendler
3rd International Semantic Web Conference (2004).

[187] Web Services / Agent-Based Model for Inter-Enterprise Collaboration
A.S. Namin and W.Shen
International Federation for Information Processing Digital Library, Emerging Solution Future Manufacturing Systems, (2005) 231-240.

[188] Development of Semantic Web-based Knowledge Management for Nuclear Reactor (KMNuR) Porta N.Madurai Meenachi and M. Sai Baba
DESIDOC Journal of Library and Information Technology, 34 (2014) 426-434.

[189] Intelligent Agents for Nuclear Domain
D. Persis, N. Madurai Meenachi and M. Sai Baba
READIT-2009, Kalpakkam.

[190] Knowledge Management Security for Nuclear Reactor
N. Madurai Meenachi, M. Sai Baba, B.Babu, B.Anandapadmanaban and V.Ramanathan
eINDIA-2009, Hyderabad.

[191] Knowledge Management and Ontology Representation For Fast Breeder Test Reactor
N. Madurai Meenachi, M. Sai Baba, B.Babu, B. Anandapadmanaban and V.Ramanathan
READIT-2009, Kalpakkam.

[192] All Standards and Drafts
W3C Updates
http://www.w3.org/TR/rdf-concepts/. 2004.

[193] The Semantic Web: An Introduction
http://infomesh.net/ 2001/swintro. 2001.

[194] The Object Management Group Ontology Definition Metamodel
R.Colomb, K.Raymond, L.Hart, P.Emery, C.Welty, G.T.Xie, E. Kendall
Ontologies for Software Engineering and Technology, (2006) 217-247.

[195] OWL Web Ontology Language
W3C Working Draft 2003
http://www.ksl.stanford.edu/people/dlm/webont/OWLOverviewJuly302003.htm

[196] TwoUse: Integrating UML Models and : Integrating UML Models and OWL Ontologies
http://www.uni-koblenz.de/~fb4reports/2007/2007_16_Arbeitsber-ichte.pdf.

[197] OWLGrEd: A UML Style Graphical Editor for OWL, Semantic Web OWL: Experiences and Directions
J. Barzdins, G. Barzdins and K. Cerans
7th International Workshop on OWL: Experiences and Directions, USA (2010) 21-22.

[198] Nuclear Reactor Engineering
S. Glasstone, and A.Sesonske
Reactor Systems Engineering, 4th Edition, Springer, 1994.

[199] FBTR System Training Manual, Primary Sodium Main Circuit
M. Sekar
SRS - 06, 2006.

[200] FBTR System Training Manual, Secondary Sodium Purification Circuit,
M. Karunakar
SRS - 14, 2006.

[201] FBTR System Training Manual, Steam Water System
C.S. Varghese, and M. Thangamani
SGP - 01, 2009.

[202] FBTR System Training Manual, Control Rod Drive Mechanism
P. Srinivasa Rao
SRS – 03, 2006.

[203] A Survey on Ontology Mapping
N. Choi, Y. Song, H. Han
ACM SIGMOD, 35 (2006) 34-41.

[204] Ontology Matching: State of the Art and Future Challenges
P. Shvaiko and J. Euzenat
IEEE Transactions on Knowledge and Data Engineering, 25 (2013) 158-176.

[205] A Survey of Schema-Based Matching Approaches
P. Shvaiko and J. Euzenat
Journal on Data Semantics, 3730 (2005) 146-171.

[206] PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment
N.F. Noy and M.A. Musen
International Journal of Human-Computer Studies, 59 (2003) 983–1024.

[207] Data Integration Using Similarity Joins and a Word-Based Information Representation Language
W.W. Cohen
ACM Transactions on Information Systems, 18 (2000) 288–321.

[208] Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa
M.A. Jaro
Journal of the American Statistical Association, 84 (1989) 414–420.

[209] UMLDiff: An Algorithm for Object-Oriented Design Differencing
Z. Xing and E. Stroulia
20th IEEE International Conference on Automated Software Engineering, (2005) 54–65.

[210] A Comparison of String Distance Metrics for Name Matching Tasks
W.W. Cohen, P. Ravikumar and S.E. Fienberg
Workshop on Information Integration on the Web, 2003.

[211] QOM - Quick Ontology Mapping
M. Ehrig and S. Staab
3rd International Semantic Web Conference, 2004.

[212] A String Metric for Ontology Alignment
G. Stoilos, G. Stamou, and S. Kollias
4th International Semantic Web Conference, 2005.

[213] Automatic Evaluation of Machine Translation Quality Using Longest Common Subsequence and Skip-Bigram Statistics
C.Y. Lin and F.J. Och
42nd Annual Meeting on Association for Computational Linguistics, 2004.

[214] Variants of Constrained Longest Common Subsequence
P. Bonizzoni, G.D.Vedova, R.Dondi and Y.Pirola
Information Processing Letters, 110 (2010) 877-881.

[215] Efficient Algorithms for Finding Interleaving Relationship Between Sequences
K.S. Huang, C.B. Yang, K.T. Tseng, H.Y. Ann and Y.H. Peng
Information Processing Letters, 105 (2008) 188-193.

[216] The Constrained Longest Common Subsequence Problem
Y.T. Tsai
Information Processing Letters, 88 (2003) 173-176.

[217] A Survey of Longest common Subsequence Algorithms
L. Bergroth, H. Hakonen and T. Raita, IEEE (2000), 39-48.

[218] Indo-European Languages Tree by Levenshtein Distance
M. Serva and F. Petroni,
Euro Physics Letter, 81 (2008) 1-8.

[219] Using Adapted Levenshtein Distance for On-Line Signature Authentication
S.Schimke, C.Vielhauer and J.Dittmann
17th IEEE International Conf. on Pattern Recognition, 2 (2004) 931 - 934.

[220] Spelling-Error Tolerant, Order Independent Pass Phrases via the Damerau-Levenshtein String-Edit Distance
M.G. Bard
Fifth Australasian Symposium on ACSW Frontiers, 68 (2007) 117-124.

[221] Plagiarism Detection Using the Levenshtein Distance and Smith-Waterman Algorithm
Z. Su, B.R. Ahn, K.Y Eom, M.K. Kang, J.P. Kim and M.K. Kim
3rd IEEE International Conference on Innovative Computing Information, (2008) 569.

[222] Quick Mapping Evaluator –Application for Ontology Mapping Evaluation: A Case Study on Nuclear Reactor Domain
N. Madurai Meenachi, N. Hari Prasad and M. Sai Baba
International Journal of Nuclear Knowledge Management, 6 (2013) 143-154.

[223] Methodology for Semantic Web Based Domain Knowledge Implementation
N.Madurai Meenachi and M. Sai Baba
READIT-2013, Kalpakkam.

[224] An Ontology Based Method for Duplicate Detection in Web Data Tables
P. Buche, J.L. Berthelemy and R. Khefifi and Fatiha Saïs.
Database and Expert Systems Applications, Springer, (2011) 511-525.

[225] Efficient Semantic – Aware Detection of Near Duplicate Resources
E. Iaonnou, O. Papapetrou, D. Skoutas and W. Nejdl
The Semantic Web: Research and Applications, Lecture Notes in Computer Science, 6089 (2010) 136-150.

[226] Steam Generator for Future Fast Breeder Reactors
R. Nandakumar, S. Atmalingam, V. BalaSubramaniyam and S.C. Chetal
Energy Procedia, 7 (2011) 351-358.

[227] Notes on Linear Algebra
http:// www.maths.qmul.ac.uk /~pjc/ notes/ linalg.pdf, 2008.

[228] Matrix Rank Based Ontology Matching: An Extension of String Equality Matching
N. Madurai Meenachi and M. Sai Baba
International Journal of Nuclear Knowledge Management (In Press).

[229] An Ontology for Engineering Mathematics
J. Doyle, P. Torasso and E. Sandewall
4th International Conf. on Principles of Knowledge Representation 1994.

[230] Twenty Five Years of Operating Experience with Fast Breeder Test Reactor
K.V. Suresh Kumar, A. Babu, B. Anandapadmanaban and G. Srinivasan
Asian Nuclear Prospects, 7 (2011) 323-332.

[231] Behavior of Steam Generator Bolted Flange Joints at Elevated Temperatures
V. Vinod, R.K. Mourya, B.K. Sreedhar, I.B. Noushad and K.K. Rajan
Annals of Nuclear Energy, 51 (2012) 307–310.

[232] Reliability Analysis of Control Rod Drive Mechanisms of FBTR for Reactor Startup and Power Control
E. Ramesh and S. Usha
2nd IEEE International Conf. on Reliability, Safety-Hazard, (2010) 431-435.

[233] Ontology Matching with Semantic Verification
Y.R. Jean Mary, E.P. Shironoshita and M.R. Kabuka
Web Semantic, 7 (2009) 235-251.

[234] Coincidence-Based Scoring of Mappings in Ontology Alignment
S.H. Haeri, H. Abolhassani, V. Qazvinian and B.B. Hariri
Journal of Advanced Computational Intelligence and Intelligent Informatics, 11 (2007) 1-14.

[235] Performance Assessment of Multiobjective Optimizers: An Analysis and Review
E. Zitzler, L. Thiele, M. Laumanns, C.M. Fonseca and V. G. De Fonseca
IEEE Transaction on Evolutionary Computation, 7 (2003) 117-132.

[236] Multi-guider and Cross-Searching Approach in Multi-Objective Particle Swarm Optimization for Electromagnetic Problems
M. T. Pham, D. Zhang, and C. SeopKoh
IEEE Transactions on Magnetic, 48 (2012) 539-542.

[237] Distributed Pareto Optimization via Diffusion Strategies
J. Chen and A.H. Sayed
IEEE Journal of Selected Topics in Signal Processing, 7 (2013) 205-220.

[238] Survey of Multi-Objective Optimization Methods for Engineering
R.T.Marler and J.S.Arora
Structural and Multidisciplinary Optimization, 26 (2004) 369-395.

[239] Using Pareto Optimality to Coordinate Distributed Agents
C.J. Petrie, T.A. Webster and M. R. Cutkosky
Special Issue on Conflict Management, 9 (1995) 269-281.

[240] Texture Image Retrieval Based on a Gaussian Mixture Model and Similarity Measure using a Kullback Divergence
H. Yuan and X.P. Zhang
IEEE International Conference on Multimedia and Expo 2004, 3 (2004) 1867-1870.

[241] Cosine Similarity for Intuitionistic fuzzy sets and their applications
J. Ye
Mathematical Computer Modelling, 53 (2011) 91-97.

[242] Scaling up Pop-K Cosine Similarity Search
S.Zhu, J.Wu, H.Xiong and G. Xia
Data and Knowledge Engineering, 70 (2011) 60-83.

[243] Multiobjective Evolutionary Algorithm for Software Project Portfolio Optimization
http://www.c2i.ntu.edu.sg/ivor /publication/PRL-MOEA.pdf

[244] Ontology Alignment using Relative Entropy for Semantic Uncertainty Analysis
E. P. Blasch, E. Dorion, P. Valin and E. Bosse
IEEE National Aerospace and Electronics Conference, (2010) 140-148.

[245] Kullback-Leibler Divergence Estimation of Continuous Distributions
F. Perez-Cruz
IEEE symposium on Information Theory, (2008) 1666 - 1670.