International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || PP. 66-75

QA System: A Perspective View

Lovely Sharma
(PCM SD College, Jalandhar, Punjab, India)

Abstract: The enormous amount of information is led to the development of Question answering (QA) dialogue system. The user hope for the relevant, accurate and precise information from the web, while the results are in the form of documents, links, etc. Thus, QA dialogue system is an emerging area of Natural Language Processing (NLP) which assists the user to find the information in an interactive way. Many tools and techniques could be applied to process the natural language query. This paper presents the review of various approaches for QA system, comparison of accuracy and limitations. Firstly, we provide the bird's eye view of the QA system and terminology to understand the NLP. Third, we describe the pocket sketch of QA architecture and forth section elaborates the history of QA system. In the last section, we conclude the QA system and give encouraging way to future researches in this field.

Keywords: Accuracy, Natural language understanding, QA approaches, Question answering dialogue system.

I. Introduction

Currently, there has been tremendous amount of information is available on the web. Users often have specific queries in their mind, for which they expect concise, comprehensible and accurate answers from the web in their native language without being restricted to a specified query language, query formulation rules, linguistic knowledge etc. while search engine provides ranked list of weblinks, documents etc. So, the need of QA system becomes more vital.

Question-Answering (QA) Dialogue System is a technique used to extract information from the web or structured database based on the Dialogue system. To achieve this, it is important to understand the query of the user in native language. The approach to understand the user's query is to comprehend the query from a linguistic point of view and to understand what the user really wants to ask. Typically, this leads to the development of various modules as explained in section 3. In general, the pipeline of these modules will interact with each other to achieve the precise answer for the user.

Recent successes are achieved because of Text REtrieval Conference (TREC) [1]. TREC is chain of workshops that focuses on the various areas of Information Retrieval (IR) and provides the support and motivate to IR community for research. Each track gives challenges to participating groups with set of problems. Researchers can present their ideas, thoughts, research work etc. in the workshop. Problems might be documents, questions, or to access any target. Owing to TREC, we have best systems which enable us to give answers of more than two thirds of factual questions.

II. An Overview of Natural Language Understanding

To process the query, system should be equipped with linguistic knowledge to understand the natural language query. Thus, various tools and techniques are available for this such as morphological analysis, syntactic and semantic analysis, shallow parser etc. as mentioned below. Moreover, to understand the query is a huge task and every researcher used its own algorithm, architecture, model and technique to process the query.

In detail, Morphological analysis is the study of internal structure of words/sentence [2]. It is important to identify the part of speech to find out the grammatical context to which it belongs. Morphological process provides the information for the syntactical stage which helps for the grammatical structure of the target language. Whereas, Semantic analysis is used to determine the meaning of the sentence [3]. It can be applied on single word or sentence. Morphology provides the structure of sentence whereas semantic implies for what is meant by a sentence [4]. Semantic analysis provides that what the object is going to represent and how a sentence is describing the action. It also contains the detail of adjectives, adverbs and propositions. In addition to this, Syntactic analysis or parsing is the process of analysing the rules of a formal grammar [5]. Its main task is to organize the data into a structured manner that can be manipulated further. It helps to create the parse tree of the given sentence. This parse tree will help to provide the meaning of sentence for helpful in semantic analysis.

III. A Pocket Sketch of QA System

The previous section described the general introduction of QA system. In this section, we will discuss the working of QA system, as keep in mind the generic view of QA architecture is useful. We quickly outline

each module of the model and discussing how the pipeline of QA architecture interacts with each other to provide succinct answer to User. Figure 1 introduces a widely used approach that receives a natural language query and provide output as an answer in a simpler and interactive manner.

The general QA architecture consists of Natural language query, Query processing, Query generation, Answer extraction and Answer display as shown in Fig (1). User asks the query by means of user query interface. Query could be processed through various tools and techniques such as shallow parser, POS tagging etc. The next stage Query processing provides the keywords, chunks etc. and helpful in query generation. After that, SQL query will strike to database and directed to the next phase i.e. SQL engine will extract all the possible answers from the database and display possible answer to the User.

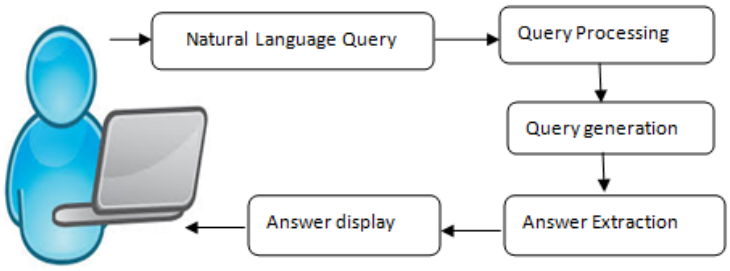

Fig.1- General architecture of QA Systems.

3.1 Query Processing

The natural language query needs to be analysed and processed before directed to the next phase. The purpose of this stage is to input the query in native language posed by the user using query interface without knowing the linguistic knowledge of that particular language and most of the important task is to process the Query. This leads to the following steps:

a) Classification of Question

b) To obtain keywords, tokens etc from the question which will further used to generate Query Frames.

In the first step, Question Classification is used to classify the question based on the entity type as shown in Table (1). The following table is referred from Kangavari et al. [6].

TABLE1- Question Classification Table

Q-type	Q- subtype	Answer type
Who		Person
Why		Reason
What	Simple what	Number/definition/title/undefined
	What – Who	Person
	What – When	Date
	What – Where	Location
Where		Location
When		Date/time
Which	Which-Who	Person
	Which-Where	Location
	Which-When	Date
	Which-what	Organization
How	Simple- how	Manner
	How –many	Number
	How –much	Value
	How- long	Distance/time.
	How –large	Number
	How- far	Distance
	How –tall	Number
	How- rich	Undefined
Whom		Person/organization

In detail, Questions can also be classified on the basis of algorithms such as Support Vector Machine [7], Template Alignment Algorithm [8] etc. The main purpose of QA System is to provide the result of WHAT, WHEN, WHERE, WHY, HOW, to WHOM type of questions [9].

The next stage is known as Query Reformulation [10] or Query Expansion [11]. Query expansion does the POS Tagging [12], Shallow Parser [13], Syntactic and Semantic analysis [14], Morphological analysis [15] etc. It also fixes the spelling or linguistic errors. Then, it extracts the keywords, generates the query frames and provides expected answer type. This stage is also used to improve the performance of information retrieval from the database.

3.1.1 Query generation

Query could be generating with the help of SQL Engine. SQL engine needs tokens, keywords, chunks etc. to generate SQL Query and foremost Query frame. These tokens and keywords are used to provide the appropriate query frame. The Selection of query frame has a great influence on the precision of the system. If query frame is not complete then, user enters into dialogue with system. System will ask the required queries from the user to fill the query frame. After completion of query frame, SQL Query is generated and passed to SQL Engine.

3.1.2 Answer Extraction

In this phase, SQL Query triggered on the database stored in SQL and SQL Engine will produce the accurate answer. The retrieved answer is used to display to the user in native language. The furnished information is updated in Dialogue History [16]. Dialogue History stores the current and previous conversation of the User and the System.

3.1.3 Answer Display

User needs always specific and precise answer from the huge amount of data which saves the time. SQL engine sends the mapped data from the database and display to the user in the appropriate form. The answer types could be classified as factoid and non-factoid. A factoid answer provides the short and precise answers such as names, timing, date etc. whereas non-factoid gives subjective answers.

IV. Related History

Only in few years, there is drastic growth in the field of NLP. This is because of the need of the User to search the precise information from the web. So, Simmons (1965) [17] introduced the survey article "Answering English Questions by Computer" in English language .He implements only fifteen questions built over the preceding five years whereas system consists of Structured Database and tries to find answers to questions from text sources, such as encyclopaedias.

To continue this tradition, another work is done by Grosz, Sparck Jones and Webber on articles is PLANES [18], LADDER [19], and TEAM [20] respectively. Also, a good review of work through 1990 is done by Copestake and Sparck Jones [21].The most notable early work here is START [22] and till now it is running on the web since 1993. It answers complex questions using approach of decomposition based on syntactic analysis.

Over the years, Question answering system is emerging very fast. There has been tremendous improvement in the field of Natural Language Processing (NLP) from the last decade e.g. research has been done on Hindi, Punjabi, Malayalam languages etc. such as Sekine et al. (2003) [23] introduced Hindi- English cross lingual QA system. This system accepts and responds in English language by accepting the input from Hindi newspaper. In the upcoming sections, we are classifying the QA system with providing the bottomless details of used techniques, approaches, models, algorithms etc.

4.1 Rule Based QA System

Rules Based Systems can apply to implement lexical analysis in natural language Processing [24]. An example of Rule based systems are domain dependent Expert System that used to specify rules or choices [25]. Various algorithms are also developed to process the documents such as Gupta et al. (2013) [26] described an algorithm for accepting user's query or document in Punjabi language. Firstly, the algorithm classifies the question on the basis of who, what, when, where and why. The next step is to discard the stop words present in query. The stop words are already developed using English and Punjabi language. Then, keywords are accepted from rest of the query as verbs, adjectives and nouns. The developed algorithm also retrieves synonyms of the keywords with the use of dictionary of Punjabi and English by applying vector space technique. The retrieved keywords and synonyms are applied on the reformulation of question. As a result, various answers are extracted

from online available web documents and set ranks for the candidate answers. The average accuracy of the system is 73%. The loophole does not return accurate answers for this type of questions.

4.2 Web Based QA System

A huge amount of data is available on web and that is considered one of the best sources to obtain the information. Web based QA systems depend on the search engines to provide the potential containing answers to the questions of the users [27]. This method will provide the answers in text documents, PDF, XML documents, Wikipedia etc. The data on the web has semi structure, heterogeneity and distributive properties.
The most popular techniques for QA systems are Open Domain System and Restricted Domain QA System while Web based QA systems [28] work for open domain [29]. TREC is the major scale evaluation environment for open domain QA system. The focus of TREC is to build a fully automatic open-domain QA System, which can answer a large variety of questions based on a huge amount of data.

By following the tradition of web based system, Kumar et al. (2005) [30] introduces a light weight stemmer for Hindi and also developed case-based rules to classify the question. The proposed architecture composed with automatic entity generator module is to recognize domain specific entities. Question classification module is used to classify the questions and questions are filtered out to remove the stop words and to identify the important keywords by recognizing the domain entities. Query formulation module reformulates the user's query by generating set of queries and is passed to SQL engine to retrieve the answer from the Database. SQL Engine provides the weightage to the passages and also provides ranks on the basis of locality. Answer selection module generates the reply by immense analysis of passage. System provides 75% accuracy by retrieving the three passages and 12% of the questions are directed to the relevant documents. 13% of the failure of results is due to lack of clear boundary and question classification.

Researchers also show the interest in document processing. Stalin et al. (2012) [27] implemented the web based QA system for retrieving the answers for user query in Hindi language from the Hindi document. The general architecture consists of various modules. Query interface accepts the query from the user in Hindi Language. Question parsing module is used to find out the domain related entities after removing the stop words. Query formulation module uses the entity file and system reads the default file and creates a hash table of these entities. Each word from the user query is mapped from the hash table to identify the entities. Query expansion module enhances the search process by including semantically related terms. System includes the Hindi world net. Information retrieval engine extracts the answer from the collection of documents. Short story is stored in database and user query maps with the DB. Sentences matching the words of questions are searched and related documents are displayed on the screen. The system works by involving the various steps as follows:

1. Users add the story and its title by clicking the add story button.
2. User then clicks on Home link to go back to the homepage and search the answer.
3. User clicks on search answer link on homepage to retrieve the relevant answer from the database. System then searches the answer from the Google if it is not able to give the answer.
4. User again clicks on home link to go on Homepage to view graph option. User can select story from the list box to generate the graph.

The result of the introduced system is inconclusive.

Again, Gupta (2013) [31] has proposed a technique of question answering for English and Punjabi text. Initially, system inputs the query from the user and stop words are removed. The list of stop words is already stored for English and Punjabi. The Keywords will be retrieved as nouns, adjectives and verbs. Synonyms of these keywords are also retrieved from bilingual dictionary using vector space model. Query is reformulated and relevant web pages are extracted by mapping string. At last, QA system returns results by using online search engine, assigns the ranks to candidate answers and finally, top 20 answers are extracted from user's query. The QA system is particularly implemented and returns 70% of the accuracy. Rest of the 30% errors are due to lack of consistence and syntax errors in the dictionary of Punjabi.

TABLE 2- Comparison Table of Various Models

SR. NO.	TITLE	ACCURACY	QA type	MODEL/TECHNIQUE FOR QUERY PROCESSING	LIMITATIONS	References
1.	BASEBALL – An Automatic Question – Answer	Not given and unsatisfied results	Restricted	Linguistic, semantic analysis	1. Questions are limited to a single clause 2. and, or, not are prohibited. 3. Semantic ambiguities unsolved	[13]
2.	SHRDLU	Quiet accurate	Restricted	Backward chaining or means-ends analysis	Not useful for commercial field	[33]
3.	Lunar	90%	Restricted	Syntactic and semantic analysis	Narrow domain	[44]
4.	Knowledge Structures in UC, the UNIX Consultant	Good efficiency	Restricted	Minsky, s theory of frames	Narrow domain	[6]
5.	The Berkeley Unix Consultant Project	Quiet satisfactory	Restricted	Language analysis using ALANA	1. Components are not integrated i.e. flow of the system is unidirectional 2. repeated question responds no answer	[42]
6.	From Sentence Processing to Information Access on the World Wide Web	Satisfactory results	Open	Morphological, semantic and syntactic analysis	Unstructured data because of open domain	[22]
7.	GENIES: A natural language processing system for the extraction of molecular pathways from journal articles	Precision=96% Recall=63%	Restricted	Own architecture using MedLEE	1. Incomplete lexicon 2.Grammar rules are incomplete	[10]
8.	Hindi-English cross-lingual question-answering system	Each component is calculated separately but accuracy is acceptable	Open domain	Monolingual QA systems with Named Entity Tagger	Errors in Keyword translation	[32]
9.	A Hindi Question Answering system for E-learning documents	Directly answered of the questions= 75% Directed to the relavant document=12%	Open domain	A light – weight stemmer for Hindi & case based rules to classify the Question	To lack of clear boundary between classes of classification & inadequate use of syntactic information.	[23]
10.	Dialogue based Question Answering System in Telugu	Dialogue success rate = 83.96% Precision = 96.34%	Restricted	Architecture consists of Query analyser, query frames, dialogue manager, dialogue history, SQL generation and answer generation	Query frames for some natural language queries were not correctly identified.	[29]
11.	A Practical QA System in Restricted Domains	precision =90.9% recall = 75.0%	IR/IE	QA system with QA system and IE engine	1. Insufficient data of weather-domain ontology 2. Error due to query frame analysis	[7]

4.3 Restricted Domain QA System

Restricted domain QA Systems are under specific domain queries and are related only to that particular domain [32, 33, 15] and based on general corpus provides a list of documents. Therefore, restricted domain gives more accurate result as compared to listing a set of related documents and websites. In recent years, open domain systems and closed domain systems has emerged very fast. Nowadays, users have no time to search the specific information from the web. So, Question answering dialogue system become worthwhile and provides unerring information but in a restricted manner. As we start from early systems, these are very restricted to the domain and provide confined information of the user's query. The best-known early program with restricted domain is BASEBALL [34], a program to provide answers about American league over one season but domain is limited to baseball only. It is primarily based on structured DB that reads the question from punch cards, lookup for keywords and idioms in dictionary. The data mining is required for specification and necessary processing is done. The answer is fetched through structured database. Thus, BASEBALL analyses the user query through linguistic knowledge against structured database having baseball data. SHRDLU [35] is perhaps the first AI System to accept the user query and send back the reply to the user. The SHRDLU [36] and GUS (Genial Understand System) [37] systems are developed to help researchers in their respective fields. The SHRDLU system performs to control the robot hand which help to arrange the blocks as requested by the user. Whereas GUS system is developed to help the travel advisor and works on restricted domain about Airlines Flights. After this system, many upcoming systems follow the same approaches to retrieve appropriate data.

						3. Low recall rate due to invalid date and topic.	
12.	Multilingual Restricted Domain QA System with Dialogue Management (Bengali and Telugu)	As shown in Table (2).	Restricted	Shallow parser	NLP queries are not perfectly identified and some queries are out of database coverage. Unable to obtain chunks, so generating wrong answers.		[12]
13.	A New Model for Question-Answering System	Version no. 1(based on fixed pattern which user cannot change)=70% Version no. 2(based on system can check the dictation of word)=90%	Restricted	Syntax n semantic analysis	Low accuracy due to query reformulation due to syntax n semantic relation		[21]
14.	Prashnottar : A Hindi question Answering system	88%	Restricted	Hindi Shallow Parser using Query logic language (QLL)	Accuracy of Question type 'Where' is low because the answer type of this question is Proper noun i.e. location.		[31]
15.	Evaluating Web-based Question Answering Systems	Good results	Open domain	proximity algorithm and probabilistic phrase ranking	Inappropriate performance of precision		[28]
16.	Web Based Application for Hindi Question Answering System	Inconclusive request	Restricted	POS tagging	Less semantic implamentation of answer extraction		[39]
17.	Algorithm for Punjabi Question Answering System	Average accuracy=73%	Restricted	Develop step algorithm and hybrid model (Pattern matching + new proposed answer finding scoring system)	Do not support which, how much, how many type of questions.		[17]
18.	A Proposed Online approach of English and Punjabi Question Answering	Accuracy of half implementation=70 %	Open Domain	Architecture consists of query analysis, query generator, open web search, snippet extractor and answer ranker	1. insufficiancy of consistency 2. Punjabi dictionary is not correct		[16]
19.	Semantic parsing for single-relation question answering	F; score=0.61	Open Domain	Convolutional Neural Network based Semantic Model to answer single- relation question	Entity linking model is not so robust.		[45]
20.	A Dialogue System for Telugu, a Resource-Poor Language	Metric Average	Restricted domain	System architecture consists of Data	Performance decreases as complexity of sentence increases		[38]

Another well remembered system in this tradition is the LUNAR system. This system provides the easy way to access, compare and assess the chemical analysis data of Lunar Rock [38]. The result is quite satisfactory that it provides the 90% accurate answers within the domain but the limitation is just to a narrow domain. Early dialogue system such as UC (Unix Consultant) [39] and The Berkeley Unix Consultant project (UC) provides the reply of user's query about UNIX operating system [40]. UNIX consultant System proposed the Minsky's theory of frames. The system is based on the slots-filler format to store all relevant information to provide the concise data to the user.

Another early dialogue system is GENIES [41], developed in 1999 extracts and structures information related to molecular pathways. It processes the complete article and achieves the high accuracy as 96% (binary relations) and 88% (nested relations) respectively. The next method implemented by Reddy et al. (2006) [42] introduced the Question answering based dialogue system for Indian railways in Telugu Language. The major module of this system is Dialogue Manger (DM).

The proposed architecture accepts the natural query in Telugu language, analyses and processes the query and provides the relevant results to the user in Telugu Language. In detail, the QA system accepts the users query in Telegu language. Query analyser accepts the query and generates the Tokens. Tokens are mapped

against the knowledgebase. The appropriate query frame is generated based on the keywords and tokens. If the query frame is complete, the associated SQL procedure is generated; otherwise system will enter into a dialogue with the user through the DM. The DM will generate the interactive messages to ask the user to complete the relevant frame. The user will send the needed information to the system. If user is unable to provide the relevant information to the system, then DM will generate the error messages to the user. After completion of query frame SQL query will be generated and answer will be retrieved from the DB using SQL query. Thus, DM is responsible to maintain the flow of data between user and system. The measurement parameters of the system are dialogue success rate and precision. Dialogue success rate is 83.96% and precision rate is 96.34%. The low dialogue success rate is due to many reasons, such as information provided by the user is not adequate, system is not able to obtain tokens correctly, system coverage of the domain is not extensive enough, and misinterpretation of dialogue history is also another problem.

While another QA system based on Hindi language is proposed by Sinha (2006) [15] presented the same approach for Railway in Hindi Domain along with system using shallow syntactic and semantic analysis on the input query. The given architecture is not implemented. Consequently, the results are not found.

Moreover, Godavarthy et al. (2007) [43] has discussed the multilingual restricted domain QA system with dialogue management. The system architecture has various components such as system maintains language related to domain and linguistic models. User will specify the query in Telugu or Bengali language and Shallow parser component will parse the input query. It will generate tokens and keywords to identify the purpose of query. Dialogue manager will generate the interactive messages to the user to complete the query input. Dialogue model defines the parameters such as clarification requests, intended dialogue acts and discourse goals. Dialogue history is used to maintain the answers for the previous queries. Answer generation module will fill the slots in the answer template. SQL generation component will generate the SQL query by retrieving the data from database and provide the relevant answer to the user.

Hence, the result is based on two parameters: Precision and Recall value with dialogue management and without dialogue management. The results are referred from Godavarthy et al. (2007) [43] as follows.

Table 3- Results of multilingual restricted domain qa system with dialogue management.

With dialogue management			Without dialogue management	
Bengali Query system	Dialogue success rate	72.91 %	Precision	85.70 %
	Precision	83.67 %	Recall	80%
Telegu Query system	Dialogue success rate	89.06 %	Precision	97.63 %
	Precision	96.49 %	Recall	93.93 %

The low dialogue success rate is due to domain coverage of the system that domain is not extensive enough due to this some queries are out of database coverage. Another reason is misinterpretation of dialogue history.

By following the same tradition, the other work in this line is done by Kangavari et al. (2008) [44] developed a model to retrieve accurate answer in "weather forecasting" domain. The system works to parse the query of user by syntax and semantic relation among words. The question is classified and reformulated to find out set of keywords that are further sent to search engine. The search engine part finds candidate answer document and pass it to answer processing module to extract correct answers. The candidate answers collection is filtered on the basis of co-occurrence patterns and it assigns the priority number to the candidate answers. At the end, system ranks these answers and sends to users to validate the retrieved answers to find out the accurate result. The accuracy for version1 (based on fixed pattern which user cannot change) and version2 (based on system can check the dictation of word) is 70% and 90% respectively. The system can be improved by improving the reformulation of query and by applying user's assessments to validate answers.

4.4 IR/IE Based QA System

Since the mid of 1950s, Information Retrieval (IR)/Information Extraction (IE) [45] is the hot area for research. Users wish to ask the query and system gives answer in the form of documents with relevant passages. The IE community plans for its own evaluation and developed the Message Understanding Conferences (MUC) which ran between 1987 and 1998.

Besides this, Chung et. al. (2007) [32] also implemented QA and IE engine to extract weather information from the website. According to developed system, user's request is analyzed by the Query analyzer, subsequent frames are generated and SQL engine will generates the relevant query. Reply of the user's query is generated from the database, which stores weather information from the web pages. The only difference to analyze the question is through named entity tagger, temporal data normaliser and inference rules. The precision and recall rates are 90.9% and 75.0% respectively. The low recall rate is due to some questions related to invalid data and topic. The primary reason of failure is due to the insufficient weather domain ontology data.

Unlike others, Sahu et al. (2012) [46] has followed the QA system using Query Logic Language (QLL) with the help of Hindi shallow parser which is developed by IIIT Hyderabad. The architecture for the QA system was proposed and involved the various modules such as Query Processing, Query Generation, DB search etc. The given QA system provides the result to users in Hindi Language. The main focus of the system was on four types of questions when, where, what & how many types of questions. The first module of architecture accepts the users Query in Hindi Language & classifies it. The second module uses the (Query Logic Language) QLL on the user's Query. QLL is a subset of Prolog used to represent the answer. Hindi shallow Parser provides the morphological analysis, Part of Speech (POS) tagging, chunking etc. of the sentence. Result generated by last module is saved as .wxi format using developed rules and is displayed to the user. The overall accuracy of this QA system is approx. 68.00 %. The accuracy of Question type 'where' is low because the question is related to location and location is proper noun and proper noun is very difficult to identify.

V. Conclusion and Future Scope

The main focus of QA system is to provide precise, definite and accurate answers to users rather than to provide documents or ranked website links. TREC also plays an important role to give pathway to researchers to evaluate the QA system. Recent successes are achieved only due to TREC. They provide critical evaluation, focuses on challenging tasks, factual answers and result comparisons. In this paper, we discussed some of existing QA systems, applied tools and techniques, various models, algorithms and approaches to retrieve unambiguous and appropriate answers.

A survey of various QA systems has been mentioned in Table (3). Many researchers used morphology, POS tagging, semantic n syntactic tools etc. to process the query input to generate the accurate answer. The produced results are quite satisfactory and acceptable. In addition to restricted domain, the results are good because of consideration of specific domain. The narrow domain filters the choice of query asked by user but the results can be improved. In many approaches, query processing and dialogue manager are the main components of the general QA architecture. We can enhance the query processing module with the addition of their own logic because as we can observe the table that their own developed architecture provides the more accurate results as compare to morphological, semantic n syntactic analysis etc. tools applied on general QA architecture.

In future, QA system will have to consider multimedia data. It should consider images, videos, audio, metadata etc. Thus, QA system is all about human-computer interaction, to understand and interpret query, integrate systems and intelligent tutor system.

References

[1]. D. Harman, "The Text REtrieval Conferences (TRECs)", *Proceedings of a workshop, Vienna*, Virginia, May 6-8, 1996.
[2]. M. Aronoff, and K. Fudeman, "What is Morphology?," Oxford: Blackwell, 2005.
[3]. L. Hirschman , R. Gaizauskas, "Natural language question answering: the view from here", *Natural Language Engineering*, Vol.-7, Issue no-4, pp. 275-300, December 2001.
[4]. M.R. Kangavari, S. Ghandchi, M. Golpour, "A New Model for Question Answering Systems", *Journal of world Academy of Science, Engineering and Technology*, 2008.
[5]. E. Brill, "A simple rule-based part of speech tagger", *Proceedings of the Third Conference on Applied Natural Language Processing*, ACL, Trento, ltaly, 1992.
[6]. M.R. Kangavari, S. Ghandchi, M. Golpour, "A New Model for Question Answering Systems", *Journal of world Academy of Science, Engineering and Technology*, 2008.
[7]. D. Zhang, W. S. Lee, "Question classification using support vector machines*", Proceedings of SIGIR*, 2003.
[8]. B. L. Webber, "Questions, answers and responses: Interacting with knowledge-base systems", Knowledge Base Management Systems: Integrating Artificial Intelligence and database Technologies, *Springer-Verlag*, pp- 365-402, 1986.

[9]. P. Moreda, H. Llorens, E. Saquete, M. Palomar, "Combining semantic information in question answering systems", *Journal of Information Processing and Management 47*, DOI: 10.1016/j.ipm.2010.03.008. Elsevier, pp. 870- 885, , 2011.

[10]. C. Lioma and I. Ounis, "A Syntactically-Based Query Reformulation Technique for Information Retrieval", *Information Processing and Management (In Press),* 2007.

[11]. P. Gupta and V. Gupta, "A survey of text question answering techniques*", International Journal of Computer Applications,* 53(4), 2012, pp. 1-8.

[12]. V. Gupta, "A Proposed Online Approach of English and Punjabi Question Answering", *International Journal of Engineering Trends and Technology*, vol.-6, pp. 292-295, 2013.

[13]. S. Sekine and R. Grishman, "Hindi-English cross-lingual question-answering system", *ACM Transactions on Asian Language Information Processing (TALIP*), Vol-2, issue no-3, pp 181–192, 2003.

[14]. W. Woods, "Progress in natural language understanding -an application to lunar geology", *American Federation of Information Processing Societies (AFIPS),* Conference Proceedings, pp 441-450, 1973.

[15]. R.M.K. Sinha, "On design of a question-answering interface for Hindi in a restricted domain", *International Conference on Artificial Intelligence, Las Vegas, Nevada, U*SA, June 26-29, 2006.

[16]. F. Landragin and L. Romary, "Dialogue History Modelling for Multimodal Human-Computer Interaction", *Proceedings of the Catalogue Workshop*, 2004.

[17]. R. F. Simmons, "Answering English questions by computer: A survey", Communications Association for Computing Machinery, Vol.-8, Issue no. - 1, pp 53-70, Jan 1965.

[18]. B. J. Grosz, , K. S. Jones and B. L. Webber, "Readings in Natural Language Processing", Morgan Kaufmann, Los Altos, CA,1986.

[19]. K. S. Jones and P. Willett, "Readings in Information Retrieval", Morgan Kaufmann, San Francisco, CA, 1997.

[20]. B. L. Webber, "Questions, answers and responses: Interacting with knowledge-base systems", Knowledge Base Management Systems: Integrating Artificial Intelligence and database Technologies, *Springer-Verlag*, pp- 365-402, 1986.

[21]. A. Copestake and K. S. Jones, "Natural language interfaces to databases", The Knowledge Engineering Review, Vol.- 5,Issue No.-4, pp 225-249, 1990.

[22]. B. Katz , "From sentence processing to information access on the world wide web", Proceedings American Association for Artificial Intelligence (AAAI) Spring Symposium on Natural Language Processing for the World Wide Web, Stanford University, Stanford CA, 1997.

[23]. S. Sekine and R. Grishman, "Hindi-English cross-lingual question-answering system", *ACM Transactions on Asian Language Information Processing (TALIP),* Vol-2, issue no-3, pp 181–192, 2003.

[24]. E. Riloff and M. Thelen, "A rule- based question answering system for reading comprehension test", Proceedings of the ANLP/NAACL, Workshop on Reading Comprehension Tests as Evaluation for Computer-Based Language Understanding Systems, Seattle, Washington, pp 13-19, 2000.

[25]. E. Brill, "A simple rule-based part of speech tagger", Proceedings of the Third Conference on Applied Natural Language Processing, ACL, Trento, ltaly, 1992.

[26]. 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, pp. 902-909, 2013.

[27]. S. Stalin, R. Pandey and R. Barskar, "Web based Application for Hindi Question Answering System", *International Journal of Electronics and Computer Science Engineering*, vol. 2, pp. 72-78, 2012.

[28]. D. R. Radev, H. Qi, H. Wu, and W. Fan, "Evaluating web-based question answering systems", *Proceeding International Conference on Language Resources and Evaluation,* 2002.

[29]. D. Azari, E. Horvitz, S. Dumais, and E. Brill, "Web-based question answering: A decision-making perspective*", Conference on Uncertainty in Artificial Intelligence (UAI)*, 2003.

[30]. P. Kumar, S. Kashyap, A. Mittal and S. Gupta, "A Hindi question answering system for E-learning documents*", Proceedings of International Conference on intelligent sensing and Information processing*, pp 80-85, 2005.

[31]. V. Gupta, "A Proposed Online Approach of English and Punjabi Question Answering*", International Journal of Engineering Trends and Technology*, vol.-6, pp. 292-295, 2013.

[32]. H. Chung, Y. Song, K. Han, D. Yoon, J. Lee, H. Rim, "A Practical QA System in Restricted Domains", *Workshop on Question Answering in Restricted Domains, 42ndAnnual Meeting of the Association for Computational Linguistics* (ACL-2004), Barcelona, Spain, pp 39–45, 2004.

[33]. S. Fitrianie, P. Wiggers, and L.J.M. Rothkrantz, "A Multi-modal Eliza Using Natural Language Processing and Emotion Recognition", Matoušek, V., Mautner, P. (eds.) TSD 2003. LNCS (LNAI), Heidelberg, vol. 2807, pp. 394–399. Springer, 2003.
[34]. B. F. Green, A. K. Wolf, C. Chomsky, and K. Laughery, "Baseball: An automatic question answerer*", Computers and Thought*, pp 207–216, 1963.
[35]. Shrdlu, Galley: Article-00056.
[36]. T. Winograd , "Understanding Natural Language", Academic Press, New York, 1972.
[37]. D. G. Bobrow, R. M. Kaplan, M. K.,Donald A. Norman, H. Thompson and T. Winograd, "GUS, a Frame-Driven Dialog System", Artificial Intelligence 8, pp 155-173, 1977.
[38]. W. Woods, "Progress in natural language understanding -an application to lunar geology", *American Federation of Information Processing Societies (AFIPS), Conference Proceedings*, pp 441-450, 1973.
[39]. D.N. Chin, "Knowledge Structures in UC, the UNIX Consultant*", Proceedings of Twenty-first ACL*, pp 159-63, 1983.
[40]. R. Wilensky, D.N. Chin, M. Luria, J. Martin, J. Mayfield, D. Wu, "The Berkeley UNIX consultant project", Computational Linguistics, Vol. -14, Issue no.-4, pp 35-84, December 1988.
[41]. C. Friedman, P. Kra, H. Yu, M. Krauthammer, A. Rzhetsky, "GENIES: A natural language processing system for the extraction of molecular pathways from journal articles", *Bioinformatics, 17*, pp S74–S82, 2001.
[42]. R. Reddy, N. Reddy and S. Bandyopadhyay, "Dialogue based Question Answering System in Telulgu*", Proceedings of EACL Workshop on Multilingual Question Answering,* Association for Computational Linguistics, pp. 53-60, 2006.
[43]. S. R. Godavarthy, P. Pakray, S. Bandyopadhyay, "Multilingual Restricted Domain QA System with Dialogue Management", Proceedings of the Workshop on Cross Lingual Information Access, *International Joint Conference on Artificial Intelligence (IJCAI 2007), Hyderabad, India*, pp. 20-27, January 2007.
[44]. M.R. Kangavari, S. Ghandchi, M. Golpour, "A New Model for Question Answering Systems", *Journal of world Academy of Science, Engineering and Technology*, 2008.
[45]. R. Gaizauskas and K. Humphreys, "A Combined IR/NLP Approach to Question Answering Against Large Text Collections", *University of Sheffield* UK, 1998.
[46]. S. Sahu, N. Vashnik and D. Roy, "Prashnottar: A Hindi Question Answering System", *International Journal of Computer Science and Information Technology,* vol.-4, Issue no-2, pp.149-158, 2012