



DESIGN AND SIMULATION OF A COOK ASSISTANCE DIALOGUE SYSTEM USING SPEECH RECOGNITION METHOD

P.Sherubha¹, Obulisathish V²,
Sudhakar P³, Prakash T⁴,

¹Assistant Professor, ^{2,3,4}UG students,
Department of information Technology,
Karpagam College of Engineering, Coimbatore.
sherubha0106@gmail.com

July 5, 2018

Abstract

This paper presents a system that helps to cook a recipe using speech recognition. This paper projects the experience of creating cook assistance system that helps the user to cook any dish by having GOOGLE API. In order to support the creation of cooking assistant, this work mainly focused on modifying paradigm i.e., instead of the system being driven by the customers, the user is instructed by speech recognition. The outcome of the system is capable of dictating generic contents to the user via speech recognition. On the top of it, the speech recognition can be used in several places where the goal is to provide GOOGLE API while user performs some procedural performance.

Keywords: Natural Language Interfaces, Speech Recognition.

1 INTRODUCTION

Cook assistance is important to everyone ends up doing sometime in user life. Somebody love it, somebody hate it. Some people have been more experience, some are just beginners. These are no matter what level the user has, most of the time user hands are very busy. This system helps the user by dictating the needed steps while his hands and eyes are occupied with the cooking tasks are much desired.

Being so, the main target is to develop a spoken dialogue system that provides assistance in reading the procedure and detailing all steps that are unclear to the user by lack of expertise. Based on our own cooking experience, a prototype system was built that helps the user to perform some task, and experimented this with the cooking domain. Also, as not everyone needs the same kind of help while cooking, we allow the system to adapt itself to the user needs and expertise. The starting point of our work was Dialogue Interface Global Access [1], an existing framework that allows the creation of spoken dialogue systems.

The Intelligent Procedure Assistant (IPA) [2] system was designed considering that the user has his/her hands busy while executing the task being tutored. IPA is a multi-modal spoken dialogue system aimed at providing guidance, support and under the execution of large functional procedures. In the proposed work cook assistance system, the tasks are very simple, and the environment is fully different. The procedures are described hierarchically and try to reuse as much as possible the steps common across different tasks. Reading a recipe instructions tree, like description of left-to-right we have the main procedures. Going top-down has the steps needed for each procedure [2].

2 RELATED WORK

Filipe Martins et al. [2009] presented a tutor dialogue system built over the DIGA framework that is designed to engage in task-oriented dialogues in one of several task domains. Modeling the types of interaction needed as services in the dialogue system eased the creation process, since only minimal architectural changes were needed to implement this new system. A little

multimedia tool was created to allow the system to show images during the tutoring session. As for the multimedia files, in the future, the framework could be adapted to open web pages. That would allow us to take advantage of information available on the internet or previously recorded cooking sessions.

Adriaan Dekker [2014] anticipated the difference that could be made with a spoken dialogue system, is remembering the user to perform actions when a certain amount of time has passed. Another thing to be researched could be the enhancement of using images, since it was suggested many times in both groups. It is also recommended to force users to interact only via spoken dialogue system, to prevent users from measure the app by its look. This can be done by hiding all text and buttons from the basic app in the enhanced app, thus enabling to measure the raw difference between the two interaction types.

Ulrich Schafer et al. [2013] described the components and implementation of a mobile, speech enabled cooking assistant for 32,000 German recipes. The system is fully implemented and runs stably and fluently. Due to its design with only a lean, well-scaling recipe text server in addition to cloud-based Google ASR, the app can easily be installed on thousands of mobile devices. Future work would extend the approach to utilizing linguistic parsing in both query and answer candidate sentences to further abstract from linguistic variants.

Ichiro Ide et al. [2010] proposed a method that detects and supplements difficult descriptions for an inexperienced user in an existing text recipe with text, image and video contents. Experimental results showed promising ability of the proposed method to assist users understand the descriptions in a recipe.

3 PROPOSED SYSTEM

In this speech recognition dialogue system [2], a conceptual model that represents a different types of recipes. The chosen representation needs enough expressiveness to allow the selection of a recipe. When mimicking the wanted system behavior, it became clear that the main instructions the user could give to the system would be: move one step forward; move one step

backward; detail the current step; return to the previous level of detail; end the tutoring task. These instructions were encoded on primitive services: next, previous that gave the user control over the dialogue.

3.1 SPEECH RECOGNITION

This work make use of voice record module for processing the input from the user. Before the input will be processed,the input must be converted into string and it should be passed to TTS [Text To Speech] [3]. If the string that contains any abbreviation means then it should be 5 x 5 diameter input and the work attempts to improve the quality of the input by adding codes and regular expression. Then it should be helpful for storing the frequencies of inputs.

3.2 STRUCTURE OF PROJECT



Figure 1: Pictorial representation of Google Speech Recognition

Fig.1 represents the external structure of the proposed work. Android logo represents the users input. Users are give the input through mobile using voice recognition the input is converting into text format using GOOGLE IPA API. After the conversion of the input, it will check out the recipe database, the recipe database was already created by the cook/chef. If the inputs are matched, that is retrieved to the conversion API. The API was convert to the text to voice recognition and displays the user [3].

Fig.2 represents the basic flow of cook assistant system. First stage user will give the input in the form of voice recognition. Second stage is Speech Recognition system. In this stage Speech recognition system verifies the input speech command. If its varying from general commands to change the responsive text or recognized text with help of Speech Engine. Third stage is Database Record, it will check the input in the database whether if it is match or not. It will provide the result through the voice recognition API [2]. It will send entire result to the user.

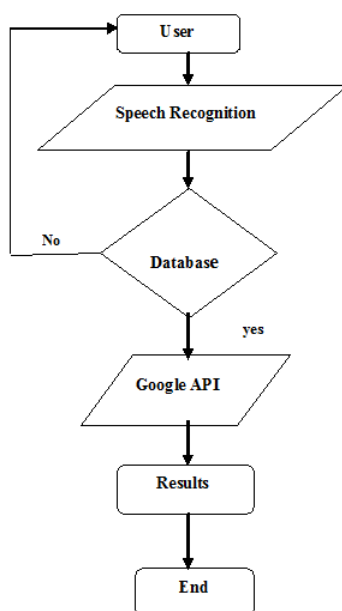


Figure 2: Flow diagram of the proposed work

3.3 STORING PROCEDURE

Table 1 display the recipe procedure and stored into database in the form of four divisions. First column denotes the recipe type and second column represents the material which contains the information to make recipe. Third column represents the

Table 1: Table for storing the recipe details

RECIPE	MATERIAL	INSTRUCTIONS	STEPS
Cake	1. Water 2. Butter 3. Sugar	Mix	3
Chicken Curry	1. Olive oil 2. Onion 3. Chicken 4. Chili(green) 5. Salt 6. Garam (masala)	Cut- vegetables And chicken Fry Add ingredients	7
Pancake	1. Bowl 2. Baking power 3. Vanilla 4. (drops) 5. Milk 6. Egg 7. Oil	Mix - egg and oil Boil (medium)	7
Pongal	1. Rice 2. Moong dal 3. Sugar 4. Ghee 5. Water 6. Cooker	Drain - rice and dal Boil Add ghee Add rest ingredients	8

instructions of making the recipe. Final column contains the no. of steps to needed to made recipe.

4 SIMULATION AND RESULTS

4.1 Markup Format

This system provides assistance that guides the user in cooking process which pre-defined the procedural list or step by step process. Several markup languages are based on XML format [2]-[4]. So, here XML markup language was utilized to store the recipe details. It contains the recipe title, ingredient lists, and cooking instructions etc.,.

4.2 Measures

There are various different approaches to measure primary ingredients, quantity/balancing amount and intervals depending on the cultural background in various geographic regions [3].

4.3 Modeling Recipes

For example, a proper model was created for a recipe where they are described as a list of ingredients, and an ordered list of steps.

Sample code for Chocolate Cake recipe in XML file format.

```
<?xml version="1.0">
<Recipe name="Cake" type="Desert">
<IngredientList>
  <Ingredient>2 Cups flour</Ingredient>
  <Ingredient>1 Cups sugar</Ingredient>
</IngredientList>
<Instruction>Sift the flour</Instruction>
</Recipe>
```

Figure 3: Sample code for recipe selection

In XML file, recipe has a set of initial steps that represents the list of ingredients needed to perform it. In Future, a multimedia

content will be added (image and playing videos for reference) for users live experience. Inside the box, sample code for retrieving the recipe data. The xml version 1 was utilized for data retrieving, and its recipe name is considered as key. User searching data will check out the keys (Recipe name).

Ingredient list are used to store the instructions as bundle. The bundles are easily through the GOOGLE API [1] and display it to user.

4.4 RESULTS AND DISCUSSION

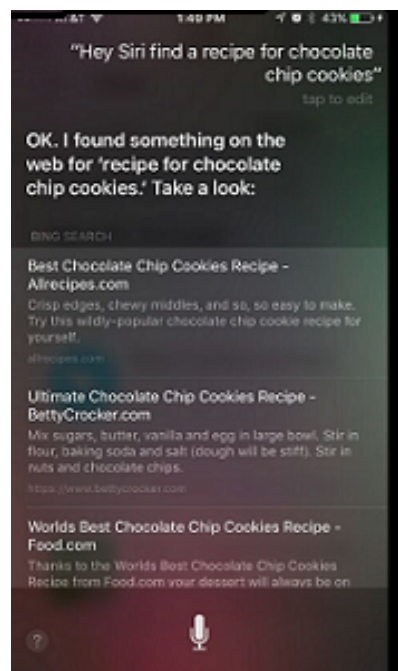


Figure 4: Recognition of input speech for execution

Fig 3 represents the input screen, this image represent a user's input recipe to search on the web. So, it will search on the web and collect the recipe's details and related recipe's. It also stored in database. Fig 4 represents the output screen, user desired recipe

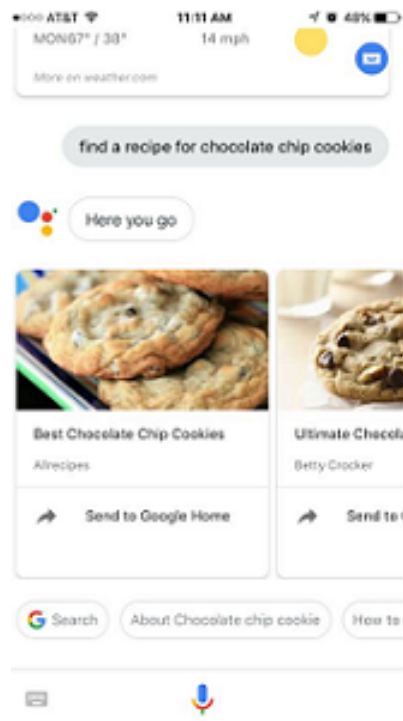


Figure 5: Representation of output screen

results show to the users. Includes the recipe's image and variety of recipe's will be shown.

5 CONCLUSION

A Speech recognition system for providing cooking assistance was designed and simulated for user desired recipes. The main part is UI design that was to make an easy-to-use application which can be controlled by using various voice/speech commands but also through standard touch interactions. In future, the work may be expanded to implement a multimedia tool that allows the system to show images during the tutoring session. That helps the user for greater experience about cooking recipe.

References

- [1] Filipe Martins, Joana Paulo Pardal, Lus Franqueira, Pedro Arez and Nuno J. Mamede "Starting to cook a tutoring dialogue system" January 2009.
- [2] Adriaan Dekker "Improvements of Adding a Spoken Dialog System to a Cooking Coach Application: a User Study".
- [3] Frederik Arnold, Simon Ostermann, and Saskia Reifers "Ingredients and Recipe for a Robust Mobile Speech-enabled Cooking Assistant for German".
- [4] Ichiro Ide, Yuka Shidochi, Yuichi Nakamura, aisuke Deguchi, Tomokazu Takahashi and iroshi Murase "Multimedia supplementation to a cooking recipe text for facilitating it understanding to inexperienced users".
- [5] P. Sherubha, "Energy Efficient and Bandwidth aggregation techniques in Wireless Heterogeneous devices: A survey" International journal of pure and applied mathematics Feb 2018.
- [6] Akey Sungheetha, Mahalakshmi. K, Sherubha. P, Narmatha K and Rajesh Sharma R "Offline Remote Path Board Structure

based on GPS/WiFi Location Access”, International journal
of pure and applied mathematics Feb 2018.

