Intelligent (Task-Oriented) Conversation Assistant for Course Selection

Progress Report



Information Technology Capstone Project

COMP5703/5707/5708

Group Members

1. Shengyuan Sun (460257820)
2. Rui Chen (470111585)
3. Quan Chen (470199228)
4. Biying Wang (470067400)

Table of Contents

[Table of Contents i](#_Toc521413584)

[1. Progress Status 1](#_Toc521413585)

[2. Roles & Responsibilities 2](#_Toc521413586)

[3. Individual Achievements 2](#_Toc521413587)

[4. Group Collaboration 2](#_Toc521413588)

# Progress Status

|  |  |
| --- | --- |
| **Project Name** | CS17 Intelligence(Task-Oriented) Conversation assistant for course selection |
| **Project Start Date** | 6/ 3/ 2019 |
| **Project Manager** | Biying Wang |

|  |  |
| --- | --- |
| Project Description | The project focuses on using NLP and other technologies to build a dialog system to analyse and answer the questions that students ask in terms of course units on University of Sydney education system |

|  |  |  |
| --- | --- | --- |
| Project Status Report | # | Date: 3/ 5/ 2019 |

|  |  |  |
| --- | --- | --- |
| **Status Item** | **Status up to last week** | **Planned for this week** |
| **Major deliverables** | Text preprocessing  Word embedding | Modify the natural language understand model.  Finfish the natural language generation part. |
| **Planned delivery date** | 3/5/2019 | 3/10/2019 |
| **Major issues** | As the intent classification need the input with simple format, so I apply three steps to change the orginal data type. | The model require a saved embedding and intent2index and slot2index. So it require we seve the index as a extral file.  The NLG part require more template support the generate process. |
| **Major risks** | 1, we discuss the necessary of remove stoping word.  2,the word embedding extend the size of input format, which require more trainng data.  3,the current model testing by batch input, I need modify it and witre a function for following part. | 1. The word embedding require a file to store. |
| **External dependencies** | Slack | slack |
| **Estimated effort (hr)** | 20hr | 20hr |
| **Recorded effort (hr)** | 20hr | 20hr |
| **Status (R, Y, G)** | green | green |

# Roles & Responsibilities

Shengyuan Sun

Role: developer

Responsibility:

* Text preprocessing
* modify the intent classification part

Rui Chen

Roles: Analyser and Developer

Responsibility:

* Handbook and CourseMap templates
* Collect data from CUSP website to fill the blank in templates
* Generate more that thirty thousands raw questions with intends and slots
* Dialog Tracker development

Biying Wang

Role: Project Manager; Developer

Responsibility:

* Check weekly deliverables
* Check meeting time, location and meeting topic
* Discuss intent classification and slot filling with team members
* Design Entity list
* Integrate template

Quan Chen

Role: Developer

Responsibility:

* Back-end database system building and data collecting
* Help to design the whole process of system (such as slot and intent)

Coding for logical part dealing with slot filling and Peewee template code matching

# Individual Achievements

NLG part is design as a template-base generater.

First part is to plan the short sentence template. A example of the design the template as:



In the figure, the template been store as a dictionary, the key is the attribute which is the query result. And the vaule of the dictionary is the answer template. Here in the template we user ‘#’ to mark the position where should been replace. To match the template, user should the all attribute shown in the tempate.

The logic inside the template combination is how to replace the mark, once we get a template we first check if there are a template can replace the ‘#’ mark in the template list, then we replace the ‘#’part by vaule we get from the query. Sencond, if therer are two sentences, then we plan the rule to conbine the two sentence, in most simple idea, we use the ‘and’ and replace the same part in two sentence.

About some specific word such as ‘yes’ ‘No’ ‘people’, we need design the rules to replace the mark position in sentence.

# Group Collaboration

