西北工業大學

Koan 4: Natural Language Understanding Robots

Group 1

2019380069 Pang Yu Cong

2019380094 Wen Bin Ian Leong

2019380141 Abid Ali

Abstract

Natural Language Processing is a subset branch of Artificial Intelligence that enables or pushes the capability of a machine to understand, interpret human languages which help to analyze emotions, actions, and thoughts. Natural Language Processing also helps to analyze data and extract information that may be needed to produce meaningful service or serve needs for a project. It is one of the most essential technologies used in intelligent robots, especially conversational AI robots. This paper reports on the future and challenges of natural language understanding robots.

1. Introduction

Robots are slowly becoming part of human daily life, as they are being marketed for commercial applications such as entertainment, services, etc. Thus, the ability to interact with non-expert users is becoming a key requirement.

1.1. What is Natural Language Processing?

Natural language processing (NLP) refers to the branch of computer science—and more specifically, the branch of artificial intelligence or AI — concerned with giving computers the ability to understand text and spoken words in much the same way human beings can.

NLP combines computational linguistics — rule-based modeling of human language — with statistical, machine learning, and deep learning models. Together, these technologies enable computers to process human language in the form of text or voice data and to 'understand' its full meaning, complete with the speaker or writer's intent and sentiment.

Some applications of NLP:

- Chatbots - Chatbots are a great example of Natural Language Processing, where it uses NLP and Machine Learning algorithms to understand and reply as best possible to the user.
- Text Analysis - Text Analysis is one of the applications of Natural Language Processing, where it enables us to get insights into the text and helps to abstract the various insights of the text, including morphological or grammatical analysis.
- Sentiment Analysis - Sentiment analysis which is a subset of Social media monitoring, Natural Language Analysis plays a huge role in analyzing the emotion of the sentence.

- Spell Check - One of the applications of NLP is the ability of Spell Check which we use in our daily life to make sure about the authenticity of any article or text blog.

- Email Classification - During the classification of email, Natural Language Processing helps to differentiate among emails to make sure that the user gets the best experience from the service.

1.2. What is Natural Language Understanding?

Natural Language Understanding (NLU) is also a branch of artificial intelligence that gives computers the ability to understand text and spoken words in much the same way human beings can. NLU enables human-computer interactions.

1.3. What is Natural Language Generation?

Natural Language Generation is a software process that automatically transforms structured data into human-readable text. This technology enables the response generation to the user. In order for the speech to be persuasive and fluent, natural answers must be produced to the user.

1.4. NLU vs NLP vs NLG

Natural Language Processing is an umbrella term that includes both Natural Language Understanding (NLU) and Natural Language Generation (NLG). NLU is more specifically about the meaning or semantics. For example, if the user is asking about today's weather or the traffic conditions on a particular route, NLU helps in understanding the intent of the user's query. NLG is invoked when framing answers in natural language.

2. Techniques and Methods

How do Natural Language Understanding Robots work? NLU Robots like Alana use technology called conversational AI, which enables automatic messaging and conversation between computers and humans. So, how does a conversational AI platform work?

In the simplest word, conversational AI processes words into action. The system of components allows it to understand, respond, and adapt to each interaction. To build a conversational AI like ALANA, there are some core components needed.

- **Inputs:** Inputs are the root of every conversational AI interaction. There are two types of inputs: **text-based input** and **voice-based input**. **Text-based input** functions by typing to interact. This method is discreet while being very consistent

due to direct input. **Voice-based input** works by speaking to interact. This input form offers more versatility via hands-free use, despite being less privacy-friendly.

- **Automated Speech Recognition (ASR):** Automated Speech Recognition allows human beings to use their voices to speak with a computer interface in a way that, in its most sophisticated variations, resembles normal human conversation. When you speak, the device you're speaking to creates a wave file of your words. The wave file is cleaned by removing background noise and normalizing volume. The resulting filtered waveform is then broken down into what are called phonemes. Each phoneme is like a chain link and by analyzing them in sequence, starting from the first phoneme, the ASR software uses statistical probability analysis to deduce whole words and then from there, complete sentences. Your ASR, now having "understood" your words, can respond to you in a meaningful way.

- **Natural Language Processing (NLP):** NLP breaks strings of dialog into cohesive sentences and shapes them to be easily read and acted upon by the AI bot. It also attributes other features like emotion to the input.

- **Natural Language Understanding (NLU):** NLU helps to understand the intent behind the text.

- **Natural Language Generation (NLG):** NLG offers a reply. It helps generate responses into human-understandable language.

- **Machine Learning (ML):** ML comprises a set of algorithms, features, and data sets that helps to learn how to better respond to the user by analyzing human agent responses.

- **Text-to-Speech (TTS):** Text-to-Speech generates texts into voices to give the reply via machine-generated voice.

There are some processes before a conversational robot gives an instant response in natural conversational language when a user speaks/types in their query.

I. **Input Generation**

Input is generated when the user inputs their query. The input could be text-based input or voice-based input.

II. **Input Analysis**

Voice-based input will have one extra step when compared to text-based input. When it is text-based input, NLP and NLU are used to break down the text into parts and pull meaning out of the words.

However, if it is voice-based input, Automatic Speech Recognition (ASR) is first applied to the voice input to parse the sound into language that the machine can understand. The ASR system should be good to understand the user's language and accent and cancel unwanted noise from the surroundings. Once that's done, NLU and NLP understand the intent behind the query, similar to the text-based input.

III. Output Generation

Now that the AI has understood the user's question, it will match the query with a relevant answer and with the help of natural language generation (NLG), it will respond to the user. The response is communicated in a conversational manner, just like how humans interact with each other.

When the output is generated, the AI interacts with the integrated systems to go through the user's profile and previous conversations to add a layer of personalisation in the response.

As in the Input Generation step, voice-based AI bots have an extra step in the Output Generation. The AI's response is converted back from text to speech. The user hears the voice response from the Voice AI, all in real-time.

IV. Reinforcement Learning

This is where the self-learning part of a conversational AI robot comes into play. The user's inputs are analysed and the AI is trained to refine its response. Over time, the user gets quicker and more accurate responses.

3. Can the Robots Really Understand?

Alana can understand context, detect sentiment and recognise patterns of speech when interacting with a human. And this is what conversational AI differs from most standard chatbots. Most standard chatbots are programmed to understand what we say, but not the meaning behind the words. However, conversational AI works differently as it uses Natural Language Understanding (NLU) and machine learning, which allows Alana or any other conversational AI robots to better comprehend what a user really wants from a conversation and to predict where it will go next, no matter how they choose to specifically word a command or phrase or if they change flow mid-sentence.

In the early stage of development of conversational robots, the AI technologies used are mostly rule-based systems or grammar-based parsers. However, slangs, abbreviations, or grammatical errors can often be found in human language. Therefore, a system designed entirely with a rule-based concept often encounters bottlenecks. Also, real human

conversations don't follow pre-written scripts or the predictable structure of a fixed decision tree.

What differs Alana from any other chatbots is that conversational AI like Alana does not follow pre-written scripts or the predictable structure of a fixed decision tree. Instead, Alana is trained to reach conclusions autonomously and spontaneously based on combinations of data, recall and learned experiences.

Even if robots could do this, will humans agree with it? The answer is no, for many reasons. Factors such as human-like gaze, blinking, gestures, vocal tone, and other emotional expressions must be varied and natural, and timed together perfectly. It would be strange, for example, if the robot never broke eye contact with you, or always said "I'm feeling great!" in the exact same way.

The flip side of this is that we don't want robots to be so realistic we're confused about how much they actually know. We don't want robots that can deceive you into thinking they're human.

4. Limitations of Natural Language Processing

NLP is a powerful tool with huge benefits, however, there are still some limitations and challenges to achieve full understanding of human language.

4.1. Contextual words and phrases and homonyms

The same words and phrases can have different meanings according to the context of a sentence and many words – especially in English – have the exact same pronunciation but totally different meanings.

For example:

*I **ran** to the store because we **ran** out of milk.*

*Can I **run** something past you real quick?*

*The house is looking really **run** down.*

These are easy for humans to understand because we read the context of the sentence and we understand all of the different definitions. And, while NLP language models may have learned all of the definitions, differentiating between them in context can present problems.

Homonyms – two or more words that are pronounced the same but have different definitions – can be problematic for question answering and speech-to-text

applications because they aren't written in text form. Usage of their and there, for example, is even a common problem for humans.

4.2. Irony and Sarcasm

The same case as Contextual words and phrases and homonyms, the same words and phrases can have different meanings according to the context of a sentence. What's limiting the NLP is that it uses words and phrases strictly according to definition. What is sarcasm? According to the Oxford dictionary, sarcasm is the use of language that normally signifies the opposite in order to mock or convey contempt. Some people could disagree about its purpose, but there is a convention in that people use positive words in order to convey a negative message.

4.3. Ambiguity

The definition of ambiguity is the state of having more than one possible meaning. It can also refer to a word or statement that can be understood in more than one way. In NLP, there are different types of ambiguity that can limit the development of NLP.

- **Lexical Ambiguity**

  Lexical ambiguity is the ambiguity of a single word. A word can be ambiguous according to its syntactic category (verb, noun, adjective etc.).

  For example, the word 'back' can be used as a verb, a noun, or an adjective.

  *He **backs** up his car to prevent blocking the entrance. (verb)*

  *After sitting all day long, my **back** hurts. (noun)*

  *Please exit through the **back** door. (adjective)*

- **Semantic Ambiguity**

  There can be different interpretations in a sentence.

  For example:

  I saw the boy on the beach with my binoculars.

  This could mean that I saw a boy through my binoculars or the boy had my binoculars with him. Semantic ambiguities arise from the fact that, in general, a computer is incapable of distinguishing between what is logical and what is not.

4.4. Colloquialisms and Slang

Colloquialisms refer to informal phrases, expressions, idioms, and culture-specific lingo that is not formal or literary and is used in ordinary or familiar conversation. For example, the word ‘stan’, normally means a very very overzealous and obsessed fan of a celebrity/band/cast of a tv show or movie, but is never a formal term.

4.5. Domain-specific language

Different businesses and industries often use very different languages. An NLP processing model needed for healthcare, for example, would be very different from one used to process legal documents. These days, however, there are a number of analysis tools trained for specific fields, but extremely niche industries may need to build or train their own models.

4.6. Low-resource languages

NLP applications have been largely built for the most common, widely used languages, especially english. However, many languages, especially those spoken by people with less access to technology often go overlooked and under processed. For example, by some estimations, (depending on language vs. dialect) there are over 3,000 languages in Africa, alone. There simply isn’t very much data on many of these languages.

5. What is the challenging problem for the future?

5.1. Legal Issues

It is an important issue we also face in the near future where if AI collects personal sensitive data, we need proper rules to regulate these sensitive datas.

- **Human Rights**

  If your dataset is biased at the outset, what happens with the output coming from the algorithm? Will certain groups/populations be disadvantaged? What will the consequences be for the company and the individuals?

- **Contracts and Liability**

  If you’re a company, you can’t possibly assume all risks related to your product (especially the unpredictable ones). How do you then draft a fair contract between your clients, yourself and ultimately, the users?

- **Data Privacy**

  You have access to millions and millions of data points – how do you protect them? Can you really protect information, or is it hopeless? If your algorithm is based on confidential or sensitive information, how do you make sure that this information is useful for the product, but also protects the owners of said information? Does AI mean that privacy will be a thing of the past?

- **Intellectual Property**

  Since AI algorithms are highly collaborative, built upon prior research and development, how do we balance the competing concerns of the public good and market advantages? How do we increase collaboration between academia and industry? How can the private legal concerns of creating IP and value be addressed, while we push forward the collective good?

5.2. Computing Problems

Due to the high availability and declining costs of in-house and on-demand cloud computing, computing has rarely been considered as a scarce resource when it comes to enabling the AI revolution. However, computer systems are facing many limitations that can slow the development of AI applications built on top. The main challenges include:

- The end of Moore's law: Due to physical limitations, transistors cannot go in size below a certain level, which is likely to invalid Moore's law sooner than we expect and limit the ability to infinitely shrink microprocessors.

- Increasing data regulation: A speedy exascale processing of data in centralized super-computers requires the data to be stored close to the processors. The increasing appetite for data regulation worldwide, including the recent adoption of GDPR in Europe, is likely to make centralized data placement more difficult and to challenge the purpose of building supercomputers in the first place.

- Lack of a computing-supportive ecosystem: As more IT talent goes into tech giants or launches AI startups, the appetite for working for traditional computer manufacturers or for launching startups focused on developing new computing hardware is on the decline. Despite accelerating the prototyping of AI products, the emergence of high-level programming languages, APIs and libraries is contributing to a less nuanced

understanding of computing architectures and basic computing operations, even among computer scientists.

5.3. Loss of Job Opportunities

As for the potential actual risks of AI nowadays, the one that seems to bring the most concerns is job loss, which in some industries seems inevitable. AI-powered employees have quite a few advantages when compared to their human colleagues. As they have no personal and emotional responses they're never exhausted, bored or distracted, not to mention that they are more productive and efficient. Furthermore, their capacity to make errors is significantly reduced. Such qualities of AI are the most likely to cause layoffs where a lot of tasks can be automated, such as the trucking, food service and retail industry, leading to millions of unemployed and an even higher income inequality.

5.4. Artificial Intelligence Bias

It is an unfortunate fact that human beings are sometimes biased against other religions, genders, nationalities, etc. This bias may unconsciously also enter into the Artificial Intelligence Systems that are developed by human beings. The bias may also creep into the systems because of the flawed data that is generated by human beings. For example, Amazon recently found out that their Machine Learning based recruiting algorithm was biased against women. This algorithm was based on the number of resumes submitted over the past 10 years and the candidates hired. And since most of the candidates were men, the algorithm also favored men over women.

In a separate example, a credit card company may use an AI algorithm that mildly reflects social bias to advertise their products, targeting less-educated people with offers featuring higher interest rates. So the question is "How to tackle this Bias?" How to make sure that Artificial Intelligence is not racist or sexist like some humans in this world. Well, for now the only way to handle this is that AI researchers manually try to remove the bias while developing and training the AI systems and selecting the data.

6. What is the right way for the future?

6.1. Challenging issues in the Future

AI is a new digital technology, with the scarcity of experts and qualified people in the future being the biggest obstacle. Legal concerns are also an important concern that we face in the near future, where we need effective regulations to monitor

these sensitive data if AI collects personal sensitive data. We will be faced with gathering and using relevant information in the future. We will need massive computational power and sufficient AI resources in the future; we need to develop required computing resources to run large quantities of knowledge and use techniques such as deep learning.

6.2. Best Path for the Future

AI would affect each and every sector of our livelihoods. It may affect travel, health care, education, media, client services, etc. AI has been the primary technology for new innovations such as IOT, big data, analytics, etc. Using AI, humans will be able to communicate in their language of choice with each other. AI could replace the world's human labor and run faster than human labor. In chemical factories, in the deep sea and even in mining, space exploration, etc., AI will take over the risky jobs.

6.3. Future NLP Possibilities

- NLP is one of the most important technologies that provide machines with the ability to interpret, understand, evaluate and gather adequate meaning from human languages. Computational Linguistics, which is a synthesis of two technologies, namely Machine Learning (ML) and Artificial Intelligence, is also known as NLP (AI).

- The future implementations of Natural Language Processing will be more user-oriented as technology continues to evolve.

- Digital assistants, for example, can solve several complex questions along with the literal, which implies the query submitted, analyzing the consequences. Not only are the NLP applications limited to answering consumer problems or providing personalized shopping, they have, however, progressed into a greater kind of technical assistance

- Natural Language Processing can be trained to include a list of errors today, if anyone uses NLP to ask, "What's wrong with my network?". NLP will be in a position to find out the real intent of the user in the coming years, as she/he needs continuous access to his/her network.

- With NLP, the future is exciting as progress would encourage people to move attention from the questions to the answers. In the exciting days yet to come, NLP will be built-in to company revenues and make them more effective and agile with various innovations such as gesture and facial recognition.

7. Conclusion

Natural Language Processing (NLP) is changing how language-based knowledge is processed and related. By having machines prepared to understand content and perform human tasks such as summary, translation, characterization, and extraction. In addition, NLP provides companies with a great opportunity to examine unstructured data, including interactions with customer service, product reviews, and social media messages, and gain useful insight into targeted consumers. The way machines comprehend human language seemed to be impossible a few years ago. In a short time, Natural Language Processing (NLP) has become one of the most influential and fastest-growing areas in Artificial Intelligence and Machine Learning. With the growing need of artificial social bots like Alana the days of artificial companionship might be closer than our speculation.

8. References

Khayrallah, H., Trott, S., & Feldman, J. (2015). Natural language for human robot interaction. In *International Conference on Human-Robot Interaction (HRI)*.

Lester, J., Branting, K., & Mott, B. (2004). Conversational agents. The Practical Handbook of Internet Computing, 220-240.

Lim, A. (2020, April 2). Robots Aren't as Smart as You Think. MIT Technology Review. https://www.technologyreview.com/2017/11/02/147995/robots-arent-as-smart-as-you-think/

Major Challenges of Natural Language Processing (NLP). (2020, December 22). MonkeyLearn Blog. https://monkeylearn.com/blog/natural-language-processing-challenges/

Pedamkar, P. (2021, March 3). Artificial Intelligence Problems. EDUCBA. https://www.educba.com/artificial-intelligence-problems/

Ray, S. (n.d.). What Is Natural Language Processing (NLP)? C # Corner. https://www.c-sharpcorner.com/article/what-is-natural-language-processing-nlp/

Smith, B. (2019, August 6). Detecting Sarcasm is difficult, but AI may have an answer. H2O.Ai. https://www.h2o.ai/blog/detecting-sarcasm-is-difficult-but-ai-may-have-an-answer/

Team, S. A. I. (2021, March 24). Exploring the Key Legal Issues in AI Today. Stradigi AI. https://www.stradigi.ai/blog/the-key-legal-issues-in-ai/

Thomas, P. (2021, September 29). Why Conversational AI isn't just another chatbot. Alana. https://alanaai.com/why-alana-conversational-ai-isnt-just-another-chatbot/

Yokota, M. (2019). Natural Language Understanding and Cognitive Robotics. Amsterdam University Press.