# Deficiencies and Upgrades in Artificial Intelligence Languages

Ruoxuan Li

Department of German, Beijing Institute of Technology, Beijing, 102401, China

1120210129@bit.edu.cn

**Abstract.** In the era of artificial intelligence, people have started to have a lot of exposure to machine output language, such as online translation robot customer service, Siri, Microsoft Ice. However, these AI languages are still in the primary stage. Compared with human languages, AI languages still have many shortcomings, for example, lack of intentionality, lack of metaphorical ability. And people need to make some changes and improvements to make them reach a higher level of intelligence, so that AI languages can be closer to natural languages in various aspects, and AI products can better facilitate human production and life, as well as better serve society. This article is written around AI language, through the existing products of AI language as an example, analyzing AI language from different levels such as linguistics and philosophy, finding out the deficiencies of AI language and the aspects that need to be improved, and finally putting forward constructive opinions on how to make today's AI language reach a higher level of intelligence, that is, paying attention to the importance of metaphor in AI, and people should solve the problem of In order to solve the problem of "intelligence", the problem of "intelligence" needed to be solved first; people should not only deal with the problem of man and the world but also deal with the problem of man, nature and the imaginary world.

**Keywords:** Artificial Intelligence; Language; Intentionality; Natural Language.

## 1. Introduction

Today, artificial intelligence is developing rapidly and is deeply connected with people's lives, and it has successfully replaced humans in many fields, for example, Microsoft Ice, Siri, Youdao Translation, and Deeply Translation. And in some aspects, they do better than humans, based on the integrated use of algorithms, cloud computing and big data, deep learning, customer service, social platforms, gradually forming humans. It can be said that without AI, today's modern society will not be able to function, but despite this, AI is still far from being a "universal machine" with the same intelligence as humans. It can be said that AI has some inherent and unavoidable disadvantages.

From the deficiency of the algorithm level, the machine has no subjective emotion and will only make choices based on the probability of the distribution of the data. The result of discrimination by machines following objective rules of output is essentially a mirror reflection of social reality by big data. To truly eliminate algorithmic discrimination, people may need to wait for a more equitable society and the accompanying language revolution. Through the cooperation of various disciplines, the research on language and discrimination will penetrate into various social fields, discover the content and structure of various biases and discrimination, and the correlation between various discrimination elements, anticipate their social impact, and continuously debug the algorithms so that AI products can truly serve society. On a philosophical level, there are many implications for improving AI. For one, it can be viewed from a super-human and super-life perspective, addressing how people represent objects. As for the significance of the application of AI technology in language learning, which is of greater concern to language learners, when innovators and change-makers in the field of language learning can gain a more accurate insight into the innovative capabilities and reasonable boundaries of technology, and have an essential knowledge and understanding of the language learning field, more innovators and innovative teams will surely be able to apply AI technology more scientifically and achieve more efficiently. Technology combinations. In the future, there will be more powerful new applications based on new technologies that will make language learning smarter and more efficient.

In this thesis, my research is organized as follows. First, the author will introduce the concept and importance of AI language, and illustrate the relationship between AI and linguistics by giving an example of assessing the capabilities of the voice assistant Siri in a linguistic perspective. Further on, the deficiencies of AI language will be presented, namely, deficiencies in foundational nature; deficiencies in the understanding of linguistic symbols; and deficiencies in philosophical sense, such as intentionality. This is followed by a description of the ways in which AI languages can grow to higher levels of intelligence, from natural language understanding to human-like language, first by learning "metaphor", then by association, and then by addressing "intelligence".

## 2. The Current Stage of Development and Results of Artificial Intelligence Language

People are now getting a lot of exposure to machine-generated language, with online translation, automated responses, and bots like Siri and Ice. So far, machines have replaced much of the work and application of words, such as weather forecasting and sports commentary. However, on the one hand, machine language is less powerful than human language. For example, lack of intentionality and metaphorical ability; on the other hand, computational linguistics holds great promise and has a long way to go in the future. Taking machine translation as an example, Chinese mathematician and linguist Professor Zhou Haizhong pointed out in his paper "Fifty Years of Machine Translation" that in order to improve the quality of machine translation, the first problem to be solved is the language itself rather than the programming problem; it is definitely impossible to improve the quality of machine translation by relying on a thousand programs to do the machine translation system alone [1]. In addition, it is impossible for machine translation to reach the level of " faithfulness, expressiveness " when the human brain has not yet understood how to carry out fuzzy recognition and logical judgment of language. The development of artificial intelligence requires the application of research results in cognitive linguistics, psycholinguistics, and sociolinguistics to the design of human-computer dialogue [2].

Take Machine Translation as an example. Due to the rapid development of information technology and increasingly frequent international exchanges, how to overcome language barriers has become a common problem faced by the international community. Today, the accuracy of AI translation has increased dramatically. Especially in Japanese and English translation, it is far beyond the level of ordinary people, and has improved to 95% accuracy even compared with professional translators. Translation is becoming increasingly important not only as a discipline but also as an activity. All activities are created by people consciously to achieve certain purposes. Then the fundamental purpose of translation activities, whether they appear in commercial trade, political meetings, academic lectures, etc., is to better communicate and express communication. Translation is apparently the conversion of one language into another, but in fact, it is - a complex process, including input (understanding), conversion, output (re-expression), and other links. In today's society, due to people's continuous pursuit of more efficient and convenient lifestyles, coupled with the increasing demand for translation in daily life, the massive use of machine translation is an inevitable result.

The use of machine translation to assist people in obtaining information quickly has become an inevitable trend, but machine translation still has some problems that affect the readability and accuracy of the translation. First of all, from the perspective of its principle, it is the use of computers to translate different languages, and is also a branch of Natural Language Processing, which is inseparable from Computational Linguistics and Natural Language Understanding. It is also a branch of Natural Language Processing, which is inseparably related to Computational Linguistics and Natural Language Understanding, and is an important field of language application in the information age [3].

Take Siri as an example, Siri can first be evaluated from a linguistic perspective, speech, vocabulary, grammar, and pragmatics. First of all, in terms of speech, Siri has a high accuracy rate in recognizing normal human pronunciation, and can basically understand what people say correctly [4].

Siri can basically understand correctly the phonetic changes in speech flow, such as the same as, weakening, and augmentation phenomena in the process of human speech. However, Siri could not understand the effect of stress on the meaning. Siri replied the same to "I want to eat an· apple" and "I want to eat an apple·" (with dots for stress). The length of pronunciation, which expresses the speaker's tone or emotion and is not related to the meaning of the distinction, does not affect Siri's understanding of the meaning of a person's speech; in other words, Siri cannot sense and respond to changes in a person's emotion. The response to "I'm going to Beijing tomorrow" (fast and excited) is the same as "I'm going to Beijing tomorrow" (slow and frustrated). In addition, Siri does not understand the effect of intonation on meaning, and responds equally to rising and falling intonation in speech. The second aspect is vocabulary. Siri has a huge vocabulary of basic words but a limited vocabulary of new words, archaic words, dialect words, foreign words, and other general words. For example, Siri knew very little about the Internet buzzwords that were very popular at the time. In addition, Siri has an average grasp of fixed phrases and usually invokes the search engine to retrieve them. Siri is able to correctly distinguish between the meanings of the user's sentences in different sequences and can respond correctly to the user's different sentences by opening the address book when it hears, "Open the address book for me." Siri can correctly respond to "Tell me a joke." Siri also understands complexly structured sentences and compound sentences. To the question, "Do you think I'm a good person," the response is, "I'd rather not comment." In addition, Siri can understand compound sentences. Siri does not punctuate the textualization of the user's speech. This shows that Siri cannot accurately distinguish which language unit the user is speaking. Siri does not handle grammatical ambiguity well. For "Set my alarm for next Thursday (it's early Thursday morning)", the response "I can't set an alarm for a day later" automatically classifies "next Thursday" as a whole and does not consider "set" as a whole. The fourth part is pragmatics, in which Siri understands the contextual context of linguistic factors in a given situation (in a conversation) [5].

Siri does not take the previous conversation into the context of the new conversation when it understands it. "Siri can understand more basic, marked indirect speech acts. For "Do you mind telling me a joke?", then a joke is answered. However, it has difficulty with unprompted signs, and is unable to understand that "I'm a little short of money these days" implies that the user or the owner want to borrow money. The Siri example shows that although Siri is a good AI voice assistant and can handle simple conversations in most cases, there is still a significant gap between Siri and human language skills.

## 3. The Shortcomings of Artificial Intelligence Language

To clarify the problems of AI language, the method of basic machine language will be presented. The process of machine translation can be generally simplified into three stages: original text analysis, original text translation transformation and translation generation. On the surface, the core issue of machine translation is the accuracy of translation; in essence, it is the methodological issue adopted by the machine translation system, which is reflected in the technical level. At the methodological level, machine translation systems can be divided into two categories: Rule-Based and Corpus-Based [6]. That is, translation knowledge base consisting of a dictionary plus a grammar rule base and linguistic materials collected for language research and kept in electronic form.

For the existence of the problem, people must first clarify the concept of natural language and artificial intelligence language. Natural language is also the daily language used by people, usually refers to a language that naturally evolves with culture, is the crystallization of human wisdom. Firstly, machine language and program language. "Machine language" and "program language". They are similar to the biological response signals of the human body, which is a kind of instruction and symbol. Signals, a purely passive, set-up symbol. The second level, which is the current stage of artificial intelligence, is the understanding and imitation of natural language. A popular research direction at the moment, it is a dialogue between machine and human, where the machine needs to understand and learn the natural language of humans, and therefore involves knowledge of human language [7].

And the third level, having language like human is the direction and goal of the improvement. On a philosophical level, it reflects the nature of artificial intelligence. Artificial intelligence does not just imitate human language, but also generate language. According to the definition of AI, Human cognition and rational behavior are imitated and automate.

For fundamental deficiencies, the language of artificial intelligence in life, the first is algorithmic bias, as opposed to "experience" in machine learning [8]. It refers to the differences in how a dataset is classified as a particular dataset, how it is given a particular weight, whether it is a test sample or not, and the final results that are extracted. Everyone has their own preferences and dislikes, and these preferences are the factors that need to be taken into account in the process of data analysis. When these preferences are different from others in a dataset that has great similarity to other datasets, this is algorithmic bias.

The second shortcoming is the understanding of linguistic symbols. According to G. Frege's division of linguistic symbol systems into linguistic level (the symbols themselves, e.g., the syntax of natural language), denotational level (the ideas expressed by the symbols), and semantic level (the objects referred to by the symbols), AI manipulates symbols according to artificial built-in programs, regardless of the prevailing symbolist or connectionist view in AI, and is limited to the linguistic level and is unable to understand symbols and language itself [8].

Third, at a deeper level, the emergence and rise of AI has been accompanied by rapid development in various fields of computer science, which, in just a few decades, have together shaped the specific behavior of this era, and even the mindset that characterizes it. Today, the fruits of the weak AI phase have led to fantasies about the coming era of strong AI. However, strong AI with mind may not come as expected, and the path to general AI is not yet clear, the key reason being that human understanding of their own intelligence and mind is still very limited. At the moment, people still generally understand language ability as one of the distinguishing characteristics of humans. According to Randall Davis, in contrast to instincts and stimulus-response associations, intelligent behavior is characterized by (1) prediction; (2) response to change; (3) intentional action; and (4) reasoning [9]. In addition to the intentionality problem, AI is now basically capable of prediction, reaction to change, and reasoning, but people still cannot admit that AI has full intelligence, and the reasons for this include the difficulty of resolving the intentionality problem.

Intentionality is the characteristic of certain mental states and events that point to, are about, involve, or manifest certain other objects and things (in the particular sense of these words below).

Jobn Searle suggests that linguistic intentionality is based on psychological intentionality, and psychological intentionality is based on human biology [8]. The semantics of language as well as action are based on intentionality, which is the mental activity of humans as organisms. Unlike a carbon-based organism, a silicon-based machine cannot generate such mental intentionality. Machines cannot generate such intentionality and ultimately cannot understand human language and behavior.

Intentionality is a more fundamental structure than intention or consciousness; intention is only one type of intentionality, and not all intentional states are conscious. That is, all human actions are in intentional states, and even if a machine realizes a variety of intelligent behaviors, it is still manipulated by humans and its actions are not intentional.

So how does an AI language reach a higher level of intelligence?

The language of AI is concerned with whether AI can own language. A large number of symbols and programs in the machine are specified by human beings. The machine is the executor of these artificial languages, and the machine lacks initiative and intentionality.

The problem of "intelligence" is needed to be solved, in order to solve the problem of "wisdom"; people do not only have to deal with the problem of human and the world, but also the problem of "intelligence" [10]. need to deal not only with the problem of people and the world, but also with the problem of people, nature and the virtual world [9].

Reaching a higher level of intelligence, that is, from level 2 understanding and imitating natural language to level 3 having language like a human being, and from having the ability to make

analogies-associations to being closely linked to the understanding of language metaphors. Analogy-association is a very important ability in human intelligence, because people can learn and experience very limited experiences to acquire knowledge. In contrast to this "small data, big task" model, AI requires a large number of learning samples and only a very limited amount of knowledge, which is a "big data, small task" learning model. This is a "big data, small task" learning model.

## 4. Solution

Human imagination, association, analogy and other thoughts reflected in language are metaphors. Meanwhile, people learn and strengthen these abilities through language metaphor learning. In other words, language also reinforces and consolidates some human cognitive abilities, a two-way feedback process that is missing in current computers and artificial intelligence. For artificial intelligence to truly understand natural language, it should understand the richness of semantics, and humans are able to understand what is beyond words because of their ability to connect apparently different words through metaphorical relationships between languages by analogy, association, and imagination.

The human-machine approach, as envisioned by Hawkins and Kurzweil and others, adds something similar to the new human cerebral cortex from which human language, memory, and self-associative abilities are derived. First of all, the establishment of corpus, keyword search and network to big data are the basic approaches, so as to sort out all the relevant semantics of human statements, artificial intelligence can master the meaning of these statements in the way of deep learning. Traditional metaphor is regarded as a rhetorical device, modern metaphor extends it to the basic features of human cognition and thinking. Its analogy-association ability is enhanced, and after the metaphors in language are used and learned by AI, and may even extend such capabilities from language to images and entire representational systems.

Some AI experts have recognized the importance of metaphor in AI technology and suggested that metaphor research is closely related to reasoning, context, and knowledge graph of AI, and Experts also believe that natural language understanding should pay attention to the metaphorical features in human thinking and language.

## 5. Conclusion

Against the background of the current stage of development of AI language and comparing the rapid development of human society, it can be found that if AI language wants to reach higher level of intelligence, it needs to improve the current problems of lack of basicity of AI language, limitations of understanding language symbols and not having intentionality. In order to reach a more advanced AI, so as to make it have the ability of analogy-association and gradually master the understanding of language metaphors, Hawkins and Kurzweil et al. proposed a new solution, that is, using a combination of human-machine approach, adding something similar to the human cerebral cortex to the machine, and collecting the relevant semantics of human utterances through big data and human search to build a corpus, and the combination of these two approaches can allow AI to use and generate metaphorical language like a human. How to reach the next stage, so that AI language can think like a human and have mental intention like a human, is the focus of this study, and this study finds that making AI language learn the use of metaphors may be a way to solve the problem. In future research, it is important to figure out how the human brain performs fuzzy recognition and logical judgment of language so that it can be applied to AI language. It is important to note that the research direction should not be put on how to make AI imitate human language more accurately, but so that AI can have its own independent logic and thinking, so that it can generate its own language. With the development of AI, and the gradual deepening of thinking about the relationship between humans and intelligence, new ways to enable the development of AI languages are bound to emerge.

## References

[1] Zhou Haizhong. Fifty years of machine translation. China Science News,2012.

[2] Hao Yue. Talking about the complementary relationship between "online web machine translation" and "human translation.North China University of Technology,2019,2(2): 39-42.

[3] Yang Qiong. Online translation quality text category differences: a comparative analysis based on manual measurement. Oriental Translation:54-66.

[4] Na Tian. Applications and Challenges of Artificial Intelligence Technologies in Language Learning. Modern Information Technology.2019,3(19):109-110.

[5] Yang Zhihui. Speech Assistant siri Language Proficiency Assessment in Linguistic Perspective. Computational Languages ,2022,6(6): 156-157.

[6] Lu Weijing. The Ethical Dilemma Facing Artificial Intelligence Robots: The Case of "Microsoft Ice". Broadcasting Realm.2018.

[7] Feng Man, Hu Dan. Focus on language and discrimination research to avoid artificial intelligence side effects. 2019.

[8] Tao Feng. A Philosophical Interpretation of the Language of Artificial Intelligence.Nankai Journal (Philosophy and Social Science Edition),2019,3(9):78-86.

[9] Zheng Wei, Yin Jiayi. Ideas and paths for linguistic research in the era of artificial intelligence. Journal Of East China Normal University.2022.

[10] Dai Yibin. Thinking about Artificial Intelligence in the Perspective of Natural Language Understanding. Journal of Donghua University (Social Science Edition),2022,22(3):26-31.