Chatbot System in using Open Source Python

[#1]Prof. Anita Mahajan, [#2]Mr. Aniket Kachare, [#3]Mr .Akshay Gorekar,
[#4]Miss. Prajakta Kunjir, [#5]Miss Aishwarya Jagtap

[1]mr.aniket007@gmail.com,
[2]akshaygorekar143@gmail.com,
[3]kunjirprajakta5@gmail.com,
[4]aishwariyajagtap@gmail.com

[#1]Assistant Professor, Department of Information Technology, G.H. Raisoni institute of
Engineering, Pune, India.
[#2345]U.G. Student, Department of Information Technology, G.H. Raisoni institute of
Engineering, Pune, India.

ABSTRACT

A chatbot is a conversational agent where a computer program is designed to simulate an intelligent conversation. It can take user input in many formats like voice, text, sentiments, etc. For this purpose, many open source platforms are available using python. Artificial Intelligence Markup Language (AIML) is derived from Extensible Markup Language (XML) which is used to build up a conversational agent like chatbot artificially. In this paper, AIML interpreter for the generation of the responses of users input is used.This method is used for developing an web application chatbot which will interact with user using text with voice.

Keywords: Artificial Intelligence, Chatbot, Web Application, AIML.

ARTICLE INFO

I. INTRODUCTION

Over the last few years, Chatbots have played a important role as human-computer interfaces. Chatbots are generally consist of three modules: the user interface, an interpreter, and a knowledge base. Laven [6] defines Chatbot as a program that attempts to simulate typed conversation, with the aim of fooling the human into thinking they were talking to another human. Basically, chatbot is a conversational agent that can interact with user in a given particular subject using the natural language. Many chatbots have been deployed on the internet for the purpose of education, customer service , guidance, entertainment. Existing and famous chatbots are ALICE [2], SimSimi and Cleverbot. Artificial Intelligence Markup Language (AIML) is derived from extensible Markup language (XML) which is used to build up a conversational agent artificially. The AIML based chatbots are light weighted, easy to configure and at minimum cost. AIML has class of data objects called AIML object which describes behaviour of computer programs. In this paper, the program which is an open source AIML engine written in PHP. It is an interpreter for the AIML scripts of the chatbot. It uses MySQL database to store the chatbot data. And all AIML scripts are store in database. When user interact with chatbot program, then according to matched reply from the AIML, the answer is formulated and send back to the user. It can be directly installed on a local server under the GNU General Public License.

II. LITERATURE SURVEY

[1] Emanuela Haller and Traian Rebedea, "Designing a Chat-bot that Simulates an Historical Figure", IEEE Conference Publications, July 2013. There are many applications that are incorporating a human appearance and intending to simulate human dialog, but in most of the cases the knowledge of the conversational bot is stored in a database created by a human experts. However, very few researches have investigated the idea of creating a chat-bot with an artificial character and personality starting from web pages or plain text about a certain person. This paper describes an approach to the idea of identifying the most important facts in texts describing the life (including the personality) of an historical figure for building a conversational agent that could be used in middle-school CSCL scenarios.

[2] Maja Pantic, Reinier Zwitserloot, and Robbert Jan Grootjans, "Teaching Introductory Artificial Intelligence Using Asimple Agent Framework", IEEE Transactions On Education, Vol. 48, No. 3, August 2005. This paper describes a flexible method of teaching introductory artificial intelligence (AI) using a novel, Java-implemented,

simple agent framework developed specifically for the purposes of this course. Although numerous agent frameworks have been proposed in the vast body of literature, none of these available frameworks proved to be simple enough to be used by first-year students of computer science. Hence, the authors set out to create a novel framework that would be suitable for the aims of the course, for the level of computing skills of the intended group of students, and for the size of this group of students. The content of the introductory AI course in question is a set of assignments that requires the students to use intelligent agents and other AI techniques to monitor, filter, and retrieve relevant information from the World Wide Web.

[3] Jiyou Jia, "The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages", This paper reports the findings of a study conducted on the application of an on-line human-computer dialog system with natural language (chatbot) on the teaching of foreign languages. A keywords-based human-computer dialog system makes it possible that the user could chat with the computer using a natural language, i.e. in English or in German to some extent.

[4] Robert P. Schumaker, Ying Liu, "Evaluating mass knowledge acquisition using the ALICE chatterbot: the AZ-ALICE dialog system", In this paper, we evaluate mass knowledge acquisition using modified ALICE chatterbots. In particular he investigate the potential of allowing subjects to modify chatterbot responses to see if distributed learning from a web environment can succeed. This experiment looks at dividing knowledge into general conversation and domain specific categories for which have selected telecommunications.

[5] Bayu Setiaji, Ferry Wahyu Wibowo, "Chatbot Using A Knowledge in Database Human-to-Machine Conversation Modeling", The sentence similarity calculation in this paper using bigram which divides input sentence as two letters of input sentence. The knowledge of chatbot are stored in the database. The chatbot consists of core and interface that is accessing that core in relational database management systems (RDBMS). The database has been employed as knowledge storage and interpreter has been employed as stored programs of function and procedure sets for pattern-matching requirement. The interface is standalone which has been built using programing language of Pascal and Java.

[6] Joseph Weizenbaum,"A Computer Program for the Study of Natural Language Communication between Man and Machine", ELIZA is a program operating within the MAC time-sharing system at MIT which makes certain kinds of natural language conversation between man and computer possible. Input sentences are analyzed on the basis of decomposition rules which are triggered by key words appearing in the input text. Responses are generated by reassembly rules associated with selected decomposition rules.
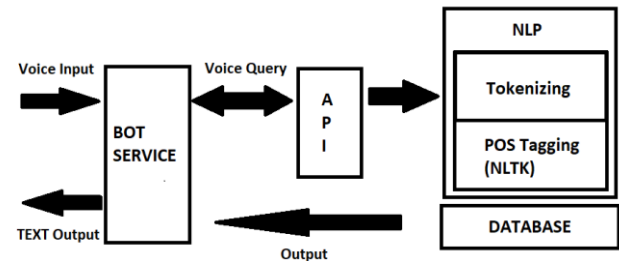
III. PROPOSED SYSTEM

Fig 1. Proposed architecture

The system recognizes user's query and understands what he wants to convey and simultaneously answers them appropriately. The questions asked by the users can be in any format. There is no specific format for users to ask questions. The built in artificial intelligence system realizes users requirements and provides suitable answers to the user.

Algorithm Used:

Algorithm:
1. Scanning the database:
    Here we will go through the database to get the table names, column names, primary and foreign keys.
2. Input:
    We will take a sentence as a input from the user (using voice) then convert into text.
3. Tokenize and Tag:
    We will tokenize the sentence and using POS tagging to tag the words
4. Syntactic Parsing:
    Here we will try to map the table name and column name with the given natural query. Also, we will try to identify different attributes of the query.
5. Filtering Redundancy:
    Here we will try to eliminate redundancy like if while mapping we have create a join requirement and if they are not necessary then we remove the extra table.
6. Query Formation:
    Here we will form a complete SQL query based on MySQL syntax.
7. Query Execution:
    Here we will execute the query on database to get results.

IV. MATHEMATICAL MODEL

Let S be a system that describes the execution of the application.
$S = \{I, O, M, S, F\}$, Identify the modules as M,
$S = \{M, ...\}$

$M = \{E, R\}$
where,
E = Predefined Questions .
R = Undefined Questions.

I = {user name, password, voice, questions}
(i) Identify input to E as Ie.

Ie= {W, n}
where,
W=Defined Questions.
n=Number of ways to ask a particular question.

(ii) Identify the modules of R as Mr
Mr= {Tl, Lv}
where,
Tl= Time required for transfer module.
Lv=Live support module.

O={Login success, Answer}

S= {when system analysis user equations properly and voice convert into the text}

F{when system can not understand user input questions and server failure}

V. CONCLUSION

The users of the computer devices are becoming more dependent on the machines. The chat bot uses artificial intelligence and hence will learn the responses of the users resulting in increasing efficiency. Chat bot has ability to respond like human being hence will reduce extra efforts required to be done by humans.

Currently chatbots have limited language support. They do not support multiple languages, dialects and do not understand colloquial usage. Hence there is a great scope for removing such language barriers in future chatbots. Also, AIML templates could be improved to include more variations for the same input.

REFERENCES

[1] Emanuela Haller and Traian Rebedea, "Designing a Chat-bot that Simulates an Historical Figure", IEEE Conference Publications, July 2013.

[2] Maja Pantic, Reinier Zwitserloot, and Robbert Jan Grootjans, "Teaching Introductory Artificial Intelligence Using Asimple Agent Framework", IEEE Transactions On Education, Vol. 48, No. 3, August 2005.

[3] J. Jia, "The Study of the Application of a keywords-based Chatbot System on the Teaching of Foreign Languages."

[4] Robert P. Schumaker, Ying Liu "Evaluating mass knowledge acquisition using the ALICE chatterbot: the AZ-ALICE dialog system"

[5] Bayu Setiaji, Ferry Wahyu Wibowo "Chatbot Using A Knowledge in Database Human-to-Machine Conversation Modeling"

[7] Weizenbaum, J. 1966. ELIZA- A Computer Program for the Study of Natural Language Communication between Man and Machine, CACM 9(7), 36-43