Automatic Knowledge Assistance System Supporting Operator Responses

Takaaki Hasegawa, Yuichiro Sekiguchi, Setsuo Yamada, and Masafumi Tamoto

Abstract

The variety and complexity of products and services handled by contact centers has increased recently, which places a heavy load on operators, as they must retain more information. This has led to decreasing operator retention rates. This article introduces an automatic knowledge assistance system that assists operators by automatically presenting appropriate information (knowledge) to them while they are handling calls.

Keywords: contact center AI, dialogue understanding, FAQ search

1. Introduction

Contact centers have the important role of being the point of contact between an enterprise and its customers. A major objective of a contact center is to increase customer satisfaction by responding to inquiries quickly and appropriately. However, the variety and complexity of products and services being handled continues to increase, requiring operators to know an increasing amount of information. This places a heavy burden on operators and has resulted in decreasing operator retention rates.

We are developing an automatic knowledge assistance system that supports operators by automatically presenting appropriate information (knowledge) to them as they respond to calls, particularly if they have little experience in the business. However, creating and maintaining the information to be presented to operators can be expensive. To reduce this cost, we are also working on technology to assist with consolidation of frequently asked questions (FAQs), which are one form of such information. This article introduces the automatic knowledge assistance system and the FAQ consolidation technology.

2. Overview of automatic knowledge assistance system

The automatic knowledge assistance system provides support to operators who are less experienced in the business while they are handling calls by presenting appropriate information based on the issue raised by the contact center (or call center) caller. Text of the ongoing dialogue between the operator and customer is shown on the left of the screen viewed by the operator, and several similar high-scoring questions with answers are displayed on the right, which are found automatically in the FAQs, based on statements by the customer regarding the issue and the operator's responses to confirm the issue.

The system performs the following steps:

(1) Conversion of the dialogue to text: Speech recognition is applied to the dialogue between operator and customer, and utterance segmentation is applied to the results so they can be presented in a textual form that is easy for the operator to read.

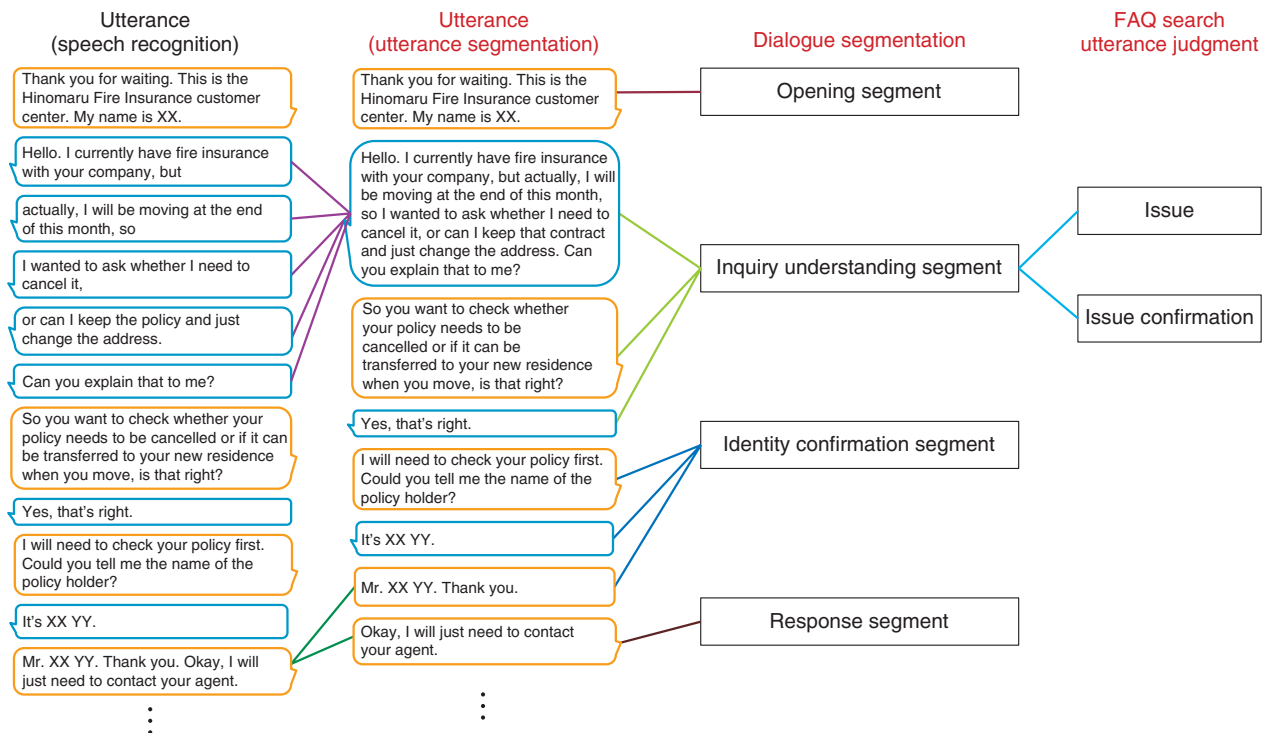(2) Dialogue structuring: Dialogue segmentation is used to apply structure to the dialogue,

Fig. 1. Automatic knowledge assistance steps.

based on features found in contact center interactions.

(3) FAQ search automation: An FAQ search utterance judgment function selects customer utterances that represent the customer's issue and operator utterances to confirm the issue. These are used to search automatically for previously created FAQ issues.

The processes from speech recognition through utterance judgment are illustrated in **Fig. 1**.

2.1 Conversion of speech dialogue to text

Speech recognition technology is used to convert the speech to text, but it is difficult for the operator to see the interaction between the operator and the customer if the results of speech recognition are simply displayed as-is. When the customer calls, they are thinking about the issue as they describe it and may speak slowly, pause, or stutter. As-is, the speech recognition results from the customer's utterances could be divided up with punctuation where there was just a pause, when they should really be displayed as a semantic unit. Thus, displaying them as-is can make the results difficult to understand.

We implemented a function called utterance seg-

mentation that displays utterances in relatively longer units. This function receives the results from speech recognition and determines whether the operator or customer has finished talking, and whether they have made a definitive statement. The decision is made using a deep neural network (DNN) trained using data of approximately 1000 conversation logs. This is complemented with a heuristic that determines whether the speaker has finished by finding points where the dialogue switches from one speaker to the other, while also ignoring confirming sounds made by the other party ("Oh," "uh huh," "Got it," etc.). Overall, this achieves very accurate results.

2.2 Structuring speech dialogue

Dialogues with contact centers tend to be centered on tasks related to the business such as inquiries regarding products or requests for procedures related to a particular service, so typical patterns appear in the flow of dialogue. For example, at a call center receiving inquiries for an insurance product, operators begin by introducing themselves, checking the reason for the call, and confirming the policy holder and policy details before handling the incident. Then they would conclude the call with some kind of

formality. Having an overall grasp of a conversation flow in this way provides strong clues for understanding the dialogue between operator and customer.

We call the function that determines this conversation flow dialogue segmentation. We designed labels suitable for response segments at an inbound call center, created training data of approximately 1000 call center conversation logs, and trained a DNN model to implement technology able to predict dialogue segments accurately.

2.3 Automating FAQ search

To search the FAQs automatically, two conditions must be satisfied. An appropriate query for the search must be selected, and the timing of the search must be appropriate. For the search query, the speech recognition results are not used as-is. The results of utterance segmentation enable us to avoid searching with statement fragments and instead perform queries for full utterances. In the example in Fig. 1, making a query using the partial utterance, "…actually, I will be moving at the end of this month, so…" might result in a search with just the keyword, "moving." Without the keywords that follow such as "cancel" or "change the address," FAQ entries useful to the operator would not be found. Use of utterance segmentation enables the system to select a query that is appropriate for the FAQ search.

Performing a search every time the customer or operator makes an utterance is also too frequent and makes it difficult for the operator to check the search results, so it is necessary to determine which utterances represent the issue. We used machine learning to implement local classifiers, including issue utterance judgment, which determines whether a customer utterance represents an issue, and issue confirmation utterance judgment, which determines whether an operator utterance is confirming the issue. By combining these with the dialogue segmentation described earlier, issue utterance judgment and issue confirmation utterance judgment are applied only to utterances in the inquiry understanding segment. This prevents the selection of irrelevant issues based on utterances in the identity confirmation segment, for example, which in turn increases the accuracy of issue selection. Through this process, queries are made with keywords selected from utterances that have been determined to be related to the issue, and FAQ searches are implemented with appropriate timing.

3. Assisting FAQ creation

Even with FAQ searches based on issue utterances and issue confirmation utterances, if the FAQs being searched are not well organized, a suitable search result cannot be returned. FAQs are composed of question and answer sentences, and since the risk is high if an operator provides an incorrect response, answer sentences must be created and organized by a contact center supervisor that is knowledgeable about the business being handled by the center. However, questions must be created based on details that customers actually query, so they must be regarding matters that are actually common in inquiries. We implemented an FAQ consolidation assistance technology that analyzes past conversation logs accumulated by the contact center, selects candidates for FAQs, and presents them to a contact center supervisor (**Fig. 2**).

Partial utterances that could represent issues are first extracted from the conversation logs using the issue utterance judgment function described above. Issues that customers ask about are many and varied, but issues that have been asked several times in the past are candidates to be added to the FAQs. Thus, issues are collected and filtered according to their frequency of occurrence. In this process, words and phrases that give clues of the issue are reduced to a multi-dimensional vector and used to cluster similar issues. Then, issues that are larger than a set scale are selected from the summarized issue sets and presented as question candidates.

In the past, when building automatic knowledge assistance systems, workers needed to search large conversation logs for question candidates and send them to the center supervisor. Automating this part of the process will contribute to reducing the cost of implementing such systems.

4. Future development

This article introduced technology that extracts issues from dialogues between operators and customers at a contact center and uses them to search FAQs. Much of the functionality introduced is implemented using machine learning, so important outstanding issues include creating large amounts of training data at low cost each time this system is introduced into a contact center, and the ability to generalize models so that they can be applied to various tasks. In the future, we will continue to work on increasing the accuracy of FAQs, and also on reducing the cost of deploying the system.
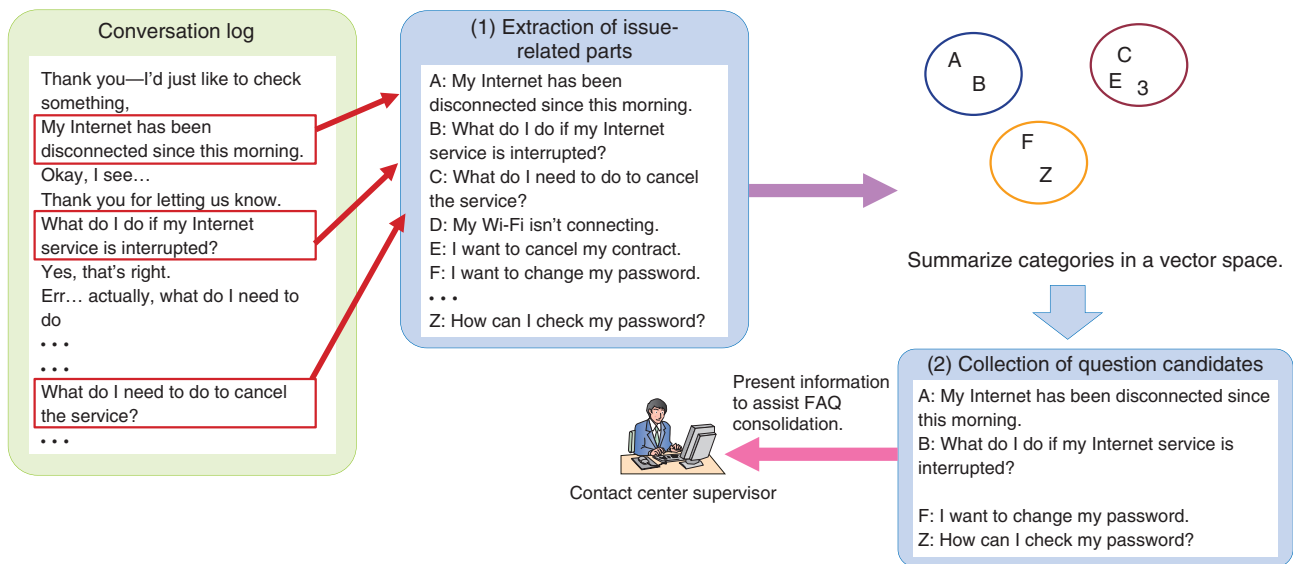
Fig. 2.   FAQ consolidation assistance technology.

Takaaki Hasegawa
   Senior Research Engineer, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
   He received a B.E. and M.E. from Keio University, Tokyo, in 1992 and 1994 and a Dr. Eng. from Tokyo Institute of Technology in 2010. Since joining NTT in 1994, he has been engaged in the research of natural language processing and intelligent information access. He was a visiting researcher at New York University from 2003 to 2004. He is a member of the Information Processing Society of Japan (IPSJ), the Japanese Society for Artificial Intelligence, and the Association for Natural Language Processing (NLP).

Setsuo Yamada
   Research Engineer, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
   He received an M.S. in information science from Tokyo Denki University in 1992 and a Ph.D. in engineering from Tottori University in 2006. He joined NTT in 1992. From 1997 to 2000, he was with ATR Spoken Language Translation Research Laboratories. He rejoined NTT in 2000. His research interests include natural language processing, artificial intelligence, and software engineering. He is a member of IPSJ and NLP.

Yuichiro Sekiguchi
   Senior Research Engineer, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
   He received a B.S. and M.S. in mechanoinformatics from the University of Tokyo in 2002 and 2004. He joined NTT in 2004 and has been researching information retrieval and natural language processing. He is a member of IPSJ and the Database Society of Japan.

Masafumi Tamoto
   Senior Research Engineer, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
   He received a B.E. and M.E. from Tokyo Institute of Technology in 1991 and 1993. Since joining NTT in 1993, he has been working on the research and development of dialogue understanding systems and knowledge media technologies. He received the Awaya Prize Young Researcher Award from the Acoustic Society of Japan (ASJ) in 1999. He is a member of IPSJ and ASJ.