Communication: Words and Conceptual Systems

Jian YU
Beijing Key Lab of Traffic Data Analysis and Mining
Beijing Jiaotong University, Beijing, China
Email: jianyu@bjtu.edu.cn

Abstract

Words or phrases play a key role in natural language. Word (or phrase) representation is the fundamental problem for knowledge representation and understanding. However, it is always a challenge that how to represent a word (or phrase) can make it easily understood. In this paper, a new representation for a category is discussed, which defines the inner name, the outer name, inner referring, outer referring and so on. In order to reduce the complexity of category representation, the economy principle of category representation is proposed. Moreover, the conceptual system and communication are also discussed. The conditions for perfect communication are discovered.

Keywords: Words, Conceptual System, Communication

1 Introduction

When studying objects in the real or virtual world, categorization is a common approach in order to communicate with each other. A word or phrase represents the corresponding category name. Every minute in daily life, we rely on words to help us to deal with everything, such as talking, listening, watching, thinking, etc. How to represent a word (phrase) become a pivotal question to understand each other. In the literature, a classical category has three representation: the corresponding name representation, the corresponding set representation, the corresponding proposition representation. In common sense, a word or phrase represent the name of a category, which is called the concept name or the category name. Set representation represents a class of objects in the world belong to the corresponding category, proposition representation represents the mental representation of the corresponding category.

Up to now, all the above three representations seem in question. Firstly, a category usually has different names in different languages. What’s more, a category has different names in the same language. For example, when seeing a specific dog, we can say it ”a dog” in English, ”” in Chinese. Clearly, the word ‘dog’ is different from that specific dog in the world, which only is the name of the category ”dog”. Hence, the name of a category cannot fully reflect the meaning of this category. In extremity, the name of a category is only considered as a symbol and has no meaning. Secondly, a category may be very fuzzy or vague such as beauty, badness and so on. Clearly, a classical set can not express this case. More surprised, Watanabe and Donovan [1969] has stated that it is impossible distinguishing a category from others without feature selection. Thirdly, Wittgenstein [1953] has claimed that important categories such as games can be defined (represented) by a proposition. Last but not least, it is taken for granted that all the men should have the same representation for a classical category. What a pity, such an assumption is also clearly not correct, otherwise, misunderstanding about the same category cannot occur among men in daily life.

Although there exist the above drawbacks, man still can use words (phrases) to express their thoughts, feelings and observations. Why? In this paper, we devote to answering this question and make three major contributions as follows.

1) Semantic set is proposed to represent a category in Section 2.
2) Conceptual system is discussed in Section 3.
3) How to communicate concepts between people is discussed in Section 4.
2 Category and Semantic Set

Let the discussed object domain be a set $O$ and $o$ represent an object in $O$, the discussed category domain be a set $L$ and $\mathfrak{A}$ denotes a category in $L$. For different categories in $L$, $o$ may have different representations. For example, when categorization by obesity, a man can be represented by his age, his weight, his height and so on; when categorization of different types of individuals, a man can be represented by his sensation, his intuition, his feeling, his thinking, his attitudes and so on. Transparently, the object representation depends on categorization. Ugly Duckling Theorem [Watanabe and Donovan, 1969] has shown that a category specializes its object representation, otherwise, objects cannot be distinguished by different categories. In this paper, when specific categorization is needed, the corresponding object representation is supposed to have been obtained. Therefore, let $\mathfrak{A}$ be the category name we will discuss, $O_A = \{o_A | o \in O\}$ where $o_A$ is the object representation of the object $o$ in the set $O$ with respect to the category $\mathfrak{A}$, $I_A$ is the corresponding membership function defined as follows.

Membership function:
$I_A: O_A \rightarrow R_+$ is called the membership function of $\mathfrak{A}$ if $\forall o_A \in O_A$, the greater $I_A(o_A)$ means the more probability $o$ belongs to the category $\mathfrak{A}$.

As $O_A$ and $I_A$ are observable, $\{O_A, I_A\}$ is called the outer set of $\mathfrak{A}$. If $\forall o \in O$, $I_A(o_A) \in \{0, 1\}$, then it is easy to know that $\forall o$, $I_A(o_A) = 1$ if and only if $o$ belongs to the category $\mathfrak{A}$, and $I_A(o_A) = 0$ if and only if $o$ belong to the category $\mathfrak{A}$. In this case, $\{O_A, I_A\}$ represents a classic set. If $\forall o \in O$, $I_A(o_A) \in [0, 1]$, then $\{O_A, I_A\}$ represents a fuzzy set [Zadeh, 1965].

Let $\overline{A}$ represent the concept corresponding to the category $\mathfrak{A}$ and the name of $\overline{A}$ be $\mathfrak{A}_3$. When the concept representation for any category is defined, objects can be categorized based on the similarity between objects and the concept representation. The category similarity mapping can be defined by computing the similarity between objects and the concept representation as follows.

Category Similarity Mapping:
$Sim_A: O_A \times \overline{A} \rightarrow R_+$ is called the category similarity mapping of $\mathfrak{A}$ if $\forall o \in O$, the increase of $Sim_A(o_A, \overline{A})$ means the greater similarity between the object $o$ and the category $\mathfrak{A}$, where $o_A$ is the object representation with respect to $\mathfrak{A}$.

Similarly, the category dissimilarity mapping can be defined as follows:

Category Dissimilarity Mapping:
$Ds_A: O_A \times \overline{A} \rightarrow R_+$ is called the category dissimilarity mapping of $\mathfrak{A}$ if $\forall o \in O$, the decrease of $Ds_A(o_A, \overline{A})$ means the greater similarity between the object $o$ and the category $\mathfrak{A}$, where $o_A$ is the object representation with respect to $\mathfrak{A}$.

Generally speaking, $(\overline{A}, Sim_A)$ or $(\overline{A}, Ds_A)$ can be unobservable, therefore, $(\overline{A}, Sim_A)$ or $(\overline{A}, Ds_A)$ is called the inner set of the category $\mathfrak{A}$, its name is called $\mathfrak{A}_3$.

For brevity, $(\mathfrak{A}, O_A, I_A)$ is called the outer representation of the category $\mathfrak{A}$, $(\mathfrak{A}_3, \overline{A}, Sim_A)$ is called the inner representation of the category $\mathfrak{A}$. In language, $\mathfrak{A}$ is the outer name of the category, $\overline{A}$ is the mental representation of $\mathfrak{A}$, $\mathfrak{A}_3$ is the inner name of of the category. Usually, $\mathfrak{A}$ corresponds to a word or (phrase) in a specific language, $\mathfrak{A}_3$ also corresponds to a word or (phrase) in a specific language. In summary, a category $\mathfrak{A}$ can be represented by a six tuple: $(\mathfrak{A}, O_A, I_A, \mathfrak{A}_3, \overline{A}, Sim_A)$ or $(\mathfrak{A}, O_A, I_A, \mathfrak{A}_3, \overline{A}, Ds_A)$. $(\mathfrak{A}, O_A, I_A, \mathfrak{A}_3, \overline{A}, Sim_A)$ or $(\mathfrak{A}, O_A, I_A, \mathfrak{A}_3, \overline{A}, Ds_A)$ is called a semantic set.

In daily life, a category $\mathfrak{A}$ is often represented by its name $\mathfrak{A}$. However, only the outer name of the category is known, it will be a puzzle for us to understand this category. Many words (phrases) in ancient classics cannot be understood by us just because their corresponding outer sets and inner sets have been lost in the long river of history.

3 Conceptual System

Every man has at least one conceptual system, which includes many categories. For the man $\alpha$, one of his conceptual system can be defined as $L^\alpha = (\mathfrak{A}^\alpha)$, where the category $\mathfrak{A}^\alpha$ can be represented by $(\overline{A}^\alpha, O_A^\alpha, I_A^\alpha, \mathfrak{A}_3^\alpha, \overline{A}_3^\alpha, Sim_{\overline{A}^\alpha})$. $\mathfrak{A}^\alpha$ is a word or phrase, $(O_A^\alpha, I_A^\alpha)$ represents the referring action with respect to $\mathfrak{A}^\alpha$, $(\overline{A}_3^\alpha, \overline{A}^\alpha, Sim_{\overline{A}^\alpha})$ represents the idea. If the above three parts of the category $\mathfrak{A}^\alpha$ are not contradictory, then $\mathfrak{A}^\alpha$ is self-consistent. How to say a category $\mathfrak{A}^\alpha$ is self-consistent?

In order to answer this question, let us define the inner referring operator as: $\overline{A}_3^\alpha = \arg\max_{\overline{A}_3} Sim_{\overline{A}^\alpha}(o_B^\alpha, \overline{A}_3^\alpha)$, where $o_B^\alpha \in L$; and the outer referring operator $\rightarrow o_B^\alpha$ as: $\overline{A}_3^\alpha = \arg\max_{\overline{A}_3} I_B(o_B^\alpha)$, where $o_B^\alpha \in \overline{A}_3^\alpha$. It is easy to know that $A_3^\alpha = \{o \in O | o \text{ is } \mathfrak{A}^\alpha\}$, $A^\alpha_3 = \{o | o \text{ is } \mathfrak{A}^\alpha\}$. If $o \in A^\alpha_3$, then $o$ can be externally called $\mathfrak{A}^\alpha$ in other words, $o$ has an outer name $\mathfrak{A}^\alpha$. If $o \in A_3^\alpha$, then $o$ can be internally called $\mathfrak{A}^\alpha_3$, in other words, $o$ has an inner name $\mathfrak{A}^\alpha_3$. Obviously, if $A_3^\alpha = A^\alpha_3$ and $\mathfrak{A}^\alpha = \mathfrak{A}^\alpha_3$, then $\mathfrak{A}^\alpha$ is self-consistent.

Self-consistent is very helpful to reduce the complexity of a semantic set. If a semantic set is self-consistent, it is the simplest because its inner representation can
be ignored in some sense. Therefore, self-consistent is the economy assumption for category representation. However, self-consistent is not always true for a conceptual system. In general cases, there exists the mapping between the inner name and the outer name as follows.

**Name Encoding Mapping:**
\( N^e_\alpha : \mathcal{A}^\alpha \rightarrow \mathcal{A}^\beta \) is called the name encoding mapping of \( L^\alpha \), where \( \forall \mathcal{A}^\alpha \in L^\alpha \).

**Name Decoding Mapping:**
\( N^d_\alpha : \mathcal{A}^\beta \rightarrow \mathcal{A}^\alpha \) is called the name decoding mapping of \( L^\alpha \), where \( \forall \mathcal{A}^\alpha \in L^\alpha \).

According to the above analysis, a conceptual system can be accurately expressed as \( L^\alpha_{(N^e_\alpha,N^d_\alpha)} = \{\mathcal{A}^\alpha\} \). For a conceptual system \( L^\alpha_{(N^e_\alpha,N^d_\alpha)} \), when its name encoding mapping and its name decoding mapping are an identity function, it can be simply denoted by \( L^\alpha \), such a conceptual system is called a plain conceptual system. For a conceptual system \( L^\alpha_{(N^e_\alpha,N^d_\alpha)} \), if its name encoding mapping and its name decoding mapping are not identity function, a category cannot be guaranteed to be self-consistent. In this case, it is more difficult to understand the conceptual system \( L^\alpha_{(N^e_\alpha,N^d_\alpha)} \) than to understand the conceptual system \( L^\alpha \).

In order to reduce the difficulty of understanding, it is natural to require that the name encoding mapping and the name decoding mapping are an identity function and all categories are self-consistent for a conceptual system. If a conceptual system satisfies the above conditions, it is called a self-consistent conceptual system. Moreover, if any concept can be defined by a proposition in a self-consistent conceptual system, such a conceptual system is called an ideal conceptual system. Certainly, an ideal conceptual system can be accurately expressed by natural language. If a conceptual system is not ideal, it cannot be accurately expressed by an natural language. Therefore, an ideal conceptual system is the most easily understood and inherited. However, it is the most difficult for man to obtain an ideal conceptual system. In general, man only can continuously approximate an ideal conceptual system through making \( \mathcal{A} \) be expressed by some proposition as accurately and precisely as possible for any category \( \mathcal{A} \) in his conceptual system.

## 4 Communication

More frustrated, even when \( \mathcal{A}^\alpha \) is self-consistent, \( \mathcal{A}^\alpha \) may be something wrong. Why? As the final goal of a category is to help communication, different personal conceptual systems concerned on the category \( \mathcal{A} \) may have different representations, which results in misunderstanding, even contradiction. 

For a communication between two men \( \alpha \) and \( \beta \) concerned on the category \( \mathcal{A} \), two simple cases about the category \( \mathcal{A} \) are discussed as follows.

One case is that only one man knows the category \( \mathcal{A} \). Without loss of generality, it is assumed that the man \( \alpha \) knows \( \mathcal{A} \), i.e. he knows \( \mathcal{A}^\alpha \) and the man \( \beta \) has no idea of \( \mathcal{A} \), then the man \( \beta \) must learn the knowledge about the category \( \mathcal{A} \) from \( \mathcal{A}^\alpha \). 

In order to simplify learning, it is supposed that the outer representation of \( \mathcal{A}^\alpha \) is known and all the relevant conceptual systems are plain. Under such assumptions, the man \( \beta \) needs to learn how to repre-
\( \alpha \to o \beta \), then it is a proper communication between the man \( \alpha \) and the man \( \beta \) concerned on the category \( A \) and the object \( o \). Transparently, a semi perfect communication on the category \( A \) can lead to misunderstanding in some cases. When communication is not perfect, misunderstanding can lead to quarrel, joke, tragedy and so on.

Sometimes, only \( A^\alpha = A^\beta \) or \( o^\alpha = o^\beta \) is true in communications. In this case, misunderstanding often occurs if one side thinks communication is right.

Furthermore, when \( o^\alpha A \) or \( o^\beta A \) is only shown although both sides know the category \( A \), communications still carry on by finding the category corresponding to \( o^\alpha A \) or \( o^\beta A \). Certainly, such a case can bring more misunderstanding and more challenge in communication.

5 Discussion and Conclusions

When ignoring category name, Yu (2015) has presented a category representation. However, everything has a name, as Cassirer (1944) pointed out. In this paper, a new approach to representing words (or phrases) is given by considering the category name. This proposed category representation solves the drawbacks of the classical category representation.

Based on this representation, we discuss the conceptual systems and communications. Without the category name, semantic set is the same as the category representation proposed by Yu (2015). Hence, it is very natural that self-consistent is another version of categorization equivalency axiom proposed by Yu (2015) and perfect communication reinterprets the uniqueness axiom of category representation. For communication, self-consistent is the economy principle, which reduces the complexity of understanding. Therefore, self-consistent can be called the economy assumption of category representation. When self-consistent is false, it is worth further investigation how to construct name decoding mapping and name encoding mapping in a conceptual system, which involves planning, reasoning and decision.

In this paper, all the analysis ignores that the influence of \( L^\alpha_{\{N^\alpha_0, N_0^\alpha\}} \) is different from that of \( L^\beta_{\{N^\beta_0, N_0^\beta\}} \). Considering different conceptual systems have different influences, more complexity will be introduced in communications. The corresponding research results can be generalized into any two agents with their own conceptual systems. Such two agents can be any two sides, maybe two men, two robots, one robot and one man, one man and one book, and so on.

In particular, word (or phrase) representation has three parts: one is about linguistic, one is about action, one is about idea. In daily life, all three parts can independently form one conceptual system, i.e. say one thing, do another, think differently from saying and doing. For instance, some men can smartly construct temporary conceptual systems to fit the local environmental requirements, such as actors, pretenders, spies or translators. When a man has several conceptual systems, it is worth studying how to use several conceptual systems in a smooth way.

Furthermore, only two simple cases about communications are discussed in this paper. More complex cases such as both communication sides only know something about \( A \) need to further study.

Acknowledgements

This work was partially supported by the NSFC grant (61370129).

References

Cassirer, E. (1944). An essay on man. Yale University Press.

Grice, P. (1975). Logic and conversation. In P.Cole and J. Morgan eds. Syntax and Semantics, vol.3, New York. Academic Press.

Watanabe, S. and Donovan, S. K. (1969). Knowing and guessing: A quantitative study of inference and information. Wiley New York.

Wittgenstein, L. (1953). Philosophical investigations. Basil Blackwell Ltd.

Yu, J. (2015). Generalized categorization axioms. *eprint arXiv:1503.09082*.

Yu, J. and Xu, Z. (2014). Categorization axioms. *eprint arXiv:1403.2065*.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.