

Argumentative dialogue system based on argumentation structures

Ryuichiro Higashinaka¹, Kazuki Sakai², Hiroaki Sugiyama³, Hiromi Narimatsu³
Tsunehiro Arimoto², Takaaki Fukutomi¹, Kiyooki Matsui¹, Yusuke Ijima¹, Hiroaki Ito¹
Shoko Araki³, Yuichiro Yoshikawa², Hiroshi Ishiguro², and Yoshihiro Matsuo¹

¹ NTT Media Intelligence Laboratories, NTT Corporation

² Graduate School of Engineering Science, Osaka University

³ NTT Communication Science Laboratories, NTT Corporation

Abstract

In this paper, we present the architecture of our argumentative dialogue system that can hold discussions with users by using large-scale argumentation structures. The system can pinpoint argumentation nodes asserted by user utterances and make supportive utterances or rebuttals. The system can be useful for decision-making support as well as promoting better mutual understanding between humans and systems.

1 Introduction

Argumentation is a process of reaching consensus through premises and rebuttals and is important for making decisions and exchanging views. Recent years have seen a large body of work on argumentation mining (Lippi and Torroni, 2016) in which elements that form arguments, such as premises and conclusions, are automatically extracted from natural language text.

Compared to the sizable work on argumentation mining, there has been little investigation in developing argumentative dialogue systems. We believe an automated agent engaging in argumentative dialogue with users will be useful for decision-making support as well as promoting better mutual understanding between humans and systems.

In this paper, we present the architecture of our argumentative dialogue system that enables a natural discussion between a user and system. The system works via text input or speech recognition. The system can be embodied or be a text-based agent. Figure 1 shows a discussion scene with two robots based on our system and a human user. To understand user utterances in a discussion domain (currently, we have five discussion domains including “The pros and cons of auto-driving”) and keep track of the discussion, the system uses large-



Figure 1: Two robots and human user engaging in argumentative dialogue

scale argumentation structures (over 2,000 argumentation nodes for each discussion domain). The system works either in English or Japanese.

2 Architecture

Figure 2 shows the overall architecture of our argumentative dialogue system. The basic flow is that the system understands a user utterance and pinpoints the argumentation node that matches the content of the user utterance in the argumentation structure. Then, the system uses the supportive or non-supportive premises of that argumentation node to utter supportive utterances or rebuttals. When speech recognition or robots are used, the system uses multimodal information to make natural turn-taking possible. We describe how each module in the architecture works below. The modules are connected using the publisher/subscriber model with activeMQ¹.

Voice Activity Detection (VAD) With this module, VAD is carried out so that the system can recognize that the user has started speaking.

User Activity Detection (UAD) With this module, UAD is carried out with a unit attached to a microphone composed of a gyro-sensor

¹<http://activemq.apache.org/>

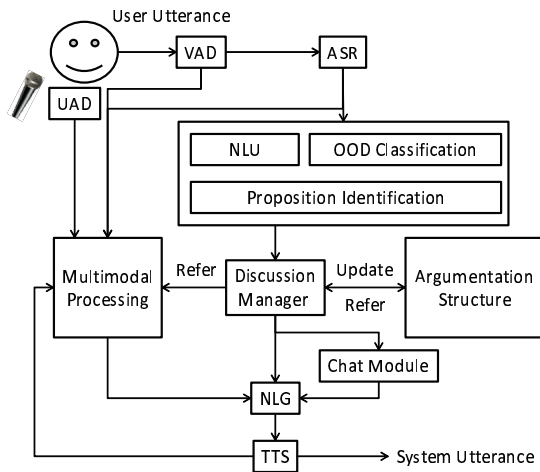


Figure 2: System architecture

and accelerometer. It is used to recognize whether the user is holding a microphone and about to make an utterance.

Automatic Speech Recognition (ASR) We use NTT’s open-vocabulary speech-recognition engine SpeechRec for this module.

Natural Language Understanding (NLU) The NLU module takes as input a user utterance and estimates its dialogue act. We have four dialogue-act types, assertion, question, concession, and retraction. We identify these types as those necessary to update the argumentation structure. We use a logistic-regression-based classifier to carry out this classification.

Out-of-Domain (OOD) Classification OOD classification module determines whether a user utterance is out-of-domain. In the case of OOD, the chat module (details below) will handle the user utterance. We use a logistic-regression-based classifier to carry out this classification.

Proposition Identification This module finds the argumentation node that has the content closest to the input user utterance. The similarity is calculated using the cosine similarity between the sentence vectors created from the averaged word vectors of the statement of an argumentation node and a user utterance. If the similarity is lower than a threshold, it is classified as OOD.

Discussion Manager The discussion manager, for an in-domain utterance, updates the argumentation structure on the basis of the un-

derstanding result and retrieves premises that can be used for support or rebuttal. In the case of OOD, the utterance is fed to the chat module.

Multimodal Processing This module tracks whether the user is speaking or is about to speak and notifies the discussion manager regarding the state of the user.

Argumentation Structure We use the model by Walton (2013) with some extensions. In this model, an argument is represented as a tree (or graph) structure composed of nodes that represent premises; the edges represent support/non-support relationships.

Chat Module The system uses a chat-oriented dialogue system (Higashinaka et al., 2014) (or Alice-based chat-engine (Wallace, 2009) for English) to respond to OOD utterances.

Natural Language Generation (NLG) We use utterances we manually created and associated with argumentation nodes for generation.

Text-to-Speech (TTS) We use NTT’s speech-synthesis engine FutureVoice.

3 Summary and Future Work

We briefly described the architecture of our argumentative dialogue system. We consider this system to be a testbed for future argumentative dialogue systems. For example, we can modify the argumentation structures and test various dialogue strategies. We plan to automatically create argumentation structures from large text data by argumentation mining (Lippi and Torroni, 2016) so that the discussion domain can be extended.

References

- Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chiaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino, and Yoshihiro Matsuo. 2014. Towards an open-domain conversational system fully based on natural language processing. In *Proc. COLING*, pages 928–939.
- Marco Lippi and Paolo Torroni. 2016. Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology*, 16(2):10.
- Richard S Wallace. 2009. The anatomy of alice. In *Parsing the Turing Test*, pages 181–210. Springer.
- Douglas Walton. 2013. *Methods of argumentation*. Cambridge University Press.