Front-line Researchers

Being Different Is a Compliment. Practice Talking Logically about Future Plans

Masaaki Nagata
Senior Distinguished Researcher, NTT Communication Science Laboratories

Overview

Research on machine translation using neural networks has made rapid progress, and the accuracy of machine translation has improved considerably. Although websites and smartphone apps providing machine translation have become more popular, issues with accuracy remain. To address such issues, researchers are studying more advanced translation technology that reflects context, situation, and culture. We interviewed Masaaki Nagata, senior distinguished researcher at NTT Communication Science Laboratories, who is researching context- and situation-based neural machine translation, about the progress of his research and what it means to be a researcher.

Keywords: natural-language processing, machine translation, neural network

Our research has been highly appreciated at one of the top international conferences on natural-language processing

—Please tell us about your current research.

Since 2013, when I was interviewed last time, I have been researching technology for translating one language into another. During this time, the target of my research changed from statistical machine translation based on translating words and grammar to neural machine translation based on context and situation (**Fig. 1**). Artificial intelligence (AI) technology has advanced rapidly since around 2014, for example, AI beat humans in the game of Go. As a result, major changes have emerged in each research field related

to AI. In speech recognition, AI can not only recognize speech, which was considered difficult, but can do so with dramatically improved quality. In machine translation, in which I have been researching, translation accuracy has rapidly improved since 2016, and human parity, namely, human-like translation, has become possible. The time has come when we can suddenly do what was previously thought impossible. In the course of searching for future research themes while considering these trends, I realized that many issues remain to be addressed. One such issue is the fact that the meaning of a translated sentence may change in accordance with the preceding and following sentences. To address such issues, I arrived at the theme of *context- and situation-based neural machine translation*.
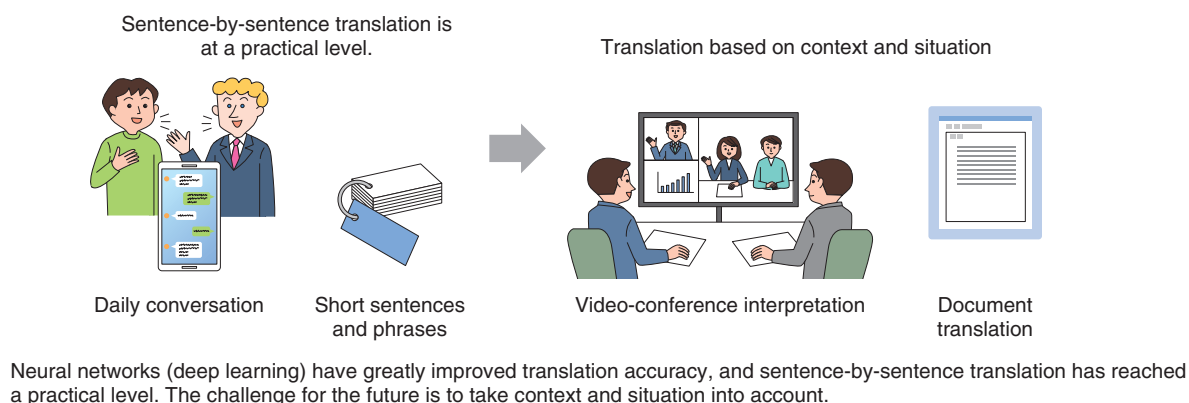
Sentence-by-sentence translation is at a practical level.

Translation based on context and situation

Daily conversation

Short sentences and phrases

Video-conference interpretation

Document translation

Neural networks (deep learning) have greatly improved translation accuracy, and sentence-by-sentence translation has reached a practical level. The challenge for the future is to take context and situation into account.

Fig. 1. Current and near-future goals of machine translation.

If you use the translation function of a website or smartphone app, the accuracy of the translation is so high that you might think that machine translation is better than humans for composing, for example, English for student homework. At a bicycle rental shop, I encountered a situation in which a Japanese shopkeeper and non-Japanese speaker were communicating via the translation function of a smartphone. As I watched the exchange, I was so impressed that I thought to myself, "Wow, translation technology has come this far!" However, the conversation between them was not smooth. This is because subtle phrases and nuances could not be translated.

This lack of smoothness is a slightly troublesome feature of neural machine translation. While the translated text might be as grammatically accurate as a native speaker's text, it may not faithfully reproduce the meaning of the source text. Conventional machine-translation systems use a sentence as the basic input unit. Accordingly, even if the translation accuracy of each sentence equates to sentences produced by humans and if a target text consists of multiple sentences, such as a document or conversation, is machine translated, the translated anaphoric relationships do not match the source text because context and situations are not taken into account. In other words, the translations are inconsistent.

—*You are advancing research to further evolve such current technology.*

To deal with the above translation inconsistency, we must address the following three issues. The first issue is that the subject and object are often omitted in Japanese conversation; therefore, they should be identified from the context and translated. The second issue is to make sure that the translation reflects the characters of the persons in question. This includes, for example, gender bias, as in associating the word "neurosurgeon" with men, even though there are female neurosurgeons. The third is to find and reflect appropriate meanings and expressions of those having multiple meanings.

Let me give an example of translation considering context and situation when translating Japanese into English (**Fig. 2**), where the English translation of the second Japanese sentence is determined in accordance with the first Japanese sentence.

In the example shown in Fig. 2, "I'm afraid that the doctor is running a bit late this afternoon. It might be about 20 minutes before he can see you." is the correct translation from Japanese. However, because the subject "doctor" is omitted from the second Japanese sentence, the machine translation often determines the subject is "we" and presents "It might be about 20 minutes before we can see you." There is also a problem of gender bias. In English, when choosing a pronoun, it is often necessary to specify the gender. Judging that "the doctor is a man," the machine translation is likely to choose the pronoun "he" for the doctor.

When Japanese is translated into English in this way, if the subject and object (which are often omitted in Japanese) are not properly interpreted, the meaning will change, and the translated words will not match the context. For example, in the Japanese sentence "昨日、渋谷に行った。すごい人だった。" ("*Kinou, Shibuya ni itta. Sugoi hito datta.*" (Romanized version) or "I went to Shibuya yesterday. There were a lot of people."), the nuance of "すごい"

	Source language	Target language
Preceding sentence	申し訳ありませんが、先生は午後少し遅れているんです。 (*Moushiwake arimasen-ga, Sensei wa gogo sukoshi okureteirun-desu.*)	I'm afraid that the doctor is running a bit late this afternoon.
Input sentence	診察するまでに 20 分ほどかかると思います。 (*Shinsatsu suru madeni 20-pun hodo kakaru to omoimasu.*)	Correct answer: It might be about 20 minutes before he can see you. Wrong answer: It might be about 20 minutes before we can see you.

	Source language	Target language
Preceding sentence	昨日、渋谷に行った。 (*Kinou, Shibuya ni itta.*)	I went to Shibuya yesterday.
Input sentence	すごい人だった。 (*Sugoi hito datta.*)	Correct answer: There were a lot of people. Wrong answer: He was a great man.

Because gender distinction applies to English pronouns, the problem of gender bias also affects translations of "doctor" and "person."

Fig. 2.   Example of translations that require contextual consideration.

("*Sugoi*" or "great") should be translated into English as "a lot of people" in this context. However, if the sentence "*Sugoi hito datta*" is interpreted in a different way, without reflecting the context from the preceding sentence, it will end up being translated into English as "He was a great man." Although that translation is also correct on a sentence-by-sentence basis, it does not correspond to the preceding sentence, "I went to Shibuya yesterday."

No matter how high the accuracy of the translation, it will never be perfect, so it is important to identify mistranslations. In statistical machine translation, if you want to find the source (Japanese) text that corresponds to the mistranslated (English) text, you can see the internal correspondence between the source and target texts by holding the cursor over the mistranslated text. However, neural machine translation interprets the entire text to create a semantic vector then translates in accordance with that vector, so it is impossible to specifically identify the corresponding text.

To solve this problem, I proceeded with my research with the goal of identifying the correspondences of the input and output sentences, which is called "word alignment" and "sentence alignment."

In 2020, at the Conference on Empirical Methods in Natural Language Processing (EMNLP), one of the world's largest international conferences in the field of natural-language processing and computational linguistics, we announced that we can identify pairs of w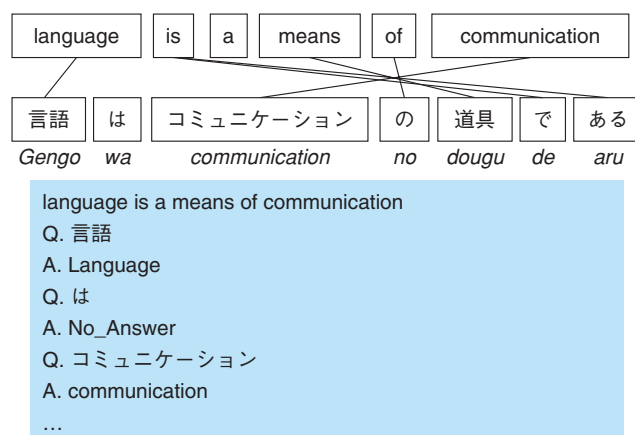ords that are translations of one another by con-sidering their context in the sentences [1] (**Fig. 3**). This "cross-language span prediction" method was recognized as pioneering and remarkably accurate.

Do something different and be as controversial as possible

—You have achieved outstanding research results. What do you value in your research activities?

I try to do things differently from other people and be as controversial—in a good way—as possible. I consider it a problem and very inconvenient that neural machine translation cannot identify word alignment. To solve this problem, I searched for and tried a different method from those of other researchers and discovered by chance that it is easy to train a neural network on whether one word and another word are semantically close, even if the words are in different languages. In fact, word alignment can be obtained with high accuracy if enough data, namely, correct word alignment between about 300 sentences, are available. During the process leading up to that discovery, I first looked at the semantic vectors of words learned by multilingual BERT (Bidirectional Encoder Representations from Transformers)*. I noticed that despite not using the so-called bilingual

---

| language | is | a | means | of | communication |

| 言語 | は | コミュニケーション | の | 道具 | で | ある |
Gengo *wa* *communication* *no* *dougu* *de* *aru*

language is a means of communication
Q. 言語
A. Language
Q. は
A. No_Answer
Q. コミュニケーション
A. communication
…

For each word (span) in the Japanese sentence, find the word sequence (span) that is a translation of each other in the English sentence. The correspondence is determined from the semantic vector representation of the two spans.

Fig. 3.   An example of word alignment based on cross-language span prediction.

corpus, which is a database of multilingual bilingual sentences, in the trained model, similar words exist in a relatively close place in the vector space.

The reason for that finding is twofold. First, even if the languages are different, some parts are written in common notations such as numbers and alphabets. Second, there are commonalities that exists between languages that have similar characteristics. For example, Japanese and Chinese overlap because both languages use Chinese characters. When I looked at that reason in more detail, I thought that there might be a situation in which common notations and words become pivots, and each language expands its own world around those pivots. From that idea, I deduced that it is actually very easy to train a neural network on whether a word and another word in the vector space are semantically close, and if there are correct-answer data for word correspondences between about 300 sentences, word alignment can be obtained with high accuracy.

This groundbreaking discovery was made possible through repeated trial and error, as well as my desire to do something different and controversial. This way of thinking may have come from my experience of studying in the United States in my early thirties. I went to the United States for research twice. I was first a visiting researcher at Carnegie Mellon University (CMU), where I learned that when they say "He is different," they use "different" as a compliment. I then spent time at AT&T, and Dr. Kenneth Church, who was my supervisor, said, "I do things differently

and try to be controversial," and I deeply related with his attitude as a researcher. Those experiences changed my view that "being different" is a good thing.

I also learned the importance of self-review after researching in the United States those two times. To do something different, you have to understand what other people do, so I am still diligently reviewing where my research stands and what I do not understand.

—*Your experience in the United States has greatly influenced your career as a researcher.*

CMU also taught me the mindset of a researcher. During the summer and fall, when new semesters start at universities in the United States, guidance sessions on research and other topics are held. At that time, I learned that research is to be at the forefront of human intelligence and use one's wisdom to elucidate things that are not yet understood. I was fortunate to have the opportunity to learn at the early stage of my research career about the basic concepts of what research entails and how I should contribute. I also remember that in the United States, the evaluation criteria for a doctoral degree, such as research methods, results, and level of achievement, were specifically indicated, so it became easier for me to set my own goals and approach to research.

NTT has an outstanding tradition called "theme planning." It involves presentations of research

projects by researchers in their first and second years after joining the company. During these presentations, the speakers present their research plans in a logical manner for the next three to four years. It is similar to the thesis proposal adopted at graduate schools in the US. Talking about what you have researched is something that you will naturally learn by gaining experience as a researcher; however, logically talking about future plans requires training. I think it is very beneficial to give that opportunity to young researchers in a systematic manner. I also think that understanding your own foothold and clarifying your mission through such activities will lead to research that is different from that of others and will produce research results worth discussing.

On the basis of the approach that I described above, I always try to look at my research objectively and work on something new. I recently started studying graph neural networks to reflect real-world circumstances beyond language in neural machine translation. Graph neural networks have been put to practical use in a variety of applications such as predicting arrival times via Google Maps. This function measures the movement of people between all road intersections and road links and predicts the time it will take to move from point A to point B at each time of day. Similarly, the recommendation function provided by Uber Eats associates information on the person placing the order with ordered items, etc. to predict what the person will order next time. I believe that the accuracy of machine translation can be further improved by reflecting information about real-world circumstances that cannot be verbalized by applying graph neural networks.

My unprecedented "I'm giving up!" declaration

—Looking back on your research career to date, what has left an impression on you, and what will you be working on in the future?

The declaration that I made in the summer of 2016, "I'm giving up research on statistical machine translation," left a lasting impression on me. I have been researching machine translation for about 30 years, and the technological changes brought about by the advent of neural machine translation have had a profound impact on the study of machine translation; in fact, its impact was so huge that it has cancelled out past research. As a result, researchers in the field of neural machine translation are all at the same starting line. Under these circumstances, I declared at a meet-ing announcing my research policy that "Statistical machine translation cannot beat neural machine translation, so I'm giving up researching it." In hindsight, saying that would have been unheard of.

I had been the leader of my team's research on statistical machine translation, so I felt as if I was retreating from the front line. However, with technological innovation, I could conduct research individually instead of in a team and was able to go back to being a researcher. It was a very emotional experience for me to be able to publish a paper as the first author again at the age when I was in a position to lead a team as a manager.

Going forward, I want to leave something behind that will serve as a foundation for research. I think that since data is the key to any research, that foundation is a bilingual corpus or translation database for research on machine translation. A bilingual corpus is a database of bilingual sentences, and a data source is necessary to create it. In the early 2000s, documents from international organizations were used as data sources; however, today, Google Translate, a typical example of an online machine-translation service, uses the web as its data source. Some people argue that it is unacceptable for a single company like Google to accumulate information exclusively. Accordingly, several attempts to create a bilingual corpus by extracting bilingual sentences on a large scale from the web archive of Common Crawl, a non-profit organization that aims to "crawl" the web and provide data to the public, have been made.

According to ParaCrawl, a project to create a large-scale bilingual corpus between European languages and English using data of Common Crawl, the accuracy of machine translation became equivalent to that of translation by humans when 40 million pairs of sentences were accumulated. That number of sentences in the Japanese language has not been accumulated, and if the current trend continues, the machine-translation industry in Japan will be left behind by that in Europe. For that reason, I want to enrich JParaCrawl, a Japanese-English bilingual corpus created by NTT, and contribute to the development of Japanese research on machine translation.

I witnessed a university professor who was about to retire organize his students in an attempt to create a large-scale database of Japanese-English sentence pairs called Tanaka Corpus, which was later incorporated into the "tatoeba" project [2]. At the time, I couldn't understand the passion of Professor Tanaka as I wondered what the point of starting something like this was when he was so close to retirement.

However, now that I am approaching retirement age, I understand the feeling. I also want to continue to be involved in activities to create a database that can be accessed by many people.

—*Do you have any advice for young researchers?*

The longer you pursue a topic of interest, the better off you will be. To that end, I want you to value two behaviors: to acquire the ability to explain your research to others in an easy-to-understand manner and to evaluate your research activities on a regular basis.

Skills to evaluate your research activities on a regular basis can also be developed through writing a research plan. At NTT, you are required to submit a one-year research plan every year. At that time, I review my research plan from the previous year, even though it was my "best effort" plan, I found that my predictions did not come true. However, after repeating the review a few times, you will learn what you can do and what you did as predicted and, conversely, what you cannot do and what you do not predict. As you gain experience, you will be able to formulate research plans on the basis of more-realistic predictions and develop the skills you need to communicate them clearly to a third party.

To work on a long-term project, the theme must be interesting. I was not so sure about my future prospects when I was younger. In the past, I tried to look ahead five or ten years, but it was in the first one or two years that the prospects held some reality. In this age of rapid change, it is extremely difficult to predict the future five or ten years down the line. Even under such circumstances, I think it is important to try to imagine as far as possible what you want to be in your imagined future.

Young researchers, there may be no other job in which you can stay active as long as a researcher. Therefore, find themes that you can research for a long time and continue your research activities focused on those themes.

References

[1] M. Nagata, K. Chousa, and M. Nishino, "A Supervised Word Alignment Method based on Cross-language Span Prediction Using Multilingual BERT," Proc. of EMNLP 2020, pp. 555–565, Nov. 2020.

[2] Tatoeba Project, https://tatoeba.org/eng/

■ **Interviewee profile**
Masaaki Nagata

Senior Distinguished Researcher, Group Leader, NTT Communication Science Laboratories.

He received a B.E., M.E., and Ph.D. in information science from Kyoto University in 1985, 1987, and 1999. He joined NTT in 1987. His research interests include morphological analysis, named entity recognition, parsing, and machine translation. He is a member of the Institute of Electronics, Information and Communication Engineers, the Information Processing Society of Japan, Japanese Society for Artificial Intelligence, the Association for Natural Language Processing, and the Association for Computational Linguistics.