Proposal for Prevention of Forgetting Shopping Items in Response to Shopping List Ambiguity

Yoshihisa Igimi^{a,*}, Lifeng Zhang^a

^aFaculty of Engineering, Kyushu Institute of Technology
1-1 Sensui-cho, Tobata-ku, Kitakyushu, Fukuoka, 804-8550, JAPAN

* Corresponding Author: igimi.yoshihisa269@mail.kyutech.jp

Abstract

Shopping plays a critical role in daily life, and also it is a time-consuming activity. For the coming elder society in Japan, the elders who live alone have to go shopping by themselves and often forget to buy the target goods in their shopping list. This forgetting is because the shopping list is often made with ambiguous expressions compared to the store registered product name. This research proposes a new shopping item forgetting prevention system that can detect the not buying item immediately after purchasing to dealing with the shopping list's ambiguity.

keywords: Shopping List, Ambiguous Expressions

1. Introduction

1.1 Background

There is an existing application to prevent forgetting to buy. Fig. 1 is screen of this application. This application creates a shopping list. The usage of this application is as follows. Input the name of the grocery should be purchased before shopping, and after shopping, compare the purchased product or its receipt with the pre-inputted grocery name displayed on the application screen. Therefore, the forgotten item can be checked.

This application is easy to understand and use by anyone. However, the withdrawal is that people have to check the items by themselves, and it takes time to check the product items printed on the receipt one by one. Also, user have to make a judgement because the shopping list is normally a genre name, but the product name printed on the receipt is normally a brand name for different company. The shopping list is not absolutely same to the receipt. Therefore, it will take time to compare and is not easy to check at the storefront where purchased the product.

Fig. 1. Application Screen

Also, if people comparing the receipt with the shopping list displayed on the application screen, it is easy to understand if the product name that appeared on the receipt belongs to genre name the prepared shopping list. However, if a computer does this task, it would not be as easy as human beings to match the equivalent of names from the shopping list and the receipt. This is because the food name input into the application by human beings is more ambiguous than the product name written on the receipt from the computer's point of view.

1.2 Purpose

This research proposes a new forget-to-buy system that can check the forgetting items by comparing the shopping list item and the receipt printed by the register. Although the ambiguity between the shopping list and the receipt is a

non-good-at task for a computer, the proposed approach can solve this problem in a new way.

2. Proposed Method

First of all, to make the expressions clear, the names on the user made shopping list are called genre name hereafter, and the names printed on the receipt called product name hereafter. This is because the genre name and the product name might different in expression and volumes. For example, user may write “eggs, one pack” on the shopping list, but buy the product named “red ball, half dozen” on the receipt.

To tell the user the forgotten genre name that does not buy from the shopping list, it must first identify the genre name purchased from the product name on the receipt. The genre names that are forgotten to buy remain after matching the purchased product belongs to the specified genre name written in the shopping list. The genre names that do not match any product names on the receipt are the forgotten items.

In this research, to realize the part of "identifying what kind of genre the purchased product belongs to" mentioned above, a method using web scraping and a method using API is proposed. Besides, both the scraping and the API for target searching in this study depend on Rakuten Ichiba. Rakuten Ichiba is an online shopping site operated by Rakuten, Inc. Rakuten Ichiba does not prohibit web scraping in its terms of use and provides an API and handles many food information, so Rakuten Ichiba web service was used in this study.

2.1 Proposed Method with API

2.1.1 Process Flow

Fig. 2 shows the process flow of the proposed system with API.

This proposal system using API is divided into two parts. One processes before shopping, and the other one processes after the payment is finished, but the purchaser is still at the storefront, and the processing operation is divided into 6 stages. Also, the programming language used in this proposed system is Python.

2.1.2 Get Genre ID

In Rakuten Ichiba, a genre ID is assigned to each product. The proposed method uses the "Rakuten product search API" to acquire the genre ID related to the entered product name. "Rakuten product search API" allows you to enter a keyword

Fig. 2. Proposal System Flowchart

and obtain information on products in the Rakuten market related to that keyword.

Explain the specific method of obtaining the genre ID. The request URL is required to use the API. The structure of the request URL for the Rakuten product search API is as follows.

```

https://app.rakuten.co.jp/services/api/
IchibaItem/Search/\20170706?format=
json&keyword=(ProductName&hits=
10&applicationId=(MyApplicationID)

```

"My Application ID" is given to the user by Rakuten when they agree to the terms of use of Rakuten Ichiba's API. Also, Since the request cannot be accepted if the "product name" is entered in Japanese, the product name is encoded using the Japanese encoding function (urllib.parse.quote) and then passed to the URL. The data of the request URL destination is acquired by using the function (requests.get) that acquires the data of the passed URL. The acquired data is hierarchical JSON data. At the top of the hierarchy is the "Items" key, the "Item" key in it, and the "genreId" key in it. The element of this "genreId" key is the genre ID of the genre related to the input product name, so this system gets the elements of the "genreId" key one by one.

2.1.3 Get Food Name

The proposed method uses the "Rakuten genre search API" to acquire the food name. "Rakuten genre search API" allows you to enter a genre ID and obtain a product name related to the genre ID.

Explain the specific method of obtaining the food name. As in Section 2.1.2, specify the request URL to use the API.

The structure of the request URL for the Rakuten genre search API is as follows.

```
https://app.rakuten.co.jp/services/api/
IchibaGenre/Search/\20140222?format=
json&genreId=(GenreID)&applicationId=
(MyApplicationID)
```

Pass the genre ID obtained in Section 2.1.2 to the "Genre ID" part. And, the data of the request URL destination is acquired by using the function (requests.get) that acquires the data of the passed URL. The acquired data is hierarchical JSON data. There are three types of keys at the top of the hierarchy: "parents", "current", and "children". The "parents" key contains the "parent" key, and the "children" key contains the "child" key. In addition, there is a "genreName" key in each of the "parent", "current", and "child" keys. The element of this "genreName" key is the food name of the genre ID passed, so this system gets the elements of the "genreName" key in the "parent" key and the "current" key one by one.

2.1.4 Matching

Match the genre name that the user plan to purchase first with the product name of the purchase obtained in Section 2.1.3 one by one.

As a specific matching method, check whether the character string of the food name entered first is included in each character string of the food name element acquired in Section 2.1.3.

3. Principle

Web API

API(Application Programming Interface) is an interface that provides functions for using service data from external applications and programs. Especially, API that exchange data via HTTP / HTTPS communication are called Web API. In the Web API, part of the software is published on the Web, and the API user can use the service of the Web API provider by calling the Web API from the outside.

4. Experiment

4.1 Experimental Method

In this experiment, the correct answer rate was calculated as to whether the proposed system correctly outputs "Purchased" when the product is purchased and "Forgot to buy" when the

Table 1. Experimental Result

Collaborator	Output Result		Correct Answer Rate[%]	Time [seconds]
	True	False		
1	7	3	70	50
2	8	2	80	52
3	6	4	60	30
4	7	3	70	33
5	7	3	70	28
6	7	3	70	45
average	7.0	3.0	70.0	39.7

product is forgotten to buy. Also, the processing time of the system was measured.

The specific experimental is performed as:

1. First, we asked six collaborators to create a shopping list with 10 different food names per person.
2. Second, I entered the food names on the shopping list into the proposed system.
3. Third, from the receipts collected from various stores, the product name corresponding to the food name listed in the shopping list are entered into the proposed system for matching and calculated the correct answer rate of the system from the output result. When entering this product name, the data of 6 collaborators were divided into 2 groups of 3 each.

And, the experiment was conducted assuming that one group bought all the products without forgetting to buy them, and the other group forgot to buy one product.

The processing time of the system is the time from the end of input of the food name and product name to the output result, and this is measured programmatically using the time measurement module (time).

4.2 Experimental Result

The results of the experiments for the 1st to 6th people are summarized in the table below. "True" of "Output Result" is the number of times the system outputs correctly, and "false" is the number of times the system outputs incorrectly.

The average is rounded to the first decimal place.

5. Consideration

Factors that Lowered The Correct Answer Rate

As shown in Table 1, the correct answer rate was 70.0%. The reason why the correct answer rate was not 100% is

considered to be the following factors.

- Lack of desired food name
- Conversion of kanji
- Lack of product corresponding to product name

Lack of desired food name

In the proposal system, the name of the food to be purchased is matched with the name of the food obtained from Rakuten Ichiba, so if the Rakuten Ichiba does not have the desired food name, the matching will not be successful. Therefore, even if you have purchased a product of the desired food name, the system will determine that you have forgotten to buy it.

Conversion of kanji

When matching the input product name with the product name acquired from Rakuten Ichiba, if the input genre name is in kanji notation and the acquired genre name is in hiragana or katakana notation, the matching will not succeed. To prevent this, if the input food name is written in Kanji, it is converted to Hiragana and then converted to Katakana. After that, each of the three types of food names, Kanji, Hiragana, and Katakana, is matched with the acquired food name.

However, there are two ways to read Kanji: on-yomi and kun-yomi. Therefore, if the entered food name is partially written in Kanji, the function that converts Kanji to Hiragana may make a mistake in selecting on-yomi and kun-yomi, resulting in matching failure.

Lack of product corresponding to product name

If there is no product that corresponds to the product name entered in Rakuten Ichiba, the food name matching will not be successful because there is no genre name to be acquired.

The remaining issues are the troublesomeness of the input section and the low rate of correct answers. Since the input of the developed system is troublesome by manual input, it is thought that the troublesomeness can be improved by incorporating voice recognition and optical character recognition into the input part. Also, regarding the correct answer rate of the system, it is considered that the failure of matching due to the conversion of kanji is one of the factors that lower the correct answer rate. About this issue, by using the API of the Information-technology Promotion Agency's character information infrastructure database, it is possible to obtain the results of on-yomi and kun-yomi readings of kanji, so it is expected that failures due to kanji conversion will be improved.

References

- (1) Shopping List, kiwi3.com, <https://www.kiwi3.com/>, Access date: 2020.11.26
- (2) How to use Rakuten Product Search API, HP-code, <https://haniwaman.com/rakuten-api/>, Access date: 2020.02.11
- (3) Try using Rakuten API with Python's Requests module, BTY Memorandum, <https://bty.sakura.ne.jp/wp/archives/1157>, Access date: 2020.02.11
- (4) What is an API? Explain difference from Web API, user tasks, TechTargetJapan, <https://techtarget.itmedia.co.jp/tt/news/1806/13/news01.html>, Access date: 2020.02.14
- (5) 9th High-speed Web API implementation and testing, Gijutsu-Hyoronsha, <http://gihyo.jp/dev/serial/01/perl-hackers-hub/000901>, Access date: 2020.02.04
- (6) MJ character information retrieval system, Information-technology Promotion Agency, <https://mojikiban.ipa.go.jp/search/help/api>, Access date: 2020.02.18

6. Conclusions

In this research, a shopping forgotten prevention system with API was developed. It can tell the genre name of a product that has been forgotten to buy while dealing with the ambiguity of the shopping list immediately after purchasing the product.

The experimental result showed that the average correct answer rate of the system was 70.0%, and the average processing time was 39.7 seconds. The processing time is fast, so it seems that the research purpose was almost achieved.