

Why is NARS good at understanding Chinese?

Bojie Feng, Chongqing University
Tao Wang, Hubei College of Economics

Abstract: On the theoretical level, the paper points out the great value of cognitive linguistics theory to natural language understanding and how to implement it in AGI system. And using the non-axiom reasoning system (NARS), a series of Chinese natural language understanding research was carried out. I hope to promote the common development of AGI and linguistics.

Key words: non-axiom reasoning system; natural language understanding; Chinese; cognitive linguistics; category

1 Why choose Chinese?

Chinese has always been known for its difficulties in the West, and there are many ancient proverbs in the West to prove this. In the field of artificial intelligence, the most well-known thought experiment "Chinese Room" is also named because of this. However, choosing Chinese as the beginning of a natural language understanding of an AGI system is not just because of its difficulty. The real reason is:

1. The goal of the AGI system is to achieve the general intelligence of human beings, so in principle it can understand any natural language.
2. According to the "Sapir-Whorf hypothesis", language reflects the user's thinking habits. The difference between Chinese and Indo-European languages (such as English) reflects the differences in thinking habits between Eastern and Western. Use the natural language understanding technology, comparing the results, finding the common features and differences between Chinese and English, we can better understand the "general intelligence boundary" from the perspective of cognitive science, that is, those thinking characteristics are shared by humans, and which are due to Formed by social culture.
3. Traditional linguistic models and formal tools (such as transformational generative grammar, mathematical logic) have many deviations and loopholes in explaining Chinese grammar and logic phenomena. So, the introduction of the AGI system as a new analytical tool is also very meaningful for linguistics.

2 NLP by NARS

2.1 Cognitive linguistics as the theoretical basis

The traditional methods of natural language processing are mainly based on statistical learning (such as deep learning) or based on formal grammar. The defects of the two are very obvious: statistical learning lacks interpretability, and there are defects in semantic understanding; The linguistic community has a lot of critical voices on formal grammar, and formal grammar still has the problem of symbolic grounding. When NARS understands natural language, it adopts a path of cognitive linguistics. The theory of cognitive linguistics has the following characteristics:

1. Investigate language as a general cognitive activity, rather than focusing on only one part of the language, such as grammar or mapping.
2. The research order is from pragmatics to semantics, and finally to grammar. In the field of philosophy and linguistics, we have long recognized the nature of language is a communicative act, such as Wittgenstein's language game theory and John Searle's speech acts theory. Therefore, pragmatics is the origin and core of language, and grammatical phenomena are derived. Understanding natural language should start with speech acts, regard language as an extension of human interaction, emphasize interactivity, and value the role of motivation and emotion in communication.
3. Emphasize the embodiment. Language is based on the sensorimotor level. Semantics is neither a mapping of the external world nor an internal grammatical rule, but a human being built on the basis of embodied experience. In psychology, many proofs can be provided for this, such as: Piaget's research on children's language development.
4. Emphasize the importance of metaphor and analogy. According to George Lakoff, metaphor is not just a rhetorical device, but a universal way of thinking. Through metaphor, human beings understand and generate more abstract concepts based on their experience. The process of metaphor mainly is composed of analogical reasoning.
5. Category and prototype theory. In cognitive language, the category is continuous, vague, and there is no obvious boundary. This is the opposite of traditional linguistics. There is a prototype in each category, and this prototype is a mental schema based on human experience and may not actually exist. The concept and the prototype in the category are family resemblance.

2.2 Advantages of NARS

NARS is a non-axiomatic reasoning system based on term logic and syllogism reasoning. It has very powerful reasoning and representation capabilities, can handle uncertainty like Bayesian reasoning, and can perform "and", "or", "greater than", "similar", "implication", "temporal", etc. logic operation. It can also represent various levels of information, such as: perception, relationship, abstract concept, procedural knowledge, psychological intention, and so on. At the same time, NARS also has a complete control mechanism that can manage the system's motivation, and attention distribution, to meet the AIKR (knowledge and resource shortage) assumptions. *The more details, please read Non-Axiomatic Logic — A Model of Intelligent Reasoning.[1]*

Because of the above characteristics of NARS, it is very suitable for natural language understanding. NARS has much in common with cognitive linguistics in cognitive theory. Moreover, for some of barriers in the way of the AGI, NARS has many technical accumulations. Just like:

1. For symbolic grounding problems and embodiment. At present, NARS has started many sensory-motion experiments, including object sensorimotor recognition, and behavior prediction. In NARS, sensory-motion is a gestalt, proactive, semantically related process that includes empirical construction and spatial reasoning. Moreover, ImageNet and ConceptNet have been imported into the system as a common sense

experience, providing sufficient corpus for the system.

2. NARS can solve the problem of reasoning in natural language understanding. NARS has strong reasoning skills, especially in non-deductive reasoning (such as analogical reasoning). Therefore, it can understand and generate metaphors. At the same time, because NARS has a complete grammar (Narsese) inside, it can represent natural language grammar and deal with the implicit logical reasoning problem. But please note that the Narsese is not a universal grammar in the sense of Chomsky, it is a thinking tool more similar to "The Language of Thought." Therefore, it cannot be transformed with natural language grammar, but reveals the law of thinking behind it.

Since NARS regards natural language understanding as a generalized intellectual activity, it will mobilize a variety of cognitive functions to deal with, so there is no special development of functional modules for NLP (some pre-processing gadgets are not listed). Due to space limitations, for technical details on how NRAS can understand natural language, see: Natural Language Processing by Reasoning and Learning.[2]

3 How does NRAS understand Chinese?

Chinese has a high degree of flexibility, and there are a large number of cases that do not conform to the theory of transformational generative grammar. Just like: double subject sentence (饭我吃了), semantic reversal (王冕死了父亲) and so on. Therefore, the Chinese linguistics community is keen to use the theory of cognitive linguistics to research Chinese and has achieved many results. The basic theory of natural language understanding in NARS is cognitive linguistics, so from a theoretical perspective, NARS is very suitable for dealing with Chinese. In conceptual representation, because there are various levels of information in NARS, the basic semantic units can be abstracted directly from the perceived information, and the vocabulary can be constructed from this. At the same time, the inference rules in NARS can also be applied to all levels of language from word formation to grammar. Below, we will demonstrate the application of NARS in Chinese word formation through a case study.

Chinese word formation is different from English. English usually has a root and an affix, such as: worker = work + er. The root can be a single word, but the affix cannot be used alone. Chinese is different. Generous words are composed of two "roots" and they can be used alone. This creates a problem that cannot understand the logical relationship between the two. But in fact, the two roots are not randomly combined, but after some reasoning, many use metaphors. For example, the word "mountain foot (山脚)" means piedmont in Chinese. The mountain is a geographical Noun, and the feet are body organs. When the two are combined into one word, metaphor is used. The reasoning process is: The lower part of the person is the foot, the mountain is similar to the person, so the lower part of the mountain is also the foot, that is, the foot of the mountain. NARS can completely realize the above reasoning process: people --> feet, mountains <-> people, so mountains --> feet, that is, the foot of the mountain. Of course, Chinese is apparently endless. The above example is just the tip of the iceberg. There is more content waiting to be studied.

4 Conclusion

As an AGI system, NARS uses cognitive linguistic theory to understand natural

language. Based on the existing research, this paper makes a preliminary experiment on the natural language understanding of Chinese using NARS, reveals some principles of Chinese word formation and the process of reasoning of metaphor. In the end, I hope that NARS can fully understand Chinese and bring inspiration for the study of natural language understanding, AGI, linguistics, and cognitive science.

References

1. Pei Wang.: Non-Axiomatic Logic — A Model of Intelligent Reasoning. Singapore: World Scientific, 2013.
2. Pei Wang.: Natural Language Processing by Reasoning and Learning. International Conference on Artificial General Intelligence, 2013 - Springer
3. George Lakoff, Mark Johnson.: Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. University of Chicago Press. ISBN 0-226-46804-6.1987.
4. George Lakoff, Mark Johnson.: With Metaphors We Live By. University of Chicago Press. ISBN 978-0-226-46801-3.1980.
5. Jean Piaget.: The Language and Thought of the Child (London: Routledge & Kegan Paul, 1926)
6. Jerry Fodor.: The Language of Thought, Harvard University Press, 1975, ISBN [0-674-51030-5](#).