

**YOOFI BROWN-POBEE**

**49142019**

**BUILDING A DOMAIN SPECIFIC TASK BASED DIALOG SYSTEM FOR  
CHALKBOARD EDUCATION**

## **Chapter 1: Introduction**

Many businesses in Ghana make use of customer service call centers to respond to the queries of customers and resolve any issues they may have. A simple search on-line would produce a list of jobs available said call centers. In principle, these centers are very instrumental to the business operations as they serve as a contact point to customers who are revenue drivers for the business. However, practicality dictates otherwise. In some cases, the numbers provided for customer service are no longer in service (some have been changed but not updated), in other cases the call experience is marred by the emotional state of calling customers or the receiving agents at a given point in time. Even beyond these issues exist a larger problem of cost. For companies, as they scale up and acquire more customers, there is a need to hire more customer service personnel to cater to the needs of customers.

An already existing solution to this problem is the use of automated conversational systems to respond to the queries of customers. Intelligent systems exist that can quickly serve numerous customers at any given time before deciding if human intervention is needed. An example of a working solution is MTN's Facebook chatbot. Querying this chatbot exposed some flaws as its responses as it could not properly understand what was being asked hence disbursing unsuitable information for the given domain. A domain in this case is simply the scope of questions, statements and classes that all communication can be categorized under.

This project's main contribution is to create an intelligent chatbot that can give suitable responses to a wide range of customer queries in a given domain. Chalkboard Education, a startup that hosts an education platform would be used as the host company. Chalkboard education allows schools to make all their content digital and enrolls their students onto a platform to access it. This results in many schools and individual students constantly calling and direct messaging Chalkboard's customer service support line for help. The company currently have only one person available for this job. By the end of this project, the goal is to build full-functioning chat bot to handle all of Chalkboard's support issues.

## **Background & Motivation**

The use of conversational agents and dialog systems can be traced as far back as 1966 with the development of ELIZA, an early natural language processing computer program to present day Siri and Alexa of Apple and Amazon respectively. Natural Language Processing and its related fields have provided a means for us humans to communicate with machines and machine systems in ways similar to how we do with each other. There are many systems with differing implementations and nuances but majority generally fall in two classes: Task oriented dialog agents and Chatbots .

The former is predominantly concerned with specific tasks and designed to have as very little interactions as possible to accomplish the task [Jurafsky and Martin, 2014]. Examples include digital assistants like the aforementioned Siri and Alexa. This class of dialogue systems are mostly employed by companies on their websites and products to enable customers and users address problems and answer questions. Key benefits of these systems especially to businesses and their customers include faster and more convenient query handling as well as lower costs relative to hiring of human agents. They are not designed to have prolonged conversations, however, as chatbots, the second class of dialogue systems, are designed for that.

Chatbots are set up to mimic the natural conversational characteristics of human beings. They are designed to have more casual and less directed conversations than task oriented dialog agents. Cleverbot is a chatbot capable of carrying on prolonged conversations with humans and exists mainly for entertainment purposes [Carpenter, n.d.]. These are more suited for social uses such as for psychological analysis and entertainment purposes. A form of evaluation for many chatbots is a test called the Turing test developed by Alan Turing in 1950. A Turing Test is a method of inquiry for determining whether or not a computer/system is capable of thinking like a human being[2]. A human should be able to find the two systems indistinguishable in order to pass the Turing test. A chatbot that is able to pass this test can be said to be a good one.

The existence of the above systems in the field of natural language processing have made it possible for the application of knowledge to different scenarios and domains. Chalkboard Education's particular situation presents an opportunity to apply this knowledge in a local context. The startup has been running for close to three years and current has over 4000 students enrolled on its platform. Customer queries are handled mainly through WhatsApp and voice phone calls and require employees to be present to respond at all times. The creation of a dialogue system on Chalkboard specific information and queries appears a potential solution to solve the problem by removing the need for the physical presence of a human and saving cost for the startup. As the startup scales, the system does not have to scale proportionally hence cost of customer care can be minimal for a given number of users.

### Related Work

#### ELIZA

ELIZA is an early natural language processing computer program created from 1964 to 1966 [Ireland, 2012]. It was created to demonstrate communication between humans and machines using natural language and was created by Joseph Weizenbaum at the MIT Artificial Intelligence Laboratory. It works by breaking down a sentence and ranking its keywords and transforming the users sentence into a suitable response using rules from its preprogrammed learning script. The system was modelled after a Rogerian psychotherapist in that it allowed the program to not need to have knowledge base of the topic the human was discussing [Weizenbaum, 1966]. By simply transforming the sentences based on key words in the users sentence, the program can appear to be engaging in conversation regardless of the topic. There were a number of reports of human test subjects developing an emotional connection to the system; a testament of how well it was able to sustain a regular conversation. This system formed the basis for the creation of other conversational agents and improvements in the field of natural language processing.

## Cleverbot

Cleverbot is a web based chat bot that uses artificial intelligence to have unsupervised conversations with humans and was created by Rollo Carpenter. It holds the distinction of having passed the Turing test. As mentioned earlier, the idea of the test is for a machine to pretend to be a human and will only pass if this pretense is found to be convincing [Gehl, 2013]. Cleverbot has performed quite well in Turing Test competitions, giving it some credibility [Gehl, 2013]. Unlike ELIZA, Cleverbot actually learns how to have a conversation as it interacts with more humans. It's responses at any given time are as a result of an analysis of previous conversations it has had with other humans. This contributes to its ability to perform well in the Turing Test. Despite the difference in how they are implemented, they are both capable of holding a conversation with a human for a considerable length of time.

## Dialogue System Evaluations

Dialogue System Evaluations are crucial to determining the quality of a dialog system by assessing its responses against human generated supervised signals such as user satisfaction scores [Walker et al., 1997]. Dialogue system models which are optimized for supervised objectives are known as supervised dialogue models and those that are not known as unsupervised dialogue models [Liu et al., n.d.]. The former requires the large scale collection of labels which can be often expensive while the latter's major concern is how to perform the evaluation automatically. N-gram based measures like BLEU exist but are found to be a poor fit for dialogue models because two different responses may have no overlap but would be good replies to a given statement [Kannan and Vinyals, 2017]. There exist many models to evaluation dialogue models like ROUGE and adversarial models to suggest the quality of a dialogue system.

## Plan for requirements analysis

- Obtaining data from Chalkboard on previous interactions with customers
- Classification of data into appropriate classes for responses.

- Ability to accept user input
- Ability to display output from task dialog system

## Requirement Specification

### Overview

The chapter seeks to give an analysis of the functionality the task dialog system would offer as well as the scope of its capability. The application would make use of frequently entered user inputs to learn and generate appropriate responses to user queries. This is to be done without human supervision. Requirements would be obtained from Chalkboard Education as well as literature such as Speech and Language Processing by Daniel Jurafsky to identify the necessary components for building a task dialog system.

### User Identification and Use Case

The application would be used primarily by students of schools the use Chalkboard Education's platform to manage their content. The service would run on Whatsapp and on Facebook Messenger hence users would need Internet access and a subscription to a mobile carrier. Some use cases are outlined below to better understand how the application will be used:

1. A student of a newly added school (newly added to Chalkboard's platform) attempts to access the platform but does not know how to log in.
2. A student of a newly added school is able to log in but has no knowledge of how to navigate the site and access resources.
3. A student of an already existing school on the platform is unable to find specific course material on the platform.

### Requirement analysis

The main function of the task dialog system is to provide appropriate generated answers to user queries through a Whatsapp Business line and Facebook Messenger without human

supervision. A key input is a set of previously asked queries to enable the system learn queries and match them to appropriate responses. All use cases of the would involve the user querying the system.

### Scope of the Project

The project would consist of a task dialog system operated through WhatsApp Business and Facebook Messenger. The front end of these applications would be used to receive user inputs,

### Functional Requirements

#### 1. Accepting user input

- 

Users should be able to type in any problems they have or queries they need answer

#### 2. Generating appropriate responses

- 

The system should be able to read the queries as input and generate a response that best answers the query

#### 3. Displaying generated responses

- 

The user should be able to read the generated response. The response should be in a form easily understandable by the user.

1. The system should be able to process queries that have the same idea but appear in different forms and contexts

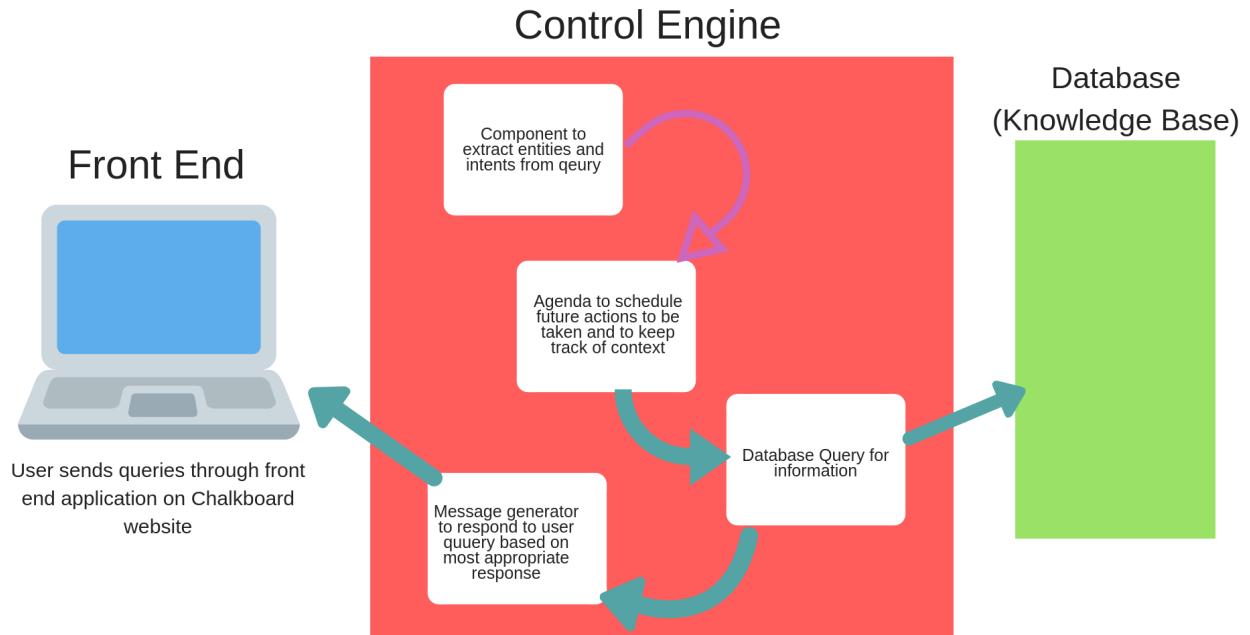
2. The system should be able to learn from queries it has not processed before to be better able to handle repeat new queries.

3. The system should Non-Functional Requirements be able to respond to a few non task oriented queries to maintain some level of interactivity with the user.

4. System needs to be able to remember responses given within the same query session and refer to previous information given within session to make a decision when necessary.

## Chapter 3: Architecture and Design

### High Level Architecture of Project



### Key Modules in the Architecture/Design

#### Front end web application

Chalkboard Education already have a website built using HTML, CSS and JavaScript. I will create an applet to run on the website to serve as the front-end that users would use to query the database for information. Like the original website, this front end would be built using HTML, CSS and JavaScript. The users query would be passed from the text box to the control engine which would then classify the issue and query the database for a response and return a response to the user.

#### Database (Knowledge Base)

This would serve as the knowledge source from which the dialog system would

- Classify user queries
- Determine appropriate responses



- Generate responses too be returned to user

The database would be a collection of formatted text files with sourced from past user queries and frequently asked questions in the past. They would also feature multiple documents to help the chatbot with classification. Also in the database would be question-answer pairs for queries that have more straightforward responses that do not require more complex logic and decision making on the part of the control engine. Since there exists a very limited scope of problems users might have pertaining to the use of Chalkboard's platform, the database would of a small size. This gives room for adding datasets that would enable the dialog system have 'small talk' conversations with users as a means to increase the user experience of the application.

### **Control Engine**

Since the dialog system has multiple knowledge based on which to evaluate queries, there needs to be some engine that decides what task to perform next. The engine would create an agenda, which would store potential processes to be undertaken based on user queries. It would examine the agenda, decide what needs to be done next and perform the function. The engine would also be responsible for maintaining context throughout the interaction with the user such that information from earlier in the conversation could be made available at a later stage. This would be necessary for maintaining flow in the conversation. The control engine includes a mechanism for extracting intents and entities from user queries and a message generator responses based on feedback from the database.

## References

[1] Daniel Jurafsky and James H. Martin. 2014. *Speech and language processing*, Harlow: Pearson.

Carpenter. Cleverbot. Retrieved October 9, 2018 from <http://www.cleverbot.com/>

[2] Anon. What is Turing test? - Definition from WhatIs.com. Retrieved October 10, 2018 from

<https://searchenterpriseai.techtarget.com/definition/Turing-test>

[3] Corydon Ireland. 2012. Alan Turing at 100. (September 2012). Retrieved October 10, 2018 from

<https://news.harvard.edu/gazette/story/2012/09/alan-turing-at-100/>

[4]Robert Gehl. Teaching to the Turing Test with Cleverbot. *The Journal of Inclusive Scholarship and Pedagogy*24, 1-2.

Anjuli Kannan and Oriol Vinyals. 2017. Adversarial Evaluation of Dialogue Models. (2017).