RECALL: A SCHEDULING SYSTEM AND QUESTION ANSWERING SYSTEM WITH USER KNOWLEDGE BASE USING KEYWORD, SYNONYM, AND RULE-BASED APPROACHES

A Capstone Project Presented to the Faculty of the

College of Information Computer and Communications Technology

University of San Jose-Recoletos

Cebu City, Philippines

In Partial Fulfillment

Of the Requirements for the Degree of

Bachelor of Science in Computer Science

By

Sanchez, Kenneth

March 2018

ENDORSEMENT

This project entitled "RECall" prepared by the following student: SANCHEZ, KENNETH E. in partial fulfillment of the requirements for the degree of BACHELOR OF SCIENCE IN COMPUTER SCIENCE MAJOR IN DATA ANALYTICS, has been endorsed and is recommended for the acceptance and approval for ORAL EXAMINATION.

MRS. JOSEPHINE E. PETRALBA, MBA
Adviser

MRS. MARISA M. BUCTUANON, MCS
Coordinator

MR. JOHN LEEROY A. GADIANE
Member

MRS. JOVELYN C. CUIZON, MSIT
Chairperson, CS/IT Department

DR. GREGG VICTOR D. GABISON, DMHRM, DIT
Dean

APPROVAL

Approved by the Tribune for Oral Examination with the grade if **PASSED**.

MRS. JOSEPHINE E. PETRALBA, MBA
Adviser

MRS. MARISA M. BUCTUANON, MCS
Coordinator

MR. JOHN LEEROY A. GADIANE
Member

MRS. JOVELYN C. CUIZON, MSIT
Chairperson, CS/IT Department

DR. GREGG VICTOR D. GABISON, DMHRM, DIT
Dean

ACKNOWLEDGMENT

For the enlightenment and blessings, thank you Almighty God.

To my family, I am grateful for the support and hard work upon providing my needs and sending me to a private school through the years. And relatives, stay in good health.

To my teachers, colleagues, supervisors, peers, strangers, and people around the world, I appreciate the little things, entertainment, learnings, inspirations, stories, and experiences that I have encountered.

To my friends, close friends, thank you for the memories that we shared, and I hope that you are all doing well.

As a new chapter will begin to unfold to my life, upon fulfilling this achievement, I hope I can bring a better change to this world. To my future self, if you ever read this, do not hesitate to move forward. Adelante.

Special Mentions:

To Junn Dobit Paras, my co-researcher in “RECALL: A MOBILE APPLICATION FOR REMEMBERING AND RECOVERING INFORMATION” paper, thank you for bringing the idea on this project, and for making the team possible.

To Ma’am Josephine Petralba, our research adviser, thank you for the encouragement, for the advice, and for defending us and the project on the panel.

To Ma’am Lorna Miro, thank you for the encouragement, for the teachings, and for taking care of us especially the CS group.

ABSTRACT

Question answering aims to answer a natural language question. The challenge, however, is how to deal with the same questions that have different ways of expressing it. Knowledge-based question answering makes use of a knowledge base as an information source where the data is structured. Structuring the knowledge representation will be essential for generating answers later on, which is done by information extraction, for a knowledge base containing logs, schedules, and medications of the user, for this study. The question answering focuses an open domain, and it uses keyword, synonym, and rule-based approaches. While, scheduling allows adding schedule, time-based moving schedule, entity-based or time-based canceling schedule, and recording medication with calendar app integration. Given a dataset of 100 knowledge and 100 questions, the question answering evaluated about 79% accuracy for classifying correct, incorrect, and null answers out of 95 valid questions, with 53% and above precision, recall, and F-measure. Although there are more rules constructed upon resolving rules in conflict, with a solid foundation and combining other approaches, it leaves the possibility of being open to more cases. Nevertheless, the study has achieved its purpose to develop a mobile application for scheduling and question answering systems with REST web services.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

RATIONALE OF THE STUDY

Memory is the faculty of the mind that encode, store, retain, and subsequently recall information and past experiences in the human brain. Humans can store a vast amount of information in memory, though we cannot avoid to forget things and can remember certain things. Fundamentally, we transfer the information on a physical material through writing or encoding them, similarly, to a computer storing data on its memory permanently, in a way to help us remember. Besides storing and retrieving data, humans had been creating machines to imitate human behavior, since the birth of artificial intelligence (AI).

AI has paved its way to make machines learn our natural language. These are natural language processing (NLP) applications, which is a tool for making the human-computer interaction possible to do tasks. There are now digital assistants like Alexa, Cortana, Google Assistant, and Siri, able to aid users to perform actions using speech commands. Question answering (QA) system retrieves and processes information into an answer through a natural language question. Over the years, there are different approaches towards QA such as statistical, pattern, or templated based [1].

Using speech that focuses strictly on pattern to distinguish a specific action, i.e. “set event” as command, and giving details, after, in a sequential input, but not “I have an event from <time> at <place> with <person>” in one interaction, becomes inconvenient, since users need to know these set of commands before they could use it. This speech processing makes it unable to recognize the user’s intention in natural language and undesirable. Most of the studies [1] focus on general knowledge which brings this study

to focus on the user knowledge prioritizing logs, schedules, and medications on his day-to-day life. Handling calendar events require more steps to manage for the user even with the digital assistant and calendar application. There is a lack of scheduling (SCH) system, as an NLP application, to create, update, and delete events on the calendar while also checking for conflict.

This paper aims to develop a scheduling system and question answering system with user knowledge base using keyword, synonym, and rule-based approaches. The user can record log, can add or move or cancel schedule, can record medication, can ask a question to retrieve answers, and can manage reminders. With these features, the users will be able to refresh their memory with ease, to keep track of their activities, and to do more and apply it on their everyday lives.

THEORETICAL BACKGROUND

Natural Language Processing

Humans and computers have their languages, and they cannot directly communicate with each other. NLP is a tool for making this human-computer interaction possible, either in speech or text input. NLP is defined as the automatic (or semi-automatic) processing of human language [2]. Some of NLP applications are information extraction and question answering. The human language, like English, has its corpora or collections of written or spoken text, and grammar. Before almost any higher level of NLP, the text undergoes normalization, part-of-speech (POS) tagging, and parsing. Whereas, context-free grammar (CFG) is the most widely used formal system for modeling constituent structure in English and other natural languages. And, regular expressions (RE) is commonly used for text processing by many computer language and word processing tools [3].

Text Normalization is the process of transforming text into a single canonical form. This process involves tokenization, normalization, and lemmatization of the words from the text. Tokenization breaks the text into basic units like words. Normalization converts words like “USA” and “The US” or “mister” and “Mr.” into one same form, and it also considers case folding and clitic (‘) contractions. Lemmatization finds root like “sang” or “sung” are forms of the verb “sing” [3].

Part-of-Speech Tagging is the process of assigning POS marker to the word tokens. In English, there are eight main parts of speech (also known as word classes): adjective, adverb, conjunction, interjection, noun, preposition, pronoun, and verb. One commonly used tokenization standard is Penn Treebank (see appendix A) [3].

Chunking, an alternative style of partial parsing, is the process of identifying and classifying the flat, non-overlapping segments of the sentence that constitute the basic non-recursive phrases corresponding to the main content-word parts-of-speech: noun, verb, adjective, and prepositional phrases. This process does not require complex and complete parse trees [3].

Context-Free Grammar, in the theory of formal language, consists of a set of rules or productions used to generate a pattern of strings. The formal definition of a context-free grammar G is defined by a 4-tuple (Ω, Σ, P, S), where:

- Ω – set of non-terminal symbols (or variables)
- Σ – set of terminal symbols (disjoint from Ω)
- P – set of rules or productions, each of the form $A \rightarrow \beta$

  where A is non-terminal,

  β is a string of symbols from the infinite set of strings $(\Sigma \cup \Omega)*$
- S – start symbol and member of Ω

Common types of sentence-level grammatical constructions in English are declarative, imperative, yes-no question, and wh-question. These can be modeled with CFGs and help implement parsers [3].

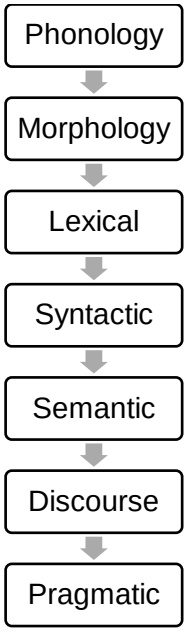

Figure 1 - Levels of NLP

Figure 1 represents the levels of language that NLP undergoes. Ambiguity is one of the major problems of natural language. It usually occurs at the syntactic level, and it has a subtask like lexical and morphology which are concerned with the study of words and word formation. Each of these levels can produce ambiguities that can be solved by the knowledge of the complete sentence [4].

Regular Expressions is an algebraic notation for characterizing a set of strings. Given a pattern expressed in RE, searching for string occurrences from the text can be done. Or, the string format can be checked by doing a full match with it [3].

Information Extraction

A lot of information is found by people on articles, journals, books, documents, scripts, or any form of written text. While, information retrieval can acquire information from

a large amount of text, it only retrieves relevant documents through term and inverse document frequency using vector models. Information Extraction (IE) distills knowledge or structured data from unstructured text by identifying references to named entities and the relationship among the entities [5]. Its main task is named entity recognition. It also involves temporal expression extraction [3].

Named Entity Recognition (NER) is the process of classifying the type of entity from spans of text that constitute proper names. A named entity is a proper name of a person, a location, an organization, etc. within the span of text [3].

Temporal Expression Extraction is the process of finding and converting time expressions into a standard format. Temporal expressions are those that refer to an absolute point in time, relative times, durations, and sets of these. Exact calendar date and time of day are absolute temporal expressions which are mapped directly. While indiscrete time expressions like the day of the week are relative temporal expressions. Lastly, durations denote a period at varying levels of granularity such as time units [3].

Knowledge-Based System

Knowledge-Based System (KB) takes knowledge from a human and stores it in such a way that it can reason with it. Commonly, KB is associated with expert systems and comprising of a subsystem, the inference engine, that applies logical rules to a knowledge base to deduce new information. The early rule-based expert systems represented facts about the world as simple assertions in a flat database and used rules to reason about and, as a result, added the rules to these assertions. The techniques representing the knowledge base became more sophisticated as knowledge-based systems became more complex. Rather than representing facts as assertions about data, the knowledge base developed its structure representing information using similar techniques to object-oriented programming such as hierarchies of classes and

subclasses, relations between classes, and behavior of objects. The independent rules and interactions within the knowledge base itself make the reasoning occur as the knowledge base became more structured [6].

Question Answering System

Question Answering, one of the applications of NLP, aims to answer a natural language question. It has two paradigms that differ on its information source; first, the Information-Retrieval-Based or Text-Based Question Answering, used on early QA, which uses textual information on the web or in collections such as documents, while the other is Knowledge-Based Question Answering (KBQA). KBQA processes a language question into an answer by mapping it to a query over a structured database. Most QA systems focus on factoid questions. Factoid questions are questions that the answers are simple facts expressed in short text answers [3].

There are two types of a domain for QA which is the open domain and closed domain. Closed domain types specialize on a type of information on one domain, which defines the scope of the QA system, while the open domain has a broader scope [1].

The question answering architecture involves three main tasks or phases which is the question processing, passage retrieval, answer processing. The question processing phase is composed of the question classification and answer type detection. After classifying a valid question, the answer type is the named entity to be detected based on the question ontology or answer type taxonomy. It also determines the query arguments for passage retrieval. The passage retrieval is a task for finding the relevant text from the information source to be processed as candidate answers. The answer processing is composed of the answer generation and validation. It extracts the answer from the passage [3].

REVIEW OF RELATED STUDIES

KBQA: Learning Question Answering over QA Corpora and Knowledge Bases

A KBQA that uses a template-based approach said that it has effectively boosted the coverage of Resource Description Framework (RDF) knowledge base 57 times with 27 million templates over 2782 intents. Their KBQA distinguishes itself from previous QA in four aspects: understanding questions with templates, using template extraction to learn the mapping from templates and predicates, using expanded predicates in RDF to expand the coverage of knowledge base, and understanding complex questions to expand the coverage of the questions [7]. The researchers categorize the previous approaches into rule-based, keyword-based, and synonym-based. They said that the rule-based approach becomes infeasible as the manually constructed rules grew as large as the number of questions (high precision but low recall for a variety of questions). While the keyword-based approach can only map one predicate representation in the KB and fail for other representations since natural language is diverse. And, the synonym-based approach is limited to understand the semantic similarity of the word as a unit without relation to other words, which fails on paraphrases since phrases can alter the original text but the meaning as a whole stays the same. Even so, their template-based approach uses large corpora, to achieve that feat; which, this study avoids, because not all knowledge from the large corpora is utilized for personal use since, practically, the user only needs his information; a user knowledge base for QA system. The templates and rules are quite similar in terms of number, but general rules will be used to limit its number and combine it with other approaches to expand its scope of use.

Large-scale Simple Question Answering with Memory Networks

A similar study develops an open-domain QA, that aims at providing exact answers to questions formulated in natural language without restriction of a domain, using memory

networks which consist of memory and neural networks. The researchers study the impact of multitasking and transfer learning for simple question answering; a setting for which the reasoning required to answer is deliverable, as long as one can retrieve the correct evidence given a question, which can be difficult in large-scale conditions [8]. Though this project does not incorporate a training-based approach, it accepts general knowledge from a user. However, to cope the difficulty on handling the large-scale conditions, limiting the passage retrieved within the specified time or next recent date traversals will recall at least an answer closer to the memory of the user since most of the knowledge relates to time.

Question answering from the web using knowledge annotation and knowledge mining techniques

Another study is about QA using a web-based approach that takes advantage of the vast amount of information on the web and checks on real-world user queries. Aranea, an open-domain question answering system, extracts answers from the Web using two different techniques: knowledge annotation and knowledge mining [9]. They describe knowledge annotation as an approach using semi-structured and structured web sources occurring on frequent questions, and knowledge mining using redundancy-based statistical techniques. Using a search engine as information retrieval backend, they were able to focus on answer extraction. Even so, this study avoids using a search engine since the passage retrieved becomes broader and involves more processing.

Natural Language Annotations for Question Answering

START, similarly, a web-based QA, answers questions through natural language annotations drawn from a set of structured, semi-structured, and unstructured multi-media information sources. START analyzes these annotations in the same fashion as any other sentences, but the system produces pointers from creating representational structures to

the information segments summarized by the annotations [10]. They make use of Omnibus, an external gazetteer for source-specific terminology, and synonyms; on the other hand, this project will use Wordnet as the dictionary corpus for the synonyms and identification of the entity in nouns instead of annotations.

A WordNet-based approach to Named Entities recognition

There is also a study that uses a WordNet-based approach to Named Entity Recognition (NER). Their NER relies on the combination of 200 language-dependent rules with a set of predicates, defined on the WordNet hierarchy, for the identification of both proper and trigger words [11]. The researchers found a total of 12,259 classes and 3,876 instances over the WordNet hierarchy with the entities for a person, location, organization, measure, money, duration, date, time, percent, and cardinal. This project will use WordNet as the dictionary corpus for the synonyms and identification of the entity in nouns.

PROJECT OBJECTIVES

This study intends to build a mobile application for scheduling and user knowledge-based question answering system. Specifically, it aims to:

- To capture the entities and intent from unstructured text or natural language input.
- To answer correctly with the specific information the user wants to retrieve from the knowledge base.
- To perform scheduling processes integrated with the calendar application.
- To explore user knowledge base, and extraction and retrieval methods with natural language processing.

PROJECT SCOPE AND LIMITATIONS

This study develops a mobile application that requires an Android operating system v.4.2 (JellyBean) and above. The accuracy of the speech recognition software and microphone hardware used in or out a noisy environment is out of the project's scope. There is no strict grammar for accepting natural language input, but some rules may apply. Compound and complex sentences and appositive phrases and verbals are not fully supported yet in this study. The system supports English language, direct questions and logs, and simple time expressions. It will not implement heuristics or discourse analysis. A simple QA, using keyword, synonym, and rule-based approaches, with no deep semantics will be developed for retrieving the answers, and cannot answer all types of questions as natural language is diverse. And, the SCH works for time-based and entity-based, and it cannot swap two schedules.

SIGNIFICANCE OF THE STUDY

QA systems were developed by humans with the goal of creating machines able to answer questions in our language. Firstly, choosing its domain determines how it can be useful as an application. However, in terms of usability and performance, a user asking a general question on QA rather than searching on the web will, for example, have similar results while QA is longer on giving an exact answer and the search engine is faster on giving the document's paragraph containing the answer. Rather than focusing on factoid questions, user knowledge such as his experiences, notes, medical history, or schedules known only from the knowledge of a user who could be a doctor, teacher, student and etc. are more suitable application for QA or the like where the domain is entity-known information or specific information from a knowledge base. The QA does not need to be able to answer all questions (even humans cannot), but rather answer most questions, and provide retrieval methods for recovering information, that would be helpful for the user.

With the scheduling, the user can store schedules and medications with calendar app integration, check conflict of schedules, and also manage the reminders.

Exploring the user knowledge base provides a better understanding of information extraction and retrieval methods with natural language processing.

RESEARCH METHODOLOGY

The study relies primarily on user knowledge. The user knowledge expected is composed of logs, schedules, or medications.

Schedules and medications are time-based knowledge for the scheduling system. Time-based knowledge is similar to events on calendar applications. Events can be recurring or one instance. One instance event is an all-day or day event and has a duration of one or more days. A schedule has a start time and end time. Medication is a specific type of knowledge compared to log and schedule which are more general. Since doctor prescribes medicine to be taken per interval, instead of manually adding each, recording medication allows adding of multiple events on the calendar for days and increasing by interval hour and minute if specified; otherwise, it treats them as past or future medication or medication log.

Logs can be a knowledge of anything. Though, it still keeps a time reference which is the current time of creation. Logs and schedules are general knowledge because its content can vary. For the user to experience the features of the app, the knowledge base was structured to be more open to different types of knowledge.

Besides storing the knowledge, it can also be retrieved to extract the answer by asking a question that is processed by the QA system. The user can ask what, where, when, who, and yes or no questions based on the log, schedule, or medication information from the user knowledge base.

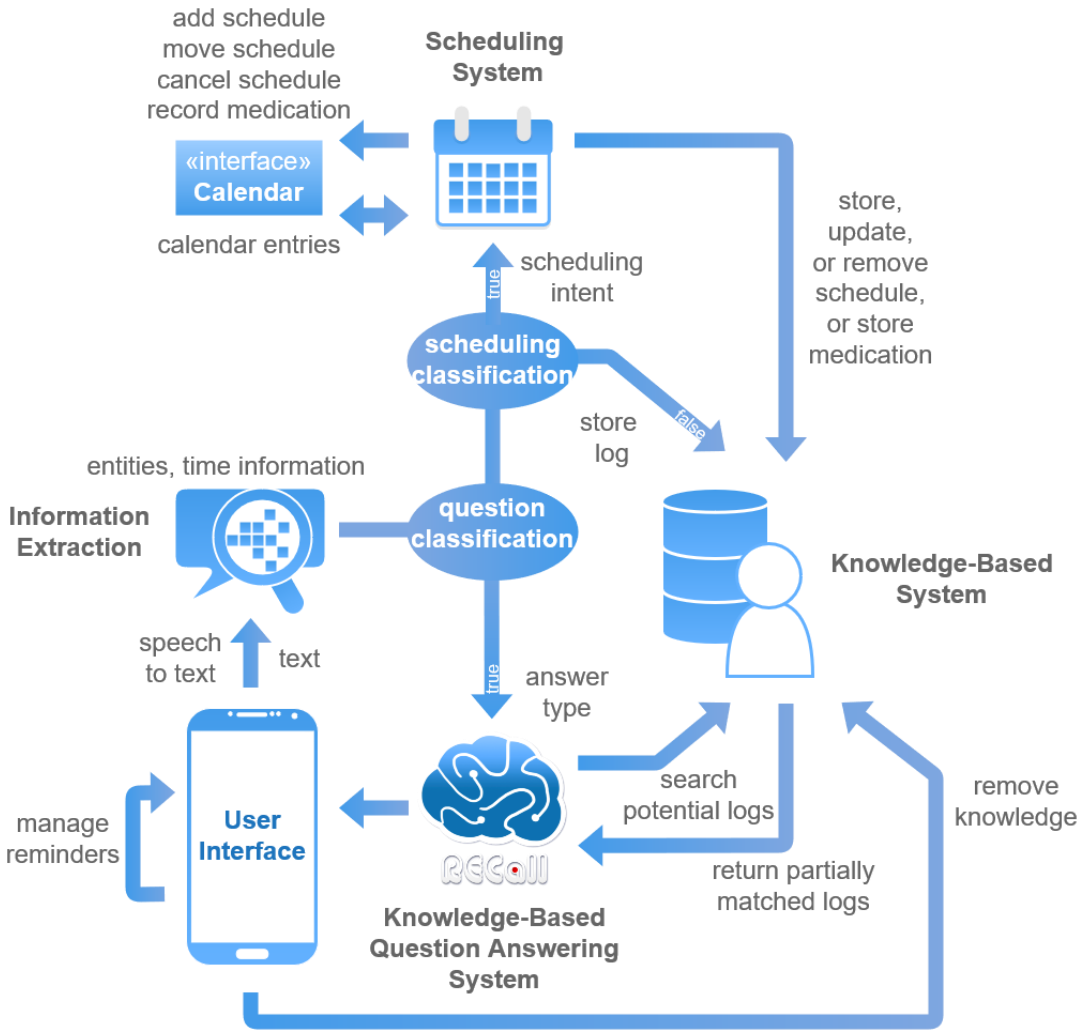

Figure 2 - Conceptual Framework

Figure 2 provides a conceptual framework of the system. The IE, KBQA, KB, and SCH are the four major system components. Through the speech or text, the user can record log, can add or move or cancel schedule, can record medication, and can ask question to retrieve answers. Also, user can manage reminders on the app settings, and remove knowledge on the app.

CHAPTER II

SOFTWARE REQUIREMENTS AND DESIGN SPECIFICATIONS

This chapter specifies the user and system requirements that are expected to be accomplished as well as the structure and process of achieving these. It contains sections for the Use Case Diagram, Use Case Narrative, Activity Diagram, Class Diagram, and User Interface Design.

USE CASE DIAGRAM

The use case diagram shows the relationship between actors and different use cases of the system in which an actor is involved. It also shows an overview of the system in a high-level design. Each use case is further detailed on its narrative.

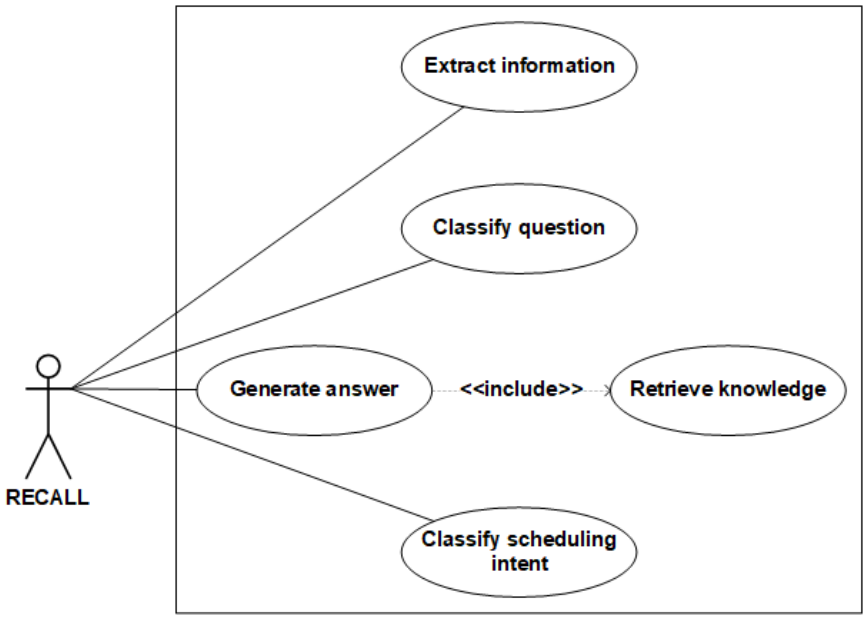

Figure 3 - Use Case Diagram

Figure 3 defines the different use cases in which the system performs when a user, the business actor, records knowledge or asks a question through the mobile app.

USE CASE NARRATIVE

The use case narrative shows the main success scenario and alternative flows of a use case. It provides more details about the use case.

<table>
<tr><td colspan="2">Number and Name: UC 101 – Extract information.</td></tr>
<tr><td colspan="2">Actor(s): System</td></tr>
<tr><td colspan="2">Description: This extracts the entities and time information from the natural language input. Entity describes the span of text in which constitutes a location, medication, person, time, and others. Time information, which consists of medication time units, and date and time expressions, is converted to system number or time representation.</td></tr>
<tr><td colspan="2">Precondition(s): None</td></tr>
<tr><td colspan="2">Postcondition(s): None</td></tr>
<tr><td>Actor Action</td><td>System Response</td></tr>
<tr><td></td><td>1. Parse medication time unit expressions.
2. Convert each unit into a number.
3. Parse date and time expressions.
4. Group each date and time into pairs.
5. Convert each pair into system time.
6. Normalize text.
7. Tag each token with part of speech.
8. Parse each tagged word into chunks.
9. Recognize entity label for each chunk.</td></tr>
<tr><td colspan="2">Exception Flow E1: System cannot extract entities information.</td></tr>
<tr><td colspan="2">1. System displays text input as user message in gray.
2. System displays unknown intent as app message in blue.
3. End process.</td></tr>
<tr><td colspan="2">Exception Flow E2: System cannot extract time information.</td></tr>
<tr><td colspan="2">1. System displays text input as user message in gray.
2. System displays unsupported time as app message in blue.
3. End process.</td></tr>
</table>

Number and Name: ***UC 102 – Classify question.***	
Actor(s): System	
Description: This classifies if the natural language input is in the form of valid question. The system finds the question parts such as the wh-word, auxiliary verb, subject, verb phrase, and keyword chunk(s), then, it checks its syntax and time, then, if valid, it detects the answer type which consists of the entity type, storage type, iteration direction, and other descriptive labels.	
Precondition(s): None	
Postcondition(s): None	
Actor Action	System Response
	1. Map question parts. 2. Check syntax with question rules. 3. Check time and verb. 4. Detect the answer type.
Exception Flow E1: System classifies invalid question form; System found ambiguity in time information and verb tense.	
1. System displays text input as user message in gray. 2. System displays unsupported question as app message in blue. 3. End process.	

Number and Name: ***UC 103 – Generate answer.***
Actor(s): System
Description: This generates an answer for the question based on the log, medication, or schedule information of the user. Null (no or none) answer may be generated. If multiple answers are generated, the system picks the top answers and formats them.
Precondition(s): None
Postcondition(s): None

Actor Action	System Response
	1. Search potential knowledge. 2. Analyze each knowledge. 3. Infer answer from each knowledge. 4. Check if answer conforms with answer type. 5. Remove no answer. 6. If no answers found, repeat step 1. 7. Sort answers. 8. Format top answers.
Alternate Flow A1: System retrieved no knowledge.	
1. System generates a null answer.	

Number and Name: ***UC 104 – Retrieve knowledge.***	
Actor(s): System	
Description: This retrieves the log, schedule, or medication information of the user based on the answer type and time information. If there is no time, the system searches the entries by next recent date and iteration direction.	
Precondition(s): None	
Postcondition(s): None	
Actor Action	System Response
	1. Get user knowledge from the database within base time based on storage type.
Alternate Flow A1: System found no base time.	
1. If the next recent date is empty, set next recent date to today. Else, get next recent date of user knowledge from the database based on storage type. 2. If the next recent date is empty, return empty knowledge. 3. If the next recent date is today and iteration direction is past, set pivot time to today (start time is current time). If next recent date is today and iteration direction is past,	

set pivot time to today (end time is current time). Else, set pivot time to next recent date (whole day start and end time). 3. Keep the next recent date. 4. Get knowledge from the database within pivot time based on storage type.

Number and Name: ***UC 105 – Classify scheduling intent.***	
Actor(s): System	
Description: This classifies if the natural language input has a scheduling intent. The system checks for scheduling verb and keyword and for time information to determine the scheduling intent. The system identifies the scheduling intent: add schedule, add past schedule, move schedule, cancel schedule, record medication interval, record medication, record past medication, or record medication log.	
Precondition(s): None	
Postcondition(s): None	
Actor Action	System Response
	1. Find the verb and scheduling entity. 2. Check for scheduling verb and keyword. 3. Inspect time information. 4. Determine scheduling intent.
Exception Flow E1: System found medication time units exceed max value.	
1. System displays text input as user message in gray. 2. System displays unsupported time as app message in blue. 3. End process.	
Exception Flow E2: User records medication interval with past start time; User cancels or moves schedule with negation; User adds schedule in past tense.	
1. System displays text input as user message in gray. 2. System displays unknown intent as app message in blue. 3. End process.	

ACTIVITY DIAGRAM

The activity diagram shows a workflow behavior of the system by describing the sequence of actions in a process.

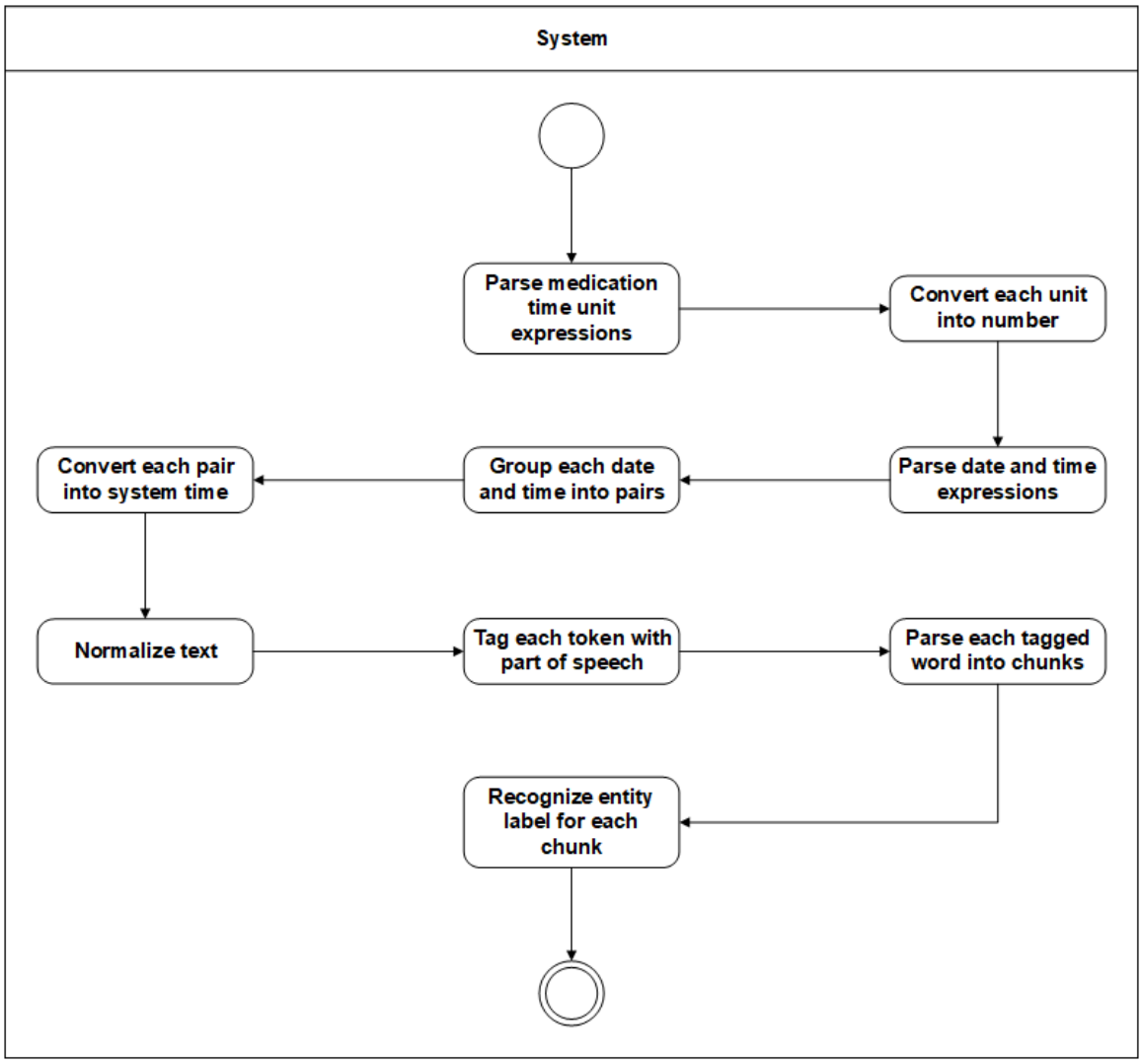

Figure 4 - Extract Information Activity

Figure 4 shows the behavior of the system upon how it extracts the entities and time information from the natural language input when the user asks a question or records knowledge. Entity describes the span of text in which constitutes a location, medication, person, time, and others. Time information, which consists of the medication time units, and date and time expressions, is converted to system number or time representation.

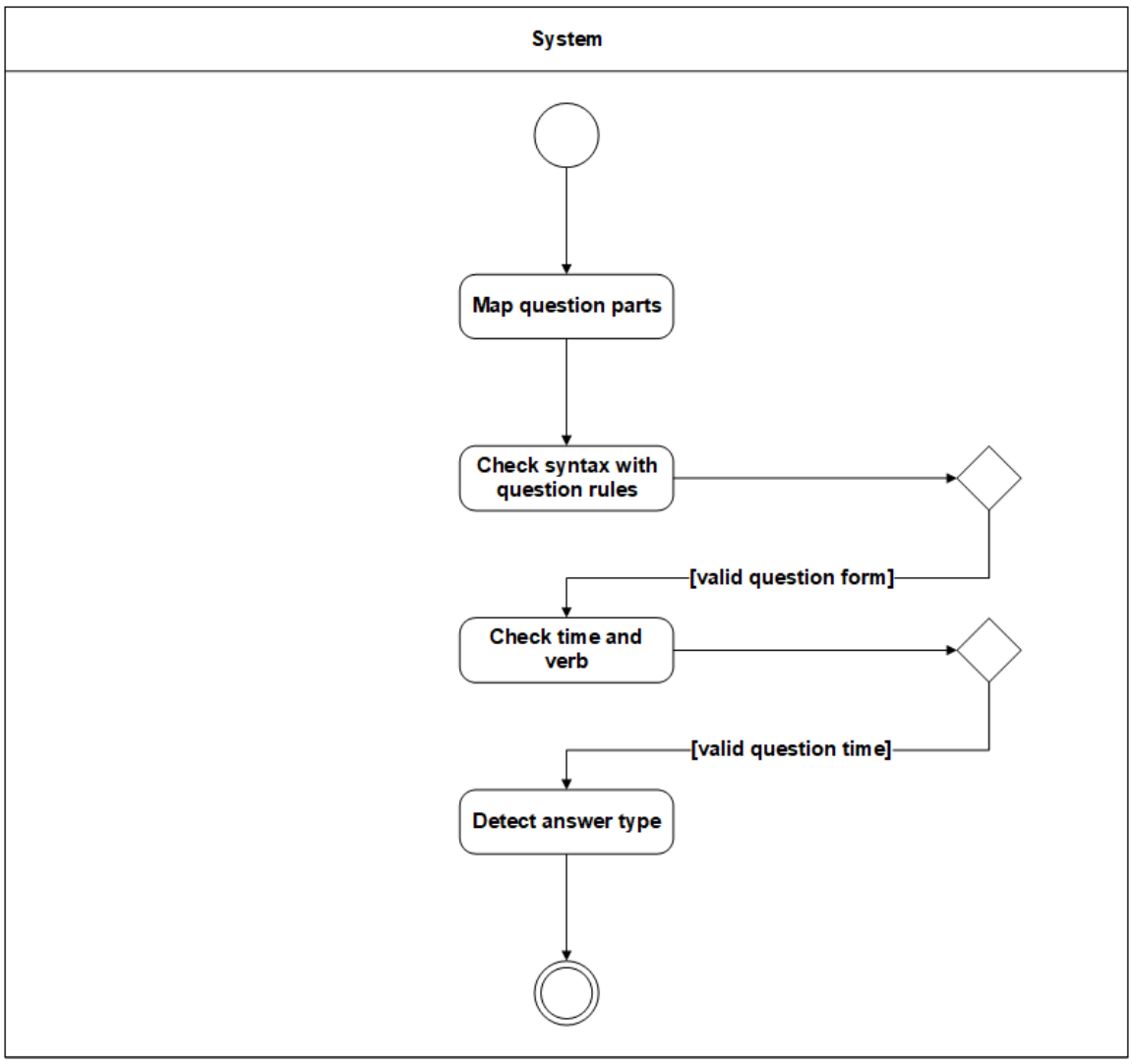

Figure 5 - Classify Question Activity

Figure 5 shows the behavior of the system upon how it classifies if the natural language input is in the form of valid question, when the user asks a question. The system finds the question parts such as the wh-word, auxiliary verb, subject, verb phrase, and keyword chunk(s), then, it checks the question structure, then, if valid, it detects the answer type which consists of the entity type, storage type, iteration direction, and other descriptive labels.

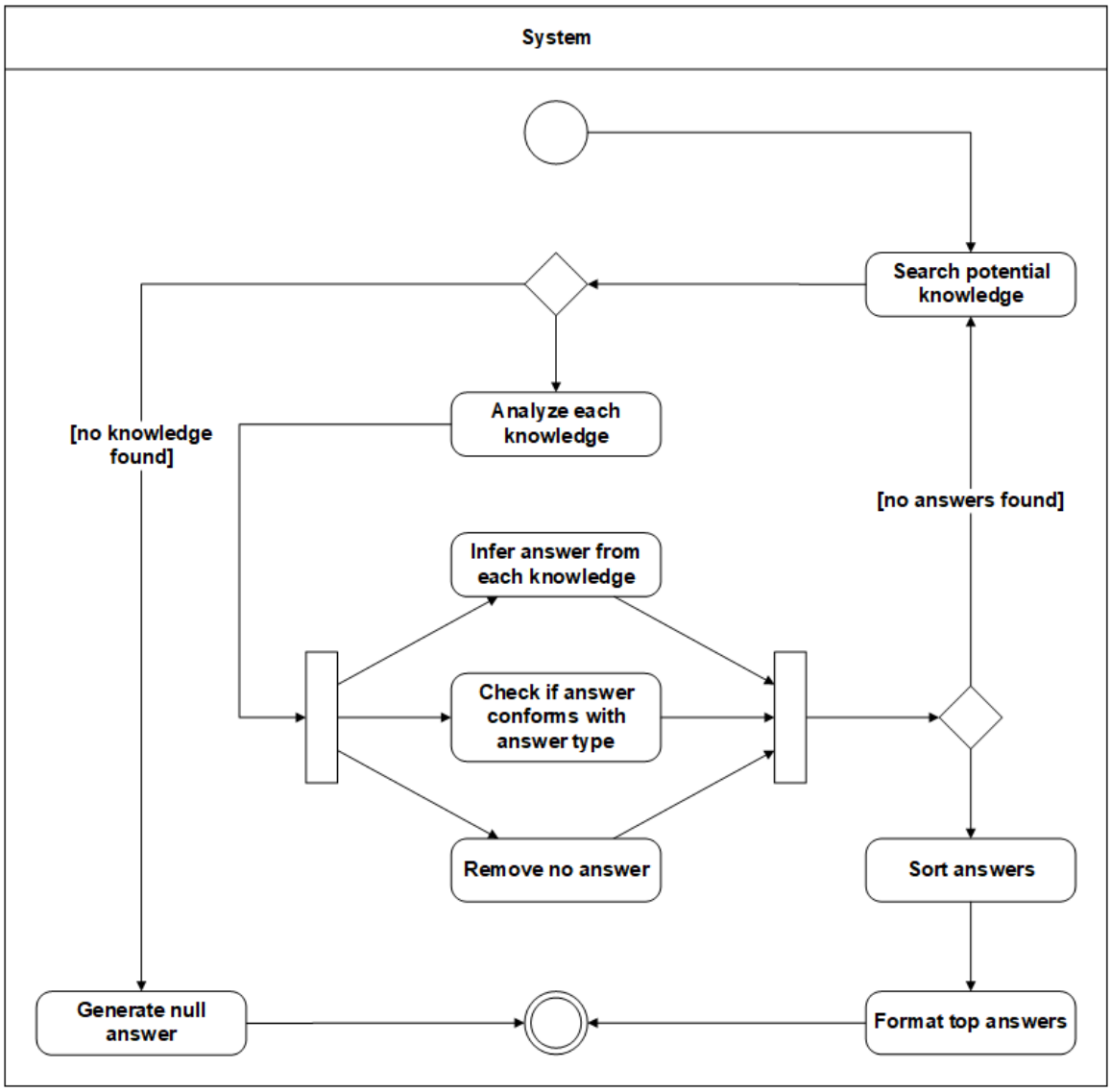

Figure 6 - Generate Answer Activity

Figure 6 shows the behavior of the system upon how it generates answer based on the log, medication, or schedule information of the user when the user asks a question. Null (no or none) answer may be generated. If multiple answers are generated, the system picks the top answers and formats them.

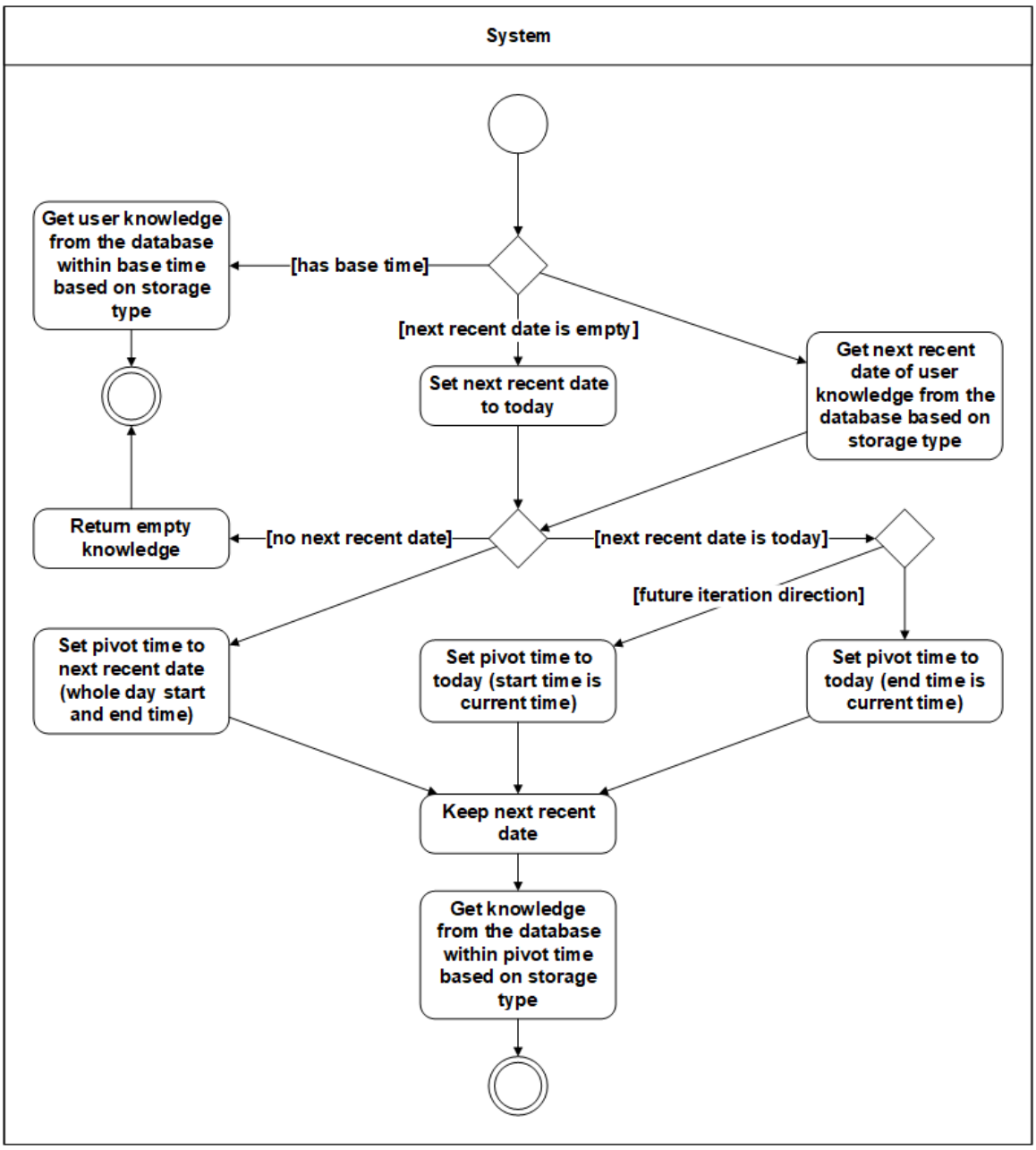

Figure 7 - Retrieve Knowledge Activity

Figure 7 shows the behavior of the system upon how it retrieves the log, schedule, or medication information of the user based on the answer type and time information when the user asks a question. If there is no time, the system searches the entries by next recent date and iteration direction.

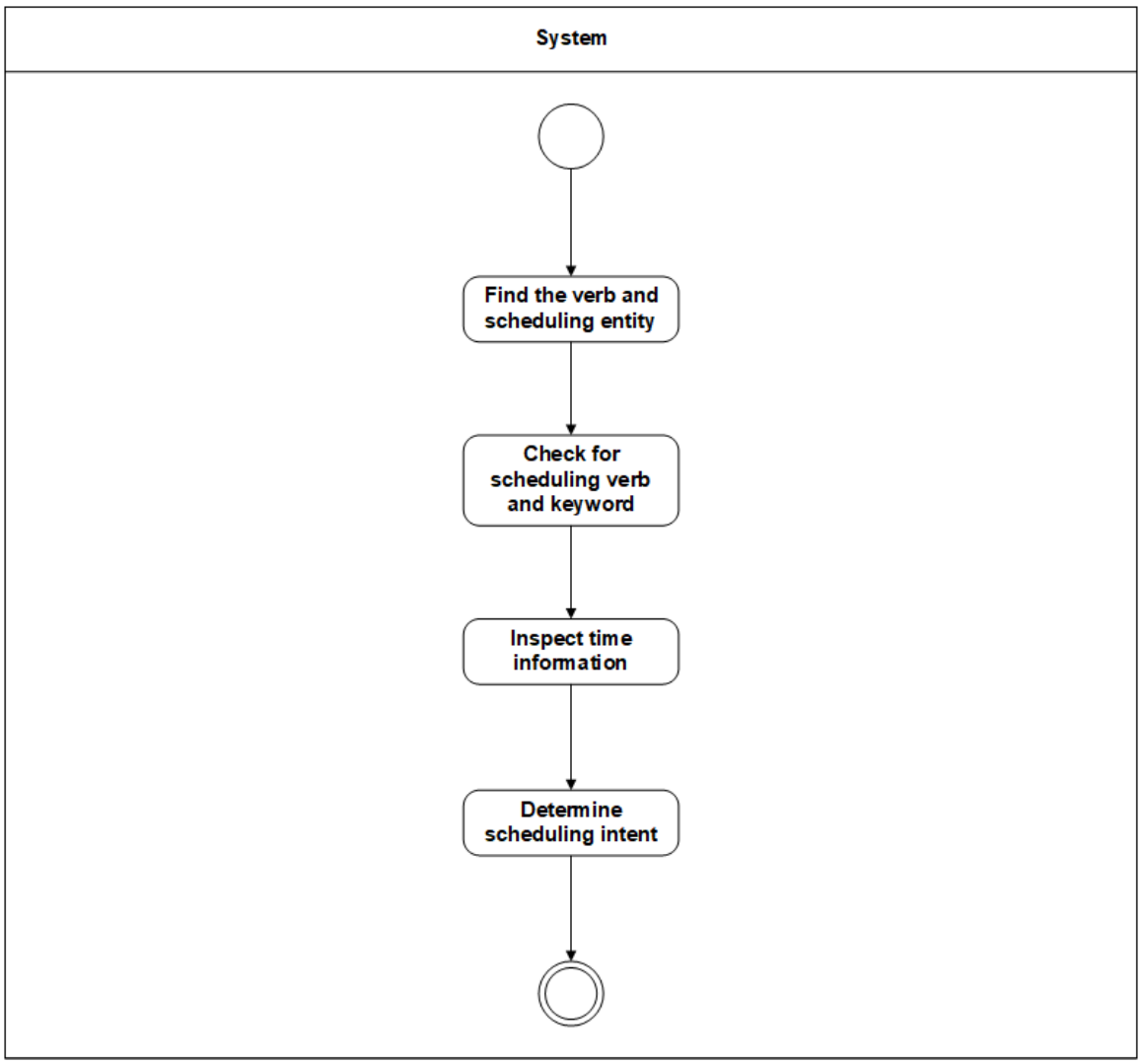

Figure 8 - Classify Scheduling Intent Activity

Figure 8 shows the behavior of the system upon how it classifies if the natural language input has a scheduling intent, when the user records medication or adds, moves, or cancels schedule. The system checks for scheduling verb and keyword and for time information to determine the scheduling intent. The system identifies the scheduling intent: add schedule, add past schedule, move schedule, cancel schedule, record medication interval, record medication, record past medication, or record medication log.

CLASS DIAGRAM

The class diagram shows the structure of the system as classes with their attributes, operations, and relationships among other classes.

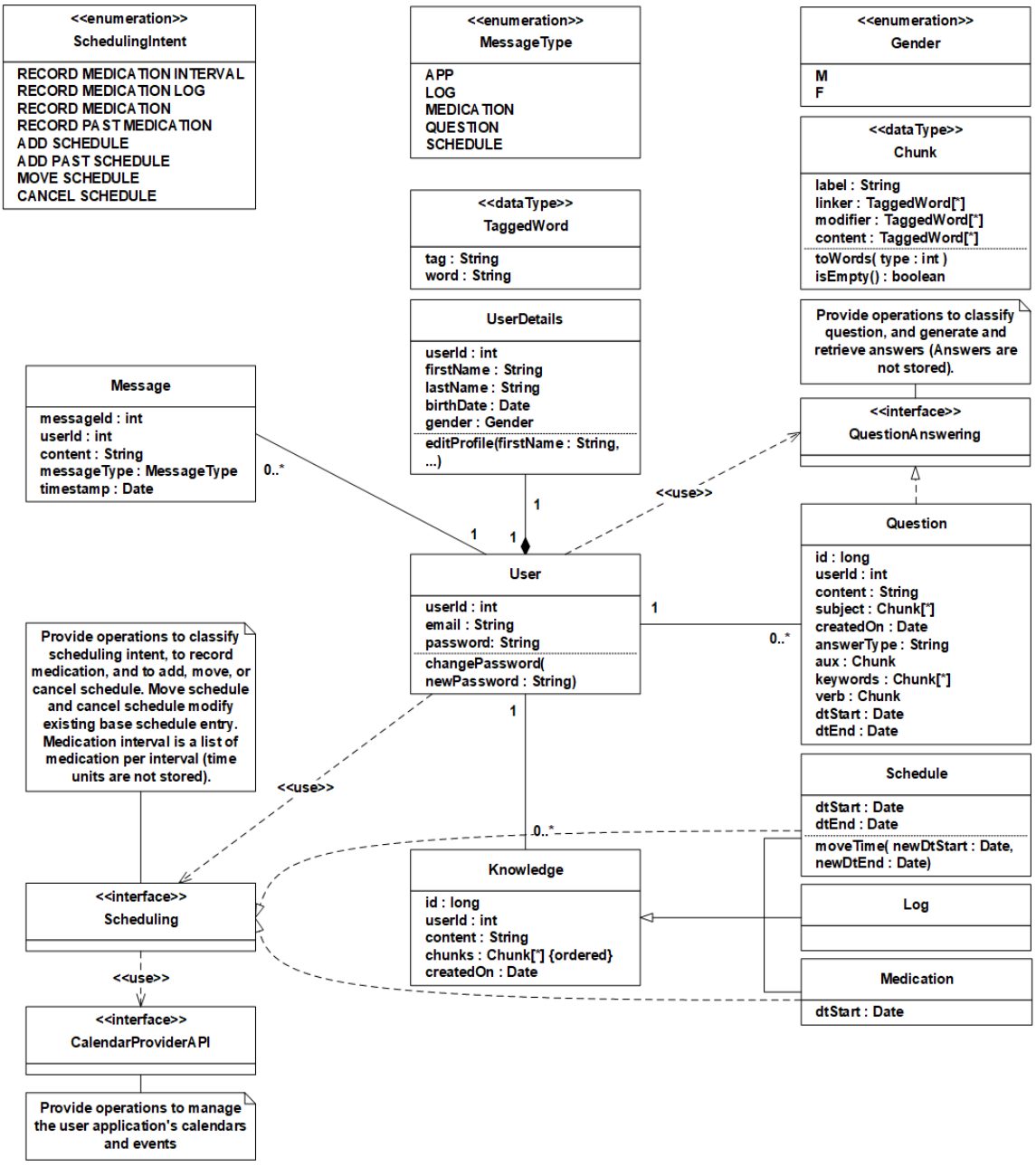

Figure 9 - Class Diagram

Figure 9 represents the structure of the designed system.

USER INTERFACE DESIGN

The user interface design is the process of creating interfaces that are expected to provides insights into how the user can interact with the system.

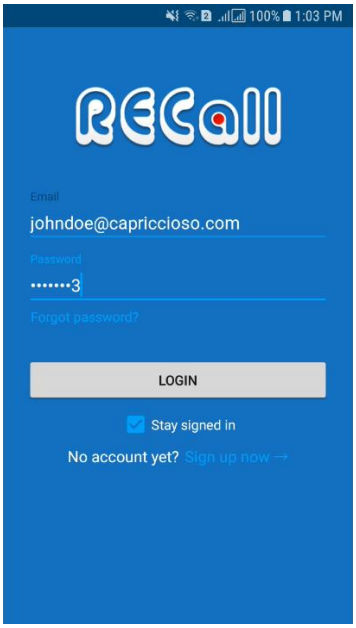

Figure 10 - Login Screen

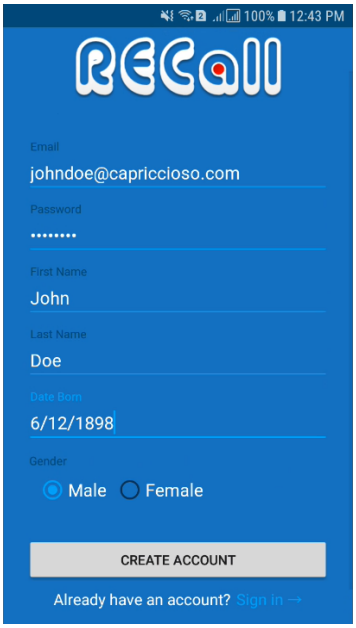

Figure 11 - Registration Screen

Figure 10-11 shows login and registration screen where a user can register for an account to login to the app. Basic information is needed for identity and security purposes.

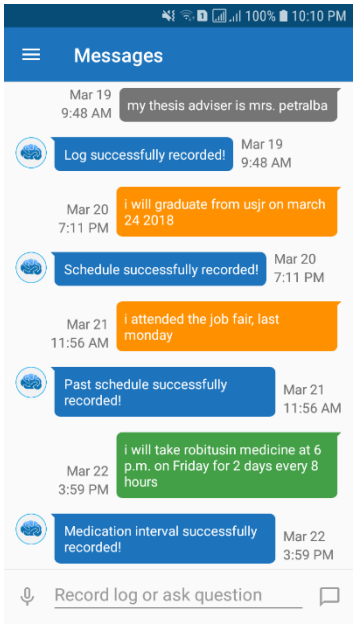

Figure 12 - Messaging Screen 1

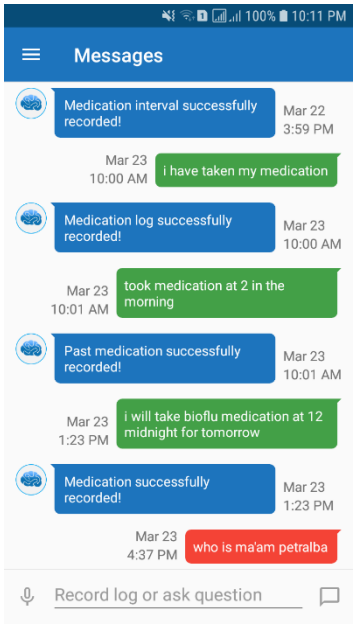

Figure 13 - Messaging Screen 2

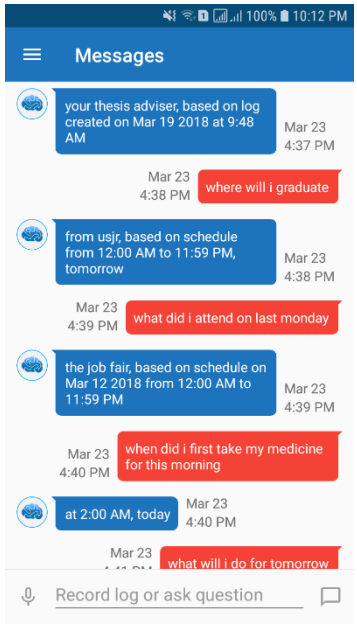

Figure 14 - Messaging Screen 3

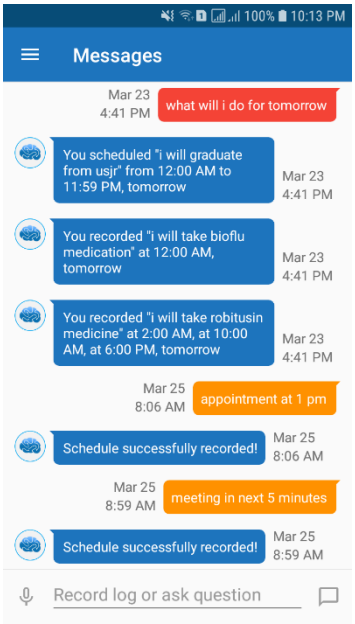

Figure 15 - Messaging Screen 4

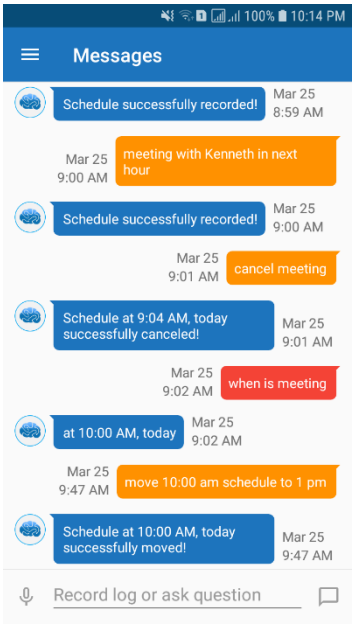

Figure 16 - Messaging Screen 5

Figure 12-16 shows the messaging screen where the user can record log, record medication, add schedule, move schedule, cancel schedule, and ask question through a messaging interface.

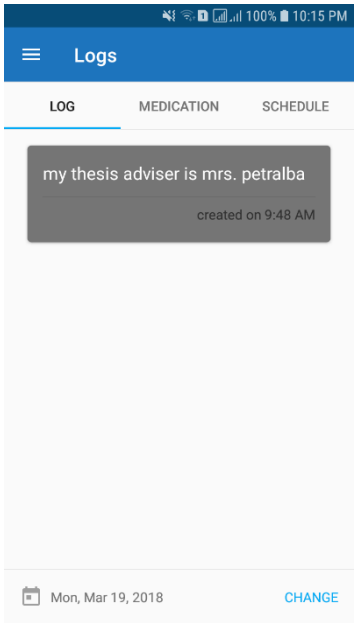

Figure 17 - Logs Screen 1

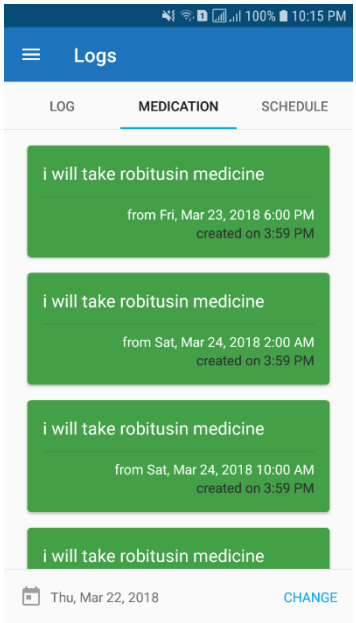

Figure 18 - Logs Screen 2

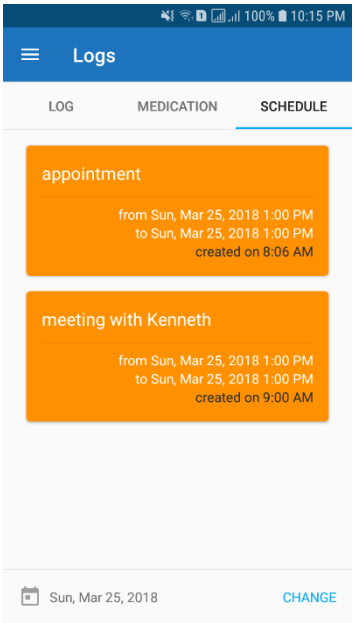

Figure 19 - Logs Screen 3

Figure 17-19 shows the logs screen where the user can view his logs or knowledge. The logs are sorted ascendingly by creation time.

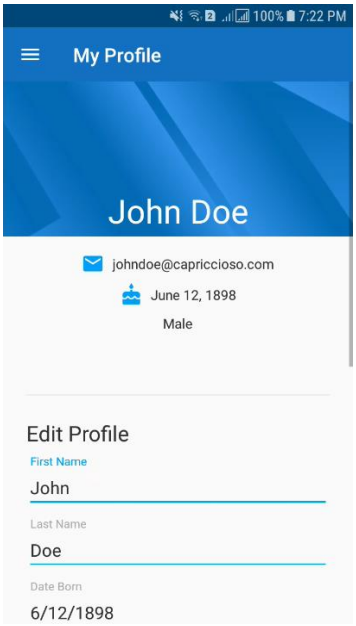

Figure 20 - Profile Screen 1

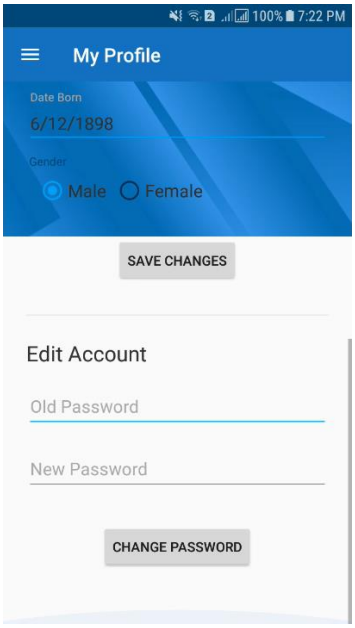

Figure 21 - Profile Screen 2

Figure 20-21 shows the profile screen where the user can edit profile and change password.

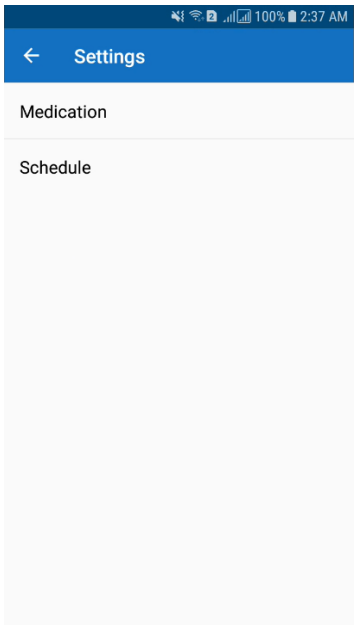

Figure 22 - Settings Screen 1

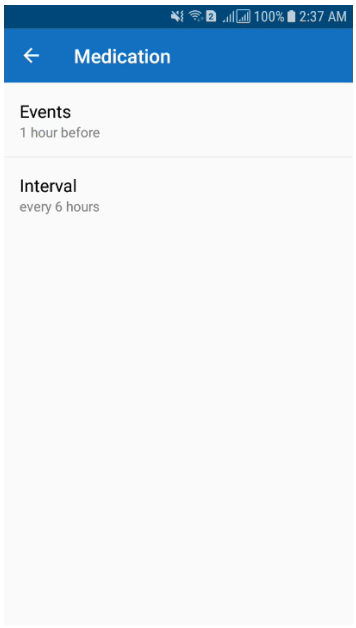

Figure 23 - Settings Screen 2

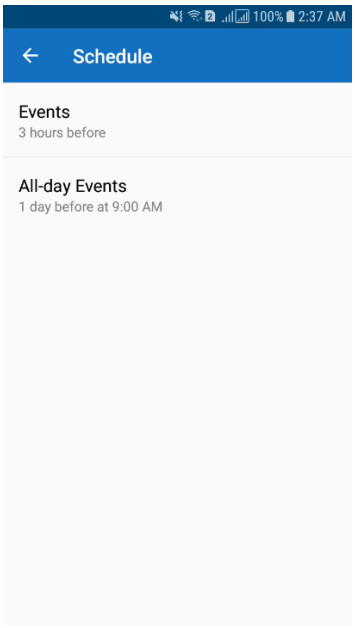

Figure 24 - Settings Screen 3

Figure 22-24 shows the settings screen where the user can modify medication settings and schedule settings shared throughout the app. This screen allows the user to set the default time of the reminder before the schedule or medication. Also, the user can

specify default interval hour and minute for recording medication interval on the settings so that the user can record the medication with the number of days as input. All-day events are schedules that start with 12 AM and ends with 11:59 PM of the day or range of day specified.

CHAPTER III

SOFTWARE DEVELOPMENT AND TESTING

This chapter describes the implementation of the project in development, the various tools used to create and run the application, and testing methods to evaluate the question answering results. It contains sections for the Development and Testing Process.

DEVELOPMENT AND TESTING PROCESS

The software is composed of the front-end which is the mobile platform application and of the back-end which is the Spring boot application with the REST web services, and database. The development process utilizes the prototyping model for breaking down the development task into smaller parts. The initial prototype is produced and is refined through stages, to help discover requirements problems and discard them later on. During the testing process, unit testing can be done earlier after building the prototype to see if the different functionalities of the system work, and then do the integration testing. A small-scale mock-up module with the prepared knowledge base and sets of questions was used to test the question answering.

For the mobile platform application, these are the development environment and tools used:

- Android Studio 3.1.3 – an IDE for Android development using Java; use to manage the client-side project workspace, to code, to design the UI, and to build, debug, and run the Android app through a connected Android device
- RoboSpice 1.4.11 – a library that eases writing asynchronous network requests and supports REST out of the box like Spring Android; use to manage

the client network requests asynchronously upon sending the request and receiving the response, during failure, progress, and success

- Spring Android 1.0.1 – a framework that is designed to provide components of the Spring family of projects for use in Android apps; use to access the HTTP methods for the Android client to make network requests
- Realm 5.7.0 – a mobile database that runs directly inside phones, tablets or wearables, which is much faster than an ORM, and often faster than raw SQLite while maintaining a powerful feature set; use to store and view messages in the local database of the Android device while offline
- Jackson 2.9.6 – a JSON library for Java that has strong data binding capabilities and provides a framework to serialize custom java objects to JSON and deserialize JSON back to Java objects; use to transfer the Java objects back and forth from client to server by network request and response
- JodaTime 2.9.9 – a standard date and time library for Java which provides provide all the required functionality for date-time and time zone calculations which are poorly done by Java date and time classes; use to manipulate date and time instances, and for testing which mocks the time according to the creation time of dataset entry

For the Spring boot application, these are the development environment and tools used:

- Eclipse Oxygen 3a for Java EE Developers – an IDE for creating Java EE and Web applications; use to manage the server-side project workspace, to code, and to build, debug, and run the Spring boot application with the REST web services

- Spring Boot 2.0.2 – a Spring project that makes it easy to create stand-alone, production-grade Spring based application, together with the dependencies Spring Boot Starter Web for full-stack web development with Tomcat and Spring MVC, and Spring Boot Starter Data JPA for Java Persistence API including Hibernate; use to build the REST web services
- MySQL Connector Java 5.1.46 – a JDBC driver for MySQL; use to create a connection by Hibernate to MySQL Server for creating queries to the server database
- MySQL Server 8.0.11 – a database server for MySQL; use to manage connections to the server database for storing and handling data
- MySQL Workbench 6.3.10 – a unified visual tool for database architects, developers, and DBAs for MySQL; use to view the data, run scripts, and model the schema of the server database
- JWI 2.4.0 – a Java library for interfacing with Wordnet; use for entity recognition and for finding the lemma of words and synonyms of verbs
- WordNet 2.1 (dictionary files) – a large lexical database of English language; use by JWI for Wordnet resources
- Stanford POS Tagger 3.5.2 (English) – a piece of software that reads the text in some language and assigns parts of speech to each word; use for text normalization and POS tagging
- Stanford Caseless Model 2015-04-20 (English) – a case-insensitive model for Stanford POS Tagger; use to ignore inconsistent capitalizations for caseless POS tagging
- Java Excel API 2.6.12 – a Java API to read, write, and modify Excel spreadsheets; use for integration testing at the backend which reads an input

excel file for the dataset, and two output excel files for the knowledge base, and question answering responses and extracted information

- JodaTime 2.9.9

Development Process

This section discusses the technical process during the development. Figure 2 shows the conceptual framework while Figure 25 provides the developed system overview. Similarly, there are four major components: IE, SCH, QA, and KB. The question classification and scheduling classification are processes of intent extraction or classifying the intent. Mainly, the KB stores the updated or new structured information or knowledge representation into the database, and fetches them for processing or viewing purposes.

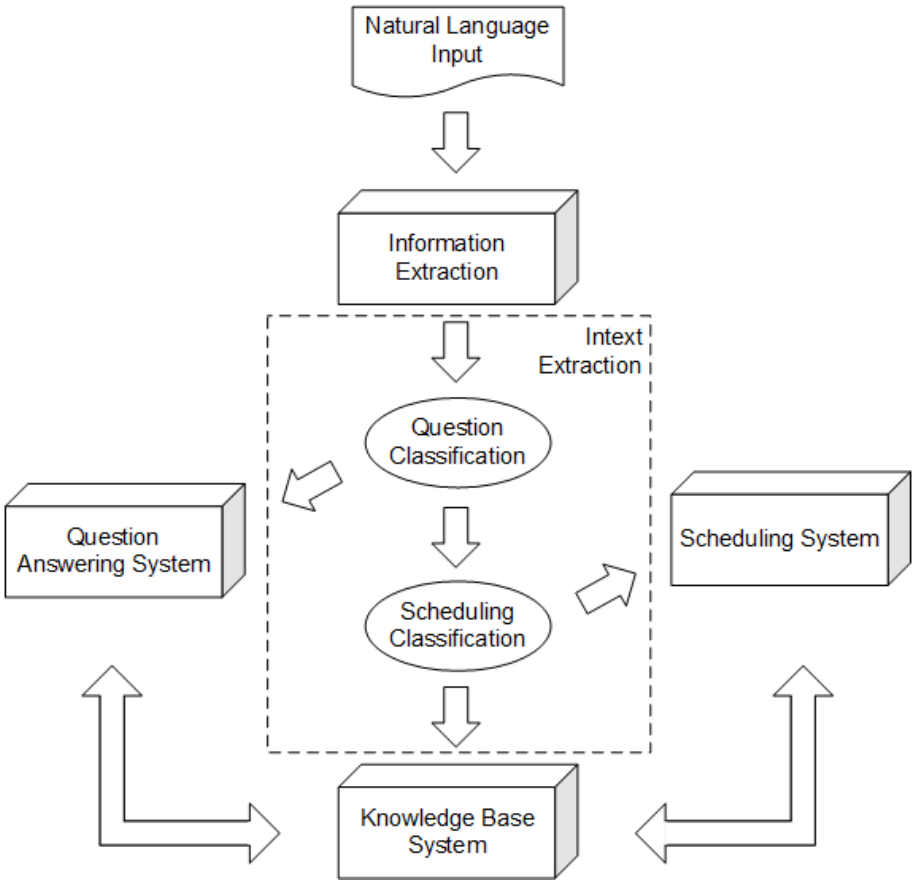

Figure 25 - System Overview

Information Extraction

This component is composed of two main processes: temporal expression extraction and entity extraction.

1. Temporal Expression Extraction

The system parses the time information from the natural language input. Time information, which consists of medication time units, and date and time expressions, is converted to system number or time representation. The parser supports temporal expressions, in the form of prepositional phrases in Table 1-3 using regular expressions; for simplification, examples are given instead of the whole pattern. This process involves medication time unit parsing and datetime parsing.

1.1 Medication Time Unit Parsing

The process follows (1) parsing medication time unit expressions, then (2) converting each to a number. It reads the natural language input and produces the system number of the medication time units.

MEDICATION TIME UNIT EXPRESSION	
day	for 1 day | for 7 days
hour	every 1 hour | every 23 hours
minute	every 1 minute | and 59 minutes

Table 1 - Medication Time Unit Expression

Table 1 shows examples of expressions for each medication time unit. The numerical values are the minimum and maximum value for each.

Given: robitussin medicine for 2 days every 6 hours and 30 minutes				
(1)	robitussin medicine	for 2 days	every 6 hours	and 30 minutes
		time unit	time unit	time unit

(2)	no. of days = 2 no. of hours = 6 no. of minutes = 30

Figure 26 - Medication Time Unit Parsing Example 1

Figure 26 is an example of a medication interval that goes through step (1-2) of this process. Also, the user can state a start time that will proceed through datetime parsing.

1.2 Datetime Parsing

The process follows (1) parsing date and time expressions, (2) grouping them into pairs, (3) converting each pair to system time. It reads the parsed medication time unit text and produces datetime (in short for date and time) pairs to represent calendar event time.

A calendar event time is composed of start time and end time. Similarly, a datetime pair consists of start datetime and end datetime.

Given: appointment from Jan. 21 at 11:00 AM to 1:00 PM of Jan. 22 start datetime pair 1 end datetime pair 1

Figure 27 - Datetime Pair Example 1

Given: meeting for tomorrow around 3:00 PM to 4:00 PM was moved to next Friday start datetime pair 1 end datetime pair 1 start datetime pair 2

Figure 28 - Datetime Pair Example 2

In Figure 27, full date and time expressions are present in the pair. In Figure 28, a second datetime pair is present for moving a schedule; hence, the maximum pairs allowed is two. During conversion, empty datetime is filled with the default or calculated value.

DATE EXPRESSION	
dd/mm/y	on 20/10/2018
dd month y	at 1st of January in 2018 | on 31 Dec. 2018 | by 20th of Oct.
month dd y	at January the 31st in 2018 | on Dec. 31, 2018 | by Oct. 20
day	for the (previous/last/next) 7 days

day adverb	from yesterday | during today | to tomorrow
day of week	for (previous/last/next) Monday | on next next Friday | by (this) Sunday
month	for (previous/last/next) January | on next next Oct. | by (this) Dec.

Table 2 - Date Expression

Table 2 shows examples of expressions for each date. The default time is a whole day duration. The relative date expression will calculate the date, like the day of the week, day adverb, and time unit.

TIME EXPRESSION	
half day hour min	12 a.m. | 11:59 PM | 12 at (midnight/noon) | 11:59 at night |
period	9:00 in the morning | 1 in the afternoon
part of the day	for (previous/last/next) (morning/noon/afternoon/evening/ night/midnight) | till next next evening | at tonight
hour	on the (previous/last/next) 1 hour
minute	for (previous/last/next) 59 minutes

Table 3 - Time Expression

Table 3 shows examples of expressions for each time. The default date will be the current day. The relative time expression will calculate the time, like time unit. Part of the day is a period with a start and end time value.

Current time: 2018-01-20 09:34:00
Given: schedule is on tomorrow in the afternoon

(1)	schedule is	on tomorrow	in the afternoon	
		date	time	
(2)		datetime start pair 1		datetime end pair 1*
(3)		= 2018-01-21 12:00:00		= 2018-01-21 17:99:00

Figure 29 - Datetime Parsing Example 1

Current time: 2018-01-20 09:34:00

Given: the flight is on Jan. 21 from 11:00 PM to 1:00 AM

(1)	the flight is	on Jan. 21	from 11:00 PM	to 1:00 AM
		date	time	time
(2)		datetime start pair 1		datetime end pair 1
(3)		= 2018-01-21 23:00:00		= 2018-01-22 01:00:00

Figure 30 - Datetime Parsing Example 2

Figure 29-30 are examples of adding a schedule that goes through step (1-3) of this process.

2. Entity Extraction

The system parses the entities information from the natural language input without the temporal expressions. Entity describes the span of text in which constitutes a location, medication, person, time, and others (see Table 4 for the full list). This process involves text normalization and POS tagging, chunk parsing, and entity recognition.

2.1 Text Normalization and POS tagging

(1) The process for text normalization follows proper encoding of quotes and expansion of clitic contractions. The clitic contractions include “not.” The text was not normalized further to maintain the original words for answer processing. It reads the parsed datetime text and produces the normalized text.

(2) The process for POS tagging uses the MaxentTagger of Stanford with the caseless model. It reads the normalized text and produces the tagged words.

Also, some fixes are done such as untagged POS before POS, gerund as GRND if not used in the verb phrase, and special characters as SYM. Custom tags include the quotes as QQ, mixed number and character name as NM for some named entity.

Given: Jose Rizal's fight wasn't implying "the sword is mightier than the pen"
Parsed time text: Jose Rizal's fight wasn't implying "the sword is mightier than the pen"

(1) Jose Rizal's fight was not implying "the sword is mightier than the pen"

(2) Jose_POS Rizal's_POS fight_NN was_VBD not_RB implying_VBG "the sword is mightier than the pen"_QQ

Figure 31 - Text Normalization and POS Tagging Example 1

Given: My first laptop was ASUS U50A during high school
Parsed time text: My first laptop was ASUS U50A during high school

(1) My first laptop was ASUS U50A during high school

(2) My_PRP$ first_JJ laptop_NN was_VBD ASUS_NNP U50A_NM during_IN high_NNP school_NNP

Figure 32 - Text Normalization and POS Tagging Example 2

Given: What did I do for today?
Parsed time text: What did I do ?

(1) What did I do ?

(2) What_WP did_VBD I_PRP do_VB ?_SYM

Figure 33 - Text Normalization and POS Tagging Example 3

Figure 31-32 and 33 are examples of log and question, respectively, that goes through step (1-2) of this process.

2.2 Chunk Parsing

The process reads the tagged words from right to left and groups them to produce the unlabeled chunks.

Chunk	
label	AUX, VB-PH, ATTR, ENTITY, LOCATION, PERSON, TIME, CD, MED, MOD, QQ, THERE, WH POS, PREP-OF, EX, PRP-I

linker	CC, IN, TO, GRND
modifier	CD, DT, JJ, JJR, JJS, LS, PDT, POS, PRP$, RB, RBR, RBS
content	EX, FW, JJ, NN, NNS, NNP, NNPS, NM, PRP, MD, VB, VBD, VBN, VBP, VBZ, VBG, WGT, WP, WP$, WRB, QQ

Table 4 - Chunk Model

Table 4 shows the structure of the chunk and a summary where the chunk parser inserts the tagged words. In the label, the 1st and 2nd lines contain the main labels while the 3rd line contains the extended labels. Note, interjection (UH) is ignored.

In expectation, the basic structure of the natural language input is a single phrase or simple sentence in active or passive voice, which is made up of at least two words. By parsing chunks, the system can accept free form speech in spite of incomplete grammar, inaccuracy of text-to-speech or POS tagging, or ambiguous input.

The kinds of phrases are a noun phrase (including appositives), prepositional phrase, verb phrase, and verbals (gerunds, participles, infinitives). Excluding appositives and verbals, the syntax used is found on appendix B.

Punctuations split the chunk upon reading. Dangling adverbs are attached to the verb phrase or auxiliary verb if found.

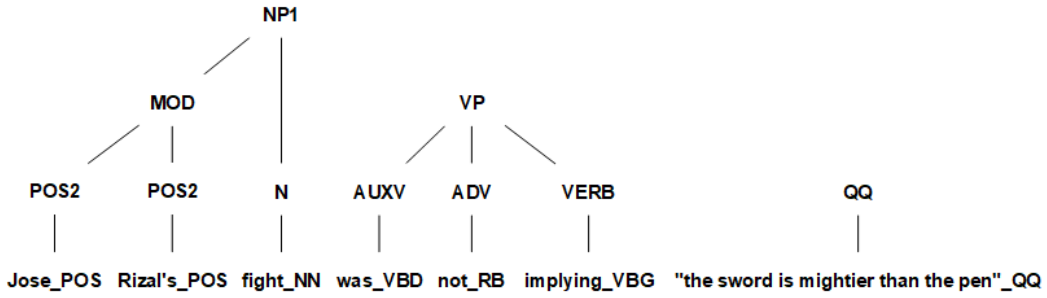

Figure 34 - Parse Tree Example 1

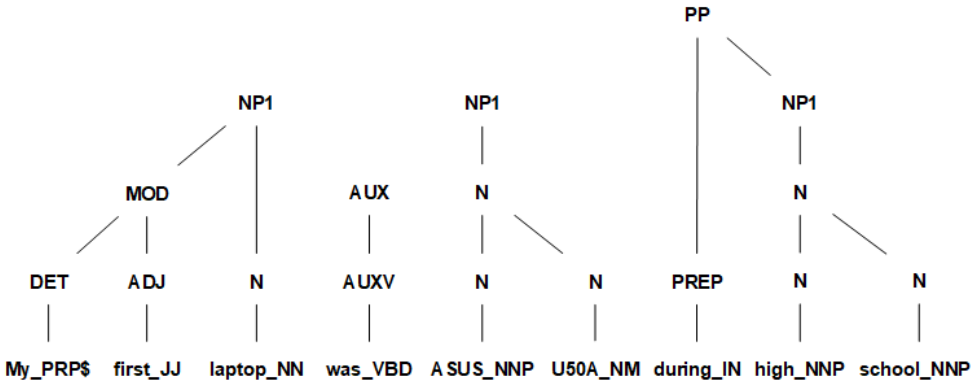

Figure 35 - Parse Tree Example 2

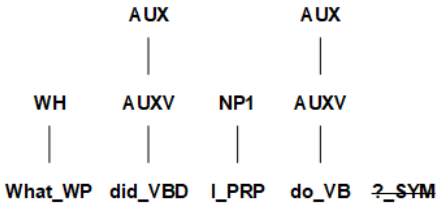

Figure 36 - Parse Tree Example 3

Figure 34-36 are examples of the generated parse tree. It only checks the syntactic structure of the chunk and the chunks are formed according to the chunk model as observed in Figure 37-39.

Note: The notation inside the box denotes *"<linker> : <modifier> : <content>"*

: Jose_POS Rizal's_POS : fight_NN	: not_RB : was_VBD implying_VBG	: : "the sword is mightier than the pen"_QQ

Figure 37 - Chunk Parsing Example 1

: My_PRP$ first_JJ : laptop_NN	: : was_VBD	: : ASUS_NNP U50A_NM	during_IN : : high_NNP school_NNP

Figure 38 - Chunk Parsing Example 2

: : What_WP	: : did_VBD	: : I_PRP	: : do_VB

Figure 39 - Chunk Parsing Example 3

2.3 Entity Recognition

The process reads the unlabeled chunks from left to right and identifies the label for each chunk to produce the entities or labeled chunks.

The three factors used for finding the appropriate label for a chunk are the entity collection, hypernyms distance path, and prepositions.

The entity collection directly labels the chunk when the keywords matched.

```
PERSON → PRSN-HONORIFIC → ma'am | maam | madame | madam | miss | mrs.
                        → mrs | mister | mr. | mr | sir | dr. | dr | master
ENTITY →     SCH-KEYWORD → appointment | schedules | meeting
   MED →     MED-KEYWORD → medicine | medications
```

Figure 40 - Entity Collection for Entity Recognition

In Figure 40, the leftmost hand side is the entity label, and the center is the entity collection name. SCH-KEYWORD and MED-KEYWORD are reserved keywords from the scheduling system.

As observed, some prepositions can determine the entity without looking at the content, while some rely on its object or function.

```
LOC-PREP → (below | beneath | down | down from | under | underneath)
         → (above | over | up) | (behind) | (in | inside | within) | …
PRSN-PREP → as | by | of | for | to | with
TIME-PREP → after | around | as of | at | before | between | during | for | from | in | of
          | past | since | till | to | until | within
```

Figure 41 - Prepositions for Entity Recognition

Figure 41 shows some of the prepositions used for entity recognition. There is limited information about these in the topic, so the values and groupings are experimental. Both preposition and hypernym distance path are evaluated to determine the entity.

Wordnet arguments are used to find the hypernyms distance path of a noun or verb in the wordnet hierarchy. The entity label divides into high category or low category. The high category is the generic term (hypernym) that represents the entity. The low category is the specific instance (hyponym) of the hypernym.

```
HI-ENTITY → W-08560226-N-1-course | W-03063333-N-1-container | …
   HI-ATTR → W-06248892-N-1-name | W-06258440-N-1-title
 LO-ATTR → W-06237055-N-1-modifier | W-04894293-N-1-visual_property | …
    HI-LOC → W-00026074-N-1-location
   LO-LOC → W-08422349-N-1-city | W-08570450-N-1-national_capital
          → W-09092958-N-1-body_of_water | W-08426193-N-1-country | …
  HI-PRSN → W-00007626-N-1-person
 LO-PRSN → W-09904388-N-1-employee | W-10082774-N-1-relative
          → W-09477889-N-1-creator | W-09686747-N-1-bad_person | …
   HI-TIME → W-14923492-N-1-time | W-14931076-N-2-time | W-15025504-N-3-time
  LO-TIME → W-15037287-N-1-season | W-14955716-N-1-time_unit | …
```

Figure 42 - Wordnet Arguments for Entity Recognition

Figure 42 shows some of the wordnet arguments used by the entity recognizer. The entity of high category is used to cancel out undesired paths, and the recognizer excludes “entity” since it is the root of the wordnet hierarchy.

Using JWI, the single noun or rightmost common noun or main verb is lemmatized, and the path is measured by the number of levels ascended from going through all the hypernyms of the lemma until it reaches the wordnet argument. Note, before the path is measured, the lemma is checked if it matches the lemma of the wordnet argument, which gives a path of zero. Also, if the original lemma is a hyponym of the wordnet argument, then the path is one.

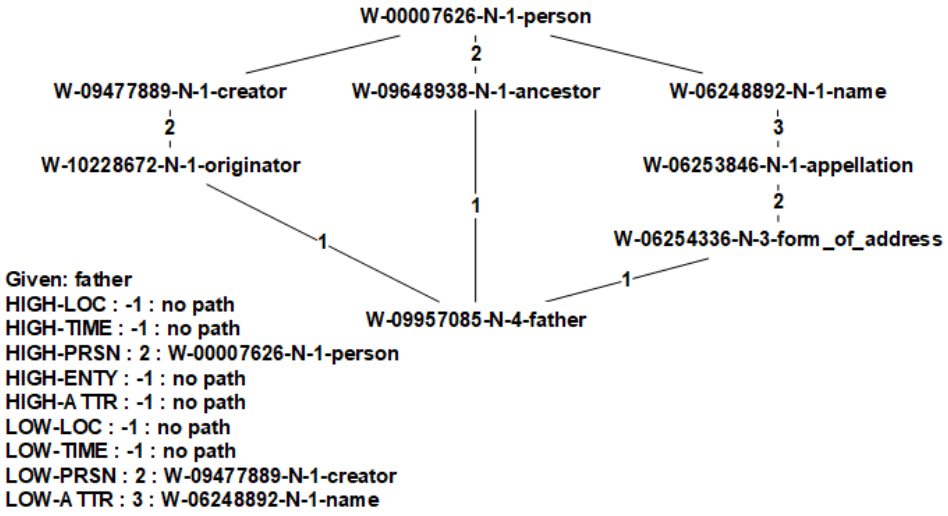

Figure 43 - Hypernym Distance Path Example

Figure 43 shows an example of finding the hypernym distance path from the wordnet hierarchy. If none of the preposition and hypernym distance path provide a clear result, “entity” is the default label.

ENTITY POS	VB-PH	QQ
: Jose_POS Rizal's_POS : fight_NN	: not_RB : was_VBD implying_VBG	: : the sword is mightier than the pen"_QQ

Figure 44 – Entity Recognition Example 1

ENTITY	AUX	ENTITY	TIME
: My_PRP$ first_JJ : laptop_NN	: : was_VBD	: : ASUS_NNP U50A_NM	during_IN : : high_NNP school_NNP

Figure 45 – Entity Recognition Example 2

WH	AUX	PERSON PRP-I	AUX
: : What_WP	: : did_VBD	: : I_PRP	: : do_VB

Figure 46 – Entity Recognition Example 3

Figure 44-46 are examples of the entities recognized.

Question Answering

This component is composed of three main processes: question processing, knowledge retrieval, and answer processing.

1. Question Processing

The system identifies a question and an answer type. It involves the question classification and answer type detection.

1.1 Question Classification

The process maps the question parts and classifies a valid question form by checking the question syntax. It reads the entities information and evaluates to either true or false.

In Figure 47, the question parts are the wh-word, auxiliary verb, and verb phrase, while also mapping the subject and keywords for the answer type detection later on.

Current time: 2018-03-23 18:39:00
Given: when will I graduate on usjr

WH	AUX	PERSON PRP-I	VB-PH	LOCATION
: : when_WRB	: : will_MD	: : I_PRP	: : graduate_VB	on_IN : : usjr_NN
↑ wh-word	↑ auxiliary verb	↑ subject	↑ verb phrase	↑ keywords

Figure 47 - Question Mapping Example

The subject is mapped from the chunks between the wh-word or auxiliary verb and verb phrase or end of the chunk, while the keywords are mapped between the verb phrase and end of chunk. Some questions may not have a verb phrase or keywords.

The question syntax is found on appendix C. SUBJ, AUX, KEYWORDS, and VB-PH represents the subject, auxiliary verb, keywords, and verb phrase string, respectively.

Upon analysis, in the context of time, the question (and knowledge) is related to either past, present or future regardless if there is time information specified. Meaning, the user may ask a question referring to past or future medication, past or future schedule, or past log. For the present time, it denotes the knowledge for today. Notice, the question is always accompanied by an auxiliary verb. An auxiliary verb has a verb tense, also, related to either past, present or future. Therefore, a question without a base time information is accepted. This behavior is incorporated into the QA.

1.2 Answer Type Detection

The process reads the question mapped chunks and identifies the answer type to produce the question.

ASKED FOR (MAIN ANSWER TYPE)	CD, ENTRY, ENTITY, LOCATION, PERSON, QQ, TIME, YES/NO
STORAGE TYPE	LOG, MED, SCH
ITERATION DIRECTION	>=, <=
EXTENDED LABELS	NAME, SUBJ-I, FIRST, LAST

Table 5 - Answer Type Expanded Form

In Table 5, the answer type consists of the asked for, storage type, iteration direction, and extended labels. The iteration direction >= means to look ahead and <= means to look behind.

For the basic question ontology in Figure 48, an answer type is associated with the wh-word or auxiliary verb.

YES/NO → auxiliary verb

ENTITY → what LOCATION → where

PERSON → (who| to whom | whose) TIME → when

Figure 48 - Basic Answer Types

For more answer types, an answer type is associated with question reserved keywords.

CD → … *AUX* … number …

ENTRY → … *AUX* i (do | say | speak | record) …

LOCATION → … *AUX* … (place | location) …

QQ → … *AUX* quote … | … *SUBJ* quoted …

TIME → … *AUX* time …

LOG → … logs *VB-PH* …

→ … *SUBJ VB-PH* … logs …

MED → … *MED-KEYWORD VB-PH* …

→ … *SUBJ VB-PH* … *MED-KEYWORD* …

→ … *SUBJ MED-VERB* …

SCH → … *SCH-KEYWORD VB-PH* …

→ … *SUBJ VB-PH* … SCH-KEYWORD …

NAME → … *AUX* … name … SUBJ-I → … *AUX* I …

FIRST → … (first | next) *VB-PH* … LAST → … (last | previous) *VB-PH* …

Figure 49 - Extended Answer Types

In Figure 49, the reserved keywords are found in the word instances on the following question pattern. The extended answer types may override or add to the answer type. For simplification, "…" is an empty or any string. Note, other verb and noun forms may be accepted. MED-KEYWORD, MED-VERB, SCH-KEYWORD are scheduling keywords (see Figure 60-61). For examples, see appendix D – #128 for CD, #119 for ENTRY, #133 for LOCATION, #193 for QQ, #197 for TIME, #191 for LOG, #20 for MED; #192 for SCH, #3 for NAME, #11 for SUBJ-I, #54 for FIRST, and #55 for LAST.

Based on the answer type, the two types of question are entry-based or entity-based. The entry-based question is whose answer is the knowledge entry itself, while

entity-based is the entity (including yes or no). Storage type provides where to look for the answer. Given the three knowledge types and their characteristics, the storage type can be narrowed down. The entry-based question explicitly tells the storage type (except in case for ENTRY which is for all). While the entity-based question relies on time based on the iteration direction.

To look behind refers from the current point of time towards the past, while to look ahead is towards the future. Therefore, logs refer to look behind, while schedules and medications refer to either to look ahead or to look behind. Furthermore, the present time denotes the current day, specifically, between the current time and start of the day (to look behind) or end of the day (to look ahead). Also, the iteration direction governs the movement of knowledge retrieval.

```
IF auxiliary verb is in future tense
    IF base time is past THEN invalid question ELSE iteration direction is look ahead
IF auxiliary verb is in past tense
    IF base time is future THEN invalid question ELSE iteration direction is look behind
IF auxiliary verb is in present tense
    IF entity type is time THEN iteration direction is look ahead
    ELSE IF base time is future THEN iteration direction is look ahead
    ELSE IF base time is today THEN iteration direction is look ahead
    ELSE IF base time is past THEN iteration direction is look behind
    ELSE iteration direction is look behind
```

Figure 50 – Determining Iteration Direction Pseudocode

In the present tense of Figure 50, questions without base time lean to look behind to prioritize log questions.

The system stores the question on the knowledge base for future use. Other steps like cleaning the question by stop word removal and text normalization (since the original words were kept during IE) can be processed for answer processing later on.

Current time: 2018-03-23 18:39:00
Given: when will I graduate on usjr
Answer Type: TIME SCH SUBJ-I >=

Figure 51 - Answer Type Detection Example

Figure 51 is an example of the answer type detected.

Question			
id:	#	userId:	#
createdOn:	2018-03-23 18:39:00	answerType:	TIME SCH SUBJ-I >=
dtStart:	null	dtEnd:	null
content:	when will I graduate on usjr		
subject:	[@PERSON PRP-I -> : : I_PRP]		
aux:	[@AUX -> : : will_MD]		
keywords:	[@LOCATION -> on_IN : : usjr_NNP]		
verb:	[@VB-PH -> : : graduate_VB]		

Figure 52 - Question Example

Figure 52 is an example of the question information.

2. Knowledge Retrieval

The process selects the potential knowledge, as candidate answers, from the user knowledge base. It reads the question and produces the knowledge retrieved.

It retrieves knowledge according to storage type either within the base time or next recent date (see Figure 7 for the process). A schedule is retrieved by start time and end time within the base time or next recent date, while medication by start time, and logs by creation time (see Figure 9 for the knowledge structure).

The selection for schedules includes those in between and overlapping the query time where

$$query\ start\ time \leq\ schedule\ time\ \leq query\ end\ time \text{ or}$$
$$schedule\ start\ time \leq\ query\ start\ time \leq\ schedule\ end\ time \text{ or}$$
$$schedule\ start\ time \leq\ query\ end\ time\ \leq schedule\ end\ time$$

Knowledge retrieved within the base time iterates once and compares the answers within the selection. While knowledge retrieved on the next recent date iterates according to iteration direction starting from the current date to the next knowledge entry's date and stops when the answer is found on that day, or the end of knowledge is reached.

	1st next recent date: 2018-03-23	2nd next recent date: 2018-03-24
	i attended the job fair	2018-03-12 00:00:00 – 2018-03-12 23:59:00
2nd pass ↓	i will graduate from usjr	2018-03-24 00:00:00 – 2018-03-24 23:59:00
X 3nd pass ↓	appointment	2018-03-25 13:00:00 – 2018-03-25 13:00:00
↓	meeting with Kenneth	2018-03-25 13:00:00 – 2018-03-25 13:00:00

Figure 53 - Knowledge Retrieval Example

Figure 53 contains schedule entries based on the scenario in user interface design (see Figure 12-16). Only the 2nd schedule entry was retrieved since the entry is within the 2nd next recent date and knowledge retrieval stops assuming the answer is found on answer processing.

3. Answer Processing

The system generates the top answers to the question. It involves the answer generation and answer validation.

3.1 Answer Generation

The process analyzes the question and knowledge retrieved to produce the candidate answers.

The entity-based question does entity matching to knowledge to find the correct one containing the answer. Also, some questions can predetermine the correct knowledge, by looking ahead the its entities if it contains the distinct answer type like quotation or cardinal.

The entry-based question skips entity matching. However, since some entry-based question has possible other entity values to match like for medication and schedule, single storage type of entry-based question will undergo entity matching as an exception.

Entity matching compares two entities or chunks (the knowledge representation) if both are the same. Question verb uses the auxiliary verb if there is no verb phrase.

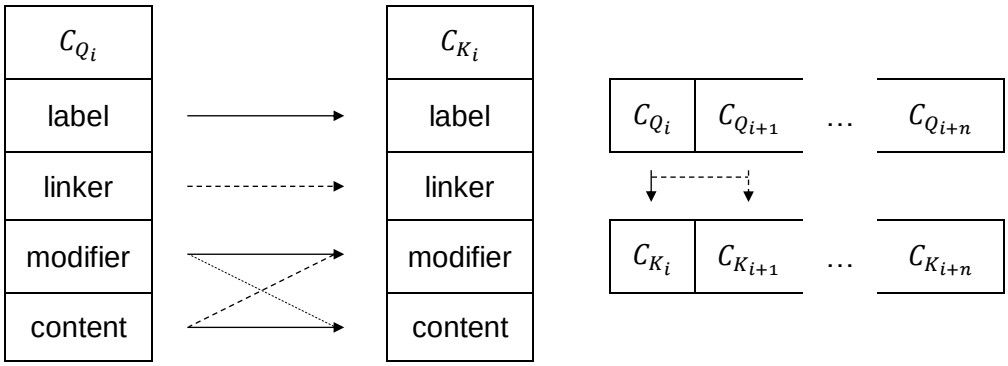

Figure 54 - Entity Matching Conceptual Model

In Figure 54, the entity matcher check for the same label then matches the entity value EV (modifier + content) through a single entity matching. Entity value includes noun lemma and main verb synonyms when matched. After a successful match, the knowledge entity is eliminated. Also, it implements the rules:

- Auxiliary verb matches an auxiliary verb
- Location preposition must belong to the same group (see Figure 41) to match
- Person honorific must belong to the same group (see Figure 40) to match
- Possessives matches through modifier to content and vice versa
- Of-phrases do entity matching by pair where C_{Q_i} and $C_{Q_{i+1}}$ (of-phrase) matches C_{K_i} or both with $C_{K_{i+1}}$ (of-phrase) and vice versa

Each candidate answer gives an answer score A. For entity matching, the formula computes the inverse score for each question chunk which gives zero for a full match

$$A = \sum_{i=0}^{n} compInvScore(C_{Q_i}) \quad compInvScore(C_{Q_i}) = \frac{EV_{mismatch}}{EV_{total\ size}}$$

Entity matching follows question subject to knowledge entity matching, question verb to knowledge entity matching, and question keywords to knowledge entity matching. Note, some questions don't have keywords which may skip the last step.

Question subject to knowledge entity matching must gain a zero score to proceed. The exceptions to this rule are:

- attribute entity NAME may not be in the knowledge representation, given all other entity value beside "name" matches for this entity
- a knowledge representation in passive voice makes "I" by default as question subject

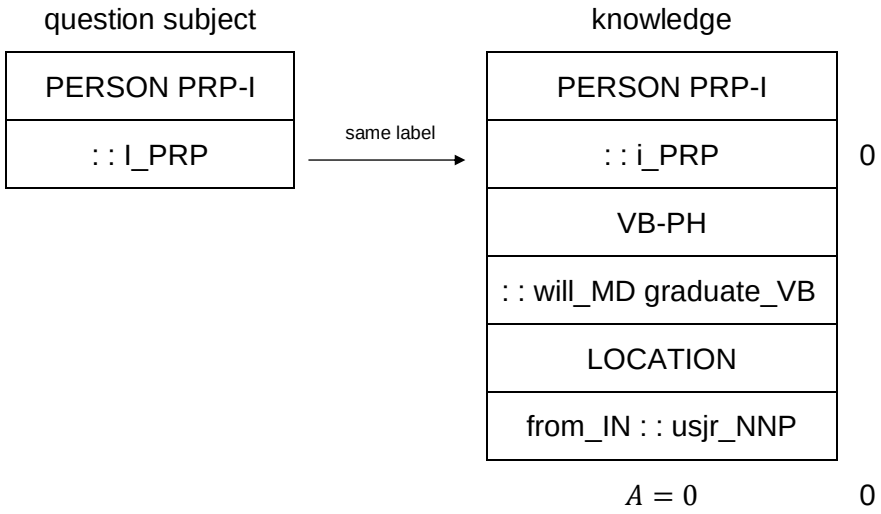

Figure 55 – Question Subject to Knowledge Entity Matching

Question verb to knowledge entity matching would accept any score to proceed unless the question verb is an auxiliary verb, and the answer type is for a person, location, or entity.

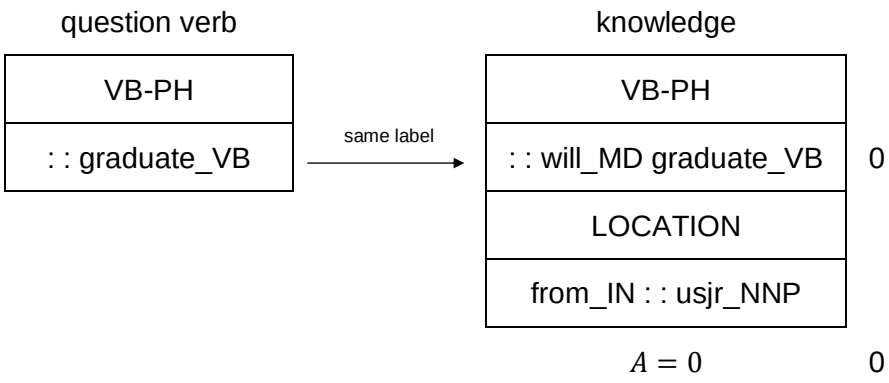

Figure 56 – Question Verb to Knowledge Entity Matching

Question keywords to knowledge entity matching would accept any score to proceed if at least one content from each keyword matches. If the knowledge entities are all eliminated and the answer type is not a time nor yes or no, hence, there is no answer, so the candidate answer is removed.

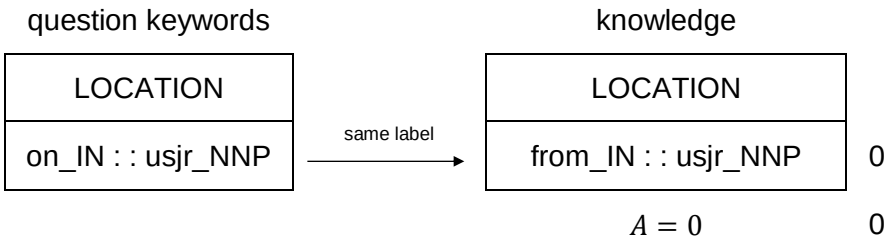

Figure 57 – Question Keywords to Knowledge Entity Matching

The candidate answer can now infer the answer to the question.

```
IF answer type is entry THEN answer is the knowledge content
ELSE IF answer type is yes or no THEN answer is yes
ELSE IF answer type is time
    IF knowledge entity contains time entity THEN answer is that entity
    ELSE IF knowledge is schedule THEN answer is from start time to end time
    ELSE IF knowledge is medication THEN answer is from start time to end time
    ELSE answer is creation time
ELSE
    answer is the knowledge entity that exactly match the answer entity type
    IF no answer found and answer type is not entity
```

```
        IF knowledge entity contains “entity” or attribute entity
        THEN answer is that entity
                add 1 to answer score
    IF answer type contains name and there is answer
    THEN answer is the answer entity that contains proper noun
```

Figure 58 - Answer Generation Pseudocode

In Figure 58, the answer is generated and validated at the same time. If no answer is found, the candidate answer is removed from the list. The answer for yes or no is yes since it has reached this point that means the knowledge is existing. Predetermined candidate answer always has an exact match. In some cases, the asked for will not exactly match due to the inaccuracy of information extraction; the closest possibility is that the answer may be in the general entity.

3.2 Answer Validation

The process verifies if the candidate answer conforms to the answer type, re-ranks the answers, and selects and formats the final answers.

The candidate answers are sorted by answer score and recency.

The top answers are selected and further narrowed down to first or last top answer if the answer type contains first or last with regards to iteration direction.

The final answers are formatted. The entry-based question has long answers (sentence form), and the entity-based question has a short answer (a fragment). Also, the knowledge type and time representation are appended on the end if the answer type is not time.

Another attempt on formatting the answers is the answer summarization. In answer summarization, the system groups the answers. It implements the rules:

- answers for time is grouped together if the answer is not a time entity

- past or future medication intake is grouped separately given the end chunk is a be verb or the verb is a medication verb (see Figure 61)

```
Current time: 2018-03-23 18:39:00
Given: when will I graduate on usjr
Answer(s): [from 12:00 AM to 11:59 PM, tomorrow]
```

Figure 59 – Final Answer Example

Figure 59 is an example of the final answer generated.

Scheduling

This component is composed of three main processes: scheduling classification and schedule process or medication process.

1. Scheduling Classification

The process follows (1) mapping the verb and scheduling entity, and (2) classifies the scheduling intent of the natural language input. It reads the entities and time information and produces the scheduling intent (see Figure 9 for the list).

Schedule and medication are time-based knowledge. Each has a collection of keywords to help identify their scheduling intent aside from time information.

```
SCH-KEYWORD → appointment | schedules | meeting
MED-KEYWORD → medicine | medications
```

Figure 60 - Scheduling Keywords

In Figure 60, the scheduling keywords include schedule and medication keyword.

```
SCH-VB-CANCEL → adjourn | cancel | postpone
SCH-VB-MOVE → change | move | reschedule | transfer | update
MED-VB → take | ingest | drink
```

Figure 61 - Scheduling Verbs

In Figure 61, the scheduling verb includes the cancel schedule verb, move schedule verb, and medication verb.

```
IF scheduling entity is medication entity
    IF there are medication time units
        IF numOfDays > 30 or intervalHour > 23 or intervalMinute > 59
        THEN unsupported time
        ELSE IF datetime start < current time
        THEN unknown intent
        ELSE scheduling intent is record medication interval
    ELSE IF datetime start > current time
    THEN scheduling intent is record medication
    ELSE IF datetime start < current time and datetime start != -1
    THEN scheduling intent is record past medication
    ELSE scheduling intent is record medication log
ELSE IF verb is move schedule verb
    IF verb contains not THEN unknown intent
    ELSE scheduling intent is move schedule
ELSE IF verb is cancel schedule verb
    IF verb contains not THEN unknown intent
    ELSE scheduling intent is cancel schedule
ELSE IF verb contains not and auxiliary verb matches “can|will|do”
THEN scheduling intent is cancel schedule
ELSE IF there is datetime start
   IF datetime start <= current time THEN scheduling intent is add past schedule
   ELSE IF main verb is not past tense THEN scheduling intent is add schedule
   ELSE IF auxiliary verb mismatches “was|were|had|did”
   THEN scheduling intent is add schedule
   ELSE unknown intent
```

Figure 62 - Classify Scheduling Intent Pseudocode

In Figure 62, the scheduling intent is determined by checking for scheduling verb and keywords and by inspecting the time information.

Current time: 2018-01-14 13:00:00

Given: robitussin medicine for 2 days every 6 hours and 30 minutes

(1)

MED
: : robitussin_NN medicine_NN

↑ scheduling entity

(2) Scheduling intent → RECORD MEDICATION INTERVAL

Figure 63 - Scheduling Classification Example 1

Figure 63 is an example of a medication interval that goes through step (1-2) of this process.

Current time: 2018-01-20 09:34:00

Given: the flight is on Jan. 21 from 11:00 PM to 1:00 AM

(1)

ENTITY	AUX
: the_DT : flight_NN	: : is_VBZ

↑ scheduling entity ↑ verb entity

(2) Scheduling intent → ADD SCHEDULE

Figure 64 - Scheduling Classification Example 2

Figure 64 is an example of a schedule that goes through step (1-2) of this process.

2.a Schedule Process

The processes involve checking for conflict if adding or moving a schedule, checking the base schedule if canceling or moving a schedule, and applying changes on the calendar and knowledge base (based on scheduling intent). It reads the entities and time information and produces the new or updated schedule as calendar entry and knowledge.

Schedule processes correspond to scheduling intent: add schedule, add past schedule, move schedule, or cancel schedule.

Cancel schedule is either time-based or entity-based, where there is no base time. Move schedule is only time-based. Time-based or entity-based retrieves the base knowledge within the base time or next recent date, respectively, similar to knowledge retrieval of QA. The retrieved base knowledge undergoes schedule to base knowledge entity matching like in QA.

Check for conflict verifies if no schedule exists at the same time with the new schedule which prompts the user to continue. Check base schedule verifies if one base schedule exists which prompts the user to choose the right base schedule from the multiple base schedule. Apply changes on the calendar may be skipped when there is no calendar entry corresponding the base schedule.

Apply changes on the calendar or knowledge base are create, delete, and update operations on the calendar or the database. Database deletion marks the schedule as not existing.

2.b Medication Process

The processes involve applying changes on the calendar and knowledge base. It reads the entities and time information and produces the medication(s) as calendar entry and knowledge.

Medication processes correspond to scheduling intent: record medication interval, record medication, record past medication, or record medication log.

Record medication interval creates multiple medication and calendar event(s) from the base medication with increasing time by interval hour and minute until the day count is greater than the number of days.

Record medication or past medication is likewise a special schedule while record medication log is a special log.

Testing Process

This section contains the results for conducting an integration test of the different components similar to the REST web service. Since scheduling process requires user prompt during a conflict of schedule and occurrence of multiple base schedule, the first schedule is selected by default while the QA shows all the answers. The dataset of 100 user knowledge and 100 questions is found on appendix D, together with the results.

The evaluation metric used for testing the performance of question classification and question answering is the accuracy since it is the most intuitive performance measure and it is a ratio of correct observations over the total observations. The precision, recall, and F-measure may be computed for validation, since accuracy is a great measure when datasets are in balance where values of false positive and false negatives have a similar cost, from the F-measure.

A confusion matrix is used to visualize the identical sets of classes in prediction and actual dimensions in a contingency table. The confusion matrix consists of the false positives (FP) or type I error – incorrect positive assignments, false negatives (FN) or type II error – incorrect negative assignments, true positives (TP) – correct positive assignments, and true negatives – correct negative assignments.

The formula for computing the accuracy A is

$$A = \frac{TP + TN}{TP + TN + FP + FN} \; or \; A = \frac{correct\ observations}{total\ observations}$$

The formula for computing the precision P, recall R, and F-measure F of the question classification are

$$P = \frac{TP}{TP + FP} \quad R = \frac{TP}{TP + FN} \quad F = 2 * \frac{P * R}{P + R}$$

Question Classification

The question classifier distinguishes a valid or invalid question; hence, the binary classification model is applicable since there are two groups or classes classified.

Actual / Predicted	valid	invalid	
valid	(TP =) 95	(FP =) 3	98
invalid	(FN =) 0	(TN =) 2	2
	95	5	100

Table 6 - Question Classification Confusion Matrix

The confusion matrix from Table 6 shows the predicted and actual number of questions classified as valid and invalid out of the 100 questions.

	valid	invalid
Precision	95 / 98 ≈ 97%	2 / 2 = 100%
Recall	95 / 95 = 100%	2 / 5 = 40%
F-measure	(2 * 0.97 * 1) / (0.97 + 1) ≈ 98%	(2 * 1 * 0.4) / (1 + 0.4) ≈ 57%
Accuracy	(95 + 2) / 100 = 97%	

Table 7 - Question Classification Evaluation Results

In Table 7, the question classification evaluated with 97% accuracy with high precision, high recall, and high F-measure for valid questions while high precision, low recall, and moderate F-measure for invalid questions.

Question Answering

Generating the answer depends on the given knowledge and the behavior of the question answering. Moreover, answers vary based on the question types or categories which are yes or no, what, when, where, and who. The test case provided consists more

of the intended usage scenarios and how to retrieve the information in case of null (no or none) answers. The evaluation is done by rate and by multiclassification model.

The answer is classified into three groups as correct, incorrect, and null, in general. To be more specific, correct answers consist of exact and fair answers, and incorrect answers consist of inexact and incorrect answers, in which, they are factors when judging the rate. The fair answers are accepted answers while inexact answers are lacking or unwanted answer, but, in a sense, answers the question.

Answer / Category	correct		incorrect		null	
	exact	fair	inexact	incorrect		
yes or no	7	1	1	1	5	15
what	22	5	1	6	2	36
who	11	3	3	1	1	19
where	5	0	1	1	1	8
when	10	1	1	0	5	17
	55	10	7	9	14	95

Table 8 - Question Answering Summary

The summary from Table 8 shows the rate of the answers as exact, fair, inexact, incorrect, and null, relating categories of the question of the 95 valid questions.

Actual / Predicted	correct	incorrect	null	
correct	59	4	14	77
incorrect	0	0	0	0
null	0	2	16	18
	59	6	30	95

Table 9 - Question Answering Confusion Matrix

The confusion matrix from Table 9 shows the predicted and actual number of answers classified as correct, incorrect, and null out of the 95 valid questions.

	correct	incorrect	null
Precision	59 / 77 ≈ 77%	undefined	16 / 18 ≈ 89%
Recall	59 / 59 ≈ 100%	0 / 6 = 0%	16 / 30 ≈ 53%
F-measure	(2 * 0.77 * 1) / (0.77 + 1) ≈ 87%	undefined	(2 * 0.89 * 0.53) / (0.89 + 0.53) ≈ 67%
Accuracy	(59 + 0 + 16) / 95 = 79%		

Table 10 - Question Answering Evaluation Results

In Table 10, the question answering evaluated with 79% accuracy with high recall, high precision, and high F-measure for correct answers while high precision, moderate recall, and moderate F-measure for null answers. As expected, the incorrect answer gave an undefined and zero probability, since incorrect answers are unpredicatable, but correct and null answers may be incorrect in the actual test.

CHAPTER IV

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

SUMMARY OF FINDINGS

This project captured 13 entities and 15 temporal expressions on information extraction, explored three knowledge representations for the knowledge base, investigated eight main answer types on question answering, and performs eight scheduling intents for the scheduling. Not having a variety of entities does not obstruct finding the intent and answer. Entity recognition for common nouns and modifiers have their uses. The use of prepositions, comma, and articles may make longer text, but it increases the accuracy of information extraction. Time analysis, from temporal expressions, verb tenses, and knowledge representation, was informative in different ways. The knowledge types were advantageous; their characteristics were practical to their usage. Given a dataset of 100 knowledge and 100 questions, the question answering evaluated about 79% accuracy for classifying correct, incorrect, and null answers out of 95 valid questions, with 53% and above precision, recall and F-measure. Retrieval methods through question answering include rephrasing the question from a specific (where) question to a generic (what) question, from an entity-based question to an entry-based question, from moving specified range of time by the last answer's time, and from no base time to specifying a base time. Combining keyword, synonym, and rule-based approaches do contribute to question answering. Constructing general or refined rules reduces the occurrence or resolving of rules in conflict. Scheduling processes are straightforward to their intent, and it relies on the temporal extraction for the time information.

CONCLUSION

The study has achieved its purpose to develop a mobile application for scheduling and question answering systems with REST web services. Natural language is diverse, and there are a lot of ways expressing the same question and knowledge. Choosing the knowledge domain determines how it can be useful as an application. This question answering is not about answering all questions but recalling information unassociated with one question but different retrieval methods.

RECOMMENDATIONS

Initially, the app was intended to be used with a wireless headset (or earphone with a microphone) to experience a hands-free interaction. Memorizing mode can allow the user to record a script or speech and rehearse through the app. Test mode can allow the user to select knowledge and the system can generate questions from it with a timer. Other features may include the user to listen to his knowledge throughout the day and to select and listen to one (or more) knowledge in sequence repetitions. In the medical field, there could be an app for the doctor and for the patient where it can monitor the patient's performance. Through data analytics, it can associate rules in what words or memories trigger the patient's memories. The QA can add more test cases by constructing more rules and knowledge representations such as the definition of terms or modifiers.

BIBLIOGRAPHY

[1] A. Pundge, K. S.A. and C. N. Mahender, "Question Answering System, Approaches and Techniques: A Review," *International Journal of Computer Applications,* vol. 141, no. 3, 2016.

[2] A. Copestake, 2003. [Online]. Available: https://www.cl.cam.ac.uk/teaching/2002/NatLangProc/nlp1-4.pdf.

[3] D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed., 2018.

[4] D. Khurana, A. Koli, K. Khatter and S. Singh, "Natural Language Processing: State of The Art, Current Trends and Challenges," 2017.

[5] R. J. Mooney and R. Bunescu, "Mining knowledge from text using information extraction," *ACM SIGKDD Explorations Newsletter,* vol. 7, no. 1, pp. 3-10, June 2005.

[6] "Knowledge-based systems," [Online]. Available: https://en.wikipedia.org/wiki/Knowledge-based_systems#cite_note-3. [Accessed 17 10 2018].

[7] W. Cui, Y. Xiao, H. Wang, Y. S. S.-w. Hwang and W. Wang, "KBQA: Learning Question Answering over QA Corpora," *Proceedings of the VLDB Endowment,* vol. 10, no. 5, pp. 565-576, 2017.

[8] A. Bordes, N. Usunier, S. Chopra and J. Weston, "Large-scale Simple Question Answering with Memory Networks," 2015.

[9] B. Katz and J. Lin, "Question answering from the web using knowledge annotation and knowledge mining techniques," *CIKM '03 Proceedings of the twelfth international conference on Information and knowledge management,* pp. 116-123, 2003.

[10] B. Katz, G. Borchardt and S. Felshin, "Natural Language Annotations for Question Answering," *Proceedings of the 19th International FLAIRS Conference,* 2006.

[11] B. Magnini, M. Negri, R. Prevete and H. Tanev, "A WordNet-based approach to Named Entities recognition," *SEMANET '02 Proceedings of the 2002 workshop on Building and using semantic networks,* vol. 11, pp. 1-7, 2007.

CURRICULUM VITAE

PROFILE

- Kenneth E. Sanchez
- Ramona Village, Dumlog, Talisay City, Cebu, 6045 Philippines
- +63 927 352 9278 / + 63 920 255 4495
- zehcnas.e.htennek@gmail.com

EDUCATIONAL BACKGROUND

- Tertiary: University of San Jose-Recoletos, A.Y. 2014 – 2017
    - Specializing as Data Analyst
    - Consistent Dean's Lister since 2014 - 2nd sem 2016
- Secondary: Talisay Malayan Academy, A.Y. 2010 - 2014

ACQUIRED SKILLS

- Cebuano, English, Nihongo, Tagalog Language
- Developing applications: desktop, game, mobile, web
- Tools Have Used: Amzi! Prolog + Server, Eclipse, Android Studio, IntelliJ, MySQL DB, Netbeans, Oracle DB, R Studio, Unity, Visual Studio, Wordpress, Emu 8086
- Formal Writing, Research Writing
- Programming in Assembly, C, C++, C#, Java, Javascript & JQuery, J2EE, Prolog, Python, R

- Other: HTML & CSS, Junit, Mockito, MySQL, Oracle SQL, Spring Boot, Spring MVC

EXPERIENCES

- OJT Web Developer | Yoshii Software Solutions Philippines Co. | May 3 - Aug. 18 2017
- Project Analyst | Accenture's Program The Future Semifinalists 2017 | June 15 - Aug. 19 2017
- Project Analyst & Designer | Microsoft Imagine Cup National Semifinalists 2017 | Feb. 11 - Apr. 3 2017

LEADERSHIP

- Information and Computer Technology Organization (ICTO) VP - Ways and Means, effective A.Y. 2016-2017
- Kabataan para sa Sambayanang Makatarungan Leaders Movement (KASAMA) CICCT HR candidate and member, 2016

APPENDIX

A. Penn TreeBank English POS Tag Set

TAG	DESCRIPTION	TAG	DESCRIPTION
CC	Coordinating Conjunction	PRP$	Possessive Pronoun
CD	Cardinal Number	RB	Adverb
DT	Determiner	RBR	Adverb, Comparative
EX	Existential There	RBS	Adverb, Superlative
FW	Foreign Word	RP	Particle
IN	Preposition or Subordinating Conjunction	SYM	Symbol
JJ	Adjective	TO	To
JJR	Adjective, Comparative	UH	Interjection
JJS	Adjective, Superlative	VB	Verb, Base Form
LS	List Item Marker	VBD	Verb, Past
MD	Modal	VBG	Verb, Gerund or Present Participle
NN	Noun, Singular	VBN	Verb, Past Participle
NNS	Noun, Plural	VBP	Verb, Present, Non-3rd Person Singular
NNP	Proper Noun, Singular	VBZ	Verb, Present, 3rd Person Singular
NNPS	Proper Noun, Plural	WDT	Wh-determiner
PDT	Pre-determiner	WP	Wh-pronoun
POS	Possessive Ending	WP$	Possessive Wh-pronoun
PRP	Personal Pronoun	WRB	Wh-adverb

B. CFG for chunk

PP → *PREP NP1*

NP1 → *MOD NPJ N CD2* | *MOD NPJ N* | *MOD N CD2* | *MOD N*

→ *NPJ N CD2* | *NPJ N* | *N CD2* | *N* | *PRP*

NP2 → *CONJ PRP* | *CONJ NP*

VP → *ADV AUXV ADV VERB* | *AUXV ADV VERB* | *ADV VERB*

→ *AUXV VERB* | *VERB*

PREP → *PREP PREP* | *IN* | *TO*

CONJ → *CONJ CONJ* | *CC*

MOD → *DET CD2 ADV ADJ* | *DET CD2 ADJ* | *DET CD2* | *DET ADJ*

→ *DET* | *CD2* | *ADJ* | *POS2*

NPJ → *NPJ NPJ* | *N JJ*

AUX → *ADV AUXV* | *AUXV ADV* | *AUXV*

CD2 → *CD2 CD2* | *CD*

POS2 → *POS2 POS2* | *POS*

DET → *PDT DT* | *PDT PRP$* | *DT* | *PRP$*

ADJ → *ADJ ADJ* | *JJ* | *JJR* | *JJS*

ADV → *ADV ADV* | *RB* | *RBR* | *RBS*

N → *N N* | *FW* | *NN* | *NNS* | *NNP* | *NNPS* | *NM*

AUXV → *MD* | am | are | can | could | dare | do | does | did | had | has | have | is | may

→ might | must | need | ought | shall | should | was | were | will | would

VERB → *VERB VERB* | *VB* | *VBD* | *VBN* | *VBP* | *VBZ* | *VBG*

WH → *WGT* | *WP* | *WP$* | *WRB*

QQ → *QQ*

THERE → *EX*

C. CFG for question

WH-Q → *WH-WORD AUX-Q*

→ *WH-WORD SUBJ AUX KEYWORDS VB-PH*

AUX-Q → *AUX-Q-DO KEYWORDS* | *AUX-Q-DO*

→ *AUX-Q-BE KEYWORDS* | *AUX-Q-BE*

→ *AUX-Q-HAD KEYWORDS* | *AUX-Q-HAD*

→ *AUX-Q-WILL KEYWORDS* | *AUX-Q-WILL*

WH-WORD → what | who | where | when | to whom | whose

AUX-Q-DO → (did | do | does) *SUBJ* (*VB* | *VBP* | have)

→ (did | do | does) *SUBJ*

AUX-Q-BE → (am | is | are | was | were) *SUBJ* (*VBG* | *VBN*)

→ (am | is | are | was | were) *SUBJ*

AUX-Q-HAD → (had | has | have) *SUBJ* (*VBD* | *VBN* | been)

→ (had | has | have) *SUBJ* been VBG

AUX-Q-BE → will *SUBJ* (*VB* | *VBP* | *VBZ* | be)

→ will *SUBJ* be *VBG*

→ will *SUBJ* have (been | *VBN*)

→ will *SUBJ* have been *VBG*

D. Testing Set Results

This is composed of the 100 knowledge and 100 questions dataset with the response. For the medication interval, the start time and end time indicates the time of the first and last intake, for display purposes, while, for logs, medication logs, and questions with no time, the start time indicates the creation time. Answer rate is equivalent to 4 as exact answer, 3 as fair answer, 2 as inexact answer, 1 as incorrect answer, and 0 as null answer.

#	START TIME	END TIME	USER MESSAGE	APP MESSAGE	RATE
1	2018/01/31 9:00 AM		my mother is Chery Lim	Log successfully recorded!	
2	2018/01/31 9:01 AM		Mr. Andrew Bacud is the name of my father.	Log successfully recorded!	

3	2018/01/31 9:02 AM		what is the name of my mother?	[Chery Lim, based on log created on 9:00 AM, today]	4
4	2018/01/31 9:03 AM		what is father's name?	[Mr. Andrew Bacud, based on log created on 9:01 AM, today]	4
5	2018/01/31 9:04 AM		my mother is a youtuber	Log successfully recorded!	
6	2018/01/31 9:05 AM		what is the job of my mother	[No answer found]	3
7	2018/01/31 9:06 AM		who is my mother	[a youtuber, based on log created on 9:04 AM, today] [Chery Lim, based on log created on 9:00 AM, today]	3
8	2018/01/31 9:10 AM		I place the keys on the cabinet	Log successfully recorded!	
9	2018/01/31 5:11 PM	2018/01/31 5:11 PM	schedule a dinner with parents for next 8 hours	Schedule successfully recorded!	
10	2018/01/31 1:00 PM		I put the keys above the fridge.	Log successfully recorded!	
11	2018/01/31 5:30 PM		did I place the keys under the fridge	[No]	0
12	2018/01/31 5:31 PM		when did I place the keys above the fridge	[at 1:00 PM, today]	4

13	2018/02/01 10:00 AM		I hid the keys under the floormat on the front door	Log successfully recorded!	
14	2018/02/01 1:00 PM		where did I place the keys?	[above the fridge, based on log created on 1:00 PM, yesterday] [on the cabinet, based on log created on 9:10 AM, yesterday]	2
15	2018/02/01 1:01 PM		where are the keys	[No answer found]	1
16	2018/02/01 11:59 PM		the medicine for my cough is robitussin	Medication log successfully recorded!	
17	2018/02/02 6:00 AM	2018/02/04 6:00 PM	I will take robitussin medicine for 3 days every 6 hours	Medication interval successfully recorded!	
18	2018/02/02 6:00 AM		I have taken my medicine	Medication log successfully recorded!	
19	2018/02/02 6:01 AM		what is the interval for taking robitussin medicine?	[No answer found]	1
20	2018/02/02 9:00 AM		when will I take my medicine	[at 12:00 PM, at 6:00 PM, today]	4
21	2018/02/05 12:00 AM	2018/02/05 11:59 PM	meeting with Sir leeroy on next Monday	Schedule successfully recorded!	
22	2018/02/02 10:00 AM		I listen to Charlie Puth's music	Log successfully recorded!	

23	2018/02/02 12:00 PM		I have taken my medicine	Medication log successfully recorded!	
24	2018/02/02 6:00 PM		took medication	Medication log successfully recorded!	
25	2018/02/03 12:00 AM		I have taken my medicine	Medication log successfully recorded!	
26	2018/02/03 6:00 AM		have taken my medication	Medication log successfully recorded!	
27	2018/02/03 9:00 AM		I buy clothes on gaisano tabunok mall	Log successfully recorded!	
28	2018/02/03 12:00 PM	2018/02/03 12:00 PM	I will buy food on the grocery at noon	Schedule successfully recorded!	
29	2018/02/03 11:59 AM		where will I buy food	[on the grocery, based on schedule at 12:00 PM, today]	4
30	2018/02/03 12:00 PM		took medication	Medication log successfully recorded!	
31	2018/02/03 1:00 PM		where do I buy clothes	[on gaisano tabunok mall, based on log created on 9:00 AM, today]	4
32	2018/02/03 1:01 PM		my favorite band is F.M. static	Log successfully recorded!	
33	2018/04/01 12:00 AM	2018/04/01 12:00 AM	I will sing happy birthday to Jem at 12 of midnight on 1st of Apr. in 2018	Schedule successfully recorded!	

34	2018/02/03 1:03 PM		I hate math.	Log successfully recorded!	
35	2018/02/03 6:00 PM		took medication	Medication log successfully recorded!	
36	2018/02/03 6:01 PM		what music do I listen	[No answer found]	1
37	2018/02/03 6:02 PM		what do I listen?	[to Charlie Puth's music, based on log created on 10:00 AM, yesterday]	4
38	2018/02/03 6:03 PM		who is my favorite band	[F.M. static, based on log created on 1:01 PM, today]	4
39	2018/02/04 12:00 AM		have taken my medication	Medication log successfully recorded!	
40	2018/02/04 6:00 AM		I have taken my medicine	Medication log successfully recorded!	
41	2018/02/04 9:00 AM		my first dog is Connor	Log successfully recorded!	
42	2018/02/04 9:01 AM		my 2nd dog is Hank	Log successfully recorded!	
43	2018/02/05 12:00 AM	2018/02/05 11:59 PM	what will I be doing for tomorrow	[You scheduled "meeting with Sir leeroy" from 12:00 AM to 11:59 PM, tomorrow]	4
44	2018/02/04 12:00 PM		took medication	Medication log successfully recorded!	

45	2018/02/03 12:00 AM	2018/02/03 11:59 PM	what are my medication for yesterday	[You recorded "took medication" or "have taken my medication" or "I have taken my medicine" at 6:00 PM, at 12:00 PM, at 6:00 AM, at 12:00 AM, yesterday] [You recorded "I will take robitussin medicine" at 6:00 PM, at 12:00 PM, at 6:00 AM, at 12:00 AM, yesterday]	4
46	2018/02/05 12:00 AM	2018/02/05 11:59 PM	no class on next Monday	Schedule successfully recorded!	
47	2018/02/04 5:31 PM		I will not meet Sir Leeroy for tomorrow	Future base schedule does not exist.	
48	2018/02/04 6:00 PM		have taken my medication	Medication log successfully recorded!	
49	2018/02/05 12:00 AM		I will take robitussin medicine at 12 a.m., tomorrow	Medication successfully recorded!	
50	2018/02/05 12:00 AM		have taken my medication	Medication log successfully recorded!	
51	2018/02/05 2:30 PM	2018/02/05 3:30 PM	I have an appointment with the doctor at chong hua hospital from 2:30 PM to 3:30 PM	Schedule successfully recorded!	

52	2018/02/05 7:00 AM		where is my meeting with sir leeroy	[No answer found]	0
53	2018/02/05 7:01 AM	2018/02/05 11:59 PM	is there a class for today	[Yes, based on schedule from 12:00 AM to 11:59 PM, today]	1
54	2018/02/02 12:00 AM	2018/02/02 11:59 PM	when did I first take my medicine on last Friday?	[on Feb 2 2018 at 6:00 AM]	4
55	2018/02/05 3:01 PM		when did I last take my medicine?	[at 12:00 AM, today]	4
56	2018/02/05 3:30 PM		the medicine for my headache is bioflu	Medication log successfully recorded!	
57	2018/02/05 3:31 PM		what was the medicine for my cough	[the medicine robitussin, based on medication on Feb 1 2018 at 11:59 PM]	3
58	2018/02/05 3:32 PM		I will take bioflu medicine starting this evening for 8 days every 24 hours 60 minutes	Time is either not supported or not recognized.	
59	2018/02/05 3:33 PM		I will take bioflu medicine for 1 week and 30 minutes	Time is either not supported or not recognized.	
60	2018/02/05 6:00 PM	2018/02/06 4:30 PM	I will take bioflu medicine starting this evening for 2 days every 7 hours every 30 minutes	Medication interval successfully recorded!	

61	2018/02/06 12:00 AM	2018/02/06 4:00 PM	starting tomorrow I will take every 8 hours robitussin medicine	Medication interval successfully recorded!	
62	2018/02/05 6:01 PM		when will I take bioflu	[at 1:30 AM, at 9:00 AM, at 4:30 PM, tomorrow]	4
63	2018/02/05 6:02 PM		what is the name of my second dog	[No answer found]	1
64	2018/02/05 8:01 PM		who are my dog	[Hank, based on log created on 9:01 AM, yesterday] [Connor, based on log created on 9:00 AM, yesterday]	4
65	2018/02/05 8:02 PM		do I hate math	[Yes, based on log created on Feb 3 2018 at 1:03 PM]	4
66	2018/02/06 8:00 AM		I am a computer science student.	Log successfully recorded!	
67	2018/02/06 8:01 AM		my name is president Jade Makig-angay	Log successfully recorded!	
68	2018/02/06 8:02 AM		I live in basak cebu city	Log successfully recorded!	
69	2018/02/06 8:03 AM		my crush is secret	Log successfully recorded!	
70	2018/02/06 9:00 AM		took bioflu medication	Medication log successfully recorded!	
71	2018/02/06 9:01 AM		had I drunk bioflu medicine	Medication log successfully recorded!	

72	2018/02/06 1:30 AM		have taken my bioflu medicine at 1:30 in the morning	Past medication successfully recorded!	
73	2018/02/05 6:00 PM	2018/02/05 11:59 PM	For last night, I watched Spiderman movie	Past schedule successfully recorded!	
74	2018/02/06 11:59 PM		I have taken all robitussin medication	Medication log successfully recorded!	
75	2018/02/07 8:00 AM		have taken robitussin medication	Medication log successfully recorded!	
76	2018/02/07 8:01 AM		Chicken tasted like salad	Log successfully recorded!	
77	2018/02/07 8:02 AM		I ate chicken for lunch on the restaurant of Ms. Jasmine	Log successfully recorded!	
78	2018/02/07 8:03 AM		feed the fish for 7 days every 8 hours	Time is either not supported or not recognized.	
79	2018/02/07 4:00 PM		have taken robitussin medication	Medication log successfully recorded!	
80	2018/02/07 9:00 PM		what did I watch	[No answer found]	1
81	2018/02/05 6:00 PM	2018/02/05 11:59 PM	what movie did I watch, last last night	[No answer found]	1
82	2018/02/05 6:00 PM	2018/02/05 11:59 PM	what did I watch, last last night	[Spiderman movie, based on schedule on Feb 5 2018 from 6:00 PM to 11:59 PM]	4

83	2018/02/07 12:00 AM		I have taken all bioflu medication for today	Past medication successfully recorded!	
84	2018/02/08 12:00 AM	2018/02/08 12:00 AM	I will take robitussin at midnight, tomorrow	Schedule successfully recorded!	
85	2018/02/07 10:32 PM		cancel schedule at midnight for tomorrow	Schedule at 12:00 AM, tomorrow successfully canceled!	
86	2018/02/08 12:00 AM		I will take robitussin medicine at 12 a.m., tomorrow	Medication successfully recorded!	
87	2018/02/07 11:30 PM		have I taken all my medication	[Yes, based on medication at 12:00 AM, today]	2
88	2018/02/07 11:31 PM		have I taken all bioflu medication	[Yes, based on medication at 12:00 AM, today]	4
89	2018/02/07 12:00 AM	2018/02/07 11:32 PM	have I taken all robitussin medication for today	[No]	3
90	2018/02/07 11:33 PM		have I taken my medication	[Yes, based on medication at 4:00 PM, today]	4
91	2018/02/08 12:00 AM		have taken robitussin medication	Medication log successfully recorded!	
92	2018/02/08 2:30 PM	2018/02/08 2:30 PM	I have an appointment with the doctor at chong hua by 2:30 PM	Schedule successfully recorded!	

93	2018/02/07 12:00 AM	2018/02/07 11:59 AM	what did I ingest in the morning, yesterday	[chicken for lunch on the restaurant of Ms. Jasmine, based on log created on 8:02 AM, yesterday] [robitussin medication, based on medication at 8:00 AM, yesterday]	3
94	2018/02/08 8:01 AM		whose restaurant did I eat chicken	[of Ms. Jasmine, based on log created on 8:02 AM, yesterday]	4
95	2018/02/08 8:02 AM		what do chicken taste like	Log successfully recorded!	
96	2018/02/09 9:00 AM		my school on college is in usj-r.	Log successfully recorded!	
97	2018/02/09 9:01 AM		my thesis Adviser's name is ma'am petralba	Log successfully recorded!	
98	2018/02/16 5:00 PM	2018/02/16 5:00 PM	I will see ma'am petralba, next Friday, at 5 pm	Schedule successfully recorded!	
99	2018/02/09 9:03 AM		I will visit my grandma tomorrow	Log successfully recorded!	
100	2018/02/12 12:00 AM	2018/02/12 11:59 PM	clan wars for next 3 days	Schedule successfully recorded!	
101	2018/02/09 9:00 PM		what is my name	[ma'am petralba, based on log created on 9:01 AM, today]	2

102	2018/02/09 9:01 PM		who am I	[a computer science student, based on log created on Feb 6 2018 at 8:00 AM]	3
103	2018/02/09 9:02 PM		where do I live	[in basak cebu city, based on log created on Feb 6 2018 at 8:02 AM]	4
104	2018/02/09 9:03 PM		who is my crush	[secret, based on log created on Feb 6 2018 at 8:03 AM]	4
105	2017/07/01 12:00 AM	2017/07/01 11:59 PM	acquaintance party with Jake and Gomez was on last July	Past schedule successfully recorded!	
106	2018/02/11 12:00 AM	2018/02/11 4:00 PM	starting tomorrow I will take every 8 hours robitussin medicine	Medication interval successfully recorded!	
107	2018/02/12 9:00 AM	2018/02/12 6:00 PM	duty at school from 9 a.m. to 6 p.m. on next Monday	Schedule successfully recorded!	
108	2019/10/01 12:00 AM	2019/10/01 11:59 PM	archcon will be held at oakridge pavillon for next October	Schedule successfully recorded!	
109	2018/02/10 9:00 PM		plate number is LUV-143	Log successfully recorded!	
110	2018/02/09 12:00 AM	2018/02/09 11:59 PM	when last is my schedule for yesterday	[No answer found]	0
111	2018/02/10 9:02 PM		when was the acquaintance party	[on Jul 1 2017 from 12:00 AM to 11:59 PM]	4

112	2018/02/10 9:03 PM		who was with me at the acquaintance party	[No answer found]	1
113	2018/02/10 9:04 PM		when will I visit my grandma at the hospital	[No answer found]	0
114	2018/02/11 9:00 AM		I have 100 dollars	Log successfully recorded!	
115	2018/02/11 1:00 PM		how much dollar do I have	Question is not supported by either its format or ambiguity or beyond the scope. Try rephrasing the question.	
116	2018/02/11 1:01 PM		what do I have	[100 dollars, based on log created on 9:00 AM, today]	4
117	2018/02/12 12:00 AM	2018/02/12 11:59 PM	do I have an appointment for tomorrow	[No]	0
118	2018/02/12 12:00 AM	2018/02/12 11:59 AM	what schedules will I have for tomorrow, morning	[You scheduled "clan wars" from 12:00 AM to 11:59 PM, tomorrow] [You scheduled "duty at school" from 9:00 AM to 6:00 PM, tomorrow]	4
119	2018/02/11 12:00 AM	2018/02/11 9:02 PM	what did I do, today	[You recorded "I will take robitussin medicine" at 4:00 PM, at 8:00 AM, at 12:00	4

				AM, today] [You recorded "I have 100 dollars" at 9:00 AM, today]	
120	2018/02/11 11:00 PM		Just believe in yourself and everything will be alright	Log successfully recorded!	
121	2018/02/12 12:00 PM	2018/02/12 5:59 PM	what is my meeting in the afternoon	[No answer found]	3
122	2018/02/12 9:01 AM		when is meeting	[No answer found]	3
123	2018/02/12 9:02 AM		who will I meet	[ma'am petralba, based on schedule on Feb 16 2018 at 5:00 PM]	4
124	2018/02/12 9:03 AM		when will I see mrs. Petralba	[on Feb 16 2018 at 5:00 PM]	4
125	2018/02/11 6:00 PM	2018/02/11 11:59 PM	what did I record at night, yesterday	[You recorded "Just believe in yourself and everything will be alright" at 11:00 PM, yesterday]	4
126	2018/02/12 9:00 PM		Ana Montecillo's friend is Dave Ryan Yuhayco	Log successfully recorded!	
127	2018/02/12 9:01 PM		Wenz Patrick is not Dave Yuhayco's friend	Log successfully recorded!	
128	2018/02/12 9:02 PM		Ms. Nicole Robledo's number is 09876543210	Log successfully recorded!	

129	2018/03/01 12:00 AM	2018/03/01 11:59 PM	I will graduate on March	Schedule successfully recorded!	
130	2018/02/13 6:00 PM	2018/02/13 11:59 PM	finish my thesis for tonight	Schedule successfully recorded!	
131	2018/02/13 12:00 AM	2018/02/13 11:59 AM	I borrowed "Safe Haven" from Mr. Nicholas Sparks in the bookstore, this morning	Past schedule successfully recorded!	
132	2018/02/13 9:00 PM		where is my school on college	[in usj-r, based on log created on Feb 9 2018 at 9:00 AM]	4
133	2018/02/13 9:01 PM		what location will the archcon be held	[at oakridge pavillon, based on schedule on Oct 1 2019 from 12:00 AM to 11:59 PM]	4
134	2018/02/13 9:02 PM		to whom did I borrow "Safe Haven"	[No answer found]	2
135	2018/02/13 12:00 AM	2018/02/13 9:03 PM	to whom did I borrow "Safe Haven", today	[from Mr. Nicholas, based on schedule from 12:00 AM to 11:59 AM, today]	2
136	2018/02/13 9:04 PM		who are my friends	[No answer found]	0
137	2018/02/13 9:05 PM		who is my adviser	[ma'am petralba, based on log created on Feb 9 2018 at 9:01 AM]	4

138	2018/02/13 9:06 PM		who is mr. dave yuhayco	[No answer found]	3
139	2018/02/13 9:07 PM		what is dave yuhayco	[Ana Montecillo's friend, based on log created on 9:00 PM, yesterday]	4
140	2018/02/13 9:08 PM		who is ana montecillo's friend	[Dave Ryan Yuhayco, based on log created on 9:00 PM, yesterday]	4
141	2018/02/13 9:09 PM		what is ana montecillo's friend	[Dave Ryan Yuhayco, based on log created on 9:00 PM, yesterday]	4
142	2018/02/13 9:10 PM		whose number is 09876543210	[Ms. Nicole Robledo's number, based on log created on 9:02 PM, yesterday]	4
143	2018/02/13 9:11 PM		Is Wenz patrick a friend of Dave Yuhayco	[No]	0
144	2018/02/13 11:59 PM		will I finish the thesis	[Yes, based on schedule from 6:00 PM to 11:59 PM, today]	4
145	2018/02/13 12:00 AM	2018/02/13 11:59 PM	what have I done, yesterday	[You scheduled "finish my thesis" from 6:00 PM to 11:59 PM, yesterday] [You scheduled "I borrowed "Safe Haven" from Mr. Nicholas Sparks in the bookstore" from 12:00	4

				AM to 11:59 AM, yesterday]	
146	2018/02/14 9:00 PM		ken means "range of what one can know or understand"	Log successfully recorded!	
147	2018/02/14 9:01 PM		karma means "the effects of a person's actions that determine his destiny"	Log successfully recorded!	
148	2018/02/20 3:00 AM	2018/02/20 3:00 AM	I will go to Manila at 3 in the morning, next Tuesday	Schedule successfully recorded!	
149	2018/02/14 9:03 PM		can't go to Japan	Future base schedule does not exist.	
150	2018/02/15 9:00 AM		Life is a journey with full of trials	Log successfully recorded!	
151	2018/02/15 9:01 AM		Everyone make mistakes.	Log successfully recorded!	
152	2018/02/15 6:00 PM	2018/02/15 6:00 PM	my shift at 5 p.m.	Schedule successfully recorded!	
153	2018/02/15 4:59 PM		my 5 p.m. shift was changed to 6 p.m.	Schedule at 5:00 PM, today successfully moved!	
154	2018/02/15 5:30 PM		when is my shift	[at 6:00 PM, today]	4
155	2018/02/15 6:46 PM	2018/02/15 6:46 PM	take a break for next 45 minutes	Schedule successfully recorded!	

156	2018/02/15 9:00 PM		when will I be going to Japan	[No answer found]	0
157	2018/02/15 9:01 PM		when will I to Manila	Log successfully recorded!	
158	2018/02/16 9:00 AM		planting beans takes an average of 7 days to fully grow	I cannot understand what you're saying.	
159	2018/02/16 12:00 PM	2018/02/16 5:59 PM	I will clean the room in the afternoon	Schedule successfully recorded!	
160	2018/02/16 9:03 AM		what is everyone	[No answer found]	0
161	2018/02/16 9:04 AM		who make mistakes	[Everyone, based on log created on 9:01 AM, yesterday]	4
162	2018/02/16 9:05 AM		what does ken mean	["range of what one can know or understand", based on log created on Feb 14 2018 at 9:00 PM]	4
163	2018/02/16 9:06 AM		what do karma mean	["the effects of a person 's actions that determine his destiny", based on log created on Feb 14 2018 at 9:01 PM]	4
164	2018/02/16 6:00 PM	2018/02/16 11:59 PM	will I clean the room on the evening	[No]	0
165	2018/02/16 1:00 PM		Albert Einstein said "I fear the day that technology will surpass	Log successfully recorded!	

			our human interaction. The world will have a generation of idiots."		
166	2018/02/16 1:01 PM		I will not move	I cannot understand what you're saying.	
167	2018/02/16 1:02 PM		I have read one piece on chapter 863	Log successfully recorded!	
168	2018/04/21 12:00 AM	2018/04/21 11:59 PM	the flight is on April 21st	Schedule successfully recorded!	
169	2018/02/17 12:00 AM	2018/02/17 9:00 PM	what did I first say for today	[No answer found]	3
170	2018/02/17 9:01 PM		what chapter am I in one piece	[No answer found]	1
171	2018/02/17 9:02 PM		where have I read in one piece	Question is not supported by either its format or ambiguity or beyond the scope. Try rephrasing the question.	
172	2018/02/18 9:00 AM		the flight was postponed on April 21	Schedule on Apr 21 2018 from 12:00 AM to 11:59 PM successfully canceled!	
173	2018/02/18 1:00 PM		when is my flight	[No answer found]	0
174	2018/02/19 9:00 PM		there are 7107 islands on the Philippines	Log successfully recorded!	

175	2018/02/19 9:01 PM		President Duterte is a strong leader of the Philippines	Log successfully recorded!	
176	2018/02/19 9:02 PM		parked the car at base floor 2	Log successfully recorded!	
177	2018/02/19 9:03 PM		Every Sunday, I attend the mass	Log successfully recorded!	
178	2018/02/19 9:04 PM		alexa will play despacito	Log successfully recorded!	
179	2018/02/19 10:00 PM		when did I park my car	[at 9:02 PM, today]	4
180	2018/01/01 12:00 AM	2018/01/07 11:59 PM	I had a world tour on 1/1/2018 to 7/1/2018	Past schedule successfully recorded!	
181	2018/02/20 9:01 AM		I said "Patience is a virtue"	Log successfully recorded!	
182	2018/02/20 9:02 AM		hello world	Log successfully recorded!	
183	2018/02/20 9:03 AM		are there 7107 islands on the Philippines	[Yes, based on log created on 9:00 PM, yesterday]	4
184	2018/02/20 9:04 AM		when was my world tour	[No answer found]	2
185	2018/02/20 9:00 PM		Was I a strong leader	[No]	0
186	2018/02/20 9:01 PM		who is the president	[a strong leader of the Philippines, based on log created on 9:01 PM, yesterday]	2

187	2018/02/20 9:02 PM		who is president Makig-angay	[your name, based on log created on Feb 6 2018 at 8:01 AM]	4
188	2018/02/21 9:00 PM		when had I been on a world tour	[from Jan 1 2018 12:00 AM to Jan 7 2018 11:59 PM]	4
189	2018/02/21 9:01 PM		am I happy	[No]	0
190	2018/02/21 9:02 PM		when did alexa play despacito	[on Feb 19 2018 at 9:04 PM]	4
191	2018/01/31 12:00 AM	2018/01/31 11:59 PM	what are my logs for Jan. the 31st	[You recorded "I put the keys above the fridge." created on Jan 31 2018 at 1:00 PM] [You recorded "I place the keys on the cabinet" created on Jan 31 2018 at 9:10 AM] [You recorded "my mother is a youtuber" created on Jan 31 2018 at 9:04 AM] [You recorded "Mr. Andrew Bacud is the name of my father." created on Jan 31 2018 at 9:01 AM] [You recorded "my mother is Chery Lim" created on Jan 31 2018 at 9:00 AM]	4

192	2018/02/05 12:00 AM	2018/02/07 11:59 PM	what are my schedules from February 5 to Feb. 7	[You scheduled "I watched Spiderman movie" on Feb 5 2018 from 6:00 PM to 11:59 PM] [You scheduled "I have an appointment with the doctor at chong hua hospital" on Feb 5 2018 from 2:30 PM to 3:30 PM] [You scheduled "meeting with Sir leeroy" on Feb 5 2018 from 12:00 AM to 11:59 PM]	4
193	2018/02/22 9:00 AM		what quote did Albert Einstein say	["I fear the day that technology will surpass our human interaction . The world will have a generation of idiots .", based on log created on Feb 16 2018 at 1:00 PM]	4
194	2018/02/22 9:01 AM		who said "Patience is a virtue"	[you, based on log created on Feb 20 2018 at 9:01 AM]	4
195	2018/02/22 9:02 AM		what is life	[a journey with full of trials, based on log created on Feb 15 2018 at 9:00 AM]	4
196	2018/02/22 9:03 AM		what is the reason of life	[No answer found]	0

197	2018/02/22 9:04 AM		what time do I attend the mass	[Every Sunday, based on log created on Feb 19 2018 at 9:03 PM]	4
198	2018/02/23 9:00 PM		what did I quote	["Patience is a virtue", based on log created on Feb 20 2018 at 9:01 AM]	4
199	2018/02/23 9:01 PM		what did I last say	[You recorded "hello world" on Feb 20 2018 at 9:02 AM]	4
200	2018/03/01 12:00 AM	2018/03/01 11:59 PM	will I graduate on March	[Yes, based on schedule on Mar 1 2018 from 12:00 AM to 11:59 PM]	4

GRAMMARLY AND PLAGIARISM TEST CERTIFICATE

UNIVERSITY OF SAN JOSE-RECOLETOS
Center for Policy, Research and Development Studies

CERTIFICATION

The study entitled **"RECALL: A SCHEDULING SYSTEM AND QUESTION ANSWERING SYSTEM WITH USER KNOWLEDGE BASE USING KEYWORD, SYNONYM, AND RULE-BASED APPROACHES"** has undergone Similarity and Grammarly tests under Turnitin and Grammarly softwares.

AUTHOR/s: **Mr. Kenneth Sanchez**

TURNITIN RESULT: 9%

GRAMMARLY RESULT: 87%

This is to certify further that the manuscript has registered an originality grade of 91% and technical writing quality of 87% which includes grammar, spelling, and punctuations, among others. Given this 21st day of February, 2019 at the Quality Assurance Unit of the Center for Policy, Research and Development Studies, University of San Jose-Recoletos, Cebu City.

Mrs. Milagros B. Baclayon
Quality Assurance Specialist