

QUESTION ANSWERING SYSTEM USING ONTOLOGY IN MARATHI LANGUAGE

Sharvari S. Govilkar¹ and J. W. Bakal²

¹Department of Computer Engineering, PCE, Mumbai, India

²Department of Computer Engineering, SJCOE, Mumbai, India

ABSTRACT

Humans are always in a quest to extract information related to some topic or entity. Question answering system helps user to find the precise answer of the question articulated in natural language. Question answering system provides explicit, concise and accurate answer to user questions rather than providing set of relevant documents or web pages as answers as most of the information retrieval system does. The paper proposes question answering system for Marathi natural language by using concept of ontology as a formal representation of knowledge base for extracting answers. Ontology is used to express domain specific knowledge about semantic relations and restrictions in the given domains. The ontologies are developed with the help of domain experts and the query is analyzed both syntactically and semantically. The results obtained here are accurate enough to satisfy the query raised by the user. The level of accuracy is enhanced since the query is analyzed semantically.

KEYWORDS

Question answering system (QAS), Ontology, Marathi Natural language QA system (NLQA), Natural language processing (NLP)

1. INTRODUCTION

With the rapid growth of the amount of online and electronic documents in Indian regional language, the keyword based approaches lack many important elements to enable QA driven process. So a system is required which can provide user with accurate answers for their queries. Question answering system provides user with functionality where they can ask questions in natural language and the system returns answer which is most accurate and precise of all the possible answers for the given input question. Question answering supports user with providing option to ask natural language query rather than traditional structured queries. A question answering system provides more accurate result when ontology is used for representation of knowledge. Ontology is a form of conceptual representation of information where relation existing between different entity and details about a particular entity is provided.

Any question answering system basically consists of three parts as question processing, answer retrieval and answer generation. In question processing users natural language question are parsed to formulate question in machine readable form using different approaches. Then in answer retrieval candidate answers are extracted based on intermediate representation of question. Finally in answer generation phase user understandable precise and accurate answer is generated and provided to user.

QA systems are classified into two main types as close domain QAs and open domain QAs. In close domain QAs, scope of user question is limited to a particular domain like sports, medicine, entertainment, history and others. An open domain QAs mostly works like search engines where scope for question is global.

Question in any question answering system can be of varying types. Question can be factoid question for which answers are simple fact about the entity in question. Some question can be of descriptive type where one needs to full detail about person, place or any event. There can be simple yes/no type of question where answers are as yes or no. A question can also be an instruction based question where answers are provided as an instruction to accomplish any task. Question in QAs can be of many other forms which provide precise answer in the same format as that of question provided.

There is very less work has been reported for creating QA system for natural languages like Hindi, Marathi etc and specifically there are no such systems available where ontology itself is represented in Marathi. Most of the QA system converts the Indian regional language data to English and the answers are extracted which many times lead to loss of morphological rich contents of Hindi or Marathi. In recent past, information extraction was based on keyword matching, but it has main drawback of semantic matching. To achieve semantic matching, ontology's with it's onto triples appeared to be efficient method. Ontology's can be general or domain specific and can be created automatically or manually.

As ontology has become trending topic now days, there are sufficient tools and information available to build a question answering system using ontology in English but hardly any ontology is created where data itself is represented in Hindi or Marathi.

The aim of this paper is to design, implement and experiment a new Marathi language QA framework based on ontologies where answers to the user's questions are provided by using predefined domain specific ontology. The overall objective is to provide user with semantically correct and accurate answer for their queries in Marathi language.

In section 2, related work and motivation is discussed in detail. Proposed system is described in section 3. Working of system is mentioned in detail in section 4. Section 5 explores performance analysis of QAs system. Finally, paper is concluded in section 6.

2. RELATED WORKS

QA systems are designed to address the problems of traditional search engines and meet the growing requirements of users searching the large amounts of information available on the web. In fact, these systems are faced with a double challenge: first processing and understanding a question in natural language and second identifying and extracting the correct answer from a set of documents also in natural language.

Sahu, Shriya, N. Vashnik, and Devshri [1] Roy have presented an approach to extract answers from Hindi text for a given question where the text is expressed in the form of query logic language and then relevant answer is extracted for the given question. The focus of the system has been basically on four kind of questions type such as: What, Where, How many, and what time. The type of question and keywords were extracted by using shallow parser, but no semantic relations are consider while extracting information. There approach uses the traditional methods i.e. to take words as independent words during matching and just check the existence of the query keywords in the stored data and no relations constraints between words in a phrase or neighborhood are extracted which leads to less accuracy.

Hindi Question Answering System is created by Stalin, Shalini, Rajeev Pandey, and Raju Barskar [2]. The system is based on searching in context by using similarity heuristic and utilizes syntactic and partial semantic information. Domain-specific and question specific entities are found out after removing the stop words and also longest phrase are extracted while processing query. Here database is used to send candidate answers collection, based on keyword present in

the question, to next answer extraction module which extract candidate answers from the retrieved documents. Building of limited words synonyms lexicon reduces the accuracy of system due to mismatch of unavailable entities.

Using locality based similarity heuristics Kumar, Praveen, et al.[3] have created Hindi search engine. It provides facility to extract correlated contents from set of e-learning contents. The architecture consists of an entity generator which generates specific domain entities. Such generated entities where corresponding to the questions of which users wanted to retrieve answers. Questions provided by the users where then classified for selection of appropriate answers. From the query stop words are removed and relevant keywords where extracted. Query was enriched with synonyms of keywords. Finally the query is passed to retrieval engine, which on basis of locality returns top passages after ranking.

To process question provided in Hindi language and retrieve answers for those question, Sharma, Lokesh Kumar, and Namita Mittal[4] have used Named Entity based n-gram approach for their question answering system. For retrieval of answers first question classified and analyzed to generate a proper query. Question classification helps to identify relevant type of answers. Then by using similarity metric relevant document is retrieved which probably contains the answer and at last by using the bigram and NER relevant answers are retrieved for the given question. Overall higher accuracy was obtained by using the bigram approach but accuracy dropped in scenario where synonyms present in document where not matched due to the use of syntactical approach.

A dialogue based question answering system which provides answers related to railway domain in Telugu language is proposed by R. Reddy, N. Reddy and S. Bandyopadhyay [5]. Question answering process is based on keyword approach where input query are tokenized and keyword are extracted using knowledge base related to railways. Tokens generally consist of train names, station names whereas keywords specify when, in, out, go and others present in the query text. Query frame is extracted by matching it with predefined procedures to generate relevant SQL query. Dialog manager task is to interact with users if more information is needed to execute SQL query to fetch relevant answer to user question.

Question answering system to produce answer of question in Punjabi and English is proposed by V. Gupta [6]. The system accept query in English or Punjabi language of which stop word is eliminated initially. Then from the query string key terms like noun, adjective, verbs or adverb are extracted. Using dictionary of Punjabi and English language synonyms of key terms is extracted. Finally query is reformulated using the extracted keywords and its synonyms. By using reformulated query various matching web pages are retrieved using a search engine. Extracted documents are summarized based on proximity of key term found in documents and finally candidate answer is provided as per its rank.

An algorithm for Punjabi question answering system is proposed by P. Gupta and V. Gupta [7]. The system provides a better approach for finding patterns and matching to extract accurate precise answers from set of possible answers. The proposed algorithm works for ਕੀ (what), ਕਦੋਂ (when), ਕਿੱਥੇ (where), ਕੌਣ (who) and ਕੌਣ (why) form of questions where first question word is extracted from question then as per different procedure create for each question type corresponding question keywords are extracted and through final answers are retrieved. The overall accuracy of system is 73 % where 4850 question where asked for over 50 documents of Punjabi language

Keyword based question answering system is developed by J. Cherapanamjeri, L. Lingareddy, Himabindu. K, [8] which provides answer to question related crop statistics in Telugu. All the key words in the user query are mapped with database and if the keyword matches then appropriate SQL queries are generated which fetch answer from the database. First the input query is

converted into WX notation and the tokenized. All the tokens are searched in knowledge base and if token is found in KB then corresponding key value pair is stored in memory which aids in development of natural language query to be provided to user. If user acknowledges the query then its corresponding SQL query is generated using the query frame and fired on database to fetch answer which is finally converted to natural language text by using predefined templates.

Chaware, S., and S. Rao [9] has discussed a system where Semantic matching is performed using ontology for Hindi and Marathi languages to infer the information from knowledge base. Knowledge is represented using ontology. The data and ontology are maintained in English for easy building and traversing, the query terms from a query matches with ontology terms semantically by using synsets for each language. Finally, ontology terms are extracted to represent knowledge as an answer for the query. The approach converts local language to English using bilingual dictionary where there is more chance of translating mismatch and losing of morphological rich words and phrases of Hindi and Marathi language, which may lead to mismatched query keywords.

Tahri, Adel, and Okba Tibermacine [10] have proposed a new architecture to develop a factoid question answering system based on the DBPedia ontology and the DBPedia extraction framework. Their system SELNI is a sentence level question answering system that integrates natural language processing, ontologies, machine learning and information retrieval techniques. Three steps are followed to build this system as the comprehension of the question, detection of its answer type, Question Processing, resources and keywords extraction to build SPARQL query and execute it by interrogating the DBPedia ontology. The result of the query is the answer of the given question. SELNI system offers encouraging results while comparing to other question answering systems.

Wang, Chong, et al [11] has created a Portable natural language interface to Ontologies, name as PANTO which accepts generic natural language queries and outputs SPARQL queries. Based on a special consideration on nominal phrases, it adopts a triple-based data model to interpret the parse trees output through parser. They have used Stanford Parser and multiple existing techniques and tools are integrated to interpret parse trees of natural language queries into SPARQL. To understand sense of the words in the NL queries and WordNet and string metrics algorithms are also integrated.

A prototype system is developed by Lopez, Vanessa, Michele Pasin, and Enrico Motta named AquaLog [12] which is a portable question-answering system which takes queries expressed in natural language and ontology as input and returns answers drawn from the available semantic markup. AquaLog uses GATE NLP platform, string metrics algorithms, WordNet and novel ontology-based similarity services for relations and classes to make sense of user queries with respect to the target knowledge base.

Architecture for ontology based natural language question is proposed by Raj, P. C. [13] where concept of semantics and ontology is used to facilitate better query construction and extraction of answer. Architecture consists of question processing, document extraction and processing and finally answers processing. Here in the question processing module the question is analyzed using NLP techniques like POS tagger, Parser, NER. In second module relevant documents are retrieved from repository based on conceptual indexing and processed to extract candidate answer set. In answer processing module candidate answers are filtered and finally answer are generated. The literature review shows that most of the existing QA systems are available for English language and some researchers have worked on Hindi, Telgu and Punjabi as Indian regional languages. Most of these algorithms have used Cross Lingual based approach to extract the information. The QA system for Telgu is based on dialogue manger which uses SQL query

generator to fetch answer. Most of the existing system mostly provide answers for “what, where, when and who” type of questions only.

Various approaches like DBPedia framework, Ontology, synonym matching, SQL query generator, Bigram, NER had been used in past to extract answer for given questions. But most of them worked well with English language only. Literature review also shows that similar work of QAs for Marathi language has been recently started. Author has used concept of Ontology but the actual ontology is created and traversed in English language so Cross Lingual based approach is used to extract the information.

3. PROPOSED SYSTEM

The proposed system provides most relevant and precise answer to the user’s natural language questions through semantic matching by using ontology. The input to the system is users question in Marathi language and output will be precise answer of the question.

Fig.1 presents proposed framework of Marathi QA system. User specifies the query in Marathi natural language in textual form. Input to the system is natural language Query in Marathi language. Input query is first tokenized to generate individual tokens and then these tokens undergo word grouping where two or three corresponding word are merged together if they are related with each other by using the available word grouped list. Part of speech (POS) tagging is performed on word grouped tokenized query text to extract relevant part of speech associated with the query text. POS tagged query text then passes through chunking process where noun and verb grouped present in the query text are extracted. Based on the extracted chunked groups initially query triples are extracted using Subject, Object and Verb (SOV). Then next process is to generate onto triples by fetching relevant onto words from ontology. Finally ontology is traversed to fetch relevant answer based on generate onto triples, if onto triple matches with any onto set in ontology then corresponding answer is fetched and passed to answer generation process to present the answer as natural as possible mostly in the form of natural language text.

Sample input and output for Marathi query:

Input Question: मुंबईची प्रमुख भाषा कोणती आहे?

Answer: मुंबईची प्रमुख भाषा मराठी आहे.

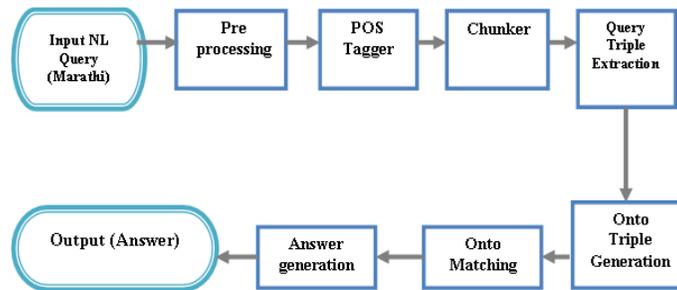


Figure 1. Proposed Framework

4. WORKING OF THE SYSTEM

Proposed question answering system is a text based question answering system where ontology is created for different domains for semantic representation of Marathi content.

Due to unavailability of ontology creation tool for Indian regional language like Hindi and Marathi, we have created a simple representation for creation of ontology in Marathi by taking into account the generalized approach used for creation of ontology in other languages like English.

After specifying the domain of ontology, stemming is performed on the document for which ontology is to be created. As Hindi and Marathi are morphologically rich languages, root word need to be extracted from the given document. After stemming is performed important terms in the document are extracted manually. These extracted terms are mainly nouns, adjective and other modifiers surrounding noun, verbs and its supporting auxiliary verbs. Form the extracted terms the nouns and verbs are the candidates to be the entity in the ontology and the modifiers associated with nouns and verbs become property or attributes of those entities. Then finally relation between entities' is extracted and stored in the ontology. The root word is useful for traversing ontology.

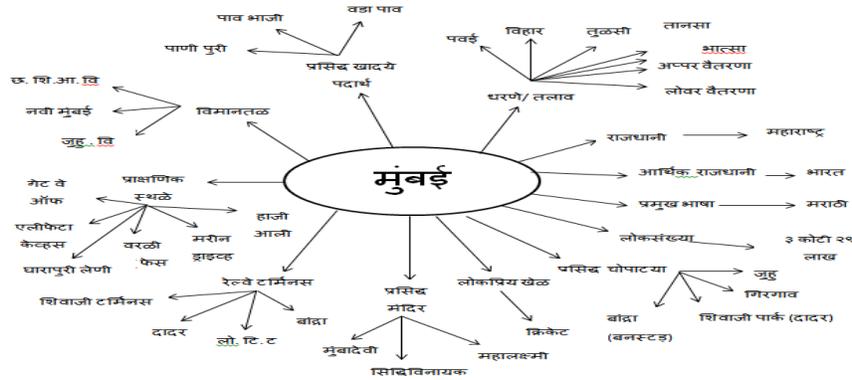


Figure 2. Sample Ontology for Mumbai City

After query is provided by user, users question text is passed onto Marathi QA module which contains tokenization, word grouping, POS tagging, Chunking, Query Triple extraction, Onto Triple Extraction, Onto matching and fetching of answer.

5. EXPERIMENTAL EVALUATION

In order to show that our proposal can have a great interest and that it can contribute to improve the performance of the Marathi QA task, we conducted various case studies and developed a prototype to show that the proposed framework can improve the performance of Marathi question answering system.

5.1. CASE STUDY

Let us consider the following scenario where user asks the question as:

Marathi Question: मुंबईतील विविध विमानतळांची नाव काय आहेत?

Input question is tokenized to generate tokens from the question, while tokenizing filtering of text is also performed to remove non Marathi tokens using UTF8 codes.

Tokenized Query:

Token 0 : मुंबईतील

Token 1: विविध

Token 2: विमानतळांची

Token 3: नाव

Token 4: काय

Token 5: आहेत

After tokenization, related tokens are combined together which on grouping constitute as single entity.

Word Grouped Tokenized Query:

Token 0: मुंबईतील

Token 1: विविध

Token 2: विमानतळांची

Token 3: नाव

Token 4: काय

Token 5: आहेत

Part of speech is generated for all the extracted tokens as Noun, Quantifier, Intensifier, Verb, Adjective, etc. are assigned to them.

POS tagged query:

मुंबईतील:: NNP विविध:: JJ विमानतळांची:: NN नाव:: NN काय:: WQ आहेत:: VAUX

Further the process of chunking is carried out to extract noun and verb group from the POS tagged question. Chunked groups can be in the form of proper noun, common noun and verb groups.

Chunked query:

[मुंबईतील NNP]:: NNPG [विविध JJ विमानतळांची NN]:: NNG [नाव NN]:: NNG [काय]:: WQ [आहेत]:: VAUX

Noun Group1: मुंबईतील NNP

Noun Group2: विविध JJ विमानतळांची NN

Noun Group3: नाव NN

Every question of Marathi language may at least contain a subject in it or it can contain combination of subject object and predicate. Subject object and predicated thus contribute for generation of query triples in the question.

Query Triple:

काय(मुंबईतील,विविध_विमानतळ,नाव)

User provided question will not always contain same terms as stored in the ontology for such scenario semantic mapping of user terms to corresponding onto term is needed. Query triple thus generated are transformed to onto triple.

Onto Triple:

काय(मुंबई,विविध_विमानतळ,नाव)

And finally matching of onto terms of question with those stored in ontology is done which leads to retrieval of accurate answer for the given question.

Answer extracted from ontology for the given question is:

मुंबईतील विविध विमानतळांची नाव छत्रपती शिवाजी आंतरराष्ट्रीय विमानतळ,जुहू विमानतळ आहेत.

5.2. EXPERIMENTAL RESULTS

In QA systems it is important to retrieve exact answer or part of the answer that will satisfy the user question. There are number of evaluation measures that can be used to compare the performance of the various retrieval techniques. Precision and Recall are the most commonly used indicators to measure Information extraction quality.

Accuracy, precision and recall are used as performance metrics which can be defined as True Positive (TP), True Negative (TN), False Positive (FP) and False negative (FN):

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN})$$

Marathi QA system accepts questions in simple sentences, analyses them, and returns answers in a single word, phrase, or sentence Here in terms of Marathi QA system, TP is number of question correctly answered, FP is number of question wrongly answered, TN is answer present in system which have no importance to context and FN specify number of answers to question present in the system but are not retrieved.

The system is evaluated to check whether the answer to the user question is relevant or not. Mostly QAS either provides relevant answer for the user question or it simply returns null if no answer is found. It is more like a Hit or Miss System i.e. either we will get answer for a question or we won't get answer.

We experimentally evaluated the performance of the proposed framework by testing it with various Marathi documents of different domains like history, festival, sports, city, politics...etc.

Table 1 shows Contingency table for history domain, where number of questions asked was 55 out of which 51 questions where correctly answered and 3 questions where either not answered or incorrectly answered.

Table 1. Contingency table for history domain

Total answers	Relevant answers	Non relevant answers
Answer Retrieved	51	3
Answers not retrieved	1	0

Here TP =51, FP =3, TN = 0, FN = 1.

Precision in % = 94.44% Recall in % = 98.07%

Accuracy in % = 92.72%

Marathi QA system (MQAS) was evaluated for different domains like history, sports, city, entertainment, politics and festival using the metrics such as precision, recall, accuracy and F-Measure. Table 2 shows the test results of Marathi QA system for a particular run.

Table 2. Experimental analysis of Ontology based Semantic information extraction system

Sr. No.	Input Domain	Precision in %	Recall in %	Accuracy in %	F- Measure
1	History	94.44	98.08	92.73	1
2	Sports	93.33	93.33	87.50	1
3	City	100.00	100.00	100.00	1
4	Entertainment	93.33	93.33	87.50	1
5	Political	91.67	91.67	84.62	1
6	Festival	90.91	90.91	83.33	1
Average		93.95	94.55	89.28	

The proposed framework's efficiency is compared with publicly available search engines like Google and Bing. Table 3 shows domain based accuracy comparison for MQAS with Google and Bing.

Here we are calculating accuracy of system by taking percentage of answer retrieved for set of question.

Figure 3 shows average accuracy comparison between MQAS, Google and Bing for various Marathi language documents.

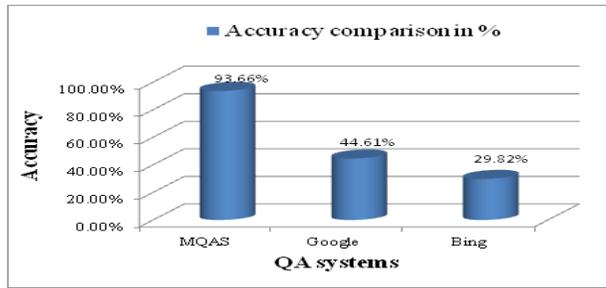


Figure 3. Overall accuracy Comparison between MQAS, Google and Bing

Sr. No.	Input Domain	Total Number of Questions fired	Total Number of questions for which semantically correct answers extracted			Accuracy in %		
			MQAS	Google	Bing	MQAS	Google	Bing
1	History	55	51	22	10	92.73	40.00	18.18
2	Sports	15	14	7	3	93.33	46.67	20.00
3	City	13	13	7	5	100.00	53.85	38.46
4	Entertainment	15	14	6	5	93.33	40.00	33.33
5	Political	12	11	5	5	91.67	41.67	41.67
6	Festival	11	10	5	3	90.91	45.45	27.27
Average Accuracy						93.66	44.61	29.82

Figure 4. Domain based accuracy comparison of MQAS, Google and Bing

Performance of MQAS was evaluated by measuring its ability to retrieve all and only relevant information. MQAS performance is strongly dependant on POS Tagging and correct processing of the queries. The system achieved an overall precision of 93.95%, recall of 94.55% and accuracy of 89.28% and F-Measure as 1. Table 3 describes the performance of MQAS based on question type. The designed system is tested with 20 different types of question types in Marathi language. Average Precision of 100.00% shows that all the answers retrieved are correct answers. Percentage of recall is 97.11%. Factoid and certain non-factoid questions were only considered in this work. Yes/No questions are not considered in the design of MQAS and hence still it remains as a research topic.

Table 3. Performance Analysis of MQAS according to Marathi Question Type

Question Type	Precision in %	Recall in %	F measure
कोण	100.00	80.00	88.89
कठे	100.00	100.00	100.00
काय	100.00	100.00	100.00
कोणत्या	100.00	90.00	94.74
कधी	100.00	88.89	94.12
कोणी	100.00	83.33	90.91
कोणाचे	100.00	100.00	100.00
कशाचे	100.00	100.00	100.00
केव्हा	100.00	100.00	100.00
कोणाबरोबर	100.00	100.00	100.00
कोणती	100.00	100.00	100.00
कशाने	100.00	100.00	100.00
कसा	100.00	100.00	100.00
कठला	100.00	100.00	100.00
कशी	100.00	100.00	100.00
किती	100.00	100.00	100.00
कोणकोणती	100.00	100.00	100.00
कोणकोणते	100.00	100.00	100.00
कोणता	100.00	100.00	100.00
कोणते	100.00	100.00	100.00

Handling of 'कसं' and 'का' type questions are the most difficult because they mostly require answers spreading over more than one sentence or paragraph. These questions sometimes require deep semantic processing of the sentences and identification of more keywords to detect the presence of explanations, intentions, justifications etc.

6. CONCLUSIONS

The system is tested with Marathi documents of various domains like History, sports, festival, politics.etc and shows an overall precision of 93.95%, recall of 94.55% and accuracy of 89.28%. The designed system is also tested with 20 different types of question types in Marathi language for ex: "कोण, कुठे, काय, कोणत्या, कधी, कोणी, कशाचे, केव्हा, कसा, कुठला, कशी, किती" and system shows precision of 100.00%, recall is 97.11%. Factoid and certain non-factoid questions were only considered in this work.

The proposed system is compared with publicly available search engines like Google and Bing. The system shows average accuracy as 93.66%, 44.61 and 29.82% for designed MQAS, Google and Bing respectively.

7. FUTURE SCOPE

At present, domain specific ontology construction is a manual task. No tool is available till date for automatic ontology construction for Marathi language. The future enhancement to the current methodology is to build the ontology automatically by using a tool. Automated tool can be developed to minimize the manual intervention in QA process.

In spite of significant contributions made by proposed system, there are number of research avenues which can be taken up in future. The dataset considered under study was very small in size and also for very few domains of Marathi language. In future, system can be tested with large dataset. Factoid and certain non-factoid questions were only considered in this work. Yes/No questions are not considered in the design. The research can be further extended for handling of 'कसं' and 'क' type questions which are most difficult type questions as these questions require deep semantic processing of the sentences to extract answer.

The systems can be scaled to cover much more domains and support of more complex natural language queries in the future.

ACKNOWLEDGEMENTS

With deep sense of gratitude I express my sincere thanks to my esteemed and worthy supervisor Dr. J.W.Bakal for his valuable guidance in carrying out this work under his effective supervision. My greatest thanks are to all who wishes me success and supports me to complete this.

REFERENCES

- [1] Sahu, Shriya, N. Vashnik, and Devshri Roy. "Prashnottar: A Hindi Question Answering System." *International Journal of Computer Science and Information Technology (IJCSIT)* 4.2 (2012): 149-158.
- [2] Stalin, Shalini, Rajeev Pandey, and Raju Barskar. "Web based Application for Hindi Question Answering System." *International Journal of Electronics and Computer Science Engineering* 2 (2012): 72-78.
- [3] Kumar, Praveen, et al. "A query answering system for E-learning Hindi documents." *South Asian Language Review* 13.1&2 (2003).
- [4] Sharma, Lokesh Kumar, and Namita Mittal. "Named Entity Based Answer Extraction form Hindi Text Corpus Using n-grams."
- [5] R. Reddy, N. Reddy and S. Bandyopadhyay, "Dialogue based Question Answering System in Teulgu," In *Proceedings of EACL Workshop on Multilingual Question Answering*, 2006, pp. 53-60.
- [6] V. Gupta, "A Proposed Online Approach of English and Punjabi Question Answering," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 6, 2013, pp. 292-295.
- [7] P. Gupta and V. Gupta, "Algorithm for Punjabi Question Answering System," *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, vol.3, 2013, pp. 902-909.
- [8] J. Cherapanamjeri, L. Lingareddy, Himabindu. K, "Keyword based Question Answering System in Natural Language Interface to Database", *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* Volume 3 Issue 12, December 2014.
- [9] Chaware, S., and S. Rao. "Ontology supported inference system for Hindi and Marathi." *Technology Enhanced Education (ICTEE)*, 2012 IEEE International Conference on IEEE. 2012.
- [10] Tahri, Adel, and Okba Tibermacine. "DBPEDIA BASED FACTOID QUESTION ANSWERING SYSTEM." *International Journal of Web & Semantic Technology* 4.3 .2013.

- [11] Wang, Chong, et al. "Panto: A portable natural language interface to ontologies." *The Semantic Web: Research and Applications*. Springer Berlin Heidelberg, 2007.
- [12] Lopez, Vanessa, Michele Pasin, and Enrico Motta. "Aqualog: An ontology-portable question answering system for the semantic web." *The Semantic Web: Research and Applications*. Springer Berlin Heidelberg, 2005.
- [13] Raj, P. C. "Architecture of an Ontology-Based Domain-Specific Natural Language Question Answering System." arXiv preprint arXiv: 1311.3175 (2013)

AUTHORS

Sharvari Govilkar is professor in Computer Engineering Department, at PIIT, New Panvel and was research scholar at TSEC, Bandra , University of Mumbai, India. She has received her M.E in Computer Engineering from University of Mumbai. And completed her Ph.D. in Information Technology from University of Mumbai. She is having twenty years of experience in teaching. She has publications in various national and international journals & conferences. Her areas of interest are Text recognition, Natural language processing, Information Retrieval, domain specific ontology construction etc.



J. W. Bakal received M.Tech in Electronics Design & Technology, from Dr. BAMU University. Later, He has completed his Ph.D. in the field of Computer Engineering from Bharati University, Pune. He is a PhD supervisor in CSE at University of Mumbai. He is presently working as principal at the S.S. Jondhale College of Engineering, Thane, India. He was a chairman of board of studies in Information Technology in University of Mumbai. His research interests are Telecomm Networking, Mobile Computing and Information Security. He has publications in journals, conference proceedings, and books in his credits. During his academics tenure, he has attended, organized and conducted training programs in Computer and Electronics branches. He is life member of professional societies such as IETE, ISTE INDIA. He is also a member of IEEE. He has prominently worked for IETE as a chairman, Mumbai section.

