Final Report: CMPE 272
Lullabyte: An IoT Android Application

By
Aditi Bhatia, <aditi.bhatia@sjsu.edu>
Akinfemi Aluko, <akinfemi.aluko@gmail.com>
Dipro Chowdhury ,<dipro.chowdhury@sjsu.edu>
Bou-Yu Chen, <bou-yu.chen@sjsu.edu>

Table of Contents

List of Figures/Tables

Lullabyte: An IoT Android Application

Aditi Bhatia[1], Akinfemi Aluko[2], Dipro Chowdhury[3], Bou-Yu Chen[4]
Software Engineering, San Jose State University
[1]aditi.bhatia@sjsu.edu
[2]akinfemi.aluko@gmail.com
[3]dipro.chowdhury@sjsu.edu
[4]bou-yu.chen@sjsu.edu

I. INTRODUCTION

Time parents spend with their children has an extremely high impact toward the children's personal development. Parents understand this, and UCI research shows that today's parents spend more time with their children compared to parents did 50 years ago [1]. Through our product, we aim to improve the quality of the time that the parents spend with their children. In the current digital era where tablets and smartphones are ubiquitous, it is not enough for children to just listen to stories. Our project is to create an app that will recognize and analyze what the parents read and transform this spoken story into an animation.

II. OUR IDEA

Our seamless design allows any parent/guardian to pick up any children's book and instantly receive relevant animations, colored mood lighting, and background color which corresponds to the content and emotion of the spoken story. For our application, we used the Google Speech-to-Text API, the IBM Watson Speech Analytics API, the Philips Hue Lightbulb API, and the Giphy API for visual representations. Basically, our application incorporates Speech Recognition, Natural Language Understanding, the Internet of Things, and Visual Representations. A combination of these allows us to create our app which brings bedtime stories to life.

III. IMPLEMENTATION

A. Google Speech Recognition API

We used the Google Speech Recognition API to translate the user's speech into text. Once the "Microphone" button on the application is clicked, it will record the user's speech until a pause or break is detected. Then, it will parse the input speech and place the output result into a string array.

B. IBM Watson Natural Language Understanding

There are two functions that we use from the Watson Natural Language Understanding, "Keyword" and "Emotions".[2] The usage of the emotions function will be discussed later in the Philips Hue section. The keyword function is used to extract keywords from the sentences the recognized by the recognition engine. These keywords are later used to search for relevant animations from GIPHY.

C. GIPHY API

Ideally we would like to have our own gallery of animation, but with the amount of time and capital we are using GIPHY, an open-source gif search engine. Using the API provided by them we were able to prototype our idea with the gallery of animation that they have.

D. Philips Hue API

The philips hue bulb is controlled by a bridge and the bridge takes REST API calls that are in JSON format.[3] Using the emotion scores (0 to 1) from IBM Watson Natural Language Understanding service we were able to different emotion to different colors. The Natural Language Understanding services provides us with five different emotions and we used three of those emotions: anger, fear and joy. We convert these three emotions to the following color: Red for anger, green for joy and blue for fear. We then mix the different levels of emotion and RGB values accordingly. The RGB values are then converted into HUE value using formulas provided by Philips.

IV. OUR DESIGN

Upon application launch, a splash screen is displayed while the application finishes loading. After the user is shown the UI which displays a microphone image to depict its request for a voice input *(See Figure 1)*. Once the users clicks the "Microphone" button, voice input is detected and passed to the SpeechToText class. Then, the speech is displayed in text format on the screen *(See Figure 2)*. This SpeechToText class is implemented with the GoogleSpeechToText API, which passes the resulting text into the Watson Natural Language Understanding (NLU) class.

Figure 1, Splash Screen: The Splash screen with the "Lullabyte" product logo.

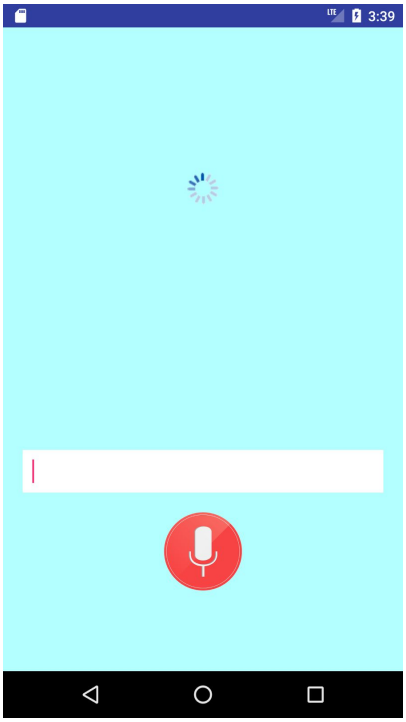

Figure 2, Startup UI: The startup UI for the user to begin speaking by clicking the "microphone" button.

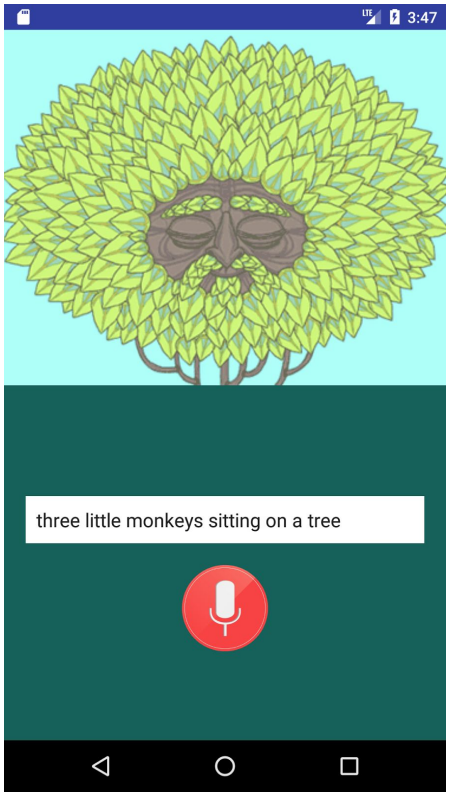

Figure 3, App in-use: An example use case: After the user speaks "Three little monkeys sitting on a tree", the application prints the text on the screen, shows a relevant animation, and changes the background color of the app to match the emotion/context of the speech.

The Watson Natural Language Understanding service, hosted by IBM on Bluemix, has a class which returns a JSON output containing the keywords and the emotion represented by the text. The keywords are extracted and passed to the Giphy API.

Meanwhile, the emotions are converted into RGB values by multiplying their values by 255, and are then used to set the background of the user interface. These values are also passed to the Philips Hue API to change the color of the Philips Hue lightbulb.

The Giphy API also returns a JSON array -- which is parsed to extract a .GIF animation, which is later passed to the Main Activity Java class. The class diagram in Figure 4 shows the workflow and connections between the various APIs we have used to build this application.

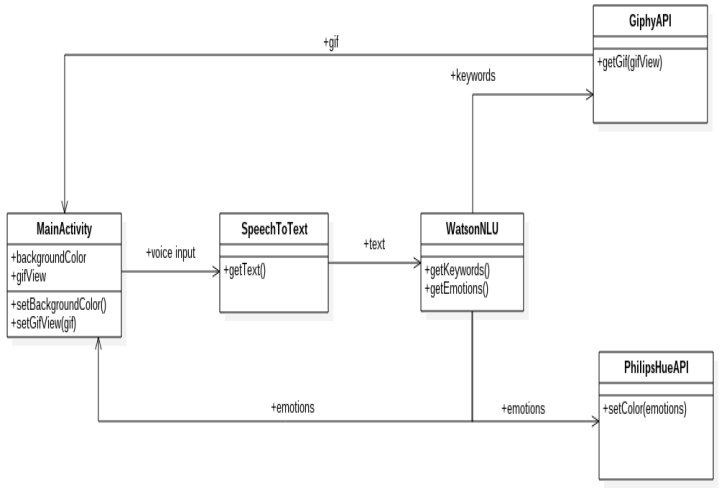

Figure 4, UML Diagram for Lullabyte: A class diagram explaining the relationships between the various APIs we have used for Lullabyte.

V. TARGET CUSTOMER

Our target customers range is broad, and includes families with young children learning how to read, as well as families with children that have learning difficulties. Our application can also be used to help children with autism and have difficulties understanding emotion, since our applications helps to interpret emotions and provides visual displays corresponding to portions of a story. In addition, children that have difficulties focusing would also benefit from our app, and would be a potential customer, since Lullabyte provides captivating animations and mood lighting to enhance story-time and reading.

VI. COMPETITION

Competition for our product is limited. One popular application currently on the market is "Disney Princess" by Disney, which does not

allow for customized story input from the user or Internet of Things support, but does allow for Speech to Text [4]. Another application is called "30 Hands" by 30 Hands learning which does not offer a mobile app, speech to text, or integrated Internet of Things support [5]. "Story Time for Kids" by Mariya Bohari is another story-time app, but it only offers pre-loaded stories and does not support speech-to-text[6]. It also does not support Internet of Things compatibility or customized stories. Overall, none of our competitors offer emotion interpretation or Internet of Things support (See Figure 5).

Features	Disney Princess	30 Hands	Story Time for Kids	LullaByte
Customized User Stories	-	+	-	+
Speech to Text	+	-	-	+
Emotion Interpretation	-	-	-	+
Animations /Images	+	+	+	+
Internet of Things Support	-	-	-	+

+ : Feature is present
- : Feature is not present

Figure 5: Competitor Analysis Chart.

VII. COST ESTIMATES

For starters, we are planning to use Firebase which is a database as a service provided by Google. We can start with the Flame Plan offered by them. Once our needs exceeds the plan we will be able to scale up and upgrade to the Blaze Plan. The Flame Plan will cost $25/month. We also hope to hire artists to provide animations for our application. One animation/gif costs about five dollars [7]. We can start with 4,000 gifs which will costs about $20,000. Assuming that we will not upscale within 3 months, we will need $20,075 in total for the application alone.

VIII. FUTURE GOALS

A. Educational Usage

To test the usability of our application, we field-tested our application with real-life users. We asked parents with young children to try our application and received valuable feedback. One of the testers suggested to give the application an education value. He suggested that we could add in features to teach children grammar and other core fundamental learning concepts. We believe that this feature would be an interesting addition to our application, and that it is also very feasible to implement this. In future developments of our app, we can develop this feature using the Natural Language Understanding Semantic Roles function from IBM Watson to recognize the subject, action, and objects in a sentence.

B. Animation Database/Expansion

For future development of this application, we hope to implement our own animation database instead of using the GIPHY API for animation results. Having our own database of images and animations will allow us to manage our server and avoid any surprises such as server down time. Most importantly, this will also mean that the application run-time will be considerably

faster since we will not have to wait for an external service provider to respond. To implement this database, we will look to hire artists to expand our animation pool, provide better services to the users, and have a quicker runtime for our app. One of our testers also suggested providing options for the user to submit their own animations/images, and by having our own database and server we will be able to achieve this function.

C. Preloaded Stories

We want users still be able to enjoy Lullabyte when there is no internet or a bad internet connection. Our solution to this is to allow for preloaded stories within the app, which will have pre-loaded text of the story, emotion analysis, and mood-lighting.

D. Expansion to Other Platforms

Currently, our application is built on the Android platform. For future development, we also plan to develop Lullabyte for iOS as well. Some might ask why are we not using a cross platform development tool such as Cordova, and this is because there will be too much overhead which will not be ideal for our app.

E. Business Model

In order to develop this prototype into a business model, we hope to add in functionality for user login, where users can subscribe to our service, save their customized stories, and upload their animations. For future development, we would like to work with someone who has relevant experience to further develop our business model.

REFERENCES

[1] University of California, Irvine. (2016, Sept 28). Today's parents spend more time with their kids than moms and dads did 50 years ago [Online]. Available: https://news.uci.edu/research/todays-parents-spend-more-time-with-their-kids-than-moms-and-dads-did-50-years-ago/

[2] IBM. Overview of the IBM Watson Natural Language Understanding service [Online]. Available: https://www.ibm.com/watson/developercloud/doc/natural-language-understanding/

[3] Koninklijke Philips Electronics N.V. Core Concepts [Online]. Available: https://developers.meethue.com/documentation/core-concepts

[4] Disney Storytime App [Online]. Available: http://lol.disney.com/games/disney-storytime-app

[5] 30Hands Storyteller Web [Online]. Available: http://30hands.com/30handsweb/

[6] Story Time for Kids [Online]. Available: https://itunes.apple.com/us/app/story-time-for-kids/id410788584?mt=8

[7] Fiverr. I WILL Do AMAZING Animated Gif Image. [Online]. Available: https://www.fiverr.com/dungdung/animate-app-icon-splash-screen-mobile-web-button?context=single_query&context_type=auto&context_referrer=search_gigs&ref_ctx_id=61ebcf56-f007-43cd-a907-9e8ac541b329&filtered_price=&pckg_id=1&pos=5&funnel=bcf9c226-1e96-40cd-8b11-fd347259581a