

Chatbot : Artificially Intelligent Conversational Agent

Ayesha Shaikh¹, Geetanjali Phalke², Pranita Patil³, Sangita Bhosale⁴, Jyoti Raghatwan⁵

U.G. Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, India¹

U.G. Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, India²

U.G. Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, India³

U.G. Student, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, India⁴

Assistant Professor, Department of Computer Engineering, RMD Sinhgad School of Engineering, Warje, Pune, India⁵

ABSTRACT: A chatbot is a humanlike conversational character. It is a computer program which conducts a conversation through auditory or textual methods. It often acts as a virtual assistant and it can have its own virtualization. Its conversational skills and other humanlike behavior is simulated through artificial intelligence. Nowadays chat-oriented dialogue systems are gaining popularity as they attempt to get into daily life and achieve some commercial success. There are different kinds of chat-bots with the aim of realizing natural human-computer interaction. As the chat-bots use various features, much handwork is needed to build the feature data. We expect that through the years every conversational chatbot will grow into a real virtual human. This paper proposes Architecture of chatbot system designed for college purpose. Students and teachers can access and use chatbot system anywhere and anytime to get required information.

KEYWORDS: Data Mining , Pattern matching, Knowledge Database, Dialogue System, Android.

I. INTRODUCTION

Recently chat-oriented dialogue systems have got common as they commit to get into everyday life and bring home the some business success. The most purpose of chatting systems is simulating the spoken language of men [2]. now of read can be stemmed from Turing take a look at, as several alternative dialogue systems do. Chatting systems area unit expected to reply to any user utterances in natural method to fulfill the requirements, it's necessary to method numerous user inputs. Also, to create users desire, they act with human, personalization methodology are often adopted to the chatting system.

Most well-known previous chat-bots were rule-based systems, that retrieve example sentences by applying easy pattern matching technique. ELIZA [3] and ALICE [4] area unit those ones of the renowned systems. Rule-based chatting systems attempt to keep voice communication going whereas avoiding inappropriate responses, although they can't find any matched pattern to the user-input. However, because the chatbots use numerous options, abundant piece of work is additionally required to make the feature information. To bring out most potential of those current systems, we propose an architecture for chatbot which is made specially for college and company. This chatbot is based on entities and intents. We used IBM's Bluemix Server for cloud storage and to run and manage our application of chatbot named as "MATE".

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

II. RELATED WORK

Jeesoo Bang, Hyungjong Noh, Gary Geunbae Lee and Yonghee Kim's [1] introduces a chat-oriented dialogue system. This system is example based system with personalization framework using long-term memory. Previous chat-bots use simple keyword and pattern matching methods. For generating number of heuristic rules, language expert knowledge is necessary. These rules are used to maintain the quality of systems.

J.Jia's [10], CSIEC adopts linguistics communication terminology as customary format of rules. It additionally uses sizable amount of linguistic options metaphysics like Chatscript. These recently developed chat-bots have powerful options and large potential to draw in the interest of users and manage the speech communication. This Chatbot focuses on supplying a virtual chatbot, which can chat in English with the users anytime. It provides response according to the user input, the user's and its own personality knowledge, common sense.

Kim, J. Bang, J. Choi, S. Ryu, and G. G. Lee's [11] This system collects user-related facts automatically from user input sentences and stores the facts in memory. If there are changes in users interest it also keeps track of them.

Chatscript [12], a chat-bot engine that's primarily the same as ALICE, suggests additional subtle options. It uses additional difficult pattern matching technique and metaphysics for understanding various intentions and topics. All of the extracted data is used once the system generates the system response.

III. SUMMARY

Our Chatbot is different than other existing chatbots because there is no special chatbot is created for particular college which can be helpful to all students. Proposed architecture is explained in fig. 1.

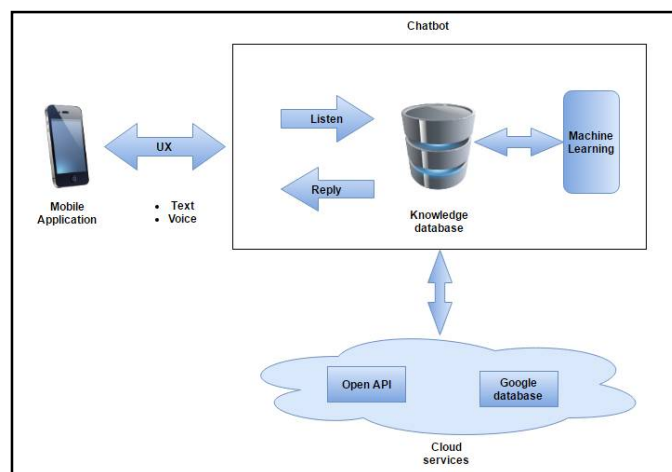


FIGURE 1. SYSTEM ARCHITECTURE

The fig. 1. Represents the system architecture of our system. In system architecture, Mobile application is android application chatbot. Machine UX is user experience design for application or it is kind of interface between user and server.

The main aim of chatbot is listen or read user's query and give response to user according to query. It needs a knowledge database which will store all queries in form of intents and entities and responses in form of dialogues because we have used IBM's Bluemix server [13] to store our database for cloud infrastructure.

Cloud computing is on demand software. We are using Software as a service (saas) in which software is provided on subscription basis to the user. IBM Bluemix gives the access to cognitive technologies like understanding

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

sentiment, keywords and entities from text. Open API allows one piece of software to interact with another piece of software. APIs are set of requirements that manage how one application can communicate and interact with another.

1. ALGORITHM STRATEGY:

Algorithm Strategy: Divide and Conquer Strategy

- Divide: The user's query will be divided into smaller tokens.
- Conquer: Tokens will be matched with the existing tokens in the database.
- Merge: Tokens will be then arranged according to its priority and the query will be reformed and answer to the query will be provided.

Proposed Algorithm:

Step-1: The user query will be classified using DA classifier i.e. Dialogue Acts classifier. DA tag or class will be assigned to the query using DA tagger.

Step-2: Query will be divided into tokens.

Let $\{w_1, w_2, w_3, \dots, w_n\}$ is a set of words.

$\{t_1, t_2, t_3, \dots, t_n\}$ is a set of POS tags,

$n = W_1, W_2, W_3, \dots, W_n$ is a sentence of n words,

$n = T_1, T_2, T_3, \dots, T_n$ is a sequence of n POS tags.

Step-3: Token will be tagged by POS-tagger and weighted.

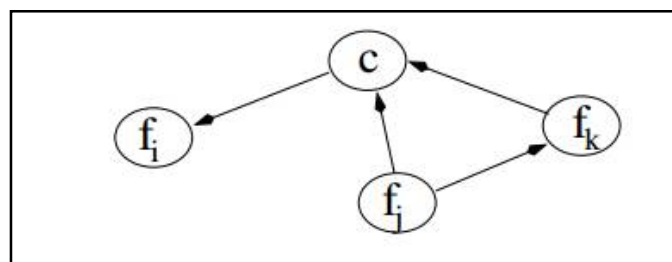


FIGURE 2. TAG AND ITS FEATURES

In fig. 2, we have to decide the max probability of the DA tag c and we are using arbitrary features $f_1, f_2, f_3, \dots, f_n$ shown.

The independent assumptions are made which make the estimation easier. In extreme case all the features are considered to be independent and the DA tag probability is estimated by

$$P(c) \prod P(f_i | c)$$

The DA probability of the graph can be estimated by:

$$P(f_i | c) \cdot P(c | f_j, f_k) \cdot P(f_k | f_j) \cdot P(f_j)$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

The probability of the most likely sequence is decided by using HMM i.e. Hidden Markov Model.

The standard HMM assume that the probability of the current tag t_i conditionally depends on only the previous K tags $t_{i-k}, -1$ and that the probability of the current word w_i conditionally depends on the current tag.

Step-4: Similar query to the users query will be matched using similarity functions by Watson Service on server.

Step-5: Final output will be answer of the matched query and chatbot will respond to user with the answer of matched query.

In this way, whole concept of chatbot can be explained in algorithm. Designing and Developing of chatbot can be complex but it is interesting to access and use.

IV. EXPERIMENTAL RESULT

In the graph shown below we compared our system with 4 different existing systems in the terms of accuracy, cost, speed and security.

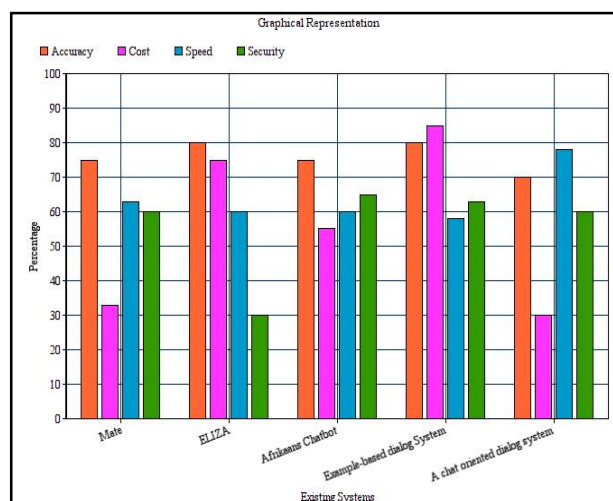


FIGURE 3. GRAPHICAL REPRESENTATION OF COMPARATIVE ANALYSIS

Fig. 3 gives us the comparative analysis of different systems. We compared the systems like Eliza, Afrikaans Chatbot, Example based dialog system in terms of accuracy, cost, speed and security. We compared our system with Eliza chatbot [3], ELIZA is a program which makes natural language conversation with a computer possible. This system has high accuracy and cost higher than us. Speed of this system is nearly same as our system. While this system is not so secured.

The another system is African Chatbot [2], in which the Java program used the Afrikaans dialogue corpus texts to generate two versions of the Afrikaans chatbot. From analysis, It has accuracy, speed and security same as Mate. Cost for making this system is higher than our system.

Example based dialogue system [5] has cost higher than us and low and chat oriented dialogue system [1] has speed greater than our system.

We conducted the experiment and did the word count of all the statements users used. Most of them are functional words like 'what', 'who', 'is', 'of' similar to the human conversations. We found that for chatbot Mate, the top ranked words to be 'subjects', 'schedules', 'classteacher', 'university' etc. We developed the system to initiate the

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

conversation regarding academics. Some of the users got annoyed and used slang languages or rude responses, even then our system responds with the appropriate answer.

We also observed the quality of the user's input is not very satisfying. Sometimes user failed to provide the right input. How to achieve high quality user input is a difficult question; very few in the chatbot community paid attention to it.

IV. CONCLUSION

Chatbot designers and developer's main aim is to build tools that help people, facilitate their work and their interaction with computers using natural language; but not imitate human conversation perfectly i.e. not to replace human role completely. Chatbot "MATE" is quick and handy and everyone can use it. The information we required we will get anywhere and anytime using this chatbot android application. Chatbots are effective tools when it comes to education, IR, e-commerce, etc. To improve performance while minimizing the human efforts, Already suggested and developed features are POS-tagged token matching, using NE-related information, and using back-off responses according to DA type. As a future development to chatbot, effort taken by human can be reduced. Additionally, more complex features such as parsing information can be used for sentence matching This chatbot application is developed on android platform, so future work can include development of this application in various platforms like ios, blackberry, etc. so everyone can use this application.

REFERENCES

- [1] Jeesoo Bang, Hyungjong Noh, Yonghee Kim, and Gary Geunbae Lee, "Example-based Chat-oriented Dialogue System with Personalized Long-term Memory" in IEEE, 2015.
- [2] <https://www.b.a.shawar.and.e.atwell>, "Using the Corpus of Spoken Afrikaans to generate an Afrikaans chatbot," Southern African Linguistics and Applied Language Studies, vol. 21, pp. 283-294, 2003. google.co.in/images/chatbot_architectures
- [3] J. Weizenbaum, "ELIZA—a computer program for the study of natural language communication between man and machine," Communications of the ACM, vol. 9, pp. 36-45, 1966.
- [4] R. S. Wallace, The anatomy of ALICE: Springer, 2009.
- [5] C. Lee, S. Jung, S. Kim, and G. G. Lee, "Example-based dialog modeling for practical multi-domain dialog system," Speech Communication, vol. 51, pp. 466-484, 2009
- [6] X. Shen, B. Tan, and C. Zhai, "Implicit user modeling for personalized search," in Proceedings of the 14th ACM international conference on Information and knowledge management, 2005, pp. 824-831.
- [7] F. Qiu and J. Cho, "Automatic identification of user interest for personalized search," in Proceedings of the 15th international conference on World Wide Web, 2006, pp. 727-736.
- [8] L. Ardissono, C. Gena, P. Torasso, F. Bellifemine, A. Difino, and B. Negro, User modeling and recommendation techniques for personalized electronic program guides: Springer, 2004.
- [9] Y. Jiang, J. Liu, M. Tang, and X. Liu, "An effective web service recommendation method based on personalized collaborative filtering," in Web Services (ICWS), 2011 IEEE International Conference on, 2011, pp. 211-218.
- [10] J. Jia, "CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning," Knowledge-Based Systems, vol. 22, pp. 249-255, 2009.
- [11] Y. Kim, J. Bang, J. Choi, S. Ryu, and G. G. Lee, "Acquisition and use of long-term memory for personalized dialog systems," in Proceedings of the 2014 Workshop on Multimodal Analyses enabling Artificial Agents in Human-Machine Interaction, 2014.
- [12] B. Wilcox, "Beyond Facade: Pattern matching for natural language applications," GamaSutra. com, 2011.
- [13] IBM's Bluemix Server[online] Available: <https://www.ibm.com>.
- [14] "Chatbot systems" [Online]. Available: <https://www.alicebot.org/chatbots/>.