

Cognitive Calculation Standard System for Language, Knowledge and Software: GLPS with GKPS

Commented [1]: supplementary

Xiaohui Zou^{1,3} [0000-0002-5577-8245] and Shunpeng Zou² [0000-0002-6544-8741]

¹ Searle Research Center, No. 100, Renshan Road, Hengqin, Guangdong 519000, China .

² Peng Ge Learning Center, Csanady utca 4/b, 1132, Budapest, Hungary

³ Interdisciplinary Knowledge Center Research Group, Peking University 100871, China.

Commented [2]: supplementary

Abstract.

The purpose of this paper is to introduce the standard system of cognitive computing for language and knowledge and software, namely: global positioning system for language and knowledge and software. The method is as follows: the first step is to construct global language positioning system from the perspective of formalization; the second step is to construct global knowledge positioning system from the perspective of conceptualization; the third step is to establish global software positioning system from the perspective of ontology and carrier respectively; It is characterized by analysis of three factors, namely: cognitive computing of language and knowledge and software. As a result, the standard system of cognitive computing for language and knowledge and software, namely: the global positioning system for language and knowledge and software, is highlighted. The significance lies in natural language understanding, expert knowledge expression and software pattern recognition, the cognitive computing standard system of trinity of language and knowledge and software, namely: global positioning system of language and knowledge and software, can first be found in higher education. A country with a popularized education has real application scenarios, and then it can be popularized to a country with a large population. Therefore, it can be used as a collaborative, intelligent, systematic, efficient and accurate cognitive computing standard and system tool to comprehensively improve the overall quality of the population.

Keywords:

Natural Language Understanding, Expert Knowledge Expression, Software Covering Hardware Pattern Recognition, Global Language Positioning System, Global Knowledge Positioning System, Cognitive Calculation Standard System.

1 Introduction

This paper describes, and is written to conform that the aim is to explore a new method of global language positioning system, that is, natural language processing by using Global Positioning System GPS sequencing positioning. GPS was developed by the United States in the 1970s. It took 20 years and cost 20 billion U.S. dollars. GPS is currently the most successful satellite positioning system and is known as a milestone in human positioning technology. In contrast, the European Union's "Galileo" system is known as the Eu-

ropean version of "GPS"; the Russian "GLONASS" system and China's Beidou satellite navigation system were developed in 1993 and 1994, respectively. They are all measurement and positioning for actual objects. It can be said that it is two completely different systems from the measurement and positioning of Chinese characters described in this article. Generally speaking, they are not comparable. However, one of the salient features of this study lies in the series of transformations made in a unique way after the analogy and association across the border. This is the standpoint of a series of unique features of this study. Turning the impossible into reality is the greatest creation. In short, the basic standpoint of this paper is the scientific and technical basis of physics, mathematics and grammar; the content of psychology, society, and humanities only plays its corresponding role under the constraints of actual use. This effect is already very satisfactory.[1-6]The purpose of this paper also is to introduce the standard system of cognitive computing for language and knowledge and software, namely: global positioning system for language and knowledge and software.

2 Method

The method is: first, it is pointed out that Saussure distinguishes language and speech although it is good for Indo-European language, but it is not enough for Chinese; then, in the process of Chinese information processing that can be divided into Yan and Yu as the formal language from Tasky and Carnap, as well as the way of combining a single element set of pure mathematics to combine the metal group set at all levels to build a broad-language formal information processing model that combines Chinese and arithmetic; finally, compare the broad-language bilingual treatment described herein, translated by Google Machines Translate as typical representatives, narrow bilingual processing, such as English-Chinese or Chinese-English machine translation, distinguish between complete matching and partial matching, and completely not matching as three types, and then through human-computer interaction and collaboration methods, the reasons why the centralized inspection is not matched, which provides a reliable basis for the improvement or optimization of the five links of classification, matching, translation, prediction and decision-making.[7-9]The method also is as follows: the first step is to construct global language positioning system from the perspective of formalization; the second step is to construct global knowledge positioning system from the perspective of conceptualization; the third step is to establish global software positioning system from the perspective of ontology and carrier respectively; It is characterized by analysis of three factors, namely: cognitive computing of language and knowledge and software.

2.1 Language and Speech as the Nature Language

Saussure's distinction between language and speech, in the author's opinion, can go further (especially for Chinese). The analogy between the language system and the chess system is considered by the author of this article to be a very good analogy among Saussure's many analogies. This not only reminds us of Wittgenstein's "language game" and Searle's "Chinese Room", but also allows us to further develop "language chess" or "formal Chinese chess", thus realizing and verifying the genius conception or prophecy of their three famous scholars. [10-13]The specific steps are as follows.

2.2 Yan and Yu as the Formal Language

Yan& monosyllabic and Yu& remixed formal language relational databases is how we fulfill the contemporary NLP or AI that the above geniuses conceived or predicted. Here, monosyllabic and mixed syllables (covering monosyllabic, disyllabic and polysyllabic) are regarded as two categories of languages that can be indirectly formalized in our research, which constitutes a feature of this research and can be regarded as the most prominent one. We know that in Tarski and Carnap, formal languages cover formal semantics

(Tarski) and formal grammars (Chomsky), and later case grammars (Fillmer) explore the relationship between syntactic structure and semantics that it is a theory of the combination of syntax and semantics and the corresponding practice of computer software engineering or language engineering. In contrast, Feigenbaum's expert knowledge base is a practice of knowledge engineering. [14-22]They (the three major projects of language, knowledge and software) have not only been inherited accordingly, but also further combined with the ideas and methods of systems engineering (Xuesen Qian) [23] to further develop what we can call "(language, Knowledge, software) three major systems engineering "AI key approaches or practices, its most notable features are as follows.

2.3 Chinese and Arithmetic

Artificial or engineering technology and humanities and arts were originally two very different systems. However, in our case, they are combined through standardized and personalized or diversified relational data. Specifically, it is the set of single elements and hierarchical tuples that we found in the digital system, and the two sets of monosyllabic (language/character) and mixed syllables (language/word) that we distinguish in the Chinese system. , not only satisfies the logical laws of sequence and position, but also satisfies the corresponding mathematical law (specifically, the synonymous parallel correspondence conversion law followed by the linkage function). As for the laws of linguistics and translation outside of logic and mathematics, they can be explained from another perspective that can become a broad cross-border perspective of generalized language, generalized bilingualism, and generalized translation. Among them, the expression form of the three principles can be simplified or unified in the scope of P base by the popular method of simplifying the complex and the two sets of symbol systems of Chinese and arithmetic. This is almost completely different from the existing machine translation paradigm, for example: the following are typical.

2.4 Google Machines Translate

We all know that the current translation quality of Google's machine translation system is very good. [24-27]When a text or even a short article within a few thousand words is input, not only can a reference translation be given in an instant, but its readability needs to be adjusted for bilinguals. There are very few places for word order or grammar. Even if there are bilingual differences in polysemy or idioms, they can be corrected by translators, i.e. bilinguals, and revised or corrected in time with the help of computer aids or LNP&AI gadgets or tools. [28-32]So we found a good way to combine the old and new methods to further improve the expert knowledge base of and the general knowledge base. Its specific practices are summarized as follows.

2.5 Matching with Three Types

For the three types of complete matching, partial matching and complete mismatch, the positive and negative extreme cases can almost be directly classified by batch processing and machine learning methods, that is: either complete matching is regarded as successful machine translation, or a complete mismatch is considered a machine translation failure; bilingual pairs (whether words, phrases, or sentences) in which there is a partial match (in other words, a partial mismatch) can be classified as requiring humans (bilinguals and experts) intervention, that is, through human-computer interaction and even interpersonal and human-computer cooperation, focus on examining the specific reasons for the mismatch, so as to further improve or optimize the five links of classification, matching, translation, prediction and decision-making, and provide reliable information in accordance with.

3 Result: Improvement or Optimization

The result is not only a significant improvement in the quality of machine translation, but also a new paradigm of natural language processing for further data mining, semantic recognition, information extraction, knowledge processing and software modeling, as well as NLP&AI technology quality improvement, namely: Global Language Positioning System. As a result, the standard system of cognitive computing for language and knowledge and software, namely: the global positioning system for language and knowledge and software, is highlighted.

3.1 Human -Computer Interaction

In our opinion, multiple choices of natural language word meanings or lexical entries are inevitable (it is one of the most prominent fundamental features of natural language), therefore, misunderstanding of word understanding due to ambiguity is a frequent occurrence. Therefore, it is usually the case that the combination of man and machine is used to resolve various ambiguity problems. Therefore, a complementary approach between human-machine interface and human-machine collaboration is necessary. To further improve the quality of machine translation and improve the quality of NLP & AI technology requires the following unremitting efforts.

3.2 Improved the Quality of Machine Translation

The key is a small number, and it is very important, especially in terms of expert knowledge such as interdisciplinary and inter-industry reading comprehension advanced knowledge or innovative ideas. Therefore, the expert knowledge acquisition is a difficulty in AI & knowledge engineering; at the same time, natural language processing is also a difficulty in AI & language engineering, and software engineering modeling is a difficulty in AI & software engineering. In our research, we found that the above three difficulties are often intertwined. Therefore, expert knowledge expression, natural language understanding, and software pattern recognition (including software requirement confirmation) are collectively referred to as the core key problems of AI. It can also be said that it is the ambiguity problem of NLP & AI (natural language understanding, expert knowledge expression, and software pattern recognition, which are commonly encountered in three aspects, including local problems and overall problems). Machine translation can be regarded as a typical, both local and global ambiguity resolution problems are covered.

3.3 Improvement of NLP & AI Technology Quality

Further data mining, semantic recognition, information extraction, knowledge processing and software modeling, as well as the improvement of the quality of NLP&AI technology, are mainly considered from several aspects: on the one hand, it is considered from the perspective of preventing problems before they occur; Starting with specific methods that have been verified in many aspects, such as data mining, semantic recognition, information extraction, knowledge processing and software modeling, as well as intelligent text analysis and knowledge module refinement, all need special attention.

Table 1. Information processing by using the way of DB as using GPS.

DB	GPS in Language and Knowledge	Information processing
A	Global Language Positioning System	Formal information batch processing
B	Global Knowledge Positioning System	Content information HCI processing

It can be seen from Table 1 that natural language processing is selected here by combining two types of information processing, formal and content, which are characterized in that: GLPS and GKPS are based on the same generalized bilingual database (DB is regarded as GPS in the dual domain of language and

knowledge) is implemented. In conjunction with the subsequent Figure 1, you can better understand our NLP&AI information processing system.

3.4 Global Language Positioning System

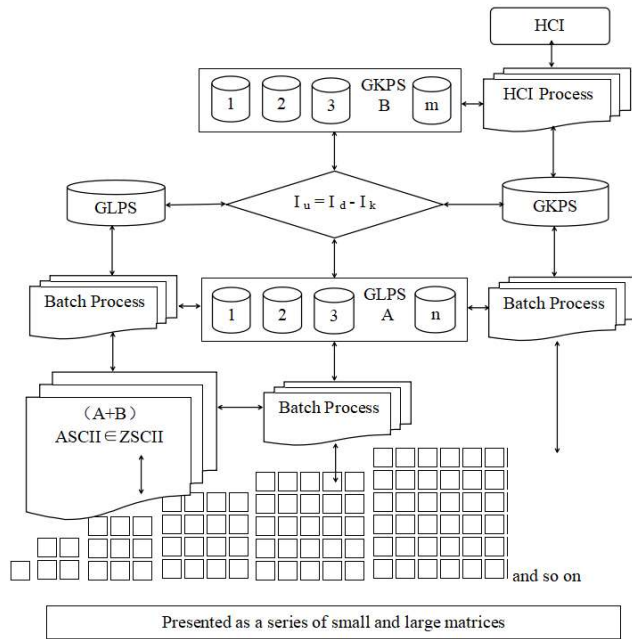


Fig. 1. A as GLPS and B as GKPS are two sets of relational databases

It can be seen from the figure 1 that A and B are two sets of relational databases with one-to-one correspondence, which are characterized by: the consistency of language and text symbol forms, expert knowledge and public common sense are the choices made in it (it is assumed here that they are in the form of also consistent) is determined by the human-computer interaction process (the corresponding background batch process is determined by pure formal consistency) where one-to-one between mono-syllabic, di-syllabic, tri-syllabic, etc., poly-syllabic the corresponding functional relationship determines its matching.

The Global Language Positioning System (GLPS) is a great idea, and with it, at least natural language formal understanding can become a reality. We have proposed such a good idea for many years. Why hasn't it attracted the attention of academic colleagues and has been implemented in Indo-European languages like English? We know that morphemes, words, phrases, and sentences are the structural units of English. From the perspective of basic linguistics, among them, words (often referring to abstract concepts or concrete objects) and sentences (subject-verb agreement) are also well recognized, especially: no word segmentation is required as in Chinese after the introduction of no words, and there is no need to break sentences like the Chinese after the introduction of the non-sentence-free text record form that was accustomed to the end of

the sentence; however, there are situations where morphemes, words, phrases, and sentences are nested or iteratively indistinguishable from each other. Especially in Indo-European languages such as German or French where the grammar is more complex and indistinguishable than English. The main thing is that they are almost all mixed-syllable languages (used to be said to be polysyllabic languages), unlike Chinese which are monosyllabic languages (this typical feature actually hides a linguist and philosopher of language as well as mathematicians, scientific philosophers, philosophers of mind, etc., have consciously or unintentionally coveted properties that we have discovered and refined and exploited scientifically without wasting time, resulting in our conceived the GLPS. See the figure below for details.

3.5 Classification, Matching, Translation, Prediction and Decision -making

With or without GLPS, the implementation of these five steps is very different. If there is, there is a unified frame of reference around the world; if there is no, there is no (but only groping in the dark)

Table 2. Batch Process in GLPS&A,HCI in GKPS&B.

DB	Classification	Matching	Translation	Prediction	Decision -making
A	Batch Process	Batch Process	Batch Process	Batch Process	Batch Process
B	HCI	HCI	HCI	HCI	HCI

The objects or ontology processed by the two groups of methods A and B in Table 1 are the same (as is the case on both sides of the formula below).

$$A + C = B \quad (1)$$

A as GLPS with C as GSPS on the formal side, B as GKPS on the content side, and the both refer to the same object or ontology. And the trend chart of the three major technical fields below tells us from another aspect: the development of each of the three aspects in the past.

It can be seen from Figure 2 that the trends in the three aspects of natural language understanding, expert knowledge expression and software pattern recognition are quite different, which shows that the three aspects of previous research methods are independent. It proves the author's point of view from another aspect, that is, looking at the three aspects as a whole, and adopting the method of the three major systems engineering, that is, the method of the three major positioning systems, is unique.

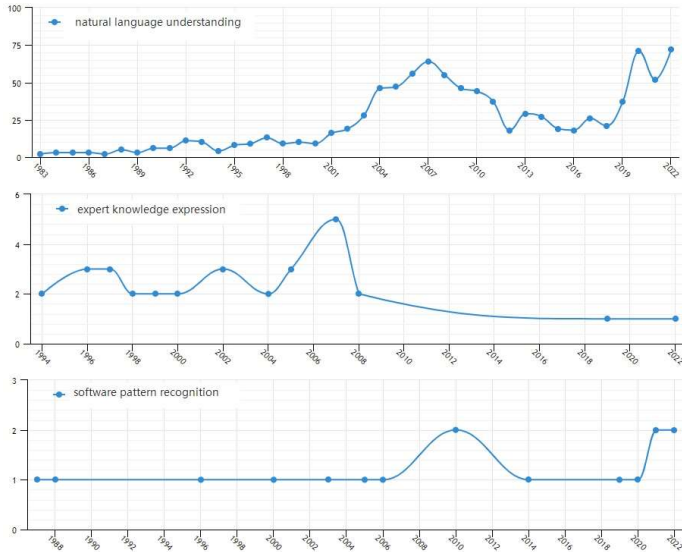


Fig. 2. Language, knowledge, software, three technical fields published trend chart.

4 Conclusions

Its significance is: this new paradigm not only practically efficient and precise, especially Chinese information processing, but also, theoretically, with the existing natural language processing methods, and it is not only the same function which can be made in three aspects of natural language understanding, expert knowledge expression, and software model recognition. The significance also lies in natural language understanding, expert knowledge expression and software pattern recognition, the cognitive computing standard system of trinity of language and knowledge and software, namely: global positioning system of language and knowledge and software, can first be found in higher education. A country with a popularized education has real application scenarios, and then it can be popularized to a country with a large population. Therefore, it can be used as a collaborative, intelligent, systematic, efficient and accurate cognitive computing standard and system tool to comprehensively improve the overall quality of the population.[33-36]

As can be seen from Table 1, Figure 1 and Table 2, the two databases of language (A) and knowledge (B) are constructed as GLPS and GKPS in our design. The batch processing of the formal information by the computer and the human-computer interaction process that could make targeted fine-tuning of the content information (for language formal characteristics, knowledge content attributes, software and hardware logic & physical signs selection and confirmation), which ensures the GLPS&GKPS described the basic features in this study. It is especially suitable for the participation of teachers and students, and it is continuously improved or optimized with the two databases.

Commented [3]: I have cited related papers published in IJHAS in our paper.

Acknowledgement

Thank Peking University for giving the first and corresponding author of this article the opportunity to personally host interdisciplinary knowledge centers and high-level interdisciplinary forums in multiple faculties. The International Conference on Intelligence Science 2018 (Peking University) Language Cognition Forum Moderator. The International Conference on Cognitive Systems and Information Processing 2018 (Tsinghua University) Rongzhixue and Cognitive Systems Forum Moderator. And in the last three years, Xiaohui taught Introduction to Rongzhixue (including linguistics and information in the Rain Classroom namely Online Classroom System of Tsinghua University) is the focus series courses, and the recent International Conference on Cognitive Computing and Systems (Cognitive Professional Committee of the Chinese Society of Automation), which is the author of many series of original results, further verified the practicability and theoretical value of this research from various aspects.

Commented [4]: supplementary

References

1. BW Parkinson and JJ Spilker (1996). Global Positioning System: Theory and Applications.
2. Baker, B. (2017). Global Positioning System. Encyclopedia of Geographic Information Systems
3. Menezes da Silva, C., Setti Júnior, PD, Alves, DB and Monico, JF (2020). Galileo: a consolidação do sistema de posicionamento europeu.
4. Zhilinsky, VO, Petrov, EN, Petrova, AF, Gagarina, LG and Lebedev, AV (2021). Research on the influence of ephemeris and clock information on GLONASS standard positioning service. 2021 IEEE Russian Conference of Young Researchers in Electrical and Electronic Engineering (ElConRus), 2311-2316.
5. Krolikowski, A. (2014). China's Beidou Vision: The Defense and Industrial Strategy Behind the World's Fastest Growing Global Navigation Satellite System.
6. Lee D. (2021). China's Beidou Satellite Navigation System: Past Achievements and Future Challenges.
7. Fournet, A. (2010). Some comparative and historical considerations regarding Ferdinand de Saussure's distinction between "language" and "speech".
8. Qin Weiyuan, & Zi Jianmin (2009). Quine et al. Misinterpretation of Tarski's Equivalence Schema (T).
9. Xiaobing (1987). On the Philosophical Significance of Carnap's Logical Syntax of Language—Also on "The Logical Syntax of Language".
10. Prosiannyk, O. (2021). Temporal states of language activity: On the distinction between ontological and epistemological concepts in F. de Saussure's scientific concepts study linguistics.
11. Chen Baoya and Chen Yue (2015). Meaning is use, and rules are analogy—starting from Wittgenstein's view of language and mathematics.
12. M. Stockhoff, Ma Minghui (translator) (2009). Meaningful-Logical Philosophy of Architecture and Stegendum of Morphology.
13. Damper, RI (2006). The logic of Searle's China room argument. Minds and Machines, 16, 163-183.
14. Al-Fedaghi, SS (2020). Towards Conceptual Modeling Semantics: Eventing Tarski's Ground-truth Patterns. ArXiv, abs/2011.01608.
15. Stockhoff, M. (2009). The Architecture of Meaning: Wittgenstein's Tractatus and Formal Semantics.
16. Simchen, O. (2019). Model truth for semantics. Analytical Philosophy.
17. JS Julian (1968). Comments: Noam Chomsky, George A. Miller, Introduction to Formal Analysis of Natural Languages; Noam Chomsky, Formal Properties of Grammars; George A. Miller, Noam Chomsky, Finite Models of Language Users. Journal of Symbolic Logic, 33, 299-300.
18. Nagy, B. (2018). Derivative "trees" and parallelism in Chomsky grammar. triangle.
19. Briseño, JL, Jimnez, MD and Olague, G. (2018). Robot behavior based on formal grammar. Global Journal of Computer Science and Technology.
20. Anderson, JM (2018). case grammar. About case grammar.
21. O'Hagan, A. (2019). Inspired by expert knowledge: subjective but scientific. American Statistician, 73, 69-81.

- 22.Chattha, MA, Siddiqui, SA, Malik, MI, Elst, LV, Dengel, AR, & Ahmed, S. (2019). KINN: Incorporating Expert Knowledge into Neural Networks. ArXiv, abs/1902.05653.
- 23.Waizu L. (2006). The four stages of systems science developed by Qian Xuesen. Journal of Xi'an Jiaotong University.
- 24.Johnson, M., Schuster, M., Le, QV, Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, FB, Wattenberg, M., Corrado, GS, Hughes, M . ., & Dean, J. (2017). Google's multilingual neural machine translation system: enabling zero-shot translation. Transactions of the Association for Computational Linguistics, 5, 339-351.
- 25.Viswanathan, A., Wang, VH and Kononova, A. (2019). Use AutoML to control the form and style of machine translation output. SIM is big.
- 26.Mitchell, SK (2017). The neural machine translation technology used by Google.
- 27.Hao C. (2021). Chinese-English machine translation.
- 28.Jawahar, G., Sagot, B. and Seddah, D. (2019). What did BERT learn from language structures? Access Control List.
- 29.Shah, D., Schwartz, HA and Hovy, D. (2020). Prediction Bias in Natural Language Processing Models: A Conceptual Framework and Overview. Access Control List.
- 30.Ben-Zaken, E., Ravfogel, S. and Goldberg, Y. (2022). BitFit: Simple parameter efficient fine-tuning of Transformer-based masked language models. Access Control List.
- 31.Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., and Zettlemoyer, L. (2020). BART: Sequence-to-sequence denoising pretraining for natural language generation, translation, and understanding. Access Control List.
- 32.Gururangan, S., Marasović, A., Swayamdipta, S., Lo, K., Beltagy, I., Downey, D., & Smith, NA (2020). Don't Stop Pretraining: Adapt Language Models to Domains and Tasks. ArXiv, abs/2004.10964
- 33.Hulle, Dirk Van. "Artificial imagination, imagine: new developments in digital scholarly editing." International Journal of Digital Humanities (2019): 1-4.
- 34.Ries, Thorsten and Gábor Palkó. "Born-digital archives." International Journal of Digital Humanities 1 (2019): 1-11.
- 35.Vlassenroot, Eveline et al. "Web-archiving and social media: an exploratory analysis." International Journal of Digital Humanities 2 (2021): 107 - 128.
- 36.Mechant, Peter. "Web 25. Histories from the first 25 years of the world wide web." International Journal of Digital Humanities 1 (2019): 133-136.