IMPLEMENTATION OF COMPUTING AND DIGITAL MEDIA DICTIONARY BY USING ONTOLOGY WEB LANGUAGE (OWL)

Nu Nan Dar Lin, Thandar Aung
Computer University (Mandalay)
nandarlin2010@gmail.com

Abstract

Ontology is becoming increasingly widespread in the computer science community, and its importance is being recognized in a multiplicity of research fields and application areas. Ontology comprises as the concept, properties, and restrictions on properties. Every object and thing in our environment that has properties can think as concept. This paper intends to take the facilities of ontology in the building of knowledge repository. This paper intends to develop online dictionary for Computer terms and Digital terms in English language. This system is nearly developed as like Wikipedia Web Site. This dictionary includes a number of words about 3000 words. This system allows the participation of people who were interested in the language of English, computer terms and digital terms. Users have opportunity to search the words; system is performed by accepting the query as input and retrieves the relevant definition, synonym, and type of word etc. The user can also see step-by-step link of term in the definition. If the user wants to give suggestion of term or definition, the user must be member of this system and then the user can give suggestion that the user want to type suggestion letter in the suggestion box. This system also takes the Ontology Web Language (OWL) for constructing Computer terms and Digital terms dictionaries (English-English).

1. Introduction

Ontology is an important emerging discipline that has the huge potential to improve information organization, management and understanding. Ontology defines the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary. Ontology is a common understanding of a domain that can be communicated between people and application systems. It has a crucial role to play in providing qualitatively new levels of services on the next generation of Web transformation in the form of Semantic Web. The use of ontology is to describe data model and to give semantics to data stored in Web pages. Also, ontologies have been used in database and information retrieval areas as a support for distributed and heterogeneous data sources interoperability.

There are roughly four kinds of ontologies: document ontology, metadata ontology, domain ontology, and service ontology. This paper is developed as Computing and Digital Media terms domain ontology. This ontology is a specification of the conceptualization and the corresponding vocabulary used to describe domain. This paper is developed by Java language and Jena API and Protégé tool for implementing OWL files.

2. Related Work

Many people use search engine to retrieve relevant information about topic. Searches based on keywords are very closely related to the spelling of the word and not to its meaning thus semantically similar queries can return different results. The main goal of semantic net is that semantic relatedness can be used to determine the meaning of words in text. There is a rich history of research in two distinct areas. First, there has been work that exploits glosses of word meanings as found in Machine Readable Dictionaries. Second, networked or hierarchical arrangements of concept information have been utilized as sources of information for word sense disambiguation.

Dictionaries have long been recognized as possible sources of information for computational methods concerned with word meanings. It proposed that the contents of a dictionary be represented in a semantic network. Each meaning associated with a word is represented by a node, and that node is connected to those words that are used to define the concept in the dictionary.

3. Motivation

There are two kinds of computer-based dictionary, online and offline dictionary although the nature and feature of offline dictionary is commercial, difficult to add new contents and also difficult to discuss on these new contents with different languages from

different place. So, this paper is developed online dictionary by using ontology web language. Ontology can be classified between Heavy-Weight and Light-Weight Ontology. Ontology Web Language (OWL) is existed nearly Heavy-Weight Ontology. So, this paper is nearly developed as Heavy-Weight and ontology.

4. Building Ontology

Building Ontology is more of a craft than an engineering process. Several research groups propose various methods more commonly known as methodologies for building ontologies. There is no consensus between these groups and each employs its own methodology. During the process of building ontology several questions arise. Some of these include:

- What domain will the ontology cover?
- What will the ontology be used for?

What types of questions should the information in the ontology provide answers for (competency questions)?

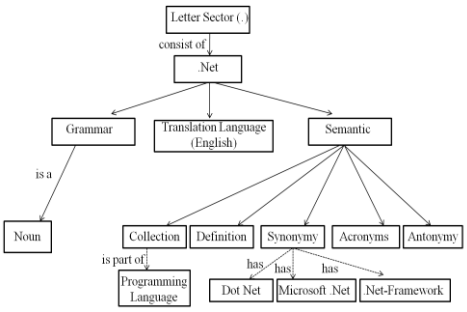

Figure1. Example of Ontology Creations for a Word

The construction of ontologies is necessarily an iterative process, and answers to these questions may change during the life cycle of the ontology [4]. The four basic aspects to consider when creating an ontology are content (the content of the ontology), the language in which it is implemented, the methodology which has been followed to develop it, and the software tools used to build and edit the ontology. Furthermore the type of language and development environment chosen is dependent on the type of ontology being built. It is important to know what the ontology is going to be used for, and how detailed the ontology should be, so that the possibility of over modeling (e.g. attempting to model the "whole world") is lessened [2].

4.1. Ontology Web Language (OWL)

Ontologies are used to capture knowledge about some domain of interest. Ontology describes the concepts in the domain and also the relationships that hold between those concepts. Different ontology languages provide different facilities. The most recent development in standard ontology languages is OWL from the World Wide Web Consortium (W3C) [2]. OWL which is short for Web Ontology Language is an ontology language designed to be compatible with the World Wide Web and the Semantic Web. The most important abstraction in OWL is concept axioms which are called classes. Each class has a list of necessary conditions and zero or more equivalent lists of necessary and suffcient conditions. A list of necessary conditions is a list of conditions that every member of the class must satisfy [3]. In the same way a list of necessary and suffcient conditions is a list of conditions that must be satisfied by every member of the class and if satisfied guarantees membership in the class [7]. OWL is based on XML, RDF and RDF-S and can be used to represent information in a way that is more accessible to applications than traditional web pages. In addition, OWL has a formal semantics, which enables logic reasoning. OWL comes in three variants: OWL-Lite ⊆ OWL-DL ⊆ OWL-Full of increasing expressive power. The variants OWL-Lite and OWL-DL are based on the description logics SHIF (D) and SHOIN (D) respectively, which guarantees that important inference problems such as satisfiability and subsumption are decidable. Since OWL is XML based, it needs an editor to create OWL ontologies. In order to regain computational efficiency, OWL DL (short for: Description Logic) is a sublanguage of OWL Full which restricts the way in which the constructors from OWL and RDF can be used [1]. This amounts to disallowing application of OWL's constructor's to each other, and thus ensuring that the language corresponds to well-studied description logic. The advantage of this is that it permits efficient reasoning support [7].

4.2. Components of OWL Ontologies

In this system used three components of OWL ontology.

4.2.1. Individuals

Individuals represent objects in the domain in which are interested. An important difference between Protégé and OWL is that OWL does not use the Unique Name Assumption (UNA). This means that two different names could actually refer to the same individual [5]. For example, "Window XP", "Window" and "Elizabeth Windsor" might all refer to the same individual. In OWL, it must be explicitly

stated that individuals are the same as each other, or different to each other otherwise they might be the same as each other, or they might be different to each other. Figure.2 shows a representation of some individuals in some domain in which represent individuals as diamonds in diagrams.

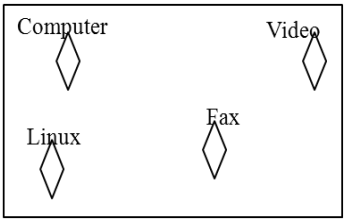

Figure2. Representation of Individual

4.2.2. Properties

Properties are binary relations on individuals i.e. properties link two individuals together. For example, the property has Sibling might link the individual Computer to the individual CPU. Properties can have inverses. Properties can be limited to having a single value i.e. to being functional. They can also be either transitive or symmetric [5]. Figure.3 shows a representation of some properties linking some individuals together.

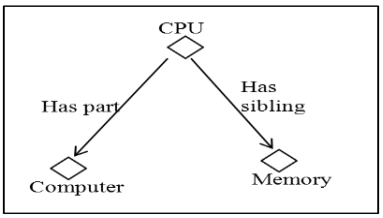

Figure3. Representation of Properties

4.2.3. Classes

OWL classes are interpreted as sets that contain individuals. OWL classes are described using formal (mathematical) descriptions that state precisely the requirements for membership of the class. One of the key features of OWL-DL is that these superclass-subclass relationships (subsumption relationships) can be computed automatically by a reasoner more on this later. Figure.4 shows a representation of some classes containing individuals classes are represented as circles or ovals, rather like sets in Venn diagrams. In OWL classes are built up of descriptions that specify the conditions that must be satisfied by an individual for it to be a member of the class [5].

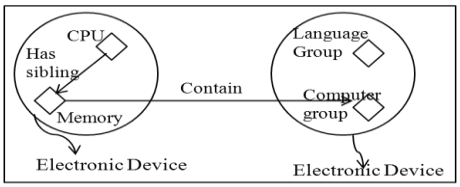

Figure4. Representation of Class(Containing Individual)

5. Ontology Expressive Spectrum

In order to classify ontologies according to their expressiveness, can be distinguished several expressiveness levels based on the ontology spectrum.

- Controlled vocabulary a list of terms;
- Thesaurus relations between terms, such as synonyms are provided;
- Informal taxonomy there is an explicit hierarchy(generalization and specializations are supported), but there is no strict inheritance; an instance of a subclass is not necessarily also an instance of the super class;
- Formal taxonomy there is strict inheritance;
- Frames a frame(class) contains a number of properties and these properties are inherited by subclass and instances;
- Value restrictions values of properties are restricted (for example by a data type);
- General logic constraints values may be constraint by logical or mathematical formulas using value from other properties;
- First-order logic constraints very expressive ontology languages such as in Ontolingua or CycL allow first order logic constraints between and more detailed relationship such as disjoint classes, disjoint coverings, inverse relationships, part-whole relationships, etc [6].

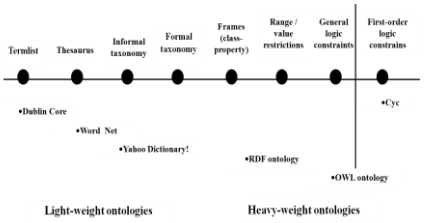

Figure5. Light-Weight and Heavy-Weight Ontology

6. Implementation

This paper is an implementation of Computer and Digital Terms dictionary using ontology web language. It uses Jena package as well as a repository of information about words that can be exploited to

distinguish their meanings in text. In this system, any user can search any terms of Computer and Digital dictionary and their semantic relation. There are three user levels. They are Simple users, Member user and Administrator user. Simple user only can search computer and digital terms. Member user can search computer and digital terms and can give suggestion for search terms. Administrator user level can search computer and digital terms and edit term's definitions and add new terms by according member user's suggestion. This admin user can build Owl classes and Owl properties into ontology database.

Input Keyword	Topics of Search Results	Descriptive Terms of Search Results	Synonyms Relationship of Search Result	Synonyms Relationship of Search Result (2)	Synonyms Relationship of Search Result by hyperlink etc.....
Hardware	**Hardware interface** [Noun] A connection point between a computer and external devices. Via hardware interfaces several hardware components like printer, mouse etc. can be connected to the computer. **Hardware** [Noun] A collective term for all material, concrete ("hard") parts and appliances in a computer. Examples for hardware are the storage media, graphics cards, output devices (screen, printer), input devices (mouse, keyboard) or drives. **Hard Disk** [Noun] Creating extra copies of files in case the original copies Creating extra copies of files in case the original copies are damaged or destroyed.	**Hardware** [Noun] **Definition:** A collective term for all material, concrete ("hard") parts and appliances in a computer. Examples for hardware are the storage media, graphics cards, output devices (screen, printer), input devices (mouse, keyboard) or drives. **Acronym:** HW **Synonyms:** Accelerator, Graphics Accelerator, Architecture, Bus, CPU **Antonyms:**	**Central Processing Unit** [Noun] **Definition:** The structure and organization of a computer's hardware, or system software, including items such as the CPU type, bus type, etc. **Acronym:** CPU **Synonyms:** Hardware, Background, Architecture, Bus, adware, Operating_System, Disk_Operating_System, System **Antonyms:**	**Operating System** [Noun] **Definition:** An application which is currently running and is usually the current window in view on the display. In other words, the active application is the one you are currently able to work in, entering text, creating pictures, browsing the Internet, etc. **Acronym:** OS, DOS **Synonyms:** Background, Window, Active_Window, Program, CPU, Software, Disk_Operating_System, System	----------

Figure6. Example of Input keyword and Search Result

6.1. Overview of the System

This system is to provide semantic interpretation of the terms for construction of domain ontology. Computing and digital terms are disambiguated to get the more relevant relations. Terms are semantically interpreted. This system consists of two parts. First part, the user constructs computing and digital terms Dictionary using OWL language and creates for viewing this dictionary.

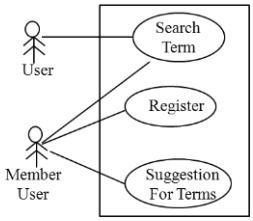

Figure7. Use Case Diagrams for the Simple User and Member User

The ontology updates facility because if the user wants to add new words, the user can do it by leaving suggestions in the suggestion box. There are three main actors in this system, (i) dictionary viewer (ii) dictionary Administrator (iii) member user. The role of each user is represented with two use case dictionary as shown in Figure.7 and Figure.8. Member user and viewer can view and search the computing and digital media terms into the dictionary software package. Member user needs to process login function. New member user also needs to process registration. Member users can suggestions of terms and definition in the suggestion box.

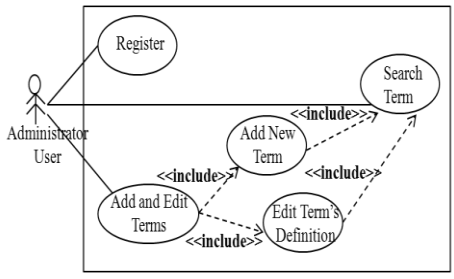

Figure8. Use Case Diagrams for the Administrator User

Administrator needs to process login function. After the login process finishing, admin can modify or add terms and then construct ontology class and

properties for the terms. Admin can also check the member users, left suggestions for modifying or adding terms and definition. If the terms already exist in the system, admin can ignore it. If not exist, admin can construct ontology for these terms.

6.2 System Flow Diagram

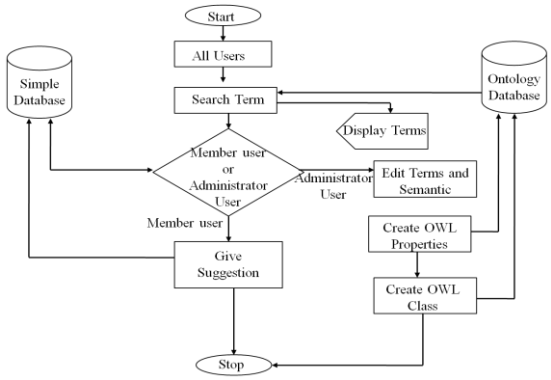

Figure9. System Flow Diagram of the System

Figure.9 shows the system flow, this system using two databases. First database, to store user data and user gave suggestions such as user name and password. In this system can search term for any user who interesting in the Computer and Digital terms. If the user wants to give suggestion of term or definition, the user must be member of this system and then the user can give suggestion that the user want to type suggestion letter in the suggestion box. These suggestions are stored in simple database. Next database is Ontology database. This database is stored OWL properties and OWL class. Only administrator user can enter both database. Administrator user can perform search terms, check member user's suggestions and edit and add new terms of definitions by creating OWL properties and OWL class into Ontology database. This system is performed as Online Dictionary.

6.3 User Interface

This paper is an implementation of computing and digital media terms dictionary using ontology web language. It uses Protégé 3.4. rc1 as well as a repository of information about words that can be exploited to distinguish their meanings in text. In this system, user can search any terms of English to English computer terms and digital terms dictionary and their semantic relation. Administrator can insert, update and delete terms, and can build Owl classes and Owl properties.

The user can search any terms of English dictionary computing and digital terms and the semantic interpretation, definition, synonyms, antonyms and acronyms of the selected features. In this step, extracted features are connected with ontology database to determine their definition and relation meaning. The user can type words that the user wants to search in the box. And then, the system shows terms and definitions of close words in the box. Users can choice, users want to watch definition of terms and simultaneously the system shows synonym, acronym and antonym of user's search term. Users can also see definition of synonym by linking. Semantic nets show generalization (isA) relation of words or phrases when the user selected.

Figure10. User's Search Result Interface

7. OWL Syntax for Class Term

In this paper, create syntax for Computer and Digital terms of OWL as follow:

```
<?xml version="1.0"?>
<owl:Ontology rdf:about=""/><owl:Class rdf:ID="Webhosting">
<rdfs:subClassOf><owl:Class rdf:about="#Online-Service"/>
</rdfs:subClassOf>
    </owl:Class>
   <owl:Class rdf:ID="Compiler-935">
     <rdfs:subClassOf>
      <owl:Class rdf:about="#Software"/>
     </rdfs:subClassOf>
    </owl:Class>
   <owl:Class rdf:ID="VOB">
     <rdfs:subClassOf>
      <owl:Class rdf:about="#File-Type/>
     </rdfs:subClassOf>
    </owl:Class>
   <owl:Class rdf:ID="Adobe Reader">
     <rdfs:subClassOf>
      <owl:Class rdf:about="# Software "/>
     </rdfs:subClassOf>
    </owl:Class>
   <owl:Class rdf:ID="Microsoft Office">
     <rdfs:subClassOf>
      <owl:Class rdf:about="# Software "/>
     </rdfs:subClassOf>
    </owl:Class>
```

8. Conclusion

The main focus of this system is to study

implementation of ontology. Since, the meaning of a word can be different based on parts of speech and the usage. This paper takes the facilities of ontology in the building of knowledge repository. This paper is developed as online dictionary. So, this paper is easy to add new contents and also easy to discuss on these new contents with different languages from different place.

9. Limitations and Further Extension of System

The users query just the words these types are Noun, Verb, Adjective from this computer and digital terms ontology dictionary. This paper is online dictionary for only computer and digital terms dictionary (English-English). And the developers who want to modify this paper, developers have to understand and be familiar with the ontology concept and OWL language. To become any English terms of Dictionary for example Yahoo!, other specific domain terms dictionary needs to be effectively developed. There will be developed by using ontology database and ontology web language (OWL).

References

[1] G. Antoniou1 and F. V. Harmelen, "Web Ontology Language: OWL"

[2] L. Casely-Hayford, S. Sufi "The ISIS Facilities Ontology and OntoMaintainer "

[3] J. De Bruij, "Using Ontologies - Enabling Knowledge Sharing and Reuse on the Semantic Web"

[4] N. Guarino "Formal Ontology and Information Systems"

[5] M.Horridge "A Practical Guide To Building OWL Ontologies Using Prot´eg´e 4 and CO-ODE Tools Edition 1.2"

[6] K. M. Htet, "Capturing And Building Knowledge Repository: Ontology Approach".

[7] http://www.w3.org/ TR/ 2004/ REC- owl- features 20040210/ "OWL Web Ontology Language"