The Analysis of Uncertain Knowledge Based on Meaning of Information

PING CANG, SUFEN WANG
Glorious Sun School of Business and Management
Doghua University
Shanghai China
http://www.dhu.edu.cn

Abstract: - The paper discusses four types of knowledge concepts, especially three kinds of uncertain knowledge, based on relationships among data, information and meaning. The goal is to fuse research results on knowledge engineering and epistemology. We analyze four kinds of basic connection forms (i.e. meaning) between information and dada by using formal context, and obtain their physical signification. Following Dretske's semantic information theory, we analyze information entropy of the four types of meaning. Then using the information entropy obtained, we look at how data convey information content, and how the types of meaning relate to different concepts in knowledge. Finally, we realize that uncertain knowledge is involved in all the above, therefore we further express the four types of meaning (i.e., four types of connections between information and data) by using Rough Set, and which deepens our understanding of uncertain knowledge.

Key-Words: - data, information, meaning, uncertain knowledge

1 Introduction

Knowledge is probably the most important subject in the research of knowledge engineering and epistemology, which have many research outcomes in each field[1][2]. It is a very meaningful theme to research how to connect the outcomes in the two fields, taking use of the outcomes of knowledge epistemology to guide the knowledge representation and reasoning of knowledge engineering, making use of the relevant research in knowledge engineering to realize the research results of knowledge epistemology[3][4]. In this paper, we try to connect the research results of knowledge engineering and epistemology by the discussion of relevant information, data and meaning in information system field to enrich the in-depth understanding of uncertain knowledge.

The following content is organized in this way: The second part analyzes the inevitable existence of the relationship between data and information meaning from the development clue of modern epistemology; The third part analyzes the four types of connections of information meaning based on data; In the fourth part, we analyze the relationships between the four types information meaning and the four types of knowledge concept by Dretske's semantic information theory, then we analyze the relationships of the four types knowledge concept in the Rough Set ; The last part is conclusion.

2 Analysis of Relationship Between knowledge and Information Meaning

2.1 What is the Meaning of Information Based on Data?

Data, information, meaning are the very important concepts in information system field all the time, while the mess phenomena exists among these concepts ([5],[6],[7]). But one thing is sure, that is there are certain relationships between these three ones. Wang et al follow the viewpoint of information philosophy ([8]), 'information can be called information' is just because we understand from the relationships between the three concepts([9]). And this relationship can only be experienced in the process of communication among people. In the repeat circulation of 'how meaning come out from information which based on data', the communication among people was finished.

Information meaning is realized by an explanation process or an explanation active. The subject of explanation process or active is human being; meaning can only be existing and valuable relative to human being. The explanation process of meaning is the realization process of data, which is the process about how people make use of

information and how to serve the people who need information. Wang considers that Heidegger's research of existance analyzes the structure and essential characteristics of interpreter which is the realization of the information meaning; find out that the basic form of information meaning is the link between information and data([10]).

If information includes the variety possibility of people's living, then the meaning of information is a specific possibility of people's survive. This possibility can be represented by data. The realization of information meaning reflects the interpreter's ability of understanding and explanation; it also reflects the horizon of the interpreter.

Information is objective, but the understanding of objective information is subjective. And this objective-subjective relationship is the belong-to-relationship of Gadamer. Information meaning reflects that this objective-subjective relationship is the belong-to-relationship, it is the link of information and data. So, the full name of information meaning should be information meaning based on data; in this paper it is short for information meaning.

2.2 The Relationship Between Knowledge and Information Meaning

To understand the relationship between information meaning and knowledge, we must first understand what knowledge is. Knowledge is used as a modern term, is one of most confusing terms. The most classic concept of the knowledge is: the verified true belief is knowledge. Gettier's counterexamples - "is the verified true belief knowledge?"- break the classic fairy tale; A new round in the research of knowledge theory started around Gettier's counterexample. Dretske believed that there are some thing missing between " the verified true beliefs" and "knowledge"[11], the missing one is information provision([12], pp.10). Dretske suggested that "to know" is naturally the information provision ability [7]. Here "information" indicates some information content which reflect some thing or matter, the so-called "information provision ability" refers the ability of data to manifest the information content, and data is "the symbols which carrying information of status of a certain event ".So, Dretske confirm there exist the relationships between the knowledge and data-information on the perspective of epistemology.

However, Dretske's, Mingers', Floridi's studies and so on, have mentioned different concepts about data, information, meaning and the information content, but the meaning of the information has not been clearly explained. Mingers consider that the transformation between meaning and information is done through a similar digitalization of the analogue [13]. From his paper we can see that in his opinion meaning is a subset of information and this subset can be obtained by people. But in Dretske's "information provision", the term of information meaning was not directly proposed, only the "information content" was emphasized; Floridi emphasized that in the definition of information, "information must be meaningful[8]. " Synthesizing the views of these researchers and the philosophical foundation of information's viewpoint [14], we propose that: Although things are objective, understanding things is subjective. The meaning of information based on data is just a representation and reflection of the essential integration of objectivity and subjectivity in people's lives. The meaning of information based on data is not all the subsets of information content relevant to information, explicitly, it is a gerund rather than a noun, and it's an exhibition or a frame in which we could catch sight of partial information from the data. And this partial information is information content. Thereby, people comprehend the world, themselves and acquire knowledge.

The meaning of information based on data is obtained through an explanation process; the process involves the relationship among data, information, knowledge and meaning. People explain information based on data, and which impact people's ability of getting knowledge. Therefore, people will have different concepts of knowledge according to the types of meaning (i.e. different types of connections between information and data).

3 The Analysis of Different Connection of Information Meaning Based on Data

3.1 Four Connections of Information Meaning Based on Data

The meaning of information is the connection between data and information, which have many different patterns.

We use a formal context (D, C, I) to denote relationship among data, information and meaning[15][16]. D is a set of data items, C is a set of information contents, and

$$I \subseteq D \times C$$

is a binary relation D and C. The binary relation reflects Galois link[10][11].So, we can define the

image of a certain data subset and the counter image of information content subset.

Definition 1: For a subset of data $A \subseteq D$, we define

$$A' := \{c \in C \mid (d,c) \in I \ \ for \ \ all \ d \in A\}$$

A′ denotes a subset of information contents C carried or displayed by all data items in A.

Definition 2: for a subset of information content $B \subseteq C$, we define

$$B' := \{d \in D \mid (d,c) \in I \ \ for \ \ all \ c \in B\}$$

B′ denotes a subset of data items D carrying or displaying by all information contents in B.

A pair <A, B> is called the meaning of information based data, there will be four possible patterns to display subset of data (A) through subset of data(B) :

1） $B = A'$ and $A = B'$

2） $B = A'$ and $A \subset B'$

3） $B \subset A'$ and $A = B'$,

4） $B \subset A'$ and $A \subset B'$

Definition 3: For the meaning of information <A, B>, if $B = A'$ and $A = B'$, then the meaning of information <A, B> is called fully matched.

Definition 4: For the meaning of information <A, B>, if $B = A'$ and $A \subset B'$, then the meaning of information is called the first partly matched.

Definition 5: For the meaning of information <A, B>, if $B \subset A'$ and $A = B'$, then the meaning of information is called the second partly matched

Definition 6: For the meaning of information <A, B>, if $B \subset A'$ and $A \subset B'$, then the meaning of information is called the unmatched.

3.2 Physical Signification Analysis of Four Types Information Meaning Connections

3.2.1 Analysis of the Fully Matched Form

Considering the meaning of information based on data is fully matched, then:

$B = A'$ and $A = B'$.

The form in the formal context is show in Fig 1.

This form is the formal meaning which defined in Chapter 3. Its physical signification of the form is that: all the data items together in A can and only can manifest the information content of subset B; meanwhile, subset B also can and only can be manifested by all data items together in subset A.

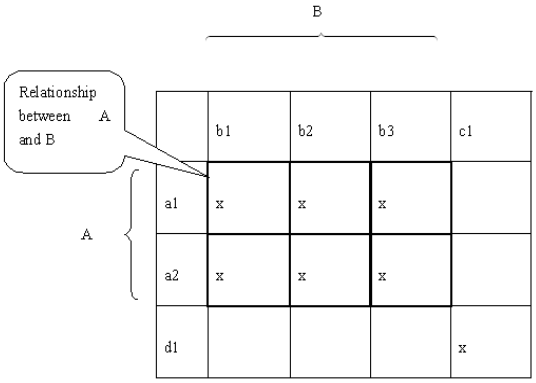

Fig.1 Fully matched form in the formal context

3.2.2 Analysis of the first partly matched form

Considering the information meaning of data is the first partly matched, then:

$B = A'$

And

$A \subset B'$

i.e.

$\exists d1 \in D, \quad d1 \notin A, \text{but} \{d1\} \subset B'$

The form in the formal context is show in Fig 2.

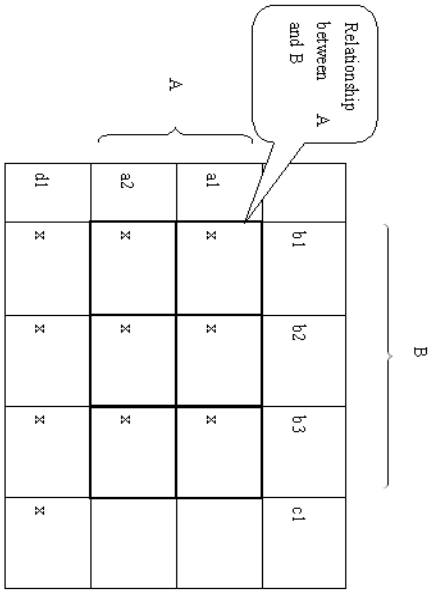

Fig.2 First partly matched form in the formal context

The physical signification of the form is that: all the data items together in A can and only can manifest the information content of subset B; But information content subset B can be manifested by much more items together. That is there is at least existing one item d1 not belonging to A and still can manifest subset B.

3.2.3 Analysis of the second partly matched form
Considering the information meaning of data is the second partly matched, then:

$B \subset A'$ and $A = B'$, i.e.

$\exists c1 \in C, \quad c1 \notin B, \text{but} \{c1\} \subset A'$

The form in the formal context is show in Fig 3.

The physical signification of the form is that: the information content of subset B can and only can be manifested by all the data items together in data subset A; But the information content manifested by data subset A is larger than information content subset B. That is there is at least existing one item c1 not belonging to B and still can be manifested by all the data items together in subset A.

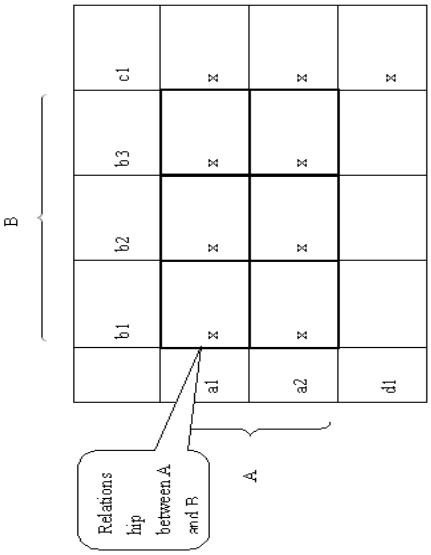

Fig.3 Second partly matched form in the formal context

3.2.4 Analysis of Unmatched Form
Considering the information meaning of data is the unmatched, then:

$B \subset A'$ and $A \subset B'$, i.e.

$\exists d1 \in D, \quad d1 \notin A, \text{but} \{d1\} \subset B'$

$\exists c1 \in C, \quad c1 \notin B, \text{but} \{c1\} \subset A'$

The form in the formal context is show in Fig 4.

The physical signification of the form is that: the information content of subset B can be manifested by all the data items together in data subset A; But the information content manifested by data subset A is larger and information content subset B can be manifested by all the data items of a larger data subset. That is there is at least existing one item c1 not belonging to B and still can be manifested by all the data items together in subset A. And there is at least existing one item d1 not belonging to A and still can manifest subset B.

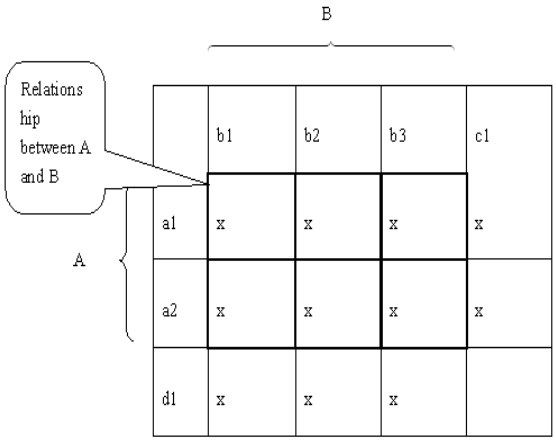

Fig. 4 Unmatched form in the formal context

3. Analysis of the Four Types Knowledge based on Information Meaning

Dretske suggested that there is tight relationship between knowledge concept and information provision which is the provision of information content. So, in this chapter Following Dretske, we analyze information entropy of the four types of meaning; on the other side, using the information entropy obtained, we look at how data convey information content, and how the types of meaning relate to different concepts in knowledge.

3.1 Dretske's Semantic Information Theory

By recomposing traditional information theory, Dretske tried to measure the amount of information involved in specific communication action between sender and receiver([11],pp.3-41).

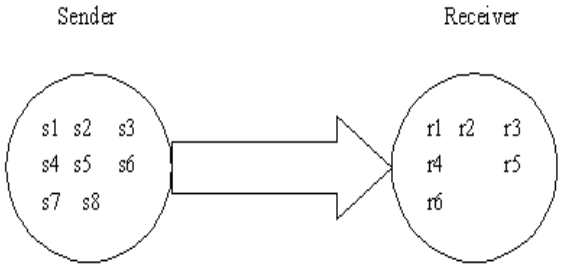

Fig 5 sender-receiver

Sender: There are some events s_1, s_2, ……,s_m may happen, the amount of information associated with, or generated by a particular event is:

$$I(s_i) = \log 1/p(s_i) = -\log p(s_i)$$

（ here, $p(s_i)$ is the probability of s_i, i=1,2,…,m ）

The average amount of information generated by S:

$$I(S) = \sum_{i=1}^{m} p(s_i) * \log(1 / p(s_i))$$

Receiver : There are some events r_1, r_2, ……,r_n may happen, the amount of information associated with particular message that was actually received is:

$$I(r_j) = \log 1 / p(r_j) = -\log p(r_j)$$

（here $p(r_j)$ is the probability of r_j, j=1,2,…,n ）

The average amount of information generated by R:

$$I(R) = \sum_{j=1}^{n} p(r_j) * \log(1 / p(r_j))$$

$Is_i(r_j)$ is a measure of the information associated with particular message r_j in situation r about a particular event s_i in situation s.

Dretesk suggested:

$$I_{s_i}(r_j) = I(r_j) - N(s_i)$$

Here: $N(s_i)$ is noise

$$N(s_i) = -\sum_{j=1}^{n} p(r_j / s_i) * \log(p(r_j / s_i)$$

$$I_{s_i}(r_j) = I(s_i) - E(r_j)$$

Here: $E(r_j)$ is equivocation

$$E(r_j) = -\sum_{i=1}^{m} p(s_i / r_j) * \log(p(s_i / r_j)$$

The following analyze data item manifesting information content and information content represented by data item in four different type forms by adapting Dretske's semantic information theory to analyze the four types of information meaning.

3.2 Analysis of the Fully Matched Form

$I_A(B) = I(B) = I(A)$

N(A)=0, and E(B)=0

It means information content subset B has the amount of information of data subset A as happening, that is the amount of information as information content B or the amount of information as data subset A happening. That is no noise and no fuzzy situations happen. If we make information content as the intension of concept and data item, then the concept manifested by <A, B> is an accurate one with clear intension and extension.

3.3 Analysis of the First partly Matched Form

N(A)=0, but E(B)≠0

$\therefore I_A(B) = I(A) - E(B)$

Seeing Fig.2:

B={b1},

A={d1,d2},

$A' = B$,

$B' = A \bigcup \{d3\}$

First, simulating Dretske's sender and receiver to show the first partly matched form （ seen in Fig.6）.

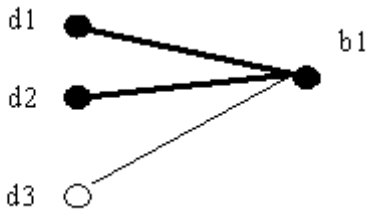

Fig. 6 the first partly matched form

Assuming that d1, d2, d3are independent, the occur probability are 1/3, then the amount of information of data subset A which is taken by information content B occurring:

$I_A(B)=I(A)-E(B)$

Here:

I(A)=-(p(d1)*log(p(d1))+ p(d2)*log(p(d2)))
=log(3)*(2/3)
=1.06 bits

E(B)=-(p(A/B)log(p(A/B)+ p(d3/B)log(p(d3/B))
=(2/3)*log(3/2)+(1/3)*log(3)
=0.92 bits

$I_A(B)=1.06-0.92$
=0.14 bits
$I_A(B)< I(A)$

Since the existence of d3 which make information content b1not can be manifested by data subset {d1, d2}, also can be manifested by d3. Therefore, data subset {d1, d2} can not completely manifest information content b1.

If we make information content as the property of concept and data item as the extension, then the concept manifested by <A, B> is an uncertain one with a clear intension but a uncertain extension1. We can use IA(B)/I(A) to manifest this uncertainty, the extent of extension A manifests intension B.

3.4 Analysis of the second partly matched form

E(B) =0, but N(A)≠0

$\therefore I_A(B) = I(B) - N(A)$

Seeing Fig.3:

B={b1, b2},

A={d1 },

$B' = A$,

$A' = B \bigcup \{b3\}$

Sameness, simulating Dretske's sender and receiver to represent the second partly matched form （seen in Fig 7）.

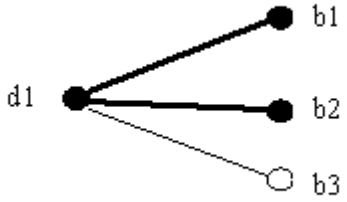

Fig.7 the second partly matched form

Assuming that d1, d2, d3are independent, the occur probability are 1/3, then the amount of information of data subset A which is taken by information content B occurring:

$I_A(B)=I(B)-N(A)$

Here:

I(B)=-(p(b1)*log(p(b1))+ p(b2)*log(p(b2)))

=log(3)*(2/3)

=1.06 bits

N(A)=-(p(B/A)log(p(B/A)+ p(b3/A)log(p(b3/A))

=(2/3)*log(3/2)+(1/3)*log(3)

=0.92 bits

$I_A(B)=1.06-0.92$

=0.14 bits

$I_A(B)< I(B)$

Since the existence of b3 which make information content d1not can be manifested by data subset {b1, b2}, also can be manifested by b3. Therefore, data subset {b1, b2} can not completely manifest information content d1.

If we make information content as the property of concept and data item as the extension, then the concept manifested by <A, B> is an uncertain one with a uncertain intension but a clear extension. We can use IA (B)/I (A) to manifest this uncertainty, the extent of intension B manifest extension A.

3.5 Analysis of Not Completely Matched Form

E(B) ≠0 and N(A)≠0

Seeing Fig.4:

B={b1 },

A={d1 },

$B' = A \bigcup \{d2\}$,

$A' = B \bigcup \{b2\}$

Simulating Dretske's sender and receiver to represent the not completely matched form (seen in Fig 8）.

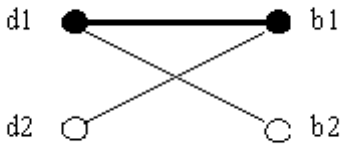

Fig.8 not completely matched form

$I_A(B)<I(B)$ and $I_A(B)<I(A)$

If we make information content B to explain the property of a concept and data subset A for the extension of the concept, then the concept manifested by <A, B> is an un-accurate one with an uncertain intension and an uncertain extension.

4 . Analysis of Four Types of Information Meaning based on Rough Set

We found that data's different provision ability of information content represent different knowledge concept in the chapter above; that is the different form of information meaning based on data reflect different knowledge form, especially the three uncertain knowledge concepts (uncertain knowledge concept with uncertain intension and uncertain extension, uncertain knowledge concept with uncertain intension and clear extension, uncertain knowledge concept with clear intension and uncertain extension), these uncertainties all reflect the so many fuzzy phenomenon existing in the true life (The uncertainty in this paper means the one occurred by fuzziness and uncertainty). These fuzzy phenomenons could not manifest by true or false simply, that is we couldn't manifest them by knowledge concept with clear intension and clear extension.

The exploration of fuzzy phenomenon can trace back to the founder of predication, Frege. In 1904, he proposed the word 'vague' and concluded it to

boundary line. 1980s, Poland's Pawlak proposed Rough Set([1][17]) aimed to boundary regional idea of Frege. He put all the uncertainty unit belong to the boundary region. And this boundary region is defined as the different set of upper approximation set and lower approximation set. Since the upper approximation set and lower approximation set all can be described by certain setmath formula through equivalence relation, then the number of fuzzy item can be counted out and the boundary regional idea of Freg can be realized([17], pp.iii).

In the classic Rough Set theory, we usually use 'un-complete information or knowledge to process some uncertain phenomenon' ([17], pp.iii). Most knowledge representing and processing in the knowledge engineering seldom consider Gettier's counterexamples of modern knowledge theory. Usually assuming that knowledge concept is an of-course and clear term. While in this paper, from Gettier's counterexamples of modern knowledge theory, integrated with the nature of 'to know' is information provision ability of Dretske, we make the following explanation about 'un-complete information or knowledge to process some uncertain phenomenon' in the classic Rough Set Theory. Here 'un-complete information or knowledge' refers to 'the category of un-complete information content' in this paper, and 'un-clear phenomenon' refers to 'unclear data set' in this paper, then we understand knowledge concept by information content provision ability of data.

Meanwhile, we adapt the Gloris connection concepts in Gloris Table to explore the analysis in two dimensions. One is analyzing data by information content, that is 'using un-complete information or knowledge to process some un-clear phenomenon' in the classic Rough Set Theory; one is analyzing information content by data. The analysis in this dimension is 'using un-clear phenomenon to process some un-complete information or knowledge' which is seldom discussed in classic Rough Set Theory. Based on the analysis of these two dimensions, we can get the representation form of the four types above in the Rough Set Theory.

4.1 Analysis of the Fully Matched Form

The fully matched form is: with clear intension and extension.

First analyze data from information content, which is confirming the representation form with clear extension.

Information table shows as Fig 1,

U={a1, a2, d1},

B={b1,b2,b3,c1},

U/B={{a1,a2}, {d1}}

Set X={a1, a2},

then

$\underline{R}(X) = \{a1, a2\}$,

$\overline{R}(X) = \{a1, a2\}$

then:

$\underline{R}(X) = \overline{R}(X)$

Because the information table of Fig 1 reflect 'clear extension', in Rough Set

$\underline{R}(X) = \overline{R}(X)$

It represents that in a certain intension, the extension is clear.

Besides, we analyze the information content by data, which is confirming the representation form with clear extension.

We transpose the information table, set

U={b1,b2,b3,c1},

A={a1,a2,d1},

U/A={{b1,b2,b3},{c1}}

Summing Y={b1, b2, b3},

then

$\underline{R}(Y) = \{b1, b2, b3\}$,

$\overline{R}(Y) = \{b1, b2, b3\}$

so:

$\underline{R}(Y) = \overline{R}(Y)$

Because the transposing of Fig 1 information table reflect 'clear intension', in Rough Set

$\underline{R}(Y) = \overline{R}(Y)$,

it represents that in a certain extension, the intension is clear.

4.2 Analysis of the First Partly Matched Form

The first partly matched form is: with clear intension and uncertain extension

First analyze data from information content confirming the representation form with uncertain extension. Information table shows as Fig 2.

U={a1, a2, d1},

B={b1,b2,b3,c1},

U/B={{a1,a2, d1}}

Set X={a1, a2}, then

$\underline{R}(X) = \varnothing$,

$\overline{R}(X) = \{a1, a2, d1\}$

so:

$\underline{R}(X) \neq \overline{R}(X)$

Because the information table of Fig 2 reflect 'uncertain extension', in Rough Set

$\underline{R}(X) = \overline{R}(X)$,

It represents the uncertain extension situation in a certain intension.

Besides, we analyze the information content by data to confirm the representation form with clear extension. We transpose the information table of Fig 2, set:

U={b1,b2,b3,c1},
A={a1,a2,d1},
U/A={{b1,b2,b3},{c1}}
Assuming Y={b1, b2, b3},
then

$\underline{R}(Y) = \{b1, b2, b3\}$,

$\overline{R}(Y) = \{b1, b2, b3\}$

so:

$\underline{R}(Y) = \overline{R}(Y)$

Because the transposing of Fig 2 information table reflects the 'clear intension' situation and it reconfirm that in Rough Set

$\underline{R}(Y) = \overline{R}(Y)$,

It represents that in a certain extension, the clear situation of the intension.

4.3 Analysis of the Second Partly Matched Form

The second partly matched form is: with uncertain intension and certain extension

First analyzing data from information content confirming the representation form with uncertain extension.

Information table shows as Fig 3.

U={a1, a2, d1},
B={b1,b2,b3,c1},
U/B={{a1,a2}, {d1}}
Set X={a1, a2}, then

$\underline{R}(X) = \{a1, a2\}$,

$\overline{R}(X) = \{a1, a2\}$

So:

$\underline{R}(X) = \overline{R}(X)$

Because the information table of Fig 3 reflect the situation of the 'uncertain extension', in Rough Set

$\underline{R}(X) = \overline{R}(X)$

it represents the uncertain extension situation in a certain intension.

Besides, we analyzing the information content by data to confirm the representation form with clear extension. We transpose the information table of Fig 3, set:

U={b1, b2,b3,c1},
A={a1,a2,d1},
U/A={{b1,b2,b3,c1}}
Assuming Y={b1, b2, b3}, then

$\underline{R}(Y) = \varnothing$,

$\overline{R}(Y) = \{b1, b2, b3, c1\}$

So:

$\underline{R}(Y) \neq \overline{R}(Y)$

Because the transposing of Fig 3 information table reflects the 'uncertain intension' situation and it reconfirm that in Rough Set

$\underline{R}(Y) = \overline{R}(Y)$,

It represents that in a certain extension, the uncertain situation of the intension.

4.4 Analysis of not Completely Matched Form

The not completely matched form is: with uncertain intension and extension.

First analyzing data from information content confirming the representation form with uncertain extension. Information table shows as Fig 4.

U={a1,a2,d1},
B={b1,b2,b3,c1},
U/B={{a1,a2, d1}}
Set X={a1, a2}, then

$\underline{R}(X) = \varnothing$,

$\overline{R}(X) = \{a1, a2, d1\}$

So:

$\underline{R}(X) \neq \overline{R}(X)$

Because the information table of Fig 4 reflect the situation of the 'uncertain extension', in Rough Set

$\underline{R}(X) = \overline{R}(X)$,

It represents the uncertain extension situation in a certain intension.

Besides, we analyzing the information content by data to confirm the representation form with uncertain intension. We transpose the information table of Fig 3, set:

U={ba,b2,b3,c1},
A={a1,a2,d1},
U/A={{b1,b2,b3,c1}}
Set Y={b1, b2, b3}, then

$\underline{R}(Y) = \varnothing$,

$\overline{R}(Y) = \{b1, b2, b3, c1\}$

So:

$\underline{R}(Y) \neq \overline{R}(Y)$

Because the transposing of Fig 4 information table reflects the 'uncertain intension' situation and it reconfirm that in Rough Set

$\underline{R}(Y) \neq \overline{R}(Y)$,

It represents that in a certain extension, the uncertain situation of the intension.

From the analysis above, we find the representation form of clear intension, uncertain intension, clear extension, uncertain extension in Rough Set (Table 1).

Table1 Representation form of uncertainty of intension and extension in Rough Set

	Clear intension	Uncertain intension	Clear extension	Uncertain extension
$\underline{R}(Y) = \overline{R}(Y)$	□			
$\underline{R}(Y) \neq \overline{R}(Y)$		□		
$\underline{R}(X) = \overline{R}(X)$			□	
$\underline{R}(X) \neq \overline{R}(X)$				□

Analyze the information content form data, we can analyze the uncertainty of the intension and find the range of the uncertain intension.

When

$\underline{R}(Y) = \overline{R}(Y)$,

it means clear intension; when

$\underline{R}(Y) \neq \overline{R}(Y)$,

It means uncertain intension. If the uncertain intension existing, then the range of uncertain intension is:

$\overline{R}(Y) - \underline{R}(Y)$.

While analyze data from information content, we can analyze the uncertainty of extension and find out the range of the uncertain extension. When

$\underline{R}(X) = \overline{R}(X)$,

It means clear extension; when

$\underline{R}(X) \neq \overline{R}(X)$,

It means uncertain extension. If the uncertain extension existing, then the range of uncertain extension is:

$\overline{R}(X) - \underline{R}(X)$.

5. Conclusion

In this paper we set four types of connection form of information meaning based on data and analyze the physical significance; then we analyze the representation ability of data in these four connections about information content and the performance of information manifested by data. Then we refined that different connections reflect data's different provision ability to information. While data's different provision ability represents different knowledge concepts, the conclusion are as follows:

1) Meaning is the conjunction between data and information. There are four different types of connect: $B = A'$ and $A = B'$, $B = A'$ and $A \subset B'$, $B \subset A'$ and $A = B'$, $B \subset A'$ and $A \subset B'$. These four connections are defined as: the fully matched form, the first partly matched form, the second partly matched form and the un-completely matched form.

2) The four different types of connection of data-information-meaning match four knowledge concepts: certain knowledge concept with certain intension and extension, uncertain knowledge concept with uncertain intension and uncertain extension, uncertain knowledge concept with uncertain intension and clear extension, uncertain knowledge concept with clear intension and uncertain extension; we call the last three uncertain concepts as three-type uncertain concept, and the uncertainty comes from the un-supplement provision ability. These uncertainty all reflect the so many fuzzy phenomenon existing in the true life (The uncertainty in this paper means the one occurred by fuzziness and uncertainty). These fuzzy phenomenons could not manifest by true or false simply, that is we could not manifest them by knowledge concept with clear intension and clear extension.

3) Based on the Rough Set Theory, integrated with the basic idea of Gloris connection, we analyze the different representation in Rough Set Theory of the intension certainty and extension certainty from two sides; Further deepen the understanding of different conjunction forms of information meaning based on data; Analyzing data from information meaning, we can analyze the certainty of the

extension and the range of uncertainty of extension; analyzing information meaning from data, we can analyze the certainty of the intension and the range of uncertainty of intension.

References:

[1] Lu, Rulin. *Knowledge engineering and knoeledge science at turn of century*. Publisher of Qinghua University,2001,9.

[2] Chen, Jiaming. *Knowledge and corroboration: Preface of morden epistemology*. Shanghai People Publishing House,2003.

[3] Zhong, Yixin. From Statistics to understanding, from convey to cognize. *Journal of Electronics*, 1998,126(7):1-8.

[4] Zhong, Yixin. Discussion on transfer rules of information, knowledge and intelligence. *Journal of Beijing Postal University*,2007,30(1):1-8.

[5] Boland, R.J, Jr. " The In-formation of Information Systems," in *Critical Issues in InformationSystems Research*, Edited by R.J.Boland Jr. and R.A.Hirschheim, John Wiley & Sons LTD, 1987, chapter 14, pp.363-379.

[6] Stamper, R. Organisation Semiotics. In *Information Systems: an Emerging Discipline?*, edt. Mingers,J. and Stowell F. McGraw-Hill Publishing Co., 1997, p.267-283.

[7] Hirschheim, R., Klein, H.K., and Lyytinen, K. *Information systems development and data modeling*, Cambridge University Press, 1995.

[8] Floridi, L. Is Information Meaningful Data?. *Philosophy and Phenomenological Research*, 2005, 70.2, 351-370.

[9] Sufen Wang, Junkang Feng, Binyong Tang. The Need of 'Information Analysis' for Information Systems and Outline of a Hermeneutic Approach to it. *9th International Conference on Enterprise Information Systems*,2007:506-511.

[10] Sufen Wang. *The Analysis and Modeling of the Process of Information Realization in Information Systems based on Hermeneutics and Channel Theory*. PhD Thesis, Donghua University,2007.

[11] Barwise.J. and Seligman.J. *Information Flow—the logic of distributed systems*. Cambridge University Press, 1997.

[12] Dretske, Fred I. *Knowledge and the Flow of Information*. MIT Press, 1981:41-52.

[13] Mingers, J. 1995. Information and meaning: foundations for an intersubjective account, *Info Systems J*, 5, 285-306.

[14] Wang, Sufeng and Wang, Yin. From empiricism to phenomenology: a new information philosophy idea. *Journal of Shanghai Jiaotong University*, 2007,3:74-80.

[15] Priss.U. Formal Concept Analysis in Information Science. http://www.upriss.org.uk/papers/arist.pdf,2006-12-16.

[16] Ganter.B. and Wille.R. *Formal Concept Analysis. Mathematical Foundations.* Springer Verlag, 1999.

[17] Liu, Qing. *Rough Set and Rough inference*. Publishing House of science,2001.