Semantic understanding processing model based on machine learning

Kunying Li *, Yu Ding, Ying Shi, Hongchao Xu, Yufei Zeng

Computer Application Technology Department,
PetroChina Research Institute of Petroleum Exploration & Development,
Beijing 100083, China
likunying@petrochina.com.cn

Abstract. Machine learning and artificial intelligence have become very diverse, and progress has been made in perception technology, which can recognize the senses of machines, including hearing, vision, smell, taste and touch. We are recently studying a technological change that has grown the dialogue information in our lives into a mainstream recognizable digital activity. This project will build a conversation subject model in the computing framework, which tracks the conversation speech behavior by the functional grammar model. The recognition of Machine Cognition and logic language is the key path of our research. The language model used is Role and Reference Grammar (RRG) to decompose sentences and understanding. In addition, we find an optimization scheme to assist the main body thinking is internal effect and external effect. (1) The subject has its own emotions (anger, joy, sadness, disappointment, joy), which can affect the subject's state at any time. (2) the external effect of the subject is determined by people and other subjects, and the intelligent conversation agent will get interaction and composite response through two effects. The RRG directional link system will be used in the language framework to build a language model in the software to help the machine analyze human language through speech behavior. This interactive language analysis model ensures that the machine can accurately define human meaning in intelligent conversation. In this project, we will improve the development of machine learning and artificial intelligence in human language recognition and analysis, and the technology is very useful in various human behaviors such as smart home, enterprise management and data analysis. There is no doubt that this technology will make a great contribution to human beings in the field of Computing Science in the future.

1. Introduction
At present, some intelligent conversational agents were released to us, such as SIRI on IOS and Android voice. Talk to Siri as you would to a friend and it can help you get things done, like sending message, retrieve some information, monitor your mobile phone applications and talk to you alone. When you are driving on the road, you can let it tell you the best way to go home and where you are now. The more apple says you use Siri, the more you realize how great it is. As you know, machine is not only a tool, but also a virtual friend in our daily life.

Speech act theory was put forward by [1] Oxford philosopher J.L. Austin (how to do things with words, 1962) and further developed by American philosopher J.R. Searle. It is believed that speech act has three levels or components: speech act, extra speech act and post speech act. We use the term speech act to describe such acts as "request", "order", "question" or "inform". For the research of speech behavior model of embedded agent, Searle (1969) method will be adopted to support human agent communication with a simple speech behaviour [2]. Therefore, in the functional behavior model...
of intelligent conversation subject, speech behavior model is the primary challenge that directly affects
the quality and effectiveness of response.

Role and Reference Grammar (RRG) discusses the extensive grammatical phenomena and
language variations of simple and complex sentences. (Nolan 2013) [3][4]. When interpreting the
"special and linguistic features" of structure, it helps to understand the relevance of structural schema
(Van Valin 2005:132) [5]. The RRG view structure has two parts: (1) keep a grammar list of all
grammar templates; (2) a dictionary containing vocabulary entries and other categories. RRG structure
also includes four templates: morphological template, syntactic template, semantic template and
pragmatic template. It is also used for a grammar function and links these templates through a robust
linking algorithm.

Next are RRG model, intelligent session broker framework and FunGramkb (Nolan, mairal uson &
peronan 2009 [6][7], perinan Pascual & mairal uson 2009, perinan Pascual & arcas tunez 20072010,
mairal uson & Ruiz de Mendoza 2009). It will form a new language analysis system to help machines
recognize and analyze natural languages. In addition, in the complexity system of language, we also
need the functional motivation model grammar dictionary structure.

We plan to build a conversation subject model in a computational framework that uses speech acts
to collect human language from texts in the functional grammar model. In this study, the key points
are the sense of the listening subject and the recognizability of the logical language: (1) the speech act
of the embedded subject is conducive to independent and rapid communication; (2) the language
model adopted is role and Reference Grammar (RRG), which is used for the analysis and
understanding of sentences in semantic and contextual emotion.

2. Language analysis model based on RPG

Intelligent Session Agent will be based on.Net program model and built on the existing computing
research of RRG and FunGramKB. The project will provide (1) a discourse workspace model (Figure
1) based on speech act and a set of functional language models, and build a central language engine
to manage language processing; (2) a personality and emotional component model, and cooperative
activities with other conversation subjects.

![Figure 1. The architecture of Language space model](image-url)

Base on the main framework is a highly complex active workspace, it involves the logic and
algorithm research before RRG and FunGrammkb. FunGrammkb is a multi-functional [8],
multilingual natural language processing system (NLPs), which can be reused in machine translation,
information retrieval, conversation and other tasks. Therefore, we have built the underlying framework
structure based on the cloud platform, and built the basic service layer, engine model layer, application
service layer and paradigm framework layer on this basis.
2.1. **Basic service layer**

In order to make semantic understanding more extensible and replaceable, we separate each module into a series of micro service structures: message interaction service, acoustic processing service, memory network service, chain value data storage service, request processing service and data storage service.

- **Message exchange service**
  
  This model can withstand large-scale service pressure. By using an optimization method based on periodic message execution and category priority to define the queue, the system defines the interaction of messages within the cycle scope. Based on the analysis of user role, system and time operation behavior, the interaction is realized far by combining the defined message importance tag End message task.

- **Voice processing services**
  
  The service uses the deep speech model to extract the features of the input speech information [9]. At the same time, a multi-channel network model is designed for the acoustic processing task under the association of the memory network service. Each task can automatically select the channel related to the context, cooperate to complete the feature extraction and semantic analysis, and further improve the performance of speech recognition is discussed.

- **Memory network service**
  
  (Weston et al, 2015) it is a service based on a neural network model, which can conduct natural language reasoning by operating the content in memory (write, read, filter and reuse). Memory Web services store and recognize long-term memories such as common sense, names, and geographic locations in the real world. It can also effectively identify short-term contexts, such as session content over a period of validity.

- **Chain value data storage service**
  
  For critical information processing, a linear table is used, but it does not store data in a linear order, but a pointer to the next node in each node. It mainly guarantees that the gradient method of critical information strategy will not be difficult to control due to multi-dimensional continuous state or action.

- **Request processing service**
  
  Its purpose is to determine which external inputs need to be initiated by the system for automatic language processing and semantic analysis, and monitor the service to output the results to the external end. For example, a simple meaningless signal input cannot trigger the system to operate. At the same time, the service link memory network map synchronously sends the previous context or special definition information related to this request to the language processing Engine.

- **Data storage services**
  
  The distributed SAN storage system defined for the cloud environment adopts the full flash architecture and RDMA network design, which can reach millions of IOPs under sub millisecond delay, breaking the performance and capacity bottleneck of distributed storage. The service stores all processed language entities in the distributed database, and provides data fragmentation and multi copy strategy to ensure strong consistency among multi copies; adopting multi-path and redundancy mechanism, each component can independently complete fault detection, repair and isolation to ensure stable operation of the system.

2.2. **Engine model layer**

Because of the role of reference grammar as a language engine, that is, structuralism formalism grammar theory, the construction of the engine has an internal structure, including (1) a unique signature, (2) a set of constraints, (3) an input behavior, (4) an output behavior, (5) a construct and (6) a workspace. All these functional modules will be used to build a schematic structure in RRG. Some functional modules will adopt parallel multithreading technology to ensure the efficiency of language engine.
2.3. **Machine learning layer**

Human conversations often contain dozens to hundreds of rounds of persistent conversations, which usually have strong contextual and contextual properties, and have a certain rule structure. For example, in a debate, the opening reading is usually used to express the context basis of the whole dialogue. In the process of communication, different contextual positions and perspectives are set up, and the critical statements of the other party need to be grasped at all times for reverse analysis. Therefore, for human beings, they are very good at tracking and controlling the flow of information in a dialogue, and continuously and efficiently transforming thinking into text language output. But for the machine, it is very difficult to maintain the conversation, define the end and start the opportunity, understand the human thinking logic, judge the semantic scene, and finally output the feedback results of human real intention.

Therefore, the construction of the service layer is carried out around the above-mentioned problems, including persistent conversation, adversarial learning, information feedback, and generalized filtering. In previous studies, researchers used LSTM to solve the problem that context cannot be effectively used due to the inability to fix the length of the session and improve the ability of context understanding. However, as a variant of RNN, it is difficult to deal with the problem that long-term information caused by gradient disappearance is difficult to store. Therefore, in consideration of the actual need to repeatedly call context for combined understanding, we use a machine learning model based on convolutional cyclic neural network to process the mapping from the digital signal received from the external end to semantic analysis and context continuous value classification, so as to effectively predict the context conference entity and effective memory network Atlas stored in the chain data storage service, Thus, the problem of persistent session is solved.

At the same time, in the process of training and learning, by using the idea of Turing test, a dialogue training model is set up. One is a dialogue generator, which defines the input session segment. The segment is composed of historical and past session information and human input learning samples. The other is a session evaluator, which marks and evaluates the sentence value sent by the generator; In the semi supervised situation, this method can effectively identify and make feedback judgments in the face of most of the individual context conversations.

In addition, for information feedback and avoiding meaningless generalization output during conversation interaction, we set a filter to optimize the output of dependency features. By calculating the maximum function of the target, we model the single dependency of feedback semantics, and exclude all kinds of meaningless feedback, such as: what's your name? I don't know, I don't know. My name is John, etc. The model filters the output. If one statement is strongly related to the input problem, the other meaningless statements are excluded.

2.4. **Paradigm framework layer**

For the paradigm framework, it will be divided into four parts to help the agent analyze natural languages, including morpheme storage [10], lexicon, warehouse and grammar. This will be based on (1) [11] the early research (Nolan & Salem 2011), the machine translation engine uniarab based on RRG (syntax to semantics, and vice versa), and (2) the recent research results, that is, understanding the structure of RRG as a syntax object, and the role of computing methods in Functional Grammar. The linguistic framework will build a grammar dictionary database, connect with the intelligent conversational language workspace and analysis results feedback information to the session manager.

3. **Semantic emotion analysis**

There are different between Austin’s and Searle’s speech act theory. Not only Searle’s theory has the characteristics of systemization and standardization, but he also thinks the speaker talks something which are implementing the speech act. Man’s use of language is an intentioned act conditioned by rules. And the basic unit of speech communication is not words or sentences but speech acts as they express the speaker’s intention. Also, the speech act is a function of sentence meaning. The construction of Searle’s speech act theory distinguish philosophy of linguistic and philosophy of
language in concepts, he thought that the philosophy of linguistic is trying to focus on individual specific words or other elements in the language of daily to solve the concrete of philosophy problem; but the philosophy of language is trying to give clear description with some common features of languages”.

Therefore, according to RPG, we have three aspects in the deep processing of conversational behavior: (1) conversational behavior refers to the behavior of making meaningful utterances in a certain context and reference; (2) conversational behavior refers to the way in which sentences are used to express states with certain functions; (3) post speech behavior is a kind of behavior or psychological state caused by speech. In order to make a clear definition of sentences, we also use several predicates to support the conversational behavior model: (a) believe which agent should trust the command of the person at any time, (b) know and represent the knowledge state of the agent to the person, (c) want, the agent wants to serve and wait for the order of the person, (d) intention, which means the agent wants to serve the person. In addition, the session behavior model will be used to establish a session behavior communication layer, which is the first interface within the session broker.

In addition, there is a framework of influencing subjects, which is also divided into four parts, including personality and emotion, task reasoning, world state reasoning and social situation. All of them are not in line with the linguistic structure and are self-defined by the subject. In addition, personality patterns affect conversational behavior and language relevance. In this paper, we will use a special processing method to carry out semantic emotional analysis of the above states, according to the organization and representation knowledge in the concept map model [12], we will use unsupervised machine learning method to study the collected unstructured data content, Then, the dynamic recognition function of text semantic emotion is realized by text analyzer, and the corresponding learning feedback results are output, so as to achieve the recognition of the influence on the subject emotion. The specific calculation method is as follows:

$$P[w \text{ in sentence } s(C_a)] = \beta p(w) + (1 - \alpha) \frac{\text{TimeL}(C_a, V_w)}{Z_{C_a}}$$  (1)

According to the formula of Semantic Smoothing, $\alpha$ and $\beta$ are constant weight parameters, $\text{TimeL}(C_a, V_w)$ is obtained by evaluating the context state, semantic emotion and social background in the same period, and then calculating the frequency of the word appearing in the article and the effective semantic vector value $\beta p(w)$.

In order to make the machine better understand and convert to vector value, the following calculation method is adopted, which will give the initial vector value to the item according to the cycle search heat and current semantics to ensure that it is within the controllable range of machine understanding.

$$\text{Ass}(w \text{ in document}) = \sum_x \exp (P_{wT}f_t(\text{max}_{T>0} \text{Proc}_y))$$  (2)

In the agent service session manager, the process is divided into four stages according to language analysis, machine transformation, environment factors and final output results. In the final output, we also fuse the results of each semantic vector and the output results of emotion vector in series through the bidirectional LSTM layer of neural network. After the output results of LSTM network are mixed in series for many times, we carry out the final emotion evaluation. See the figure below for details:

Figure 2. Design of series fusion model
4. Results and Analysis

In order to verify the validity and reliability of this study, the validity of this study will be evaluated by the following factors: (1) the effectiveness of natural language collection; (2) the effectiveness of natural language analysis; (3) the effectiveness of instructions and responses; (4) external environmental factors. Therefore, this experiment will be verified by four aspects: feature word extraction, emotion word extraction, association combination extraction and overall theme emotion analysis.

The comparison of the accuracy and recall rate of the above key parts extraction calculation with traditional SVM and CNN is shown as follows, in which the abscissa is the number of theme (T) and the ordinate is the accuracy (P).

![Figure 3. Trend chart of accuracy and number of themes](image)

It can be seen from the figure 2 that SVM and CNN model are at the same level when the topic is initialized, but the intelligent session agent model we build can filter and filter the external environment and mutual relevance once, making the average value higher than the first two models. When the number of topics is greater than 50, SVM is difficult to capture those low-frequency characteristic words with practical significance, which leads to The decrease and increase of the accuracy rate are obviously slower than that of CNN and our mixed periodic calculation model; In order to the new model adds time period and external impact assessment, it has a global impact on the topic selection of context words. Secondly, it can ensure the integrity of those low-frequency emotional words for the topic conversation in the later stage through the analysis of relevance. Finally, it can make the accuracy of machine recognition topic conversation increase rapidly, which has more advantages than the analysis results of SVM and CNN.

5. Conclusion

This kind of intelligent conversation agent can be regarded as a new human brain. They know how to act, how to think, how to understand emotions and how to speak. Under the control of human, the agent will greatly reduce the workload and risk of human. This new technology will bring a new two-way dialogue robot. Therefore, these "brains" will help people to complete certain tasks and understand the meaning of human beings.

From the perspective of the development of information and communication technology, many artificial intelligence technologies have been at the forefront of innovation, which will greatly affect the development direction of the information industry in the future. Artificial intelligence involves many research fields, including symbol computing, language recognition, pattern recognition and computer vision, as well as machine learning, intelligent information retrieval, machine translation problem solving and expert system, logical reasoning and logical proof, natural language processing, etc., which gradually become a more extensive information science. Our research also extends the depth of AI and language analysis, including RRG model, intelligent session agent, FunGrammkb, etc.
As the particularity of our study, the conversational subject will respond with complex social interaction. The framework of agent language model not only needs language knowledge, lexical grammar semantic constructor, role grammar and reference grammar engine, but also needs to consider the intention structure of agent, which is composed of belief, desire and intention expressed by discourse elements. Other dialogue actors and personality and emotion models will be modified as needed. Because this kind of interaction technology realizes the state of agent with its own feeling and emotion, the external environment will also affect the judgment and result of the machine, so different agents will produce various responses according to different current state, which is an important development and innovation research in the field of artificial intelligence.

Acknowledgment
I would like to take this opportunity to express my gratitude to Dr. Qiao Dexin, who is the initiator of the project, for all his efforts in the process of writing my thesis, including helping me to select topics, decide research questions and topic lines, revise our project through face-to-face guidance, and make critical and meticulous corrections and effective suggestions in the whole process of project implementation.

Finally, I would like to thank my colleagues for their unfailing support and patience in answering my questions in the interviews during my thesis writing.

REFERENCES
[1] J.L. Austin, J. O. Urmson, Marina Sbasà. (1962) How to Do Things With Words, Second Edition. Harvard University Press, Cambridge Massachusetts.
[2] Searle, John R. (1969) Speech Acts. Cambridge University Press.
[3] Nolan, Brain. (2012) Constructions as grammatical objects: A new perspective on constructions in RRG. Paper presented at the 1st International Conference on Knowledge Representation and Meaning in FunGramKB and LCM. Facultad de Filologia, at the Universidad Nacional de Educacion a Distancia(UNED), in Madrid, Spain.
[4] Nolan, Brian & Diedrichsen, Elke. (2013) Linking Constructions into Functional Linguistics. The Role of Constructions in Grammar[Studies in Language Companion Series 145]. Amsterdam: John Benjamins.
[5] Van Valin, Robert D. (2005) Exploring the Syntax-Semantics Interface. Cambridge: CPU.
[6] Nolan, Brian, Mairal-Uson, Ricardo & Periñán, Carlos. (2009) Natural language processing applications in an RRG Framework. In Proceeding of 10th International Role and Reference Grammar Conference. University of California, Berkeley USA.
[7] Perinan-Pascual, Carlos & Arcas Tunez, Francisco. (2010) The architecture of FunGramKB. In Proceedings of the Seventh International Conference on Language Resources and Evaluation, European Language Resources Association (ELRA).
[8] Nolan, Brian & Carlos Periñán-Pascual. (2014) Language Processing and Grammars - The role of functionally oriented computational models, John Benjamins Publishing Company.
[9] J. Nilsson, Nils n. (2014) Principles of artificial intelligence. Stanford University. Morgan Kaufmann Publishers, Inc.
[10] Van Valin, Robert D., Jr. (2004) Lexicon Representation, Co-composition, and Linking Syntax and Semantics.
[11] Nolan, Brian & Salem, Yasser. (2011) UniArab: RRG Arabic-to-English Machine Translation software. Based on Role and Reference Grammar. Institute of Technology Blanchardstown, Dublin, Ireland.
[12] Kunying Li, Xiaolian Li, Yu Ding. (2019) Message Queue Optimization Model Based on Periodic Execution and Category Priority. CIIS 2019.