International Conference On Emanations in Modern Technology and Engineering (ICEMTE-2017) ISSN: 2321-8169
Volume: 5 Issue: 3 305 - 308

Use of Natural language processing for accessing the Graphical User Interface and Relational Database

Sampada Watane
Department of Computer Engineering
Shree L R Tiwari College of Engineering, Mira road

Riteshri Nagmote
Dr. Ambedkar Institute of Management and Research, Nagpur

Mangesh Watane
Regional Sales Manager
Cabot Sanamr

Abstract - Data is info which might be translated into a lot of convenient type to maneuvers or method. Information has been keep within the info and major sources of knowledge is info. In way of life info plays terribly very important role. Within the growing use of pc and web, info technology has major impact thereon. The employment of direction system is for accessing, storing and retrieving information. All cannot simply perceive info system. For those people that not having any information of info Language they will realize it tough to access info. Therefore, there's have to be compelled to decide one simple technique on ways for accessing the info by victimization tongue process. Therefore, this concept of victimization tongue rather than SQL triggered the event of a unique sort of process technique known as tongue Interface to info (NLIDB). Wherever user doesn't have needing to find out the formal language, they'll offer question in their language. For the people that are comfy with the Hindi language want this application to just accept Hindi sentence as a question, method it and when execution offer result to the user within the same language that is nothing however the Hindi Language Interface to direction System.

Index Terms - *DBMS, HLIDBMS, NLP, NLIDB, SQL.*

*****

I. INTRODUCTION

The requirement of facts and statistics may be very important a part of existence. there are numerous assets of statistics however the most important one is the databases. Database allows us to store, access and retrieve statistics. Any enterprise or enterprise is possible without the usage of database. every n each laptop programs are dependable on database to get entry to the information. For that it's miles vital to have know-how of formal question language like square but it is very tough for anybody to examine and write square queries. to overcome this problem many researchers have delivered out to apply herbal Language (NL) i.e. English, Hindi, Marathi, Bengali, Arabic etc. in vicinity of formal question language which can be best interface among an application of pc and non-technical consumer.

1.1 NATURAL LANGUAGE PROCESSING (NLP):

Natural language processing (NLP) is a subject of laptop technology, synthetic intelligence, and computational linguistics involved with the interactions among computer systems and human (natural) languages. As such, NLP is associated with the area of human–computer interaction. Many demanding situations in NLP contain natural language expertise, that is permitting computers to derive that means from human or natural language input, and others contain natural language generation. In idea, NLP is a very appealing method of human computer interaction. Natural Language expertise is on occasion called an AI-entire hassle because it seems to require an intensive understanding approximately the out of doors global and the potential to govern it. NLP has appreciably overlapped with the sphere of computational linguistics, and is often considered a sub-subject of artificial intelligence. the muse of NLP lies in some of disciplines' like computer and statistics sciences, linguistics, mathematics, electrical and electric engineering, artificial intelligence and robotics, psychological, agriculture, weather forecasting and many others. [1]. Applications of NLP consist of some of fields of research, together with device translation, natural language interface to database, herbal language text processing and summarization, person interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, AI and professional machine, and so forth. [2].

1.2 NATURAL LANGUAGE INTERFACE TO DATABASE (NLIDB)

Someone without an understanding of database language may find it tough to get admission to database effortlessly. consequently, square tutor become evolved for analyzing the capability of Natural Language Processing to expand merchandise for human beings to engage with database in simple English. these, products have created a revolution in extracting statistics from databases. they have got discarded the fuss of gaining knowledge of sq. and time is likewise stored in mastering in query language.

II. RELATED WORK

Development work in NLIDB has begun in early Nineteen Seventies. on the grounds that then many systems have been evolved. Early systems have many flaws then some structures

had been developed to conquer those flaws. a number of the evolved NLIDB structures are mentioned below.
Following are a few evolved NLIDB systems are given-

2.1 LUNAR SYSTEM

W. Woods etal [5] has given records approximately LUNAR system that solutions question on samples of rocks delivered returned from the moon. The meaning of device's name is that is with regards to the moon. To accomplish its function, the LUNAR machine uses databases; one for the chemical evaluation and the opposite for literature references. The LUNAR device makes use of an Augmented Transition network (ATN) parser and Woods‟ Procedural Semantics. W. Woods [6] have also given the examine of the LUNAR gadget overall performance which became pretty stunning; it controlled to address seventy-eight% of requests without any mistakes and this ratio rose to ninety% when dictionary mistakes had been corrected. but those figures can be misleading due to the fact the system become no longer difficulty to intensive use because of the hassle of its linguistic talents.

2.2 LADDER

It became designed as a NLIDB of information about US military ships. in keeping with G. Hendrix etal [7], the LADDER system makes use of semantic grammar to parse inquiries to question a allotted database. even though semantic grammars helped to put into effect structures with spectacular traits, the resulting structures proved hard to port to unique utility domain names. indeed, a unique grammar had to be developed each time LADDER changed into configured for a new utility [5]. The system uses semantic grammars approach that interleaves syntactic and semantic processing. The query answering is executed through parsing the input and mapping the parse tree to a database query. The device LADDER turned into applied in LISP. on the time of creation of the LADDER device, it was capable of process a database that is equal to a relational database with 14 tables and 100 attributes

2.3 RENDEZVOUS

Inside the device advanced and studied with the aid of E. Codd [8] users should get right of entry to databases via extraordinarily unrestricted herbal language. in this Code's gadget, special emphasis is located on query paraphrasing and in attractive users in rationalization dialogs while there is issue in parsing user input.

2.4 PLANES D.

D. Waltz stated [9] the Programmed Language-based Enquiry device (PLANES) at the college of Illinois Coordinated technological know-how Laboratory. PLANES consist of an English language the front give up with the ability to understand and explicitly answer consumer requests. It includes out clarifying dialogues with the consumer as properly as solution indistinct or poorly described questions. This painting is being executed the usage of database primarily based upon records of the U.S. military Three-M maintenance and fabric management, it's miles a database of plane preservation and flight statistics, although the thoughts can be directly implemented to different non-hierarchic record based databases.

2.5 PHILIQA

This was referred to as Philips question Answering gadget (PHILIQA) defined via R. Scha [10], makes use of a syntactic parser which runs as a separate skip from the semantic know-how passes. This device is mainly concerned with issues of semantics and has three separate layers of semantic knowledge. The layers are known as "English Formal Language", "international version Language", and "data Base Language" and seem to correspond more or less to the "external", "conceptual", and "internal" views of information.

2.6 CHAT-80

CHAT-80 changed into carried out entirely in Prolog and it is the fine NLIDBs machine. It converted English questions into Prolog expressions, which have been evaluated towards the Prolog database. The code of CHAT-80 became circulated extensively and formed the basis of several other experimental NLIDBs. The database of CHAT-80 includes data (i.e. oceans, main seas, foremost rivers and main cities) about one hundred fifty of the international locations international and a small set of English language vocabulary which can be sufficient for querying the database [24].

2.7 TEAM

B. J. Gross has given a paper on group (portable herbal Language Interface system). A massive a part of the research of that time was dedicated to portability issues. crew changed into designed to be without difficulty configurable by way of database administrators without understanding of NLIDBs [11,12].

2.8 ASK

Allowed quit-users to train the gadget new phrases and ideas at any point all through the interplay. ASK was definitely an entire records management gadget, supplying its personal integrated database and the ability to interact with more than one external databases, e-mail applications and different computer packages. all of the applications connected to ask have been accessible to the cease-user through herbal language requests. The consumer stated his/her requests in English and Ask transparently generated suitable requests to the best underlying systems.

2.9 JANUS

P. Resnik studied [13] that it had similar skills to interface to a couple of underlying structures (databases, professional structures, snap shots devices, and many others). all of the underlying systems could participate inside the evaluation of a Natural language request, without the consumer ever becoming aware about the heterogeneity of the general system. JANUS is also one of the few structures to support temporal questions.

2.10 EUFID

M.Templeton etal [14] has given that the EUFID gadget includes 3 essential modules, no longer counting the DBMS.

First is analyser module, 2nd is mapped module and third is translator module.

2.11 DATALOG

it's far an English database question device primarily based on Cascaded ATN grammar. with the aid of offering separate illustration schemes for linguistic know-how, widespread [14] international know-how, and alertness domain information, DATALOG achieves a excessive degree of portability and extendibility [15]. systems that also appeared in mid-Eighties were LDC [16], TQA [17], TELI [18] and lots of others.

2.12 SQL TUTOR

SQL may be very hard for beginner users to recognize. The sq.-teach application tutors students by means of helping the scholars thru some of database questions from 4 distinct databases. A scholar version is stored for each student primarily based on query constraints (every constraint represents a part of the query that is necessary to answer the question). Each time a particular query constraint is used, square-show records whether or not it became used successfully or unsuccessfully. in this way a model of a scholar's strengths and weaknesses is generated and sq.-tutor can pick questions which re-put into effect problem areas or introduce new query ideas [22].

3.1. PROBLEM STATEMENT

Hindi language interface to relational database is absolutely primarily based at the regulations through which we are going to carry out the operations like choose, insert, update, delete. We also are running to provide the boost question operation including functionality of combination capabilities such as MIN (), MAX (), SUM () and AVG (). The user will type the question in Hindi language and that natural language has been processed and could give the output in Hindi language simplest. Time distinction has been calculated, machine will deliver translation time and execution time in milliseconds in addition to in nanoseconds.

3.2. METHODOLOGY

To acquire the above goal technique used is given as- we are going to use the rule of thumb primarily based gadget so that it will observe and execute each and every query as consistent with the rules made for it. First it'll identify the nature of the question i.e. pick, update, delete, create, insert and additionally it'll pick out that the question is with aggregation functions or no longer. we're using the relational database so it's far very much bendy we are able to without problems save all Hindi in addition to English values in it and also we will effortlessly retrieve it. Randomize automatic document generation method is also there so that we will effortlessly generate maximum variety of information in very much less time. suitable mapping of tokens with database values ought to be executed by way of extracting desk, columns facts from enter Hindi sentences. With the assist of saved values of databases generate square question by way of mapping input query. finally, we are able to execute the Hindi question and additionally get the output in Hindi language itself.

3.3 IMPLEMENTATION & STRUCTURE OF THE SYSTEM

Architecture of Hindi language interface to relational database using NLP is given and explained below from fig1. This architecture is known as HLIDBMS i.e. Hindi Language Interface to Database control gadget. There are critical phases i.e. Tokenizer, question kind rule, question table rule, basic queries and its sub rules, query generator engine DBMS & database server. In tokenize section Hindi sentence is cut up into tokens. this is completed with fact that everyone the tokens are separated via an area hole from each different. all of the tokens which we get in this segment are stored in an array. Tokens are phrases of Hindi language. Token can be a table name, column name, situation, any value, command name, operation call or any non-beneficial word.

सभी विद्यार्थी के का नाम ,अंक बताओ . This Hindi sentence has 7 tokens. First token is सभी that's the beginning of sentence. Now सभी approach it is reflecting like select all i.e. in sq. we are saying 'pick *', any other token is विद्यार्थी के it's far reflecting the call of the database table i.e. 'pupil desk' a few tokens can be fields name as in the above query नाम and अंक are the field names. there is conjunctions additionally like का as well as we also included the comas (,) inside the listing of

tokens & ultimately closing issue is बताओ that is reflecting as the 'pick question' consequently after this step we have all the tokens from which the sentence consists of. After that we are able to follow the question type rule. Question type rule is a rule on the way to discover which sort of query it's far whether it's far choose, insert, replace, delete type of question. we are given with the query properties through which we can effortlessly identify the related Hindi word that's given within the sentence within a Question.

Later it will become aware of the desk call with the assist of question table policies. it will simply see whether the given desk is gift there or now not. those each the things had been feasible because of the tokenizer and its tokens which we're matching beneath each rule. as soon as the query policies and table policies has been applied then we can proceed with the further tokens and we will apply the sub policies of the selected query. If the query in Hindi could be the choose question then it will search for the policies like column policies, mixture feature rule, in which clause and wherein circumstance rule. it'll work with the assist of tokens simplest. It is like column regulations it's going to choose the variety of columns given within the Hindi query. combination characteristic will perceive whether or not it is min (), max (), sum (), avg () query or not. other policies like wherein clause for that we are given with the homes i.e. it will become aware of all of the related Hindi English phrases.

CONCLUSION

Rule based graphical user interface to relational database is provided on this paper. The system will accept Hindi sentence as a question and offers output in Hindi itself. it's far very a whole lot beneficial for the individuals who do not have any previous knowledge of database and square queries languages. we're the use of exclusive rule together with the NLP to

perform operation together with insert, replace, delete, pick out as well as the combination capabilities consisting of min (), max (), sum (), avg () and so forth. This system may be greater by using making it extra regular. We can also put into effect it for very complicated queries like be part of operations & order through operations (queries). To make the gadget extra pleasant the dialogue based device can be used wherein consumer will offer the enter Hindi question thru speech
interface. In addition wherein condition is likewise there it's going to work like identical as given above it's far inclusive of all of the conditional part and its related Hindi phrases together with <,>,=,logical and ,or not and so on. further for update query it is having update column rule, in which clause rule and where condition rule and its running is same as explained above. The same manner insert and delete additionally paintings. At final there may be question generator so that you can generate query from Hindi sentence. that query generated will be fired to database and all the selected records decided on rows has been displayed in Hindi Language. square is generated in this segment in line with Hindi sentence. Execute question and display result to consumer the above square question is performed and end result of which in Hindi language is displayed to consumer. The output is in the form of Hindi language and we're giving question also in Hindi language and processing of all this has been finished with the aid of inner module as explained above. As soon as the question has been performed and the end result has been proven aspect by means of aspect it will additionally display the timing result which encompass whether or not the query has been correctly performed or now not if it is failed it will display . It'll additionally supply the interpretation time in milliseconds as nicely as nanoseconds to notice the minute difference throughout conversion and same inside the case of execution time also, it's going to display the time required to execute the
query.

REFERENCES

[1] Akshar Bharthi, Y. Krishma Bhargava and Rajeev Sangal, "References and Ellipsis in an Indian Languages Interface to Databases", Journal of Computer Science of India, VOL 23, no. 3, pp 60-82, Sep. 1993.

[2] Gobinda G. Chowdhury, "Natural Language Processing", Annual Review of Information and Sciences Technology, VOL 37, no. 1, pp 51-89 ,2003.

[3] Amandeep Kaur, "Punjabi Language Interface to Database", ME Thesis, Thapar University, Jun 2010.

[4] H. R. Tennant, K. M. Ross, M. Saenz, C.W. Thompson and J.R. Miller, "Menu Based Natural Language Understanding", Proceeding of the 21st Annual Meeting of ACL, Cambridge, Massachusetts, pp. 151-158, 1993.

[5] W. Woods, R. Kaplan and B. Webber, B., "The Lunar Sciences Natural Language Information System", Final Report. B. B. N. Report No 2378, USA, 1972.

[6] W. Woods, "An experimental parsing system for transition network grammars. In Natural Language Processing", Algorithmic Press, New York, USA, 1973.

[7] G. Hendrix, E. Sacrdoti, D. Sagalowicz, and J. Slocum, "Developing a natural language interface to complex data", ACM Transactions on Database Systems, Volume 3, No. 2, pp. 105 – 147, USA, 1978.

[8] E.F. Codd, "Seven steps to rendezvous with the casual user", In IFIP Working Conference Data Base Management, pp 179–200, 1974.

[9] D.L. Waltz., "An English Language Question Answering System for a Large Relational Database", Communications of the ACM, pp 526–539, 21(7): July 1978.

[10] R.J.H. Scha., "Philips Question Answering System PHILIQA1", In SIGART Newsletter, no.61. ACM, New York, February 1977.

[11] B.J. Grosz, "TEAM: A Transportable Natural-Language Interface System", In Proceedings of the 1st Conference on Applied Natural Language Processing, Santa Monica, California, pp 39–45, 1983.

[12] B.J. Grosz, D.E. Appelt, P.A. Martin, and F.C.N. Pereira, "TEAM: An Experiment in the Design of Transportable Natural-Language Interfaces", Artificial Intelligence, pp 173–243, 32: (1987).

[13] [P. Resnik, "Access to Multiple Underlying Systems in JANUS", BBN report 7142, Bolt Beranek and Newman Inc., Cambridge, Massachusetts, September 1989.

[14] M. Templeton and J. Burger, "Problems in Natural Language Interface to DBMS with Examples from EUFID", In Proceedings of the 1st Conference on Applied Natural Language Processing, Santa Monica, California, pp 3–16, 1983.

[15] C.D. Hafner, "Interaction of Knowledge Sources in a Portable Natural Language Interface", In Proceedings of the 22nd Annual Meeting of ACL, Stanford, California, pp 57–60, 1984.

[16] B.W. Ballard, J.C. Lusth, and N.L. Tinkham, "LDC-1: A Transportable, Knowledge based Natural Language Processor for Office Environments", ACM Transactions on Office Information Systems, pp 1–25, January 1984. [17] F. Damerau, "Operating statistics for the transformational question answering system", American Journal of Computational Linguistics, pp 30–42, 7: 1981.

[17] B. Ballard and D. Stumberger, "Semantic Acquisition in TELI", In Proceedings of the 24th Annual Meeting of ACL, New York, pp 20–29, 1986. [19] B. Sujata, S. Viswanadha Raju and Humera Shaziya, "A Servey of Natural Language Interface to Database Management System" International Journal of Science and Advance Technology", vol.2, no. 6,june 2012 .

[18] Mohit dua, Sandeep Kumar, Zorawar Singh Virak, "Hindi Language Graphical User interface to Database Management System" 12th International Conference of Machine Learning and Application 2013.

[19] Mohit dua, Sandeep Kumar, Zorawar Singh Virak, "Hindi Language Graphical User interface to Database Management System" 12th International Conference of Machine Learning and Application 2013.

[20] Seymour Knowles, "A Natural Language Database Interface For SQL-Tutor", 1999.

[21] Abhijeet R. Sontakke, Prof. Amit Pimpalkar "A Review Paper on 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 .

[22] M. E. Saleh, "Semantic Based Query in Relational Database Using Ontology", Canadian Journal on Data, Information and Knowledge Engineering, vol.2, 2011.

[23] Seymour Knowles, "A Natural Language Database Interface For SQL-Tutor", 1999.

[24] Abhijeet R. Sontakke, Prof. Amit Pimpalkar "A Review Paper on 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 .

[25] M. E. Saleh, "Semantic Based Query in Relational Database Using Ontology", Canadian Journal on Data, Information and Knowledge Engineering, vol.2, 2011.