ON INTELLECTUAL QUESTION-ANSWERING WITH INCOMPLETE INFORMATIONS

Sakai, Hiroshi
Department of Information Systems, Interdisciplinary Graduate School of Engineering Sciences,
Kyushu University

ON INTELLECTUAL QUESTION-ANSWERING WITH INCOMPLETE INFORMATIONS

By

Hiroshi SAKAI*

Abstract

Semantic issues connected with incomplete knowledge bases are discussed. In general, information incompleteness means the incompleteness of given values in specified attribute domains, e.g., Codd's null value and Lipski's partial information. Another information incompleteness connected with the lack of primary keys may also be considerable, and in this case the closed world assumption should not be assumed. This paper uniformly formulates these two types of incomplete knowledge base and precisely defines interpretations of queries. In order to realize more intellectual question-answering, some conditions that queries are not affected by the incompleteness are discussed.

1. Introduction

Several types of incomplete informations have already been discussed, for example Codd's null value [1] and Lipski's partial information [2, 3]. In these researches, the incompleteness connected with given values in specified attribute domains is discussed. On the contrary, researches concerned with the lack of primary keys or the lack of information are rare. It seems that only the informations which can assume *the closed world assumption* (CWA) [4] are discussed. But, in general such informations are special cases of ones which can not assume the CWA. Recently, completeness information [5] is introduced instead of the CWA. Completeness information indicates the partial informations which can assume the CWA. Motro extended a relational algebra by the completeness information [5].

In this paper, we discuss the semantic issues based on two types of incomplete informations uniformly. More precisely, the problem is how to realize a question-answering method based on an incomplete information. First, we define a lower bound of a response and an upper bound of a response based on an incomplete information, which are the sure answer and the possible one to a query based on an incomplete system, respectively. Next, a method to compute two responses to a query is discussed. An answer to a query based on a complete information is between the lower bound and the upper bound of the response based on an incomplete system. If the two responses are equal, the query is not affected by the incompleteness of informations in the question-answering. These discussions lead to the realization of a more intel-

* Department of Information Systems, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

lectual question-answering system.

2. Definition of Information Systems

In this section, we clarify basic definitions and two types of incompleteness of an information. First, we show an information system.

DEFINITION 1. *An information system* (or a system) Ω is specified by a triplet,

$$\Omega = \langle X, (D_i)_{i \in I}, (f_i)_{i \in I} \rangle,$$

where,

1) X is a finite set, whose elements are called *objects*.
2) I is a finite set, whose elements are called *attributes*.
3) $D_i (i \in I)$ is a non empty set, which is called *the domain of an attribute i*.
4) $f_i (i \in I)$ is a total function from X to D_i, and called *the assignment function*.

In the question-answering, a query which indicates some conditions on the assignment values is called *term*, and the response is a subset of the object set X. The minimal elements of query languages like $\langle i, E \rangle (i \in I, B \subset D_i)$ are called *descriptors*. One descriptor $\langle i, B \rangle$ implies that the attribute value of i is an element of the set B.

Based on descriptors, terms are defined.

DEFINITION 2. Terms are recursively defined as follows.

1) All descriptors, 0 and 1 are terms.
2) If t and s are terms, $t+s$, $t*s$, $\sim t$ and $t \rightarrow s$ are terms.
3) Terms are defined only by the rule 1) and 2).

The intuitive meaning of connectives '+', '*', '$\sim$' and '$\rightarrow$' are OR, AND, NOT and IMPLY, respectively. Further, '$\cup$', '$\cap$' and '$\setminus$' imply union, intersection and difference of set.

Several problems based on a system are solved by Marek and Pawlak [6]. They defined one interpretation of terms, and showed some equivalent transformation rules for terms based on the interpretation. The one is as follows.

INTERPRETATION 1. ([6]; Marek and Pawlak) The value of a term t based on a system Ω denoted by $\|t\|_\Omega$, is inductively defined as follows. If it is needless to specify the subscript Ω we omit it.

1) $\|\langle i, B \rangle\| = \bigcup_{b \in B} \{x \in X | f_i(x) = b\}$.
2) $\|\langle i, D_i \rangle\| = \|1\| = X$ (object set).
   $\|\langle i, \phi \rangle\| = \|0\| = \Phi$ (ϕ implies an empty set of domains,
   Φ implies an empty set of object.).
3) $\|\sim t\| = X \setminus \|t\|$.
4) $\|t+s\| = \|t\| \cup \|s\|$.
5) $\|t*s\| = \|t\| \cap \|s\|$.
6) $\|t \rightarrow s\| = \|\sim t + s\|$.

According to [6], a question-answering based on a system can be realized by Interpretation 1, further, the equivalent transformation rules for terms can be realized by the boolean algebra.

Next, we define the incompleteness of information systems by the following two types of incompleteness,

1) *Incompleteness of the assignment function,*

2) *Incompleteness of the object set.*

Further, from now on, we assume each element of the system in Definition 1 is fixed, and we call this system Ω *truth complete system.* The incompleteness implies the incompleteness of systems for the truth complete system Ω.

DEFINITION 3. *An incomplete system Ω_1 with respect to the assignment* is such that,

$$\Omega_1 = \langle X, (D_i)_{i \in I}, (f_i^1)_{i \in I} \rangle,$$

where,

1)~3) is the same as Definition 1.

4) $f_i^1 (i \in I)$ is a partial function from X to D_i and if $f_i^1(x)(x \in X)$ is defined, then $f_i^1(x) = f_i(x)$ (f_i is the assignment function of truth complete system Ω.).

If the assignment function f_i^1 is undefined for an object x, then we denote $f_i^1(x) = \perp$. The truth value of $f_i^1(x)$ is $f_i(x)$, but it is not known in the system Ω_1. By the definition of f_i^1, there exist several incomplete systems wrt the assignment, and hereafter Ω_1 implies one of them.

DEFINITION 4. *An incomplete system Ω_2 with respect to the object* is such that,

$$\Omega_2 = \langle Y, (D_i)_{i \in I}, (f_i^2)_{i \in I} \rangle,$$

where,

1) Object set Y is a subset of X (X is the object set of truth complete system Ω.).

2)~3) is the same as Definition 1.

4) $f_i^2 (i \in I)$ is a total function from Y to D_i and $f_i^2(y) = f_i(y)$ for any i and any object y in Y.

In this system, there is no incompleteness with respect to the assignment, but it is known that there are some lack with respect to the object set. By the definition, there also exist several incomplete systems, and also Ω_2 implies one of them.

DEFINITION 5. *An incomplete system Ω_3 with respect to the assignment and the object* is such that,

$$\Omega_3 = \langle Y, (D_i)_{i \in I}, (f_i^3)_{i \in I} \rangle,$$

where,

1)~3) is the same as Definition 4.

4) $f_i^3 (i \in I)$ is a partial function from Y to D_i and if $f_i^3(y)(y \in Y)$ is defined, then $f_i^3(y) = f_i(y)$.

In this system, there is incompleteness with respect to the assignment, and it is known that there are some lack with respect to the object set.

Here, the relations of four systems are as follows.

$\Omega \longrightarrow \Omega_1$

incomplete wrt the assignment

incomplete wrt the object

incomplete wrt the assignment and the object

$\Omega_2 \qquad \Omega_3$
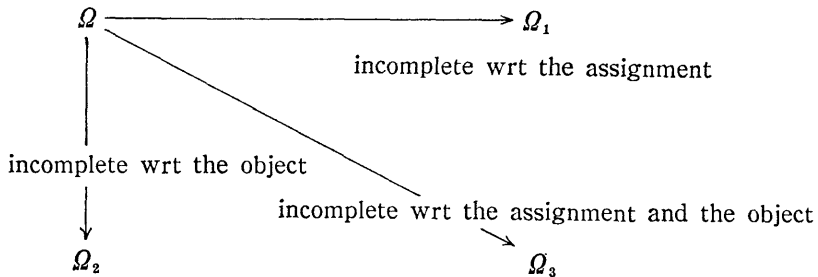
Ω : truth complete system,

Ω_1 : incomplete system wrt the assignment,

Ω_2 : incomplete system wrt the object,

Ω_3 : incomplete system wrt the assignment and the object.

Since the question-answering of the truth complete system Ω is discussed in [6], from now on we discuss three types of incomplete systems.

First, we discuss how to realize a question-answering based on an incomplete system with respect to the assignment, namely the question-answering based on a system Ω_1. For the preparation, we define a concept of extension.

DEFINITION 6. For any two incomplete systems Ψ and Ψ' with respect to the assignment,

$$\Psi = \langle X, (D_i)_{i \in I}, (g_i)_{i \in I} \rangle,$$

$$\Psi' = \langle X, (D_i)_{i \in I}, (g'_i)_{i \in I} \rangle,$$

let take any object $x (\in X)$ and any attribute $i (\in I)$. If $g_i(x) \neq \perp$ then $g'_i(x) = g_i(x)$ holds, then the system Ψ' is called *an extension of system Ψ with respect to the assignment*.

If a system Ψ' whose assignment function is a total function and it is an extension of a system Ψ, then Ψ' is called *a complete extension of system Ψ*. Generally, there exist several complete extensions of Ψ. The set of complete extensions of a system Ψ is denoted by $COEX(\Psi)$, (*COEX* is a concatenation of the underlined parts in complete and extension). One element of $COEX(\Omega_1)$ is constructed by replacing $f_i^1(x) = \perp$ with $f_i^1(x) = d$ (d is one element of D_i). In general, $COEX(\Omega_1)$ may be infinite set. By the definition, truth complete system Ω is one element of $COEX(\Omega_1)$, but the system Ω_1 can not decide which element of $COEX(\Omega_1)$ is the truth complete system Ω.
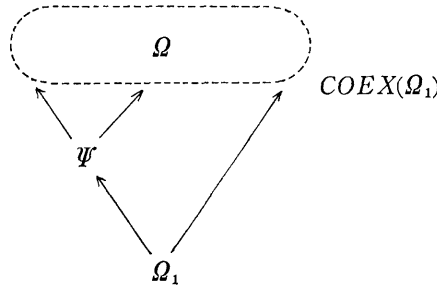

Fig. 1. Extensions of system Ω_1 wrt the assignment.

3. One Problem Concerned with an Incompleteness of the Assignment Function

By the definition of an incomplete system wrt the assignment in Definition 3, we may consider that one truth complete system is in $COEX(\Omega_1)$. We want to realize the question-answering based on an unknown truth complete system Ω. But, we do not know which element of $COEX(\Omega_1)$ is the truth complete system.

Clearly, the following relations hold for any term t.

$$\bigcap_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi} \subset \|t\|_{\Omega} \subset \bigcup_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi}.$$

According to Section 2, we can now evaluate any query based on a complete system with respect to the assignment, i.e., we can calculate $\|t\|_{\Psi}$ for any $\Psi \in COEX(\Omega_1)$. Therefore, we can not evaluate a term based on an unknown truth complete system, i.e. $\|t\|_{\Omega}$, but we can define the lower bound and the upper bound of a response of $\|t\|_{\Omega}$ by the incomplete system Ω_1.

DEFINITION 7. For any incomplete system Ω_1, we define *the lower bound and the upper bound of a response* for a term t based on an unknown truth complete system Ω as follows.

    1) The lower bound of a response is

$$\bigcap_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi}.$$

    2) The upper bound of a response is

$$\bigcup_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi}.$$

In general, $COEX(\Omega_1)$ may be infinite set. Therefore, it is impossible to compute the lower and the upper bounds by set operations. The following problem arises.

PROBLEM 1. How to compute the lower and the upper bounds of a response?

For the problem, we want to give one interpretation, whose evaluation of a term agrees with the lower and the upper bound of a response. Similar to [6], we define the second interpretation and study its property.

INTERPRETATION 2. The values of a term t of an incomplete system Ω_1, denoted by $inf\|t\|_{\Omega_1}$ and $sup\|t\|_{\Omega_1}$, are inductively defined as follows. If it is needless to specify the subscript Ω_1, we omit it.

    1) $inf\|\langle i, B \rangle\| = \bigcup_{b \in B} \{x \in X \mid f_i^1(x) = b\}.$
       $sup\|\langle i, B \rangle\| = \bigcup_{b \in B} \{x \in X \mid f_i^1(x) = b\}$
             $\bigcup \{x \in X \mid f_i^1(x) = \perp\}.$
    2) $inf\|0\| = sup\|0\| = \Phi$ (empty set).
    3) $inf\|1\| = sup\|1\| = X$ (total object set).
    4) $inf\|\sim t\| = X \backslash sup\|t\|.$
       $sup\|\sim t\| = X \backslash inf\|t\|.$
    5) $inf\|t+s\| = inf\|t\| \cup inf\|s\|.$
       $sup\|t+s\| = sup\|t\| \cup sup\|s\|.$
    6) $inf\|t*s\| = inf\|t\| \cap inf\|s\|.$
       $sup\|t*s\| = sup\|t\| \cap sup\|s\|.$
    7) $inf\|t \rightarrow s\| = inf\|\sim t+s\|.$
       $sup\|t \rightarrow s\| = sup\|\sim t+s\|.$

DEFINITION 8. For any system Ω_1 and any term, we define $=_i^1$, $=_s^1$ and $=^1$.

    1) If $inf\|t\|_{\Omega_1} = inf\|s\|_{\Omega_1}$, then we denote $t =_i^1 s$ and call t and s are equivalent with respect to $inf\| \cdot \|_{\Omega_1}$.
    2) If $sup\|t\|_{\Omega_1} = sup\|s\|_{\Omega_1}$, then we denote $t =_s^1 s$ and call t and s are equivalent with respect to $sup\| \cdot \|_{\Omega_1}$.
    3) If $t =_i^1 s$ and $t =_s^1 s$, then we denote $t =^1 s$ and call t and s are equivalent with respect to Ω_1.

PROPOSITION 9. *For any descriptors $\langle i, A \rangle$ and $\langle i, B \rangle$, the following equations hold.*

1) $\langle i, A \rangle + \langle i, B \rangle =^1 \langle i, A \cup B \rangle$.

2) $\langle i, A \rangle * \langle i, B \rangle =^1 \langle i, A \cap B \rangle$.

3) $\sim \langle i, B \rangle =^1 \langle i, D_i \backslash B \rangle$.

[Proof] Let Ω_1 be an incomplete system wrt the assignment.

1) $inf \| \langle i, A \rangle + \langle i, B \rangle \|_{\Omega_1} = inf \| \langle i, A \rangle \|_{\Omega_1} \cup inf \| \langle i, B \rangle \|_{\Omega_1}$

$= \bigcup_{a \in A} \{ x \in X \mid f_i^1(x) = a \} \cup \bigcup_{b \in B} \{ x \in X \mid f_i^1(x) = b \}$

$= \bigcup_{c \in A \cup B} \{ x \in X \mid f_i^1(x) = c \} = inf \| \langle i, A \cup B \rangle \|_{\Omega_1}$.

$sup \| \langle i, A \rangle + \langle i, B \rangle \|_{\Omega_1} = sup \| \langle i, A \rangle \|_{\Omega_1} \cup sup \| \langle i, B \rangle \|_{\Omega_1}$

$= \bigcup_{a \in A} \{ x \in X \mid f_i^1(x) = a \} \cup \bigcup_{b \in B} \{ x \in X \mid f_i^1(x) = b \}$

$\cup \{ x \in X \mid f_i^1(x) = \perp \}$

$= \bigcup_{c \in A \cup B} \{ x \in X \mid f_i^1(x) = c \} \cup \{ x \in X \mid f_i^1(x) = \perp \}$

$= sup \| \langle i, A \cup B \rangle \|_{\Omega_1}$.

2) and 3) are similarly proved.

PROPOSITION 10. *For any terms t and s, the following equations hold.*

1) $\sim (t + s) =^1 \sim t * \sim s$.

2) $\sim (t * s) =^1 \sim t + \sim s$.

3) $\sim (\sim t) =^1 t$.

4) *Concerning to '+' and '*' operators, commutative, associative and distributive laws hold.*

[Proof] The proof is clear by Interpretation 2.

By Proposition 9 and 10, for any system Ω_1, any term t can be equivalently transformed to a term s in the form of $\Sigma_l \Pi_m \langle i_{lm}, B_{lm} \rangle$.
Here, the symbol 'Σ' implies a concatenation of finite terms by '+', and the symbol 'Π' implies a concatenation of finite terms by '*'.

PROPOSITION 11. *Let Ψ be one extension wrt the assignment of a system Ω_1 and t be any term, the following relations hold.*

1) $inf \| t \|_{\Omega_1} \subset inf \| t \|_\Psi$.

2) $sup \| t \|_\Psi \subset sup \| t \|_{\Omega_1}$.

3) $inf \| t \|_{\Omega_1} \subset sup \| t \|_{\Omega_1}$.

[Proof] 1) Generally, it is sufficient to consider a term in the form of $\Sigma_l \Pi_m \langle i_{lm}, B_{lm} \rangle$.

$inf \| \Sigma_l \Pi_m \langle i_{lm}, B_{lm} \rangle \|_{\Omega_1}$

$= \bigcup_l \bigcap_m inf \| \langle i_{lm}, B_{lm} \rangle \|_{\Omega_1}$

$\subset \bigcup_l \bigcap_m inf \| \langle i_{lm}, B_{lm} \rangle \|_\Psi$

$= inf \| t \|_\Psi$.

2) and 3) are similary proved.

PROPOSITION 12. *If a system Ψ is complete wrt the assignment, then the following equations hold.*

$inf \| t \|_\Psi = \| t \|_\Psi = sup \| t \|_\Psi$ for any term t.

[Proof] By Proposition 11, it is clear.

Proposition 12 shows Interpretation 1 is a special case of the Interpretation 2.

PROPOSITION 13. *For any term t, the following equations hold.*

1) $t + 1 =^1 1$, $t + 0 =^1 t$.

2) $t * 1 =^1 t$, $t * 0 =^1 0$.

3) $t+(\sim t)\neq\frac{1}{4}1$, $t+(\sim t)=\frac{1}{8}1$.

4) $t*(\sim t)=\frac{1}{4}0$, $t*(\sim t)\neq\frac{1}{8}0$.

[Proof] Let Ω_1 be an incomplete information system.

3) $inf\|t+(\sim t)\|_{\Omega_1}=inf\|t\|_{\Omega_1}\cup(X\backslash sup\|t\|_{\Omega_1})$. In general, $inf\|t\|_{\Omega_1}\subset sup\|t\|_{\Omega_1}$ by Proposition 11, therefore $inf\|t+(\sim t)\|_{\Omega_1}\subset X=inf\|1\|_{\Omega_1}$.

These propositions show that the query transformation based on psuedo boolean algebra keeps the equivalency of terms. The most important cause that the complementary lows in Proposition 13 do not hold is as follows. In general,

$\|\langle i, D_i\rangle\|_{\Psi}=X$ (total object set) for any complete system,

and $\bigcap_{\Psi\in COEX(\Omega_1)}\|\langle i, D_i\rangle\|_{\Psi}=X$ can be derived. On the other hand, $inf\|\langle i, D_i\rangle\|_{\Omega_1}$ is the set of the objects whose assignment is definite in the incomplete system. It shows that for some incomplete systems,

$$\bigcap_{\Psi\in COEX(\Omega_1)}\|t\|_{\Psi}\neq inf\|t\|_{\Omega_1}.$$

In real, only for two types of trivial descriptors $\langle i, \phi\rangle$ and $\langle i, D_i\rangle$, there exists a difference between Interpretation 2 and two bounds. By revising this point, we get the following theorem.

THEOREM 14. *For any system Ω_1, the following equations hold.*

1) *If a term t is the form of $\Pi_l\Sigma_m\langle i_{lm}, B_{lm}\rangle$,*
if $m\neq m'$ then $i_{lm}\neq i_{lm'}$ for any l, and
$\phi\neq B_{lm}\subsetneqq D_{i_{lm}}$ for any l and m, then

$$\bigcap_{\Psi\in COEX(\Omega_1)}\|t\|_{\Psi}=inf\|t\|_{\Omega_1}$$

$$(=\bigcap_l\bigcup_m inf\|\langle i_{lm}, B_{lm}\rangle\|_{\Omega_1}).$$

2) *If term t is the form of $\Sigma_l\Pi_m\langle i_{lm}, B_{lm}\rangle$,*
if $m\neq m'$ then $i_{lm}\neq i_{lm'}$ for any l, and
$\phi\neq B_{lm}\subsetneqq D_{i_{lm}}$ for any l and m, then

$$\bigcup_{\Psi\in COEX(\Omega_1)}\|t\|_{\Psi}=sup\|t\|_{\Omega_1}$$

$$(=\bigcup_l\bigcap_m sup\|\langle i_{lm}, B_{lm}\rangle\|_{\Omega_1}).$$

[Proof] 1) We show mutual inclusions.

First, we show $\bigcap_{\Psi\in COEX(\Omega_1)}\|t\|_{\Psi}\supset inf\|t\|_{\Omega_1}$ for any term t.

By Proposition 11, $inf\|t\|_{\Omega_1}\subset inf\|t\|_{\Psi}$ for any extension Ψ. Namely,

$inf\|t\|_{\Omega_1}\subset inf\|t\|_{\Psi}$ for any system Ψ in $COEX(\Omega_1)$, and

$inf\|t\|_{\Psi}=\|t\|_{\Psi}$ for any system in $COEX(\Omega_1)$

by Proposition 12. Therefore,

$$\bigcap_{\Psi\in COEX(\Omega_1)}\|t\|_{\Psi}\supset inf\|t\|_{\Omega_1}.$$

Next, take a term t which satisfies the condition of 1). we show $\bigcap_{\Psi\in COEX}\|t\|_{\Psi}\subset inf\|t\|_{\Omega_1}$. We suppose $x_0\in(\bigcap_{\Psi\in COEX(\Omega_1)}\|t\|_{\Psi}\backslash inf\|t\|_{\Omega_1})$, then

$$x_0\in\|t\|_{\Psi}\quad\text{for any}\quad\Psi\in COEX(\Omega_1)$$

and

$$x_0\notin inf\|t\|_{\Omega_1}.$$

$x_0 \in \|t\|_\Psi$ for a complete system Ψ is equivalent to there exists m such that $x_0 \in \|\langle i_{lm}, B_{lm}\rangle\|_\Psi$ for any l.

Here, $(f_i^1)_{i\in I}$ is the assignment function of the system Ω_1. If there exists m such that

$$f^1_{i_{lm}}(x_0) \in B_{lm} \quad \text{for any } l, \text{ then } \quad x_0 \in inf\|t\|_{\Omega_1} \text{ is trivial},$$

and it contradicts to $x_0 \overline{\in} inf\|t\|_{\Omega_1}$.

If there exists m such that

$f^1_{i_{lm}}(x_0) = \perp$ and term t does not satisfy the above case, then $x_0 \overline{\in} inf\|t\|_{\Omega_1}$.
But in this case, there exists a complete system Ψ such that

$$x_0 \overline{\in} \|t\|_\Psi,$$

because there exists one extension $g_{i_{lm}}(x_0) \in D_{i_{lm}}\backslash B_{lm}$. g_i is the assigment function of Ψ. (Since the term t satisfies $\phi \neq B_{i_{lm}} \subsetneqq D_{i_{lm}}$ for any l and m, there exists at least one element in $D_{i_{lm}}\backslash B_{i_{lm}}$ for any l and m.)

If a term t is in the form of $\Sigma_l \Pi_m \langle i_{lm}, B_{lm}\rangle$, then there may be a term $\langle i, B\rangle + \langle i, D_i\backslash B\rangle$ in the term t, i.e., the term t may contain $\langle i, D_i\rangle$. In order to exclude this case, we must constrain the term t in the form of $\Pi_l \Sigma_m \langle i_{lm}, B_{lm}\rangle$.

After all, there is no x_0 such that,

$$x_0 \in (\bigcap_{\Psi \in COEX}\|t\|_\Psi \backslash inf\|t\|_{\Omega_1}).$$

Therefore,

$$\bigcap_{\Psi \in COEX(\Omega_1)}\|t\|_\Psi \subset inf\|t\|_{\Omega_1}.$$

Namely,

$$\bigcap_{\Psi \in COEX(\Omega_1)}\|t\|_\Psi = inf\|t\|_{\Omega_1}.$$

2) $sup\|t\|_{\Omega_1} = X\backslash inf\|\sim t\|_{\Omega_1}$ is the definition of Interpretation 2. Here, term t is in the form of

$$\Sigma_l \Pi_m \langle i_{lm}, B_{lm}\rangle$$

by the condition, and then a term $\sim t$ is in the form of

$$\Pi_l \Sigma_m \langle i_{lm}, D_{i_{lm}}\backslash B_{lm}\rangle.$$

This term $\sim t$ satisfies the condition 1), therefore the result of 1) can be applicable to this case.

Let $t = \Sigma_l \Pi_m \langle i_{lm}, B_{lm}\rangle$, then

$$sup\|t\|_{\Omega_1}$$

$$= X\backslash(inf\|\sim t\|_{\Omega_1}),$$

$$= X\backslash(inf\|\Pi_l \Sigma_m \langle i_{lm}, D_{i_{lm}}\backslash B_{lm}\rangle\|_{\Omega_1}),$$

$$= X\backslash(\bigcap_{\Psi \in COEX(\Omega_1)}\|\Pi_l \Sigma_m \langle i_{lm}, D_{i_{lm}}\backslash B_{lm}\rangle\|_\Psi)$$

$$= X\backslash(\bigcap_{\Psi \in COEX(\Omega_1)}\|\sim t\|_\Psi)$$

$$= X\backslash(\bigcap_{\Psi \in COEX(\Omega_1)}(X\backslash\|t\|_\Psi))$$

$$= X\backslash(X\backslash(\bigcup_{\Psi \in COEX(\Omega_1)}\|t\|_\Psi))$$

$$=\bigcup_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi}.$$

According to [6], any term t can be equivalently transformed to terms which satisfy the conditions of Theorem 14 based on a complete system. Therefore, Theorem 14 is applicable to any term. Namely, for any term t and an incomplete system Ω_1,

$$\bigcap_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi} = \bigcap_{\Psi \in COEX(\Omega_1)} \|s\|_{\Psi} \quad \text{(equivalent transformation based on a complete system)}$$

$$= \inf \|s\|_{\Omega_1}. \qquad \text{(Theorem 14)}$$

$$\bigcup_{\Psi \in COEX(\Omega_1)} \|t\|_{\Psi} = \bigcup_{\Psi \in COEX(\Omega_1)} \|r\|_{\Psi} \quad \text{(equivalent transformation based on a complete system)}$$

$$= \sup \|r\|_{\Omega_1}. \qquad \text{(Theorem 14)}$$

4. Two Problems Concerned with the Incompleteness of the Assigment Function and the Lack of Objects

Up to Section 3, the incompleteness of informations depends upon the incompleteness of the assignment function. But, from now on, we also consider the incompleteness of the object set. Namely, we also consider the question-ansewring based on two systems Ω_2 and Ω_3 in Definition 4 and 5. In this case, we always feel anxious that there may really be other objects which are not in the current object set but satisfy a condition of a term, because a range of a response is beforehand limited.

In our question-answering, the response is a subset of object set, therefore there may be other objects which satisfy the condition of a term, but is not in the response. In general, these problems are rarely discussed, because the CWA for the object set is usually assumed.

In this section, we can not assume the CWA for the object set. Instead, we introduce the completeness information [5]. The intuitive meaning of the completeness information is not to assume the CWA for the object set, but to assume the CWA for the objects which take some fixed attribute values. We leave the precise definition later.

First, we discuss the question-answering based on a system Ω_2 in Definition 4. In the system Ω_2, there is no incompleteness with respect to the assignment. Therefore, the result of [6] is applicable to Ω_2. But, there may exist objects which satisfy the condition of a query without a response of system Ω_2.

Next, we discuss the system Ω_3 in Definition 5, and define an extension with respect to the object.

DEFINITION 15. Let Ψ and Γ be any two systems,

$$\Psi = \langle Y, (D_i)_{i \in I}, (g_i)_{i \in I} \rangle$$

and

$$\Gamma = \langle Z, (D_i)_{i \in I}, (h_i)_{i \in I} \rangle.$$

If the following conditions hold, then the system Γ is called *one extension of Ψ with respect to the object.*

1) $Y \subset Z$.

2) $h_i(y)=g_i(y)$ for any attribute i and any object $y \in Y$.

3) $(h_i)_{i \in I}$ is a total function for any object in $Z \setminus Y$.

In real, when we construct one extension with respect to the object from a current system, the lacked object itself can not be known, but the assignment values of it is one element of D_i for each attribute. Let a system Ψ be

$$\Psi = \langle Y, (D_i)_{i \in I}, (g_i)_{i \in I} \rangle.$$

In this case, take an unknown object symbol m and define the assignment function h_i for any i as follows.

$$h_i(y) = \begin{cases} g_i(y) & (y \in Y), \\ an\ element\ of\ D_i & (y=m), \end{cases}$$

and set the system Γ as follows,

$$\Gamma = \langle Y \cup \{m\}, (D_i)_{i \in I}, (h_i)_{i \in I} \rangle.$$

We consider this system is one extension of system Ψ. Therefore, there may exist false informations in an extended system with respect to the object.

Next, we closely define completeness information [5].

DEFINITION 16. Let Ω_3 be any system in Definition 5.

$$\Omega_3 = \langle Y, (D_i)_{i \in I}, (f_i^3)_{i \in I} \rangle.$$

Completeness information for an attribute i, denoted by $COMP(D_i)$, is a subset of D_i and satisfies the following conditions.

1) For any $y \in Y$, if $f_i^3(y)=\perp$, then the truth value of the object y (i. e., $f_i(y)$), is in $D_i \setminus COMP(D_i)$ for an attribute i.

2) For any extension system Ψ of Ω_3 with respect to the object, $\Psi = \langle Z, (D_i)_{i \in I},$ $(g_i)_{i \in I} \rangle$, the truth value of any object $z(\in Z \setminus Y)$, i. e. $f_i(z)$ is in $D_i \setminus COMP(D_i)$ for any attribute i.

We consider there exist an attribute i such that $COMP(D_i)=D_i$ in truth complete system Ω and incomplete system Ω_1, but we consider there is not such attributes in the system Ω_2 and Ω_3.

After this, we consider the case that these completeness informations are given, namely the new system Ω_3 is specified by a quadruplet as follows,

$$\Omega_3 = \langle Y, (D_i)_{i \in I}, (f_i^3)_{i \in I}, (COMP(D_i))_{i \in I} \rangle,$$

Other notations with respect to a new system are the same as those of the former one.

Introduction of the completeness informations to an incomplete system with respect to the assignment and the object causes the following problem.

PROBLEM 2. How to compute the lower and the upper bounds of a response of a term in Definition 7 based on a new system?

The difference between the Problem 1 and 2 is that, in Problem 1, if $f_i^1(x)=\perp$ then the truth value of the object x is in D_i, but in Problem 2, if $f_i^3(x)=\perp$ then the truth value of the object x is in $D_i \setminus COMP(D_i)$. For the problem, we want to give

one interpretation, whose evaluation of a term agrees with the lower and the upper bound of a response, respectively. Similar to [6], we define the third interpretation and study its property. In the third interpretation, the following should be noticed. If a descriptor $\langle i, B \rangle$ satisfies $B \neq D_i$ and $B \cup COMP(D_i) = D_i$, then clearly the truth value of an object x which is not definite is in B, and $x \in inf \| \langle i, B \rangle \|_{\Omega_3}$ should hold. Next, if a descriptor $\langle i, A \rangle$ satisfies $\phi \neq A \subset COMP(D_i)$, then the truth value of an object x which is $f_i^3(x) = \perp$ is not in A, and $x \notin sup \| \langle i, A \rangle \|_{\Omega_3}$ should hold.

INTERPRETATION 3. The values of a term t in a new incomplete system Ω_3, denoted by $inf \| t \|_{\Omega_3}$ and $sup \| t \|_{\Omega_3}$, are inductively defined as follows. If it is needless to specify the subscript Ω_3, we omit it.

1) $inf \| \langle i, B \rangle \|$ is as follows;
If $B \neq D_i$ and $B \cup COMP(D_i) = D_i$ then

$$\bigcup_{b \in B} \{ x \in X \mid f_i^3(x) = b \}$$

$$\cup \{ x \in X \mid f_i^3(x) = \perp \}.$$

else

$$\bigcup_{b \in B} \{ x \in X \mid f_i^3(x) = b \}.$$

$sup \| \langle i, B \rangle \|$ is as follows;
If $\phi \neq B \subset COMP(D_i)$ then

$$\bigcup_{b \in B} \{ x \in X \mid f_i^3(x) = b \}.$$

else

$$\bigcup_{b \in B} \{ x \in X \mid f_i^3(x) = b \}$$

$$\cup \{ x \in X \mid f_i^3(x) = \perp \}.$$

2)~7) is the same as Interpretation 2.

The following propositions hold based on Interpretation 3.

PROPOSITION 17. *For any descriptors $\langle i, A \rangle$ and $\langle i, B \rangle$ the following equations hold. Here, $=_i^3$, $=_s^3$ and $=^3$ are similarly defined as $=_i^1$, $=_s^1$ and $=^1$ in* Definition 8.

1) $\langle i, A \rangle + \langle i, B \rangle \neq_i^3 \langle i, A \cup B \rangle$.
$\langle i, A \rangle + \langle i, B \rangle =_s^3 \langle i, A \cup B \rangle$.
2) $\langle i, A \rangle * \langle i, B \rangle =_i^3 \langle i, A \cap B \rangle$.
$\langle i, A \rangle * \langle i, B \rangle \neq_s^3 \langle i, A \cap B \rangle$.
3) $\sim \langle i, B \rangle =^3 \langle i, D_i \backslash B \rangle$.

PROPOSITION 18. *Concerning to the evaluation of any terms t and s based on* Interpretation 3, *the same equations of* Proposition 10, 11 *and* 12 *hold.*

THEOREM 19. *Concerning to the evaluation of any term t based on* Interpretation 3, *the same equations of* Theorem 14 *hold.*

Next, let take the system Ω_3 in Definition 5 and Ω_1 in Definition 3. In this case, Ω_1 is a truth extension wrt the object of Ω_3, and generally,

$$inf \| t \|_{\Omega_3} \subset inf \| t \|_{\Omega_1} \quad \text{and} \quad sup \| t \|_{\Omega_3} \subset sup \| t \|_{\Omega_1}$$

hold. Because, the object set of system Ω_1 is larger than the object set of system Ω_3.

The following problem arises.

PROBLEM 3. What are the conditions of a term t such that,

$$inf\|t\|_{\Omega_3}=inf\|t\|_{\Psi} \quad and \quad sup\|t\|_{\Omega_3}=sup\|t\|_{\Psi},$$

for any incomplete system Ω_3 with respect to the assignment and the object, and its any extension system Ψ with respect to the object?

The purpose of this problem is to clarify some terms which is not affected by the incompleteness of object set in the question-answering, and to assure that there is no object which satisfy the condition of the term. In general, this assurance can not be assumed because of the incompleteness of object set. Namely, for the term t which satisfies the condition of Problem 3,

$$inf\|t\|_{\Omega_3}=inf\|t\|_{\Omega_1} \quad and \quad sup\|t\|_{\Omega_3}=sup\|t\|_{\Omega_1}$$

can be derived. Next, for an unknown truth complete system Ω

$$inf\|t\|_{\Omega_3}=inf\|t\|_{\Omega_1}\subset\|t\|_{\Omega}\subset sup\|t\|_{\Omega_1}=sup\|t\|_{\Omega_3}$$

hold. Namely,

$$inf\|t\|_{\Omega_3}\subset\|t\|_{\Omega}\subset sup\|t\|_{\Omega_3}$$

can be derived from current incomplete system Ω_3.

PROPOSITION 20. *For any system Ω_3 and its any extension Ψ with respect to the object,*

$$\Omega_3=\langle Y, (D_i)_{i\in I}, (f_i^3)_{i\in I}, (COMP(D_i))_{i\in I}\rangle,$$

$$\Psi=\langle Z, (D_i)_{i\in I}, (g_i)_{i\in I}, (COMP(D_i))_{i\in I}\rangle,$$

we make a new complete system Γ with respect to the assignment as follows

$$\Gamma=\langle Z\backslash Y, (D_i)_{i\in I}, (g_i)_{i\in I}, (COMP(D_i))_{i\in I}\rangle$$

Then, for any term t the following equations hold.

$$inf\|t\|_{\Psi}=inf\|t\|_{\Omega_3}\cup\|t\|_{\Gamma}.$$

$$sup\|t\|_{\Psi}=sup\|t\|_{\Omega_3}\cup\|t\|_{\Gamma}.$$

[Proof] Any term t can be transformed to

$$\Sigma_l\Pi_m\langle i_{lm}, B_{lm}\rangle$$

by the Interpretation 3. And then,

$$inf\|t\|_{\Psi}=inf\|\Sigma_l\Pi_m\langle i_{lm}, B_{lm}\rangle\|_{\Psi}$$

$$=\bigcup_l\bigcap_m inf\|\langle i_{lm}, B_{lm}\rangle\|_{\Psi}.$$

Since there is no element of $Y\cap(Z\backslash Y)$,

$$=\bigcup_l\bigcap_m inf\|\langle i_{lm}, B_{lm}\rangle\|_{\Omega_3}$$

$$\cup\bigcup_l\bigcap_m inf\|\langle i_{lm}, B_{lm}\rangle\|_{\Gamma}$$

$$=inf\|t\|_{\Omega_3}\cup inf\|t\|_{\Gamma}.$$

The system Γ is a complete system with respect to the assignment,

$$inf\|t\|_\Gamma = \|t\|_\Gamma$$

can be derived. It implies

$$inf\|t\|_\Psi = inf\|t\|_{\Omega_3} \cup \|t\|_\Gamma .$$

The proof for $sup\|t\|_\Psi$ is similarly shown.

This proposition shows that the conditions of Problem 3 are $\|t\|_\Gamma = \Phi$ for any complete system Γ in Poposition 20.   Therefore the following theorem can be derived.

THEOREM 21.   *For any system* Ω_3

$$\Omega_3 = \langle Y, (D_i)_{i \in I}, (f_i^3)_{i \in I}, (COMP(D_i)_{i \in I} \rangle ,$$

and any extension Ψ *with respect to the object, the necessary and sufficient conditions of a term t that satisfy*

$$inf\|t\|_{\Omega_3} = inf\|t\|_\Psi \quad and \quad sup\|t\|_{\Omega_3} = sup\|t\|_\Psi ,$$

are such that
  when the term is transformed by Interpretation 1 *to*
  $\Sigma_l \Pi_m \langle i_{lm}, B_{lm} \rangle$ *which satisfies*
  $i_{lm} \neq i_{lm}$, *for* $m \neq m'$ *and any l, and*
  $\phi \neq B_{lm} \subsetneqq D_{i_{lm}}$ *for any l and m*
  (*According to* [6], *this transformed term can always be derived.*)
  Then, there exist a descriptor $\langle i_{lm}, B_{lm} \rangle$ *such that* $B_{lm} \subset COMP(D_{i_{lm}})$ *for any l.*
[Proof] Let Γ be a complete system with respect to the assignment, and whose object set does not contain any object of the system Ω_3.   Then,

$$\|t\|_\Gamma = \|\Sigma_l \Pi_m \langle i_{lm}, B_{lm} \rangle \|_\Gamma$$

$$= \bigcup_l \bigcap_m \|\langle i_{lm}, B_{lm} \rangle \|_\Gamma$$

If $B_{lm} \subset COMP(D_{i_{lm}})$ for any l, then

$$\bigcap_m \|\langle i_{lm}, B_{lm} \rangle \|_\Gamma = \Phi .$$

Therefore, $\|t\|_\Gamma = \Phi$.

On the contrary, if $\|t\|_\Gamma = \Phi$ for any complete system Γ and there exists a descriptor $\langle i_{l'm'}, B_{l'm'} \rangle$ such that $B_{l'm'} \not\subset COMP(D_{i_{l'm'}})$, then there exists one complete system Λ whose assignment g_i is

$$g_{i_{l'm'}}(x) \in B_{l'm'} \backslash COMP(D_{i_{l'm'}}),$$

$$g_{i_{l1}}(x) \in B_{l1} \quad for \ any \ l(\neq l')$$

Clearly, $\|t\|_\Lambda \neq \Phi$, it contradicts to $\|t\|_\Gamma = \Phi$ for any complete system whose object set does not contain any object in system Ω_3.

5. Concludng Remarks

In this paper, we discussed the semantic issues connected with the incompleteness of the assignment and the object.   Namely,
  1) Concerning to the incompleteness of the assignment, the calculation method of

the lower and the upper bounds of a response is shown.

2) Concerning to the incompleteness of the object set, some conditions of terms which is not affected by its incompleteness are shown.

More concretely, four systems were defined as follows.

Ω : truth complete system,

Ω_1: incomplete system with respect to the assignment,

Ω_2: incomplete system with respect to the object,

Ω_3: incomplete system wrt the assignment and the object.

In Problem 1, the relation of Ω and Ω_1 is discussed.

Where, Ω_1 is known and Ω is unknown. For any term t,

$$inf\|t\|_{\Omega_1}\subset\|t\|_{\Omega}\subset sup\|t\|_{\Omega_1}$$

can be derived based on Interpretation 2.

In Problem 2, the relation of Ω_2 and Ω_3 is discussed.

Where, Ω_3 is known and Ω_2 is unknown. For any term t,

$$inf\|t\|_{\Omega_3}\subset\|t\|_{\Omega_2}\subset sup\|t\|_{\Omega_3}$$

can be derived based on Interpretation 3 and the completeness informations.

In Problem 3, first the relations of Ω and Ω_2, Ω_1 and Ω_3 are discussed. Where, eather Ω_2 or Ω_3 is known. In general, for any term t,

$$inf\|t\|_{\Omega_3}\subset inf\|t\|_{\Omega_1} \quad and \quad sup\|t\|_{\Omega_3}\subset sup\|t\|_{\Omega_1},$$

$$\|t\|_{\Omega_2}\subset\|t\|_{\Omega}$$

can be derived. In general, if only Ω_3 is known then we can not prescribe the upper bound of $\|t\|_{\Omega}$. But for a term which satisfies some conditions in Theorem 21,

$$inf\|t\|_{\Omega_3}=inf\|t\|_{\Omega_1} \quad and \quad sup\|t\|_{\Omega_3}=sup\|t\|_{\Omega_1}$$

hold. In this case, the relation of Ω and Ω_3 can be prescribed. Namely, in this case,

$$inf\|t\|_{\Omega_3}\subset\|t\|_{\Omega}\subset sup\|t\|_{\Omega_3}$$

can be derived, i.e., the upper bound of $\|t\|_{\Omega}$ can be calculated based on the incomplete system Ω_3 with respect to the assignment and the object.

By these discussion, we prescribed what is the sure response and what is the possible response, furthermore we can clarified terms which is not affected by the incompletencess.

There exist some other problems.

1) In this paper, we suppose that each object surely takes one attribute ualue. We did not consider the case that some objects take several attribute values or take no value.

2) How can we use and apply constraints like functional dependency in our system?

Acknowledgement

his guidance and encouragement.

References

[1] CODD, E. F.: *Extending the Database Relational Model to Capture More Meaning*, ACMTDS, **4**, 4 (1979), 397-434.

[2] LIPSKI, Jr. W.: *On Semantic Issues Connected with Incomplete Information Databases*, ACMTDS, **4**, 3 (1979), 262-296.

[3] LIPSKI, Jr. W.: *On Databases with Incomplete Information*, J. ACM, **28**, 1 (1981), 41-70.

[4] REITER, R.: *On Closed World Databases*, In Logic and Databases, Gallaire, H., Eds. (1979).

[5] MOTRO, A.: *Completeness Information and Its Application to Query Processing*, Proc. 12th Int'l Conf. on VLDB. (1986), 170-178.

[6] MAREK, W. and PAWLAK, Z.: *Information Storage and Retrieval Systems: Mathematical Foundations*, TCS, 1 (1976), 331-354.