European Journal of Academic Essays 1(3): 84-88, 2014
ISSN: 2183-1904
www.euroessays.org

A Survey: Concepts of Ontology

< Tiwari Sujeet Kumar >[1], < Tiwari Neelu >[2] <Kumar Raghvendra>[3] <Upadhyay Ashish Kumar>[4]

[1]Dept of computer science and engineering, LNCT group of college, Jabalpur, India
sujeet.tiwari08@gmail.com
[2]Dept of computer science and engineering, Hitkarini College of Engineering & Technology, Jabalpur, India
neel_31@radiffmail.com
[3]Dept of computer science and engineering, LNCT group of college, Jabalpur, India
Raghvendraagrawal7@gmail.com
[4]Dept of computer science and engineering, LNCT group of college, Jabalpur, India
ashishup6@gmail.com

Abstract: This article proposes an approach of construction an ontological intentions based on the work study suggested in the literature. Many methods were proposed in the literature. We are interested to the methods applying to the texts, and more particularly, with the methods to learn ontology starting from textual corpus. This paper presents a construction of an intentions ontology. We state some fundamental principles to respect to build an ontology. We will build the ontology of specific domain which comprises an abstraction levels to knowing a linguistics level, and understanding the effective concepts of the domain and the relations which links them, will be useful for the knowledge extraction between the various semantic contexts, knowing a semantic context is define as being a subset of concepts and relations which link them. The Ontology of Verb Concepts combines linguistic and psycholinguistic classification. The information for verbs encodes typical associations with actions and events. We propose a method by using domain ontology and syntactic-semantic analysis. This article presents experimentation on scientific publications article in the domain of computer science.

Keywords: Information research, Term Extractions, Intention Ontology, Syntactic-Semantic Analysis, Knowledge Modeling.

1. Introduction

Today, the Web is an essential tool tries to obtain quickly the relevant information. The Web represents an immense source of data to get information, search, learn and discover new knowledge. The Web allows to better answer to our needs always more growing of information and knowledge. But while bringing a solution, it brings by the fact even a new problem. How we can fmd the relevant information among million documents? To facilitate the information retrieval, the users often call already available techniques with the conventional search engines such as Google, AltaVista and others. But these solutions only based on the research by key words allow solving only part of the enormous complexity of the information search problem. The access, the extraction, interpretation and the maintenance of information are left to the user [I]. This can pose many problems because it is very frequent that a research does not us the documents which we needed. Moreover, even if the turned documents contain what was searched, the user must later search the text to find the relevant information which he needs. The information extraction methods aim at improving the research experiment of the user while trying to obtain a semantic representation of the sense of executed requests with a search engine. With these search engines we are able to find document which not contain the key words requested by the user. The information extraction consists in extracting knowledge from various documents by using linguistic techniques. The development of such system is a long task which often requires a domain expert which they work on linguistics knowledge. The proposed method uses the formalism of the conceptual graphs to obtain a semantic representation of scientific publications documents a specific domain by using syntactic and semantic analysis coupled with the use of an ontology built starting from the terms extracted from textual corpus. Ontology aims to answering to this problem. Although ontology is an essential element if it is wanted that one day the semantic Web becomes a reality. We tried to automate the maintenance process of ontology by using automatic learning techniques and statistical methods. We tried to extract the link between the concepts contained in textual corpus. This enables us to have an ontology created according to the domain which we tried to treat. The rest of this paper is organized as follows. The second section, we describe the notion of intentions and several systems based on this notion for extracting the relevant information from textual corpus. In the third section, we describe the creation of an intention ontologism used to recognize the authors intentions in a specific domain based on the analysis of the semantic relations between terms to extract new relations between the concepts as well as new concepts order to lead to a semantic network of concepts, this section also we

Corresponding Author: Kumar Raghvendra, Email: Raghvendraagrawal7@gmail.com

present an example of this ontology built manually and how the verbs make the link between the concepts. The fourth section we finish by conclusion and we give ideas for future research.

1. The Intention Concept

The intention corresponds to a mental state of any actor who executes an action. We are interested exclusively in the intentional actions. Several works attempt to account the relations between an action undertaken by a human being and the mental state which guides this action. Searle remains a main reference on the matter [2]. He distinguishes between two types of intention: intentions in the course of action and the 'preformulated'intentions --Pre-formulated intentions represent a condition of satisfaction of the intention. Intentions in the course of action are those which represent these intentions. Several systems were developed by our research team based on the concept of intention. System SABRE [3] and XSEdit system [4], related to the information research in a textual corpus of documents, and system Pero [5]-[6], for the learning by observation through the reasoning of intentions. The concept of the intentionality indicates two minimal reasons for an intentional action: that the agent has a desire for results and that the agent has a belief which the action envisaged leads to these results. To be able to exploit the notion of the intention, it is necessary to formalize it according to a formal language. By definition we represent an intention as: Where: I 0*, M*, R*)

Where: '1' represents the intention carried out by an action A; A is the action that expresses what the author of the intention wants to do; G represents the goal to achieve by performing the action; M represents the means to express how the action is accomplished; R represents the reason to express why the author chooses this action and for which reasons, *indicate that the number of the acts composing the intention can be or N [7]-[8]. Four types of questions are posed on this level: What are we doing? A Model, represent and process intentions of the writing actions and derma new structures of document: Intentional Structure (reference).

Why? G Enable the access to documents in terms of author's intentions. How M Specifying a conceptual model of intentions Coding intentions as Xml metadata inside documents to recognize these intentions (semi automatically). For what reason? R Understand human cognitive processes Contribute to the progress of this new area have more funded projects, Have more interested students, Etc. our RICAD System which based on the model of intention we created an intentions ontology to extract the intentional authors' information from scientific publications articles [9]. It is based on algorithms which facilitates the intentions research and its principle of operation closer to domain expert. The RICAD recognizes the sentences and the verbs using an ontology of verbs (Figure 1). the following section we will present the creation of this intentions ontology in a specific domain by using a semantic and syntactic method to constructed, an example will be presented explain how this ontology is built and how the verbs link. Between two concepts.

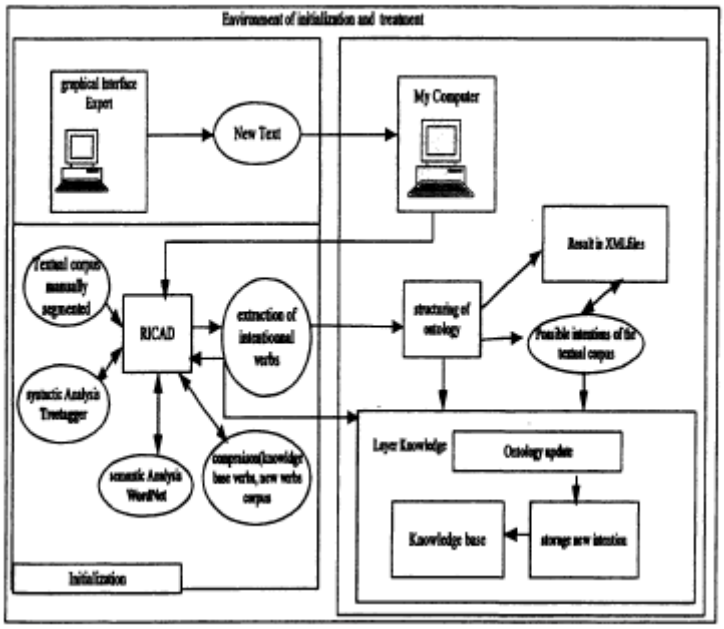

Figure1:Represent the direction of data flow

III. CREATION OF AN INTENTIONS ONTOLOGY

Many methods aim at building ontologies starting from corpus. The majorities are based on the contents of the texts to build ontology; the texts are then the principal source of knowledge for information acquisition [10]. The set of concepts of the model as well as their relations exclusively come from an analysis of texts, without external contribution of knowledge. Aussenac-Gilles register in this step if it is recognized that the texts can not constitute the only source of knowledge [11].

This endogenous approach decomposes into several stages: the extraction of terms referring to basic concepts (primitive conceptual [12]), then lexical relations which they maintain (creation of a terminological base [12]- [13]) in order to make emerge the fIrSt relations inter-concepts. The following stage is based on the analysis of the semantic relations between terms to extract new relations between the concepts as well as new concepts in order to lead to a semantic network of concepts. The semantic .network must be validated by an expert of domain order to specify the significant relations (normalization [14]).

In each term a concept of initial ontology does not correspond. It is necessary to be able to connect the terms of the corpus to the concepts defined this ontology. A process of normalization is necessary. It is held two stages: the first one extension of initial ontology, and the second one definition of a terminological base. -Initial ontology is revised with the assistance of domain experts. New concepts supplement hierarchies.

-According to the same experts, and by pressing us on scientific documents specific to the domain, we specify starting from the terminology another sets of concepts, the basic concepts. Recursively of new concepts are defined by heritage starting from the basic concepts. The result is a sets of small hierarchies having each one a single ancestor of which the last descendants are basic concepts. These hierarchies are normalized, i.e. organized a systematic way, because each father decomposes into sons according to a single criterion. They also respect the criterion of rigidity of Guarino [15].

The two preceding processes give on the one hand an ontology related to the treated problem and set of sub hierarchies directly related to the terminology of the text.

We obtain a modelling of the domain covering all the concepts which interest us for the information search. This modeling is described by a relational diagram formalized by sets of directed graphs. It establishes a described ontology in a semi-formal way because independent of a language of representation [16].

In order to produce the ontological representations by using the classes and relations of an ontology or several ontologies a preparative step is necessary which links the entries of our semantic thesaurus with the ontological elements. This mapping should be automated the most possible.

We exploited the names of the classes and ontological relations (subclasses, sub-relations etc). For example the main class is the concept ontology and the sub class for this concept is ontology, Knowledge Representation, RDF and etc. we obtain a list of sub classes for each ontological concept by using semantic thesaurus.

We will build the ontology of specific domain which comprises an abstraction levels to knowing a linguistics level, and understanding the effective concepts of the domain and the relations which links them, will be useful for the knowledge extraction between the various semantic contexts, knowing a semantic context is define as being a subset of concepts and relations which link them. We will start with the construction of an ontology which will be extended according to a process of incremental enrichment. When with the extraction of the concepts, the relations and the axioms, the idea is to combine the linguistic techniques and the techniques of training.

Initially, we plan to use the Text-To-Onto tool which adopts a multistrategic approach which places at our disposal a bookshop of algorithms of extraction of the concepts and relations, to enrich our ontology. Then, in the second time, we will develop our own algorithms of extraction of concepts (techniques of clustering), to build a complete ontology of domain.

The use of ontologies in our case will facilitate the recognition of the intentions. Ontologies will help the authors of the documents to specify their intentions, and they intervene in the intentions recognition of document [17]. The knowledge representation of an intentional ontology that we use is as follows. existing thesaurus contains the intentions of a precise domain; these intentions were put by an expert of the domain during the method of our system analyzer. The conception of ontology is generated in parallel with the execution of different the stages from the analyzer, After having introduced the corpus of a document, the extraction of the concepts and verbs of each segment, with the assistance of thesaurus intentions, the analyzer identifies the intentional verbs. The ontological representation should further go that a simple list from key words, but also represent the relations between classes (ontological). For the construction of our ontology, we frrst determine the domain of ontology the filed of computer science by using the scientific publication, and secondly we define the most significant concept of ontology, fmally we used the top down methods to define the most general concept, and the sub classes. We chose the top down methods 'which is simplest and more methodical.

We used the predicate isa(X, Y) like formalism for the representation of hyponymy relation. For the relation of synonymy, we allocated all lexicalization in every concept. We considered that the lexicalization of every concept corresponds to these synonyms (quasi synonymous) since it regroups all terms which allow to express it. We also defined the concept of heritage properties between concepts. The properties defined at the level of each concept father are inherited by his descendants, but they can be redefined on a lower level.

A. Syntactic and semantic analysis

We have several ways of generating ontological representations starting from the texts in natural language. A simple way would be the extraction of the verbs as key words to make the relation between the concepts as you will see the following sections. It enables us to link only the semantic data with the classes and ontological relations. First we perform a syntactic analysis by using Treetagger, for extracting the verbs from the corpus.

1. (Treetagger) allows the system to make a syntactic analysis on each logic element of the document. Our analyzer recognizes the sentences and the verbs by using ontology of verbs. Starting from the syntactic trees, we create semantic representations. These semantic representations or graphs are built starting from the predicates (thesaurus). Second we used Semantic Analysis (Wordnet) after generation of the textual files which contain lists of the verbs, the semantic analysis uses Wordnet to fmd the synonyms of verbs in order to avoid the redundancy of the verbs, and at the same time to fmd the other synonyms of these verbs.
2. WORDNET is a lexical data base. WordNet [18] is a semantic lexicon of the English language. The terms are organized there the forms of sets of synonyms, the synsets. Each synset is a lexiconized concept (by taking again the terms of Mr. Slodzian [19]). These lexiconized concepts are connected by conceptual relations (is-a has-a. The originators of WORDNET thus affrrm to build a linguistic ontology. If we think that the purpose of ontology is to conceptualize a domain and we maintain it, then the process which leads to his ontology must clearly take the direction of it. It is not the case ofWORDNET which juggles briskly between conceptual relations and terms. Let us specify well the motivation of these note: WORDNET is an. enormous dictionary of English American (more than 100 000 synsets and his richness and its accessibility make of them an interesting tool for information search or other tasks as the treatment of the natural language but it is not an ontology, that does not take the way of it and to to use it just as it is or with a minimum of modification a formal system is dedicated to the failure. Once arrived at such a semantic representation, which ignores information judged not much important, we can build an ontological representation which uses only the ontological vocabulary (classes and relations). This transformation is done according to rules created during alignment between the semantic data and ontologies [20].

B. The Ontology ofVerb Concepts

this section we present techniques of constitution of our classes and their justifications, we were based on the classification of the verbs of WordNet given by [21]. Initially, we present a defmition of the thematic schema as well as the main thematic roles. Then, we will present for each class some examples of verbs belonging to that followed by a general thematic schema. According to our definition of intention an intention is composed of acts considered as action goal means and reason. We consider that each act of an intention represented by the most significant verb of a segment. Each segment is dependent with a single intention. The name of an action is obtained from that the verb from which it rises; we could therefore recover the list of used actions by analyzing the chain verbe1 till verb N. However, this analysis should be carried out with each analysis of a new corpus; it is therefore preferable to memorize the link between an action and a verb from which it rises. We are interested in this work in the intentional verbs, Le. the verbs which translate intentional actions (Verb => Action In our case the verbs describe the relation between the concepts (entities). For example if we take the verb identify and these two concepts or entities Ontologies and Relations· this verb make the links between these two concept (Ontologies define relations). The main class is Identify and the sub classes of this verb are Define, Describe, Recognize, etc. the next section we will show you an example of our ontologies of verbs and their relations between the concepts.

C. *An example ofontology*

this section, we present an example of ontology built from textual corpus in scientific document (Figure 2), which show us how the verbs link between two entities (concepts) in our case. We chose the scientific documents to limit the infinity of intentions. The domain we considered for this study is publications articles in computer science. One of the reasons why we chose to work with scientific articles is the practical value of better document retrieval environments for scientists. Scientific papers are different from other text types with respect to their overall structure, an aspect we are particularly interested This kind of documents is justified by the fact that Intentions scientific documents are deterministic.

We took the first sentences of this abstract above (Biomedical ontologies define entities and relations in order to represent knowledge in the bioDlcdical dOlnain.), and I represented it by ontology of verbs and concepts. Our case we take into account a scientific article in specific domain (ontologies field) computer science. The Figure 3 below is composed of three main concepts and the main class verbs which contain sub classes, each concept contains instances if we take for example the Concept ontology; the instance or the sub classes of this concept is ontology, Knowledge Representation, RDF, etc. The two others concept is Class and Link. The Concept Verb this ontology is to make a relation between other concepts, as you will see in the figure 3, the relation between concept ontology (ontology) and concept class (Entity) is the verb Define which is the instance of verb Identify (Synonymous); the result is the ontology defines Entity. And the same analysis for the others sentences. The updating of our ontology in the sub specific domain is made we find new concepts and the synonym of verbs which belongs to this domain.

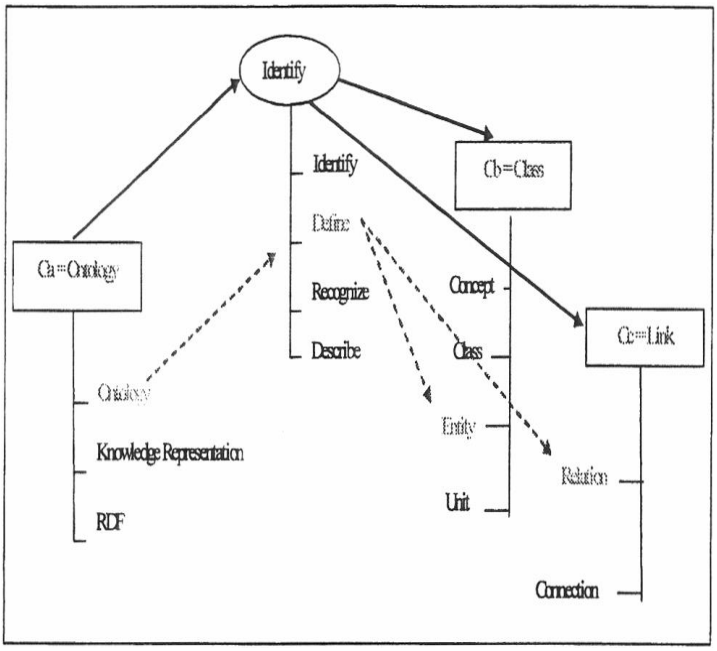

Figure 2: Ontology of verbs concept

The perspective of our work is to consider richer formal languages to represent intentions ontologies and to experiment the developed prototype in other domains.

CONCLUSION

Many methods aim at building ontologies starting from corpus. The majority are based on the contents of the texts to build ontology. This article presents methodologies of construction an intentions ontologies. this paper we make an analysis of the semantic relations between terms to extract new relations between the concepts as well as new concepts in order to lead to a semantic network of concepts.

The semantic network must validated by an expert of domain in order to specify the significant relations. First we perform a syntactic analysis by using Treetagger, for extracting the verbs from the corpus. (Treetagger) allows the system to make a syntactic analysis on each logic element of the document. Our analyzer recognizes the sentences and the verbs by using ontology of verbs. Starting from the syntactic trees, we create semantic representations. We used also wordnet to find the list of synonymous for the terms. We build an ontological representation which uses only the ontological vocabulary (classes and relations).

We chose the scientific documents to limit the infinity of intentions. The domain we considered for this study is publications articles in computer science. We present an example of ontology built from textual corpus in scientific document, which show us how the verbs link between two entities (concepts). We based on the concept of intentional structure to establish a semi-automatic system of segmentation according to the authors' intentions. The updating of our ontology in the sub specific domain is made we find new concepts and the synonym of verbs which belongs to this domain. Within this framework, we presented architecture of ontological components. Initially, we will work to automate the ontology construction ofthe domain.

References

[I] A. Maedche. Ontology Learning for the Semantic Web. Kluwer Academic Publishers, Norwell, MA, USA; 2002.

[2] J. Searle, Intentionality", Cambridge: Cambridge University Press, 1983.
[3] Y. AI-Tawki, "creation par reutilisation de documents d6crits par lesintentions de l'auteur", Doctorat de I'Universitl! de Toulouse I, 2002.
[4] S. TAZI, K. DRIRA and ESSAJIDI, "Enabling consistency of communicative intentions in cooperative authoring", 2nd International Conference on Innovative Views of NET Technologies (lVNET'2006), Florianopolis (Bresil), 21 October 2006.
[5] A. EI Hore and S. Tazi, "Pero a planning system for the explanation of problem solving in physics", Mixed Language Explanation in Learning Environment (MLELE'05) in conjunction with AIED'05, Amsterdam, p. 80-81, 18-22 juillet 2005.
[6] A. EI Hore and S. Tazi, "Planning and explaining solutions for physics learning", international conference on Machine Intelligence (http://www.acidcaicmi2oo5.orgl) The 2nd ACIDCA-ICMI'2oo5. Tozeur- Tunisia, November 5-7, 2005.
[7] H. Kanso, C. Soule-Dupuy and S. Tazi, "Reconnaissance intentions de communication Ccrite dans des corpus de documents scientifiques", geme edition de la conference H2P'IM : Hypertextes, Hypermed.ias 29,30 et 31 Octobre 2007 - Tunisie.
[8] H. Kanso, C. Soule-Dupuy and S. Tazi, "Representing author's intentions of scientific documents", Funchal, Madeire (portugal), 12-16 June 2007.
[9] 9 H. Kanso, A. Elhore, C. Soule-Dupuy, S. Tazi. "RICAD:Towards an Architecture for Recognizing Author's Targets". 10th International Conference on Enterprise Information Systems (ICEIS): Barcelona,' Spain, 12 - 16 June 2008.
[10] 1. NobCcourt. A method to build formal ontologies from text. In: EKAW-2000 Worlcshop on ontologies and text, Juan-Les-Pins, France 2000.
[II] N. Aussenac-Gilles, B. Biebow and S. Szulman. Corpus Analysis For Conceptual Modelling.Workshop on Ontologies and Text, Knowledge Engineering and Knowledge Management, 12th International Conference EKAW, Juan-Ies-pins, France 2000a.
[12] N. Aussenac-Gilles, B. Biebow and S. Szulman. Revisiting Ontology Design: A MethodologyBased on Corpus Analysis. In 12th International Conference Knowledge Engineering and Knowledge Management (EKAW). Juan-Les-Pins, France 2ooob.
[13] G. Lame, Knowledge acquisition from texts towards an ontology of French law, Proceedings ofEKAW'2OOO, 53-62.
[14] 12 J. Bouaud, B. Bachimont, J. Charlet, P. Zweigenbaum. Methodological Principles for Structuring an Ontology, Proceedings of UCAI-I995.
[15] N. Guarino, W. Welty. A Formal Ontology ofProperties, Proceedings of the ICAI-OO Workshop on Applications of Ontologies and ProblemSolving Methods, 12/1-12/8, 2000.
[16] 5 C. Barry, C. Cormier, G. Kassel, J. Nobecourt. Evaluation de languages operationneIs de representation d'ontologies, Actes d'IC'2ool, 309-327.
[17] James FIELDING, Dirk MARWEDE. The Image as Spatial Region: Location and Adjacency within the Radiological Image. Formal Ontology Information Systems B. Bennett and C. Fellbaum (Eds.) lOS Press, 2006.