DREAM Final Report

Introduction

English is currently the most spoken language worldwide as well as the most learned language worldwide. According to the CIA World Factbook 16.5% of people in the world speak English despite only 5.1% speaking it as their first language. Given this, there is high demand for English learning platforms, with applications such as Duolingo and gaining widespread popularity. These types of applications focus on correcting users' grammar and vocabulary but tend to ignore teaching them more nuanced conversational skills which are key to sounding natural when speaking in a foreign language. Language learners instead tend to learn their conversational skills by speaking with other learners or with native speakers, often in a classroom setting. However, these kinds of resources are not available to many learners and therefore many of them lack the conversational skills necessary to engage in day to day English conversation. To this end we introduce EduBot, an English learning application that uses AI-powered dialogue systems to teach learners English conversational skills without the need to speak with a native speaker or language instructor.

EduBot combines state of the art dialogue systems with templates developed based on English textbook curriculum to give learners the opportunity to practice their conversational skills on a variety of topics. Users can chat with a purely template-based bot which can help them prepare for more rigid conversational settings such as the IELTS speaking test. A purely model-based chatbot is also provided so that users can speak with the bot about any topic in a free-flowing conversation. Lasty, the application includes chatbots which combine both template-based and model-based dialogue systems to allow users to converse about a specific topic while still enjoying the more adaptive and flexible nature of a neural language model. We hope that this application will provide learners with a natural environment to practice their conversational skills without the need to engage with a native speaker or language instructor.

Related Work

Previous research has been directed towards designing chatbots to teach students English. Researchers have had success both by using existing chatbot systems such as Cleverbot (Fryer et al. 2020, Kim et al. 2019) or by designing their own systems (Ayedoun et al. 2015). The chatbots built on existing systems tend to be focused more on open conversation. Users can chat with these bots about any topic and can guide the conversation in whichever direction they choose. The self-designed systems tend to be more focused with the chatbot guiding learners to more narrow topics such as thesis writing (Lin and Chang 2020).

While these chatbots have seen some success, they are often handicapped by technological limitations. Communication issues between the bot and learners can occur in cases where students respond with incomplete sentences or when the bot gives a nonsensical output

(Huang et al. 2022). The former issue is particularly important as natural conversation between native English speakers can often involve exchanges of incomplete sentences.

Most of the existing chatbot systems do not use state of the art dialog models which may be contributing to many of the technical problems currently being faced. Large language models have been evolving rapidly in recent years and are able to perform well on a variety of tasks. The EduBot application leverages these recent advancements to create a chatbot that is able to resolve some of the technological limitations faced by previous attempts to create English learning chatbots.

Approach

The EduBot application includes three different classes of chatbots: 1) Purely template-based chatbots, 2) Purely model-based chatbots, and 3) Hybrid chatbots that incorporate both models and templates. Each of these three classes are designed to help users with different aspects of their learning. The details behind the implementation of these three types of chatbots are discussed in the following sections.

Template-Based Chatbots

EduBot's template-based chatbots are designed around the curriculum in *Conversations in Class, 3rd Edition*. There are a total of eight templates, each of which is based around the model sentences provided in each textbook unit. The chatbot interacts with these users by uttering these template phrases along with acknowledgments that are based on the natural language features of user responses. The natural language features are extracted from user responses using a combination of different apis and regular expressions. Any questions asked by the user are handled by a model-based dialogue system and the response is inserted as an acknowledgement before the next template phrase is uttered. The conversations provided by these chatbots are more rigid and suited more towards settings such as speaking tests where the conversations are heavily directed and more emphasis is placed on grammar and vocabulary as opposed to more subtle conversational skills.

Model-Based Chatbots

The EduBot application also includes a model-based chatbot which is built as a modified version of Meta's BlenderBot 3 (Roller et al. 2020). This bot is designed to speak to users about any topics they like for however long they like. Users can guide the bot to any topics they like and practice their general conversational skills. The user studies that we conducted show that BlenderBot 3 can encounter issues due it repeating itself as well as contradicting itself. To resolve the repetition issues we added n-gram blocking to the BlenderBot 3 decoder, specifically tri-gram blocking. This method seemed to relieve the repetition to some degree although more user testing is required to further test its effectiveness. Contradiction is a more difficult issue to resolve and cannot be rectified by simple decoder modifications. Instead our plan is to train a

text classifier and use it to rank BlenderBot 3's response candidates. Using this we can select the least contradictory response and avoid contradiction issues. As of writing this report this solution has not been extensively tested and will require further analysis to determine its effectiveness.

Hybrid Chatbots

The final class of chatbots in the EduBot application are hybrid chatbots that mix both template phrases and model based responses. These bots are designed so that users can have natural, varied conversations on specific topics within the textbook curriculum. The first issue that arises when creating these bots is deciding whether or not the templates or the model should drive the conversation. Using templates as the conversation driver allows for the bot to easily stay on topic and follow the textbook curriculum more strictly. However it has the drawback of making the overall conversation more rigid and less natural since it is often difficult to smoothly insert template phrases given only the basic natural language features that can be extracted from a user's response. Letting the model be the driver has the opposite problem since it can provide a more natural and adaptable conversational experience but is more likely to shift the conversation away from the textbook curriculum. Based on the user tests we performed, learners seemed to strongly prefer the model driven chatbots however did point out some issues with the model switching topics. In order to alleviate this we took a similar approach to resolving contradictions in the purely model based chatbots. Using a semantic search we were able to create a classifier to detect when the bot was giving an off topic response. This is then used to rank the bot's response candidates based on their topic relevance and output the most relevant one.

Ultimately the most successful of the hybrid bots was a chatbot that was "seeded" with a template phrase followed by BlenderBot 3 continuing the conversation. The template phrase is written to BlenderBot 3's memory so that it can continue the conversation with the proper context. This bot was shown to be preferred by users in our tests and our empirical testing of the bot showed that it is able to stay on topic fairly well. To test the bot's ability to stay on topic, we conducted 30 different conversations with the chatbot each of which was seeded with a different phrase. Then the first three turns of the conversation were annotated on how often the bot was able to remain on topic, Figure 1 shows the full results of this analysis. Overall the bot was able to give an on topic response 80.0% of the time on its first response. This dropped to 70.0% on the second turn and finally 63.3% on the third turn. These results were obtained before the topic detection classifier was implemented and we are hoping to improve these rates after incorporating and testing the classifier further.

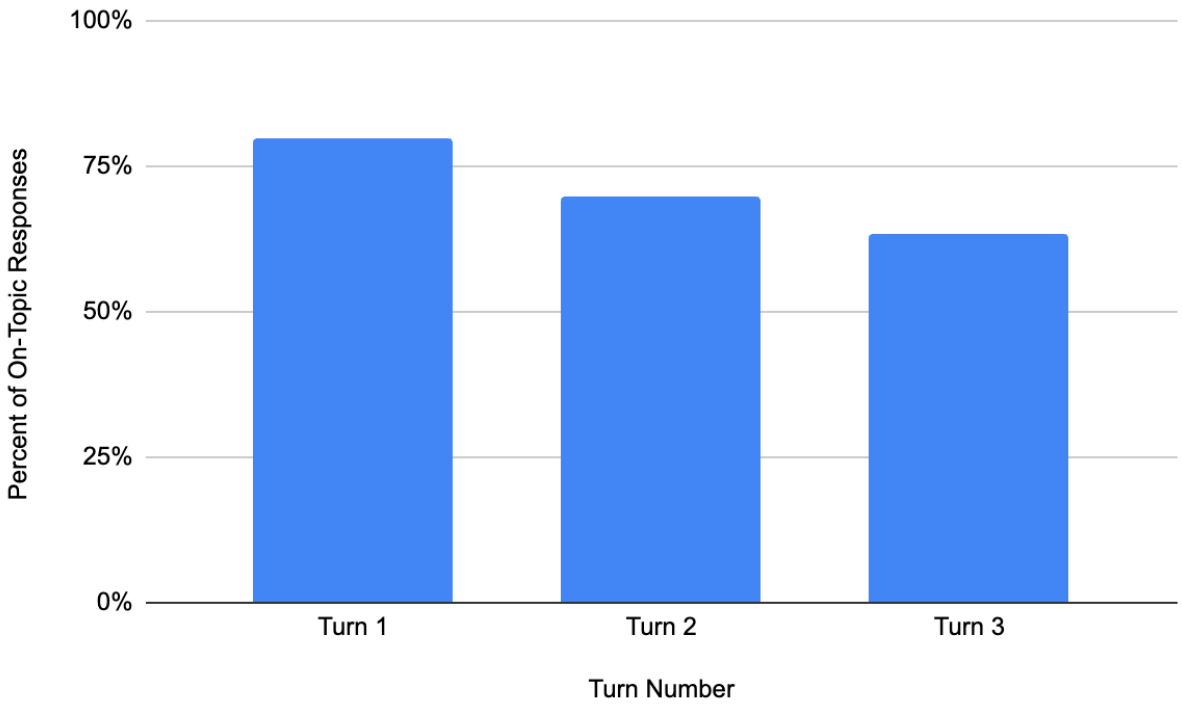

Figure 1

Future Work

Despite the steps outlined in this report, EduBot is still a work in progress. The classifier methods to detect topic switching and contradiction have yet to be fully tested and may not be accurate methods of detection. Beyond that, the classifiers can only use the response candidates output by the bot. In cases where all the bots response candidates are contradictory or off topic, the classifiers will not be able to make the model provide an on topic response. Other methods that may be worth exploring to fix these issues are regenerating model responses in cases where all candidates are bad or using constrained decoding methods like neurologic decoding (Lu et al. 2020). These methods and other potential solutions will continue to be explored as EduBot evolves as an application.

Works Cited

Ayedoun, Emmanuel, et al. “A Conversational Agent to Encourage Willingness to Communicate in the Context of English as a Foreign Language.” *Procedia Computer Science*, vol. 60, no. 1, 2015, pp. 1433-1442.

Fryer, Luke K., et al. “Bots for language learning now: Current and future directions.” *Language Learning & Technology*, vol. 24, no. 2, 2020, pp. 8-22.

Huang, Weijiao, et al. “Chatbots for language learning-Are they really useful? A systematic review of chatbot-supported language learning.” *Journal of Computer Assisted Learning*, 2022.

Kim, Na-Young, et al. “Future English Learning: Chatbots and Artificial Intelligence.” *Multimedia-Assisted Language Learning*, vol. 22, no. 3, 2019, pp. 32-53.

Lin, Michael Pin-Chuan, and Daniel Chang. “Enhancing Post-secondary Writers’ Writing Skills with a Chatbot: A Mixed-Method Classroom Study.” *Educational Technology & Society*, vol. 23, no. 1, 2020, pp. 78-92.

Lu, Ximing, et al. “NeuroLogic Decoding: (Un)supervised Neural Text Generation with Predicate Logic Constraints.” *arXiv:2010.12884*, 2020.

Roller, Stephen, et al. “Recipes for building an open-domain chatbot.” *arXiv preprint arXiv:2004.13637*, 2020.