Interactive Question Clarification in Dialogue via Reinforcement Learning

Xiang Hu[†] **Zujie Wen**[†] **Yafang Wang** [†*] **Xiaolong Li**[†] **Gerard de Melo**[‡]

Ant Financial Services Group[†]
Hasso Plattner Institute, University of Potsdam[‡]

Abstract

Coping with ambiguous questions has been a perennial problem in real-world dialogue systems. Although clarification by asking questions is a common form of human interaction, it is hard to define appropriate questions to elicit more specific intents from a user. In this work, we propose a reinforcement model to clarify ambiguous questions by suggesting refinements of the original query. We first formulate a collection partitioning problem to select a set of labels enabling us to distinguish potential unambiguous intents. We list the chosen labels as intent phrases to the user for further confirmation. The selected label along with the original user query then serves as a refined query, for which a suitable response can more easily be identified. The model is trained using reinforcement learning with a deep policy network. We evaluate our model based on real-world user clicks and demonstrate significant improvements across several different experiments.

1 Introduction

In real-world dialogue systems, a substantial portion of all user queries are ambiguous ones for which the system is unable to precisely identify the underlying intent. We observed that many such queries in our question answering (QA) system exhibited one of the following two characteristics.

1. Lack of semantic elements such as subject, object, or predicate, e.g. “How to apply”, “Credit card”.
2. Ambiguous entities, e.g. “My health insurance” (because health insurance consists of numerous sub-categories).

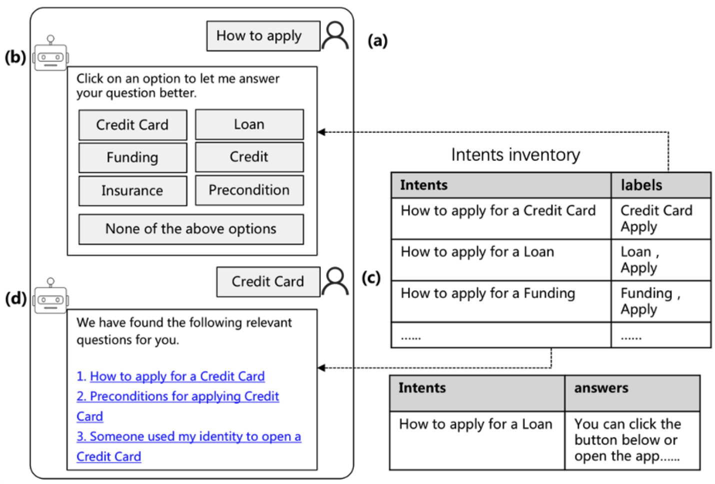

Figure 1: Interactive question clarification example. a) The user provides an incomplete or ambiguous question. b) The agent suggests pertinent labels. c) The user confirms by selecting one such label. d) The agent considers the label in conjunction with the original query as a refined query and responds to it.

*corresponding author, email: yafang.wyf@antfin.com

Given such limited information, it is difficult for a system to accurately respond to a user's ambiguous queries, often resulting in that the user's needs cannot be addressed. For example, the specific intent underlying an utterance such as "How to apply?" remains obscure, because there are too many products related to the action of "applying". In practice, one often needs to fall back to human agents to assist with such requests, increasing the workload and cost. The main purpose of deployed automated systems is to reduce the human workload in scenarios such as customer service hotlines. The lack of an ability to deal with ambiguous questions may directly lead to these sessions being transferred to human agents. In our real-world customer service system, this affects up to 30% of sessions. Hence, it is valuable to find an effective solution to clarify such ambiguous questions automatically, greatly reducing the number of cases requiring human assistance.

Automated question clarification involves confirming a user's intent through interaction. Previous work has explored asking questions (Radlinski and Craswell, 2017; Quarteroni and Manandhar, 2009; Rao and Daumé, 2018; Rao and Daumé, 2019). Unfortunately, clarification by asking questions requires substantial customization for the specific dialogue setting. It is challenging to define appropriate questions to guide users towards providing more accurate information. Coarse questions may leave users confused, while overly specific ones may fail to account for the specific information a user wishes to convey.

In our work, we thus instead investigate interactive clarification by providing the user with specific choices as options, such as intent options (Tang et al., 2011). Unlike previous work, we propose an end-to-end model that suggests labels to clarify ambiguous questions. An example of this sort of approach is given in Figure 1. Here, we consider a closed-domain QA system, where a typical method is to build an intent inventory to address high-frequency requests. In this setting, the set of unambiguous candidate labels for an ambiguous user utterance corresponds to a set of frequently asked questions covered by the intent inventory. In a closed domain, we consider the candidate set to be finite. For example, in Figure 1, there are three specific intents corresponding to the ambiguous question "How to apply".

Our approach induces phrase tags as *labels* for each intent. Thus, we have a catalog of intents with corresponding labels that can be presented to the user. The challenge lies in selecting a suitable list of labels that can effectively clarify the ambiguous question. In our approach, the problem of finding the label sequence is formulated as a collection partitioning problem, where the objective is to cover as many elements as possible while distinguishing elements as clearly as possible. The task of question clarification thus amounts to obtaining a suitable set of labels.

The main contributions of our work are:

1. We formulate interactive clarification as a collection partitioning problem.
2. We propose a novel reward function to evaluate the clarification ability of phrase collections and an end-to-end sequential phrase recommendation model trained with reinforcement learning.
3. Both offline and online experiments confirm that our method outperforms pertinent baselines significantly.

2 Related Work

Query Refinement. Several works explore the use of clarification questions for query refinement (Kotov and Zhai, 2010; Sajjad et al., 2012; Zheng et al., 2011; Ma et al., 2010; Sadikov et al., 2010). For instance, Kotov and Zhai (2010) and Sajjad et al. (2012) use question templates to generate a list of clarification questions. Elgohary et al. (2019) rewrite questions using the dialogue context. Zhang et al. (2019) invoke graph edit distance for query refinement. Other studies rely on reinforcement learning to refine user queries (Nogueira and Cho, 2017; Buck et al., 2018; Liu et al., 2019), but consider queries that are unambiguous (though possibly ill-formed or non-standard). Accordingly, they seek to increase the recall, while in our setting, we consider ambiguous user queries, and our model primarily seeks to address the task of question clarification.

Dialogue. Boni and Manandhar (2003) developed an algorithm to recognize clarification dialogue, rather than for asking clarification questions. Varges et al. (2010) found that the use of clarification has a positive effect on concept precision in task-oriented dialogue. Li et al. (2017) focus on clarification in the specific circumstance of a bot not understanding a teacher because of spelling mistakes, which is a sub-problem of our setting. Zhang et al. (2018) generate clarification questions using language patterns with predicted aspect. They do not use reinforcement learning to optimize the order of the questions. Wang et al. (2018) devised soft and hard-typed decoders to generate good questions by capturing different roles of different word types. Aliannejadi et al. (2019) designed a two-stage retrieval and ranking model to rank clarification question candidates generated by human annotators, different from our end-to-end reinforcement learning approach. Korpusik and Glass (2019) construct clarification questions from a food attribute list (brand, fat, etc.). They rely on a hybrid reinforcement learning approach to select the order of clarification questions to ask, while we present an end-to-end reinforcement learning method.

Question Answering. Some studies focus on clarification questions in a community question answering setting (Braslavski et al., 2017; Rao and Daumé, 2018; Rao and Daumé, 2019). These share in common that they seek to rank or generate clarification questions, while our approach uses reinforcement learning to perform sequential label recommendation for question clarification. The key differences between our work and Tang et al. (2011) are three-fold. First, they rely on an ontology, which limits the applicability of their approach in real-world deployments and prevents us from being able to compare against their approach in our experiments, since each domain requires a custom ontology. Second, they cluster the keywords through the ontology, based on templates to achieve a refinement of questions, without using machine learning. Third, they rely on clustering to increase the keyword diversity, while we design a reward with an information gain term that automatically encourages diversity.

3 Preliminaries

System overview. In order to provide a more concrete picture of our approach, we first briefly describe our QA system, illustrated in Figure 2, as an example of how this approach can be instantiated.

When the conversation exceeds a certain number of rounds or the user explicitly requests human service, the conversation is transferred to a human customer service agent. In this setting, our clarification method chiefly serves to reduce the workload of those human agents. In our real system, there are two stages: label clarification and intent retrieval as illustrated in Figure 1. The label clarification stage provides 6 labels for the user to confirm. Upon selecting one of the suggested labels, the user question is concatenated with the selected label phrase as a new query input. The intent retrieval stage seeks to provide 3 relevant intents for the user to select according to the concatenated query. These additional labels can help clarify and improve the relevance.

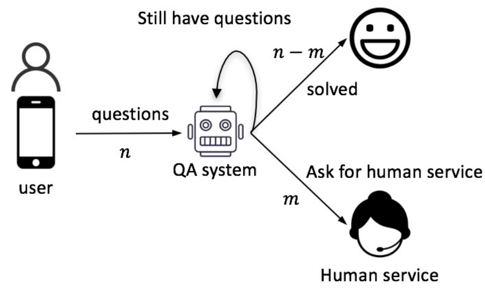

Figure 2: Pipeline of our QA system. $\frac{m}{n}$ is the rate of transferal to human agents (THA).

Intent and Label Inventory. Our system relies on a closed-domain intent and label inventory. The intents along with their corresponding answers are compiled by human experts. The set of labels is a collection of words or phrases that are manually constructed from intents by marking up keywords such as suitable predicates, subjects, or objects. As shown in Figure 3, there is a many-to-many relationship between intents and labels. Note that there is substantial synonymy among the set of labels, which may result in numerous repetitive recommendation results. Thus, ensuring the diversity of the results ought to be a factor in the design of the policy model.

Dataset Setup. In order to solve the cold start problem and evaluate the effectiveness of each model offline, we constructed a benchmark corpus. This annotated corpus consists of 40k ambiguous questions and their potential intents. For this, ten experts were divided into five teams. The two experts in each team annotate the same corpus. Data on which there are disagreements are

Recall	**Valid**
Can't transfer money using Alpha	No
How to apply for a Credit Card	Yes
How to apply for a Loan	Yes
...	...

Table 1: Example of related intent annotation for user question "How to apply".