ISSN:0975-9646

K. Karthikeyan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , 2014, 5801 - 5806

Markov Logic Network: Unify Framework for Ontology Learning

K. Karthikeyan[1], Dr. V. Karthikeyani[2]

[1]*Research and Development Centre, Bharathiar university, Coimbatore*

[2]*Thiruvalluvar Govt. Arts College, Rasipuram*

Abstract **- Real-world data is most often presented in inconsistent, noisy, or incomplete state. Probability is suitable framework for uncertainty and Logic theory is suitable for complexity. But Statistical Related learning based Markov Logic is suitable for both uncertainty and Complexity. The Markov logic is an advanced and encouraging method to handle this kind of uncertainty presented in the structured data. Markov logic joints the gap between the first order logic and then the probabilistic theory. A Markov logic network (MLN) is collection of first-order logic formulas called rules. This each rule is connected with a numerical weight. Markov logic provides facility to specify probability distributions among complex relational domains and Deterministic dependencies are specified by formulas with infinite weight.**

Keywords – Markov Logic Network, Ontology Learning, Semantic Web

1. INTRODUCTION

The Markov Logic Networks[1] provide a strong probabilistic modeling framework based on First-Order Logic. The statistical relative learning combines the communicatory power of data representation formalisms with probabilistic learning approaches, so facultative one to represent syntactical dependencies between words and capturing applied mathematics information of words in text. Here, the MLN process should be combined with the Markov Random Fields. The weights placed in a MLN process it may be either positive or negative. The MLN process has two kinds of constraints. There are hard constraints and soft constraints. The set of possible worlds are placed in hard constraints and also the set of impossible worlds could be placed in soft constraints.

The Markov Logic Network having two types of inference tasks. There are maximum a posteriori (MAP) and probability inference[2]. The aim of MAP inference is to find the most probable state of the world given some evidence. According to the truth assignments we have to maximize the sum of weights in the network. There are two approaches for learning the weights of a given set of formulas: generative and discriminative learning. Generative learning aims at maximizing the joint likelihood of all predicates while discriminative, at maximizing the conditional likelihood of the query predicates given the evidence ones. Probabilistic logical thinking aims at determinative the likelihood of a formula given a set of constraints and may be alternative formulas as proof. The likelihood of formula is that add of the possibilities of the worlds wherever it holds.

Normally, Markov Logic Networks consists of a weighted first order formule is also called as clauses or rules. In that we have to describe the truth associated weight. The probability of the particular truth assignment can take the variable of x, its probability can be described as follows,

$$P(X=x) = \frac{1}{\mathcal{Z}} \exp\left(\sum_{f_i \in \mathcal{F}} w_i \sum_{g \in \mathcal{G}_{f_i}} g(x)\right)$$
$$= \frac{1}{\mathcal{Z}} \exp\left(\sum_{f_i \in \mathcal{F}} w_i n_i(x)\right)$$

Where, g(x) is 1 means value of g has to be satisfied or otherwise not satisfied with every values in the process.

For learning markov logic networks, it consists of two tasks. There are structure learning and weight learning. The weight learning is the independent component that can be learn weights for clauses written by a human expert. In weight learning we have to use two types of learning approaches named as generative learning and discriminative learning. The structure learning can be performed using an algorithm. The process behind the structure learning based on search methods. Beam search or shortest first search can be used to develop the candidate clauses. All candidate clauses are evaluated and added into the markov logic network.

The Alchemy software[3] is mainly used to learning the weights. Using this kind of method we have to produce the modification of finding inference. Here, we can use the exact probabilistic method of learning weights and produce the good results of inference. In that the markov blanket of a query atom can only contains the evidence atoms. The conditional patterns of this method can be described as,

$$\log P(Y=y|X=x) = \log \prod_{j=1}^{n} P(Y_j=y_j|X=x) = \sum_{j=1}^{n} \log P(Y_j=y_j|X=x)$$

Where, X is the set of evidence atoms and Y is the set of query atoms.

This process can help us to reduce the size of markov blanket, when the clauses are satisfied by the evidence. Exact inference is very fast because the MLN contains thousands of clauses.

The MLN weights can be derived from more relational databases. MLN weights can be calculated by the log likelihood manner.

$$\frac{\partial}{\partial w_i} \log P_w(X=x) = n_i(x) - \sum_{x'} P_w(X=x')\, n_i(x')$$

Here, the sum is overall possible databases x' and compute the probability by using current weight vector ω.

MLN can formulate their features into social networks, language processing and spatial statistics. To optimize those process use pseudo-likelihood described as,

$$P_w^*(X=x) = \prod_{l=1}^{n} P_w(X_l=x_l|MB_x(X_l))$$

Where, M B_x (X_l) is the state of the markov blanket of a data. For first order logic we have to compute the probability. In first order logic attributes have one variable for each pair (a,b), where a is an argument of the query predicate and b is the argument of the query predicates with some same values of each pair. Each and every set of predicates have the truth assignment values. According to those values we construct the network inference. Learning weight process could be done by using the network inference values computed with the help of log likelihood method and their possible probabilities.

2. Related Work

2.1. Ontology Learning

Ontology learning [6] will be outlined because the set of methods and techniques used for building metaphysics from scratch, enriching or adapting associate in nursing existing meta-physics during semi-automatic fashion victimization many sources. Ontology may be an illustration of information formal. It offer a transparent and consistent illustration of nomenclature and ways that facilitate individuals to watch issues and managing affairs, offer public vocabulary of areas and outline totally different levels of formal meanings of terms and relationship between terms. The meta-physics will be considered as vocabulary of terms and relationships between those terms during a given domain. Ontology learning use strategies from a various spectrum of fields like machine learning, information acquisition, natural language process, information retrieval, computer science, reasoning and database management.

2.2. Semantic Web

Semantic internet technologies are aimed toward providing the mandatory representation languages and tools to bring linguistics to the current internet contents. The semantic Meta data can be producing the language of Resource Description language. The RDF [7] we are having three kinds of elements. There are resources, literals and properties. The resources are also called as web objects. The web objects are identified through URI. Literals are the atomic values it contain the values of strings, dates and numbers. Properties also identified through URI. Properties are the binary relationship between resources and literals. The linguistics internet offers the potential for facilitate, permitting a lot of intelligent search queries supported websites marked up with linguistics data.

2.3. First-Order Logic

A first order knowledge base [5] could be a set of sentences or formulas in first-order logic. Formulas area unit created victimization four kinds of symbols. There are constants, variables, functions and predicates. Constant symbols represent objects within the domain of interest. Variable symbols vary over the objects within the domain. Operate symbols represent mappings from tuples of objects to things. Predicate symbols represent relations among objects within the domain or attributes of objects. A first-order computer memory unit is often seen as worlds, if a world violates even one formula, its zero likelihood. The basic idea in markov logic is to melt these constraints. Once a world violates one formula within the computer memory unit it's less probable, however not possible.

3. Markov Logic Network is a Unify Framework for Ontology Learning

Ontology is a specification of conceptualization it specifies the meanings of the symbols in an information systems. This ontology has three kinds of components like individuals, classes and properties. Individuals are nothing but the things or objects in the world. Classes are a set of individuals. Properties are between individual and their values. Ontology learning things about with information discovery numerous knowledge sources associated with its illustration through an onto logic structure and together with metaphysics population, constitutes associate approach for automating the information acquisition method [8]. Ontology may be an illustration of information formal. It offer a transparent and consistent illustration of nomenclature and ways that facilitate individuals to watch issues and managing affairs, offer public vocabulary of areas and outline totally different levels of formal meanings of terms and relationship between terms.

Ontologies are square measure the backbone of the linguistics internet also as of a growing range of knowledge-based systems [9]. A long standing goal of Artificial Intelligence is to create an autonomous agent which will scan and understand text. The method of meta-physics learning from text includes three core subtasks. There are learning of the ideas which will represent the meta-physics, learning of the linguistics relations among these ideas and at last one is learning of a set of abstract through rules. Ontology learning tools discover binary relations not just for lexical things however additionally for ontological ideas. Ontology learning is bothered with data acquisition and within the context of this volume additional specifically with data acquisition from text. Ontology learning is inherently multi-disciplinary attributable to its robust affiliation with the semantic web.

In ontology learning had two kinds of evaluation procedures. There are manual evaluation and posteriori evaluation. Manual evaluation has an advantage which is supposed to know the concepts and their relationships in their domain of expertise, they are supposedly able to tell either a given ontology represents a domain or not. The manual evaluation is subjective and time consuming. The posteriori evaluation also same like manual evaluation but small change in that it is more time consuming and represent the evaluation properly. The grammatical relations are arranged in a hierarchy, rooted with the most generic relation, dependent. When the relation between a head and its dependent can be identified more precisely, relations further down in the hierarchy can be used.

Ontologies became omnipresent in current generation information systems. Associate meta-physics is associate explicit, formal illustration of the entities and relationships which will exist during a domain of application. The term meta-physics came from philosophy and was applied to data systems within it have to characterize the formalization of a body of information describing a given domain. One amongst the key drivers for the popularity of that idea was the conclusion that many of the foremost difficult issues within the data technology field, cannot be resolved while not considering the linguistics intrinsically embedded in every explicit data system.

Ontology learning uses methods from a diverse spectrum of fields such as machine learning, knowledge acquisition, natural language processing, information retrieval, artificial intelligence, reasoning, and database management. Ontology learning framework for the Semantic Web that included ontology importation, extraction, pruning, refinement, and evaluation giving the ontology engineers a wealth of coordinated tools for ontology modeling.

Semantic ability means that not simply that two distinct knowledge systems are ready to exchange information during a given format, that the token should have an equivalent which means for each parties [10]. Customary illustration languages for ontologies like the raptorial bird net meta-physics language and publicly obtainable ontologies will greatly facilitate the event of practical systems, but the process of desegregation and reusing ontologies remains fraught with issue. Strategies for machine-controlled discovery of information and construction of ontologies will facilitate to beat the tedium and mitigate non-compliance however gaps and inconsistencies are inescapable.

Learning semantic resources from text instead of manually creating them might be dangerous in terms of correctness, but has undeniable advantages: Creating resources for text processing from the texts to be processed will fit the semantic component neatly and directly to them, which will never be possible with general-purpose resources. Further, the cost per entry is greatly reduced, giving rise to much larger resources than an advocate of a manual approach could ever afford.

Semantic annotation of a corpus will be performed semi-automatically by varied annotation tools that speed up the entire procedure by providing a friendly interface to a website knowledgeable. A manually annotated corpus will be wont to train associate degree info extraction system. The aim of those approaches is that the exploitation of associate degree initial small-sized lexicon and a machine learning-based system for the lexicon enrichment through associate degree repetitive approach.

Producing robust components to process human language as part of applications software requires attention to the engineering aspects of their construction. For that purpose we have to use the GATE (General Architecture of Text Engineering) tool [11]. This tool is used to perform the dataset progress procedure. According to this tool we had to load our xml dataset. This could be processed to generate the input files for further processing. This GTAE tool shall be used in initial stage of ontology process. The ontology learning process mainly focused this kind of tool for enlarge the ontology process. The framework provides a reusable design for an LE software system and a set of prefabricated software building blocks. As a development process it helps its users to minimize the time they spend building new LE systems or modifying existing ones, by aiding overall development and providing a debugging mechanism for new modules.

The GATE framework contains a core library and a set of reusable lupus modules. The framework implements the design and provides facilities for process and visualizing resources, together with illustration, import and export of information. GATE element is also enforced by a spread of programming languages and databases, however in every case they are delineate to the system as a Java category.

Statistical Relative Learning (SRL) [20] focuses on domains wherever knowledge points square measure not freelance and identically distributed. It combines ideas from applied mathematics learning and inductive logic programming and interest in its adult rapidly. One in every of the foremost powerful representations for SRL is Markov logic, that generalizes each markov logic random fields and first-order logic.

Weight Learning in Markov logic could be a bell-shaped improvement drawback, and therefore gradient descent is absolute to realize the worldwide optimum. However, convergence to the present optimum is also very slow. MLN's square measure exponential models, and their decent statistics square measure the number of times every clause is true within the information. The Markov Logic random fields computing the chance in MLNs needs computing the partition functions, that is mostly stubborn. This makes it troublesome to use ways that require activity line searches, that involve computing the perform as well as gradient.

Wordnet was accustomed notice nouns that area unit derived from verbs and to filtrate words that are not in the noun information. Noun in wordnet area unit joined to their derivationally connected verbs, however there is no indication regarding that springs from that. WordNet, that uses documents retrieved from the Web. However, in their approach, the query strategy is not entirely satisfactory in retrieving relevant documents which affects the quality and performance of the topic signatures and clusters. Using Word Net to enrich vocabulary for ontology domain, we have presented the lexical expansion from Wordnet approach providing a method of accurately extract new vocabulary for an ontology for any domain covered by WordNet.

The vocabulary of associate degree object language for a given domain consists of names representing the people of the domain, predicates standing for properties and relations and of logical constants.

The meaning of a predicate is in general not independent of the meaning of other predicates. This interdependence is expressed by axioms and intensional and extensional definitions[16]. An extensional definition of a predicate is simply the list of the names of the individuals that constitute its extension. An intensional definition of a predicate (definiendum) is the conjunction of atomic sentences (definientia) stating the properties an individual must possess for the predicate to apply.

The axioms picture structural properties of the domain and limit the possible interpretation of the primary terms. The intensional and extensional definitions are terminological.

There are containing two ways that conceiving ontology construction, very cheap up predominant within the methodology of arithmetic and a prime down approach that's predominant disciplines wherever the domain consists of objects of the planet a given as in science.

Alchemy [13] is a software package providing a series of algorithms for statistical relational learning and probabilistic logic inference, based on the Markov logic representation. Alchemy may be a computer code tool designed for a large varies of users. It assumes the reader has noises of classical machine learning algorithms and tasks and is accustomed to first-order and markov logic and a few likelihood theory. Alchemy may be a add progress and is frequently being extended to satisfy these desires.

During weight learning every formulas are reborn to conjunctive traditional form (CNF) and a weight is learned for every of its clauses. If a formula is preceded by a weight within the inpu.mln file, the load is split equally among the formula's clauses. The load of a clause is employed because the mean of a Gaussian previous for the learned weight.

This paper told Anaphora resolution system. This anaphora lets be the front step of Natural Language Processing (NLP). Understanding the Text means, understanding the meaning of context or concept [22]. Anaphora system produces most likelihood antecedent after development of machine learning approach. The knowledge poor strategy provides best results compare to previous knowledge rich strategy. The computational strategy provide maximum share to produce most accurate antecedent. But not least, preprocess task is base for computational strategy perform well good manner.

The name of animal or human being is more concentrate in AR community to categorization of candidate set. It is easy to accept or reject the discourse in candidate set. So, we concentrate much effort to be taken against to recognize the animacy agreement. But this system constructed rule based AR system [14].

Learning from text and natural language is one of the great challenges of Artificial Intelligence and machine learning. The Probabilistic Latent Semantic Analysis (PLSA) [15], [16], [17], [18], [19] is a statistical model which has been called as Aspect Model. The aspect model is a latent variable model for co-occurrence data which associate an unobserved class variable. The PLSA model has an advantage of the well established statistical theory for model selection and complexity control to determine the optimal number of latent space dimensions.

Commonsense knowledge parsing can be performed using a combination of syntax and semantics, via syntax alone (making use of phrase structure grammars), or statistically, using classifiers based on training algorithms. Probabilistic inference aims at determining the probability of a formula given a set of constants and, maybe, other formulas as evidence. The probability of a formula is the sum of the probabilities of the worlds where it holds.

Ontology learning can be dividing the portion of relationship mechanism. This process called as semantic relation extraction. This kind of ontology learning process usually takes the steps of finding weight learning. The first process behind in this stage is to construct the predicates and first order logic formula. A first order logic formula can be constructing from the first order logic knowledge base using Markov Logic Network. According to the predicates we have to build the MLN file. After that build the training file and also the test files. These all based on the input we are taking at the time initialization. Here, we have to find the inference values and last predicated number etc., and then we have to learn the parameter weights of each and every document progressed in our corpus.

4. Learning Ontology Using Markov Logic Network

The process of finding inference can be done with the help of markov logic network. The MLNs inference including maximum likelihood inference calculates the marginal probabilities and calculates the conditional probability. Maximum likelihood reasoning process can be expressed a predicate formula x and find the relative formula y which can be express the probability as

$$\max_y P(y/x)$$

Then according to the markov logic network's joint probability formula it can be transformed to

$$\max_y \sum_i w_i n_i\left(x, y\right)$$

According to this transformation we have to find the inference value for each every document in the corpus.

Next, we do the process of weight learning. For finding weight learning we have to take the discriminative learning based MLN method. Then we assume the evidence predicate x and query predicate y, this could be the large collections of y produce the probability condition of,

$$p_w(y|x) = \frac{1}{Z_x}\exp(\sum_{i\in F_r} \omega_i n_i(x,y))$$

Here, we take the greater than true value number. According to that find the weights in the learning process.

Originally, proposed learning weights generatively using pseudo-likelihood. Pseudo-likelihood is the product of the conditional likelihood of each variable given the values of its neighbors in the data. While efficient for learning, it can give poor results when long chains of inference are required at query time. Pseudo-likelihood is consistently outperformed by discriminative training, which minimizes the negative conditional likelihood of the query predicates given the evidence ones.

In an MLN, the derivative of the negative conditional log-likelihood (CLL) with respect to a weight is the difference of the expected number of true groundings of the corresponding clause and the actual number according to the data. Use this approach and discuss how the resulting method can be viewed as approximately optimizing log-likelihood. However, the use of discriminative MLNs is potentially complicated by the fact that the MPE state may no longer be unique.

4.1. Concept Extraction

In concept extraction first we have to remove the noise presented in our system. Noise means spelling error, abbreviations and word variants. After that we move on the co-occurrence analysis, which means analyze the co occurrences of the data. Next process is the concept identification. In concept identification we have to use the Markov Logic Networks for processing the learning weight and inference. In our concept we use the discriminative learning process for finding the learning weight.

This learning weight maximizes the conditional likelihoods of the query predicates of given evidence and the atoms of unknown truth values handled with EM (Expectation maximization). The Markov logic network is used in our concept. We use the MCMC (Markov chain and Monte Carlo) combination with MC-SAT algorithm to find the probabilistic inferences. Deterministic dependency produce disconnected regions, with out support of probability distribution, it seems to be complicated in design Markov chains for MCMC inference.

An MCMC algorithm using SampleSAT procedure to solve deterministic and near-deterministic dependency in proper way and switch over between isolated or near-isolated region with non-zero probability. This MC-SAT get input from Markov logic, this Markov logic has Markov network and first-order logic and used to calculating conditional probability in graphical model using Markov Chain Monte Carlo. So, MC-SAT used to process the sample into Markov logic using SampleSAT to generate new state for given variables.

4.2. Concept Hierarchy Extraction

In concept Hierarchy extraction, we use MLN method to learning the weight of words for Link Prediction. For that purpose we can use the simple weight learning method to produce the good results. In that situation, need to do the efficient way of hierarchy extraction. For this reason we predict the links in the social networks. The link prediction process is to be done with method of Markov Logic Network (MLN). The main progress include in that is to find the inference values. For finding the inference we could use alchemy process. It may be the software to produce the optimized values of every word in the corpus. Alchemy Packages also used for implement the concept identification process. Alchemy packages are used for make the perfect inference process.

4.3. Semantic Relation

It consists of the steps of first-order logic predicates and finding inference and learning weight based semantic relation extraction. Here the construction of predicates mainly focuses the Knowledge Base of the formulas. According to the knowledge base we have to construct the first-order logic predicates. After that we have to build the .mln file, build training file and also build the test file. Then next, we have to find the inference and weight learning of the contents, this is done with the help of Markov Logic Network. These are performed by the same method taken from the state-of-art method.

5. Conclusion

This process encourages the method of knowledge base. According to this knowledge base we have to construct the predicates. Because it might be used the knowledge base formula symbols. The symbols are carried out by many parts. There are constants, variables, functions and predicates. A term is any expression representing an object in the domain. It can be a constant, a variable, or a function applied to a tuple of terms. After that we have to build some kind of files such as .mln file, training file and also test files. Second process is find inference and weight learning we use MLN method to learning the weight of words. For that purpose we can use the simple weight learning method to produce the good results. The main progress include in that is to find the inference values. For finding the inference we could use alchemy process. It may be the software to produce the optimized values of every word in the corpus. Alchemy Packages also used for implement this identification process. Alchemy packages are used for make the perfect inference process.

Reference

[1] Lucas Drumond and Rosario Girardi, “An Experiment Using Markov Logic Networks to Extract Ontology Concepts from Text,” ACM Special Interest Group on Applied Computing, pp. 1354-1358, 2010.

[2] Daniel Lowd and Domingos.P, ”Efficient Weight Learning for markov logic networks”, Proceedings of the Eleventh European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, pp. 200-211, 2007

[3] K.Karthikeyan and Dr.V.Karthikeyani, “Migrate Web Documents into Web Data,” Electronics Computer Technology (ICECT) 3rd International Conference, 2011, Vol. 5, pp. 249 - 253.

[4] Tuyen N. Huynh and Raymond J. Mooney, ”Max-margin weight learning for Markov logic networks”, Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD-09), Bled, Slovenia, pp.564-57, 2009

[5] Lucas Drumond and Rosario Girardi, “A Survey of Ontology Learning Procedures,” WONTO, volume 427 of CEUR Workshop Proceedings, CEUR-WS.org, vol. 427, 2008

[6] Chris Biemann, “Ontology Learning from Text: A Survey of Methods,” LDV Forum, vol. 20, no.2, pp.75-93, 2005

[7] Paul Buitelaar, Philipp Cimiano and Bernardo Magnini, “Ontology Learning from Text: An Overview,” Ontology Learning from Text: Methods, Evaluation and Applications, IOS Press, pp. 3-12, 2005

[8] Cimiano P, Hotho A, Staab S, “Learning Concept Hierarchies from Text Corpora usingFormal Concept Analysis,” Journal of Artificial Intelligence Research, vol. 24, no.1, pp. 305–339, 2005

[9] Cimiano P, Hotho A, Staab S, “Clustering Concept Hierarchies from Text,” Proceedings of the Conference on Lexical Resources and Evaluation (LREC), pp. 1721–1724, 2004

[10] Parag Singla and Pedro Domingos, “Discriminative Training of Markov Logic Networks,” Proceedings of the 20th National Conference on Artificial Intelligence, vol. 2, pp. 868-873, 2005

[11] Rohit Kate and Ray Mooney, “Probabilistic Abduction using Markov Logic Networks”, Proceedings of the IJCAI-09 Workshop on Plan, Activity, and Intent Recognition, 2009.

[12] K.Karthikeyan and Dr.V.Karthikeyani, “PROCEOL: Probabilistic Relational of Concept Extraction in Ontology Learning”, International Review on Computers and software, Vol.9, No.4, 2014.

[13] Kaustubh Beedkar, Luciano Del Corro, Rainer Gemulla, “Fully Parallel Inference in Markov Logic Networks”, BTW, pp.205-224, 2013

[14] Hassan Khosravi, ”Fast Parameter Learning for Markov Logic Networks Using Bayes Nets”, 22nd International Conference, Dubrovnik, Croatia, pp.102-115, September 17-19, 2012

[15] Biba, M., Ferilli, S., and Esposito, F.,”Discriminative Structure Learning of Markov Logic Networks,” Proceedings of the 18th international conference on Inductive Logic Programming (ILP’08), Czech Republic. Springer-Verlag, pp. 59–76, 2008

[16] Tuyen N. Huynh and Raymond J. Mooney, “Discriminative Structure and parameter learning for Markov Logic Networks”, Proceedings of the 25th International Conference on Machine Learning (ICML), New York,USA, Finland, pp. 416-423, July 2008

[17] Shalini Ghosh, Natarajan Shankar, Sam Owre, “Machine Reading Using Markov Logic Networks for Collective Probabilistic Inferences”, Proceedings of ECML-CoLISD, 2011

[18] Thomas Hofmann, “Probabilistic Latent Semantic Analysis,” Proceedings of 15th Conference on Uncertainty in Artificial Intelligence UAI’99, Stockholm, Sweden, pp.289-296., 1999

[19] Emhimed Salem Alatrish, “Comparison of Ontology Editors”, eRAF Journal on Computing, vol. 4, pp. 23 – 38, 2012

[20] Khondoker, M. Rahamatullah, Müller, Paul, “Comparing Ontology Development Tools Based on an Online Survey”, World Congress on Engineering 2010 (WCE 2010), London, UK, March 2010.

[21] Mark Sanderson and Bruce Croft, “Deriving concept hierarchies from text”, SIGIR '99 Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, pp. 206-213, 1999.

[22] Maryam Hazman, Samhaa R. El-Beltagy and Ahmed Rafea, “Ontology Learning from Domain Specific Web Documents,” International Journal of Metadata, Semantics and Ontologies, vol. 4, no. 1/2, pp.24 – 33, 2009

[23] Zellig Sabbettai Harris, “Mathematical Structures in Language,” Interscience Publishers, p. 230, 1968

[24] Karthikeyan.K and Dr.V.Karthikeyani, “Understanding text using Anaphora Resolution”, Pattern Recognition, Informatics and Mobile Engineering (PRIME) 2013 International Conference, 2013, pp- 346 – 350.