

Toward Natural Language Understanding by Machine Reading Comprehension

*Kyosuke Nishida, Itsumi Saito, Atsushi Otsuka,
Kosuke Nishida, Narichika Nomoto, and Hisako Asano*

Abstract

The ability of artificial intelligence (AI) to comprehend text is becoming a major topic of discussion. While natural language understanding by AI poses a difficult problem, a significant improvement in AI reading comprehension has been achieved in recent years through the application of deep learning techniques. This article introduces machine reading comprehension technology now under research and development at NTT with the aim of achieving an agent that can understand business knowledge written in manuals, understand the language used by customers, and provide appropriate answers to questions.

Keywords: machine reading comprehension, natural language understanding, deep learning

1. Machine reading comprehension and natural language understanding

NTT Media Intelligence Laboratories is undertaking the research and development (R&D) of machine reading comprehension to support customer reception at contact centers through artificial intelligence (AI). Machine reading comprehension is a technology involving the use of AI to read the content of manuals, contracts, and other documents and reply to questions. It takes up the challenge of achieving *natural language understanding*, that is, the understanding of everyday human language. The aim here is to support operators at a contact center by having AI interpret the content in manuals and find exact answers even if FAQs (frequently asked questions) have not been prepared beforehand from manuals for search purposes (Fig. 1).

Machine reading comprehension is a new research field that is developing rapidly thanks to progress in deep learning and the preparation of large-scale datasets. It attracted much attention in January 2018 when AI achieved a higher score than humans on the Stanford Question Answering Dataset (SQuAD) [1]—a

machine reading comprehension dataset—prepared by Stanford University. However, the problem setup in SQuAD is relatively simple, and for more difficult problem setups, AI still comes up short against human reading comprehension. While continuing to refine machine reading comprehension technology amid academic competition using datasets for research purposes, researchers at NTT Media Intelligence Laboratories are working to solve key problems with the aim of providing a practical system for contact centers. This article introduces our research results to date in machine reading comprehension.

2. Large-scale machine reading comprehension technology for finding answers from many documents

Problem setups in early research of machine reading comprehension limited the knowledge source to the text in a single document, but actual application scenarios including contact centers require that answers to questions be found from many documents. However, comprehending in detail the content of all such documents based on a machine comprehension

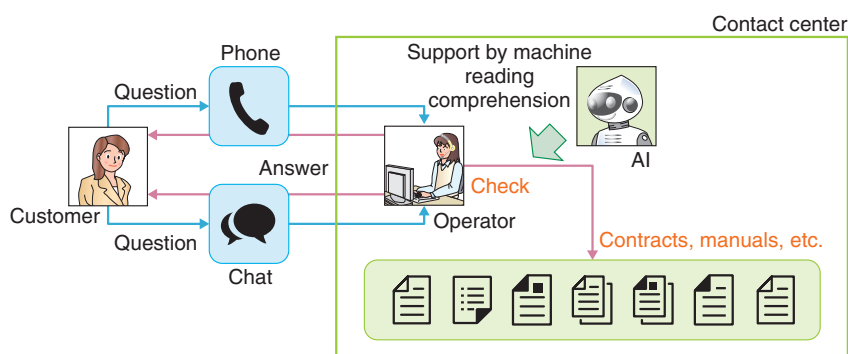


Fig. 1. Contact center and machine reading comprehension.

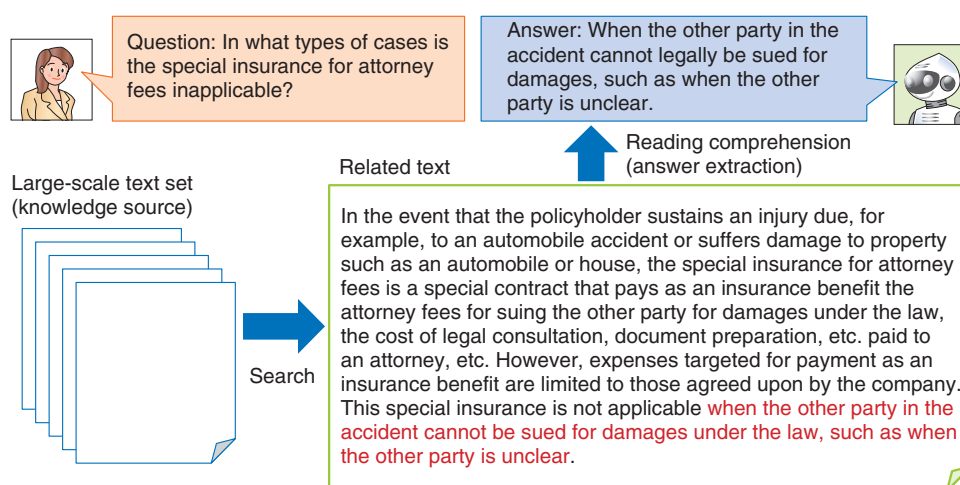


Fig. 2. Large-scale machine reading comprehension.

model would slow down system operation, so there is a need for narrowing down the documents needed for finding an answer in a high-speed and accurate manner.

In response to this need, NTT Media Intelligence Laboratories established large-scale machine reading comprehension technology [2]. This technology first narrows down the documents related to the input question in a broad manner using high-speed keyword search technology. It then further narrows down related documents using a neural search model, and finally, it finds an answer using a neural reading comprehension model (Fig. 2). Thus, we have greatly improved information retrieval accuracy by having AI learn neural information retrieval and neural reading comprehension simultaneously with one model. As a result, we achieved the world's highest question

answering accuracy [2] in a machine reading comprehension task against a set of five million Wikipedia articles in English.

3. Explainable machine reading comprehension technology for understanding and presenting evidence written at multiple locations

It is difficult to guarantee that answers will be 100% correct by AI based on machine learning, so it is important to have a function that enables humans to check the validity of an answer that is output by a machine reading comprehension model. A machine reading comprehension model that can present where in the document set information that serves as evidence for an answer is written would therefore be useful. However, there are cases in which information

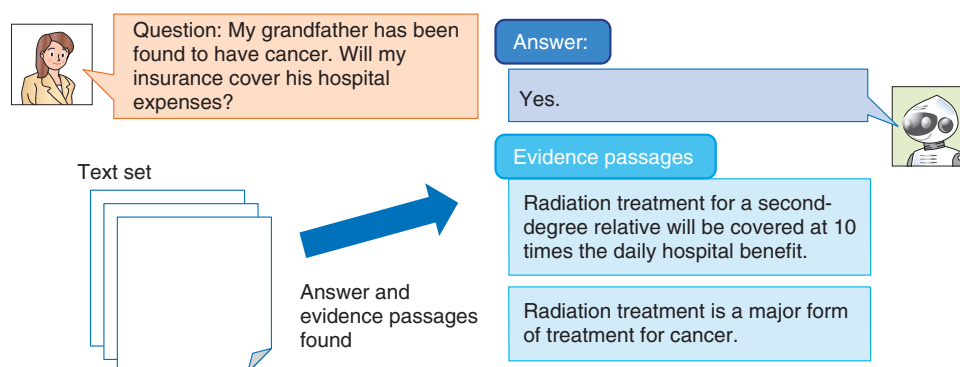


Fig. 3. Explainable machine reading comprehension.

that provides evidence for an answer is divided among multiple passages within the document set. Understanding and extracting all instances of such information is known to be a formidable problem [3].

At NTT Media Intelligence Laboratories, we have addressed this problem by establishing explainable machine reading comprehension technology that combines a model inspired by document summarization technology with a machine reading comprehension model (**Fig. 3**) [4]. Document summarization technology extracts important sentences within text as a summary, but our proposed technique extracts sentences that are important for the answer as evidence sentences. Extracting evidence sentences using a neural evidence extraction model while simultaneously finding answers with a neural reading comprehension model has enabled us to present both an answer and evidence with good accuracy even for difficult problems in which the evidence is divided among multiple passages. Consequently, on applying this technology to the HotpotQA [3] task of outputting an answer and its evidence with respect to questions made against Wikipedia articles in English, we were able to take first place on the leaderboard.

4. Specific query generation technology to clarify a vague question

Datasets for research use in the field of machine reading comprehension are commonly prepared in such a way that an answer can be uniquely specified. However, in an actual contact center or similar operation, there are cases in which the intent of the question input by a customer is vague, preventing an answer from being found. In such situations, more natural communication could be achieved if AI could

appropriately ask the customer about the intent of the question much like a human operator does.

With this in mind, NTT Media Intelligence Laboratories established specific query generation technology to rewrite a vague question into specific questions that can be answered by machine reading comprehension (**Fig. 4**) [5]. The proposed technique extracts candidate answers based on machine reading comprehension with respect to the input question and generates revised questions that remove question ambiguity for each candidate answer.

The example in Fig. 4 is on cancellation fees. As another example, we take the question “What is the maximum insurance benefit provided by special insurance for attorney fees?” In specific query generation, the technique generates multiple specific queries instead of immediately presenting an answer, such as “What is the maximum insurance benefit provided by special insurance for attorney fees for one accident?” and “What is the maximum insurance benefit provided by special insurance for attorney fees for legal consultation and document preparation?” The customer is then asked to select the question nearest to the information desired. In this example, the answer when selecting the former question is “3,000,000 yen,” while the answer when selecting the latter question is “100,000 yen.” This technique enables machine reading comprehension to be applied to a broader range of questions.

5. Generative machine reading comprehension technology for generating multi-style answers

In machine reading comprehension, the most common type of problem setup involves extracting an answer from text serving as a knowledge source.

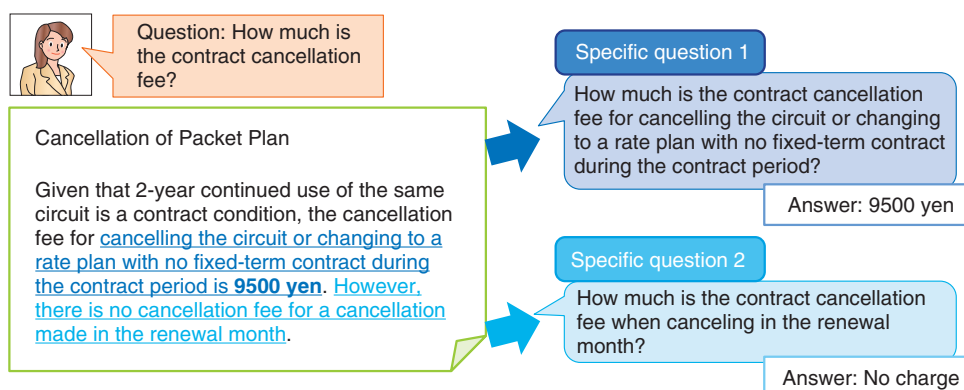


Fig. 4. Specific query generation technology.

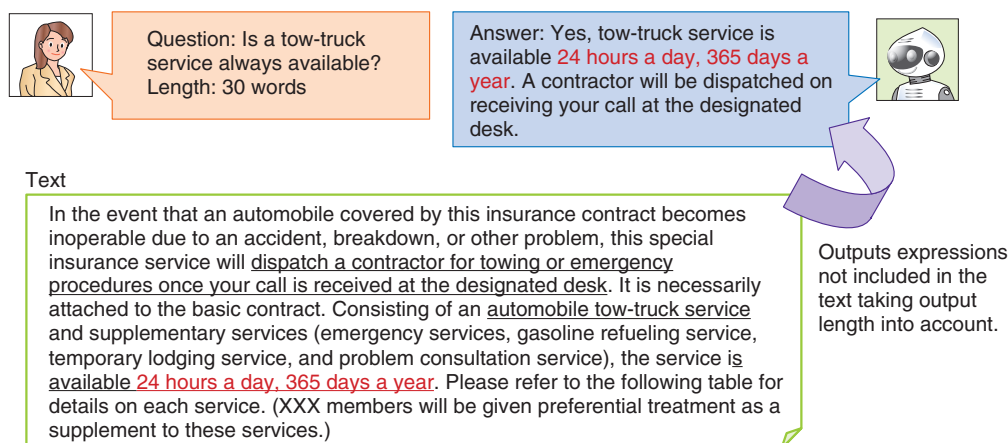


Fig. 5. Generative machine reading comprehension and answer summarization.

Nevertheless, there is a need to generate a more natural style of writing in the case of smart devices, chatbots, and other types of advancing technology. However, the degree of difficulty in generating such an answer is high, and the amount of training data is still insufficient, so worldwide research in this area has yet to advance.

In response to this problem, NTT Media Intelligence Laboratories established generative machine reading comprehension technology that can generate answers with expressions not found in the source text (Fig. 5) [6]. For example, given the question “When is a tow-truck service available?” the proposed technique can generate a natural-sounding answer in the manner of “Tow-truck service is available 24 hours a day, 365 days a year,” which includes not just source text but also appropriate content from the question

itself. This technique can simultaneously learn from machine reading comprehension data with different answering styles and enables the answering style to be selected at the time of generation. This alleviates the problem of having insufficient training data. We applied this technique to two tasks with different answering styles against the MS MARCO [7] dataset that performs open-domain question answering using actual search engine logs and were able to take first place on the leaderboard in that task.

6. Answer summarization technology for controlling answer length

For a chatbot that replies to questions from customers, presenting long answers output from a machine reading comprehension model in their original form

makes for difficult reading by customers. There is therefore a need for appropriately adjusting the length of an answer. In addition, adjusting the length of answers according to present conditions, such as when reading answers on a smartphone or when reading on a personal computer, enables flexible answering tailored to the customer or the current device.

In recognition of this need, NTT Media Intelligence Laboratories established answer summarization technology for controlling answer length using a neural network (Fig. 5). This technology achieves a function for appropriately summarizing an answer to a question by combining a neural network model that identifies important words in text and another neural network model that summarizes (generates) an answer from the words. Here, giving information on length in embedding vector form when summarizing an answer enables that answer to be output in any specified length.

7. Future development

Going forward, we plan to identify problems hindering practical application of machine reading comprehension technology to question answering based on manuals in a contact center. We also plan to feed these problems back to R&D to make technical improvements with the aim of supporting operators through AI and achieving automatic answering. Machine reading comprehension—the understanding of natural language used by humans—is a difficult challenge, but it should foster innovation in various

NTT Group agent-based AI services for contact centers and beyond. NTT Media Intelligence Laboratories is committed to further R&D in this field with the aim of achieving AI that can communicate with human beings using natural language.

References

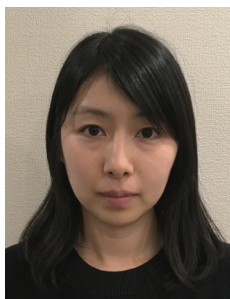
- [1] P. Rajpurkar, J. Zhang, K. Lopyrev, and P. Liang, “SQuAD: 100,000+ Questions for Machine Comprehension of Text,” Proc. of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), pp. 2383–2392, Austin, Tx, USA, Nov. 2016.
- [2] K. Nishida, I. Saito, A. Otsuka, H. Asano, and J. Tomita, “Retrieve-and-Read: Multi-task Learning of Information Retrieval and Reading Comprehension,” Proc. of the 27th ACM International Conference on Information and Knowledge Management (CIKM 2018), pp. 647–656, Torino, Italy, Oct. 2018.
- [3] Z. Yang, P. Qi, S. Zhang, Y. Bengio, W. W. Cohen, R. Salakhutdinov, and C. D. Manning, “HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering,” Proc. of EMNLP 2018, pp. 2369–2380, Brussels, Belgium, Oct./Nov. 2018.
- [4] K. Nishida, K. Nishida, M. Nagata, A. Otsuka, I. Saito, H. Asano, and J. Tomita, “Answering while Summarizing: Multi-task Learning for Multi-hop QA with Evidence Extraction,” The 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019), Florence, Italy, Jul./Aug. 2019.
- [5] A. Otsuka, K. Nishida, I. Saito, H. Asano, and J. Tomita, “Specific Question Generation for Reading Comprehension,” AAAI 2019 Reasoning for Complex Question Answering Workshop, Honolulu, HI, USA, Jan./Feb. 2019.
- [6] K. Nishida, I. Saito, K. Nishida, K. Shinoda, A. Otsuka, H. Asano, and J. Tomita, “Multi-style Generative Reading Comprehension,” ACL 2019, Florence, Italy, Jul./Aug. 2019.
- [7] P. Bajaj, D. Campos, N. Craswell, L. Deng, J. Gao, X. Liu, R. Majumder, A. McNamara, B. Mitra, T. Nguyen, M. Rosenberg, X. Song, A. Stoica, S. Tiwary, and T. Wang, “MS MARCO: A Human Generated MACHine Reading COMprehension Dataset,” arXiv, 1611.09268v3, 2018.



Kyosuke Nishida

Distinguished Researcher, Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.E., MIS, and Ph.D. in information science and technology from Hokkaido University in 2004, 2006, and 2008. He joined NTT in 2009. He was a chief engineer at NTT Resonant from 2013 to 2015. He received the Yamashita SIG Research Award from the Information Processing Society of Japan (IPSJ) in 2015 and the Kambayashi Young Researcher Award from the Database Society of Japan (DBSJ) in 2017. His current interests include natural language processing, artificial intelligence, and web and spatial data mining. He is a member of the Association for Computing Machinery (ACM), IPSJ, the Association for Natural Language Processing (NLP), the Institute of Electronics, Information and Communication Engineers (IEICE), and DBSJ.



Itsumi Saito

Researcher, Knowledge Media Project, NTT Media Intelligence Laboratories.

She received a B.E. from Waseda University, Tokyo, in 2010 and an M.S. from the University of Tokyo in 2012. She joined NTT in 2012. Her research interests include natural language processing. She is a member of IPSJ and NLP.



Atsushi Otsuka

Researcher, Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.S. and M.S. in information science from University of Tsukuba, Ibaraki, in 2011 and 2013. He joined NTT Media Intelligence Laboratories in 2013. His research interests include natural language processing, particularly dialogue processing, and information retrieval. He is a member of IPSJ, the Japanese Society for Artificial Intelligence, and DBSJ.



Kosuke Nishida

Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.E. in engineering and MIS in information science and technology from the University of Tokyo in 2015 and 2017. He joined NTT in 2017. His research interests include natural language processing and mathematical informatics. He is a member of the Association for Computational Linguistics and NLP.



Narichika Nomoto

Senior Research Engineer, Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.A. and M.M.G. in information science and technology from Keio University, Tokyo, in 2005 and 2007. He joined NTT in 2007 and studied speech and emotion recognition. He worked as a data scientist at NTT Communications from 2012 to 2016. He is currently engaged in R&D of natural language processing at NTT Media Intelligence Laboratories.



Hisako Asano

Senior Research Engineer, Supervisor, Knowledge Media Project, NTT Media Intelligence Laboratories.

She received a B.E. in information engineering from Yokohama National University, Kanagawa, in 1991. She joined NTT Information Processing Laboratories in 1991. Her research interests include natural language processing, especially morphological analysis, and information extraction. She is a member of IPSJ and NLP.