

HINDI LANGUAGE AS A GRAPHICAL USER INTERFACE TO RELATIONAL DATABASE FOR TRANSPORT SYSTEM

Rupesh Nichante¹, Sagar Giripunje², Ankita Nikam³, Shivani Arsod⁴

Prof. Nutan Sonwane⁵

^{1,2,3,4}Student, CSE, Dr. Bababsaheb Ambedkar College of Engineering and Research, Maharashtra, India

⁵Assistant Professor, Department of Computer Science And Engineering, DBACER, Nagpur

Abstract - Database management systems is used for storing and retrieving the data. Databases are very hard to handle thus their interfaces is difficult in collaboration with users. Computers are used in accessing of information from database. Most of all the IT applications are accessing, storing, retrieving and analyzing the information from databases using the web as well as offline sources. Retrieving information from database needs knowledge and information of database languages like Structured Query Language (SQL) Everybody is not able to write SQL queries as they are not known to the syntax of SQL structure of the database. Thus, the idea of using natural language interface to provoke the development of new kind of processing method of Natural Language Interface to Database. Users do not need to learn any other formal language, they can execute query in their native language. Hindi language required this method to accept Hindi sentence as a query, process it and after execution make the result available to the user in the same language which is nothing but the Hindi Language Interface to DBMS.

Key Words: DBMS, HLIDBMS, SQL, NLI.

1.INTRODUCTION

Database Management System is a collection of interconnected data and set of programs to get those data. Database systems are designed to manage large bodies of information. To access this information, we must have the knowledge and information of Structured Query Language (SQL).

Internet provides the largest data in current date to users of all kinds. The vastness of data makes it compulsory that data is saved in an organized manner so that it is easy to search, retrieve and maintain. For this purpose the most logical and commonly used storage method is by the use of databases. In order to efficiently use or maintain any database the knowledge of languages such as SQL becomes essential. This would limit the use of data to only those users who

have the knowledge of these languages. Hence, an easy to use user interface comes into picture which would facilitate diverse users to access data. There is a requirement to design and develop an interface in the local or native language so that user can easily use that system without any knowledge of English as well as query language.

The need of user to retrieve the database in local language is fulfilled by using Natural language Interface to Database. With the help of this interface, the end user can query the system in natural language like English, Hindi, marathi, Gujrati, etc., and can see the result in same language. NLIDB system is proposed as a solution to the problem for accessing information in a easy way, allowing any type of users, mainly non technical people to retrieve information from a database using natural language (NL).

We are going to develop a system for people who know Hindi language. User can access database using Hindi language and get the result in the same language.

For example :-

उन सभी भेजनेवालो का नाम बताओ जिनका शहर नागपुर है. For our system we are going to use relational database for transport system.

2. Literature Review

There is enormous improvement in the zone of NLIDB. Researchers are taking a shot at it from numerous a long time. Researchers like Androustopoulos, G.D. Ritchie and P.Tanisch gives different designs for NLIDB which is given underneath. Designs as example coordinating pattern matching systems, syntax based

system, semantic grammar frameworks are created by the specialists that are clarified.

Natural Language Query Interface for Database, the framework that turns out is expected to help a beginner client, without the learning of database language like SQL language, to interact with the database altogether in native language. This particular framework that the review creates can have the accompanying favorable circumstances or significances for the expected client.

TITLE	TECHNOLOGY	ADVANTAGES	DISADVANTAGES
1. Natural Language Interface to Database- An Introduction (2016).	Syntax-Based System	Give the detail information about the structure of the sentences.	n/a
2. A Rule Based Graphical User Interface to Relational Database using NLP (2015).	Rule based system	Flexible	Works for the rules made for it
3. A Comprehensive Study of Natural Language Interface To Database (2014).	Pattern matching system	No parsing and module needed and system can be easily implemented.	n/a
4. Hindi Language Graphical User Interface to Relational Database using NLP (2014).	Natural language processing using Phases of compiler	Time complexity ,Reliable handling	n/a

3. Proposed System

3.1 Problem Statement

The Hindi language Graphical user interface is developed using Natural Language Processing (NLP). This interface is developed for the non technical users who are unaware about the query languages. The TRANSPORT database is used as a case study for building this interface.

The input for this system is the Hindi language sentence. This sentence is translated into English language by using ‘semantic matching technique’. The equivalent database query is build from that English sentence which on execution provides result to user. The output of this system is obtained in Hindi language.

In the system provides the facilities of accessing, inserting, updating and deleting the data from the Transport database. For the Hindi language interface, a Hindi based compiler which is generalized for any application is develop in Microsoft Visual Studio using C#.net. This compiler is a generalized, means it will be applicable to any database.

3.2 Methodology

Approach to accomplish the above goals some strategy has been taken after and it is given underneath:-

- 1]Make transport database which will store data about transport services.
- 2]Identify the way of queries i.e. select, update, delete, insert by queries and aggregation functions.
- 3]Appropriate mapping of tokens with database values be finished by extracting table, columns data from input Hindi sentences.
- 4]With the assistance of stored values of databases create SQL queries by mapping input queries.
- 5]Execute the query and give result in Hindi Language.

3.3 Architecture of the system

This system consists of two databases. First is the “Compiler Database” and second is the “Transport Database”. The column names of the Transport database will be stored in English language but the data in that columns will be stored in Hindi language. The user will enter the query in Hindi, this query will be processed and translated into its corresponding English query by using compiler database and depending upon that query result is calculated and provided to user. The result provided to user is also in Hindi language.

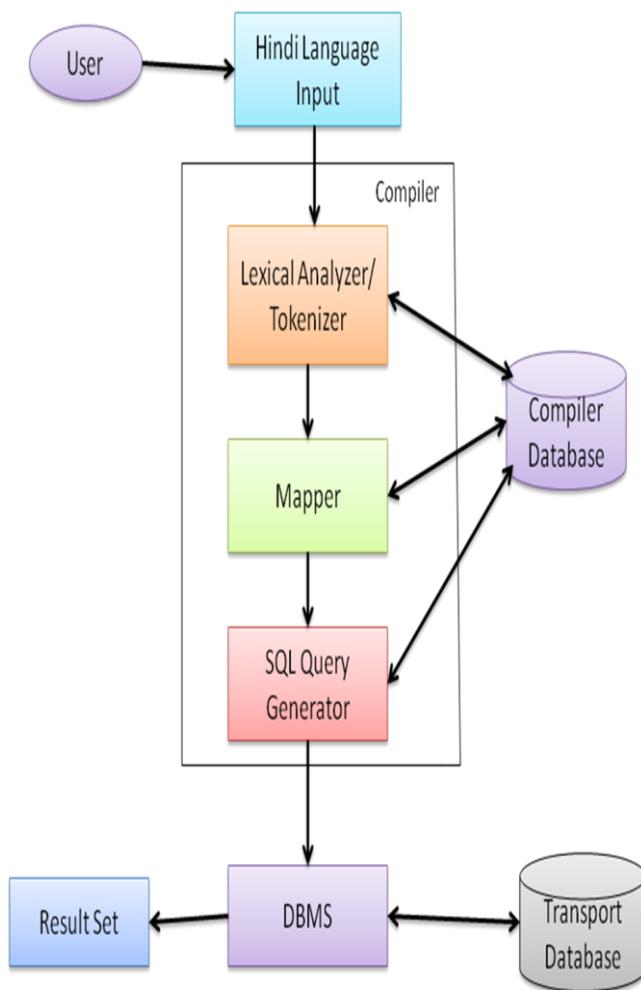


Figure : Block diagram of proposed system

There are some important phases i.e. Tokenizer, Mapper, SQL Query Generator and DBMS. while executing the query it is given in the HINDI language which is a hindi sentence. This hindi sentence in tokenize phase split into tokens and all tokens are separated by space gap and this tokens are stored into array. In lexical analysis, tokenization is the process of breaking a stream of input text into words, phrases and symbols called tokens. These Hindi tokens are stored in the lexicon(system dictionary)with their corresponding english words. The extracted tokens matched with the tokens stored in lexicon in sequence where mapping is performed. If these tokens matches the corresponding English words then it is saved with its type and all the useless tokens are discarded in this phase and useful tokens are stored. With column name, table name the SQL query is to be formulated. The SQL query is generated. The generated query will be fired to the database and Hindi language is displayed to user on the interface. The query was given in the Hindi language and we get the output in the form of Hindi language.

Some HDLIDBs Words :-

Token Word	Mapped Word
बताओ	Select
करो	Update
हटाओ	Delete
सामान	=
समान नहीं	!=
ज्यादा	>
कम	<

Token Word	Mapped Word
जिसका	Where
जिनका	Where
जिनकी	Where
जिसकी	Where
जिसके	Where
जिनके	Where
जो	Where
जोकि	Where

3. CONCLUSION

This concept is on Hindi Language Interface to Relational Database utilizing NLP acknowledges the question in Hindi Language and gives yield in the same language. With the assistance of tokenizer, mapper, question generator, DBMS and etymological segment it is capable to play out the suitable operation and gives yield in Hindi language which is anything but difficult to peruse and comprehend for the general population who don't have any learning of SQL language or other inquiry language. It is particularly helpful for non-specialized individual to recover information from database and get learning from it. With the assistance of this Hindi Language Interface to Social Database utilizing NLP we can play out all the operation, for example, select, refresh, erase, make and furthermore the request by inquiries, for example, rising request and dropping request and total capacities.

REFERENCES

[1] Abhijeet R. Sontakke, Prof. Amit Pimpalkar "Hindi Language Graphical User Interface to Relational Database using NLP" *International Journal of Advanced Research in Computer*

Engineering & Technology (IJARCET) Volume 3 Issue 10, October 2014.

- [2] MOHIT DUA, SANDEEP KUMAR, ZORAWAR SINGH VIRAK, "HINDI LANGUAGE GRAPHICAL USER INTERFACE TO DATABASE MANAGEMENT SYSTEM" 12 TH INTERNATIONAL CONFERENCE OF MACHINE LEARNING AND APPLICATION 2013.
- [3] Mrs. Neelu Nihalani. Dr. Sanjau Silkari and Dr. Mahesh Motawani, "Natural Language interface for Database: A brief Review", IJCSI International Journal of Computer Science Issues, vol. 8,no. 2, pp. 600-608, March 2011.
- [4] Faraj A. El-Mouadib, Zakaria S. Zubi, Ahmed A. Almagrous and Irdess S. El-feghi, "Generic interactive natural Language Interface to Databases(GINLIDB)", International general of Computers, vol. 3, no. 3, 2009.
- [5] M. JOSHI, R. A. AKERKAR, "ALGORITHM TO IMPROVE PERFORMANCE OF NATURAL LANGUAGE INTERFACE", INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND APPLICATIONS, VOL. 5, NO. 2, PP. 52-68,2008.
- [6] Androutsopoulos, G.D. Ritchie, and P. Thanisch, Natural Language Interfaces to Databases - AnIntroduction, Journal of Natural Language Engineering 1 Part 1 (1995), 29-81.
- [7] D.L. Waltz, "An English Language Question Answering System for a large Relational Database", Communications of the ACM, pp. 526-539, 1978.

- [8] G. Hendrix, E. Sacerdoti, D. Sagalowicz, and J. Slocum, "Developing a Natural Language Interface to Complex Data", ACM Transaction on Database system", communications of ACM, pp. 526-539,1878.
- [9] W.Woods, "An experimental parsing system for transition net grammar in NLP", Algorithmic press, New York, USA, 1973.
- [10] L.R. Harris, "Experience with Intellect: Artificial Intelligence Technology Transfers", the AI magazine, pp. 43- 50, 1984.
- [11] "START natural Language Query Answering System"[online].Available:
<https://start.csail.mit.edu/>