Description, Ethics, Objectives and Resources

Description

User-Centric Natural Language Generation

User-Centric Natural Language Generation (UC-NLG) will be a means of allowing Chatbots and Virtual Assistants(VAs), such as Google Assistant, Alexa, Siri, and Bixby, to respond to queries presented to them with a perceived charisma surrounding them; one which is relatable to the user.

UC-NLG will be capable of conveying the information relevant to a query, however, it will ensure that the response structure and vocabulary it uses to do so will be tailored specifically for the user. Additionally, the possibility exists that it will even become dependent on what the system has learnt about the user (in terms of their dialect) and the situation regarding the supplied query.

Context and background

Current Chatbot and VA implementations diligently follow a defined structure when sending the result of a query to the user, typically utilising string interpolation. This method of responding to clients leads to VAs becoming predictable in their responses, given their rigid response format dictated by a single team of programmers. UC-NLG ensures that, whenever a user receives a response, it is one which they find relatable; as if it was created exclusively for them.

UC-NLG will ultimately have the ability to transform a consumer's favourite VA into something much more than a mere "robot" which is used intermittently. Instead, Siri becomes "one of us", after all, UC-NLG could be used to enable Siri to communicate in a manner which consumers are much more accustomed to. By tailoring responses to each user, hopefully, VAs - and other voice interfaces - could become a much more viable and natural means of interacting with a computer. Almost as natural as interacting with a human.

Development

In order for UC-NLG to suitably form its own sentences, I propose the use of a Sequence-to-Sequence model, commonly seen in Neural Machine Translation. This model would initially be non-deterministic as one query/intent (in the Natural Language Processing sense) would lead to multiple possible responses. However, once state (composing of a suitable number of factors to personalise users) is considered, the model will become fully deterministic

The proposed Sequence-to-Sequence model would be trained on various databases in order to learn how to respond to questions in general, as well as build a mapping between questions and answers. Using this mapping from prior exposure to varying questions, the system will, ideally, be able to computationally associate the user's speech patterns with one which it had previously encountered and respond as appropriate

Objectives

Primary

By the end of the project, I will have primarily produced:

- A Sequence-to-Sequence RNN (or the like) to allow for the distinguishing users apart, for the purposes of appropriating responses
- A program capable of providing responses to a restricted domain of queries, which are influenced by factors surrounding individual users
- A means of processing speech to use with said program, utilising NLP to identify user intents
- A means of outputting information to the user, such as through the use of a Speech Synthesis program/API

Secondary

Time permitting, additional aims are possible in the event that the primary objectives are adequately satisfied:

- Manipulate a Speech Synthesiser to emulate a user's dialect
- Integrate a VA with UC-NLG
  - i.e. interact with an Alexa or Google Assistant API, for example, to create an app which can be called using the relevant VA, but uses a response generated through UC-NLG
- Expand the domain of queries for UC-NLG
- Allow UC-NLG to be contextual
  - Ensures that users may ask follow-up queries, which are resolved in the context of previous queries

Ethics

While various datasets may be used in this project, none possess any private or sensitive user information which may require prior approval. Additionally, whilst this project is user-centric, the project itself does not pose any potential to cause distress towards users. The potential to learn from users should not also be considered a breach of ethics as this learning would be based entirely on the user's interaction with the project, as opposed to any additional sources which may be seen as invasive.

Resources

I anticipate that this project should be reasonable in terms of resources required. The most demanding area of the project concerns the use of a Sequence-to-Sequence model. As a result compute servers with appropriate processing power should suffice. Machines more powerful than those running Scientific Linux in the Jack Cole labs may be needed. However, this is rather cautionary, and solely influenced by the amount of time which could be spent training the model, especially when the amount of data that may be utilised is taken into account.