Chefbot: A Novel Framework for the Generation of Commonsense-enhanced Responses for Task-based Dialogue Systems

Carl Strathearn and Dimitra Gkatzia
Edinburgh Napier University
{c.strathearn,d.gkatzia}@napier.ac.uk

Abstract

Conversational systems aim to generate responses that are accurate, relevant and engaging, either through utilising neural end-to-end models or through slot filling. Human-to-human conversations are enhanced by not only the latest utterance of the interlocutor, but also by recalling and referring to relevant information about concepts/objects covered in the conversation so far. Such information may contain recent referred concepts, commonsense knowledge and more. A concrete scenario of such dialogues is the cooking scenario, i.e. when an artificial agent (personal assistant, robot, chatbot) and a human converse about a recipe. We will demo a novel system for commonsense enhanced response generation in the scenario of cooking, where the conversational system is able to not only provide directions for cooking step-by-step, but also display *commonsense* capabilities such as offering explanations on object use and recommending replacements of ingredients.

1 Introduction

Although conversational User Interfaces (CUIs) have gained popularity with the introduction of commercial personal assistants, these CUIs are mostly retrieval-based question answering (QA) systems that are incapable of holding a multi-turn conversation or providing follow-up information on the same topic or task. In addition, they do not incorporate commonsense capabilities, i.e. the ability to understand how an object is used, understand/infer other non-obvious properties such as its weight and materiality, or generally make inferences about ordinary tasks in our daily lives (Davis and Marcus, 2015)). Significantly, these CUIs cannot provide "how-to" instructions when performing practical tasks that require conversation over multiple steps, such as cooking a recipe or building
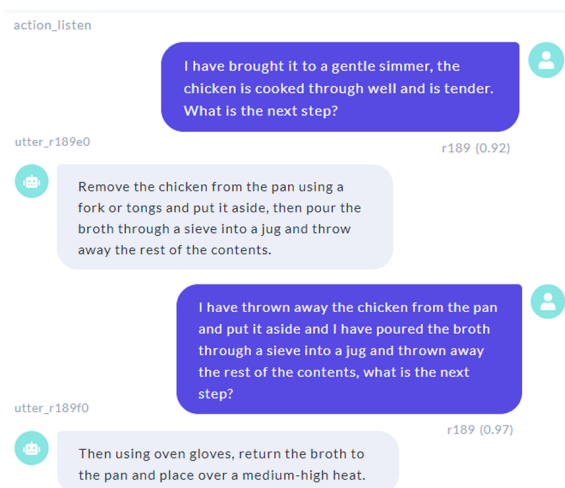

Figure 1: Excerpt from dialogue.

furniture. This work is inspired by the human ability to read a document with instructions on how to perform a task and recall and rephrase the instructions to someone else. We consider a practical task where two people cook together, where one user, the Information Giver (IG) has access to the recipe and provides instructions to the second user, the Information Follower (IF), as shown in Figure 1.

In our setup, the IG is required to understand the recipe, split it into manageable steps, possibly rephrase it as well as being able to describe objects, their use and common storage locations. The IF can ask questions such as repeating an instruction, clarification, and confirmation. Replicating these innate cognition processes in artificial agents that function in real-world conditions is highly challenging due to the complexity and interdisciplinary nature of the problem. Unlocking these challenges will enable artificial agents to operate with greater levels of common-sense reasoning. This demo paper presents a novel task-based CUI which combines knowledge-grounded dialogue and commonsense-enhanced response generation.

46

Figure 2: Examples of commonsense capabilities: On the right, Chefbot is suggesting an ingredient substitution and on the left, Chefbot explains how an object is used.

2 Task Description

Task-oriented dialogue is concerned with helping users achieve specific goals, by understanding user intends, state tracking and generating responses based on next actions (Hosseini-Asl et al., 2020). On the other hand, open-domain dialogue systems aim to converse over different topics within the same dialogue. Our proposed task is situated between these two tasks: firstly, the goal of the system is to help the user prepare a recipe by providing instructions; secondly, the system aims to converse about related concepts, such as ingredients and objects' utility in an open domain fashion (Fig. 2).

3 Chefbot

To demonstrate the use of commonsense enhanced dialogue in a practical task-based challenge, Chefbot was designed using RASA X[1]. Annotated sample conversations between the IG and IF are modelled as two modes of question and answer pairs. The first series of utterances are open-domain and the second set are domain specific. In the dialog flow, forms were used for each recipe to force the sequence between the two series allowing for both domain and non-domain utterance classification. This produced a more robust structure and contextual awareness for state tracking and response generation. The Chefbot is able to handle questions that are not represented in the sample dataset with the help of two commonsense databases. The first database provides the user with appropriate alternative ingredients for a specific recipe and the second explains the use, handling, alternative names and

typical storage locations of kitchen utensils. A combination of rules, checkpoints and custom actions, allow the user to ask questions at any stage in the task and then on fulfilment, return the the next logical step in the recipe. From this framework we create a multi-intent / multi-turn policy model that permits adaptability to cope with the variable conditions of real-world tasks.

4 Future Work

In future, we aim to extend our system so it can be used in a situated Human-Robot Interaction scenario, where the conversation will take place as a spoken conversational interaction.

5 Conclusions

This demo paper describes a commonsense-enhanced chatbot for task-based dialogue. At INLG, we will demo the chatbot and will discuss initial findings.

Acknowledgements

References

Ernest Davis and Gary Marcus. 2015. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Commun. ACM*, 58(9):92–103.

Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, and Richard Socher. 2020. A simple language model for task-oriented dialogue. In *Advances in Neural Information Processing Systems*, volume 33, pages 20179–20191.

---

[1] https://rasa.com/docs/rasa-x/