© 2018 IJSRSET | Volume 4 | Issue 4 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Development of Context Ontology Using Context Assessment

Ram Kumar[1], Shailesh Jaloree[1], R S Thakur[1], Manit
[1] Applied Mathematics & Computer Application, SATI Vidisha, Madhya Pradesh, India
[2]Bhopal, Madhya Pradesh, India

ABSTRACT

In this paper, Information Extraction addresses the intelligent access to document contents by automatically extracting information applicable to a given task. This paper focuses on how ontologies can be exploited to interpret the contextual document content for IE purposes. It makes use of IE systems from the point of view of IE as a knowledge-based NLP process. It reviews the dissimilar steps of NLP necessary for IE tasks: Rule-Based & Dependency Based Information Extraction, Context Assessment.
Keywords : NLP, Rule-Based & Dependency Based Information Extraction, Context Assessment

I. INTRODUCTION

Ontology represents the complete semantics of content information must be associated with context information. Moreover the index considering only semantics ignores the context information regarding that data. Unfortunately a single data object which is separated from its context has less capability of conveying semantics. This includes contextual information regarding the data, such as date, time, and place of date taken. In the sports data context details, abstracts complete information of that context, i.e., the match, team, tournament, stadium, umpire and player personal details, etc .

In this work we construct the context ontology, to support Context based information extraction. The context information associated to the data can be classified in to context independent information and context dependent information. Context information is the information associated to the data. In cricket domain the information associated to the data can be of two types, One, specific information associated to the occasion, event, game, concepts for which the data has been taken, known as Context dependent information. Second, the general information associated to the occasion, event, game, concepts, ie the general information not pertaining to a specific game is known as Context independent information. Cricket concepts are players, stadium, umpire, match, sponsors, score, records, commentators etc. Cricket domain ontology contains these concepts, relations between concepts, concepts attributes. The concept attributes share extraction system for cricket which meets the end user requirements domain knowledge is essential.

II. RELATED WORK

In Conceptual clustering, the concepts are grouped according to the semantic distance between each other to make up hierarchies. But because of lack the domain context to instruct in the process of distance computation, the conceptual clustering process can't be efficiently controlled. Furthermore, by this method, only taxonomic relations of the concepts in the ontology can be generated [1].

Concept learning, a given taxonomy is incrementally updated as new concepts are acquired from real-world texts. Concept learning is a part of the process of ontology learning [2].

Text2Onto creates ontology from annotated texts. This system incorporates probabilistic ontology models (POMs). It shows a user different model ranked according to the certainty ranking and does linguistic preprocessing of the data. It also finds properties that distinguish a class from another.Text2Onto uses an annotated corpus for term generation [3].

Association rules, the association rules have been used to discover non-taxonomic relations between concepts. Association rules are most used on the data mining process to discover information stored on database. Ontology learning mostly uses unstructured texts but not the structure data in database. So, association rule is just an assistant method to help the ontology generation [4].

Wen Zhou have proposed a semi-automatic technique that starts form small core ontology constructed by domain experts and learns the concepts and relations by use of the general ontology .In his paper ,WordNet and event based NLP technologies are used that automatically to construct the domain ontology [5].

H. Kong gave the methodology for building the ontology automatically based on the frame ontology from the WordNet concepts and existing knowledge data. The ontology building method is divided into two parts. One part is to make the possibility for building the ontology automatically based on the frame ontology from the WordNet concepts that are the standard structured knowledge data [6].

Iqbal proposed a semi automated algorithm to transform data to the ontology language, OWL. They described Ontology, as the vocabulary and core component of the Semantic Web, provides a re-usable representation of real-world things in a particular domain or application area [7].

In Sowa's Top Level Ontology, the ontology has a lattice structure where the top level concept is of universal and low level concepts are absurd type [8].

OntoEdit provides an environment for the development of ontologies .It provides a user interface. The concept hierarchy can be edited or created .The decision of making direct instances of a concept depends upon the type of the concept [9].

OntoLT allows a user to define mapping rules which provides a precondition language that annotates the corpus. Preconditions are implemented using XPATH expressions and consist of terms and functions. According to the preconditions that are satisfied, candidate classes and properties are generated. Again, OntoLT uses pre-defined rules to find these relationships[10].

J. Wang used rule-based information extraction as a method to learn ontology instances. It automatically extracts the wanted factors of the instances, with the help of the definition in domain ontology [11].

Wu yuhuang proposes a web based ontology learning model. This approach concerns realizing the ontology's automatic extraction from the Web page and exploring the pattern and the relations of the ontology semantics concept from the Web page data. It semi-automatically extracts the existing ontology through the analysis of Web page collection in the application domain[12].

III. DEVELOPING CONTEXT ONTOLOGY FOR CRICKET NEWS DATA

This section discusses the methodology for construction of cricket context ontology. The annotation generator should support the context based data information extraction. Context search and extraction based on various concepts details like player, umpire, stadium, commentator, match, events, action etc. Context dependent and context

independent details are extracted and stored using object oriented concepts. Our understanding of the overall process for cricket context ontology construction is summarized below. The general steps involved in the practical approach of context ontology construction are:

Step1: BBC Cricket news data are taken as the source of the semantic data.

Step2: Identify the context concepts, concept hierarchy. Identify the abstract, concrete concept classes in cricket context.

Step4: Concepts are classified into disjoint concepts, overlapping concepts, range concepts …etc.

Step5: Build the ontology using ontology construction tool -protégé and the ontology graph to visualize and evaluate the ontology.

The construction of cricket context ontology is been done using Protégé as an Ontology developing tool. Protégé was developed by Mark Musen's group (http://protoge.stanford.edu) at Stanford University.

Context ontology graphical notation is shown in Figure 1.Context concept hierarchy is shown in Figure 2.We selected OWL, as the ontology language, which is standard ontology language recommended by W3C. OWL/XML format for context ontology is shown in Figure 3.

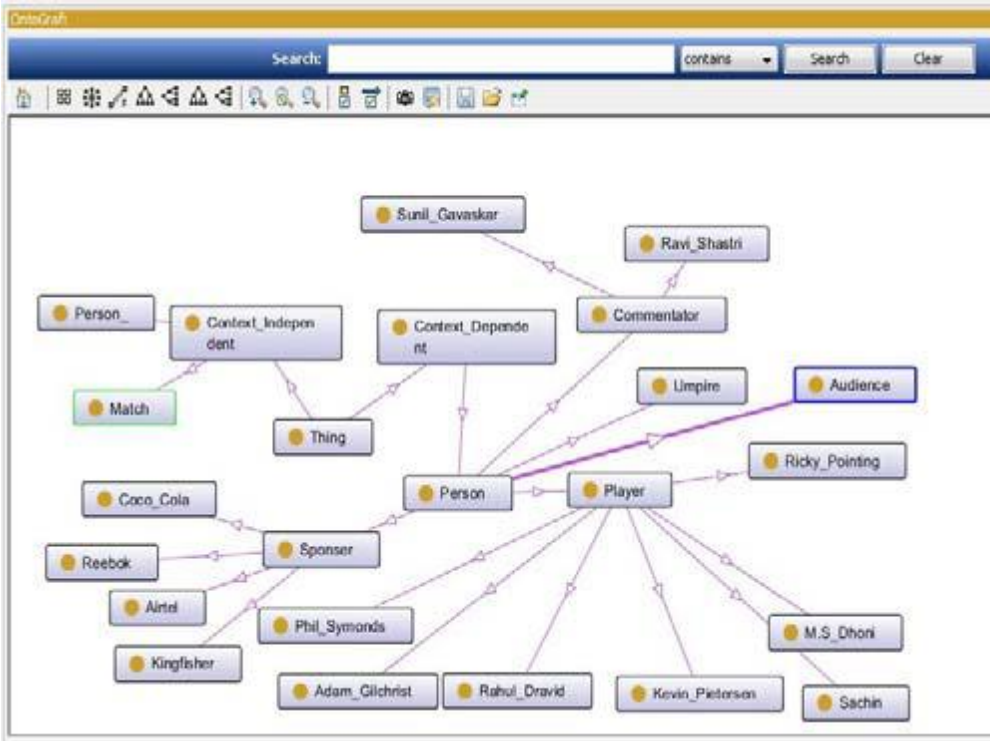

Figure 1. Onto Graph picture of context ontology

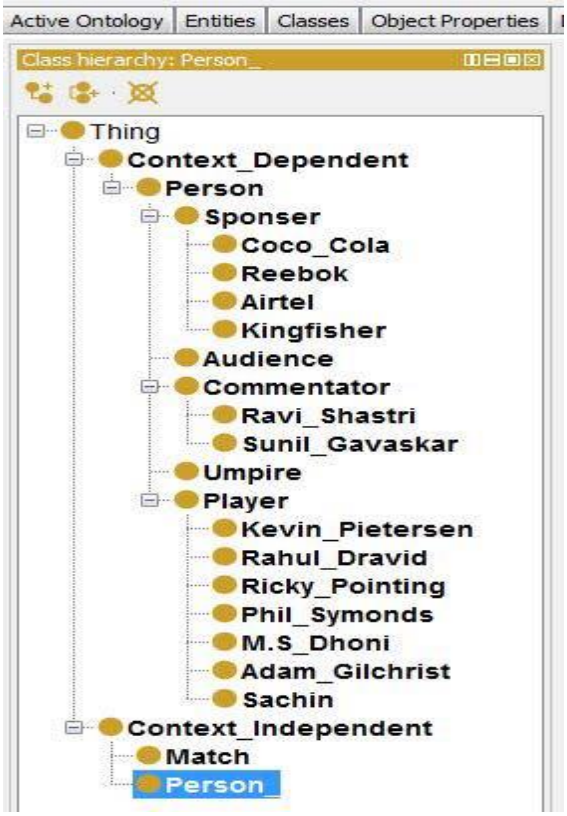

Figure 2. Concepts in cricket context

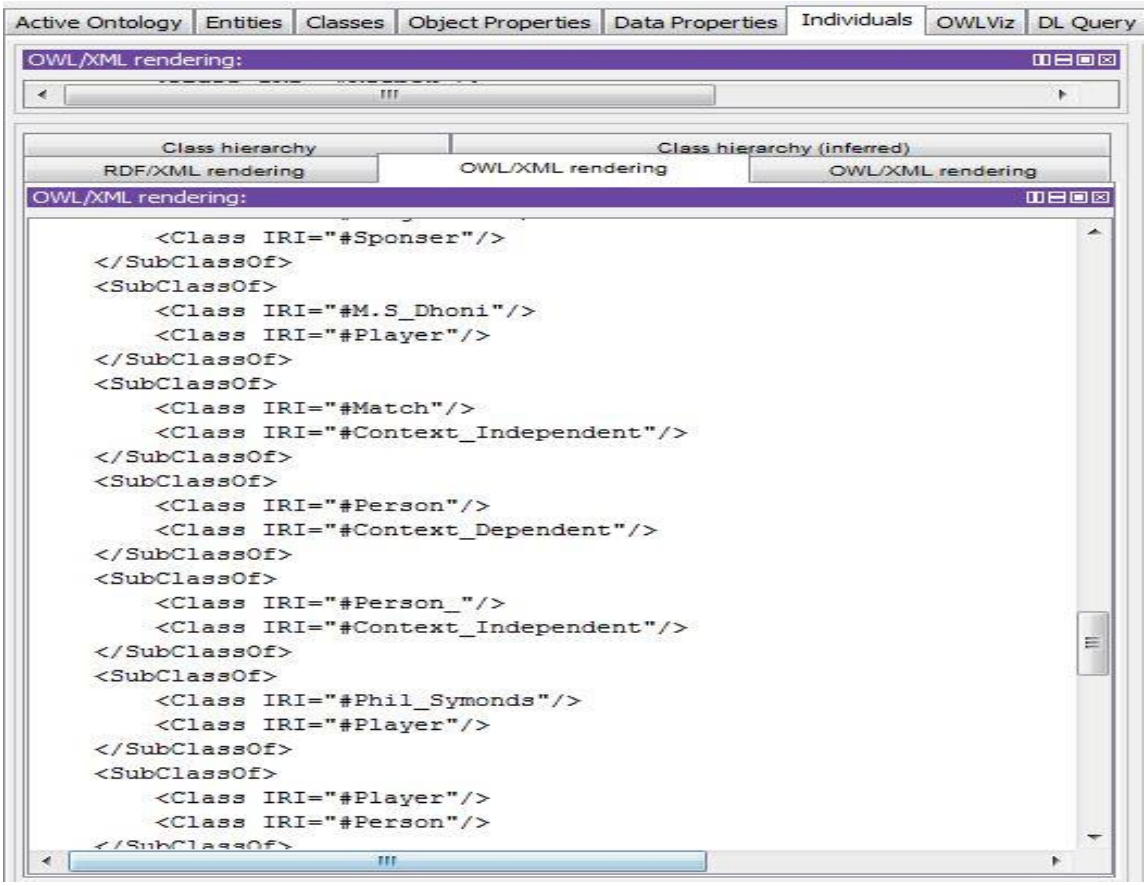

Figure 3. OWL/XML rendering for context ontology

IV. CONCLUSIONS

Since IE is an ontology-based activity and we suggest that future effort in IE should focus on formalizing and reinforcing the relation between the context extraction and the ontology model.. The results obtained over data are quite encouraging; to improve the accuracy of the extraction procedure.

V. REFERENCES

[1]. ArhA. Ozgur, "Supervised and unsupervised machine learning techniques for text document categorization," B.Sc., Institute for Graduate Studies 2004.

[2]. U. Hahn and K. Schnattinger, "Towards text knowledge engineering," in Am. Assoc. for Artificial Intelligence (AAAI-98) 1998.

[3]. A. Maedche and S. Staab, "Semi-automatic engineering of ontologies from text," in 12th international conference on software engineering and knowledge engineering(ECAI-00), Amsterdam, 2000, pp. 231-239.

[4]. A. Maedche and S. Staab, "Ontology learning for the semantic web," IEEE Intelligent systems, vol. 16, pp. 72-79, 2001.

[5]. W. Zhou, Z. Liu, Y. Zhao, G. Chen, Q. Wu, M.-l. Huang, et al., "A semi-automatic ontology learning based on wordnet and event-based natural language processing," in International Conference on Information and Automation, 2006, pp. 240-244.

[6]. K. Hyunjang, H. Myunggwon, and K. Pankoo, "Design of the automatic ontology building system about the specific domain knowledge," in 8th International Conference Advanced Communication Technology, 2006, pp. 4-14.

[7]. A. M. Iqbal, A. R. Mohd, and Z. Khan, "A semi-automated approach to transforming database schemas into ontology language," in 24th Canadian Conference on Electrical and Computer Engineering (CCECE), 2011, pp. 930-933.

[8]. J. F. Sowa, "Top-level ontological categories," International journal of human-computer studies, vol. 43, pp. 669-685, 1995.

[9]. Y. Sure, J. Angele, and S. Staab, "OntoEdit: Guiding ontology development by methodology and inferencing," in International Conferences on the Move to Meaningful Internet Systems, 2002, pp. 1205-1222.

[10]. P. Buitelaar, D. Olejnik, and M. Sintek, "A protege plug-in for ontology extraction from text based on linguistic analysis," in European Semantic Web Symposium, 2004, pp. 31-44.

[11]. J. Wang, C. Wang, J. Liu, and C. Wu, "Information extraction for learning of ontology instances," in International Conference on Industrial Informatics, 2006, pp. 707-711.

[12]. W. Yuhuang and L. Yusheng, "Design and realization for ontology learning model based on web," in International Conference on Information Technology and Computer Science, 2009, pp. 485-488.