ISSN (e): 2250 – 3005 || Volume, 09 || Issue, 3 || March– 2019 ||
International Journal of Computational Engineering Research (IJCER)

"Implementation of Question and Answering Retrieval System over Watch Word & Head Word"

Atul Kumar Mishra[1], Prof. Amit Saxena[2], Prof. Kaptan Singh[3]
[1] Dept. of CSE, TIEIT, Bhopal, M.P., India
[2] Prof. Dept. of CSE, TIEIT, Bhopal, M.P., India
[3] Prof. Dept. of CSE, TIEIT, Bhopal, M.P., India
Corresponding Author: Atul Kumar Mishra

ABSTRACT: *In Processer System, Query Replying System continuously assistances to the user to find the answers of the given question. Information area is a new area for a humanoid. With the assistance of web flatterers, we can get any data gamely accessible. We are only a tick far from receiving to a page at a remote corner of the world. We have regularly needed PCs to act savvy [1]. To achieve this undertaking the arena of Artificial Intelligence looked. One of the key obstructions in making PCs clever is knowledge of Natural Language. Normal dialect administration which manages to understand of dialects is a subdivision of Artificial Intellect. Square plan of Query Replying Scheme is looking in underneath figure1.1.It is inordinate NLP request. A task that an inquiry noticing framework admits is given an autopsy and meeting of records, finds the correct response for the question. It joins two integral objects: first, understand the different issues in usual dialect understanding and depiction than the second to strategy steady dialect interface to PCs [2].*
KEYWORDS: *Natural Language processing, information retrieval, semantic similarity, delimited area, riposte removal, answer ranking.*

Date of Submission: 28-03-2019 Date of acceptance: 08-04-2019

I. INTRODUCTION:

Though usual of leaflets which are saved by the search train shelter a ration of information around the search them nonetheless it might not cover exactly the material [1].The basic idea behind the inquiry replying system is that the operators just have to arrive the query and the scheme will recover the most appropriate and exact answer for that question and revisit. For small and exact reply, query answer scheme theater a huge position quite than hunt Engines, which usually provide a large set of links of those web pages which valor contain the answer to that question. A characteristic Query Replying scheme alienated into 3 components specifically: Query Dispensation module, in order repossession component and rejoin processing module. Each Processing and Information Retrieval module covers numerous subunits and these units usage numerous Query Replying scheme is intended to response humble wh-questions like "who", "what", "when", "where", etc. But the new QA investigation emphases on spreading the system to answer complex questions, summary questions, opinion questions etc. The daily proposes a Query Replying arrangement that answers humble factoid, wh-questions by a method termed Semantic Role Labeling.

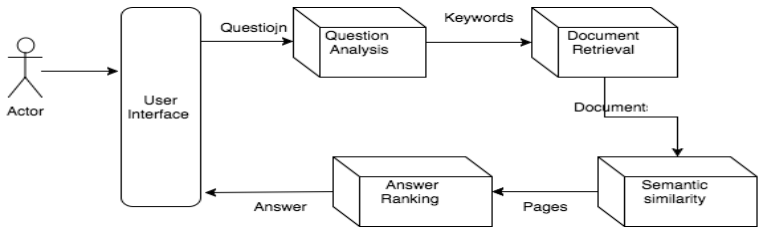


Figure 1. Figure for query answer scheme

The respite of the weekly is ordered as follows. The ensuing segment calls the broad manner of a Query Replying Scheme. Unit 3 debates specifically of the connected everything in this part. The future scheme building is labeled in section 4. The paper concludes with the new setup and results.

II. ARCHITECTURE OF A QUESTION ANSWERING

In this segment, we label the building of our scheme. The universal structure of the scheme can be split into three main components: (1) pre-processing, (2) question stencil matching, and (3) answering. Each module is categorized in detail in the following subsections.

Query Replying Schemes can be secret on the foundation of the domains over which it has been constructed.

- Open Field Query Replying
- Close Field Query Replying
- Limited Area Query Replying

Exposed area query replying schemes are area self-governing. Usually, these systems have a big group of data from where the required answer is to be found out. Since in case of unlocking area query answer information content is not of a particular domain it can answer questions of various fields however here deep reasoning is not possible [3].

Near domain query replying schemes contract with queries in a specific domain [3]. In case of close area query replying schemes bottom less cognitive is likely but the problem with these schemes was that due to the actual minor size of statistics set they are not additional than a 'Toy Schemes"[4].

Investigation in restricted-domain query replying (RDQA) speeches glitches connected to the combination of area- exact gen into current state-of-the-art QA knowledge with the confidence of reaching deep intellectual abilities and reliable accuracy performance in real-world applications. In fact, as a not too-long-term vision,

III. LITERATURE SURVEY:

In investigation identifications [4, 5, 6] LUNAR [7] and BASEBALL [8] have been deliberated as the previous industrialized query replying schemes. Though numerous inquiry responding schemes which have been industrialized with dissimilar ideas since the knowledge of QA Scheme has been invented.

In a scheme settled Athira P. M, Et.al [10 Applies semantics and domain knowledge to improve both query construction and answer extraction.

an additional scheme urbanized by Pragisha K. Et.al [11], labeled around them. It gets Malayalam natural language investigations from the operator and excerpts suitable reply by examining a group of Malayalam leaflets. The scheme grips four each question.

Investigation and evaluations in inquiry answering system established through Sanjay K Dwivedi Et.al[12] suggest classification for typifying Query Response (QA) schemes, review of main QA schemes labeled in works and deliver a qualitative analysis of them.

Table [I] gifts contrast around dissimilar kinds of question answering system [22].

S. No	Type of Question and Answering System	Question and Answering System Methods
1	Multilingual Question/Answering	Tokenization and pos tagging., Word sense disambiguation, Answer type identification, Keywords expansion, Semantic Disambiguation
2	Analysis of the Asks Question-Answering System	Query Reformulation, NGram Mining, N-Gram Sifting, N-Gram Tiling.
3	Multilinguality, Spatial-sequentialsettingconsciousness, Textual entailment	Answering architecture
4	A Question Answering System based on Information Retrieval and Validation	Expected Answer Type, Named Entities Presence,
5	A Hybrid QueryReplying System grounded on Material Retrieval and Reaction Validation	Module, Hypothesis Generation Module, Document Processing and Indexing
6	A specifiable domain multilingual Question	Answering architecture

Query replying is a tough formula of evidence reclamation branded by material needs that are at least somewhat expressed as natural language Template Corresponding Automatic Answering System For natural idioms queries future by Pachpind Priyanka Et.al [17], Regularly Inquired QA Scheme that answers with pre-stored responses to worker queries requested in steady English, rather than keyword or sentence structure based retrieval mechanisms.

IV. PROBLEM IDENTIFICATION

As initial as 2002 a collection of researchers3 wrote a roadmap of the investigation in the field of query answering. They also recognized the issues associated to question answering. The subsequent conversation is based on the issues they acknowledged during their research.

1. Question classes
2. Question processing
3. Context and QA
4. Databases aimed at QA
5. Answer extraction
6. Answer training
7. Real-time query replying
8. Multilingual (or cross-lingual) question answering
9. Interactive QA
10. Advanced cognitive for QA
11. Gen clustering for QA 12. User summarizing for QA.

V. PROPOSED SYSTEM:

The architecture consists of various steps.

1. Query Section:- The User enters a Question in the section.
2. Preprocessing:- In the step, the question entered by the user undergoes three methods

i. Tokenization- Here the question entered is converted into tokens or single words
ii. Stop word removal- All the stop words such as is, am are, etc., are removed in this process.
iii. Stemming-Stemming refers to the reducing of the term to its origin by filtering out prefix and suffix of the word.

3. Token Identification: - The Following stage afterward preprocessing is Token Identification. This is the important step for answer extraction where tokens present in the question are identified for an efficient answer extraction process.
4. Question Analysis:- This phase is broadly divided into three categories :

i. Definition Type:- Definition Type of question requires one or two sentence as an answer
ii. Descriptive Type: -Descriptive Type of question requires a few sets of sentences or a paragraph as an answer
iii. Factoid Type: - These require one or two-word answer. Forex. Why, How and Explain Question are asked for the descriptive type of answers. Who, When, Where, What, Which are generally asked for Factoid type of answers here who signifies the name of a person, When signifies Time/Date, Where signifies Place/Location.

5. Headword Selection: - After the Question Analysis and Token generation the following stage is Head word Generation. Here the Tokens which are generated in the third phase (Token Identification) are chosen as a Head Word. This Head word will be useful in the next phase of Clustering.
6. Clustering: - Here Crowding performance is pragmatic on Wikipedia documents set for answer mining process. Clustering is the job of association a set of substances in such a way that gears in a similar cluster are additional alike to each other than to persons in another group. The Clustering Technique used in this paper is K-Means. K-means Clustering is a method of Vector Quantization. It objectives to barrier n Explanations into k bands in which each remark fits the cluster with the nearest mean.

The Head Word Selected in the above phrase is used in this phase to form clusters. To apply K-Means on the data set TF-IDF has been used. Three clusters are formed using a K-means algorithm, from where and the responses will be retrieved.

7. Templates Matching: - . Templates are the predefined format of the answer which will be presented in front of the user. Templates are formed by the Headword, selected in the 5th phase and by the Question Format decided in the Question analysis Phase.

The Answer Extraction process will match the following Templates in its Database to give an exact answer. Templates such as:

i. "Head Word" is . . . . . .
ii. "Head Word" means . . . . . .
iii. "Head Word" is known as . . . . . .
iv. "Head Word" is called. . . . . .
v. "Head Word" can be defined as . . . . . .

8. Answer Extraction: - In this Phase, the models caused in the above phase will be matched to that in the Clusters formed in the clustering phase.

VI. RESULT ANALYSIS

Question and Answering System is developed in this research with help Java (JDK1.8) and Net Beans IDE8.02 on window operating sytem7. All methods of Query Replying Organization design in Swing. The graph plotted

for computation time, type of question and memory management using JFree Chart Library. In Result, Analysis compares Planned Query Replying scheme with present Query Replying scheme in the period of computation time and memory.
In Query, Replying Scheme took each kind of queries for an experiment like Factoid Question, Descriptive and Definition. Wikipedia used as a dataset for search Question answers. Below figure 5.1 shows the home screen of the project.

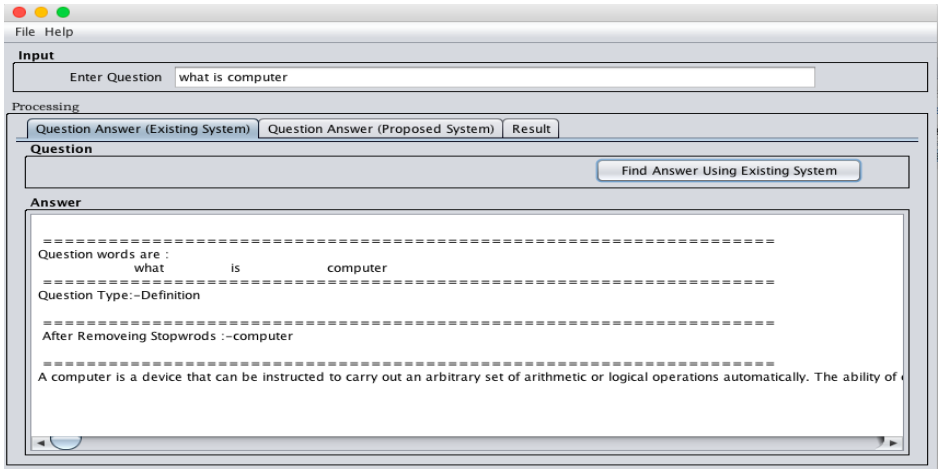


Figure 5.1 Home screen of project

6.1.1 Evolution Parameters

In Query Replying scheme focus on following parameters

- Question type
- Computation Time
- Memory Management

6.1.1.1 Question Types

Find type of Question corresponding to Enter Question for Answer. Using type of question design template that helps to find more accurate answer for given entered Question.

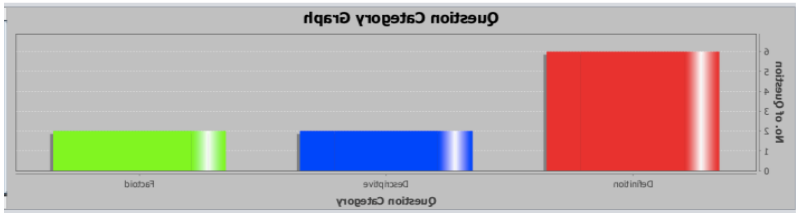


Figure 5.2 Question Types

Table 5.1 Number of Question in Types

S.No.	Question Type	No. Of Questions
1	Definition Type	6
2	Description Type	2
3	Factoid Type	2

Table 5.2 Question Type of each Question

S.No.	Question Number	Question Type
1	Question Number 1	Description
2	Question Number 2	Definition
3	Question Number 3	Definition
4	Question Number 4	Definition
5	Question Number 5	Description
6	Question Number 6	Definition
7	Question Number 7	Factoid
8	Question Number 8	Description
9	Question Number 9	Factoid
10	Question Number 10	Definition

6.1.1.2 Computation Time

We calculate computation time for Exiting QueryReplyingscheme and PlannedQueryReplying system. And results shown with help of graph. From experiments found that Proposed Question Answering system less computation time compare to Existing Question Answering system.

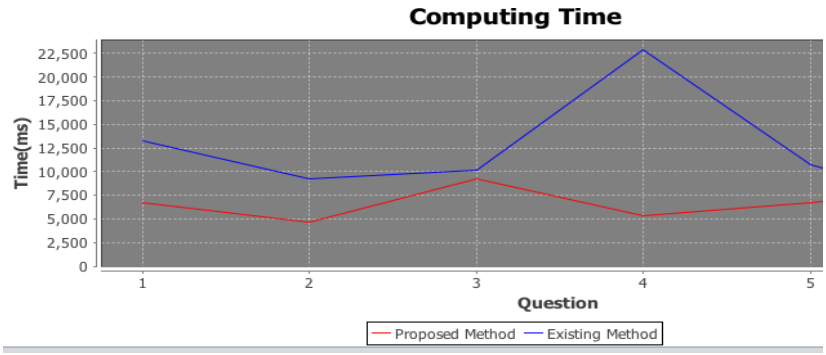


Figure 5.3 Computation time for Existing and Proposed System

Table 5.3 Computation time for Existing and Proposed System

S.No.	Question Number	Computation Time of Existing Question Answering System (MS)	Computation Time of Proposed Question Answering System (MS)
1	Question Number 1	13203	7215
2	Question Number 2	9853	4978
3	Question Number 3	10340	9734
4	Question Number 4	22565	5123
5	Question Number 5	11287	7460

6.1.1.3 Computation Memory

We calculate computation memory for Exiting Query Replying scheme and Proposed Query Replying system. And results shown with the help of graph. From experiments found that Proposed Question Answering system less computation memory compare to Existing Question Answering system.

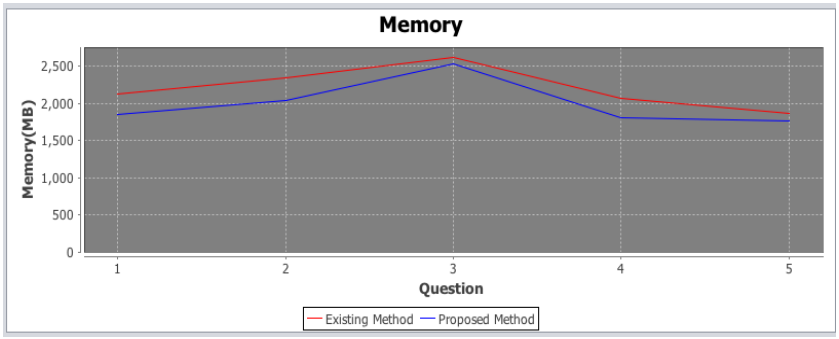


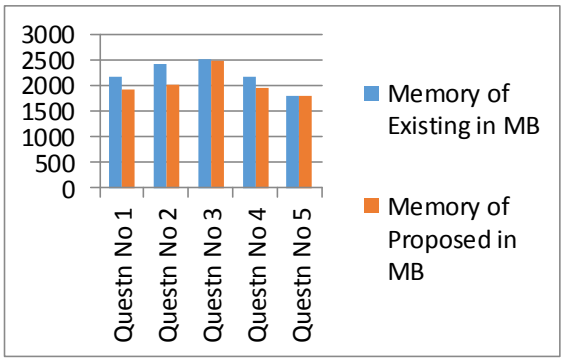
Figure 5.3 Computation Memories for Existing and Proposed System

5.2 Output Screen

Table 5.4 Computation memory for Existing and Proposed System

S.No.	Question Number	Computation Memory of Existing Question Answering System (MB)	Computation Memory of Proposed Question Answering System (MB)
1	Question Number 1	2183	1940
2	Question Number 2	2434	2015
3	Question Number 3	2507	2473
4	Question Number 4	2164	1946
5	Question Number 5	1798	1816

Comparative Analysis by Chart:-



VII. CONCLUSION:

Here we have proposed an outline for limited area query Replying Scheme using progressive NLP tools and software. This outline can be rummage-sale to mature a Query Replying Scheme for removing careful and detailed response from limited domain textual data set. The proposed framework not only provides a simple and implementable agenda for evolving question Answering System but also runs a suitable drift of data for riposte extraction.

Since the suggested classic all over keywords and headword and is independent of the question or sentence structure, it has reduced the overhead of question normalization. Moreover, meanwhile the outline is given for restricted domain, it also handles the issue of word sense disambiguation. The main problem which exists with the proposed agenda is that its presentation is in need of on the presentation of the hunt engine and the used NLP tools.

REFERENCES

[1]. Sreelakshmi V, Sangeetha Jamal, Survey Paper : Question Answering Systems, in National Conference on Computing and Communication - (NCCC), March 2014, GEC Idukki.
[2]. M Ramprasath, S Hariharan Improved Question Answering System by semantic reformulation, IEEE- Fourth International Conference on Advanced Computing, 2012.
[3]. Ali Mohamed Nabil Allam, and Mohamed Hassan Haggag, The Question Answering Systems: A Survey, International Journal of Research and Reviews in Information Sciences (IJRRIS), September 2012 Science Academy Publisher, United Kingdom
[4]. Molla D., and Vicedo J., "Question answering in restricted domains: An overview", Computer Linguist, ppAI-6 1, 2007
[5]. Moreda P., Llorens H., Saquete E., & Palomar M., "Combining semantic information in question answering systems", Information Processing & Management, pp.870-885, 20 1 1.
[6]. Svetlana Stoyanchev, and Young Chol Song, and William Lahti, "Exact Phrases in Information Retrieval for Question Answering", Coling 2008: Proceedings of the 2nd workshop on Information Retrieval for Question Answering (IR4QA), pp. 9- 16 Manchester,UK. August 2008".
[7]. Woods W.A, Kaplan R.A, Nash-Webber.B, "The lunar sciences naturallanguage information system" , Final report: BBN Report #2378. Technical report, Bolt Beranek and Newman Inc.,Cambridge, MA., June 1972.
[8]. Green RF, Wolf A.K., Chomsky, K. Laughery, "BASEBALL: An automatic question answerer", in: Proceedings of Western Computing Conference, vol.19, pp. 2 19-224, 196 1.
[9]. Ittycheriah A, Franz M, Zhu WJ, Ratnaparkhi A and Mammone RJ. IBM"s statistical question answering system. In Proceedings of the Text Retrieval Conference TREC-9, 2000.
[10]. Athira P. M., Sreeja M. and P. C. Reghuraj Department of Computer Science and Engineering, Government Engineering College, Sreekrishnapuram, Palakkad, Kerala, India, 678633. Architecture of an Ontology-Based Domain-Specific Natural Language Question Answering System.
[11]. Pragisha K. "design and implementation of a QA system in Malayalam".
[12]. Sanjay K Dwivedi, Vaishali Singh. Research and reviews in question answering system Department of Computer Science, B. B. A. University (A Central University) Luck now, Uttar Pradesh, 226025, India.
[13]. Poonam Gupta, Vishal Gupta Assistant Professor, Computer Science & Engineering Department University Institute of Engineering & Technology Panjab University, Chandigarh.
[14]. Kolomiyets, Oleksander. And Moens, Marie-Francine. "A survey on question answering technology from an information retrieval perspective". Journal of Information Sciences 181, 2011.5412-5434. DOI: 10.1016/j.ins.2011.07.047. Elsevier.
[15]. Ko, Jeongwoo., Si, Luo., and Nyberg Eric. "Combining evidence with a probabilistic framework for answer ranking and answer merging in question answering" Journal: Information Processing and Management 46, 2010 541-554. DOI: 10.1016/j.ipm.2009.11.004. Elsevier.
[16]. Pachpind Priyanka P, BornareHarshita N, KshirsagarRutumbhara B, Malve Ashish D BE Comp S.N.D COE & RC, YEOLA," An Automatic Answering System Using Template Matching For Natural Language Questions".

Atul Kumar Mishra "Implementation of Question and Answering Retrieval System over Watch Word & Head Word" International Journal of Computational Engineering Research (IJCER), vol. 09, no. 3, 2019, pp 77-82