APPROXIMATION OF INDISTINCT CONCEPTS

Zhang Ningyi
Institute of Applied Mathematics
Guizhou Academy of Sciences
40 East Yanan Road, Guiyang, China
(Completed March, 1987.)

ABSTRACT

This theory on semi-equivalence relations is an important and useful tool for investigating classification, pattern recognition, polling and inference etc. Based on it, this paper presents a new framework, in which an indistinct concept, that is one with incomplete or undetermined information about the objects, can be represented approximately. Such an approximate representation will reflect deep structures of concepts which are meaningful for the system. Clearly, the work we present here is to a great extent inspired by general discussions of knowledge engineering research. The theory developed here seems to be of interest in knowledge representation and natural language processing. From the implementation point of view, this theory can be realized by various AI techniques.

0. INTRODUCTION

Two major issues of knowledge engineering are representation and utilization of knowledge. Following Orlowska and Pawlak [1], anything that can be spoken about in the subject position of a natural language sentence is an object, properties of which are fundamental elements of the knowledge of a given domain; then concepts are more complex elements of knowledge. This gives the possibilities of representing the concepts related to a given domain. To represent indistinct concepts——the information about a set of the objects represented by such a concept is undetermined or incomplete in a sense,——Pawlak [2] introduced the rough sets concepts defined by union of some equivalent classes of an equivalence relation. Unfortunately, in many real situations it is not sufficient to consider equivalence relation only. In fact, a lot of relations determined by the attributes of objects do not satisfy transitivity. This limits the expressive power of rough sets.

The Preceding fact forced us to extend Pawlak's works. Semi-equivalence relation theory [3] just offers one of the possible research directions in this field. The original idea of the theory was suggested by Poincare'. Wu Xuemou and his colleagues have established and developed the theory [3], [4]. In its framework, we give interior and exterior approximations of indistinct concepts respectively. Its gradual approximations defined in terms of a family of semi-equivalence relations are also given in it. Such approximate representations will reflect deep structures of concepts and improve the expressive power of Pawlak's knowledge representation system. The work we present in this paper provides a powerful tool for incomplete knowledge representation and utilization, and develops some new researches in AI, e.g. pattern recognition, automated deduction, search methods, etc.——these will be discussed in other papers.

1. PRELIMINARIES

In this paper, we will use almost the same terminology and notations as in [4]. First, we will give a brief account of semi-equivalence relations.

<u>Definition 1. 1</u> [3] A (binary) relation 6 on a nonempty set G is called a semi-equivalence relation

if the following conditions hold for any a , b and c in G,

1. $a \delta a$ (reflexivity)

2. $a \delta b \Rightarrow b \delta a$ (symmetry).

If it further satisfies

3. $a \delta b$ and $b \delta c \Rightarrow a \delta c$ (transitivity)

then we call δ an equivalence relation.

Let $Es[G] = \{\delta, \delta$ is a semi-equivalence relation on $G\}$

$E[G] = \{\epsilon, \epsilon$ is an equivalence relation on $G\}$

Obviously, $E[G] \subseteq Es[G]$

Theorem 1.1 [4] $(Es[G], \cup, \cap)$ is a complete lattice, where G^2 (complete relation) is the greatest element and I (equality relation) is the least element.

Lemma 1.1 [4] $(E[G], \cap)$ is a complete lower semi-lattice. $E[G]$ does not close under $\cup$, but $(E[G], <)$ is a complete lattice.

Definition 1.2 For any $\delta \in Es[G]$ and any given $a \in G$, we call the set $\{b, a \delta b, b \in G\}$ a relative class of a to δ, in symbol $[a]_\delta = \{b, a \delta b, b \in G\}$. The family of sets $\{[a]_\delta, a \in G\}$ is called a relative quotient set of G and is denoted as G_δ.

Definition 1.3 [3] For $\delta \in Es[G]$, a set $Q \subseteq G$ with $Q^2 \subseteq \delta$, maximal with respect to inclusion,

$Q = \max\{A \subseteq G, A^2 \subseteq \delta\}$

is a semi-equivalence class of G relative to δ ; a family of sets $\{Q, Q$ is a semi-equivalence class of G relative to $\delta\}$ is the semi-equivalence quotient set of G relative to δ and is denoted as G/δ.

From the above definition, it is easy to verify the following facts.

Corollary 1.1 If $\delta \in E[G]$ and $a \in Q \in G/\delta$, then $[a]_\delta = Q$, for each $a \in G$.

Corollary 1.2 For any $Q \in G/\delta$, if $a \in Q$, then $Q \subseteq [a]_\delta$. Therefore, $|G/\delta| < |G_\delta| < |G|$. (Here $|A|$ denotes the cardinal of A, for any set A).

Theorem 1.2 For any $a \in G$, there is $b \in G$ such that $a \in [b]_\delta$, there also is $Q \in G/\delta$ such that $a \in Q$.

Theorem 1.3 For any $\delta \in Es[G]$, $a, b, c \in G$,

1. $b \in [a]_\delta$ iff $a \delta b$

2. $b \in Q \in G/\delta$ iff $\forall c \in Q \Rightarrow b \delta c$

3. $\forall Q \in G/\delta$, the restriction $\delta|Q$ of δ to Q is an equivalence relation on Q.

The proofs of theorem 1.2 and 1.3 are trival.

Theorem 1.4 [3], [4] $\bigcup_{a \in G} [a]_\delta = G$, $\bigcup G/\delta = G$.

Theorem 1.5 For any $\delta \in Es[G]$,

$$\bigcup_{Q \in G/\delta} Q^2 = \bigcup_{a \in G} [a]_\delta^2.$$

Proof. Immediate.

Definition 1.3 Let I be an index set. Suppose for any $i \in I$, $\delta_i \in Es[G_i]$. Define $\delta = \prod \delta_i \subseteq \prod G_i \times \prod G_i$ as follows,

$(\vec{a}, \vec{b}) \in \delta$ iff $(a_i, b_i) \in \delta_i$ for any $\vec{a}, \vec{b} \in \prod G_i$ and each $i \in I$, where a_i, b_i is i-th component of $\vec{a}$ and $\vec{b}$ respectively.

Theorem 1.6 [3] For $\delta = \prod \delta_i \in Es[\prod G_i]$, we have

$\prod G_i / \prod \delta_i = \{\prod Q_i, Q_i \in G_i/\delta_i, i \in I\}$

$[\prod G_i]_{\pi\delta_i} = \{[\vec{a}], \ldots, [\vec{a}],$

$= \{\vec{b}, a_i \delta_i b_i, i \in I\}\}$.

Remark. If $\delta_i \in E[G_i]$, then the above facts still hold and are transformed into ones corresponding to the theory on the equivalence relations. Further discussions on these works are given in [3], [4]

2. APPROXIMATE DEFINABILITY

In general, we are not able to distinguish all the objects by means of properties of these objects, informations about which are incomplete or undetermined. To deal with such cases we introduce notions of approximate definabilities of sets. Definitions and inferences, introduced in this and the next sections, are applied to that case in which $G = \bigcup\{Q_i, Q_i \in G/\delta, i \in I\}$.

Definition 2.1 For any given $\delta \in Es[G]$, a set $A \subseteq G$ is δ-definable, if there is $I_0 \subseteq I$ such that $A = \bigcup\{Q_i, i \in I_0\}$, where $G = \bigcup\{Q_i, Q_i \in G/\delta, i \in I\}$. Denote $Def[G] = \{A, A \subseteq G$ and A is δ-definable $\}$.

Clearly, both the empty set and the universal set G are δ-definable. By the definition we easily obtain,

Theorem 2.1 $(Def[G], \cup)$ is a complete upper semi-lattice. In general, $Def[G]$ is not closed under $\cap$ or c, where $\cup$, $\cap$, c are union, intersection and complement respectively.

Definition 2.2 A set G is δ-selective iff any $Q \in G/\delta$ is a set containing a single element.

Corollary 2.1 A set G is δ-selective iff any $A \subseteq G$ is δ-definable.

Proof. $\Rightarrow$, Since any set $Q \in G/\delta$ contains only a single element, so 6 is an equality relation on G. It implies Def [G] $=\{A, A\subseteq G\}$. Therefore, any $A\subseteq G$ is 6 -definable.

$\Leftarrow$, By hypothesis $\{a\}\in$ Def [G] holds for any $a\in G$. It implies $G/\delta=\{\{a\}, a\in G\}$.

Definition.2 3 For any $\delta\in$ Es [G] , $A\subseteq G$, we say that

1. the set $\overline{A}=\cap\{B, A\subseteq B$, $B\in$ Def [G] $\}$ is an exterior approximation of set A, and the set $\underline{A}=\cup\{B, B\subseteq A, B\in$ Def [G] $\}$ is an interior one;

2. a set A is approximately δ -definable if $\overline{A}\neq G$ and $\underline{A}\neq\phi$;

3. a set A is internally 6 -nondefinable if $\underline{A}=\phi$, and A is externally δ -nondefinable if $\overline{A}=G$. A is totally δ - nondefinable if $\underline{A}=\phi$ and $\overline{A}=G$.

Roughly speaking, δ -definability gives us a possibility to answer such membership question as x ?$\in$ A precisely. Approximate definability enables us to decide that an element x more definitely belongs to A or not to A; or is in the borderline case, which depends on the information provided by the objects.

Theorem 2. 2 For $\delta\in$ Es [G] , $A\subseteq G$, we have§

1. A is δ -definable iff $\underline{A}=\overline{A}=A$

2. $\underline{A}\subseteq A\subseteq\overline{A}$

3. $A\subseteq B\Rightarrow\overline{A}\subseteq\overline{B}$, $\underline{B}\subseteq\underline{A}$,

4. $\underline{\underline{A}}=\overline{(\underline{A})}=\underline{A}$, $\overline{\overline{A}}=\overline{A}\subseteq\underline{(\overline{A})}$

5. $\overline{A\cup B}=\overline{A}\cup\overline{B}$, $\underline{A\cup B}=\underline{A}\cup\underline{B}$.

Proof. We should prove $\underline{A\cup B}=\underline{A}\cup\underline{B}$ as an example only. The others are trival.

By definition 2.3, it is clear that

$\underline{A}\subseteq A$, $\underline{B}\subseteq B\Rightarrow\underline{A}\cup\underline{B}\subseteq A\cup B\Rightarrow\underline{A}\cup\underline{B}\subseteq\underline{A\cup B}$.

On the other hand, we suppose $A\cap B=\phi$ without loss of generality, and let J be a subset of I such that $\underline{A\cup B}=\cup\{Qj$, $Qj\in G/\delta$, $J\in J\}$. Then $\underline{A\cup B}\subseteq A\cup B\Rightarrow\cup\{Qj, j\in J'\}\subseteq A$, $\cup\{Qj, j\in J''\}\subseteq B$, where $J'\cup J''=J$.

Therefore, $\underline{A\cup B}=\underline{A}\cup\underline{B}$

Example 2. 1. Let us consider a set G, which consists of six people. Let δ be a relation on G such that

$(o_i, o_j)\in\delta$ iff o_i and o_j are familiar with each other, for any i , j $\in\{1,2,3,4,5,6\}$.

The relation 6 is defined by the following table,

S	o_1	o_2	o_3	o_4	o_5	o_6
o_1	1	0	1	1	0	0
o_2	0	1	0	1	1	1
o_3	1	0	1	0	0	0
o_4	1	1	0	1	0	1
o_5	0	1	0	0	1	0
o_6	0	1	0	1	0	1

where if o»6 ojthen write 1 on the crossed point of i-th line and j-th column, otherwise write 0. So, we get G/δ = {{ o, , o,}, { o_3 , o_4 , o,), { o_t , o,})
For A- { o, , o $_t$,o,}, we have $\overline{A}$ = (o, , o, , o,}, A- (o, , o.)

3. GRADUAL APPROXIMATIONS OF INDISTINCT CONCEPTS

By theorem 1.1 (Es [G] , $\cup$, $\cap$) is a complete lattice. For the sake of convenience, let $(J, \vee, \wedge)$ be such an algebra that the following conditions are held,

$\delta i\cap\delta j=\delta k$ iff $i\vee j=k$,
$\delta i\cup\delta j=\delta k$ iff $i\wedge j=k$

for any i, j, k $\in J$, δi, δj, $\delta k\in$ Es [G] . In particular, δ_0 is a complete relation on G, δ_1 is an equality relation on G where 0 and 1 are the least and the greatest elements of J respectively.

We consider a family of relations $\{\delta k, k\in K\subseteq J\}$ $\subseteq$ Es [G] . Without loss of generality, we suppose that $\{\delta k, k\in K\}$ is a monotone decreasing sequence, by which gradual approximations of the indistict concepts will be established.

Definition 3. 1 For any $i\in K$, a function $f_i, G^2\rightarrow\{0,1\}\subseteq J$ is defined by

$$f_i(a,b)=\begin{cases}1 & \text{if } (a, b)\in\delta i\\ 0 & \text{otherwise}\end{cases}$$

for both a and b in G.

Definition 3.2 A function $f, G^2\rightarrow K$ is defined by $f(a,b)=\vee\{(f_i(a,b)\wedge i), i\in K\}$

And let $\delta\subseteq G^2$ be a relation such that $a\delta b$ iff $f(a,b)>0$, for any a,b in G .

Lemma 3.1 For any i, j ∈ K if i < j then $f_i(a,b) > f_j(a,b)$, ∀ a, b ∈ G.

Proof. Immediate from monotony of {δk, k ∈ K} and definition 3.1

Theorem 3.1 δ = ∪{δi, i ∈ K} ∈ Es [G].

Proof. For all a, b in G.

$(a,b) \in \delta \iff f(a,b) > 0$
$\iff \exists\, i \in K,\ f_{i}(a,b) \wedge i > 0$
$\iff \exists\, i \in K,\ f_{i}(a,b) = 1 \text{ and } i > 0$
$\iff (a,b) \in \cup\{\delta i,\ i \in K\}.$

It is now evident that δ ∈ Es [G].

Lemma 3.2 If λ, δ ∈ Es [G], and λ < δ, then for any Q ∈ G/λ, there is P ∈ G/δ such that Q ⊆ P.

Proof. It is immediate from setting
P = max {A, Q ⊆ A and A' ⊆ δ}.

For any given set A ⊆ G, let $\overline{A}$ and $\underline{A}$ be exterior and interior approximations of A with respect to δ, respectively; and let $\overline{A_i}$ and $\underline{A_i}$ be exterior and interior approximations of any grade i ∈ K of A, with respect to δ_i, respectively. It seems true that the sequences $\{\underline{A_i}, i \in K\}$ and $\{\overline{A_i}, i \in K\}$ should satisfy montoneity. Unfortunately, the following example shows that neither the sequence $\{\underline{A_i}, i \in K\}$ nor the sequence $\{\overline{A_i}, i \in K\}$ satisfies monotoneity —— monotone increasing or decreasing. So, we will consider only the case of {δi, i ∈ K} ⊆ E [G].

Example 3.1 let G = {a, b, c, d}, δ = {a, b, c}' ∪ {(b, d), (d, b), (d, d)}, λ = {a, c}' ∪ {b, d}'.

Obviously, λ < δ. So we have,

$\overline{\{a,b\}}_\lambda = G \supseteq \overline{\{a,b\}}_\delta = \{a,b,c\}$

$\underline{\{a,c\}}_\lambda = \{a,c\} \supseteq \underline{\{a,c\}}_\delta = \phi$,

but then we also have,

$\overline{\{a,c\}}_\lambda = \{a,c\} \subseteq \{a,b,c\} = \overline{\{a,c\}}_\delta$

$\underline{\{a,b,c\}}_\lambda = \{a,c\} \subseteq \{a,b,c\} = \underline{\{a,b,c\}}_\delta$.

Even so, it is quite a useful tool for gradual approximations of the indistinct concepts, specially when we try to simplify our problems. Generality speaking, a concept can be represented by listing the attributes of objects. The more of the attributes we list, the better the approximations are But this is usually to be done only in an extent.So, we can use $\cup\{\underline{A_i}, i \in K\}$ $(\cap\{\overline{A_i}, i \in K\})$ as an approximation of $\underline{A}$ $(\overline{A})$ in most cases.Here, we give a simple illustration of exploiting the theory.

Example 3.2 Let G be a set constituted by a given group of peysons. Suppose R(o), R(s) and R(t) are three relations on G. i.e. R(o) is the neighbor relation, R(s) the same schoolmate relation,and R(t) the townsman relation.Given a set A ⊆ G, our task is to find out that in G who are closely related to one another among themselves and who are closely related to a person in A.

To do so, we first get interior and exterior approximations $\underline{A}(o)$ and $\overline{A}(o)$ of A with respect to R(o). Similarly, we have $\underline{A}(s)$, $\overline{A}(s)$, $\underline{A}(t)$ and $\overline{A}(t)$ Secondly, let's set

$\underline{B} = \underline{A}(o) \cup \underline{A}(s),\ \overline{B} = \overline{A}(o) \cap \overline{A}(s)$

Then $\underline{B}$ and $\overline{B}$ are the approximations of interior and exterior approximations of A with respect to R (o) ∩ R (s) respectively. In the same way. we obtain ,

$\underline{C} = \underline{A}(o) \cup \underline{A}(t),\ \overline{C} = \overline{A}(o) \cap \overline{A}(t)$

$\underline{D} = \underline{A}(s) \cup \underline{A}(t),\ \overline{D} = \overline{A}(s) \cap \overline{A}(t)$

$\underline{E} = \underline{A}(o) \cup \underline{A}(s) \cup \underline{A}(t)$

$\overline{E} = \overline{A}(o) \cap \overline{A}(s) \cap \overline{A}(t).$

Finally, we can choose a rational solution based on our understanding of the saying" Be closely related to a person in A". Such an idea seems useful to machine cognition, natural language understanding and automatic theorem proving, etc.

Of course, the situations become cleaer if we limit ourselves to the case of the set of equivalenc relations E[G]. Now, $\underline{B}$ abd $\overline{B}$ in example 3. 2 are really interior and exterior approximations of A with respect to R(o) ∩ R(s) respectively.The reason for this lies in the following theorems.

Lemma 3.3 Suppose any δ, λ ∈ E [G] and let δ < λ. For each Q ∈ G/λ there is {Pj, j ∈ J} ⊆ G/δ such that Q = ∪{Pj, j ∈ J}.

Proof.Clearly δ|Q ∈ E [Q]. Let Q/δ = {Pj, j ∈ J} It is sufficient to prove that Pj ∈ G/δ for each j ∈ J. In fact, if it is not true, then we can suppose that there is some Pj ∉ G/δ. Thus there is always Pj* ∈ G/δ such that Pj ⊊ P*. So (P*j ∪ (Q−Pj))' ∈ G/λ, it is contradictory with maximality of Q (see definition 1.3).

From above lemma it is easy to establish the following facts ,

Theorem 3.2 Let {δi, i ∈ K} ⊆ E [G] be a monotone dcreasing sequence, then the sequence $\{\overline{A_i}, i \in K\}$ is monotone decreasing and $\{\underline{A_i}, i \in K\}$ is monotone increasing. Moreover, $\underline{A_i} \subseteq \underline{A} \subseteq \overline{A} \subseteq \overline{A_i}$ for each i ∈ K.

Corollary 3.2 $\overline{A} = \cap\{\overline{A_i}, i \in K\}, \underline{A} = \cup\{\underline{A_i}, i \in K\}$. That is

$a \in \overline{A}$ iff for all $i \in K$ there is $b \in G$ such that $(a, b) \in \delta_i$,

$a \in \underline{A}$ iff there is $i \in K$ such that for all $b \in G$ and $(a, b) \in \delta_i$.

Remark. If a relation $\delta \in E[G]$ is given, we choose a monotone decreasing family of relations, $\{\delta_i, i \in K\} \subseteq E[G]$, according to the practical considerations, such that $\cup \{\delta_i, i \in K\} \supseteq \delta$. Let $\delta_i^* = \delta \cap \delta_i$ for $i \in K$. Then by what is mentioned before we obtain exterior and interior approximations $\overline{A}_i$ and $\underline{A}_i$ of any grade $i \in K$ of A, with respect to δ_i^*. In fact, such an idea has been realized in a Computer Diagnosing System [5]. The Logical formalism that provides tools for the examination of expressive power of the system in terms of approximate definability is discussed further in other papers.

REFERENCES

1. Orlowska E .and Pawlak Z., Expressive Power of Knowledge Representation Systems, Int. Journ. of Man-Machine Studies, 20, 1984.
2. Pawlak Z., Rough sets, Int. Jour, of Computer and Information Science, 11 (5), 1982.
3. Wu Xuemou, Pansystems Analysis, Some New Investigation of Logic Observ-Controllability and Fuziness, Journ. Huazhong I. T. ,1981.
4. Zhang Mingyi, Pansystems Es-homorphism and Semi-congruence, Guizhou Science, 2, 1984.
5. Zhang Mingyi, Li Danning, A Computer Diagnosing System and its Application to Diagnosing Patients with Cardiac Pacer, Procce. 5—th Chinese Symposium on Artificial Intelligence, Beijing, 1986.