An Intelligent IR for Natural Language Query Processing Using Fuzzy Logic

Ashish Tamrakar[1] Vishnu Mishra[2]
[1]Dr. C. V. Raman University, Bilaspur(C.G.)

Abstract— NLP (Natural language processing) is a technique for the mechanical analysis and illustration of human language. NLP is a IR (Information retrieval) systems which are shabby for discovery, within an immense text database, containing details desirable by a user. Natural language based IR system will be much constructive. Natural language processing based IR systems are extremely competent to indicate and manipulate the obscure query as complex and uncertain relationship presented among them. The conventional query in relational database management system is not capable of satisfying the needs for dealing with queries which are in natural language. In this paper we have discussed about a Intelligent IR's and some fuzzy based query system, we have proposed an intelligent IR system based on natural language processing which uses fuzzy logic for query processing for complex queries. We have implemented a software tool which can execute complex query on natural language and we can fetch the information from the database very effortlessly which uses NLDBR architecture.

Key words: IR, NLDBR, Fuzzy sets

I. INTRODUCTION

Relational Database management systems (RDBMS) are extensively used software products in numerous types of systems. As we all are familiar with natural language which is the main communication means for humans, but this causes it not easy to handle damaged information. Unsatisfactory information can be incoherent, inexact, unclear, uncertain or vague. However, the Complexity is limited in specific data processing and is not capable of directly expressing fuzzy concepts of natural language.

Earlier system work based on the architecture as fig. 1 in which user get the database information in his/her language i.e. in natural language, user query is passed to spell correction module then query string is divided into tokens, afterwards query mapping has been done and corresponding SQL (Structured Query Language) query will be generated.

Earlier, research work while querying data from relational databases frequently goes through one of two ways: the keyword-based approach and the structured query approach. Both the ways have their recompense and disadvantages. The structured query approach, while expressive and powerful, is not easy for naive users. The keyword-based approach is very friendly to use, but cannot express complex query intent accurately. In contrast, natural language has both advantages to a large extent: even naive users are able to express complex query intent in natural language [Fei Li et. al. 2014].

A. Natural Language Query Processing Using Semantic Grammar

Mrs. Gauri Rao proposed that provide search interface and reduced the part of user for recalling the tedious syntax of databases, this system provided the user to get the database

information in his/her language. The drawback of this system was that user has to fire queries in WH type question, it also did not supported single word query, ambiguous words care was taken while processing query. A limited data dictionary was used in which words were updated after regular period of time. All the name in the Natural language query had to be in double quotes (""). The system reduced the part of user but not to that extent that user required again he had to keep some of the things in mind that the name should in double quotes, query should consist of words that were present in the data dictionary, and the query had to be in WH question type and so on.

In our proposed work we have taken advantages of earlier research and to reduce the complexity of natural language user query we have adopted fuzzy set approach.

In next section we will go through different literatures of natural language query processing, section 3 we will discuss the problems of earlier research work, section 4 we will describe our proposed work, further in result and discussion section we will describe experimental evaluations, at last we will conclude our research work.
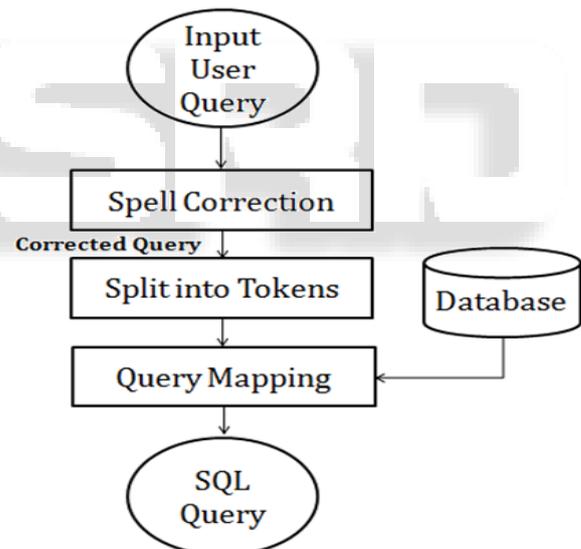

Fig. 1: Earlier System Architecture

II. LITERATURE SURVEY

We have gone through various literature of natural language query processing research work.

Fei Li et. al. presents NaLIR, a generic interactive natural language interface for querying relational databases. NaLIR can accept a logically complex English language sentence as query input. This query is first translated into a SQL query, which may include aggregation, nesting, and various types of joins, among other things, and then evaluated against an RDBMS. In this demonstration, we show that NaLIR, while far from being able to pass the Turing test, is perfectly usable in practice, and able to handle even quite complex queries in a variety of application domains. In addition, we also demonstrate how

carefully designed interactive communication can avoid misinterpretation with minimum user burden.

S.No.	Author/Paper title/Year	Name of Algorithm/Tool/Method	Description
1.	Mathias Soeken et. al. Automating the Translation of Assertions Using Natural Language Processing Techniques FDL Proceedings | ECSI 2014	High abstraction level and low abstraction level assertions.	Author presented an algorithm that automates the translation of natural language assertions into SystemVerilog Assertions using natural language processing techniques. Instead of manually translating each assertion separately.
2.	Ryuichiro Higashinaka et. al. Towards an open-domain conversational system fully based on natural language processing Proceedings of COLING 2014	Rule-based system.	This paper proposes an architecture for an open-domain conversational system and evaluates an implemented system. The proposed architecture is fully composed of modules based on natural language processing techniques.
3.	Anupriya1 et. al. Fuzzy Querying Based on Relational Database IOSR-JCE Jan-2014	Fuzzy Logic	This paper mainly discusses the realization of fuzzy query through fuzzy theory and SQL combined C#. Also, a real life application of fuzzy query based on relational database (the Patient Information database) is provided.
4.	Lei Zou et. al. Natural Language Question Answering over RDF — A Graph Data Driven Approach SIGMOD 2014	Graph Data Driven Approach	Author proposes a semantic query graph to model the query intention in the natural language question in a structural way, based on which, RDF Q/A is reduced to subgraph matching problem. More importantly, author resolve the ambiguity of natural language questions at the time when matches of query are found. The cost of disambiguation is saved if there are no matching found

Table 1:

III. PROBLEM IDENTIFICATION

The intricate and weak understood semantics of documents and user queries has prepared feedback and alteration important distinctiveness of any IR systems. Hence natural language based IR system will be much favorable. Natural language processing based IR systems are enormously capable to symbolize and manipulate the intricate query as complex and uncertain relationship presented among them. We have gone through various literatures in section 2 and found some bottleneck in natural language query processing which are as follows:

– Difficulty of translating user-specified query structure to the actual schema structure in the database.

- Spelling correction for mistakes made by the user while firing query map the natural language query into database query language.
- Semantic checking over user query.
- In earlier system user has to fire queries in WH type question.
- Earlier system did not supported single word query.
- A limited data dictionary was used in which words were updated after regular period of time.
- Natural language query had to be in double quotes ("").
- Earlier system reduced the part of user but not to that extent that user required again he had to keep some of the things in mind that the name should in double quotes, query should consist of words that were present in the data dictionary.

Question words are also called wh questions because they include the letters 'W' and 'H'. Some examples are as follows:

Question words	Meaning	Examples
who	person	Who's that? That's Nancy.
where	place	Where do you live? In Boston
why	reason	Why do you sleep early? Because I've got to get up early

Table 2:

IV. PROPOSED ARCHITECTURE

The conventional query in relational database management system is not capable of satisfying the needs for dealing with queries which are in natural language. After going through different literature survey we have concluded that the earlier systems were not that much vital to as per the user's prospect. It condensed the user part of recalling the syntax but not to that amount. The system only provided search interface for only partial database due to that user was obligatory to input queries according to that database.
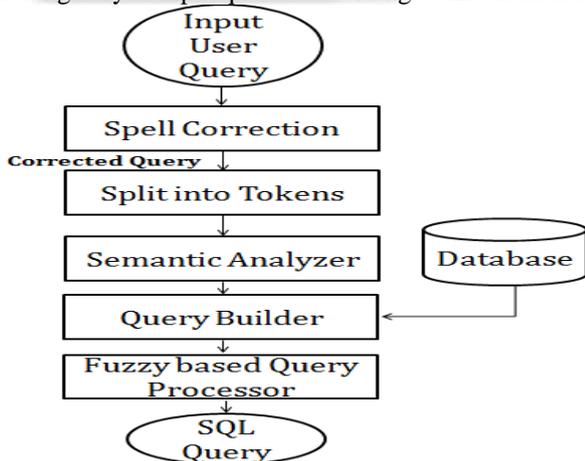

Fig. 2: Proposed System Architecture

We have proposed an intelligent IR system based on natural language processing which uses fuzzy logic for query processing for complex queries. We have implemented a software tool which can execute complex query on natural language and we can fetch the information from the database very effortlessly.

The main aspect of fuzzy logic is that it is able to deal with imprecise linguistic information which makes it attractive for automatic text summarization from the mathematical viewpoint; complex-valued fuzzy sets are natural. In classical (2-valued) logic, every statement is either true or false. In the computer, "true" is usually represented as 1, and "false" as 0. As a result, in the 2-valued logic, the set of possible truth values is a 2-element set f0; 1g.

The traditional 2-valued logic is well equipped to represent:

- Situations when we are completely sure that a given Statement is true.
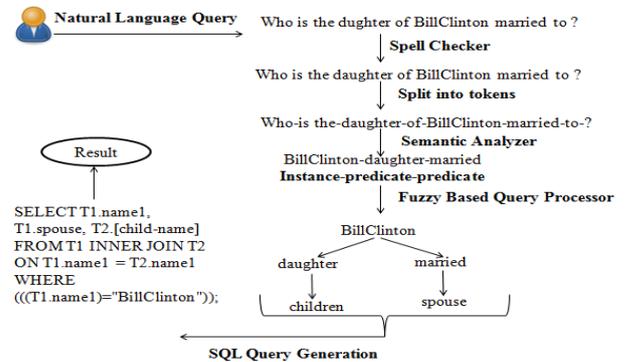- Situations when we are completely sure that a given Statement is false.

Fig. 3: NL Processing an example execution

V. RESULT AND DISCUSSION

For implementation of our project we have used java platform. We have overcome the some drawbacks of earlier system by applying fuzzy set based query processor.
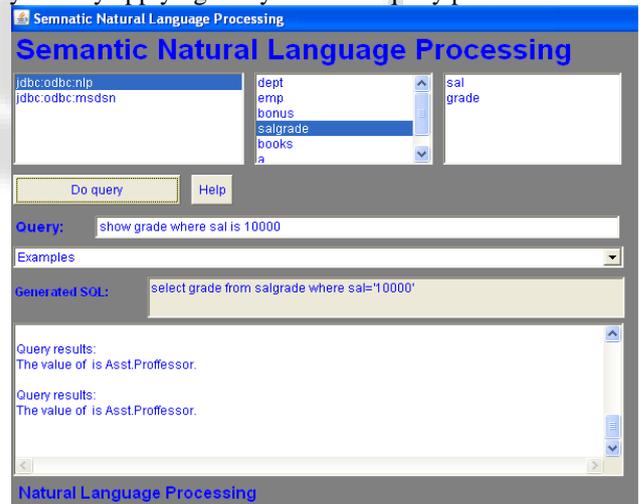

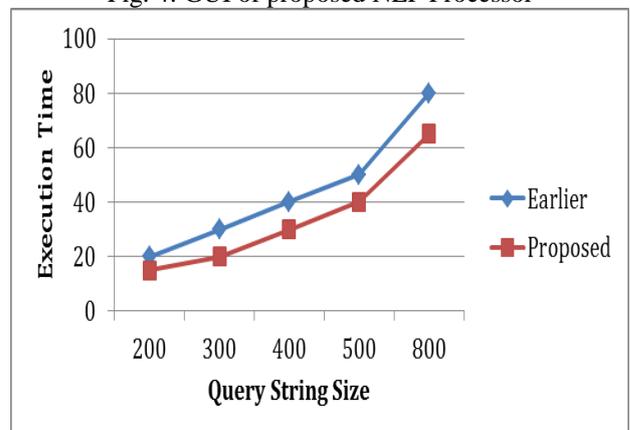Fig. 4: GUI of proposed NLP Processor

Fig. 4: Performance Comparison

VI. CONCLUSION

The conventional query in relational database management system is not capable of satisfying the needs for dealing with queries which are in natural language. In this paper we have discussed about some Intelligent IR's and some fuzzy based query system, we have proposed an intelligent IR system based on natural language processing which uses fuzzy logic for query processing for complex queries. We have implemented a software tool which can execute complex query on natural language and we can fetch the information from the database very effortlessly.

REFERENCES

[1] Mathias Soeken, Christopher B. Harris, Nabila Abdessaied, Ian G. Harris, Rolf Drechsler, Automating the Translation of Assertions Using Natural Language Processing Techniques FDL Proceedings | ECSI 2014.

[2] Ryuichiro Higashinaka1, Kenji Imamura1, Toyomi Meguro2, Chiaki Miyazaki Towards an open-domain conversational system fully based on natural language processing Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pages 928–939, Dublin, Ireland, August 23-29 2014.

[3] Anupriya, Prof. Rahul Rishi Fuzzy Querying Based on Relational Database OSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver. I (Jan. 2014), PP 53-59.

[4] Antonio Lieto, Andrea Minieri1, Alberto Piana1 and Daniele P. Radicioni A Knowledge-Based System for Prototypical Reasoning ISSN: 0954-0091 print/ISSN 1360-0494 oct 2014.

[5] Lei Zou, Ruizhe Huang, Haixun Wang Natural Language Question Answering over RDF — A Graph Data Driven Approach SIGMOD'14, June 22–27, 2014.

[6] Andre Freitas, Edward Curry Natural Language Queries over Heterogeneous Linked Data Graphs: A Distributional-Compositional Semantics Approach ACM 978-1-4503-2184-6/14 Feb 2014.

[7] Fei Li, H. V. Jagadish NaLIR: An Interactive Natural Language Interface for Querying Relational Databases SIGMOD'14, June 22–27, 2014.

[8] Akshay G. Satav, Archana B. Ausekar, Radhika M. Bihani, Mr Abid Shaikh A Proposed Natural Language Query Processing System WARSE Volume 3, No.2, March - April 2014.

[9] Huangi,Guiang Zangi, Phillip C-Y Sheu "A Natural Language database Interface based on probabilistic context free grammar", IEEE International workshop on Semantic Computing and Systems 2008.