International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

# Implementation of a Chatbot using Natural Language Processing

**Niranjan Dandekar[1], Suyog Ghodey[2]**

*[1,2]Department Of Computer, Pimpri Chinchwad College of Engineering*

**ABSTRACT**

*Over the past few years, Artificial Intelligence has grown and leaps and bounds. We are heading towards a society in which machines will take care of the most complex issues that we are confronting today. Before long, the greater part the assignments on the planet will be computerized. This will prompt a huge increment in connection amongst people and machines. This collaboration will be interceded by Natural Language handling. A chatterbot or chatbot intends to make a discussion between both human and machine. The machine has been implanted the learning to distinguish the sentences and choosing itself as reaction to answer a question. The reaction rule is coordinating the info sentence from client. The information of chatbot is put away in the database. Clearly, there is an expansion in the request of talk mechanization on the grounds that a) it expels the human element and b) it can give a 24-hour benefit which will multiplicatively affect the income era.*

***Keywords*: *Chatbot, Database, Natural Language Processing, Response principle***

## I. INTRODUCTION

The development of Artificial Intelligence applications is challenging because computers traditionally require humans to speak to them in a programming language that is precise, unambiguous and highly structured or, perhaps through a limited number of clearly-stated voice commands.

Natural language processing(NLP) is a branch of artificial intelligence, and machine linguistics that enables computers to derive meaning from human or natural language input. It is used to analyze text, allowing machines to understand human's language. NLP considers the hierarchical structure of human language in which several words make a phrase, several phrases make a sentence and, ultimately, sentences convey ideas. However, the ambiguity of language in which humans speak is what makes natural language processing a difficult problem for computers to undertake.

The computer performs Natural Language Understanding(NLG) to overcome this obstacle. It is the process of disassembling and parsing input because of the occurrence of unknown and unexpected components in the input and the need to determine the appropriate syntactic and semantic schemes to apply to it. A Chatbot is a computer program which conducts conversation with a human using auditory or textual methods. Chatbots are based on two basic principles Natural Language Processing and Pattern matching. We aim towards creating a conversational Chatbot with the help of NLP as well as pattern matching.

## II.NATURAL LANGUAGE PROCESSING

It is the process of producing meaningful phrases and sentences in the form of Natural Language. It involves text Planning, Sentence Planning and Text Realization. Text planning includes retrieving the relevant content from knowledge base. Sentence Planning includes choosing required words, forming meaningful phrases and setting tone of the sentence. Text Realization is the process of mapping the sentence plan into sentence structure.There are two broad types of Natural Language Processing (NLP) algorithms: traditional and deep learning .

Traditional NLP involves a lot of coding. You have to anticipate words and phrases a user could say in each context, recognize patterns of speech, extract predefined keywords, and so on. In other words: There are a lot of rules, a lot of regular expressions, and a lot of hard coding. It takes a lot of time to write and debug these programs. And the rules quickly break down when the user deviates from the expected path. This is the reason many bots feel so dumb.

We have to build millions of intelligent virtual assistants capable of maintaining natural language dialog. Can we do it with traditional NLP? The answer seems to be no. It's too hard and too time-consuming. The world doesn't have enough programmers to do it.

Deep Learning is a concept in development. With the advent of GPU's, Deep Learning has come closer to becoming a reality. However, as of today we still have to use the traditional methods of Natural Language Processing.

The core approach in the method is the use of dynamic programming to efficiently parse/tag a piece of text. So, the algorithm derives parse tree for a text by iteratively adding edges to chart or graph, where each edge represents a hypothesis about the parse tree structure for a subsequence of the text, and the chart itself represents a substrate for composing and combining the given data and producing a response.

When the algorithm begins parsing a text, it creates a new chart, spanning the text. It then incrementally adds new edges to the chart, where a set of the determined or discovered grammar rules specifies the conditions under which new edges should be added to the chart. Parsing is completed once the chart reaches a stage where none of the rules allow the addition of any new edges.

## III.CHATBOTS

A Chatbot is a computer program that tries to conduct a natural language conversation with a human user. It lets the user ask normal questions and statements such as "How are you today?" and "How is the weather today?" and then responds in a similar natural manner . The ultimate aim for chatbots is when the user cannot tell the difference between talking to a chatbot and talking to a real human. This will mean that the user will pass the Turing Test. However, we have not yet ascended that peak yet. Turing Test is also known as the Imitation Game. In this test, the goal for the chatbot is to maintain a conversation which is indistinguishable from a human conversation. The usual way to apply the test is that there is a human observer (judge), who is asking questions or having a conversation with someone over the computer link. That someone can be a computer that is chatbot or a person. If on the other side there is a chatbot and the judge would think it is a person, then the chatbot would pass the test.

Chat boats are typically used in dialog systems for various practical applications including customer service or information acquisition. Some chatterbots use sophisticated natural language processing systems, but many simpler systems scan for keywords within the input, then pull a reply with the most matching keywords, or the most similar wording pattern, from a database.

Currently the biggest challenge that existing chat-bots have is maintaining of the context and understanding the human inputs and its responses. Most of the existing bots still work just on the pattern matching of inputs and then trying to find a scripted response which matches the input. But this approach cannot result in a fully satisfying conversation or lead a conversation with some specific purpose.

The historic chatbots are ELIZA (1966) and PARRY (1972). More recent programs include A.L.I.C.E., Jabberwacky and D.U.D.E. While ELIZA and PARRY were used exclusively to simulate typed conversation, many chatbots now include functional features such as games and web searching abilities. Recently, with the advent of smartphones, chatbot virtual assistants such as Google Assistant, Siri, Amazon's Alexa and Bixby have kickstarted the golden age of chatbots.

## IV.TECHNIQUES FOR NATURAL LANGUAGE PROCESSING

### 4.1 Pattern Matching

This is the most common technique and approach used in chatbots. Variations of some pattern matching algorithm exist in every existing chatbot system.

### 4.2 Parsing

Textual Parsing is a method which takes the original text and converts it into a set of words (lexical parsing) with features, mostly to determine its grammatical structure. On top of that, the lexical structure can be then checked if it forms allowable expression (syntactical parsing). The earlier parsers were very simple, looking for recognizable keywords in allowed order. Example of such parsing would be that sentences "please take the gold" and "can you get the gold" would be both parsed into "take gold". With this approach the chatbot with a limited set of patterns can cover multiple input sentences. The complicated parsers used in latter chatbots do the complete grammatical parsing of the natural language sentences.

### 4.3 A.I.M.L

To build a Chatbot, universal language is needed which should be flexible and easy to understand. AIML, a derivative of XML is one of the widely used approaches that satisfies the requirements. AIML represents the knowledge inserted into Chatbots and is based on the software technology developed for A.L.I.C.E. (the Artificial Linguistic Internet Computer Entity). It has the ability to characterize the type of data object (AIML objects) and describe partial conductance of the programs that it processes. These objects consist of two units: topics and categories; the data contained in these categories is either parsed or unparsed.

### 4.4 Deep Learning

Deep learning is subset of machine learning where a machine uses huge amounts of data and highly complex algorithms to 'learn' and to simulate human-like decision-making. Deep learning is the deeper part of the AIs brain and the conduit to the NLP allowing the machine to learn from everything and improve upon itself for next time, just as humans would.

## V.APPLICATIONS AND FUTURE OF NLP

The examples given below show a few of the methods in which a chatbot may be deployed. A chatbot can act as a guide to a website. In it's most primitive form this can assume the mold of a simple descriptive guiding hand to the subject matter of all the pages – highlighting the important content of each page. In it's most complex form this can be a complete chatbot proficient to handle the questions that a person visiting may inquire the site, the organization, or its products or services – feeding the visitor with the web pages relevant to the inquiry. This can be a useful method to improve the experience for those who are visually impaired without compromising the accessibility of the website.

### 5.1 Machine translation

Machine Translation transcribe from one natural language to another. This problem is part of a class of problems requiring complete various types of education that humans have such as grammar, semantics, certainty about the world we live for solving the problems.

### 5.2 Named Entity Recognition (NER)

Given a stream of text, figure out which items in the text map to appropriate names, such as people or places, and what the category of each such name is (e.g. person, place, organization). note that, although capitalization can help in determining named entities in languages such as English, this dossier cannot aid in recognizing the type of named entity, and in any case is usually inappropriate or insufficient.

### 5.3 Frequently Asked Questions (FAQ)

Guide during a Site Guide will be present on many pages on a site, a FAQ Guide will only be present on the Questions and Answers page of a site. As knowledge has earlier been assembled in question and answer form for the page it is simple to translate this in to cases for the chatbot.

### 5.4 Support Agent

Even more specialized appearance than Site Guide is where the chatbot is working as a virtual support agent. Here the chatbot would have access to your support knowledge base and answer any questions the user has with the correct reply from the knowledge base. The bot can also give links to relevant web pages, display diagrams, or provide files for download to aid the user. If the user needs human help the chatbot can use email or SMS to alert a live support person. Virtual Sales Agent The Virtual Sales Agent is optimized

to detect customer needs, match them to a solution, provide information and navigate them to your ordering system. As well as accessing web pages and alternate documents about your products and services, the chatbot can also access stock and pricing information from your current systems.

### 5.5 Form Guide

Many users think form filling a intimidating activity. An oft quoted figure is that successful form completion halves for each and every question on the form. Chatbots can achieve two roles to help users' complete forms. First they can be on hand to answer user questions as they fill out a form. Second they can actually transmit out the form-filling themselves, interrogating each question in turn – constructing form-filling more enjoyable experience.

### 5.6 Survey Taker

Natural elongation of the form-filling role, chatbots is an excellent way to carry out simple surveys on your web site. The chatbot can ask the questions, store answers, and use logic to ask questions related to previous inputs.

### 5.7 E Learning Tutor

On a more sincere note, chatbots are proving to have a symbolic effect in the field of e-learning. Many conventional e-learning systems have fallen short of assumption because of the relatively sterile nature of the experience – there is no bond between student and computer in the way that there is between student and teacher.[9][10]

## VI.PROPOSED SYSTEM

As we know the use of automated machines is increasing day by day i.e., the machine-human interaction is also increasing day by day. The case might as well come true where one day humans would be replaced by automated bots for performing specific functioning of the system. Where, a central system will be a chatbot which will receive the input from the user which will be in the form of an audio/ voice. The input can be received using a mic and then this voice input will be converted to the corresponding textual input using pre-existing python libraries.

Then this textual input will be provided to the system which consists of the chatterbot which will perform Natural Language Processing on the provided input and search within its database for the query. Two cases will be generated:

(1) When the input entered is an identifiable query, to which already an answer is known to the system.
This corresponding answer will be directly provided to the user in this case.

(2) When the input entered is not recognized, this is where the intelligence of the bot comes into play and the system will understand the input and answer it using its intelligence. This is termed as "Artificial Intelligence".

International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

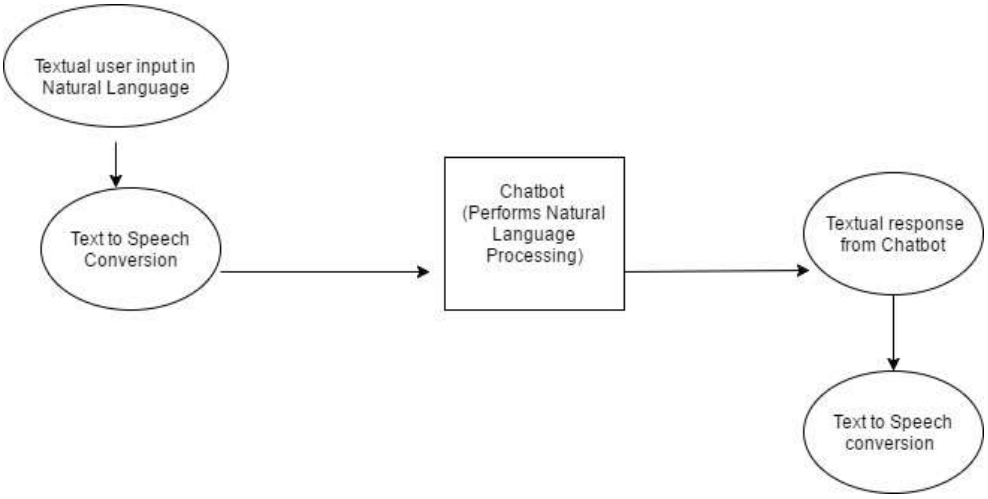


**Figure No. 1: System Architecture**

Once the system knows what output is to be given to the user, the output is generated which is in a textual format. This textual format is then converted to voice using pre-existing libraries and a voice note is generated. This voice clip will be given to the user as an output.

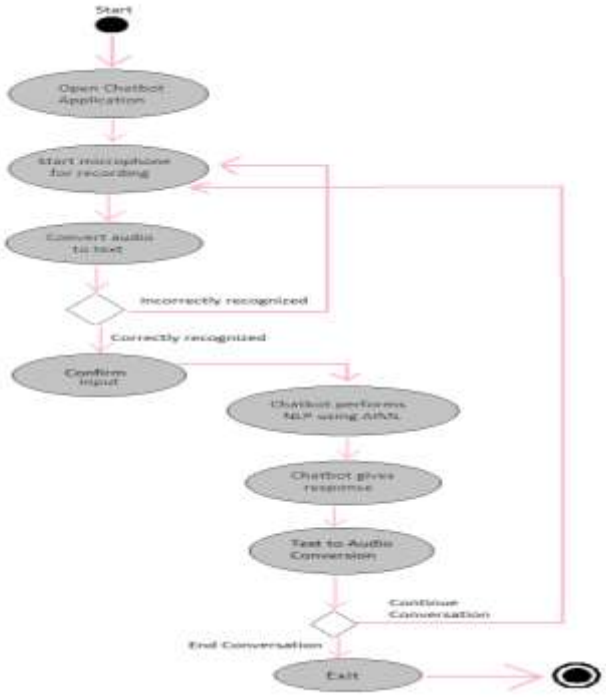


**Figure No. 2: System Flowchart**

## VII.CONCLUSION

The idea of chatbots has been around for a long time. Only now are they emerging from the universities and research laboratories and becoming ready for primetime use. Apart for the general advances in technology there are three main reasons why we believe that Chatbots are now ripe for exploitation by businesses:

First the technology to deliver chatterbots as dynamic and speaking avatars is now ready for mass-market use making chat bots a far richer and engaging experience than lines of text on a screen.

The richer bots (with animation facilities) require fair amount of bandwidth for serving all its purposes. The widespread adoption of broadband by businesses, and domination of broadband in the domestic market means that this richness can now be delivered across the net to almost any user.

In future the bot will be able answer accurate and somehow it can also replace humans in live support that will save lot of infrastructure and resource cost. User analytics can have the way to track where the bot did not help and analyze over time.

Thus, to conclude, demand for chatbots will keep on increasing day by day and there will be a huge scope for them in the near future.

## REFERENCES

[1] Hirschberg, Julia, Bruce W. Ballard, and Donald Hindle. *"Natural language processing."* AT&T technical *journal 67.1(1988): 41-57.*

[2]Brandao, Cesar, Luis Paulo Reis, and Ana Paula Rocha. *"Evaluation of Embodied Conversational Agents."* Information Systems and Technologies (CISTI), 2013 8th Iberian Conference on. IEEE, 2013.

[3]Angga, P. Antonius, et al. *"Design of chatbot with 3D avatar, voice interface, and facial expression."* Science in Information Technology (ICSITech), 2015 International Conference on. IEEE, 2015.

[4]Satu, MdShahriare, and MdHasnatParvez. *"Review of integrated applications with AIML based chatbot."* Computer and Information Engineering (ICCIE), 2015 1st International Conference on. IEEE, 2015.

[5]Hameed, Ibrahim A. "Using natural language processing (NLP) for designing socially intelligent robots."*Development and Learning and Epigenetic Robotics (ICDL-EpiRob), 2016 Joint IEEE International Conference on*. IEEE, 2016.

[6]Shawar, Bayan Abu, and Eric Atwell. *A comparison between ALICE and Elizabeth chatbot systems*. University of Leeds, School of Computing research report 2012.19, 2012

[7]Yin, Fu-Lian, et al. *"Deep neural network language model research and applicationoverview."* Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2015 12th International Computer Conference on. IEEE, 2015.

[8]Merrouni, Zakariae Alami, Bouchra Frikh, and Brahim Ouhbi. *"Automatickeyphrase extraction: An overview of the state of the art."* Information Science andTechnology (CiSt), 2016 4th IEEE International Colloquium on. IEEE, 2016.

[9]Thomas, N. T. "An e-business chatbot using AIML and LSA."*Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on*. IEEE, 2016.

[10]Perera, Maria João, et al. *"Chatbots' Greetings to Human-Computer Communication." arXiv preprint arXiv:1609.06479* (2016).