A Knowledge Based Approach for Recognizing Textual Entailment for Natural Language Inference using Data Mining

Dr. Arti Arya[1], Vishwanath Yaligar[2], Ramya D. Prabhu[2], Ramya Reddy[2], Rohith Acharaya[2]
[1]Department of MCA, PESSE, VTU, Bangalore, India
[2]Dept. of Computer Science, PESSE, VTU, Bangalore, India

***Abstract*— Recognizing Textual Entailment (RTE) is a relatively new problem necessary for Natural Language Understanding (NLU) and Automated Knowledge Discovery. Natural Language Inference has dealt with approaches like the bag of words approach, formal methods (First order Logic) and pattern relation extraction which usually do not show satisfactory results. In this paper, Knowledge based approach has been proposed, utilizing data mining concepts on large text which is appropriately classified. Different Lexical resources like WordNet, VerbNet, ConceptNet have been integrated into a rich knowledge base, to provide semantics and structural information on English words. The data mined is used by an Inference system to give the output to the problem. The complete concept presented in the paper has been implemented in the form of a movie search engine wherein the knowledge based RTE concept has been employed on "summaries or plots of the movies" internally to get best possible classification of the movies. The experiments have shown encouraging results, reduced the time of search and provided more accurate results. To the best of our knowledge, it is the first time that RTE concept has been implemented to Information Search in the form of Movie Search Engine.**

Keywords- Recognizing Text Entailment (RTE), natural language processing,data (text) mining, knowledge base.

I. INTRODUCTION

Textual Inference plays an important role in many Natural Language Processing (NLP) [16] tasks. In recent years, Recognizing Textual Entailment (RTE) [3], which focuses on detecting semantic inference, has attracted a lot of attention. The main idea behind RTE is concerned with inferring the meaning of the text from that of another larger text. This has now become a focus area for the natural language processing community and is known as Recognizing Textual Entailment (RTE). The concept of RTE explores the relationship between the hypothesis and text.

Given a hypothesis for a sentence in natural language (English), the system must be able to identify if the text entails the hypothesis, i.e., if the hypothesis can be inferred from the text or not or is the data insufficient to conclude. The relationship is denoted by T → H. For instance, given H = "The actor of movie "Notorious" is Craig Grant." and T = "The movie Notorious was a superhit having Craig Grant in the main lead." the relation T → H holds true.

Successful automation of natural language applications helps in an accurate understanding of the underlying meaning (semantics) of texts with different syntaxes by machines [5]. This becomes a challenging task when different sentences with different words or phrases express the same meaning. The exponential increase in web data creates new challenges for Information Retrieval [15]. Automated search engines that are based on the concept of keyword matching usually return poor quality results in case of same meaning inferred from different linguistic expressions.

In this paper, different approach is explored for improving an inference rule collection and its application to the task of recognizing textual entailment and thus implementing the concept to a movie search engine. Two main concerned issues are:

- To improve the Natural Language Inference ability of the computer.
- To make use of it.

Recognizing Textual Entailment (RTE) also known as Natural Language Inference is a relatively new problem necessary for Natural Language Understanding (NLU) and Automated Knowledge Discovery. NLI has a spectrum of approaches like the bag of words approach, formal methods (First order Logic) and pattern relation extraction which do not show satisfactory results.

In this paper, the Knowledge based approach is proposed, utilizing data mining tools on large text appropriately classified. Given a hypothesis for a sentence in natural language (English), the system must be able to recognize if the text entails the hypothesis, i.e., if the hypothesis can be inferred from the text or not or is the data insufficient to conclude.

The main objective of making a three-way decision of "YES", "NO" and "UNKNOWN" is to drive systems for making accurate informational distinctions. If the outcome of a hypothesis is "unknown" because of insufficient text then that text must be segregated from the hypothesis having outcomes showing "no" .

Different Lexical resources like WordNet[8, 13], VerbNet[14, 25] and ConceptNet[7, 23] have been integrated into a rich knowledge base, to provide semantics and structural information on English words. The data mined is used by an Inference system based on mathematical logic to give the output to the problem. This paper aims to address these issues, by

(a) Enhancing the Natural Language Inference (NLI) ability of the computer by designing and implementing the knowledge based RTE system.

(b) Implementing movie search engine and making use of the RTE system in constructing summaries of movies plots in the database for fast, efficient and accurate search.

The rest of the paper is organized as follows: Section II provides the details of different approaches to RTE and describes some of the work done in the area of RTE. Section III gives the in depth idea of the proposed approach, architecture and construction of Movie Search Engine. Section IV describes the performance related expectations and results of the proposed approach. In section V conclusion and future scope of the approach is discussed.

II. RELATED WORK

A few approaches to RTE have been developed during recent years which are shown in figure 1.

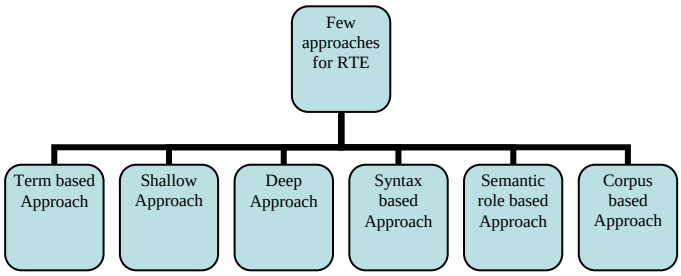

Figure 1. Various RTE based approaches.

In Term-based approach [5], most of the systems consider morphological and lexical variations of the terms (words) in texts and hypotheses and determine the existence of entailment between the texts and hypotheses by means of their lexical similarities. Shallow approach [5], relies on lexical or semantic overlap, pattern-based relation extraction, or approximate matching of predicate-argument structure. Such methods while robust and often effective, are at best partial solutions, unable to explain even simple forms of logical inference. Deep approach[5] considers inference as logical deduction, building on work in theoretical semantics to translate sentences into first-order logic (FOL), and then applying a theorem prover or model builder. Complexities include full semantic interpretation, including tense, aspect, causality, intentionality, vagueness, idioms, and many other issues like high precision but poor recall.

Syntax-based approach carries out a similarity analysis between the dependency trees [5] extracted from the texts and hypotheses in order to identify the entailment relationships. There are also systems that take a paraphrase detection strategy to generate a set of different styles of the hypotheses with the aim of searching for a subset of which may occur in the texts. Semantic role-based approaches annotate the sentences of the texts and hypotheses with semantic roles (using shallow semantic parsers) and then analyze the coincidences between sets of assigned semantic roles.

Corpus based approach [12] makes use of large data in form of text available in form of corpus on which data mining techniques are applied to build probability based models and knowledge base is taken. Lexical database like WordNet is used to provide semantics and structural information on English words. Standard problems sets used worldwide for testing RTE systems are used to test the accuracy and precision of the results. The data mined is used by an Inference system based on mathematical logic to give the output to the problem.

Many approaches are available in literature [5, 12, 18, 19] for handling the problem of textual entailment. Some of the approaches have used the concept of estimating distance between the pair of T and H as main aspect which separates the entailment classes. In one of the algorithms [19], Tree Edit Distance is implemented. In this algorithm distance is computed in terms of cost of insertion, deletion and substitution operation that converts the text T to hypothesis H. The one limitation of this method is that it converts T to H and do not generate inference from it.

Mehdad et. al [18] have proposed textual entailment recognition using Particle Swarm Optimization (PSO). In this paper, PSO technique is applied where learning process for each particle happens with in a swarm (population) and it learns from the behavior of other particles. This technique is based on social sharing of information. This paper makes use of Bhattacharya distance [22] to determine the similarity between two discrete probability distributions.

Haghighi A.D. et. al [20] have proposed a graph matching model that makes textual inferences. Each sentence is represented as a directed graph and nodes of the graph represent words or phrases and links represent semantic and syntactic relationships.

Minqing Hu and Bing Liu [21] have proposed an opinion mining concept of customer reviews. In this paper, they have summarized the reviews of the customers about a product based on specific features of the product. They have not written the summary or review of the products just from the customers review but considered certain features of the product and extracted customer's opinion regarding those features.

Jijkoun and Rijke[26] have used the concept of "directed sentence similarity score" between two sentences. To verify the entailment, they have compared the above similarity score with a threshold value. They have assigned weights to words depending on the importance of the word for similarity identification. They have also used word similarity measure for similarity identification.

Tatu and Moldvan [27] have suggested a technique which employs a set of axioms to extract semantic knowledge from text and they have used a logic prover, which received as input the logic forms of two texts enhanced with semantic relation instances.

In this paper, Knowledge-based approach is proposed, wherein Wordnet, Verbnet and Conceptnet are integrated into a knowledge pool. The utilization of knowledge in RTE systems facilitates recognizing entailment relationships where existing lexical or semantic knowledge is not adequate for confidently inferring the relationships. Any previous RTE system that uses knowledge based approach has not yet been found.

A critical look at the existing systems highlight that the following points needs to be addressed, some of which are discussed in literature but still needs attention [17]:

(a) The documents so retrieved, contains a high percentage of irrelevant documents. The search is based on words drawn from documents and document title. Such searches are unable to list documents wherein the words are embedded only in document. Most of the available search engines do not perform semantic search.

(b) Since HTML doesn't provide any standard method to identify contents of documents, it is extremely difficult for search engine to identify contents of web page to index them.

Many approaches to Text Inference exist but there are still a lot of challenges faced in entailing a piece of text from another. The relatively low accuracy achieved by the existing systems [10] suggests that the entailment task is surely a tedious one. So, there lies a good scope of improvement. To differentiate unknown from false/contradictions also create a challenge.

In this paper, a novel knowledge based semantic search method has been proposed using the concept of RTE for faster results. Proposed method has been implemented in the form of movie search engine, which shows very encouraging results.

III. PROPOSED METHOD AND SYSTEM ARCHITECTURE

The System Architecture gives the overview of the final working system. The user is a person using the query interface provided. The user can query by inputting the different attributes of the movie. The query to be executed depends on the user input attributes. The information of movies is displayed by searching the knowledge base.

The architecture that is followed is a variant of Blackboard architecture [4]. The Blackboard, as an architectural pattern, has been traditionally used in the development of systems with artificial intelligence techniques. Blackboard architectures are particularly suited to solve problems, such as decision support, signal processing and speech recognition [4]. In Blackboard architectures, a central Blackboard (BB) data structure holds the entire state of a solution.

A. System Architecture: Blackboard architecture variant

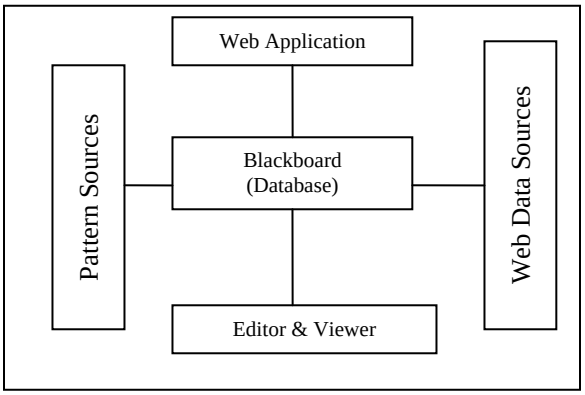

Figure 2. System Architecture

A BB is often hierarchically structured and contains non-homogeneous content. Domain and world knowledge are represented in separate independent Knowledge Sources (KS), which hold computations that respond to changes in BB, and with direct access to it; KS interact through the BB to yield solutions. A Control (C) monitors changes in BB and determines the next action to be executed, plans evaluations and activations according to a strategy that decides which KS is the one that will next change the BB.

Three different lexical resources have been integrated into a knowledge pool: VerbNet [14], ConceptNet [7] and WordNet [13], into a unified, richer knowledge-base, in order to facilitate more robust semantic parsing [6].

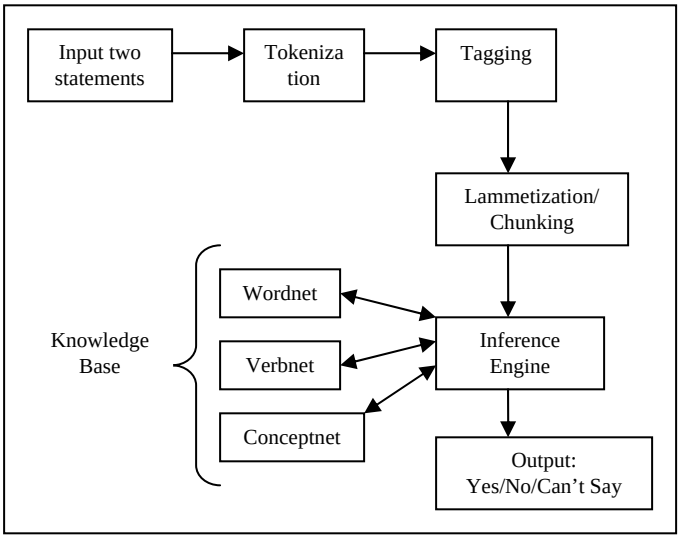

Figure 3. The proposed Knowledge Based Complete RTE System

The first type of knowledge is drawn from WordNet [13], a large lexical database with rich information about words and concepts. This is referred to as word-level knowledge. The latter is derived from ConceptNet [7] – an openly available large-scale commonsense knowledge base with an integrated natural-language-processing tool-kit that supports many practical textual-reasoning tasks over real-world documents, and from VerbNet [14], a verb lexicon which provides

sentence-level knowledge. Knowledge is acquired and represented using various knowledge representation techniques rules, frames and scripts. The WordNet analyses all words in a sentence for all possible senses it can represent in a sentence. Based on this it will deduce which sense each word is representing. WordNet excels at lexical reasoning, ConceptNet's forte is contextual commonsense reasoning — a research area that is poised to redefine the possibilities for intelligent information management.

The proposed RTE system is depicted in figure 3. Tokenization means a text in terms of smaller components, such as words or sentences. The nltk.token [2, 9] module defines classes for representing and processing these smaller elements. The term word can be used either to refer to an individual occurrence of a word or to refer to an abstract vocabulary item.

For example, the sentence "my dog likes his dog" contains five occurrences of words, but four vocabulary items. To avoid confusion more precise terminology [2] is used:

Word token: an occurrence of a word

Word Type: a vocabulary item

The simplest way to represent a text is with a single string. It is difficult to process text in this format. The task of converting a text from a single string to a list of tokens is known as tokenization. The simplest approach is to use "graphic words" (i.e., separate words using white space). Another approach is to use regular expressions to specify which substrings are valid words. NLTK [9] provides a generic tokenization interface: Tokenizer [1, 2]. It defines a single method, tokenize, which takes a string and returns a list of tokens.

Extraction of propositions from each sentence in the hypothesis and the text is called lemmatization. A proposition is a singular representation of concepts in the texts in which there are no clauses or dependent parts of texts included. For instance, from the sentence "The girl watching movie is not my friend." the proposition "girl watching movie" can be extracted. Chunking is an analysis of a sentence which identifies the constituents (noun groups, verbs, verb groups, etc.), but does not specify their intrinsic structure, nor their role in the main sentence.

The proposed method of inferencing textual entailment using knowledge base has been implemented for constructing a Movie Search Engine (MSE).

B. Flow of data in the proposed system

Flow of data in the entire process is shown in figure 4. Intially user will input the query and after processing the query using knowledge base, the movie database is queried and output in the form of images (relevant movies) are displayed to user.

When knowledge base is used for query processing, at that time it interacts with the RTE system for obtaining relevant movies based on the similarity in plots of the movies.

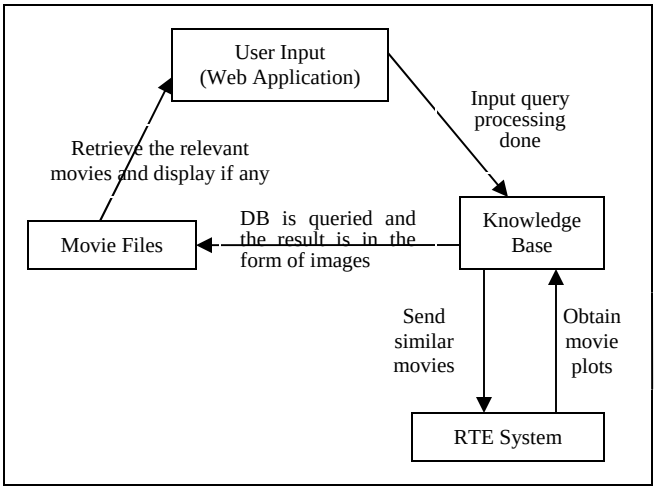

Figure 4. Data Flow Diagram to indicate the entire process.

C. Movie Search Engine(MSE) implemetation using RTE

The Movie search engine is a web application developed in such a way that anyone with a browser can search for movies or patterns among them easily through a good interface. The RTE system so developed constructs "reviews of the movies". The Movie Search Engine so proposed, generate a minimized and most relevant set of results based on the user input. The main users of this search engine may be:

- People who search for movies using various combinations of attributes.
- People who study movies for extracting common patterns among them.
- People searching for the overview or plot of a movie.
- People looking for well constructed summary of reviews.

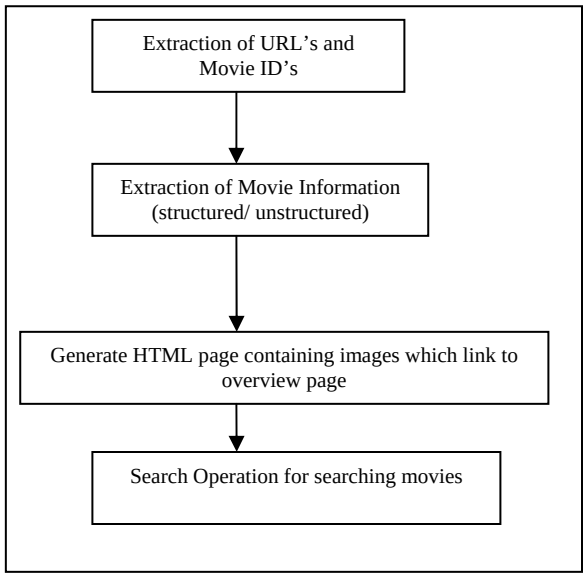

Figure 5. Working of Movie Search Engine

Two steps involved in implementing the search engine are:

1. Construction of movie database by extracting the movies from wikipedia [24].
2. Search operation for searching the movies from the database as per the user requirement as shown in figure 5.

D. Working of RTE for Search Engine

The most exciting part of the RTE Movie Search Engine is the pattern search; it finds similarities based on movie attributes generated by RTE. The RTE system probes the entire script for keywords and develops similarities between movies.

Input to the system: Common Actors or Genre

Output: Movies

Or

Input: Movies

Output: Similar Pattern

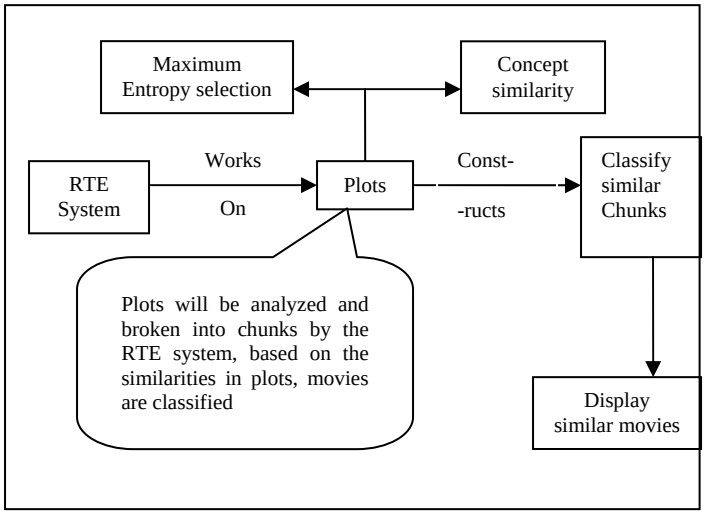

Figure 6. Working of RTE system on plots of the movies

Figure 6 clearly shows the complete working of RTE system on the plots of movies. Plots of the movies are analyzed by the RTE system based on Maximum Entropy selection measure and Concept similarity measure. Thus movies are classified into different classes and movies belonging to same class (similar movies) are displayed.

E. Concept Similarity and Maximum Entropy Selection

Motivation behind Concept similarity and Maximum Entropy Selection is depicted in figure 7 and figure 8 respectively.

Consider two movies M1 and M2 as sets of concepts (words). These concepts are the words occurring in the plots of the two movies. E.g.

M1 (C_1, C_2, C_3, …….,C_n) and M2 (C'_1, C'_2, C'_3,…….C'_n)

The similarity of concepts of both M1 and M2 are compared with a similarity measure "Concept Similarity". Time complexity of this operation is O (n2), where n is the maximum number of concepts used for comparison. Some concepts are given more importance than others based on how much information those concepts can provide.

Here, it is been considered that if the frequency of occurrence of C_1 is greater than the other concept C_2, then C_1 is capable of conveying less information about the characteristics of the movie. So C_1 is assigned lower priority.

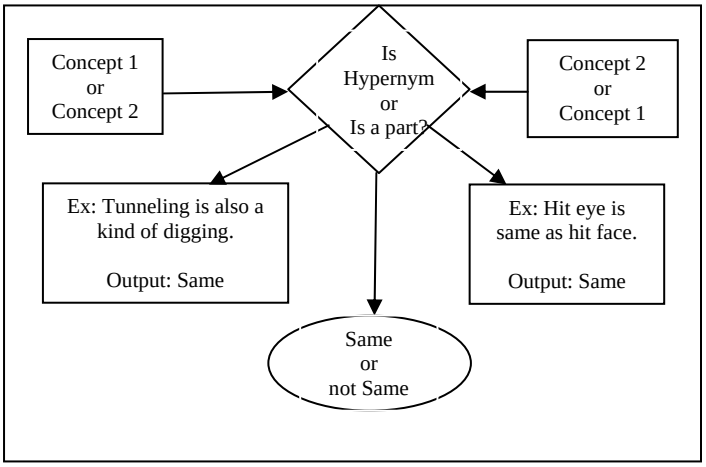

Figure 7. Concept Similarity

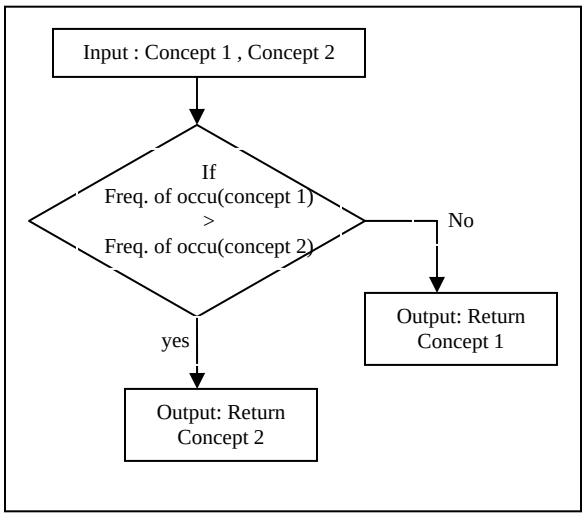

Figure 8. Maximum Entropy Selection

Consider a universal set of all concepts, set of concepts of M1 and M2. Frequency of occurrence of concept in the universal set of concepts is calculated and importance is given according to:

```
If
    Frequency of C1 > Frequency of C2
                    Return C2
else
                    Return C1.
```

The concept returned by the above function gives more information about the movie than the other concept.

For example, consider two concepts like bottle and imprisonment, imprisonment gives more information if its frequency is lower and this supports any human understanding of that if there is bottle in a movie then that is of no use to know for the viewers compared to having imprisonment.
All these concepts are ordered in the descending order of frequency and top n/3 concepts are used for comparison and decision of similarity. During experiments, it is been found that meaningful information about the plots can be obtained from comparing first n/3 ordered concepts and thus comparison of remaining concepts is not required.
In the entire process of comparison, first the concepts occurring with in the same movie are ordered using concept similarity and after ordering the concepts of both the movies, the concepts of both the movies are compared for the similarity of plots.

IV. PERFORMANCE OF SEARCH ENGINE USING RTE

Some of the performance expectations are:

The number of movies to be stored in the database can be extended, in order to provide flexibility and extend the search to a wide range of movies. Time taken to search for a movie is approximately 30 milli seconds, thus giving instant delivery effect and reducing waiting time of the user. Users can use this search engine for the following purpose:

- Search for movies of their taste
- Study common patterns among movies
- To obtain summarized review of desired movies

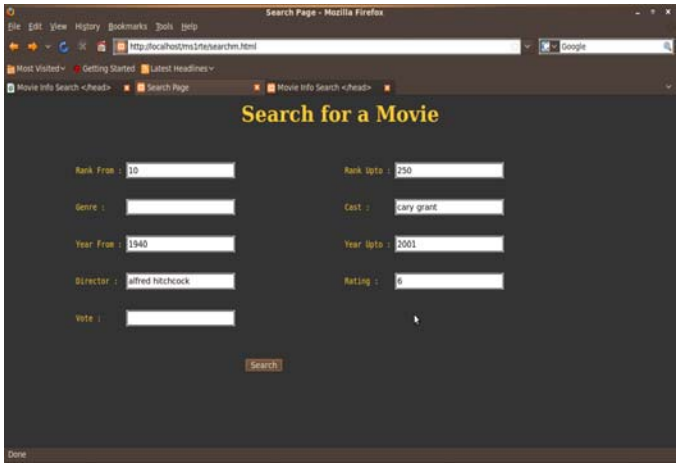

Figure 9. Search Page for entering query

Initially user can input the query in the screen shown in fig10 in order to search a movie of his interest. Based on the criteria entered by the user the search results are shown with a negligible waiting time for the user as shown in fig 11.

Overview page for each movie is generated from the database. An example of movie “Notorious” having Cary Grant in the main lead is shown in figure 11.

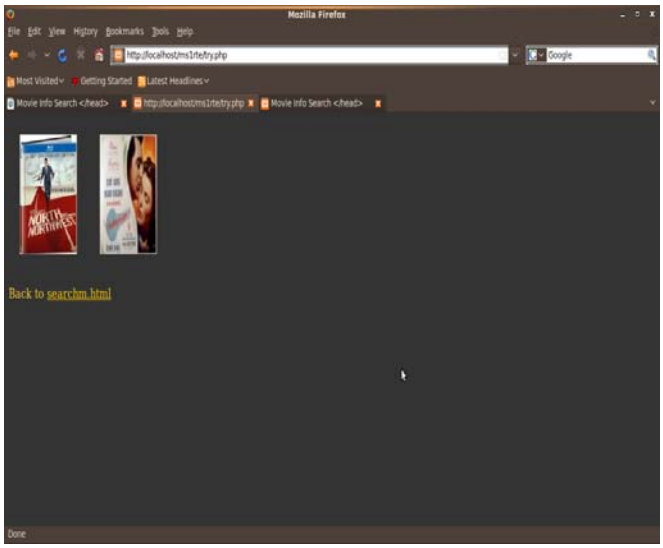

Figure 10. Screen showing search results

The graphical user interface for the RTE system shown in the figure 12, is developed so that the RTE system can work on the plots of the movies at the back ground on the movies database so that movies can be classified as per the requirement.

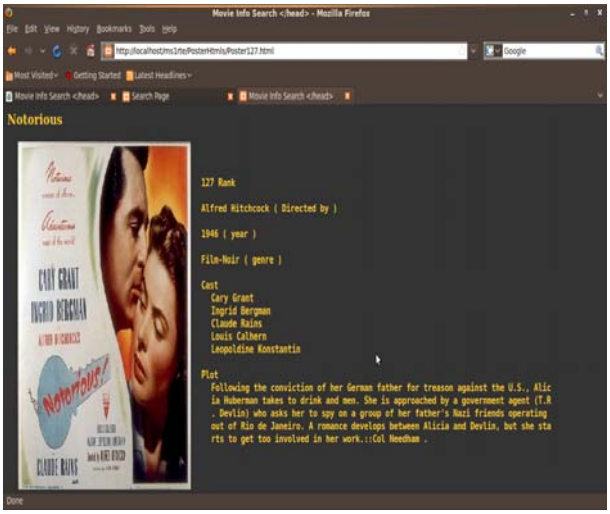

Figure 11. Overview page

For searching similar movies the concept of text mining i.e. Maximum entropy selection and concept similarity are being used.

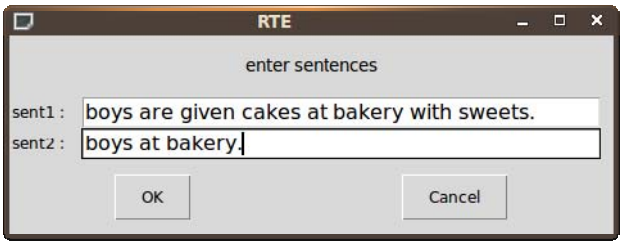

Figure 12. GUI of RTE system

The screen in figure 13 shows the movie database. After inputting the search creiteria in the search page we have obtained a list of all the movies where plots of the movies are similar to first movie in the database as shown in fig 15.

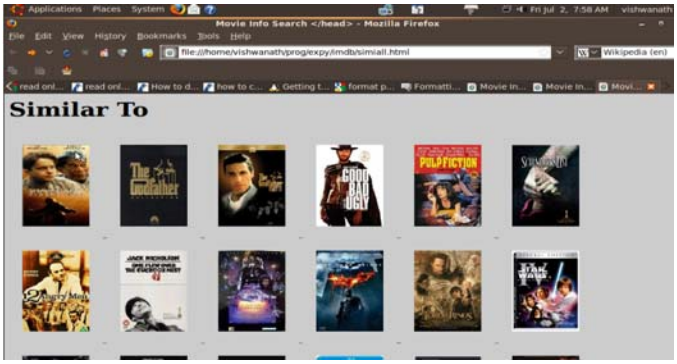

Figure 13. Screen showing all the movies in the database

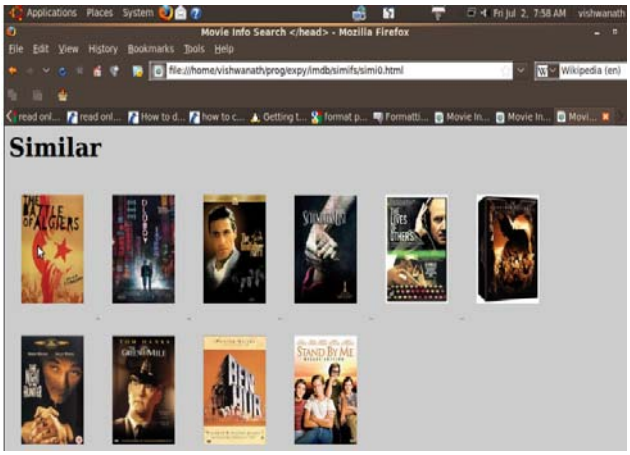

Figure 14. Screen showing movies similar to first movie in the database.

For getting the search result for the movies similar to first movie in movie database, the RTE system had worked on the plots of the various movies and segregated them depending on the similarity in their concepts.

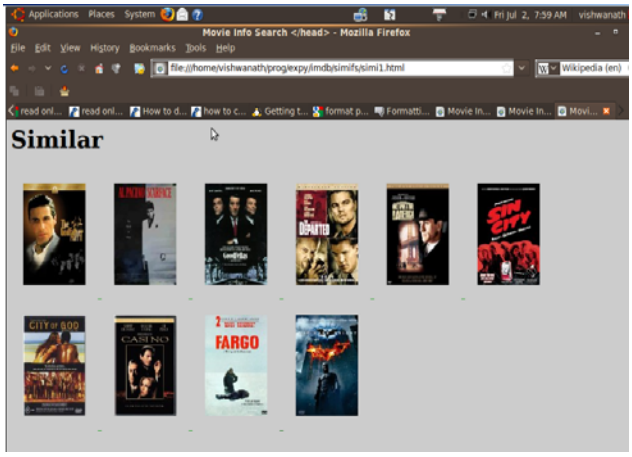

Figure 15. Screen showing movies similar to second movie in the database.

This similarity is adjudged based on the Maximum Entropy Selection and Concept similarity.

The snapshot shown in figure 15 shows the movies which are similar to second movie in the database.

The results shown by the search engine are highly appropriate as far as similarity of plots of the movies is concerned.

V. CONCLUSION AND FUTURE SCOPE

By incorporating this knowledge based approach, it becomes possible to implement the concept of RTE system to Summarization (as of now out of Machine Translation, Question Answering, and Summarization only Machine Translation is implemented). Knowledge Based approach integrates three different lexical resources Wordnet, Verbnet and Conceptnet and is a good solution to Natural Language Inference.

In this paper, a knowledge based RTE system has been developed and made use of it by applying it to a Movie Search Engine. The system has been tested with various information retrieval tasks, and it was able to recognize and lend assistance to almost all queries issued to it, and significantly improved retrieval performance.

The Web Application that makes use of RTE displays summarized reviews of movies. These reviews are worked upon by the RTE system and stored in the database.

The RTE system was developed with the intention of implementing any of the applications of Natural Language Inference. In future, this could also be developed into a Question Answering system.

This system can search and retrieve information of about 250 movies. In future people can use it to find movies to watch — movies that are right for their taste. In this paper, search is based on few selected attributes, so search can be improved in future by including many more attributes to improve the inference level. Moreover, in Maximum entropy selection criterion, the condition where the frequency is same for two concepts is not considered. This small direction will be explored in further work.

The RTE system (knowledge based approach) can be extended to other applications like Question Answering, Information Retrieval and Machine Translation. “Cannot say” output of the RTE system can be worked upon using Fuzzy Logic Theory. The Lexical resources included are Wordnet, Verbnet and Conceptnet. It can be extended to Framenet.

References

[1] NLTK Tutorial: Introduction to Natural Language Processing, Steven Bird, Ewan Klein, Edward Loper, Revision 1.66, 7 Apr 2005

[2] Python for NLP and the Natural Language Toolkit, CS1573: AI Application Development, Spring 2003(modified from Edward Loper’s notes)

[3] I. Dagan, O. Glickman, and B. Magnini., “The PASCAL Recognising Textual Entailment Challenge. In Machine Learning Challenges,” LNCS, volume 3944, pages 177-190. Springer, 2006.

[4] C. Metzner, L. Cortez, and D. Chacín Universidad Central de Venezuela, Caracas Venezuela, Using a Blackboard Architecture in a

Web Application", In issues of Informing Science and Information Technology, 2005, pp.743-755

[5] B. Ofoghi, J. Yearwood, "From Lexical Entailment to Recognizing Textual Entailment Using Linguistic Resources," Centre for Informatics & Applied Optimization, University of Ballarat Victoria 3350, Australia

[6] L. Shi and R. Mihalcea, "Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing," In LNCS, Springer Verlag, 2005, pp. 100-111.

[7] H. Liu and P. Singh, "ConceptNet — a practical commonsense reasoning tool-kit," in BT Technology Journal, 2004, pp. 111-126.

[8] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Five papers on WordNet." Technical report, Cognitive Science Laboratory, Princeton University, 1993.

[9] NLTK website at http://nltk.org/

[10] D. Milne, I. H. Witten, D. M. Nichols, "A Knowledge-Based Search Engine Powered by Wikipedia," In proc. of ACM Conference on Information and Knowledge Management, 2007, pp. 445-454

[11] en.wikipedia.org/wiki/Web_spider

[12] A. Koller, S. Schulte, I.M. Walde, "Introduction to Corpus-based Computational Semantics", University of Stuttgart, World Knowledge, Introductory Course ESSLLI 2007 Trinity College, Dublin, Ireland

[13] http://wordnet.princeton.edu/wordnet/

[14] en.wikipedia.org/wiki/**VerbNet**

[15] S. Brin, L. Page, "The anatomy of a large-scale hypertextual Web search engine", In *Computer Networks and ISDN Systems*, Vol. 30, No. 1--7. (1998), pp. 107-117.

[16] nlp.**stanford**.edu/software/**Stanford**

[17] http://www.searcheasy.htmlplanet.com/concl.html

[18] Y. Mehdad , B. Magnini., "Optimizing Textual Entailment Recognition using Particle Swarm Optimization", In Proc. of 2009 Workshop on Applied Textual Inference, ACL-IJCNLP 2009, pp. 36-43

[19] M. Kouylekov, B.Magnini, "Recognizing textual entailment with tree edit distance algorithm," In PASCAL Challenges on RTE, 2005, pp 17-20.

[20] A.D Haghighi, Y.Ng Andrew, D.M. Christopher, "Robust Textual Inference via Graph matching," In proc. of conf. on Human Language Technology and Empirical Methods in Natural Language Processing, 2005, pp. 384-397.

[21] M.Hu, B. Liu, "Mining Opinion features in customer reveiws," In proc. of AAAI conf. on Artificial Intelligence, California, 2004, pp. 755-760.

[22] http://en.wikipedia.org/wiki/Bhattacharyya_distance

[23] http://imdbpy.sourceforge.net/?page=main

[24] http://verbs.colorado.edu/~mpalmer/projects/verbnet.html

[25] V. Jijkoun, M de Rijk, " Recognizing textual entailment, is word similarity enough?" In LNCS, 2006, Vol. 3944/2006, pp. 449-460.

[26] M. Tatu , D. Moldvan " A semantic approach to recognize textual entailment" In proc. of Conf. on Humn Language Technology & Empiricl Methods in NLP, 2005, pp. 371-378.

AUTHORS PROFILE

Arti Arya has completed B.Sc (Mathematics) in 1994 and M.Sc(Mathematics) from Delhi University in 1996. She has completed her Doctorate in Philosphy in Computer Science Engineering from Maharishi Dayanand University, Rohtak, Haryana in 2008. Her areas of interest include spatial data mining, text mining, web mining, data management and knowledge based systems. She is a life member of CSI and member IEEE.

She is currently Prof. and HOD of MCA department, PES School of Engineering, Bangalore. She has approximately eleven years of teaching experience at undergraduate level which also includes around four years of research experience.

Vishwanath Yaligar has completed his BE from PESSE, VTU in 2010 and currently he is working as Consultant with a software company in Bangalore. His areas of interest are NP-Complete problems and Natural Language Processing. Other interests include Game Design, Speech Recognition and Synthesis, Movie Analysis. He is currently invovled in Software Development.

Ramya Prabhu has completed her BE (CSE) from PESSE, VTU in 2010. She is currently working with INFOSYS at Mysore, India.

Ramya L has completed her BE (CSE) from PESSE, VTU in 2010 and currently working as lecturer in PESSE , CSE department.

Rohith Acharaya has completed his BE (CSE) from PESSE, VTU in 2010 and currently working as Software tester in a software company in Bangalore.