SURJ

Sindh Univ. Res. Jour.

TOWARDS AN AGENT-BASED INFERENCE RULE ENGINE MODEL FOR KNOWLEDGE CLASSIFICATION IN CYBERSPACE

G. A. Mallah, N. A. Shaikh, M. S. Memon* and Z.A. Shaikh**
ghulam.ali@salu.edu.pk, noor.shaikh@salu.edu.pk, sullemanhalai@yahoo.com
zubair.shaikh@nu.edu.pk

Department of Computer Science, Shah Abdul Latif University, Khairpur Mirs, Pakistan

Abstract
The World Wide Web has become a primary source of information gathering. By increases of its use knowledge become scattered over WWW. Peoples go online to get their desired information but due to absence of appropriate classification, desired information is being retrieved after putting some additional effort. The challenge of classification of knowledge is under research because of dynamicity and type of contents knowledge contained in. On the other hand software agents are special and intelligent programs having ability of working through a wide variety of devices and on a bigger data communication protocol canvas to achieve assigned task. This paper presents the process of knowledge classification through software agents in a modular fashion with detailed discussion on the process of Analysis/Evaluation of knowledge item object using Inference Rule Engine.

1. Introduction

Agent-based systems are becoming commonplace since last few years and getting more attention in development of intelligent systems and technologies. Whereas the information being gathered from World Wide Web is large enough hence some efficient methodology is required for its archiving and management. Knowledge classification is one of the solutions of the identified issue. Though it is not the only solution of this problem but it also solves the problem of the relevance of information retrieval in searching and many more. New information on World Wide Web is constantly being generated and updated. To address this problem efficient technique is required to classify the knowledge item got from cyberspace that may continuously needed continual learning process to carry on the operations. This will facilitate to acquire knowledge step by step and synchronize dynamic and adaptive scenario of information in cyberspace (Park, *et al.,* 2004).

In order to classify knowledge automatically, we must have knowledge of rules by the help of which we can enable a system to classify knowledge automatically (Andrei, *et al.,* 2007). Knowledge classification and knowledge categorization is the term synonymously used in this paper. This paper focuses on knowledge classification resided in any document type instead of only textual document classification. (Jones, *et al.,* 2009). In entire process of knowledge classification only Analysis/Evaluation of knowledge item using Inference Rule Engine (MCRDR) has been focused. Knowledge item will be treated as object or data structure having list of attributes with document type. It can be any type of document like text documents, images, videos, audios etc (Macskassy Sofus, Provost Foster, 2008). It will not contain the actual data or documents in any case rather summary or keywords upon which inference rule engine will work. Knowledge items other than textual

---

*Department of CSE/IT, Quaid-e-Awam University of Engineering Sciences and Technology, Nawabshah.
**Director FAST-National University, Karachi campus.

documents require meta information associated with file format for its tagging (Li Tao, *et al.,* 2009).

2. Software agents

Software agent is defined as "an umbrella term for a heterogeneous body of research and development" (Hyacinth, 1996). Software agents are not merely confined to computer science but also involve diverse fields such as sociology, psychology, etc (Tosic, *et al.,* 2004). According to Russell and Norvig agent performs two tasks: It senses its surrounding environment through sensors and performs actions in it with its effectors. One of the familiar definitions of the agents is known as the special software that involve in communications, bargaining, coordination, and perform so many other actions autonomously same as being done in real life (Michael Pazzani and Daniel Billsus, 2004).

3. Knowledge classification process

A process of knowledge classification consists on the following sub-processes, shown in **(Fig.1)**.
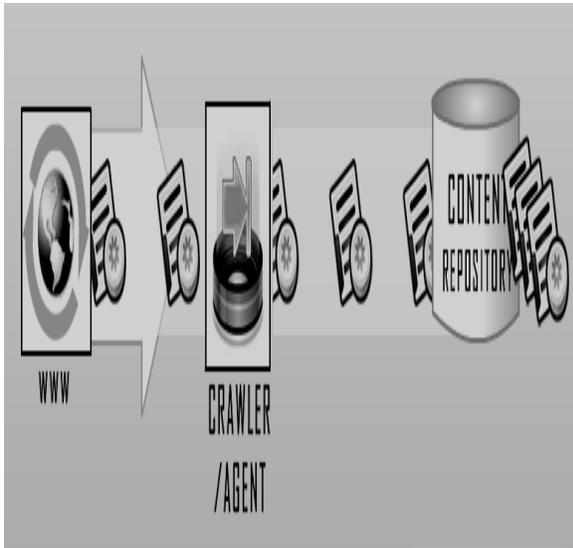

Fig. 1: Process of Knowledge Classification

a   Knowledge acquisition

b   Transformation of knowledge item into unified format.

c   Analysis/Evaluation of knowledge item object using Rule Based Engine

a   Analysis of knowledge item object for its possible classification

b   Placement of knowledge item object in knowledge repository under certain classes identified by defined rules.

3.1 Knowledge acquisition. It is a process in which knowledge items are acquired from WWW through software agents available at web and web crawlers.   In the entire process of knowledge classification this acquisition can easily be automated which will gather information item of any type for knowledge classification.

3.2 Transformation of knowledge item into unified format. It is the process in which any document type will be converted into unified knowledge item object by the help of transformation layer. Transformation layer will deal the complexity of reading meta-information associated with the document as shown in figure-2. By keeping in mind the dynamicity of WWW and type of documents, file formats on WWW are introduced at this layer (Xu, J., 2000). By unifying the knowledge item object, addition of new types of document will not have effect on the implementation of Inference Rule Engine as this complexity will be dealt by transformation layer. This process will work automatically and just a plug-in/driver requires when any new type of document is introduced.
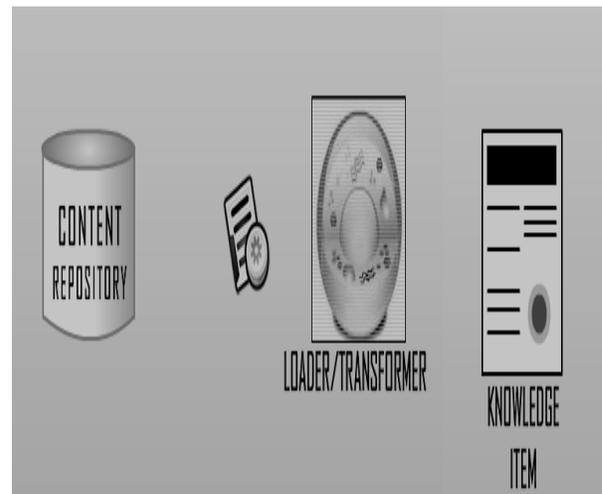

Fig. 2: Transformation of Knowledge in Specified formats

3.3 Analysis/Evaluation of Knowledge item. It is a process which works on the attributes of knowledge item object in order to identify its class to place at appropriate position in the hierarchy of knowledge repository.

3.4 Knowledge Repository. It is a hierarchical data structure in which knowledge items are organized in term of their classes after process of Analysis/Evaluation of Knowledge item object.

The most crucial process in the entire process of knowledge classification is the Analysis/Evaluation of knowledge item object through inference rule engine, MCRDR as mentioned in (**Fig. 3**).
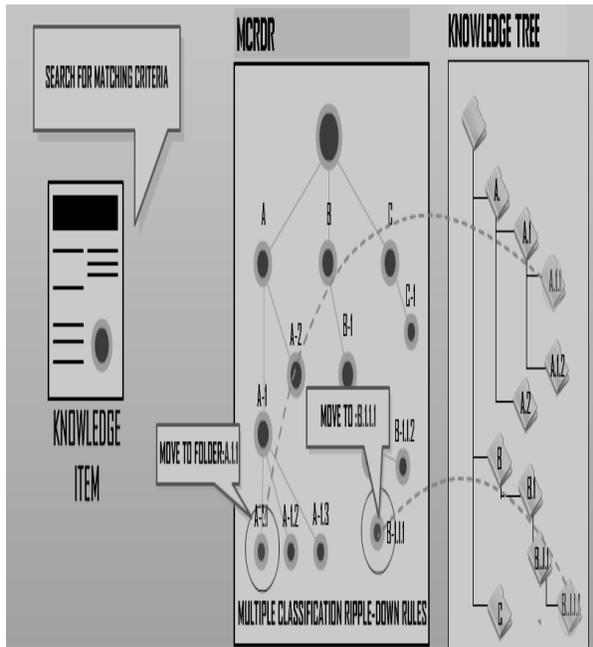

Fig. 3: Analysis and Evaluation Process

In order to implement this process we will need knowledge to define and maintain Inference Rule Engine. Performance of knowledge base is based upon Effectiveness of Domain Expert and knowledge engineer. Knowledge tree has been proposed as data structure to store knowledge classification but in order to cope up the situation where a knowledge item belongs to more than one classes, its reference should be stored instead of complete knowledge item. This functionality

can either be handled by File System or implementation of MCRDR inference Rule Engine. If we go for our own way to store knowledge item on storage media then we can index the knowledge item in term of most frequently used items, to have better searching. Knowledge repository then will be the tree structure which only holds the reference of actual knowledge item, where as actual knowledge item will be stored somewhere in file system (J.I. Col, 2003).

4. Proposed Technique For Implementation

The Multiple Classification Ripple-Down Rules (MCRDR) has been found useful technique for knowledge identification. Opposite to ML (Machine Learning) process the Ripple-Down rules builds incremental base to classify knowledge over time based on the process of differentiation through specialist. That specialist and skilled is capable to identify major properties when the Case-Based Reasoning (CBR) system of Ripple-Down rule accesses and points out the problems that are recognized by the specialists as unsuitable. Through this process knowledge acquisition is applicable and hence some other rules are formed. This type of methodology and algorithm makes possible Ripple-Down rules feasible to design and dynamically adjusted to match the adaptable nature of cyber information (Park, *et al.,* 2004).

Rule Engine Knowledge base inference is an   n-ary tree as shown in (**Fig.4**), where every node in the tree represents a rule having a relative problem. The inference process will involve all rules involved to actual and targeted parents focused against the list of attributes in knowledge item object. Therefore, the process is initiated by analyzing the root rule and the moving downward stage by stage till a leaf node is appeared or none of the child node evaluates to true (Park, *et al.,* 2003). Since various pathways of refinement may be gone through and various findings can be considered and opted amongst the list. This shows that ending and conclusive rule on each pathway gives final result in form of conclusion for the considered case (Dazeley, R. and B., Kang, 2003).
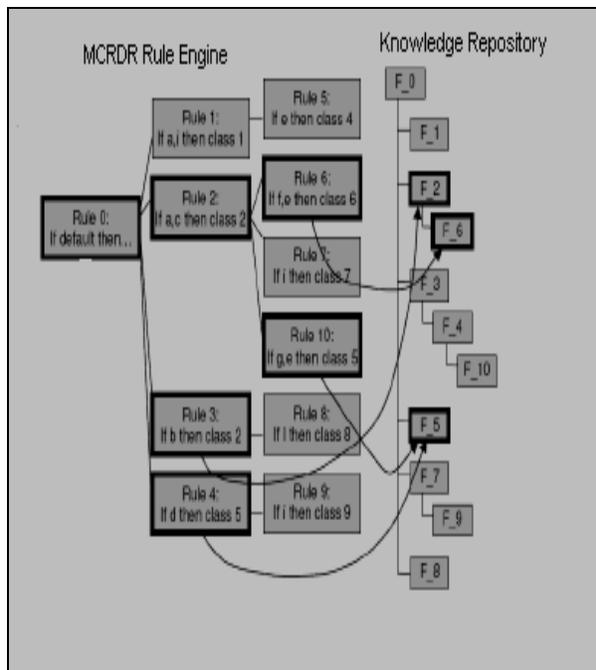
Fig. 4: MCRDR Knowledge base inference

5. Conclusion

The main focus of this research is to build a process for knowledge classification or categorization. In entire process the concept of knowledge identification using MCRDR technique as inference rule engine has been implemented. This research does not deal with the knowledge classification specifically with respect to contents in which knowledge resides; rather process decouples the logic of inference rule engine by introducing a unified knowledge item. MCRDR inference rule engine has been proposed in the model that analyzes the unified knowledge item, through the rules defined by domain expert and knowledge engineer. Finally knowledge item is placed under the classifications identified.

Another useful aspect of the process of knowledge classification mentioned in the model is modularization. An implementation and improvement has been done in the desired area like knowledge acquisition, transformation of knowledge item into unified format and Analysis/Evaluation of Knowledge Item. Transformation layer takes care of the representation of knowledge item in a form as

the MCRDR inference rule engine may perform its operations. Magnificence of MCRDR is that it can work on incremental knowledge, which can cope up the World Wide Web dynamicity.

Future work for this research will be improvement of sub-processes of knowledge classification process. Another aspect in which future work focuses is defining relationship between user searching goal and knowledge item other than keyword based approach. Since user searches are aimed for particular goals and these goals are achieved by retrieving desired knowledge item.

References

Andrei, Z., Fontoura M. and Z. Tong (2007) Robust classification of rare queries using web knowledge, Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, Amsterdam, The Netherlands, 231–238.

Col, J.I. (2003) The UCLA Internet Report Surveying the Digiral Future Year vol. (**3**): UCLA Center for Communication Policy.

Dazeley, R., and B. Kang (2003) Weighted MCRDR: Deriving Information about Relationships between Classifications in MCRDR published in 16th Australian Joint Conference on Artificial Intelligence, Perth, Australia.

Hyacinth, S. (1996) Software agents: An Overview, Knowledge Engineering Review, at Cambridge University Press. Vol. 11, No (**3**): 1-40.

Jones, J., and A. Goel (2009) Meta reasoning for adaptation of classification knowledge, Proceedings of The 8th International Conference on Autonomous Agents and Multiagent systems, Budapest, Hungary, 1145-1146.

Macskassy Sofus, Provost Foster (2008), Classification in Networked Data: A Toolkit and a Univariate Case Study, The Journal of Machine Learning Research, Vol. (**8**): 935–983.

Michael P. and D. Billsus (2004) Learning and Revising User Profiles: The Identification of Interesting Websites, Journal of Machine Learning, Vol. (**27**): 313-331.

Park, S.S., and Y.S. Kim (2003) Web Information Management System: Personalization and Generalization, published in IADIS International Conference WWW/Internet.

Park, S. S., and Y. S. Kim (2004) Adaptive Web Document Classification with MCRDR. International Conference on Information Technology: Coding and Computing ITCC, Orleans, Las Vegas, Nevada, USA.

Park, S. S., and Y. S. Kim (2004) Web Document Classification: Managing Context Change, published in IADIS International Conference WWW/Internet Madrid, Spain.

Tao, L., V. Sindhwani, C. Ding and Y. Zhang, (2009) Knowledge transformation for cross-domain sentiment classification, Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, Boston, MA, USA, 716-717.

Tosic, P. and G. Agha (2004) Towards a Hierarchical Taxonomy of Autonomous Agents, published as proceedings of IEEE International Conference on Systems, Man and Cybernetics (IEEE-SMC'04), The Hague, The Netherlands.

Xu, J. (2000) Multilingual search on the World Wide Web, published in HICSS-33, January Maui, Hawaii 4-7.